Copyright
by
Suryansh Purwar
2008
A Study on the Use of Proxy Reservoir Models in Field Development Optimization and Value of Information Problems

by

Suryansh Purwar, B.Tech.

Thesis
Presented to the Faculty of the Graduate School of
The University of Texas at Austin
in Partial Fulfillment
of the Requirements
for the Degree of

Master of Science in Engineering

The University of Texas at Austin
August 2008
A Study on the Use of Proxy Reservoir Models in Field Development Optimization and Value of Information Problems

Approved by
Supervising Committee:

Christopher J. Jablonowski

Quoc P. Nguyen
Dedication

To my late grandfather Seth Hari Krishna for his inspiration

To my parents Arun & Jyotsna Purwar

and my sisters Shubhra & Shilpi Purwar

for their love and support
Acknowledgements

This thesis is the product of research performed by myself and my advisors Dr. Christopher J. Jablonowski and Dr. Quoc P. Nguyen at the University of Texas at Austin.

I am grateful to Chris for his support and guidance throughout these two years. Without his direction, I would not have been able to work in this exciting new field which tries to bring engineers and economists on a common page. I am also grateful to him for his dedication in reviewing and correcting the thesis.

I would also like to thank my co-supervisor Dr. Quoc Nguyen for giving valuable inputs and also for reviewing this thesis. I thank Dr. Lasdon from the McCombs School of Business for his teachings on optimization theory and practice. I thank my colleagues Chai, Neal, Mark and Hariharan for their work, ideas and inputs which have been instrumental in getting this thesis to be what it is now. I owe a big thanks to the University of Texas at Austin, its professors and my colleagues for these fruitful and enjoyable two years which have been a great experience personally and professionally.

I also acknowledge the financial support given for this research by Marathon Oil, Maersk Oil, OMV, and the Cockrell School of Engineering at The University of Texas at Austin without which this project would not have been possible.
Finally, I would like to express my gratitude to my parents, sisters and my undergraduate supervisor S. Ray at the Indian Institute of Technology, Kharagpur for motivating me to pursue higher studies and for their constant encouragement.

August, 2008
Abstract

A Study on the Use of Proxy Reservoir Models in Field Development Optimization and Value of Information Problems

Suryansh Purwar, M.S.E.
The University of Texas at Austin, 2008

Supervisor: Dr. Christopher J. Jablonowski
Dr. Quoc P. Nguyen

Mitigating uncertainty in subsurface variables to optimize field development is an intriguing problem still lacking a robust solution. The complexity is further increased by design constraints and quasi-irreversible capital investments. Although versatile optimization techniques and powerful reservoir simulators are available, the primary reason for the lack of a comprehensive solution is the disconnect between optimization and reservoir simulation studies. In this work, we propose an approach that offers a reasonable way to do this. For handling the uncertainties, the proven approach of Experimental Design was used to create cases of possible combination of reservoir variables which were then run in CMG simulator to generate production profiles. Response Surface methodology was employed to generate numerical or ‘proxy’ oil
production models from the reservoir simulation results. The numerical models thus generated were plugged into an optimizing model in GAMS to get the desired facility plan. A workflow with these components also allows us to look at the value of buying more information to mitigate uncertainty and the value of incorporating a ‘real’ option of modifying facilities after $n$ years to support initial capital investment decisions.
Table of Contents

List of Tables ..............................................................................................................................................xii

List of Figures ...............................................................................................................................................xiii

Chapter 1: Introduction .................................................................................................................................1

Chapter 2: Theoretical Background .............................................................................................................5

2.1 Response Surface Methodology (RSM) .................................................................................................5

2.1.1 Design of Experiments ......................................................................................................................8

2.1.1.1 Full Factorial Design ..................................................................................................................9

2.1.1.2 Fractional Factorial Design .......................................................................................................10

2.1.1.3 Central Composite Design .....................................................................................................10

2.1.1.4 Box-Behnken Design ..............................................................................................................15

2.1.1.5 D-Optimal Design ....................................................................................................................16

2.1.1.6 Plackett-Burman Design .........................................................................................................18

2.1.1.7 Limitations of Design of Experiments ......................................................................................19

2.1.1.8 Selection of Design of Experiments ..........................................................................................20

2.1.2 Reservoir Simulation .........................................................................................................................22

2.1.3 Ordinary Least Squares (OLS) Regression .......................................................................................25

2.2 Field Development Optimization .........................................................................................................26

2.2.1 Value of Information .......................................................................................................................28

Chapter 3: Problem Definition ...................................................................................................................32

3.1 Reservoir Description ............................................................................................................................32
3.1.1 Subsurface Uncertainties.........................................................37
3.1.2 Design Variables.................................................................42
3.2 Capital Investment Decisions......................................................44

Chapter 4: Reservoir Simulations....................................................47
4.1 Reservoir Simulator.................................................................48
4.2 Design of Experiments............................................................48
4.3 Specification for the Simulations .................................................49
4.4 Simulation Results.................................................................51

Chapter 5: Proxy Model Estimation..................................................54
5.1 Brief History of Proxy Models..................................................55
5.2 Model Specification and Results...............................................60

Chapter 6: Economic Optimization................................................65
6.1 Algorithm..................................................................................66
6.2 Model Equations........................................................................67
  6.2.1 Objective Function..............................................................68
  6.2.2 Cost Functions......................................................................69
  6.2.3 Reconfiguration Equations....................................................73

Chapter 7: Analysis and Results......................................................75
7.1 Planning with Current Information (Determining Optimal Initial Facility
  Configuration)..............................................................................75
7.2 Modeling Alternate Outcomes Regarding Reservoir Uncertainty........80
7.3 Planning for Alternate Outcomes: The Value of Information.............81
# List of Tables

2.1 Selection Criteria for DoE

4.1 BBD Table

4.2 Value of Parameters Used for Simulation Runs

7.1 NPV Distribution for Base Case

7.2 NPV Distribution for Alternate Hypothesis

7.3 NPV and Standard Deviation Values for Base case and Alternate Hypothesis

7.4 NPV Distribution for Base Case and Alternate Hypothesis for $60/bbl oil price

7.5 NPV Distribution for Base Case and Alternate Hypothesis for $60/bbl oil price

7.6 NPV Distribution for Base Case and Alternate Hypothesis for $60/bbl oil price

7.7 Scenarios for Analysis

7.8 Reconfiguration Options for High Uncertainty and High Oil Price Scenario

7.9 Reconfiguration Options for Low Uncertainty and High Oil Price Scenario

7.10 Value Addition by Reconfiguration for Case A

7.11 Value Addition by Reconfiguration for Case B

7.12 NPV and Standard Deviation Values for Double the Cost Estimates

A.1 Naïve Estimator

A.2 AR-1 Estimator
List of Figures

1.1 Workflow Chart.................................................................4
2.1 Graphical Representation of CCD.........................................12
2.2 Comparison of Three Types of CCD.....................................14
2.3 Value of Information Tree..................................................31
3.1 Reservoir Configuration....................................................35
3.2 Relative Permeability Curve ..............................................38
3.3 Relative Permeability Curves for Oil and Gas.........................39
3.4 Different Well Patterns....................................................43
4.1 Simulation Results..........................................................52
4.2 Reservoir After Flooding..................................................53
7.1 NPV Distribution for Base Case..........................................78
7.2 NPV Distribution for Alternate Hypothesis..........................80
7.3 NPV Values for Base Case and Alternate Hypothesis..............82
7.4 NPV Distribution for Base Case and Alternate Hypothesis for $60/bbl oil price.85
7.5 NPV Distribution for Base Case and Alternate Hypothesis for $K_{rg}$.................87
7.6 NPV Distribution for Base Case and Alternate Hypothesis for $K_{ro}$.................90
7.7 Graphical Distribution of NPV and Standard Deviation Values for Double the Cost Estimates.................................................................99
CHAPTER 1: INTRODUCTION

The goal of field development optimization is to determine the development plan that will enable the operator to extract the maximum potential value from an asset. That said, optimizing a field for an exploratory block involves a multitude of inherent challenges. There are always uncertainties associated with key reservoir parameters due to reasons like inconclusive tests or insufficient availability of data. This leads to uncertain, imperfect and incomplete forecasts\textsuperscript{18-22} that manifest in numerical simulations as well. The uncertainties never cease to exist even when a secondary or tertiary recovery plan is being implemented. Such uncertainties impair the judgment of a reservoir engineer for developing field plans. In some cases, these uncertainties do not affect decision-making economically. In other cases, they may introduce significant engineering and/or economic risk to the project, and mitigation investments may be warranted. The constant questions that the engineer seems to be asking is “What if this scenario is not true and the other one is” or “If only I knew this with certainty I could optimize better” or “How much should I pay to get this information with certainty”. Mitigating such uncertainty often corresponds to investments in additional analysis and new information. Determining whether to make such investment is not a trivial exercise and hence analyzing, describing, mitigating and/or accommodating uncertainty prior to making initial capital decisions becomes imperative. One of the objectives of this thesis is to develop a workflow to facilitate such an analysis.
One of the popular approaches to incorporate more uncertainty into project analysis is to define NPV as a function of certain key subsurface parameters and then run a Monte Carlo analysis to estimate the net present value (NPV) distribution. The scope of the analysis is often limited. Recently there has been a renewed interest in having a proper methodology in place to handle uncertainties while planning and designing oil fields. Design of experiments (DoE) and Response Surface Methodology (RSM) have been used quite regularly to manage such scenarios and such usage in literature has been discussed in detail in Section 5.1.

In spite of significant improvements in the evaluation process and more rigorous handling of field development plans, current processes still have a major fallacy in non-implementation of real options in the initial analysis that supports major capital decisions in the field. When there are options (i.e. expand/abandon/alter field) in the future to respond to outcomes (i.e. revealed uncertainties), a sophisticated model of decision-maker behavior is warranted. A static view of the project will both undervalue the project and can lead to suboptimal initial investment decisions. Such an option can add significant value to the project which is best demonstrated through an example. Let us consider a project. An offshore field has been discovered and preliminary tests indicate a presence of 20 million BOE recoverable oil but the uncertainties present also result in a 40% standard deviation. If we go ahead and build the facilities with an estimate of 20 million BOE, there is a reasonable chance that the actual reserves are less than that and we lose some value in our project. Alternately, we could also end up with a discovery
higher than what we had imagined. In such a situation, the engineer chiefly concerns himself with one question “How big should be the facility”. If we include an option in our analysis, exercising which is not an obligation, that we can abandon or expand the field after three years depending on the size of our discovery, the value proposition changes. The option to start with a small facility and expand later might be the best alternative. Now, the questions one seeks to answer are “What is the value addition of having such an option in our plan” and “What should be the optimum facility size if this option is included in the calculations”. This example encapsulates the basic idea behind this work. Inclusion of real options could change the optimized values obtained through standard uncertainty analyses.

This thesis uses the concepts of DoE and RSM to outline an approach to solve such decision-making problems. Together, these components define a workflow (shown in Fig. 1) that can be used to answer various questions and support recommendations for initial capital investment and uncertainty mitigation. Ultimately, the ideal solution would be the integration of reservoir simulation and project optimization, thus doing away with all the intermediate steps shown in the workflow.
Fig 1.1 Workflow Chart

Chapter 3

Chapter 4

Chapter 5

Chapter 6

Chapter 7
CHAPTER 2: THEORETICAL BACKGROUND

2.1 RESPONSE SURFACE METHODOLOGY (RSM)

Response surface methodology (RSM) is used to model complex phenomena like production from oil and gas reservoirs by relating multiple explanatory variables to one or more response variables. RSM is useful for modeling, improving, and optimizing processes, and it has been widely applied in several areas, such as agriculture, industrial production engineering, and petroleum reservoir engineering. The initial development of RSM is generally attributed to Box and Wilson in 1951. The great advantage of RSM is that it is easy to apply even when little is known about the phenomena being modeled. There are various ways to implement RSM, but there are three typical components of RSM in applied analysis:

- The Design of Experiments for defining the set of explanatory variables
- A proxy model which relates the response variable to the explanatory variables
- An optimization method to estimate the marginal impacts of the explanatory variables on the response variable.

One of the criticisms of RSM is that it is too abstract and there is variability and the potential for bias in the estimation of the relationships and therefore in predicted values. However, more detailed modeling entails costs (in resources and time), and the value derived from additional effort may not exceed the costs. RSM seeks to balance these
costs and benefits, ideally on a case by case basis. The optimal design depends on the
objective of the problem, the data available, and on the magnitude of the decision being
made. In some cases, less detail and less iteration will be acceptable, while in other cases,
the need for additional accuracy will justify more detailed specifications and more
iteration. Several methods for RSM and Design of Experiments have been proposed in
the literature and the most popular methods are described in the remainder of this chapter.
Most methods have been devised to address specific scientific and engineering problems.

Here we describe the mathematical formulation behind Response Surfaces. Assume there
are k independent variables given by the notation $X_1, X_2, ..., X_k$ which affect the
phenomenon under consideration. Response refers to the measurement of the effect of
change in these k independent variables on the value of the variable(s) of interest, or
response variables. The true value of this response would be devoid of any error or
uncertainty. However, the observed value would contain some error originating from
measurement error or from relevant but excluded $X$ variables (knowingly or unknowingly
excluded):

$$Y = \eta + \varepsilon$$  \hspace{1cm} (2.1)

where

$Y = $ the response variable,

$\eta = f(X_1, X_2, ..., X_k)$

$\varepsilon$ is a random error term.
The response function enables an evaluation of $Y$ for any definition of $\eta$. In practice, this response function is estimated using regression analysis. It is assumed to be a continuous function within the range of the $X$ variables. Parameters or regression coefficients, $\beta_k$ represent the rate of change of $Y$ with respect to $X_k$.

The error term is assumed to have a normal distribution with mean zero and variance $\sigma^2$. The explanatory variables can be transformed into dimensionless variables having mean zero and the same variance by using the following formula:

$$X_i^n = \frac{X_i - \frac{\text{max}(X) + \text{min}(X)}{2}}{\frac{\text{max}(X) - \text{min}(X)}{2}}$$

(2.2)

This transforms all the values to a range of -1 to +1. While some analysts prefer dimensionless variables, this transformation is not required.

A typical response surface is estimated using the following steps:

- Define a set of combinations of explanatory variables $X_1, X_2, \ldots, X_k$ using DoE.
- For each combination defined in Step 1, obtain the respective values of $Y$ using experiments or simulations. In this study, a commercial reservoir simulator is
used to generate the response variable and hence the term “experiment” in this work would mean “numerical simulation.”

- Specify the response function as a linear combination of the $K$ explanatory variables.
- Use regression analysis to estimate the coefficients of the $X$ variables.

A common and flexible specification of the response function that is often used in practice includes the variables directly, their squares, and the cross-products. For $K = 3$, the specification would be as follows:

$$Y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \beta_3 x_3 + \beta_4 x_1^2 + \beta_5 x_2^2 + \beta_6 x_3^2 + \beta_7 x_1x_2 + \beta_8 x_1x_3 + \beta_9 x_2x_3 + \varepsilon$$ \hspace{1cm} (2.3)

2.1.1 DESIGN OF EXPERIMENTS (DoE)

The concept of DoE refers to the process of defining a set of experiments or simulations in a systematic, pre-defined and statistically correct way. For complex phenomena, the number of experiments required to completely characterize and explain the phenomena is very large and impractical (resource constrained). For example, reservoir simulations are time-consuming and computational intensive (expensive). Several methods of DoE have been developed to balance the experimental or simulation effort with the accuracy of the model.
In the field of DoE, the word “level” has a specific meaning. A variable’s level is an indication of the variation of the explanatory variable. For example, if a variable varies such that it can take on one of two values, a low value and a medium value, the variable is modeled using a 2-level design. Similarly, if a variable has a low value, a medium value and a high value, it is a 3-level design. The remainder of this section provides a description of the major approaches to DoE.

2.1.1.1 Full Factorial Design

The total number of experiments required for this type of design is given by $k^N$ where, $N$ denotes the level and $k$ denotes the number of explanatory variables. The number of experiments increases exponentially as the number of variables increases. This design is often unsuitable for a large variable space because it is too computationally demanding. There is also the risk of running more experiments than may be required to obtain a satisfactory response function. Even though its accuracy is to its advantage, it seldom has been used in the E&P industry.\(^9\)
2.1.1.2 Fractional Factorial Design

This design often is used for conducting preliminary experiments with a large number of variables and then using the result to screen significant variables and use this subset of variables in a more comprehensive design approach. This method is suitable if one can assume that certain high-order interaction terms in a response function are negligible. They can be used for 2-level as well as a 3-level design.

2.1.1.3 Central Composite Design

The Central Composite Design (CCD) method was introduced by Box and Wilson (1951). It is one of the most popular methods of second-order, or quadratic, designs. The design consists of three sets of experimental runs:

1. A two-level factorial design
2. A set of center points
3. A set of axial points, experimental runs replicating the center point except for one factor.

Hence, the DoE matrix for k variables is a combination of three distinct matrices:

1. Matrix F obtained from factorial experiment where entries are coded as +1 and -1 to indicate maximum and minimum values
2. Matrix C from the centre point where entries are zeros to denote the mean value
3. Matrix $A$ from the axial point with $2k$ rows where one factor is sequentially coded $\pm \alpha$ while all others are coded zero.

The design matrix is merely the vertical concatenation of these matrices. So, the number of runs required for this design is $2^k + 2k + 1$. The value of $\alpha$ as mentioned in the second bullet above varies with parameter. For $k=2$, it is $\sqrt{2}$ while for $k=3$, it is $\sqrt{3}$. For example, the design matrix $D$ for 3 variables using CCD would look like:

$$
D = \begin{bmatrix}
-1 & -1 & -1 \\
1 & -1 & -1 \\
-1 & 1 & -1 \\
1 & 1 & -1 \\
-1 & -1 & 1 \\
1 & -1 & 1 \\
-1 & 1 & 1 \\
1 & 1 & 1 \\
-1.732 & 0 & 0 \\
1.732 & 0 & 0 \\
0 & -1.732 & 0 \\
0 & 1.732 & 0 \\
0 & 0 & -1.732 \\
0 & 0 & 1.732 \\
0 & 0 & 0
\end{bmatrix}
$$

Two-level factorial points

Axial Points

Center Point
A graphical representation of the points is given by the following figure:

Figure 2.1: Graphical representation of CCD

Some key observations regarding this design are as follows:

- The two-level factorial points (Matrix F) have a major role to play in estimating linear terms and two factor interaction terms
• The axial terms (Matrix A) contribute in a large way in the estimation of the quadratic terms; their contribution in the estimation of quadratic terms is negligible and without those axial terms, only the sum of quadratic terms can be estimated.

• The center point is more of an error estimate but it also contributes to the estimation of the quadratic terms.

CCD is an efficient design that covers a large sample space and provides accurate results. Compared to a factorial design, fewer runs are required and this saves time and other resources. Three types of CCD are most common in engineering practice.

**Face-Centered Cube Design (FCD):** This design is appropriate when one cannot determine the responses outside the cubical region because of some impracticality or restriction in the ranges of control or design variables. The responses inside and on the perimeter of the cube are used. The value of $\alpha$ is 1 which means that the axial points are inside the cube rather than outside as was the case for a CCD.

**Circumscribed (CCC):** In this design, the value of $\alpha$ varies from 1 to $k$. This means that the axial points are outside the cube. The center points are still at the origin and the distance of axial points from the center point is equal to $\alpha$. The advantage of this method is that it allows exploration of a large sample space.
**Inscribed (CCI):** This design is more suited for an evolving experiment. The design variables initially are restricted to certain ranges, but as the experiment proceeds, the points at the corners of the cube may not give acceptable responses and hence these may need to be excluded from the design. This is done by scaling the axial points such that they take the values of -1 and +1 and the cube points lay in the interior of the cube. The disadvantage is that this design involves the smallest sample space. Also, because the points at the corners of the cube are not included in the design, the predictability at these points is poor. A CCI design can be said to be a scaled down CCC design with each factor level of the CCC design divided by $\alpha$ to generate the CCI design. This design also requires 5 levels of each factor. The following figure reproduced from the Engineering Statistics handbook\textsuperscript{28} shows the difference between the three designs:

![Diagram of CCI design](image)

Fig. 2.2  Comparison of three types of CCD
2.1.1.4 **Box-Behnken Design (BBD)**

This design was developed by George E.P. Box and Donald Behnken in 1960\(^3\). This method is applicable for a three-level design for approximating a quadratic response function. Each design can be thought of as a combination of a two-level factorial design with an incomplete block design. In each block, a certain number of factors are put through all combinations for the factorial design, while the other factors are kept at the central values. For instance, the Box-Behnken design for 3 factors involves three blocks, in each of which 2 factors are varied through the 4 possible combinations of high and low(±1). It is necessary to include center points also (in which all factors are at their central values). A BBD for three factors can be depicted as follows:

\[
D = \begin{pmatrix}
\pm 1 & \pm 1 & 0 \\
\pm 1 & 0 & \pm 1 \\
0 & \pm 1 & \pm 1 \\
0 & 0 & 0
\end{pmatrix}
\]

This design would result in \(n.2^{k+1}\) number of runs where \(k\) is the number of parameters being varied. Usually it is half the number of design variables.
The **D** matrix for 6 control variables of the BBD is:

\[
D = \begin{bmatrix}
\pm1 & \pm1 & 0 & \pm1 & 0 & 0 \\
0 & \pm1 & \pm1 & 0 & \pm1 & 0 \\
0 & 0 & \pm1 & \pm1 & 0 & \pm1 \\
\pm1 & 0 & 0 & \pm1 & \pm1 & 0 \\
0 & \pm1 & 0 & 0 & \pm1 & \pm1 \\
\pm1 & 0 & \pm1 & 0 & 0 & \pm1 \\
0 & 0 & 0 & 0 & 0 & 0
\end{bmatrix}
\]

The BBD method can be used when there are three evenly spaced levels. As described earlier, the BBD is the three-level design, but the sampling method is not uniform like the FCD. However, the use of the BBD should be confined to situations where approximating response at the extremes is less important because the corner points are not included in the design matrix.

### 2.1.1.5 D-Optimal Design

Experimental design like Full Factorial or Fractional Factorial Design are more suitable in experimental settings where factors are relatively unconstrained in the region of interest. That may not be the case in certain experiments. In other scenarios certain combinations of factors may be infeasible and non-linear or to say that the possible cases are pre-defined and cannot be generated by combination of all factors as seen in the
previous cases. *D-optimal designs* are model-specific designs that rectify these limitations. The two main reasons for using D-optimal designs are:

1. Standard factorial or fractional factorial designs require too many runs for the amount of resources or time allowed for the experiment.
2. The design space is constrained meaning that some runs are not feasible or are impossible to run or the sample space is irregularly spaced. This means it can also work when one of the parameters varies across five levels while others do not. This is in contrast with other designs which relied on parameters varying across two or three levels and forming a regular sample space (cube).

For a certain experiment, let us consider there are only $K$ possible experiments. Since, we cannot run all experiments we need to choose $N$ experiments which give us the best possible estimate. A *D*-optimal design employs an iterative search algorithm that minimizes the covariance of the parameter estimates for a specified model. This is equivalent to maximizing the determinant $D = |X^TX|$, where $X$ is the design matrix of model terms. Thus, out of all possible ways of selecting $N$ out of $K$ observations, forming the matrix $X$, the combination with the highest numerical value of the determinant of the variance-covariance matrix $X^TX$ will be selected. This also improves the accuracy of the regression coefficients. Unlike traditional designs, *D*-optimal designs do not require orthogonal design matrices, and as a result, parameter estimates may be correlated.
This design is extensively used because not only does it require fewer computations, there is also no restriction in the type of model (first order, quadratic etc.) you want to fit. Also, this design offers more flexibility for the user.

2.1.1.6 Plackett-Burman Design

The Plackett-Burman Design (PBD) was developed by Robin L. Plackett and J. P. Burman in 1946\textsuperscript{35}. This design is applicable for two levels only and is particularly useful for economically detecting large main effects, assuming all interactions are negligible when compared with the few important main effects. This design has the least number of possible runs and can accommodate a large number of factors and hence it is most widely used for screening purposes. The one disadvantage with this design is that it does not take into account quadratic interactions. The number of runs are given by \( n \) such that \( n > 4p \) where \( n \) is the number of parameters and \( p \) is any integer. This limits its use in actual model development.

To illustrate its utility, as many as 19 factors can be studied in a mere 20 simulations. The PBD table for 11 parameters is given by (+ denotes the highest value and – the lowest):
2.1.1.7  LIMITATIONS OF DESIGN OF EXPERIMENTS

Most experimental design can be constructed in commercial statistical software. In some situations, standard designs are not appropriate and may sometimes call for a design like D-optimal design. These situations include the following:

1. The sample space of the values of the variable in the experiment does not fit into a standard blocked regular design
2. Not all combinations of the factor settings are feasible
3. Original design combinations may contain forbidden or unreachable combinations that were not considered before the design was generated.

4. There is a nonlinear model.

5. A quadratic or response surface design is required in the presence of “qualitative” factors.

6. There are multiple sources of variation leading to nested structures; a standard fractional factorial design requires too many treatment combinations for the given amount of time and / or resources.

2.1.1.8 SELECTION OF DESIGN OF EXPERIMENTS APPROACH

Each DoE method has advantages and disadvantages relative to other designs. The selection is a function of the time and resources available and the discretion of the experimenter. Criteria for selecting a specific design can be summarized as follows:\textsuperscript{28}

- **Comparative objective**: When the primary task is to figure out how significant a factor is vis-à-vis other factors, there is a comparative problem requiring a comparative design solution. “Significant” here implies a noticeable change in the response for different levels of the factor.
• **Screening objective**: When the primary purpose of the experiment is to select or screen out the few important main effects from the many less important ones, there is a *screening problem*.

• **Response Surface objective**: When the primary purpose of the experiment is to estimate interaction and quadratic effects for different levels of the parameters, thereby finding shape of the response surface, there is a *Response Surface problem* requiring a *Response Surface solution*.

The handbook\textsuperscript{28} also includes a table that summarizes the method of choosing an Design of Experiments for comparative, screening, and response surface designs. It is reproduced below in the Design Selection Table (Table 2.1):

<table>
<thead>
<tr>
<th>Number of factors</th>
<th>Comparative Objective</th>
<th>Screening Objective</th>
<th>Response Surface Objective</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1-factor completely randomized design</td>
<td>-</td>
<td>-</td>
</tr>
<tr>
<td>2-4</td>
<td>Randomized block design</td>
<td>Full or fractional factorial</td>
<td>Central composite or Box-Behnken</td>
</tr>
<tr>
<td>5 or more</td>
<td>Randomized block design</td>
<td>Fractional factorial or Plackett-Burman</td>
<td>Screen first to reduce number of factors</td>
</tr>
</tbody>
</table>

Table 2.1 Selection criteria for DoE
2.1.2 Reservoir Simulation

The purpose of ED is to remove the bias from the variables while generating cases to be run by a reservoir simulator. It is a logical way to set up the “experiment” while the actual “experiment” is done in the reservoir simulator. In this section, the fundamentals of reservoir simulation are discussed.

Most commercial reservoir simulators are essentially based on the fundamental law of Conservation of Mass, Momentum, and Energy. These equations are numerically solved using the IMPES (IMplicit in Pressure and Explicit in Saturation) scheme. The oil production from a reservoir can be characterized into three phases – primary, secondary and tertiary. Primary production is when the reservoir essentially produces under the effect of its own pressure. After a while, the reservoir production declines to such an extent that economical oil production is not possible. This is when secondary production begins usually with a water flood. In this case, there are injector well(s) and producer(s). Water is injected in the reservoir which displaces the oil in the reservoir thereby causing oil production. Initially when water is injected the amount of oil produced is the amount of water injected. As water displaces oil, it creates a shock front. The shock front is produced because not all saturation waves travel with the same velocity and hence there is a possibility of many waves reaching the same point at the same time implying different saturations at the same point if it is a piston-like displacement. With time, the
A shock front moves to the end of the core and when it does, the water ‘breaks through’ and then there is simultaneous production of oil and water. In other words:

\[
q_t = q_o = q_w \quad t \leq t_d \quad (2.7)
\]

\[
q_t = q_o + q_w \quad t > t_d \quad (2.8)
\]

Where \( q_t \) is the total flow rate coming out of the reservoir

\( q_w \) is the flow rate of water injected

\( q_o \) is the flow rate of oil produced

\( t_d \) is the breakthrough time which is defined as:

\[
t_d = \frac{1}{\frac{df_w}{dS_w}} \quad (2.9)
\]

Where, \( f_w \) is the fractional flow of water and \( S_w \) is water saturation

Because water is injected to produce oil after there is no natural drive, we need to measure how much oil is being displaced. The fractional flow of water \( (f_w) \) gives the amount of oil displaced.
Mobility Ratio (M) = \frac{\text{Mobility}_{\text{Displacing}}}{\text{Mobility}_{\text{Displaced}}} = \frac{k_{rw} \mu_o}{\mu_w k_{ro}} \tag{2.10}

If water is displacing phase and ignoring capillary pressure effects,

\[ f_w = \frac{1 - N_o \sin \alpha}{1 + M} \tag{2.11} \]

\( k_{ro} \) and \( k_{rw} \) are the relative permeabilities of oil and water respectively. From the Corey-type relative permeability curves, these can be represented as:

\[ k_{ro} = (1 - S)^m \tag{2.12} \]
\[ k_{rw} = 0.2(S)^n \tag{2.13} \]

Where \( m \) and \( n \) are Corey parameters and vary depending on rock type

Before breakthrough, the cumulative oil recovery in pore volumes is given by:

\[ N_{pd} = f_o t_d \tag{2.14} \]

Where \( t_d \) is the breakthrough time and \( F_{oil} \) is the oil cut \( (f_o) \) at initial water saturation.
A pore volume is a dimensionless measurement of volume which is given by dividing the volume injected in barrels by total pore volume in barrels.

After breakthrough, cumulative oil recovery in pore volumes is given by:

\[
N_{pd} = S_w - S_{wi} + \frac{1 - f_w}{df_w/dS_w}
\]  

(2.15)

Where \(S_{wi}\) is the initial water saturation.

### 2.1.3 ORDINARY LEAST SQUARE (OLS) REGRESSION

Equation (2.3) specifies a response function. The parameters of this equation can be estimated using the ordinary least squares (OLS) regression method. Consider a general function as:

\[
Y = xb + e
\]

where,

\[
Y = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad x = \begin{bmatrix} 1 & x_{11} & x_{12} & \cdots & x_{1k} \\ 1 & x_{21} & x_{22} & \cdots & x_{2k} \\ \vdots & \vdots & \vdots & \ddots & \vdots \\ 1 & x_{n1} & x_{n2} & \cdots & x_{nk} \end{bmatrix}, \quad b = \begin{bmatrix} b_o \\ b_1 \\ \vdots \\ b_k \end{bmatrix}, \quad \text{and } e = \begin{bmatrix} e_1 \\ e_2 \\ \vdots \\ e_n \end{bmatrix}
\]

\(Y\) is a vector of the measurements, \(x\) is a \(n\)-by-\(K\) matrix of the explanatory variables, \(b\) is a vector of the parameters, and \(e\) is a vector of residuals.
OLS is based on minimizing the sum of the squared residuals. Numerically, it can be represented as:

\[ L = \sum_{i=1}^{n} e_i^2 = e'e = (Y - xb)'(Y - xb) \]

\[ = Y'Y - 2b'x'Y + b'x'xb \]

\( b'x'Y \) is scalar, and thus its transpose \((b'x') = Y'xb\) is also scalar.

The least squares estimators must satisfy:

\[ \frac{\partial L}{\partial b} \bigg|_b = -2x'Y + 2x'xb = 0 \]

which can be simplified as:

\[ x'xb = x'Y \]

The least square estimator can be determined by multiplying both sides of the above equation with the inverse of \((x'x)\). This gives:

\[ b = (x'x)^{-1}x'Y. \]

2.2 Field Development Optimization

Optimizing field development is a complex problem. Estimates are required for reserves and for production rates. In practice, expected values are used to plan a base case, and sensitivity analysis is conducted on the major uncertainties to examine the range of potential outcomes. This analysis yields a set of recommendations on the scope,
configuration, size, and timing of capital investment. The scope of the analysis is often limited by data availability, human resources, time constraints, and bounded rationality. These constraints can lead to suboptimal initial configurations. Well designed workflows and computer aided optimization modeling can support development optimization by relaxing some of the constraints itemized above.

A current challenge in field development optimization is the link between modeling reservoir phenomena and broader (non-subsurface) aspects of field development (wells and facility investment and timing, costs, market dynamics, etc.) This challenge can be met in two ways:

- Establish an explicit link between existing reservoir simulators and an economic optimization model, endogenizing all decision variables;
- Develop simplified representations of reservoir phenomena and workflows that integrate the simplified models into economic optimization algorithms.

There have been efforts to integrate reservoir simulators with optimization algorithms, but no significant headway in this regard has been reported in the literature because of stability and convergence issues. One of the solutions that have been proposed by White et al.\textsuperscript{1,2}, which gives us a significant advantage on accuracy as well as simplicity in implementation in optimization modules, is the use of Response Surfaces, which have been discussed in Section 2.1. The proxy model concept offers the following benefits:
• It allows the analyst to change our reservoir production estimates according to some chosen ‘key’ reservoir parameters.

• Proxy Models are also simple enough to be easily integrated into optimization packages to produce the desired results quickly and efficiently.

• Their accuracy is generally consistent with the reservoir simulation models and adequate to support decision-making.

It is precisely because of this advantage over other methods that Response Surfaces become a critical component of any field development optimization studies and Integrated Asset Models (IAM).

2.2.1 **Value Of Information**

The workflow chart in Fig. 1 implies that RSM methodology would be used to generate proxy models from reservoir simulations to be integrated in an optimization model and Value of Information analysis would be done on the model. The above sections detail the concepts behind DoE, RSM and reservoir simulations. This section elaborates on the concept of VoI.

The value of information (VOI) is defined as the maximum willingness to pay (WTP) for acquiring new information. This information can be perfect or imperfect. Perfect implies that the new information is completely accurate while imperfect implies there is a chance
that the information may be inaccurate. Hence, there is value of perfect information as well as imperfect information.

A simple example best illustrates the concept. Consider a decision situation with one decision *Drill a well* and one uncertainty *Oil Reservoir underneath* which will be resolved only after the *Drill a well* decision has been made. Now,

- Value of perfect information on *Drill a well* captures the value of being able to know *Oil Reservoir underneath* even before making the *Drill a well*. It is quantified as the highest price one is willing to pay for being able to know *Oil Reservoir underneath* before making *Drill a well* decision.
- Value of imperfect information on *Oil Reservoir underneath*, however, depends on another possible related uncertainty, e.g., *Seismic Test*, instead of *Oil Reservoir underneath* itself before making *Drill a well* decision. It is quantified as the highest price one is willing to pay for being able to know *Seismic Test* before making *Drill a well* decision. The assumption here is that *Seismic Test* may be inconclusive.

There are two extremely important characteristics of VOI that always hold for any decision situation;

- VOI can never be less than zero; one can always go ahead without bothering to acquire new information;
• To have value, information must influence a decision; reducing uncertainty is not a sufficient condition for positive VOI

For this thesis, we assume a risk-neutral decision-maker and that the probability of the alternate hypothesis being true to equal 1. VOI is computed using the decision tree in Fig. 2.3: \( \text{VOI} = C - B \).
Figure 2.3  Value of Information Tree
CHAPTER 3: PROBLEM DEFINITION

3.1 RESERVOIR DESCRIPTION

In the scope of this work and in general in the upstream industry, a reservoir is a generic term for a subsurface accumulation of hydrocarbons in porous rock formations. However, production is determined on the basis of economic feasibility which depends on the combination of several factors like presence of a trap and seal, sufficient presence of hydrocarbons and rocks porous enough to allow extraction of oil. An economically feasible reservoir has been assumed for this work.

Oil production in an oil reservoir normally includes three distinct phases:

- Primary recovery
- Secondary recovery
- Tertiary recovery

During primary recovery, the natural pressure of the reservoir or gravity drives the oil into the wellbore to bring the oil to the surface. As a rough estimate, only about 10% of a reservoir’s original oil in place (OOIP) is typically recovered during this phase.
Secondary recovery techniques add to the field’s productive life generally by injecting water to displace oil and drive it to the wellbore. This process is called water flood and results in the recovery of 20-40% of OOIP.

Tertiary or enhanced oil recovery can be initiated to recover an additional 10-20% of the OOIP. There are three major types of tertiary recovery – thermal, gas flood and polymer/surfactant flood. The time to begin secondary or tertiary recovery begins depends on reservoir performance and reservoir engineering and economic analysis.

Light oil reservoirs make good candidates for water/gas flood. The recoverable reserves and production rates from such processes depend on parameters like permeability, porosity, pay thickness, fault boundary, anisotropy etc., each of which can have significant uncertainty associated with it. In some cases, these uncertainties may not effect decision-making in a material way and are harmless. In other cases, the uncertainty may introduce significant engineering and/or economic risk to the project, and mitigation investments may be warranted. Mitigating uncertainty often corresponds to investments in additional analysis and new information. Determining when to make such investment is not a trivial exercise. One objective of this thesis is to develop a workflow to facilitate VOI analysis.

For this work, we concentrate on a reservoir which is a candidate for CO₂ gas flooding.
We define the reservoir as follows:

- Physical Dimensions: 3500 ft. X 3500 ft. X 100 ft.
- Grid Blocks: 7 X 7 X 3
- Grid Block Size: 500 X 500 X (20,30,50)
- Permeability ($K_x$): 200, 50, 500 md
- Permeability ($K_y$): 200, 50, 500 md
- Permeability ($K_z$): 25, 50, 50 mD
- Porosity: 30%
- Compressibility: 5E-06 psi$^{-1}$
- Reservoir Pressure = 3000 psi
- Irreducible Water Saturation = 0.2
- Water Saturation across all grid blocks when CO$_2$ flood begins = 0.3
- Terminal run time = 3000 days
- Minimum bottomhole pressure = 1000 psi
- Wellbore radius = 0.25 ft.
- Skin = 0

The following figures illustrate the reservoir:
Subsurface uncertainties derive from the state of the reservoir and are beyond our control.

In this study, the analysis is limited to three sources of subsurface uncertainty:

1. End-point relative permeability for oil ($K_{rwo}$)

2. Residual oil saturation $S_{orw}$

3. End-point relative permeability for gas ($K_{rg}$)

These three parameters are discussed in detail in the following section.
Design variables are elements of the field development that are under the control of the decision-maker. In this study, the analysis is limited to two design variables:

- Well Pattern
- Gas injection rate

These design variables are discussed further below.

### 3.1.1 Subsurface Uncertainties

As discussed, these uncertainties in subsurface parameters are inherently present in the reservoir and cannot be controlled. Based on our engineering judgment and from literature, we have selected three parameters - $K_{rg}$, $K_{rwo}$ and $S_{or}$. In a green field or virgin field; most of the parameters like pay thickness, permeability and porosity would also be uncertain. Even the viscosity of oil would not be known with certainty and hence the mobility ratio, which is the most important scaling group in a water flood, would be uncertain. Because we are dealing with a mature reservoir, we assume that analysis of the production history has significantly reduced the uncertainty in some of the above parameters. That is, we assume that by the time the reservoir enters tertiary production, most of the reservoir parameters are well described and the remaining uncertainty resides in estimation of the relative permeability curves, especially the end-points of the curve. In simple words, these curves give the dimensionless measure of permeability for each
phase in a multi-phase flow in porous media. A typical relative permeability curve is shown in Fig. 3.2.

Fig. 3.2  Relative Permeability Curve (Courtesy: www.spec2000.net)
The difference in the residual oil saturation and the irreducible water saturation gives us an indication of the ‘recoverable oil window’. Any reduction in the residual oil saturation would enable more oil recovery from the reservoir.

The figure below shows the two relative permeability curves – $K_r$ vs. $S_w$ and $K_r$ vs. $S_g$:

![Relative Permeability Curves for oil and gas](image)

Fig. 3.3 Relative Permeability Curves for oil and gas

The most reliable method to estimate these curves is to perform a core flood on a core of the reservoir rock. However, obtaining a core is expensive and often these curves are
estimated based only on existing data and estimates. As far as these curves are concerned, there are three major uncertainties:

- **Residual Oil Saturation** ($S_{or}$): This point determines where the oil curve (blue curve in Fig. 3.4) terminates. The numerical value of that point is $1 - S_{or}$. For a physical understanding, let us assume an extremely water-wet rock where the surface is covered with water. In that case oil will be located in the centre of the pores. In the laboratory, when water-filled water-wet rock is brought to irreducible or connate water saturation ($S_{cw}$) by oil flooding, this water will remain a continuous phase covering the pore walls. Therefore, in theory, by ongoing oil flooding of a perfect water-wet rock it is possible to "scratch-out" the last bit of water, albeit at infinitely long displacement time, so that $S_{cw} = 0$ for a perfect water-wet system. Consequently, for reservoir rock, a general feature of water-wetness is that $S_{cw}$ is low, say 10-20% or so of the pore space. During a subsequent water-drive to produce the oil, a significant amount of oil eventually will remain capillary trapped, floating as disconnected blobs in the centre of the pores. This remaining oil is measured by the parameter Residual Oil Saturation ($S_{or}$) and is determined by the topology of the pore space and is usually higher than $S_{cw}$: around 20% and up. In practice however, only remaining saturations are reached, because of capillary end-effects or extremely small mobility of the displaced phase.
The lesser the value of $S_{or}$, more mobile oil is available and hence, the oil recovered would be higher. This is possibly the biggest uncertainty in terms of its impact. Another reason that makes this uncertainty important is its variation with surface tension or Capillary Number. Although the manipulation of a favorable $S_{or}$ by changing the interfacial tension has not been very successful on a commercial scale, it does represent an intriguing scientific problem in itself. For our simulations, the value of $S_{or}$ has been varied from 0.1 to 0.2.

- **End-point relative permeability for oil ($K_{ro}^o$):** This point determines the value oil curve taken at the respective connate water or gas saturation. Physically, it represents the ease with which oil moves in comparison to other phases. In the curves shown above, it is given by the maximum value the oil curve would take on the relative permeability axis (y axis). The higher the value, the more oil can flow and hence recovered. For our simulations, the value has been varied from 0.8 to 1.0.

- **End-point relative permeability for gas ($K_{rg}^o$):** This point gives a measure of the permeability of gas relative to the medium. Physically, it represents the ease with which gas phase travels over other phases. The higher it is, gas would move faster in the rock and hence the sweep would be lower. It is better to reduce the value so that gas moves with less ease thereby displacing oil from the pores more efficiently. On the gas-oil relative permeability curve it is given by the end point
of the gas curve (red curve) on the relative permeability axis (x axis). For our simulations, it has been varied from 0.8 to 1.0.

Fluid flow in a reservoir is such a complex process that it would not be possible to define and simulate all uncertainties in our analysis, nor would it be possible to mitigate all of the uncertainty. Ideally, the decision-maker estimates the optimal level of new information and/or mitigation, and accommodates for the remaining uncertainties.

3.1.2 DESIGN VARIABLES

Two design variables are considered: Well Pattern and the Gas Injection Rate. For a \( \text{CO}_2 \) flood, these are two important design decisions. Although gas injection pressure is also important, there is little incremental recovery when gas is injected at a pressure more than its minimum miscibility pressure.\(^{36}\) Therefore, we have decided to concentrate only on two factors.

Well pattern is an important design variable because it determines the number of wells and their configuration. This affects both flood performance and capital costs. If the initial configuration is found to be suboptimal after the subsurface uncertainties are resolved (say, after a few years of production), then changes will be required in the well pattern. At the least, this results in deferred production and a probable loss in NPV. The initial sub optimality may also lead to lower ultimate recovery.
Fig. 3.4 shows some of the different well patterns.

Fig. 3.4 Different Well Patterns

Increasing the number of wells augments production, but also increases costs. The well spacing determines the pressure drop in adjacent wells because the wells affect, or ‘interfere’ with, each other. If the wells are too widely spaced, they could effectively be producing independent of each other and the pattern would cease to have any affect whatsoever on the oil production from wells.
Gas injection rate is another important design variable for gas floods. The compressors required to inject gas into the reservoirs are huge machines and are costly. The benefit of higher rates is that more oil can be swept. A rule of thumb in most gas flooding projects is to inject 10000 cubic ft. per barrel of oil produced per day.

3.2 **Capital Investment Decisions**

Section 3.1 illustrated that uncertainty in key subsurface parameters and design variables can have a significant impact on overall project performance. Analyzing, describing, mitigating and/or accommodating uncertainty prior to making initial capital decisions are therefore important. Some important questions that must be asked when making capital investment decisions could be: “What are the most important subsurface parameters to know before doing a gas flood?”; “What initial pattern or well spacing is optimal given the uncertainty?”; “How will certain uncertainties impact project value?”; or “Is it worthwhile to reduce the uncertainty or to just go ahead with the design on current assumptions?”

As discussed in Chapter 1, a sensitivity analysis of NPV to key uncertainties forms the basis of most capital investment decisions. In many cases, this analysis is sufficient and will not introduce significant suboptimalities. In other cases, especially when there is
significant subsurface or other uncertainty, or when the initial capital investment
decisions are large and quasi-irreversible, a basic sensitivity analysis may be insufficient.

One approach to incorporate more uncertainty into project analysis is to define NPV as a
function of certain key subsurface parameters and then run a Monte Carlo analysis to
estimate the NPV distribution. The advantage of this approach is that it also allows the
analyst to assess uncertainty in a more continuous fashion rather than running a finite
number of cases on the key uncertainties. But in most cases, the implementation of this
approach does not include future options, that is, there is no endogenous optimizing
behavior in the analysis.

When there are options in the future to respond to outcomes (i.e. revealed uncertainties),
a sophisticated model of decision-maker behavior is warranted. A static view of the
project will both undervalue the project and can lead to suboptimal initial investment
decisions. A simple example can demonstrate the value of allowing options in
optimization procedure. Suppose there was a new field discovery five years ago which
had small reserves. The NPV analysis with the then oil price of $20 per barrel would
certainly have rendered the venture uneconomical and hence sold off. However, if the
analysis was done with an option to develop that field five years later if the price went up,
the result could have indicated a profit especially with the oil price now.
When a mature field is to be gas flooded (immiscible CO₂ flood), two important design decisions to be made are the number of wells in a pattern and the gas injection rate. If the subsurface parameters show a wide range of uncertainty, it might be a profitable alternative not to build a large facility initially because there is a chance of small recovery and overinvestment could lead to losses. Intuitively, it makes sense to start with a small initial facility configuration and have the option to expand later. On the other hand, it would be worthwhile to build a facility for the expected case if the uncertainty is less. The idea is to optimize capital investment decisions by deciding which initial well pattern and gas injection rate should be installed given that there is an option to switch to another pattern and rate after $n$ years.
CHAPTER 4: RESERVOIR SIMULATION

The objective of reservoir simulation is to model reservoir phenomena. In this thesis, reservoir simulation is used in an iterative fashion to generate multiple potential production profiles given specific reservoir uncertainties and facility constraints. The reservoir simulation is complex and time consuming. Because of this, the number of simulations is reduced via Design of Experiments. To further simplify and expedite the analysis, the resulting production profiles are then used as inputs to estimate simplified proxy and auxiliary equations that can be easily inputted into a comprehensive project optimization model to determine the optimal initial configuration of the project.

Ultimately, the industry would prefer a solution where the reservoir simulation and the project optimization are completely integrated, and where all of these intermediate steps would not be required. But until the time that such integrated algorithms are developed and available, the sequential workflow as proposed and implemented here is viewed as the second best solution to this type of project optimization problem.

In this Chapter, the results of the reservoir simulation are discussed.
4.1 Reservoir Simulator

A reservoir simulator is a computational tool that honors the fundamental laws of fluid flow in porous media (See Chapter 2.1.2) and with the help of certain simplifying assumption helps in predicting the fluid flow inside the reservoir. It is an important tool in reservoir management and is used extensively in monitoring reservoirs and developing field plans. All values of reservoir parameters and the production constraints can be specified in an easy-to-use graphical interface and the various output parameters can be observed. Currently, a plethora of commercial simulators are available for use. For this work, CMG/STARS was chosen because it is not too complex to use, is used by practicing reservoir engineers, and is efficient in modeling water and gas floods.

4.2 Design of Experiments

A discussion of DoE has already been provided in Chapter 2. We have decided to use Box-Behnken Design. This is largely because:

- It requires a reasonable number of simulations – for three parameters, thirteen simulation runs are required
- It accounts for quadratic effects resulting from interaction between parameters
- It does not require points to be taken outside the given range of parameters.
4.3 Specification of the Simulations for Each Case

In addition to the three uncertain subsurface parameters defined above: \( S_{or} \), \( K_{pwo} \) and \( K_{rg} \), an additional variable of Water Saturation, \( S_w \) has also been included in the Design of Experiments to allow for re-initialization of the reservoir for the economic optimization. This aspect will be discussed in more detail in Chapter 7. The PVT properties have been kept constant for all cases. The terminal run time of the simulation is 3000 days. The oil production values are generated at equi-spaced time intervals of 100, 200, 300,\ldots, 2900, 3000 days. The simulations produce estimates of cumulative production and auxiliary data that is required to control specific aspects of the economic optimization.

The following tables define the cases that were run:

<table>
<thead>
<tr>
<th>Case</th>
<th>Krow</th>
<th>Sor</th>
<th>Krg</th>
<th>Sw</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>-1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>3</td>
<td>-1</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>4</td>
<td>-1</td>
<td>-1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>5</td>
<td>0</td>
<td>1</td>
<td>-1</td>
<td>0</td>
</tr>
<tr>
<td>6</td>
<td>0</td>
<td>1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>-1</td>
<td>1</td>
<td>0</td>
</tr>
<tr>
<td>8</td>
<td>0</td>
<td>-1</td>
<td>-1</td>
<td>0</td>
</tr>
<tr>
<td>9</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>-1</td>
</tr>
</tbody>
</table>
Table 4.1  BBD Table

The values of the variables have been divided into three levels – High denoted by 1, Medium denoted by 0 and Low denoted by -1. The following table shows the individual values taken for each variable in a simulation run:

<table>
<thead>
<tr>
<th>Case</th>
<th>Krow</th>
<th>Sor</th>
<th>Krg</th>
<th>Sw</th>
</tr>
</thead>
<tbody>
<tr>
<td>10</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>11</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>1</td>
</tr>
<tr>
<td>12</td>
<td>0</td>
<td>0</td>
<td>-1</td>
<td>-1</td>
</tr>
<tr>
<td>13</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>-1</td>
</tr>
<tr>
<td>14</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>15</td>
<td>-1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>16</td>
<td>-1</td>
<td>0</td>
<td>0</td>
<td>-1</td>
</tr>
<tr>
<td>17</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Case</th>
<th>Krow</th>
<th>Sor</th>
<th>Krg</th>
<th>Sw</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>0.1</td>
<td>0.9</td>
<td>0.325</td>
</tr>
<tr>
<td>2</td>
<td>1</td>
<td>0.2</td>
<td>0.9</td>
<td>0.325</td>
</tr>
<tr>
<td>3</td>
<td>0.8</td>
<td>0.2</td>
<td>0.9</td>
<td>0.325</td>
</tr>
<tr>
<td>4</td>
<td>0.8</td>
<td>0.1</td>
<td>0.9</td>
<td>0.325</td>
</tr>
<tr>
<td>5</td>
<td>0.9</td>
<td>0.2</td>
<td>0.8</td>
<td>0.325</td>
</tr>
<tr>
<td>6</td>
<td>0.9</td>
<td>0.2</td>
<td>1</td>
<td>0.325</td>
</tr>
<tr>
<td>7</td>
<td>0.9</td>
<td>0.1</td>
<td>1</td>
<td>0.325</td>
</tr>
</tbody>
</table>
Table 4.2  Values of parameters used for simulation runs

<table>
<thead>
<tr>
<th>Case</th>
<th>Krow</th>
<th>Sor</th>
<th>Krg</th>
<th>Sw</th>
</tr>
</thead>
<tbody>
<tr>
<td>8</td>
<td>0.9</td>
<td>0.1</td>
<td>0.8</td>
<td>0.325</td>
</tr>
<tr>
<td>9</td>
<td>0.9</td>
<td>0.15</td>
<td>1</td>
<td>0.3</td>
</tr>
<tr>
<td>10</td>
<td>0.9</td>
<td>0.15</td>
<td>1</td>
<td>0.35</td>
</tr>
<tr>
<td>11</td>
<td>0.9</td>
<td>0.15</td>
<td>0.8</td>
<td>0.35</td>
</tr>
<tr>
<td>12</td>
<td>0.9</td>
<td>0.15</td>
<td>0.8</td>
<td>0.3</td>
</tr>
<tr>
<td>13</td>
<td>1</td>
<td>0.15</td>
<td>0.9</td>
<td>0.3</td>
</tr>
<tr>
<td>14</td>
<td>1</td>
<td>0.15</td>
<td>0.9</td>
<td>0.35</td>
</tr>
<tr>
<td>15</td>
<td>0.8</td>
<td>0.15</td>
<td>0.9</td>
<td>0.35</td>
</tr>
<tr>
<td>16</td>
<td>0.8</td>
<td>0.15</td>
<td>0.9</td>
<td>0.3</td>
</tr>
<tr>
<td>17</td>
<td>0.9</td>
<td>0.15</td>
<td>0.9</td>
<td>0.325</td>
</tr>
</tbody>
</table>

4.4  SIMULATION RESULTS

The above sets of seventeen cases correspond to a particular set of well pattern and injection rate. Three different injection rates (4, 5, and 6 MSCF/D) and two different patterns (5- and 7-spots) have been considered as the possible design alternatives.

While there is no restriction on the number of design alternatives, injection rates of 4, 5, and 6 MSCF/D and 5 and 7 spot patterns have been considered as the possible design alternatives in this thesis. In all, there were 102 simulations run – seventeen for each case.
of injection rate and well pattern. Figure 4.1 shows the production profiles given by different combinations of wells and injection rates for Case 1 as mentioned in Table 4.2:

![Comparison of cumulative oil production for different parameters (Case 1)](image)

**Fig. 4.1** Simulation Results
Similar production profiles were obtained for other cases. The following figures illustrate how the reservoir looked at terminal time of flooding:

Fig. 4.2  Reservoir at the end of flooding

The above figures illustrate how the oil has been swept out of the reservoir for one of the cases run. Initially, the reservoir is at a given initial water saturation. CO\textsubscript{2} is injected from the four injectors at each corner of the reservoir and it then displaces the recoverable oil. As the gas saturation figure indicates, the oil from the top most layer (the most
permeable) has been displaced to the point of residual oil saturation. However, the middle layer has a permeability of just 50 mD and hence, the sweep efficiency is not that good in the region. The simulation results show some areas being virtually unswept as indicated by the oil saturation at the end of simulation being the same as initial oil saturation for some blocks in Fig. 4.2. Even though the basic idea behind doing such reservoir simulations was to get the approximate estimate for oil production from the field, such simulation results also give us an idea of the sweep efficiency of the enhanced oil recovery process.
CHAPTER 5: PROXY MODEL ESTIMATION

In this Chapter, we use the simulation results to estimate simple proxy models for cumulative production that will be used in the economic optimization. We also estimate several auxiliary equations. The estimation is accomplished with ordinary least squares regression (OLS) already discussed in Chapter 2. In this Chapter we provide the following:

(i) a brief history of proxy modeling in oil industry,
(ii) a description of how the current method compares to earlier work,
(iii) model specifications,
(iv) estimation results.

5.1 BRIEF HISTORY OF PROXY MODELS

Proxy Reservoir models are simplified mathematical expressions which give reasonable results as would be given by a reservoir simulator. The mathematical expressions are formed by linear or non-linear regressions on important reservoir parameters. One of the criticisms of proxy model has been that it is not the true representation of the physics underlying the fluid flow. However, the great advantage is that they are simplified expressions and provide a way to circumvent the very difficult problem of integrating an optimization module with a reservoir simulator. This flexibility has made proxy models a vital component of most field development optimization processes in industry.
Though proxy models have been used as an accepted method in various other fields, they were not used in the oil and gas industry problems till the early 90s. Here is a brief history on the usage of proxy models in E&P industry:

- Chu⁹ was one of the first papers to use a full factorial design on pay thickness, porosity, permeability, oil viscosity, oil saturation, area, steam injection and steam quality to predict steam flood performance in a heavy oil reservoir. However, full-factorial design is rarely used because of the large number of computations required. Damsleth et al¹⁰ used D-optimal design to model recovery as a function of vertical permeability, skin and relative permeability for a North Sea development study. In many ways, this research set the trend for using proxy models for developmental studies in exploration and production. Egeland et al¹¹ used a proxy model on the same parameters as Damsleth to illustrate better decision making methods but they did not use any particular Design of Experiments.

- Dejean and Blanc¹² detailed how to manage uncertainties using methods like design of experiments and statistics. They used a Face Centred Design on reservoir length, fault transmissibility, aquifer and oil well co-ordinates. Although they used a proxy model for cumulative recovery, they did not have a time variable and used a separate model for cumulative recovery at certain time steps. A similar strategy was used by Corre et al¹³ who used D-optimal design to develop models for cumulative oil recovery at certain time steps. They used eight
parameters – permeability, fault transmissibility, PVT properties, relative permeability, fracture pressure, injector skin and position of horizontal barriers. Kabir et al\textsuperscript{15} also followed the above approach and chose Plackett-Burman Design to model recovery factor at the end of 5,10 and 15 years. They considered seven parameters – well type, number of wells, permeability, residual oil saturation, anisotropy multiplier, well type and injector location. They justified using Plackett-Burman method of design of experiments on account of having many parameters. Plackett-Burman design is popular chiefly because it is requires fewer computational resources.

- Aigbe\textsuperscript{4} et al. used Plackett-Burman it to quantify uncertainties in a little produced Nigerian field using just six parameters – permeability ratio, anisotropy multiplier, fault integrity, residual oil saturation, Oil viscosity and end-point relative permeability. Another approach was illustrated by Peake et al.\textsuperscript{17} They used Plackett-Burman Design on twelve parameters - pay thickness, porosity, horizontal & vertical Permeability, oil-water contact, irreducible water saturation, geologic contrast, sealing faults, tar zone thickness, rock compressibility, oil properties and relative permeability. This was used to shortlist parameters on the basis of their significance and a D-optimal design was used to regress on cumulative oil recovered. A similar approach was also used by Li and Friedmann\textsuperscript{7}, Gupta et al.\textsuperscript{16} and Carreras et al.\textsuperscript{3} Li and Friedmann used a proxy model on cumulative oil recovery to do an uncertainty analysis on reservoir simulation forecasts. After short listing the parameters, they did a D-optimal
design on five parameters – permeability, aquifer pore volume, permeability multiplier, viscosity and fault transmissibility. Gupta et al performed water flooding screening and sensitivity analysis using this approach. They used Plackett-Burman design on fourteen parameters - Gross Rock Volume, Sandy Facies proportion, Channel Belt width, Channel-belt sinuosity, Channel width-thickness ratio, Permeability, Petrophysics variogram range, aquifer strength, mobility ratio, End-point saturation, Vertical/Horizontal permeability, Compressibility and production wells. The preliminary analysis allowed them to screen parameters with a significant effect and then use a three level fractional factorial design on the cumulative recovery after 25 years. Carreras et al.\(^3\) made an assessment of uncertainty in the Gulf of Mexico Tahiti reservoir using this approach. They performed a Plackett-Burman Design of Experiments on fourteen parameters - Earth model, Salt-face location, WOC location, Faulting and compartmentalization, \(K_v/K_h\), Aquifer support, Water-oil relative permeability, Pore volume compressibility, rock compaction & dilation, Fluid characterization, Skin Factor and Well Pressure Drawdown. This analysis revealed four significant parameters – Earth Model, Faults, Water-oil relative permeability and oil-water contact. A D-optimal design was then performed on these parameters.

- Some important white papers on use of proxy model have also been given by White et al.\(^1,2\) The authors showed the identification of important parameters using a Box-Behnken Design of Experiments and how well location in a turbidite reservoir could be optimized. One of the important features was showing the
orthogonality of the elements of the proxy model which is an important assumption in developing the model.

- Of late, people using proxy models have been regressing on NPV rather than recovery factor or cumulative oil recovery at terminal time as shown by Esmaiel et al.⁷ and Yamali et al.⁸. Esmaiel et al.⁷ used a D-optimal design on four parameters – permeability ratio streak, relative pore volume of streak, revenue of oil per barrel and pore volume injected per day to optimize for smart wells while Yamali et al. used a Central Composite Design on normalized pressure values at different values to reduce unwanted water production in stratified gas reservoirs and hence optimize NPV.

Besides quantifying uncertainties for field development, response surface methodology and Design of Experiments have been applied to several other reservoir engineering applications including performance prediction and sensitivities studies.

To summarize, there is one underlying theme in all the proxy models developed so far. Most of the models have been developed using Design of Experiments and Response Surface methodology. Design of Experiments methodology is basically a set of simulations run in a statistically pre-defined and correct manner. Plackett-Burman design is the most widely used because of the fewer computations involved. Because it does not account for interaction between parameters and can only be used for a linear regression, others⁵,¹⁵,¹⁶,¹⁷ use this method to select the most important parameters and then use D-
optimal Design\textsuperscript{3,5,7,10,13,16,17} on those selected parameters. Box-Behnken method\textsuperscript{2} is also used in conjunction with Plackett-Burman Design. Another important observation from the literature is that all models only predict ultimate oil recovery\textsuperscript{7,12,16} or recovery factor\textsuperscript{2,9,10} or NPV\textsuperscript{5,8} at terminal time and do not give predictions on production at specific intervals. Some papers\textsuperscript{12,13,15,17} however generate proxy model for cumulative production/recovery factor after specific time intervals and study the effect of production on parameters with time. However, when we are modeling real options and trying to find value of information, it would be helpful to find the NPV at all times which would require estimating a proxy model that can predict production as a function of time.

\section*{5.2 Model Specification And Results}

Initially, we tried to specify instantaneous oil production rate as a function of reservoir parameters:

\[ q(t) = f(S_{or}, K_{rw}, K_{rg}) \]

However, there were cases that exhibited an initial plateau region and hence Ordinary Least Square (OLS) method was not appropriate because in the plateau region we get the same production regardless of time or reservoir parameters. Even when the production was unconstrained so as to get rid of the plateau region, the curve generated was not smooth enough to give a reasonably good fit. Hence, it was decided to regress on
cumulative oil production values because the curve is smoother. Also because the underlying statistics of time series panel data are not trivial and are outside the scope of this thesis (see Appendix A.3), a more conventional approach is adopted here:

\[
Q_{\text{cum}} = f(S_{or}, S_w, K_{rw}^o, K_{rg}^o) \tag{5.1}
\]

where,

- \(Q_{\text{cum}}\) is the cumulative oil production in STB/D
- \(S_{or}\) is the residual oil saturation
- \(S_w\) is the water saturation
- \(K_{rw}^o\) is the end-point relative permeability for water
- \(K_{rg}^o\) is the end-point relative permeability for gas

The regressions were done for cumulative oil produced at 1000 days, which is the time the engineer may or may not change the facility configuration, and at 2000 days, the time between terminal time and the time of switching or reconfiguration (1000 days). The model is the same as Equation (5.1). In this thesis, we do not include estimation of time series proxy models. A proposal for such models is included in Appendix A.3. The following results were obtained for the 1000 and 2000 days models:
Cumulative oil production after 1000 days:

- 5 spot pattern & 4 mSCF/D:
  \[ 8719290 + 8050075*K_{ro} - 2842775*S_{or} - 408650*K_{rg} - 9218550*S_w - \\
  2124850*K_{ro} * K_{ro} - 3344000*S_{or} * S_{or} - 1232500*K_{rg} * K_{rg} + 0*S_w * S_w + \\
  235000*K_{ro} * S_{or} + 0*K_{ro} * K_{rg} - 7246000*K_{ro} * S_w + 1730500*S_{or} * K_{rg} + 0*S_{or} * S_w + \\
  5053000*K_{rg} * S_w \\
  \]

- 5 spot pattern & 5 mSCF/D:
  \[ 8959539 + 9859700*K_{ro} + 452475*S_{or} + 779875*K_{rg} - 8542100*S_w - \\
  2744938*K_{ro} * K_{ro} - 6159500*S_{or} * S_{or} - 1332563*K_{rg} * K_{rg} + 0*S_w * S_w - \\
  981000*K_{ro} * S_{or} + 0*K_{ro} * K_{rg} - 8971000*K_{ro} * S_w - 325500*S_{or} * K_{rg} + 0*S_{or} * S_w + \\
  4025000*K_{rg} * S_w \\
  \]

- 5 spot pattern & 6 mSCF/D:
  \[ 13642775 + 1959875*K_{ro} + 994000*S_{or} - 1020125*K_{rg} - 2868000*S_w + \\
  1378750*K_{ro} * K_{ro} - 1860000*S_{or} * S_{or} + 1386250*K_{rg} * K_{rg} + 0*S_w * S_w - \\
  4290000*K_{ro} * S_{or} + 0*K_{ro} * K_{rg} - 6050000*K_{ro} * S_w + 330000*S_{or} * K_{rg} + 0*S_{or} * S_w - \\
  8770000*K_{rg} * S_w \\
  \]

- 7 spot pattern & 4 mSCF/D:
  \[ 16518914 + 1918338*K_{ro} + 12146375*S_{or} - 17000000*K_{rg} + 6230600*S_w + \\
  6270687*K_{ro} * K_{ro} - 20000000*S_{or} * S_{or} + 7697312*K_{rg} * K_{rg} + 0*S_w * S_w - \\
  8704000*K_{ro} * S_{or} + 0*K_{ro} * K_{rg} - 30000000*K_{ro} * S_w - 350500*S_{or} * K_{rg} + 0*S_{or} * S_w + \\
  7506000*K_{rg} * S_w \\
  \]
• 7 spot pattern & 5 mSCF/D:

\[ 10399412 + 12061500 K_{ro} + 1108500 S_{or} + 929125 K_{rg} - 17000000 S_{w} - \]
\[ 4943125 K_{ro} K_{ro} + 15755000 S_{or} S_{or} - 754375 K_{rg} K_{rg} + 0 S_{w} S_{w} + \]
\[ 3090000 K_{ro} S_{or} + 0 K_{ro} K_{rg} - 3130000 K_{ro} S_{w} - 13000000 S_{or} K_{rg} + 0 S_{or} S_{w} + \]
\[ 1520000 K_{rg} S_{w} \]

• 7 spot pattern & 6 mSCF/D:

\[ 7225500 + 31606125 K_{ro} + 1516000 S_{or} - 4427875 K_{rg} - 31000000 S_{w} - \]
\[ 12000000 K_{ro} K_{ro} + 48230000 S_{or} S_{or} - 24850000 K_{rg} K_{rg} + 0 S_{w} S_{w} - \]
\[ 10000000 K_{ro} S_{or} + 0 K_{ro} K_{rg} - 13000000 K_{ro} S_{w} - 9710000 S_{or} K_{rg} + 0 S_{or} S_{w} + \]
\[ 25710000 K_{rg} S_{w} \]

Cumulative oil production at 2000 days:

• 5 spot pattern & 4 mSCF/D:

\[ 11023312 + 16720125 K_{ro} - 7273500 S_{or} + 4451000 K_{rg} - 20000000 S_{w} - \]
\[ 6375625 K_{ro} K_{ro} + 6505000 S_{or} S_{or} - 4021875 K_{rg} K_{rg} + 0 S_{w} S_{w} - \]
\[ 55000 K_{ro} S_{or} + 0 K_{ro} K_{rg} - 7230000 K_{ro} S_{w} + 455000 S_{or} K_{rg} + 0 S_{or} S_{w} + \]
\[ 2200000 K_{rg} S_{w} \]

• 5 spot pattern & 5 mSCF/D:

\[ 19062475 + 16672250 K_{ro} - 5612000 S_{or} - 2688250 K_{rg} - 37000000 S_{w} - \]
\[ 7256250 K_{ro} K_{ro} + 590000 S_{or} S_{or} - 1648750 K_{rg} K_{rg} + 0 S_{w} S_{w} - \]

63
\[
215000*K_{ro} * S_{or} + 0*K_{ro} * K_{rg} - 190000*K_{ro} * S_{w} - 445000* S_{or} * K_{rg} + 0* S_{or} * S_{w} + \\
9130000*K_{rg} * S_{w}
\]

- 5 spot pattern & 6 mSCF/D:
\[
15736087 + 19865375*K_{ro} - 2971250*S_{or} - 817250*K_{rg} - 13000000*S_{w} - \\
2924375*K_{ro} * K_{ro} + 4375000*S_{or} * S_{or} - 2085625*K_{rg} * K_{rg} + 0*S_{w} * S_{w} - \\
13000000*K_{ro} * S_{or} + 0*K_{ro} * K_{rg} - 26000000*K_{ro} * S_{w} + 6350000*S_{or} * K_{rg} + \\
0*S_{or} * S_{w} - 510000*K_{rg} * S_{w}
\]

- 7 spot pattern & 4 mSCF/D:
\[
13816875 + 14829250*K_{ro} + 19985250*S_{or} - 23000000*K_{rg} + 39379500*S_{w} + \\
9071250*K_{ro} * K_{ro} - 30000000*S_{or} * S_{or} + 11286250*K_{rg} * K_{rg} + 0*S_{w} * S_{w} - \\
22000000*K_{ro} * S_{or} + 0*K_{ro} * K_{rg} - 72000000*K_{ro} * S_{w} + 4115000*S_{or} * K_{rg} + \\
0*S_{or} * S_{w} - 2170000*K_{rg} * S_{w}
\]

- 7 spot pattern & 5 mSCF/D:
\[
16894900 + 21748175*K_{ro} - 6786000*S_{or} - 1252875*K_{rg} - 26000000*S_{w} - \\
5563750*K_{ro} * K_{ro} - 5580000*S_{or} * S_{or} - 3576250*K_{rg} * K_{rg} + 0*S_{w} * S_{w} - \\
8010000*K_{ro} * S_{or} + 0*K_{ro} * K_{rg} - 20000000*K_{ro} * S_{w} + 8230000*S_{or} * K_{rg} + 0*S_{or} * S_{w} + \\
9250000*K_{rg} * S_{w}
\]

- 7 spot pattern & 6 mSCF/D:
\[
-1402150 + 33126000*K_{ro} + 21269500*S_{or} + 10728750*K_{rg} + 16173000*S_{w} - \\
13000000*K_{ro} * K_{ro} - 29000000*S_{or} * S_{or} + 2303750*K_{rg} * K_{rg} + 0*S_{w} * S_{w} - \\
44000000*K_{ro} * S_{or} + 0*K_{ro} * K_{rg} + 7280000*K_{ro} * S_{w} + 23850000*S_{or} * K_{rg} + \\
0*S_{or} * S_{w} - 69000000*K_{rg} * S_{w}
\]
CHAPTER 6: ECONOMIC OPTIMIZATION

The goal of field development optimization is to determine the development plan that will enable the operator to extract the maximum potential value from an asset. In this thesis, we focus on two initial capital investment decisions: well pattern and injection rate. This decision making process is implemented by linking detailed technical analyses via a logical workflow. So far, we have described several of these steps: Design of Experiments, reservoir simulation, and proxy model estimation. The final step is to use the proxy equations to represent reservoir behavior within a comprehensive field development optimization model. The optimization model was written in GAMS (General Algebraic Modeling Systems).

Any generic GAMS model consists of five basic components: Sets, Variables, Parameters/Tables, Scalars and Equations. Within GAMS, sets are equivalent to subscripts in algebra and are a series of items that can be simultaneously operated over; including summed over or looped over among other possibilities. Parameters and Tables are used to enter items defined with respect to sets. Scalars define values which remain constant throughout the optimization while Variables assume changing values. Equations define the optimum, relationships, and constraints.
6.1 Algorithm

The following is the step-by-step procedure used to perform the optimization:

- Define Pattern and Injection Rate as sets
- Define $S_{or}$, $K_{ro}$, $K_{rg}$, $S_w$ and $S_w'$ (Water Saturation at time of altering patterns) as parameters
- Define Cumulative oil Production at $t=1000$ and $t=2000$ days as Parameters e.g. $Q1$ (Pattern, Injection Rate) and $Q2$ (Pattern, Injection Rate) respectively
- Define Switching/Reconfiguration time and terminal run time
- Define the twelve proxy model equations as a function of $S_{or}$, $K_{ro}$, $K_{rg}$, $S_w$ and $S_w'$
  e.g.

\[
Q1(5\text{spot}, 4mSCF / D) = f(S_{or}, S_w, K_{ro}^{\alpha}, K_{rg}^{\alpha})
\]

\[
Q2(5\text{spot}, 4mSCF / D) = f(S_{or}, S_w', K_{ro}^{\alpha}, K_{rg}^{\alpha})
\]

- Define a Dummy Variable array DV as a function of Pattern and Injection Rate
- Initialize Cumulative Oil (QC)
- Cumulative Oil in the post switching regime is given by: $DV \times Q2$, where $Q2$ is computed on the basis of updated water saturation, $S_w'$
- Sum of all entries of DV matrix equals 1
- $NPV = NPV(1000 \text{ days}) + NPV(3000 \text{ days})$
- Run a loop on $n$ – the number of trials by a random variable
• In the loop, define a distribution for $K_{rg}$, $K_{ro}$, $S_{or}$ and $S_w$ and then define

\[ S'_w = S_w + \frac{QC}{Area \times PayThickness \times Porosity} \]  

according to the mass balance and in consistent units.

A Mixed Integer Non-linear Programming Solver (MINLP) is invoked in GAMS to solve the above problem. The final output gives us the optimized DV matrix which indicates whether to switch or not and also the maximized NPV values and the corresponding probabilistic distribution for NPV.

### 6.2 Model Equations

The optimization module assumes the following:

• The decision maker is a risk-neutral profit (NPV) maximizer

• For simplicity of calculation, the cash flows are calculated at only two times – the terminal time and the time when we decide to reconfigure

• The produced natural gas generated is not taken into account to calculate NPV

• There is no recycling of gas injected
• The saturations are constant across all grid blocks and only change with time. This is physically incorrect but in terms of the numerical output, we find that there is no significant difference.

The various parts of the optimization module are described as under:

6.2.1 **Objective Function**

\[
z = \sum \left[ (p(t) \times q(t) - \text{opex}(t) - \text{capex}(t)) \times \frac{1}{(1 + r)^t} \right]
\]  

(6.2)

where,

z is the Net Present Value in $

p(t)$ is the average oil price in $/bbl which is held constant at $60/bbl

q(t) is the cumulative oil production in STB

opex(t) is the operating expense at time t in $

capex(t) is the capital expense at time t in $

r is the discount rate which is kept constant at 15%
6.2.2 **Cost Functions**

Costs are comprised of operating expenses and capital investment in initial facilities, facility expansion and wells. We have found the two costs at two different times – the time of reconfiguration (which is 1000 days in our case) and terminal time (3000 days in our case). As mentioned in the algorithm in Section 6.1, we have introduced a Dummy Variable DV as a function of Pattern (defined as ‘pat’) and Injection rate (defined as ‘IRate’). This Dummy Variable assumes only Binary Variables (0 & 1) and the optimization run assigns a value of 1 to the optimum reconfiguration and the value of 0 to the rest of the combinations. A sample value of DV after one of the runs would look like:

\[
\begin{array}{ccc}
4 & 5 & 6 \\
5 & 0 & 0 & 0 \\
7 & 1 & 0 & 0 \\
\end{array}
\]

The columns indicate the Injection Rate in mSCF/D and the rows indicate the type of pattern. The particular above result shows that the optimum configuration in the post switching regime is a 7 spot pattern and 4 mSCF/D injection rate.

**Capital Expense (capex)**

Capital expense when the system is initialized is composed of two components as follows:
\[ \text{Capex} = C_{\text{well}} + C_{\text{plat}} \]

where,

- \( C_{\text{well}} \) is the cost of drilling the wells
- \( C_{\text{plat}} \) is the cost of building the infrastructure or support facility

This \( (C_{\text{plat}}) \) is defined as:

\[ C_{\text{plat}} = b_0 + b_1 \left( \frac{\text{cap}_0}{365} \right)^{e_{\text{plat}}} \]  \hspace{1cm} (6.3)

where,

- \( b_0, b_1 \) and \( e_{\text{plat}} \) are cost paremeters
- \( \text{cap}_0 \) is the Initial Facility Capacity in mbopy

The cost of drilling wells is defined as:

\[ C_{\text{well}} = n_{\text{wells}} \times f_{\text{cwell}} \]  \hspace{1cm} (6.4)

where, \( n_{\text{wells}} \) is the number of wells drilled found depending on the pattern while \( f_{\text{cwell}} \) is the cost of drilling a well which is fixed at 3 million $/well.

Once the field has been reconfigured after a certain time (1000 days in our case), the cost of building a platform is not included as it is only an initial cost. There is no cost
associated with abandoning a well and only incremental cost would be drilling a new well if required and the cost to increase the injection rate by adding more compressor trains. That would depend on the initial facility design and the design switched to later. For example, if we initialize with a 5 spot and 4 mSCF/D initialized and then we switched to a 7 spot and 6 mSCF/D configuration, then the increase in capital expenditure would be the cost of drilling two wells and the cost of increasing the injection capacity of wells by an additional 2 mSCF/D. On the other hand, if we need to switch from 7 spot and 6 mSCF/D to 5 spot and 4 mSCF/D, there would be no incremental cost as we just have to abandon a couple of wells and compressor trains. This has been incorporated in the code in the form of a table called CapexSwitch which uses different values depending on the initial configuration. For a full list of values generated for different configurations, see Appendix A.1

**Operating Expense (opex):**

*Opex* represents the variable cost of production. In the field, we are injecting gas and producing oil. Hence, the operating cost will also be a function of the production volume as well as the amount of gas injected. Operating expenses while producing oil would include cost of manhours, equipment operating costs and maintenance costs. All these factors have been taken into account by defining a variable *vcprod* which when multiplied by the cumulative oil produced gives us the operating cost. The other major contributor of operating cost is the gas injection. Had it been constant, it could have been incorporated by a proportionate change in the variable *vcprod*. However, in this thesis,
the gas injection is not constant and hence, that component is defined separately. A variable $f_{copex}$ is used which when multiplied by the total gas injection would give the operating cost for gas injection. The variable $f_{copex}$ is a function of the pumping, transportation, electricity and other operation costs for compressor trains used for gas injection. The total operating cost can thus be defined as:

$$\text{opex} = v_{prod} \times Q_{cum} + f_{copex} \times \text{GasInj} \quad (6.5)$$

where,

$v_{prod}$ is the variable cost associated with oil production

$Q_{cum}$ is the cumulative oil produced in STB

$\text{GasInj}$ is the total gas injected in mSCF/D from all the wells

$f_{copex}$ is the operating cost in $ for every mSCF of gas injected.

The total gas injection defined by the variable $\text{GasInj}$ is a function of the well pattern and the amount of gas being injected from each well. Each facility configuration (x,y) corresponds to a different value of $\text{GasInj}$.

Once reconfiguration is done, equation (6.5) is modified to calculate the $\text{opex}$ accordingly:
\[ \text{opex} = \text{vcprod} \times \sum_{\text{pat}} \sum_{\text{Irare}} \text{DV} \times Q_{\text{cum}} + \text{fcopex} \times \text{GasInj} \] (6.6)

The \( Q_{\text{cum}} \) value is given by the proxy models defined in Chapter 5 while \( DV \) is the dummy variable defined earlier to facilitate reconfiguration calculations.

### 6.2.3 Reconfiguration Equations

As discussed in Section 6.1, the reconfiguration or switching is done in the optimization model by re-initialization of the proxy models. The reservoir has been water flooded and the point when the gas flooding design begins, the time zero of our model, the water saturation in the reservoir is \( S_w \), which is given by a normal distribution like the other three uncertain variables discussed in Chapter 4. After the reservoir has produced for 1000 days, the saturations will change. We cannot simply use the proxy model that gives the cumulative oil produced at 3000 days and use the initial \( S_w \). To overcome this hurdle, we suggest using a simple overall material balance to find the new \( S_w \) and re-initialize the proxy model with that value \( S_w' \) as described in equation (6.1):

\[
S_w' = S_w + \frac{5.615 \times Q_{\text{cum}}}{V \times \phi \times B_o}
\]

where,

\( V \) is the Reservoir Volume in bbls
\( \phi \) is the porosity

\( B_o \) is the formation volume factor

This value of \( S_w' \) when entered in place of \( S_w \) in the proxy model that gives the cumulative oil production at 2000 days gives the cumulative oil produced between time of reconfiguration and terminal time (1000 days and 3000 days for our case).

There is only one constraint while optimizing the reconfiguration that only one pattern or injection rate can be used at one time. In terms of numerical definition, it is given by:

\[
\sum \sum DV = 1 \quad \text{(6.7)}
\]

If one was to look at this in terms of the physical processes involved, it is a simplification of the actual processes. However, the objective of defining a proxy model is not to have a perfect representation of reservoir physics but rather to find a middle ground to facilitate implementation. Because the proxy models are estimated based on results of a reservoir simulator, this approach is reasonable. Also, the focus of this thesis is on workflow development, and the scope of the reservoir and facility modeling was constrained so that emphasis could be placed on finding solutions to problems in the workflow.
All of the necessary components to implement an integrated field development optimization process have been defined: Design of Experiments to define the scope of the analysis, reservoir simulation to generate production forecasts, regression analysis to estimate proxy reservoir models, and optimization modeling to simulate decision-maker behaviors. Together, these components define a workflow that can be used to answer various questions and support recommendations for initial capital investment and uncertainty mitigation. In this section, we discuss how to use the model to find the optimum course of action and value of information as discussed in Section 2.4.

7.1 Planning With Current Information (Determining Optimal Initial Facility Configuration)

This is the first step in the decision making process. A reservoir engineer has some data regarding subsurface parameters and their uncertainty. If additional information will not be obtained, the task is to determine the optimal course of action given the current information. In practice, expected values are used to plan a base case, and sensitivity analysis is conducted on the major uncertainties to examine the range of potential outcomes. This analysis yields a set of recommendations on the scope, configuration, size, and timing of capital investment. The scope of the analysis is often limited by data availability, human resources, time constraints, and bounded rationality.
A more sophisticated method is to define each uncertain variable as a probability density function (PDF) and then do a Monte Carlo Analysis on those variables and choose the option with the maximum expected value of NPV. However, this is still not widely used as it requires considerable time and resources. This approach should deliver a more robust description of the range of possible outcomes than a finite case-based sensitivity analysis type of approach.

A different approach can lead to even better results. Let us consider a base case $\Omega_1$ where each of the subsurface variables is defined as a normal probability distribution:

$$K_{or} = \text{normal} (0.9, 0.02)$$

$$K_{rg} = \text{normal} (0.9, 0.02)$$

$$S_{or} = \text{normal} (0.15, 0.015)$$

$$S_{w} = \text{normal} (0.35, 0.02)$$

where, \(\text{normal}(x,y)\) denotes a distribution with a mean of \(x\) and standard deviation of \(y\)

As defined above, the design variables to be investigated are well pattern (5 wells or 7 wells) and injection rate. Using Design of Experiments, subsurface cases are defined as explained above. Reservoir simulation is performed given each combination of initial well pattern and injection rate. These results are used to estimate proxy models which are
entered into the optimization model. Recall that the optimizer allows the decision-maker the option to switch from the initial configuration to an alternate configuration.

For example, assume the field is initialized with a 5 spot pattern and 4 mSCF/D. When the optimizer solves, the subsurface uncertainties are sampled from the predefined PDFs, production commences, the option to change configuration is exercised or not, and the iteration produces a final facility configuration and associated NPV. If this process is repeated many times, an empirical PDF of NPV can be developed for this case of initial configuration. Note, the final configuration in any iteration is not necessarily the same as the initial configuration; it depends on whether or not the option to switch is exercised. Similar results are obtained for other cases and are depicted in Table 7.1.

<table>
<thead>
<tr>
<th>(P,I)</th>
<th>Case A</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
</tr>
<tr>
<td>5,4</td>
<td>31.68</td>
</tr>
<tr>
<td>5,5</td>
<td>32.42</td>
</tr>
<tr>
<td>5,6</td>
<td>33.6</td>
</tr>
<tr>
<td>7,4</td>
<td>22.09</td>
</tr>
<tr>
<td>7,5</td>
<td>20.35</td>
</tr>
<tr>
<td>7,6</td>
<td>21.33</td>
</tr>
</tbody>
</table>

Table 7.1 NPV Values for Base Case
For a risk-neutral decision-maker, the initial configuration with the largest NPV is selected (5, 6). This initial configuration is denoted as $Fac_0$. 
7.2 Modeling Alternate Outcomes Regarding Reservoir Uncertainty

In practice, it is possible that there is uncertainty in the specification of the PDF’s. That is, an engineer may hold a null hypothesis or base case for the set of assumptions regarding the subsurface parameter PDF’s ($\Omega_1$); he may also hold an alternate hypothesis for the set of assumptions on the PDF’s ($\Omega_2$). To simplify the analysis, we adopt the view that either the null or alternate hypothesis is true. PDF Therefore, if the project is planned based on the null hypothesis and an optimal is implemented, it is possible that this $Fac_o$ could be suboptimal if the alternate hypothesis is realized.

Let us consider that there is a possibility of $\Omega_2$ being realized where each of the subsurface variables is defined by the following PDFs:

$$Krg = normal (0.85, 0.035)$$

$$Kro = normal (0.85, 0.035)$$

$$Sor = normal (0.15, 0.035)$$

$$Sw = normal (0.35, 0.02)$$

If the workflow is repeated, a similar set of results is obtained and these are depicted in Table 7.2. The initial configuration, $Fac_o$, may be different. As expected, different assumptions regarding the uncertainties may yield different recommendations regarding the initial facility configuration.
Fig. 7.2 NPV Distribution for alternate hypothesis

<table>
<thead>
<tr>
<th>(P,I)</th>
<th>Case B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
</tr>
<tr>
<td>5,4</td>
<td>31.05</td>
</tr>
<tr>
<td>5,5</td>
<td>33.2</td>
</tr>
<tr>
<td>5,6</td>
<td>33</td>
</tr>
<tr>
<td>7,4</td>
<td>22.64</td>
</tr>
<tr>
<td>7,5</td>
<td>20.47</td>
</tr>
</tbody>
</table>
Table 7.2 NPV Values for the alternate hypothesis

7.3 **Planning for Alternate Outcomes: The Value of Information**

Based on the preceding analyses, it is clear that one can use this workflow to estimate the value of mitigating the uncertainty regarding the subsurface assumptions. The previous sections illustrate the effect different subsurface assumptions can have on the optimum initial facility configuration and hence it becomes an intriguing problem to estimate the value in knowing which set of assumptions will hold prior to making the initial investment decision. Fig. 2.3 illustrates how this value of information is found.
Superimposing the graphs for Section 7.1 and 7.2 would give us the following curve which shows a different optimum for both cases. The values are shown in Table 7.1.

![NPV Distribution Graph](image)

The values are shown in Table 7.3 below:

<table>
<thead>
<tr>
<th>(P,I)</th>
<th>Case A ($\Omega_1$).</th>
<th>Case B ($\Omega_2$).</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Stdev</td>
</tr>
<tr>
<td>5,4</td>
<td>31.68</td>
<td>2.38</td>
</tr>
<tr>
<td>5,5</td>
<td>32.42</td>
<td>2.85</td>
</tr>
<tr>
<td>5,6</td>
<td>33.6</td>
<td>3.35</td>
</tr>
<tr>
<td>7,4</td>
<td>22.09</td>
<td>3.05</td>
</tr>
<tr>
<td>7,5</td>
<td>20.35</td>
<td>4.26</td>
</tr>
<tr>
<td>7,6</td>
<td>21.33</td>
<td>4.08</td>
</tr>
</tbody>
</table>

Table 7.3 NPV and Standard Deviation Values for base case and alternate hypothesis
As Table 7.1 shows, the optimum configuration for $\Omega_1$ is (5, 6) while for $\Omega_2$, it is (5, 5). Herein lays the conflict that warrants a more sophisticated analysis of uncertainties. Both $\Omega_1$ and $\Omega_2$ are possible scenarios and designing the facility on one particular assumption could result in an economic sub-optimality. In such a scenario, the decision-making process could be easier if one knew for certain which of the above scenarios would be realized and what was the value of that information.

This VoI could be found using the model. Another simulation is run where the system is initialized with 5 spot & 6 mSCF/D, the $Fac_0$ of Case A while the PDFs entered in the optimization module are those of Case B. The simulation runs give a mean NPV of $32.9$ million.

So, in this case we find the maximum VoI in accordance with Fig. 2.3:

$$VoI = $33.2 - $32.9 = $0.3 \text{ million}$$
7.4 **Main Effects on VoI**

Sections 7.1, 7.2 & 7.3 explain the three steps taken to find the Value of Information in accordance with Fig. 2.3. This three-pronged approach can be applied to generate valuable results and analysis to support the decision-making process for field development planning. In this section, some of those results are discussed:

### 7.4.1 Effect of Oil Price on VoI

The first three sections of this Chapter assume an oil price of $20/bbl. With the price of crude oil continually increasing, it would be logical to see what effect oil price could have on our decisions. The idea was to find the value of knowing the correct estimate of the value of oil price. The VoI computations were repeated with an oil price of $60/bbl, which is what most companies use for evaluating commercial viability of projects. The cases remain the same as described in Sections 7.1 and 7.2. The results are shown in Fig 7.4 and Table 7.4:

<table>
<thead>
<tr>
<th>(P,I)</th>
<th>Case A</th>
<th>Case B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Stdev</td>
</tr>
<tr>
<td>5,4</td>
<td>401</td>
<td>16.36</td>
</tr>
<tr>
<td>5,5</td>
<td>423.6</td>
<td>16.56</td>
</tr>
<tr>
<td>5,6</td>
<td>441.5</td>
<td>19.1</td>
</tr>
<tr>
<td>(P,I)</td>
<td>Case A Mean</td>
<td>Stdev</td>
</tr>
<tr>
<td>------</td>
<td>-------------</td>
<td>-------</td>
</tr>
<tr>
<td>7.4</td>
<td>409.2</td>
<td>18.3</td>
</tr>
<tr>
<td>7.5</td>
<td>418.5</td>
<td>23</td>
</tr>
<tr>
<td>7.6</td>
<td>458.34</td>
<td>20.12</td>
</tr>
</tbody>
</table>

Table 7.4  NPV and Standard Deviation values for base case and alternate hypothesis for $60/bbl oil price

Fig. 7.4  NPV Distribution for base case and alternate hypothesis for $60/bbl oil price
The graph and the table clearly show that the optimum in both the cases is the same 7 spot and 6 mSCF/D configuration. Since the optimum does not change even in the face of competing hypothesis, the VoI in this case is zero. There is no inherent value in mitigating uncertainty in estimates and the rational decision supported by economics is to go ahead with the current development plans.

A possible explanation for zero VoI when the oil price was changed to $60/bbl could be that the field in consideration had tremendous recoverable oil and any value increments due to the field optimization were easily offset by the profits owing to large scale crude production. Hence, it is a possible conclusion that VoI analysis assumes more significance in small to marginal fields rather than large fields.

7.4.2 VoI of $K_{rg}$

In the demonstration in the first three sections of this Chapter, the subsurface assumptions were defined as a group. An appropriate follow-up analysis would investigate the value of knowing each of these subsurface parameters independently. The idea was to find the value of knowing the correct estimate of the value of relative permeability of gas. The methodology to find this is the same as described in Section 7.4.1. There are two competing hypothesis for the value of relative permeability:

Case A: $K_{rg} = normal (0.85, 0.035)$
Case B: $K_{rg} = normal (0.9, 0.02)$

Other distributions are known:

- $K_{ro} = normal (0.85, 0.035)$
- $S_{or} = normal (0.15, 0.035)$
- $S_{w} = normal (0.35, 0.02)$

The results are shown in Fig. 7.5 and Table 7.5:

![NPV Distribution](image)

Fig. 7.5  NPV Distribution for base case and alternate hypothesis for $K_{rg}$
<table>
<thead>
<tr>
<th>(P,I)</th>
<th>Case A</th>
<th></th>
<th>Case B</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Stdev</td>
<td>Mean</td>
<td>Stdev</td>
</tr>
<tr>
<td>5,4</td>
<td>31.08</td>
<td>2.85</td>
<td>31.05</td>
<td>2.74</td>
</tr>
<tr>
<td>5,5</td>
<td>33.61</td>
<td>3.45</td>
<td>33.2</td>
<td>2.78</td>
</tr>
<tr>
<td>5,6</td>
<td>33.83</td>
<td>3.17</td>
<td>33</td>
<td>3.83</td>
</tr>
<tr>
<td>7,4</td>
<td>22.9</td>
<td>3.24</td>
<td>22.64</td>
<td>2.9</td>
</tr>
<tr>
<td>7,5</td>
<td>20.72</td>
<td>3.53</td>
<td>20.47</td>
<td>3.48</td>
</tr>
<tr>
<td>7,6</td>
<td>23.73</td>
<td>3.69</td>
<td>22.56</td>
<td>3.9</td>
</tr>
</tbody>
</table>

Table 7.5 NPV and Standard Deviation values for base case and alternate hypothesis for $K_{rg}$

The system is initialized with 5 spot and 6 mSCF/D, the optimum for Case A and then a run is done with estimate of Case B to find the NPV. The NPV is $32.93$ million.

\[ \text{Vol} = \$33.2 - \$32.93 = \$0.27 \text{ million} \]

7.4.3 \textbf{VoI of } $K_{ro}$

The idea was to find the value of knowing the correct estimate of the value of relative permeability of oil. The methodology to find this is the same as described in Section 7.4.1 and 7.4.2. There are two competing hypothesis for the value of relative permeability:

Case A: $K_{ro} = \text{normal} (0.85, 0.035)$

\(^1\text{A configuration mentioned as } x,y \text{ in this work means } x \text{ spot and } y \text{ mSCF/D}\)
Case B: $K_{ro} = normal (0.9, 0.02)$

Other distributions are known:

$K_{rg} = normal (0.85, 0.035)$

$S_{or} = normal (0.15, 0.035)$

$S_{w} = normal (0.35, 0.02)$

The results are shown in Fig. 7.6 and Table 7.6:

<table>
<thead>
<tr>
<th>(P,I)</th>
<th>Case A</th>
<th>Case B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Stdev</td>
</tr>
<tr>
<td>5,4</td>
<td>30.49</td>
<td>2.93</td>
</tr>
<tr>
<td>5,5</td>
<td>32.4</td>
<td>3.43</td>
</tr>
<tr>
<td>5,6</td>
<td>32.87</td>
<td>3.41</td>
</tr>
<tr>
<td>7,4</td>
<td>22.7</td>
<td>3.28</td>
</tr>
<tr>
<td>7,5</td>
<td>19.43</td>
<td>3.43</td>
</tr>
<tr>
<td>7,6</td>
<td>21.04</td>
<td>3.93</td>
</tr>
</tbody>
</table>

Table 7.6 NPV and Standard Deviation values for base case and alternate hypothesis for $K_{ro}$
Fig. 7.6    NPV Distribution for base case and alternate hypothesis for $K_{ro}$

The system is initialized with 5 spot and 6 mSCF/D, the optimum for Case A and then a run is done with estimate of Case B to find the NPV. The NPV is $33$ million.

- VoI = $33.2 - 33 = 0.2$ million

### 7.4.4 Value of Reconfiguration

Chapter 6 outlined the concept of reconfiguring patterns for maximizing the value. The option of reconfiguration is one of the unique points of this work. The results generated in the sections above give the values of NPV after *including* the option to reconfigure.

So, the next questions that arise are “If I don’t want to reconfigure, how much do I lose
by not doing such an analysis” and “If I start with a certain field plan, what alterations in the field later would maximize my return”.

To answer the former, four scenarios were considered for analysis:

Table 7.7 Scenarios for analysis

<table>
<thead>
<tr>
<th>High Oil Price, High Uncertainty</th>
<th>Low Oil Price, High Uncertainty</th>
</tr>
</thead>
<tbody>
<tr>
<td>High Oil Price, Low Uncertainty</td>
<td>Low Oil Price, Low Uncertainty</td>
</tr>
</tbody>
</table>

The high oil price is taken as $60 /bbl while low oil price is taken as $20 /bbl. The high uncertainty case is defined as:

\[
K_{rg} = \text{normal (0.85, 0.035)}
\]
\[
K_{ro} = \text{normal (0.85, 0.035)}
\]
\[
S_{or} = \text{normal (0.15, 0.035)}
\]
\[
S_{w} = \text{normal (0.35, 0.02)}
\]

While the low uncertainty case is defined as:

\[
K_{rg} = \text{normal (0.9, 0.02)}
\]
\[
K_{ro} = \text{normal (0.9, 0.02)}
\]
\begin{align*}
S_{op} &= normal(0.15, 0.015) \\
S_w &= normal(0.35, 0.02)
\end{align*}

For each of the scenarios mentioned in Table 7.5, the system was initialized with one given configuration and the model was run for 100 iterations. The results showed the configuration switched to maximum number of times.

For both the low price scenarios, the optimum reconfiguration was (5, 4) or 5 spot, 4 mSCF/D 100% of the time. Even though for the high uncertainty case, the optimum was (5,5) while for low uncertainty it was (5,6) as shown in Table 7.4, these optimum values are realized only when the field is reconfigured to (5,4) at the set time. A possible explanation could be that continuing depletion of field and marginal returns due to low oil price make it worthwhile to downsize operations to maximize value.

For high oil price scenarios, the uncertainty has a significant impact on the reconfiguration options. For the high uncertainty case as Table 7.6 shows, there is no unanimously preferred option like in the low oil price scenarios. Depending on the reserve estimate, the facility reconfiguration could be different. The most preferred options are (5, 6), (7, 4) and (7, 6). From Section 7.4.1, we can recall that the optimum initial configuration for this case was (7, 6). Although the value-maximizing option is initializing with that configuration and reconfiguring later (the last entry in the table below), if the field was indeed initialized with a sub-optimal configuration it is possible to salvage some value by reconfiguration. For example, if the field was initialized with
(5, 4) configuration, the most profitable scenario would be to switch to (5, 6) configuration most (60%) of the times. Similarly, (5, 5) reconfigured to (7, 4), (5, 6) reconfigured to (7, 4), (7, 4) remained unchanged, (7, 5) reconfigured to (7, 4) and (7, 6) also reconfigured to (7, 6) most of the times.

<table>
<thead>
<tr>
<th>Initial Configuration</th>
<th>%age of Cases</th>
<th>Reconfigured Facility</th>
</tr>
</thead>
<tbody>
<tr>
<td>(5,4)</td>
<td>60</td>
<td>(5,6)</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>(7,4)</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>(7,6)</td>
</tr>
<tr>
<td>(5,5)</td>
<td>30</td>
<td>(5,6)</td>
</tr>
<tr>
<td></td>
<td>50</td>
<td>(7,4)</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>(7,6)</td>
</tr>
<tr>
<td>(5,6)</td>
<td>10</td>
<td>(5,5)</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>(5,6)</td>
</tr>
<tr>
<td></td>
<td>55</td>
<td>(7,4)</td>
</tr>
<tr>
<td></td>
<td>15</td>
<td>(7,6)</td>
</tr>
<tr>
<td>(7,4)</td>
<td>5</td>
<td>(5,6)</td>
</tr>
<tr>
<td></td>
<td>65</td>
<td>(7,4)</td>
</tr>
<tr>
<td></td>
<td>30</td>
<td>(7,6)</td>
</tr>
<tr>
<td>(7,5)</td>
<td>15</td>
<td>(5,6)</td>
</tr>
<tr>
<td></td>
<td>55</td>
<td>(7,4)</td>
</tr>
</tbody>
</table>

\(^2\) (x,y) facility to be read as x spot and y mSCF/D throughout this work.
Table 7.8  Reconfiguration options for high uncertainty and high oil price scenario

One possible indication from the result could be that in most cases, it is more profitable to alter the injection rate rather than change the pattern in cases of high uncertainty. Cases (5, 4), (7, 4), (7, 5) and (7, 6) indicated alteration in injection rate as the preferred reconfiguration, subject to our assumptions regarding associated costs.

A similar table was generated for low uncertainty cases:

<table>
<thead>
<tr>
<th>Initial Configuration</th>
<th>%age of Cases</th>
<th>Reconfigured Facility</th>
</tr>
</thead>
<tbody>
<tr>
<td>(5,4)</td>
<td>15</td>
<td>(5,5)</td>
</tr>
<tr>
<td></td>
<td>85</td>
<td>(5,6)</td>
</tr>
<tr>
<td>(5,5)</td>
<td>25</td>
<td>(5,5)</td>
</tr>
<tr>
<td></td>
<td>70</td>
<td>(5,6)</td>
</tr>
<tr>
<td>---------------</td>
<td>----</td>
<td>-------</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>(7,4)</td>
</tr>
<tr>
<td>(5,6)</td>
<td>40</td>
<td>(5,5)</td>
</tr>
<tr>
<td></td>
<td>60</td>
<td>(5,6)</td>
</tr>
<tr>
<td>(7,4)</td>
<td>25</td>
<td>(5,5)</td>
</tr>
<tr>
<td></td>
<td>65</td>
<td>(5,6)</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>(7,4)</td>
</tr>
<tr>
<td>(7,5)</td>
<td>20</td>
<td>(5,5)</td>
</tr>
<tr>
<td></td>
<td>65</td>
<td>(5,6)</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>(7,4)</td>
</tr>
<tr>
<td></td>
<td>5</td>
<td>(7,6)</td>
</tr>
<tr>
<td>(7,6)</td>
<td>20</td>
<td>(5,5)</td>
</tr>
<tr>
<td></td>
<td>25</td>
<td>(5,6)</td>
</tr>
<tr>
<td></td>
<td>45</td>
<td>(7,4)</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>(7,6)</td>
</tr>
</tbody>
</table>

Table 7.9  Reconfiguration options for low uncertainty and high oil price scenario

It may be recalled that the optimum initial configuration for such a case from Section 7.4.1 was (7, 6) configuration. However, if the system was not initialized with this field plan and initialized with (5, 4) configuration, the maximum value is realized by switching to (5, 6) most (85%) of the times. Similarly, (5, 5) would reconfigure to (5, 6), (5, 6) should not be altered, (7, 4) would reconfigure to (5, 6), (7, 5) would reconfigure to (5, 6) and (7, 6) would reconfigure to (7, 4) most of the times. One possible conclusion that can
be drawn from the above table is that even though the uncertainty is lesser as compared to the case above, the expansion option is not always preferred which is a bit counter-intuitive. Neither does low uncertainty imply a reduction in possible reconfiguration options.

Another question arising from the above analysis is what if we decide not to reconfigure. In such a scenario, how much value is being lost? To find data to answer that question, the optimization model was run twice for each configuration – once it was allowed to reconfigure and another time it was not. With the oil price fixed at $60/bbl, the calculations were done for both the high and low uncertainty cases and the results are tabulated in Tables 7.10 and 7.11:

<table>
<thead>
<tr>
<th>Facility</th>
<th>NPV in million $ without Reconfiguration</th>
<th>NPV in million $ with Reconfiguration</th>
<th>%age Project Value lost</th>
</tr>
</thead>
<tbody>
<tr>
<td>(5,4)</td>
<td>392.1</td>
<td>400.6</td>
<td>2.15%</td>
</tr>
<tr>
<td>(5,5)</td>
<td>412.1</td>
<td>425.1</td>
<td>3.15%</td>
</tr>
<tr>
<td>(5,6)</td>
<td>442.7</td>
<td>448.6</td>
<td>1.35%</td>
</tr>
<tr>
<td>(7,4)</td>
<td>408.0</td>
<td>410.0</td>
<td>0.49%</td>
</tr>
<tr>
<td>(7,5)</td>
<td>409.9</td>
<td>425.6</td>
<td>3.85%</td>
</tr>
<tr>
<td>(7,6)</td>
<td>449.3</td>
<td>459.9</td>
<td>2.35%</td>
</tr>
</tbody>
</table>

Table 7.10  Value addition by reconfiguration for Case A
Table 7.11 Value addition by reconfiguration for Case B

The above tables clearly show that the option to reconfigure does have significant value addition. Inclusion of such real options in the analysis for initial capital investment decisions can potentially realize up to 3% of project value. The absolute value of such an incremental realization could run in millions of dollars which is a significant sum of money.
7.4.5  **Effect of Cost on Optimum Facility**

The above sections in this chapter made certain analyses that were based on cost equations as described in Chapter 6 and values as shown in Appendix A.2. However, those costs can vary quite a bit given onshore and offshore drilling and different geological environments. The idea is to see whether a difference in costs affects our choice of facility. The two uncertainty cases – high & low as mentioned in Section 7.4.4 were redone with the cost estimates of platform and drilling wells doubled. The results are shown in Fig. 7.12 and Table 7.13:

<table>
<thead>
<tr>
<th>(P,I)</th>
<th>Case A</th>
<th>Case B</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Mean</td>
<td>Stdev</td>
</tr>
<tr>
<td>5,4</td>
<td>375.06</td>
<td>16.01</td>
</tr>
<tr>
<td>5,5</td>
<td>399.38</td>
<td>16.37</td>
</tr>
<tr>
<td>5,6</td>
<td>416.54</td>
<td>18.64</td>
</tr>
<tr>
<td>7,4</td>
<td>375.97</td>
<td>17.42</td>
</tr>
<tr>
<td>7,5</td>
<td>395.97</td>
<td>20.33</td>
</tr>
<tr>
<td>7,6</td>
<td>430.82</td>
<td>20.83</td>
</tr>
</tbody>
</table>

Table 7.12  NPV and Standard Deviation Values for Double the Cost Estimates
Fig. 7.7  Graphical Distribution of NPV and Standard Deviation Values for Double the Cost Estimates

The results show the optimum as (7, 6) configuration which was what was found with the other cost estimate in Sec 7.4.1. The data shows that the Value of Information for Cost estimates is zero in this case. Based on this case alone, it would be wrong to draw a general conclusion that price estimates do not affect optimum facility configuration. However, one possible conclusion could be that in the case of fields with very significant recoverable reserves, as in this case, the effect of costs on optimum design is insignificant as it is offset by the high profit margins from the recovered oil. This would also be true for very high oil prices.
CHAPTER 8: CONCLUSIONS

8.1 KEY FINDINGS

One of the objectives of this thesis was to develop a workflow to facilitate a VoI analysis which in the presence of uncertainty in subsurface estimates would assist in making initial capital decisions. The results discussed in Chapter 7 clearly establish that there is a significant value associated in including real options while performing facility optimization. A standard Monte Carlo approach to handle uncertainty fails to realize the potential value that could be secured by including the option to respond to future outcomes like revealed uncertainties. The results could be summarized thus:

- The Value of Information on reservoir parameters is higher for small or marginal fields than for larger fields for a fixed oil price.
- Cost estimates do not significantly impact optimum facility sizing especially for large reserves.
- Allowance for a real option to expand/abandon facilities in initial field development analysis can realize as much as 3% of the project value.
- The NPV realized by including option to expand/abandon facility prior to designing fields is higher than the NPV realized by expanding or abandoning facilities after it has been built and some uncertainties have been revealed.
- The VoI of important subsurface parameters, though significant in absolute terms, constitutes less than 1% of the NPV of the project.
• Price fluctuations are more critical in determining the reconfigured field than the uncertainty in reservoir parameters. Given everything else remains constant; as the oil price increases, different field plans could be the optimum reconfigured field depending on the revealed uncertainty.

8.2 FUTURE WORK

The work detailed in this thesis is essentially the demonstration of a method and proof of its concept. The reservoir used for the pilot study is at best a simplified representation to the reservoirs encountered in the field. As a next step forward in this research, it is desired to make the reservoir more ‘realistic’ by inclusion of more heterogeneities in data for permeability and porosity and also in re-defining the problem in the more complex domain of foam floods, with the inclusion of more relevant design parameters like SAG/WAG ratio.
APPENDIX

A.1 CAPEXSWITCH VALUES FOR EACH CASE

Cost for 5 spot and 4 mSCF/D
TABLE CapexSwitch(pat,IRate) Capital cost for Switching in million dollars

\[
\begin{array}{ccc}
4 & 5 & 6 \\
5 & 0 & 2 \\
7 & 10 & 13 & 16 \\
\end{array}
\]

Cost for 5 spot and 5 mSCF/D
TABLE CapexSwitch(pat,IRate) Capital cost for Switching in million dollars

\[
\begin{array}{ccc}
4 & 5 & 6 \\
5 & 0 & 0 & 2 \\
7 & 10 & 10 & 13 \\
\end{array}
\]

Cost for 5 spot and 6 mSCF/D
TABLE CapexSwitch(pat,IRate) Capital cost for Switching in million dollars

\[
\begin{array}{ccc}
4 & 5 & 6 \\
5 & 0 & 0 & 0 \\
7 & 10 & 10 & 10 \\
\end{array}
\]
### Cost for 7 spot and 4 mSCF/D

TABLE CapexSwitch(pat,IRate) Capital cost for Switching in million dollars

<table>
<thead>
<tr>
<th></th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0</td>
<td>3</td>
<td>6</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>3</td>
<td>6</td>
</tr>
</tbody>
</table>

### Cost for 7 spot and 5 mSCF/D

TABLE CapexSwitch(pat,IRate) Capital cost for Switching in million dollars

<table>
<thead>
<tr>
<th></th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>0</td>
<td>3</td>
</tr>
</tbody>
</table>

### Cost for 7 spot and 6 mSCF/D

TABLE CapexSwitch(pat,IRate) Capital cost for Switching in million dollars

<table>
<thead>
<tr>
<th></th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>5</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>7</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
</tbody>
</table>
A.2 Values of Parameters in Optimization Module

- Fixed cost of well in million dollars, $fcwell = 3$
- Operating cost in million dollars per mSCF injected per day, $fcopex = 0.6$
- Coefficient for facility cost (does not vary with size), $b0plat = 20$
- Coefficient for facility cost (vary with size), $b1plat = 0.001$
- Exponent term for facility cost, $explat = 2.486$
- Variable cost of production in dollars per bbl, $vcprod = 10$
- Expansion cost multiplier, $m = 1.10$
- Discount rate, $r = 15\%$
- Time of reconfiguration, $tswitch = 1000$ days
A.3 AN ALTERNATE APPROACH TO PROXY MODELING

Most of the proxy models developed in the literature are static in nature in the sense that they either predict ultimate recovery at terminal time or they use different proxy models at different fixed points in time. None of the proxy models in the literature estimate a time series production function. Such a model would be helpful in field development optimization.

We first describe a naïve time-series model. Let Y be the cumulative oil production rate, X contains a polynomial series for time and a selection of subsurface variables. Assume we have multiple time series observations of oil production from a reservoir simulator where:

\[ y_{it} = \text{the value of the dependent variable for time series } i \text{ at time } t \text{ where } i = 1 \ldots n \text{ and } t = 1 \ldots T. \]

\[ X_{it}^j = \text{the value of the } j\text{th explanatory variable for time series } i \text{ at time } t, \text{ where } j = 1 \ldots K. \]

\[
y_i = \begin{bmatrix} y_{i1} \\ \vdots \\ y_{iT} \end{bmatrix}, \quad X_i = \begin{bmatrix} X_{i1}^1 & X_{i1}^2 & \cdots & X_{i1}^K \\ X_{i2}^1 & X_{i2}^2 & \cdots & X_{i2}^K \\ \vdots & \vdots & \ddots & \vdots \\ X_{iT}^1 & X_{iT}^2 & \cdots & X_{iT}^K \end{bmatrix}, \quad e_i = \begin{bmatrix} e_{i1} \\ \vdots \\ e_{iT} \end{bmatrix} \tag{A.1} \]
If \( n \) individual time series are stacked:

\[
\begin{bmatrix}
y_1 \\
y_2 \\
\vdots \\
y_n
\end{bmatrix}_{n \times T_x} ; \quad
\begin{bmatrix}
X_1 \\
X_2 \\
\vdots \\
X_n
\end{bmatrix}_{n \times T_x} ; \quad
\begin{bmatrix}
e_1 \\
e_2 \\
\vdots \\
e_n
\end{bmatrix}_{n \times T_x}
\]

We can then express the standard linear model as:

\[
y = X \beta + e, \quad \text{where } \beta \text{ is a } k \times 1 \text{ vector.} \quad (A.3)
\]

Equation (A.3) can also be specified as \( \ln(y) = X \beta + e \).

For the current application, each time series is the result of a reservoir simulation based on revealed values of (uncertain) subsurface attributes given a fixed facility configuration. The minimum number of time series in each case is determined by Design of Experiments. The goal is to obtain a “composite” production function for each collection of simulations via regression analysis.

Regression on this type of panel data structure requires some advanced statistics to treat for various specification problems (e.g. heteroscedasticity and autocorrelation). For the naïve estimator, we proceed as follows:
• Assume that the $e_{it} \sim iid(0, \sigma^2)$. This is called the “naive estimator” because it ignores likely serial correlation within each series, and assumes constant variance of the errors across series and time [the coefficient estimates are unbiased, only the variance-covariance matrix is affected].

• In most applications, the assumption that $e_{it} \sim iid(0, \sigma^2)$ will be violated. This can be tested via a Durbin-Watson statistic. A generalized Durbin-Watson statistic is calculated as follows:

$$DW = \frac{\sum_{j=1}^{n} \sum_{t=2}^{T} (e_{it} - e_{i,t-1})^2}{\sum_{j=1}^{n} \sum_{t=1}^{T} (e_{it})^2}$$

Some investigation was done to estimate the inefficiency of the naïve estimator relative to a more advanced model. An alternate model was specified to account for serial correlation within series by assuming a first-order autoregressive process (AR-1) for the error term. Assume we have a time series observation of oil production where:

$$y_{it} = \alpha + X_{it}\beta + u_i + e_{it}$$

\(y_{it}\) = the cumulative production volume for time series \(i\) at time \(t\) where \(i = 1…n\) and \(t = 1…T\).
\( \alpha \) = intercept term that is constant over all \( i \) and \( t \)

\( X_{it} \) = a \( T \times k \) matrix of explanatory variables

\( \beta \) = a vector of coefficients to be estimated

\( u_i \) = random effect for series \( i \)

\[ e_{it} = \varphi e_{i,t-1} + z_{it}, \text{ where } z_{it} \sim iid \left( 0, \sigma_z^2 \right). \]

It is assumed that \(|\rho| < 1\) and the \( u_i \) are independent of the \( X_{it} \). A random effect (versus fixed effect) model is required because some of the \( X_{it}^j \) are constant across individual time series and therefore collinear with the \( u_i \). As before, \( n \) individual time series are stacked. The procedure for estimating this model is not trivial; see Baltagi and Li (1991) for a detailed exposition of the problem\(^3\).

We can compare the results of the naïve estimator and the AR-1 model. Table A.1 and Table A.2 report the regression results of the naïve estimator and the AR-1 estimator for one case of facility configuration.

---

\(^a\) For additional technical background, see Baltagi and Li (1992) and Baltagi (2001). For examples of empirical applications see Lillard and Willis (1978) and Berry, Gottschalk and Wissoker (1998).
Table A.1: Naïve Estimator

<table>
<thead>
<tr>
<th>Source</th>
<th>SS</th>
<th>df</th>
<th>MS</th>
<th>Number of obs = 390</th>
</tr>
</thead>
<tbody>
<tr>
<td>Model</td>
<td>1.0187e+16</td>
<td>7</td>
<td>1.4553e+15</td>
<td>F( 7, 382) = 15769.11</td>
</tr>
<tr>
<td>Residual</td>
<td>3.5254e+13</td>
<td>382</td>
<td>9.2289e+10</td>
<td>Prob &gt; F = 0.0000</td>
</tr>
<tr>
<td>Total</td>
<td>1.0222e+16</td>
<td>389</td>
<td>2.6279e+13</td>
<td>R-squared = 0.9966</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Adj R-squared = 0.9965</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td>Root MSE = 3.0e+05</td>
</tr>
</tbody>
</table>

| qsim | Coef.  | Std. Err. | t    | P>|t| | [95% Conf. Interval] |
|------|--------|-----------|------|-----|---------------------|
| t    | 1609536| 41078.19  | 39.18| 0.000| 1528769 - 1690304 |
| t2   | -65821.34| 5266.147 | -12.50| 0.000| -76175.61 - -55467.08 |
| t3   | 1598.558| 253.2813  | 6.31 | 0.000| 1100.558 - 2096.558 |
| t4   | -15.67142| 4.05568 | -3.86 | 0.000| -23.64567 - -7.697167 |
| krow1| 4243892| 196096.5 | 21.64| 0.000| 3858328 - 4629456 |
| sor2 | -4405783| 392193.1 | -11.23| 0.000| -5176911 - -3634656 |
| krg3 | -2531983| 196096.5 | -12.91| 0.000| -2917547 - -2146420 |
| _cons| -534944.4| 273548.9 | -1.96| 0.051| -1072794 - 2905.678 |
Table A.2: AR-1 Estimator

|                   | Coef.   | Std. Err. | z      | P>|z|   | [95% Conf. Interval] |
|-------------------|---------|-----------|--------|-------|----------------------|
|                   |         |           |        |       |                      |
| qsim              | 1420751 | 23446.27  | 60.60  | 0.000 | 1374797 – 1466705   |
|                   | -43847.65 | 3185.419  | -13.77 | 0.000 | -50090.96 – -37604.35 |
|                   | 669.9391 | 156.4845  | 4.28   | 0.000 | 363.2351 – 976.6431 |
|                   | -2.685256 | 2.493716  | -1.08  | 0.282 | -7.572849 – 2.202338 |
|                   | 4423464 | 916788.8  | 4.82   | 0.000 | 2626591 – 6220338   |
|                   | -4646073 | 1833577   | -2.53  | 0.011 | -8239818 – -1052328 |
|                   | -2786451 | 916788.8  | -3.04  | 0.002 | -4583324 – -989577.8 |
|                   | -20637.45 | 1202636   | -0.02  | 0.986 | -2377761 – 2336486  |
|                   |         |           |        |       |                      |
| rho_ar            | .9626613 | (estimated autocorrelation coefficient) |
| sigma_u          | 0       |           |        |       |                      |
| sigma_e          | 87998.422 |          |        |       |                      |
| rho_fov          | 0       | (fraction of variance from to u_i) |
| theta            | 0       |           |        |       |                      |

modified Bhargava et al. Durbin-Watson = .09480491
Baltagi-Wu LBI = .25425279
The overall fit of both models is excellent, and this is often the case when regressing with time variables. The coefficient estimates for the time variables differ statistically between the two models but the estimates for the subsurface variables are not statistically different. Both models are adequate predictors and residuals are negligible for the time periods of most concern in the economic optimization.

The objective of the regression analysis is to obtain a reasonable representation of likely production profiles so that economic optimization can proceed without recycling to the reservoir simulator. The naïve estimator is acceptable for this purpose notwithstanding probable violations of OLS assumptions. That is, deficiencies in the naïve estimator are unlikely to affect the general conclusions derived from the economic optimization. Also, the naïve estimator is easy to understand and can be estimated in common software packages like Excel, increasing the probability of its implementation among practicing engineers.
REFERENCES


28. Engineering Statistics Handbook, 


33. McCarl GAMS Guide

34. George Box, Donald Behnken, "Some new three level designs for the study of quantitative variables", Technometrics, Volume 2, pages 455-475, 1960.


37. Yamali, N., Optimum control of unwanted fluid production in stratified gas and oil reservoirs, MS Thesis 2007, University of Texas at Austin