

**USE OF ORTHOGONAL ARRAYS, QUASI-MONTE CARLO
SAMPLING, AND KRIGING RESPONSE MODELS FOR
RESERVOIR SIMULATION WITH MANY VARYING FACTORS**

A Thesis
Submitted to the Graduate Faculty of the
Louisiana State University and
Agricultural and Mechanical College
in partial fulfillment of the
requirements for the degree of
Master of Science in Petroleum Engineering

in

The Department of Petroleum Engineering

by
Subhash Kalla
B.Tech., Regional Engineering College, Warangal (India), 2002
August 2005

ACKNOWLEDGEMENTS

The author expresses his sincere thanks to Dr. Christopher White for his enlightenment, guidance, patience and encouragement. His insights and suggestions have always been valuable and beneficial during the course of this research. The author owes a lot to him.

This project was funded by Landmark Graphics Corporation and Unocal. I thank Dr. David Heath, Dr. Stan Cullick, and Keshav Narayanan for their support and guidance. Without their support, this work would not have been possible.

Thanks to Dr. Julius Langlinais and Dr. Anuj Gupta in the Department of Petroleum Engineering who were enthusiastic to serve on the examining committee and gave valuable suggestions.

Thanks are extended to all the other faculty members and students who have offered help and been kind to the author and made the past several years enjoyable and memorable.

TABLE OF CONTENTS

ACKNOWLEDGEMENTS	ii
LIST OF TABLES	v
LIST OF FIGURES	vi
ABSTRACT.....	vii
CHAPTER 1. INTRODUCTION	1
1.1 Experimental Design and Response Models	2
1.1.1 Experimental Design.....	3
1.1.2 Response Models	4
CHAPETR 2. RESPONSE SURFACE METHODS	7
2.1 Effects, Interactions and Interpretation	7
2.2 Fitting the Polynomial Models	10
2.3 Alternate Regression Method	14
2.4 Uncertainty Assessment.....	18
CHAPETR 3. SAMPLING TECHNIQUES AND EXPERIMENTAL DESIGN ...	20
3.1 Factorial Designs	20
3.2 Fractional Factorial Designs	21
3.2.1 Aliasing and Confounding Effects	23
3.3 Geometric and Optimal Designs	26
3.4 Why Another Design?.....	29
3.5 Orthogonal Arrays.....	29
3.5.1 Benefits of OA/NOA	32
3.5.2 Algorithm for OA/NOA.....	32
3.6 Latin Hypercubes and the Hammersley Sequence	33
3.6.1 Latin Hypercube Sampling	33
3.6.2 Hammersley Sequence Sampling.....	36
3.7 Inducing Correlation Structure into the Designs.....	37
3.8 Number of Samples or Runs	38
CHAPTER 4. COMPARISON OF DESIGNS	40
4.1 Comparing Orthogonal Arrays and Latin Hypercubes	40
4.2 Comparing Orthogonal Arrays and Full Factorial Designs	42
4.2.1 <i>F</i> Test to Compare Two Variances.....	45
4.3 Comparing Different Sampling Techniques.....	46
CHAPTER 5. APPLICATION OF RESPONSE SURFACES	50
5.1 Simulation Model	51
5.1.1 Reservoir Geometry and Properties.....	51
5.1.2 Parameters under Consideration	53

5.2 Response Model Building	58
5.3 Automation	59
5.4 Fitting Response Surface Models Using Regression	59
5.4.1 Gas Recovery	59
5.4.2 Water Breakthrough	61
5.5 Kriging Surfaces.....	62
5.6 Comparison between Polynomial Regression and Kriging.....	65
5.6.1 Bootstrapping the Polynomial Model.....	65
5.6.2 Jackknifing the Kriging Model.....	66
5.6.3 Comparison of Errors.....	67
CHAPTER 6. RESERVOIR APPLICATIONS.....	68
6.1 Uncertainty Analysis.....	68
6.1.1 Robustness of Uncertainty Estimate	68
6.1.2 Response Surfaces for Uncertainty Estimation.....	70
6.2 Optimization of Engineering Factors.....	71
6.2.1 Formulation of the Objective Function.....	72
6.2.2 Optimization Method and Results.....	72
6.2.3 Discussion of Optimization.....	73
6.3 Statistical Comparison of Designs	75
6.4 Economic Analysis	77
CHAPTER 7. CONCLUSIONS AND RECOMMENDATIONS	79
REFERENCES.....	81
VITA.....	86

LIST OF TABLES

2.1 Design Matrix of a 2^2 Factorial Design	8
2.2 Data for Multiple Linear Regression	12
3.1 2^{4-1} Fractional Factorial Design	23
3.2 Example of an OA (8, 4, 2, 3)	31
3.3 Latin Hypercube LHS (5, 4).....	34
4.1 D-optimality Test Results for OA and LHS	42
4.2 F Test Results for OA and Full Factorial Models.....	46
5.1 Factors considered in the Gas Reservoir Problem.....	54
5.2 Regression Statistics for the Gas Recovery	60
5.3 Analysis of Variance for the Gas Recovery	60
5.4 Regression Statistics for the Breakthrough Time.....	62
5.5 Analysis of Variance for the Breakthrough Time.....	62
6.1 Monte Carlo Simulation Statistics.....	71
6.2 Gas Recovery Optimum Values.....	73
6.3 Comparison of NOA(36, 14, $3^6 4^8$, 2) and NOA(108, 14, $3^6 4^8$, 2)	76
6.4 Comparison of NOA(72, 14, $3^6 4^8$, 2) and NOA(108, 14, $3^6 4^8$, 2)	77
6.5 Net Present Value Optima.....	78

LIST OF FIGURES

2.1 Graphical Representation of the Design Matrix	8
2.2 Example Semivariogram	16
2.3 Monte Carlo Workflow	18
3.1 2^3 Factorial Design	22
3.2 2^{3-1} Fractional Factorial Design	22
3.3 Central Composite Design.....	28
3.4 Box-Behnken Design.....	28
4.1 Sample Points on a Unit Square	41
4.2 Sample Points (100) on a Unit Square for Monte Carlo, LHS and HSS.....	47
4.3 Convergence Properties of Monte Carlo, LHS and HSS	49
5.1 Sketch of the Radial Reservoir Model	52
5.2 Gas Recovery Sensitivity to Pressure and Permeability	61
5.3 Estimation of Correlation Function Parameters	64
5.4 Gas Recovery Sensitivity to Pressure and Permeability	65
6.1 Monte Carlo Simulation Distributions.....	70
6.2 Objective Function Sensitivity to the Engineering Factors.....	74

ABSTRACT

Asset development teams may adjust simulation model parameters using experimental design to reveal which factors have the greatest impact on the reservoir performance. Response surfaces and experimental design make sensitivity analysis less expensive and more accurate, helping to optimize recovery under geological and economical uncertainties. In this thesis, experimental designs including orthogonal arrays, factorial designs, Latin hypercubes and Hammersley sequences are compared and analyzed.

These methods are demonstrated for a gas well with water coning problem to illustrate the efficiency of orthogonal arrays. Eleven geologic factors are varied while optimizing three engineering factors (total of fourteen factors). The objective is to optimize completion length, tubing head pressure, and tubing diameter for a partially penetrating well with uncertain reservoir properties. A nearly orthogonal array was specified with three levels for eight factors and four levels for the remaining six geologic and engineering factors. This design requires only 36 simulations compared to $4^6 \times 3^8$ (26,873,856) runs for a full factorial design. Hyperkriging surfaces are an alternative model form for large numbers. Hyperkriging uses the maximum likelihood variogram model parameters to minimize prediction errors. Kriging is compared to conventional polynomial response models. The robustness of the response surfaces generated by kriging and polynomial regression are compared using jackknifing and bootstrapping. Sensitivity analysis and uncertainty analysis can be performed inexpensively and efficiently using response surfaces.

The proposed design approach requires fewer simulations and provides accurate response models, efficient optimization, and flexible sensitivity and uncertainty assessment.

CHAPTER 1. INTRODUCTION

Reservoir simulation and modeling are complex because they involve integration of geological properties of the reservoir, drilling and production parameters, and fluctuating economic parameters. Complexity of integration is further increased, as many of the parameters like permeability, gas price, and fluid saturations are uncertain. Many geological parameters are obtained by predicting reservoir properties at unknown locations in reservoirs by conditioning imprecise data (such as seismic surveys) on relatively precise well data and possible prior training images. The precise data are sparse and the estimation at most reservoir locations is uncertain.

In exploration and production decisions, there are many alternatives such as well placement, drainage strategies, artificial lift, and capital investment for reservoir development. Development studies examine these alternatives under uncertain geologic, engineering, and economic parameters to formulate and optimize production plans (Narayanan *et al.*, 2003). Hence, many technical and economic decisions must be made despite uncertainty in the reservoir. As a result, reservoir studies may require many expensive simulations to understand the effects of variables on reservoir responses such as production rate, net present value, and breakthrough time. Despite exponential growth of computational memory and speed, computing accurate solutions is still expensive, so that it may not be feasible to consider all alternative models.

Thus, simulation runs must be chosen and analyzed efficiently. Experimental design and response models can improve study efficiency and have been applied in engineering and sciences (Sacks and Welch, 1997). These statistical methods are

becoming widely used in reservoir engineering (White *et al.*, 2003; Peng and Gupta, 2004; Peake *et al.*, 2005) to

1. Minimize computational costs by choosing a smaller but statistically representative set of simulation runs for predicting responses (e.g., recovery)
2. Improve accuracy by decreasing the possible errors caused by using arbitrary or nonoptimal simulation model choices
3. Understand factor sensitivity to response (sensitivity analysis)
4. Translating uncertainty in input factors to uncertainty in predicted performance (uncertainty analysis)
5. Analyze uncertainty to identify factors that most strongly influence the model predictions, so that resources can be focused on reducing uncertainty in factors that will most effectively decrease the uncertainty and aid optimization of engineering factors

1.1. EXPERIMENTAL DESIGN AND RESPONSE MODELS

Experimental design and response models use interpolation to minimize the number of simulation runs required for a given level of accuracy and inference. Interpolation computes values at unknown points using known surrounding values. By using simulated results, interpolation predicts outputs for factor combinations not yet simulated. The factor combinations should be chosen so that simulated factor combinations fill the multidimensional space. Thus, interpolation avoids running expensive simulations at *all factor combinations*. Choosing factor combinations is critical and is called *design of experiments* or *experimental design*. Experimental designs with less redundancy between points fill the factor space and thus improve the response

interpolation. Design of experiments selects the combinations of points for simulation models in this research.

1.1.1. Experimental Design

There is a range of designs in the literature. Designs should span the factor space completely with the fewest possible runs; good designs will also simplify response model calculations. Classical designs from experimental statistics are orthogonal and rotatable (Box and Hunter, 1978). Monte Carlo designs fill the design space and they are generalized as sampling methods (Sandor and Andras, 2004). True Monte Carlo methods use random numbers, but extensions uses deterministic numbers to sample the design space. These deterministic number sequences, called quasi-Monte Carlo samples, typically constructed using number theoretic methods (Halton, 1960; Sobol, 1967; Kalagnanam and Diwekar, 1997). Parallel to development of quasi-Monte Carlo samples, McKay *et al.* (1979) developed a stratified Monte Carlo sampling method called Latin hypercubes based on random numbers. A generalized Latin hypercube method called orthogonal arrays is orthogonal, which makes it similar to classical experimental designs. Actually, orthogonal arrays are a type of fractional factorial design, which are fractions of full factorial design (Chapter 3). In this thesis, designs including orthogonal arrays, factorial designs, Latin hypercubes and Hammersley sequences are compared and analyzed.

In statistical applications, classical two-level factorial designs or space filling Latin hypercube (Aslett *et al.*, 1998) designs are frequently used. To approximate nonlinear oil and gas reservoir responses accurately, designs must consider factors at more than two levels, not just high and low values as in two-level factorial or fractional

factorial designs (Box and Hunter, 1961). These high level designs can relate responses and factors more accurately by using higher-order effects in polynomial regression. However, multilevel designs increase the computational burden because they require more simulations; this is especially burdensome if many factors are being considered.

Alternatives to multilevel full factorial designs include partial factorials and a variation of the number of levels used per factor (mixed level designs). Unfortunately, mixed multilevel partial factorial designs are difficult to formulate and have not been used in reservoir engineering. Orthogonal arrays and nearly orthogonal arrays provide the required design properties and can handle many parameters (Hedayat *et al.*, 1999; Xu, 2002). These arrays span the design space with fewer runs, can be manipulated easily, and are appropriate for analysis of computer experiments. Mixed-level orthogonal arrays can also be generated, and they are good designs for reservoir simulation studies.

1.1.2. Response Models

Prediction of production and reserves is an important part of reservoir development and management. The production of gas reservoirs that have no associated aquifers is relatively simple to predict by analytical models. However, gas recoveries from water-drive reservoirs are hard to predict using analytical solutions because of water influx causing trapped gas and water coning. This problem was therefore selected as appropriate for study using experimental design, simulation, and response models.

In this study of water-drive gas reservoir behavior, fourteen geological and engineering factors are considered to investigate water-drive gas reservoir responses like gas recovery and breakthrough time. Different designs for the problem are analyzed and compared.

The concepts of uniformity, space filling characteristics, D-optimality, orthogonality, and

univariate and multivariate characteristics (Myers and Montgomery, 1995) are used in choosing better designs for reservoir applications. Numerical experiments clarify the possible computational savings from better designs.

Polynomial models relate the factors and responses. Kriging is examined as an alternative interpolation technique to build the response models. Kriging uses a linear combination of weights at known points to interpolate, similarly to inverse-distance weighted averaging but treating data clustering or redundancy more effectively (Goovaerts, 1997). In kriging, a variogram is similar to an autocovariance or autocorrelation function (Deutsch and Journel, 1998). These functions define the continuity of a factor with lag or separation distance in a particular direction. Variograms are simple linear transforms of the covariance between all possible pairs of points, which is well-defined for any variogram with a sill. Kriging has several important properties:

1. Kriging predictions are weighted linear combinations of observations.
2. Weights of surrounding data points depend on how close the other data points are. That is, kriging considers redundancy.
3. Most kriging methods assume second order stationary (expectations of predictions and variances do not depend on prediction point location).
4. Kriging is an “exact” interpolator; the interpolated value equals the data value (which is not true for an unsaturated polynomial).
5. Far from data, the kriged estimate reverts to the mean, whereas polynomials perform more erratically.

Kriging in the high dimensional factor space is compared with regression surfaces using bootstrapping and jackknifing (Efron and Tibshirani, 1993).

The benefits of this study are improved understanding of water encroaching gas reservoir and efficient methods to examine the sensitivity and uncertainty of complex reservoir models with large number of influential factors.

CHAPTER 2. RESPONSE SURFACE METHODS

Response surface methodology (RSM) uses statistical techniques like experimental design and regression analysis (chapter 1) to evaluate alternatives and optimize processes. Response surfaces are approximations to summarize the response of interest. RSM has been extensively used in industrial applications, particularly where there are several input variables, or factors, that are potentially influencing the output, or response. The *factors* are the *inputs* that are varied, and the *responses* are dependent variables, or *outputs*.

2.1. EFFECTS, INTERACTIONS AND INTERPRETATION

To generate response surfaces, the numbers of *levels* for each variable are selected to explore a region in the factor space. Factor spaces are an interval for a single variable and a hypervolume for multiple variables. As the number of levels increases, more values within the range of a factor are evaluated, increasing the accuracy of the model. Usually the levels are denoted by whole numbers from 0 to $n-1$ where n is the number of levels. The whole range of a factor is transformed and rescaled between 0 to $n-1$ and level values are selected (Box *et al.*, 1978). For example, if the factor is two-level then the levels are 0 and 1. The 0 and 1 can be mapped back to the original values and they represent the minimum and maximum values in the range of that factor. Sometimes scaling is done between $-(n-1)/2$ and $(n-1)/2$ if n is odd and $-(n-2)/2$ and $n/2$ if n is even. For example, if the factor has three levels then the level values are 0, 1, and 2 or -1, 0 and 1. If there are l_1 levels for the first variable, l_2 for the second... and l_k for the k^{th} variable then experiments could be performed at possible combination of all the factor levels obtain a vector of responses. Usually, these levels comprise a set of points on a

rectangular lattice. A set of factor levels intended to estimate the effects of varying factors on the observed response is called a *design* and the display of levels as a matrix is called *design matrix* (Box and Hunter, 1978; Box and Draper, 1987).

What can we find from these designs? For example, if the effect of tubing head pressure (p_{tf}) and completion length (h_p) on cumulative gas production (G_p) is to be examined, it could be done using designs. If both the factors p_{tf} and h_p are varied at two levels 0 and 1 then the possible combinations at which experiments could be done are (0, 0), (0, 1), (1, 0), and (1, 1). Running all possible combinations of these two factors each at two levels is called a 2^2 factorial design. The design matrix for factorial combinations having two factors is shown below (Table 2.1; Figure 2.1).

Table 2.1 Design Matrix of a 2^2 Factorial Design

Run	Factors	
	p_{tf}	h_p
1	0	0
2	0	1
3	1	0
4	1	1

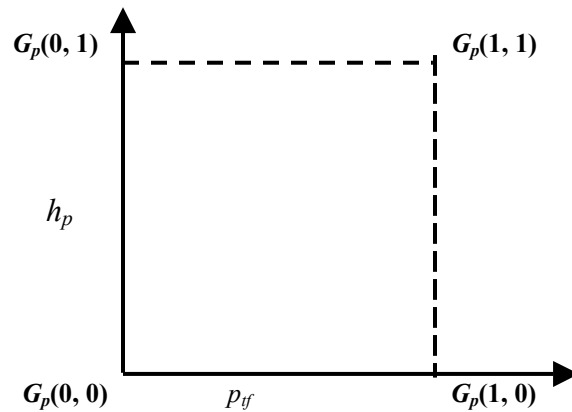


Figure 2.1 Graphical Representation of the Design Matrix

$G_p(x, y)$ is the response at the treatment combinations evaluated at chosen p_{tf} and h_p level values.

The factor effect is the change in the response as we move from a lower level to a higher level of a factor, keeping all the other factors constant. What is the effect of completion length on the response G_p ? Consider the first two runs in Table 2.1. Aside from experimental error (which is reasonable to neglect for deterministic computer experiments), the response G_p differs only because of the completion length, h_p . The p_{tf} is constant for both the experimental runs.

$$e_y = (G_p(0,1) - G_p(0,0))$$

where e_y is the effect of the factor h_p on response G_p at low factor setting of p_{tf} . There are two measures of the effect of factor h_p . The average of these two measures is the main effect (E_y) of factor h_p . It measures the average effect of completion length when all the other factors are unchanged.

$$E_y = \frac{(G_p(0,1) - G_p(0,0)) + (G_p(1,1) - G_p(1,0))}{2}$$

Because of symmetry the tubing head pressure will also have two differences that would specify the main effect of p_{tf} on G_p .

The main effects of variables p_{tf} and h_p may not explain the total variability of G_p . Interaction between factors may play a significant role in explaining how the factors affect the response. In other words, if one factor change influences the other factor's effect in a consistent way, then they are said to be interacting. Interaction between two is one-half the difference in the average p_{tf} effects at the two levels of h_p (Myers and Montgomery, 1995).

$$E_{xy} = \frac{(G_p(1,1) - G_p(0,0)) - (G_p(0,1) - G_p(1,0))}{2}$$

These coefficients measure the effect of each factor or interaction term on the response and they are a measure of factor influence on the response. Coefficients can be obtained more easily and generally by regression, which uses matrices and the matrix algebra (below).

2.2. FITTING THE POLYNOMIAL MODELS

The relation between the input variables ($\zeta_1, \zeta_2 \dots \zeta_m$) and response (η) may be a function that is exactly known (Myers and Montgomery, 1995)

$$\eta = \varphi(\zeta_1, \zeta_2 \dots \zeta_m) + \varepsilon$$

where ε is the sum of the model and measurement errors. It is assumed that ε has mean zero and variance σ^2 . In general, the experimenter approximates the system function φ with an empirical model f of the form

$$\eta = f(\zeta_1, \zeta_2 \dots \zeta_k) + \varepsilon$$

where f is a first or second order polynomial and $k \leq m$. The variables are known as *natural variables* (Box and Hunter, 1978) since they are expressed in physical units of measurement. In the response surface method (RSM), the natural variables are transformed into coded variables $x_1, x_2 \dots x_k$, where the coded variables are usually defined to be dimensionless with mean zero and have the variance equal to the original variable. The response function now becomes

$$y = f(x_1, x_2 \dots x_k) + \varepsilon$$

where y is the expected value of η , which is denoted as $E(\eta)$. The successful application of RSM relies on the identification of a suitable approximation for f . This is commonly approximated as a first-order model with only main effects,

$$y = \beta_0 + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_k x_k$$

a first-order model with main effects and interactions,

$$y = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{i=1}^k \sum_{j>i} \beta_{ij} x_i x_j$$

or a second-order model of the form

$$y = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{i=1}^k \beta_{ii} x_i^2 + \sum_{i=1}^k \sum_{j>i} \beta_{ij} x_i x_j$$

Higher orders of x_i can be taken to mimic the response y . When considering the first-order or second-order model the coefficients β comprise the unknown parameter set, which can be obtained by running some experiments and tabulating the factors (inputs) and responses (output). This data can either be from physical experiments or from computer experiments. The parameter set can be estimated by regression analysis based upon these experiments.

The method of least squares is typically used to estimate the regression coefficients (Montgomery, 2001).

Each experiment gives a response y , results for n experiments can be written as $y_1, y_2 \dots y_n$. The number of experiments n must be at least as large as the number of coefficients to be estimated because the degrees of freedom available must be larger than the number of coefficients. For a linear model n should be at least $k+1$. If n is equal to $k+1$ the design is called a *saturated design*. For each response y_i there is a set of k

variables as the input. Let the j^{th} variable of i^{th} observation is denoted as x_{ij} , all the data for k variables and n runs can be tabulated in a standard form (Table 2.2).

Table 2.2 Data for Multiple Linear Regression

x_1	x_2	\dots	x_k	y
x_{11}	x_{12}	\dots	x_{1k}	y_1
x_{21}	x_{22}	\dots	x_{2k}	y_2
\vdots	\vdots	\dots	\vdots	\vdots
x_{n1}	x_{n2}	\dots	x_{nk}	y_n

For observation i a linear model is written

$$\begin{aligned}
 y_i &= \beta_0 + \beta_1 x_{i1} + \beta_2 x_{i2} + \dots + \beta_k x_{ik} + \varepsilon_i \\
 &= \beta_0 + \sum_{j=1}^k \beta_j x_{ij} + \varepsilon_i, \quad i = 1, 2, \dots, n
 \end{aligned}$$

Assumptions are that the error term ε in the model has the properties, $E(\varepsilon) = 0$, $Var(\varepsilon) = \sigma^2$ and $\{\varepsilon_i\}$ are uncorrelated random variables. Matrices provide a convenient way to write sets of equations. The linear model can be written in matrix form as

$$\mathbf{y} = \mathbf{X}\boldsymbol{\beta} + \boldsymbol{\varepsilon}$$

where

$$\mathbf{y} = \begin{bmatrix} y_1 \\ y_2 \\ \vdots \\ y_n \end{bmatrix}, \quad \mathbf{X} = \begin{bmatrix} 1 & x_{11} & x_{12} & \dots & x_{1k} \\ 1 & x_{21} & x_{22} & \dots & x_{2k} \\ \vdots & \vdots & \vdots & & \vdots \\ 1 & x_{n1} & x_{n2} & \dots & x_{nk} \end{bmatrix}$$

$$\boldsymbol{\beta} = \begin{bmatrix} \beta_0 \\ \beta_1 \\ \vdots \\ \beta_k \end{bmatrix}, \quad \text{and} \quad \boldsymbol{\varepsilon} = \begin{bmatrix} \varepsilon_1 \\ \varepsilon_2 \\ \vdots \\ \varepsilon_n \end{bmatrix}$$

where \mathbf{y} is an $(n \times 1)$ vector of the responses observed, \mathbf{X} is an $(n \times (k+1))$ matrix containing column of ones and the $(n \times k)$ design matrix, $\boldsymbol{\beta}$ is a $((k+1) \times 1)$ vector of the regression coefficients, and $\boldsymbol{\varepsilon}$ is an $(n \times 1)$ vector of random errors. The method of least squares chooses the $\boldsymbol{\beta}$ so that the sum of the squares of the errors, ε_i are minimized. The

least squares function $L = \sum_{i=1}^n \varepsilon_i^2$ should satisfy the condition

$$\left. \frac{\partial L}{\partial \boldsymbol{\beta}} \right|_{\mathbf{b}} = \mathbf{0}$$

which means first derivative of the function L by all the coefficients estimators (b_i) should be equal to zero. After simplifying the result and solving the equations, the least square estimator of $\boldsymbol{\beta}$, \mathbf{b} is obtained by

$$\mathbf{b} = (\mathbf{X}^T \mathbf{X})^{-1} \mathbf{X}^T \mathbf{y}$$

The fitted regression model is

$$\hat{\mathbf{y}} = \mathbf{X} \mathbf{b}$$

In scalar notation, the fitted model is

$$\hat{y}_i = b_0 + \sum_{j=1}^k b_j x_{ij}$$

The difference between the actual value y_i and the estimated value \hat{y}_i of an observation is called residual error and it is denoted as e_i .

$$e_i = y_i - \hat{y}_i$$

and in matrix form the vector \mathbf{e} is of size $(n \times 1)$

$$\mathbf{e} = \mathbf{y} - \hat{\mathbf{y}}$$

The sum of squares of the residuals gives variance of the estimate.

$$\begin{aligned} SS_E &= \sum_{i=1}^n \left(y_i - \hat{y}_i \right)^2 \\ &= \sum_{i=1}^n e_i^2 \\ &= \mathbf{e}^T \mathbf{e} \end{aligned}$$

The expected value of the SS_E is the variance estimate and it is given by

$$\hat{\sigma} = \frac{SS_E}{n - p}$$

where $p = k+1$.

2.3. ALTERNATIVE REGRESSION METHOD

DOE/RSM is used to construct the models or approximations for the solutions. Polynomial regression is frequently used for such approximations. Kriging is presented as an alternative approximation for the design and analysis of computer experiments (Kleijnen and Beers, 2003; Simpson *et al.*, 1997). Both polynomials and kriging are applied to analysis of reservoir performance in this study. Landa *et al.* (2003) provide a good reference for kriging applications in reservoir engineering.

Deterministic computer experiments, which lack random error, could be approximated by a more appropriate statistical method like kriging (Sacks *et al.*, 1989). Kriging is maximum likelihood interpolation technique to predict regionalized variables (Simpson *et al.*, 1997). More precisely, a kriging prediction is a weighted linear combination of all output values already observed (Chapter 1). The weights depend on the distances between the locations to be predicted, the locations already observed, and a

model for how the response is correlated in the factor space (that is, covariance). Kriging assumes closer data are more positively correlated. This assumption is modeled through the semivariogram, discussed below.

Kriging predictions typically assume second-order stationarity. That is, the covariances of the observations depend only on the distance between the data and not on the absolute location. Usually, covariances decrease if distance between the observations increases. This concept is used to build a variogram as explained below. Kriging minimizes the mean squared prediction error using the variogram to assign greater weights to observations that are closer to the prediction point and incorporates data redundancy. This ensures that when predicting the output for a location that has already been observed the prediction equals the observed value; this property is known as exact interpolation. This property is important in deterministic simulation. Kriging is the best linear unbiased (BLUE) estimator for correlated data.

Kriging relies on a variogram. The assumption of a second-order stationary implies that the variogram is a function of the separation vector \mathbf{h} between two locations only. When the distance $|\mathbf{h}|$ increases, the variance of the increment may asymptotically reach a constant value. A semivariogram (γ) is shown in Figure 2.1 and defined as

$$\begin{aligned} 2\gamma(\mathbf{h}) &= \text{Var}(Z(\mathbf{u} + \mathbf{h}) - Z(\mathbf{u})) \\ \gamma(\mathbf{h}) &= C(0) - C(\mathbf{h}) \end{aligned}$$

where Z is a random variable like permeability or porosity with \mathbf{u} as the coordinate vector and $C(\mathbf{h})$ being the stationary covariance.

The relation $\gamma(\mathbf{h})$ is modeled using various analytic forms; in this study, exponential and Gaussian functions are considered (Deutsch and Journel, 1998; Goovaerts, 1997).

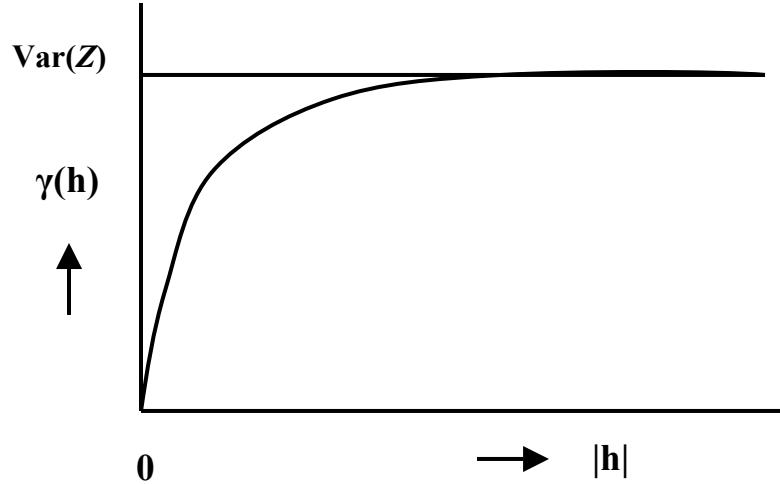


Figure 2.2 Example Semivariogram

A kriged estimate $y(\mathbf{x})$ can be conceived as a departure from a global model, $f(\mathbf{x})$, and the estimate can be written in the form (Chen and Simpson, 2000; Simpson *et al.*, 1997):

$$y(\mathbf{x}) = f(\mathbf{x}) + z(\mathbf{x})$$

The $f(\mathbf{x})$ term is similar to the polynomial model in a response surface and provides an approximation in the design space. In many cases $f(\mathbf{x})$ is approximated to be a constant term β_0 . $z(\mathbf{x})$ is a realization of a stochastic process assumed to be normally distributed with mean zero and variance σ^2 . The covariance matrix of $z(\mathbf{x})$ is given by

$$\text{Cov} [z(\mathbf{x}^i), z(\mathbf{x}^j)] = \sigma^2 \mathbf{R}$$

where \mathbf{R} is the correlation matrix. $R(\mathbf{x}^i, \mathbf{x}^j)$ is the correlation function which gives relation between observations (or runs or rows in design) \mathbf{x}^i and \mathbf{x}^j . $R(\mathbf{x}^i, \mathbf{x}^j)$ is commonly modeled as a Gaussian, an exponential or any other positive definite function. For example, a Gaussian function can be in the form

$$R(\mathbf{x}^i, \mathbf{x}^j) = \exp \left[- \sum_{m=1}^k \theta_m |x_m^i - x_m^j|^2 \right]$$

where $\boldsymbol{\theta}$ is a vector of parameters used to fit the model, x_m^i and x_m^j are the m^{th} factor levels of design runs \mathbf{x}^i and \mathbf{x}^j , n is the number of runs, and k is the total number of factors. The values of $\boldsymbol{\theta}$ can be found using the ranges obtained in the variogram and inverting those ranges. The estimator \hat{y} of the response $y(\mathbf{x})$ at untried values of \mathbf{x} is given by

$$\hat{y} = \hat{\beta}_0 + \mathbf{r}^T(\mathbf{x})\mathbf{R}^{-1}\left(\mathbf{y} - \hat{\beta}_0 \mathbf{1}\right)$$

where $\mathbf{1}$ is a column vector of ones of size n , \mathbf{y} contains the sample values of the response ($n \times 1$) and $\mathbf{r}^T(\mathbf{x})$ is the correlation vector of length n between an unsampled data at \mathbf{x} and the sampled data points (design points):

$$\mathbf{r}^T(\mathbf{x}) = \left[R(\mathbf{x}, \mathbf{x}^1), R(\mathbf{x}, \mathbf{x}^2), \dots, R(\mathbf{x}, \mathbf{x}^n) \right]^T$$

β_0 is estimated using

$$\hat{\beta}_0 = \left(\mathbf{1}^T \mathbf{R}^{-1} \mathbf{1} \right)^{-1} \mathbf{1}^T \mathbf{R}^{-1} \mathbf{y}$$

The estimate of the variance, σ^2 is

$$\sigma^2 = \frac{\left(\mathbf{y} - \hat{\beta}_0 \mathbf{1} \right)^T \mathbf{R}^{-1} \left(\mathbf{y} - \hat{\beta}_0 \mathbf{1} \right)}{n}$$

where $f(\mathbf{x})$ is assumed to be a constant value β_0

Validation of polynomial regression models is based on hypothesis testing like t -tests, f -tests, plotting residuals, and computing sum of squares (R^2) statistics. As explained in Sacks *et al.* (1998), this way of testing for consistency is inappropriate when analyzing deterministic computer experiments. It could be more suitable to analyze the fit by cross validation and integrated mean square error.

2.4. UNCERTAINTY ASSESSMENT

Uncertainty assessment is done to quantify how variability in the input parameters causes uncertainty in the response. This variability in the inputs may be caused due to insufficient data availability or approximations used in the model fitting. Monte Carlo analysis has been used to translate randomness in inputs to uncertainty in outputs by approximating the input probability density functions (*pdfs*) (Cullick *et al.*, 2003). That is, Monte Carlo sampling methods have been used to randomly vary parameters over ranges of values from specified probability distributions to generate response distributions.

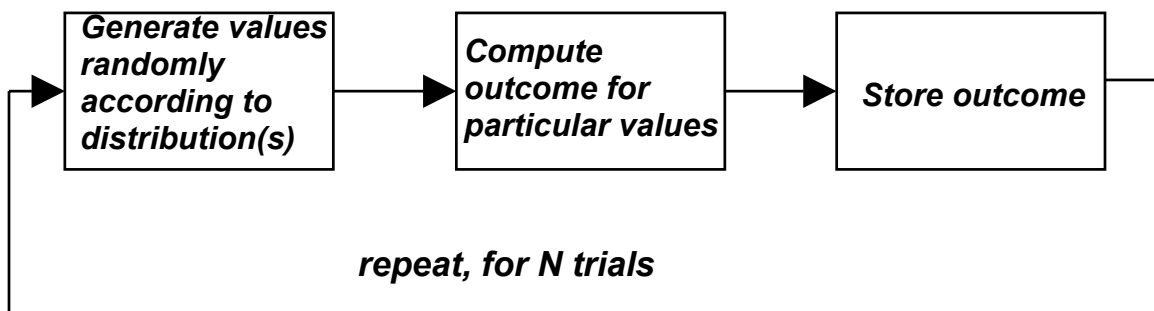


Figure 2.3 Monte Carlo Workflow

Sampling requires a sufficiently large number of model runs with randomly sampled inputs. Stochastic risk analysis could need thousands of simulation runs, especially if there are many uncertain variables present and an accurate response distribution is required. For complex systems, the cost of running so many simulations could be prohibitively expensive. But the model must assess a wide range of possible situations; wrong data assumptions can cause bad decisions and could have serious outcomes. A modified Monte Carlo method called *Latin Hypercube Sampling* (LHS)

uses stratified sampling to reduce the required number of runs. Hammersley sequences are an efficient, deterministic alternative to true (random) Monte Carlo; such methods are called quasi-Monte Carlo. Sampling techniques like the Latin hypercubes and Hammersley sequence can be used for risk and uncertainty analysis in complex systems.

CHAPTER 3. SAMPLING TECHNIQUES AND EXPERIMENTAL DESIGN

Experimental design chooses parameter combinations and the number of combinations required for efficient sets of experiments; this framework also aids parameter sensitivity analysis and response optimization. Response models can be fitted to relate the factors and responses. Response model accuracy depends on the number of levels of the variables and the pattern of experimental design points chosen – that is, how comprehensively the design fills the factor space. In addition, some models (e.g., appropriately scaled quadratic models) may be more accurate than others (e.g., a linear model). Choosing the models that best replicate the physical system at hand mainly is mathematically equivalent to selecting the model matrix (\mathbf{X}), which has rows corresponding to each design point and columns corresponding to each regressor in the model. Techniques to find a suitable model matrix include criteria such as orthogonality or minimizing the prediction variance of estimated parameters.

3.1. FACTORIAL DESIGNS

The design that includes all possible combinations of all levels all the factors is called a complete or full factorial design. If all the factors have two levels then they are called two-level full factorial designs, denoted as 2^k where k is the number of factors (Figure 3.1 is a 2^3 factorial). For a 2^k factorial and a linear model without interactions, the model matrix consists of $k+1$ columns (one column for each variable, plus one for the mean) and 2^k rows (one row for one experiment). Two level designs cannot capture nonlinear behavior but they can indicate major effects and so may guide future experiments. Full factorial designs having k factors require 2^k runs, which increases

exponentially as k is increased. Estimation redundancy often occurs, as some factors do not significantly affect the response; nonetheless, each added factor to a 2^k design doubles the number of runs required.

Higher-level designs can capture higher-order terms like quadratic and cubic terms in a response model. For example, at least three levels are required to infer quadratic terms. Even more levels are needed for higher-order terms, but many higher-order terms may be negligible. Thus it is reasonable to use quadratic models for most problems.

3.2. FRACTIONAL FACTORIAL DESIGNS

Full factorial designs form the basis for fractional factorial designs, which are less expensive alternatives obtained by *confounding* interaction terms. Confounding is a design property where the effects of the some influential factors and higher order effects are impossible to differentiate (Myers and Montgomery, 1995). In general, increasing fractionation increases confounding between effects of various orders.

Confounding creates fractional factorial designs that require fewer experiments (compared with full factorials). In some situations, only a few of the variables influence the response; in such a case fractional factorial designs can be much more efficient than full factorials. Some parameters that do not contribute to the response can be identified in a process called screening.

A complete factorial requires 2^k runs. But with a 2^k factorial design the grand average, the main effects, and the interactions with (total degrees of freedom, $\nu = 1 + k + k(k - 1) / 2$, can be estimated assuming the remaining higher-order interactions are

negligible. If the high-order interactions are negligible, the main effects and low-order interactions can be estimated from only a fraction of the factorial design.

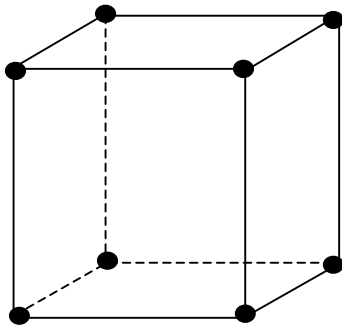


Figure 3.1 2^3 Factorial Design

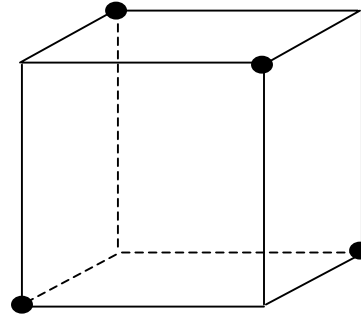


Figure 3.2 2^{3-1} Fractional Factorial Design

Consider a situation with four factors, each at two levels is of interest. But the resources allow only eight runs, which is a one-half fraction of the possible $2^4=16$ combinations required for a factorial design. The eight combinations (Table 3.1) give a half fraction of the complete 2^4 design. A $\frac{1}{2}$ fraction of a 2^k factorial design is called 2^{k-1} fractional factorial design (Figure 3.2 is a 2^{3-1} fractional factorial design). Changing the signs of one column and repeating the signs of other columns will give the other fraction. Similarly, a $(\frac{1}{2})^p$ fraction of a 2^k design is called a 2^{k-p} fractional factorial design. It is convenient to represent the levels of variables as -1 and +1 to understand these designs (Box and Hunter, 1961).

The first three columns and eight runs form a full factorial and the main effect d is identical to that to the three factor interaction abc (Table 3.1). The product abc refers to the multiplication of individual elements in the corresponding columns a , b , and c . Since the columns of the design matrix (Table 3.1) are identical for factor d and factor

interaction abc , these quantities have identical settings for all experiments. Thus, effects of d and abc cannot be distinguished and they are said to be confounded. The d effect really estimates the sum of d and abc effects (Chapter 2). As $d = abc$, abc is called the generator of this particular fraction. Only one generator exists for this design and the defining relation, $I = abcd$ is obtained by multiplying both sides of generator by d and using the condition that $d^2 = I$. The number of generators for a design is equal to the number of fractions p (Montgomery, 2001).

Table 3.1. 2^{4-1} Fractional Factorial Design

Runs	Design Matrix			
	a	B	c	$abc=d$
1	-1	-1	-1	-1
2	+1	-1	-1	+1
3	-1	+1	-1	+1
4	+1	+1	-1	-1
5	-1	-1	+1	+1
6	+1	-1	+1	-1
7	-1	+1	+1	-1
8	+1	+1	+1	+1

3.2.1. Aliasing and Confounding Effects

Multiplying any column by the defining relation yields the aliases for that effect. The alias structure for this design can be determined using the defining relation $I=abcd$. For example, the confounding of the main effect of a with an interaction (or interactions,

for more general partial factorials) can be identified by multiplying a by either side of the defining relation (or relations, for partial factorials that are smaller than half fractions).

$$a \cdot I = a \cdot abcd = a^2bcd = bcd$$

Similarly, the aliases of b and c are

$$b \cdot I = b \cdot abcd = b^2acd = acd$$

$$c \cdot I = c \cdot abcd = c^2abd = abd$$

This aliasing implies that the main effects of a , b , and c estimated using a 2^{4-1} factorial design are in fact the sum of main effects and three factor interactions $a + bcd$, $b + acd$, and $c + abc$ respectively. As a result, one cannot differentiate between main effects and 3-term interactions. These are called Resolution IV designs: main effects are not confounded with two factor interactions. Designs are often described using a notation such as 2_{R}^{k-p} where k is the number of factors, p is the fraction of the factorial, and R is the resolution. For this particular example design is denoted as 2_{IV}^{4-1} design.

Fractional factorial designs are classified into different resolution designs depending on the type of confounding. Frequently used resolutions are of type III, IV, and V. The higher the resolution lesser is the confounding, the lesser the assumptions, and clearer the interpretation. The three types of designs are

1. Resolution III: Main effects are not aliased with each other, but are aliased with two-factor interactions, and two-factor interactions are aliased with each other. Example for this resolution design is 2_{III}^{3-1} .

2. Resolution IV: Main effects are not aliased with each other or two-factor interactions but two-factor interactions are confounded with each other. An example of a Resolution IV design is a 2_{IV}^{4-1} , which is explained above.

3. Resolution V: Main effects or two-factor interactions are not aliased with any of the main effects or two factor interactions, but are confounded with three-factor interactions, e.g., 2_V^{5-1} .

Two-level designs can estimate main effects and interactions only. To estimate quadratic or higher-order effects higher-level designs are to be used. The simplest of these higher-order designs is a three-level factorial design where three levels are denoted as (0, 1, 2) or (-1, 0, 1). Full factorial three-level designs require 3^k experiments. In mixed-level design different factors have different levels. If k_i is the number of factors having level s_i then the total number of experimental runs is the sum of s^{k_1}, s^{k_2}, \dots , and s^{k_q} , where $k = k_1 + k_2 + \dots + k_q$ is the total number of factors.

Many factors can be investigated using fractions of full factorials, allowing screening of many factors to find the critical ones, which may be a small portion of the total factor set. Main effects (designs with resolution, $R \geq III$ only) and interactions (designs with resolution, $R \geq IV$ only) can be estimated unequivocally. Fractional factorial designs ranging from a $\frac{1}{2}$ replicate to a $\frac{1}{2048}$ replicate of 15 factors could be attained. These designs, like full factorials, are orthogonal. Aliasing structure of the factors gives information of confounding factors.

A (fully saturated) resolution III fractional factorial is easy to prepare by aliasing all interactions, and a resolution IV design can be generated by aliasing odd interactions with three or more terms (Box and Hunter, 1961). These designs are useful and easy to

generate. If the R-III design results indicate that some other effects which are confounded are important, other R-III designs can be added to generate a to R-IV design to identify the effects more accurately; this process is known as foldover. Similar techniques can be used to unconfound R-IV designs. Switching signs of a single column in the original design generates a new design of same resolution; this can be used as the foldover, with the column with changed signs selected to have the desired effect on the aliasing structure.

3.3. GEOMETRIC AND OPTIMAL DESIGNS

Optimal designs are developed using an optimality criterion. Because these designs are usually multilevel, they can inspect main effects and interactions as well as quadratic effects. Optimal designs can also increase the number of runs in the course of analysis to improve model fitting and also to fit polynomial models over irregular factor domains (*i.e.*, they are not restricted to hypercubes).

Several optimal criteria have been proposed; one of the most popular design criterions is to maximize the determinant of $\mathbf{X}^T\mathbf{X}$, where \mathbf{X} is the design tableau with column rank equal to then number of coefficients to be estimated. If the determinant $|\mathbf{X}^T\mathbf{X}|$ is maximized then the regression estimation variance is minimized. A design for which $|\mathbf{X}^T\mathbf{X}|$ is maximized is called a D-optimal design. Unlike factorial designs these designs are not orthogonal; this implies that the coefficients estimated in the regression model will be correlated. A design is A-optimal if it minimizes the sum of the diagonal elements (also called the trace) of $(\mathbf{X}^T\mathbf{X})^{-1}$. A-optimality minimizes the variances of the regression coefficients. Another optimality criterion minimizes the maximum variance of the predicted response over the design region; this property is called G-optimality.

Optimal designs are based on algorithms that try to keep the distance between design points as high as possible; in general, increasing the moments of the design improves estimation error (Myers and Montgomery, 1995). The computer algorithm generally uses an exchanging process to select the optimal design. That is, given the number of runs and the model that needed to be fitted; the best possible design is chosen depending on the criterion (can be multi-level and mixed-level designs).

The Plackett-Burman designs are two-level fractional factorial designs for studying up to $k = N - 1$ variables in N runs, where N is a multiple of four. These Resolution III designs are saturated designs, as all the degrees of freedom are used to estimate effects. So, these designs are perfect for screening to determine the factors dominating the response. For $N = 12, 20, 24, 28,$ and 36 the Plackett-Burman designs are especially useful (Hedayat *et al.*, 1999). These designs have complicated alias structures. For example, in the 12-run design every main effect is partially aliased with every two-factor interaction not involving itself and also each main effect is partially aliased with 45 two-factor interactions. Because of the complex aliasing structure interpreting the results with confidence is difficult – thus they are limited to screening.

A central composite design (CCD) contains an imbedded factorial design with center and a group of axial points that allow estimation of curvature (Feng and White, 2002). CCD could be three or five level designs depending on axial point locations. If the axial points are located on the faces of the cube then it is a three level design. Otherwise, it is a five level design (Figure 3.4). If the distance from the center of the design to an edge is ± 1 then the distance from the center to an axial point is $\pm\alpha$ where $\alpha \geq 1$. There are normally five levels in the CCD for each design factor, $(-\alpha, -1, 0, 1, \alpha)$. The precise

value of α depends on certain properties preferred for the design like rotatability and the number of factors present (e-Handbook of Statistical Methods, 2005). The number of axial points is double the number of factors involved in the design. To maintain rotatability, the value of α depends on the number of factorial runs in the central composite design.

$$\alpha = [\text{number of factorial runs}]^{1/4}$$

The Box-Behnken design is for quadratic regression models and is a three level design (Figure 3.4). In this design the treatment combinations are at the midpoints of hyperedges of the design space and at the center (Box and Behnken, 1960). These designs are rotatable. For three factors or lower, the Box-Behnken design requires fewer runs than a fractional factorial or Central Composite Designs. For four or more factors, this advantage disappears (e-Handbook of Statistical Methods, 2005). This design has a very good distribution of data points in the design space. Box-Behnken design is preferred over CCD if the vertices of the factor space are not influential.

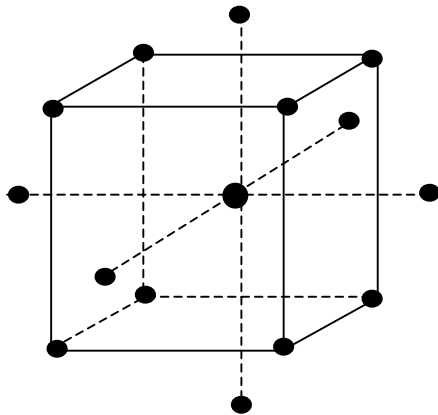


Figure 3.3 Central Composite Design

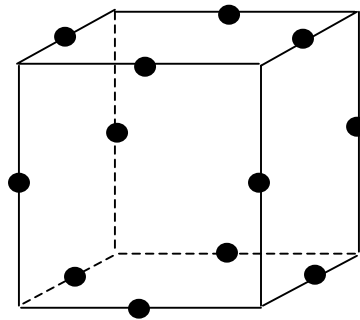


Figure 3.4 Box-Behnken Design

3.4. WHY ANOTHER DESIGN?

The idea of replications to assess random experimental errors is essential in the classical design of experiments (DOE), but this focus becomes irrelevant for computer experiments as explained by Sacks *et al.* (1989). The main requirement in deterministic computer experiments is that the design should span the design space. The replicated design points (as in CCD and many Box-Behnken designs) do not give extra accuracy in computer experiments or allow better error estimates. The three-level/five level designs of the CCD/Box-Behnken design can cause problems as the numbers of factors increases, as they require at least as many runs as two-level full factorial designs. The Plackett-Burman design has to be used only when the confounding situation is known, or for screening.

Another design problem is selecting the number of runs for a particular level combination, so that some n parameters are estimated. Suppose there are k factors under investigation and we plan to study the main effects and q interactions (where $q \leq \binom{k}{2}$). So the number of parameters to be estimated are $n = 1 + k + q$. In general, a saturated design with n runs is hard to get from CCD/Box-Behnken designs. Methods like graph-aided design by Taguchi (1960), and Hedayat and Pesotan (1992) fractionate and get supersaturated designs or very few runs from full factorials, but generating such designs remains difficult.

3.5. ORTHOGONAL ARRAYS

Orthogonal arrays offer an alternative approach. Orthogonal arrays are a generalized form of Latin hypercube sampling (McKay *et al.*, 1979). Orthogonal arrays

share certain desirable properties with Latin hypercubes while allowing additional desirable statistical properties to be obtained. Most noteworthy is that for certain response models they are universally optimal for the estimation of some or all the effects that appear in the model.

Orthogonal arrays (OA) are combinatorial structures that have nonnegative integers as its structure elements. Since their introduction by C. R. Rao in the 1940's, these design matrices have been extensively used in statistics, mainly in experimental design. OA's are considered the most efficient designs for a particular number of runs. Other OA applications include universal hashing, authentication codes, derandomization of algorithms, and perfect local randomizers (Hedayat *et al.*, 1999).

An orthogonal array is an $n \times k$ collection of elements in n rows and k columns in which the i^{th} column has s^i levels. The array $n \times k$ is arranged such that all sets of t columns in the array have equal number of rows with the same permutation of elements. The OA is called mixed if the levels of different factors (or columns) are different. Orthogonal arrays can be denoted as $OA(n, k, s, t)$ (Hedayat *et al.*, 1999). For mixed designs, the notation is $OA(n, k, s^{k_1} s^{k_2} \dots s^{k_q}, t)$ where $k = k_1 + k_2 + \dots + k_q$. According to the definition n should be equal to λs^t where λ is called the index of orthogonal array. Arrays of strength $t > 2$ require of more runs and hence would seem to be of less practical interest right now. Strength is related to the resolution and the projection properties of a design. Strength is equal to resolution minus one. For example, Resolution IV design has strength three. The higher the strength, the higher the resolution and the better the design.

Hedayat *et al.* (1999) gives a necessary and sufficient condition for a matrix $n \times k$ to be an orthogonal array:

$$\sum_{u=\text{row of } A} \zeta uv^T = 0$$

where vector v contains w non-zero entries from s , $\zeta = e^{(2\pi i/s)}$, and uv^T is evaluated at modulo s . An example OA (8, 4, 2, 3) is given below in Table 3.2.

Table 3.2. Example of an OA (8, 4, 2, 3)

Runs	Design Matrix			
	1	2	3	4
1	0	0	0	0
2	0	0	1	1
3	0	1	0	1
4	0	1	1	0
5	1	0	0	1
6	1	0	1	0
7	1	1	0	0
8	1	1	1	1

Sometimes the OA available for a set of factors and level combinations can have an acceptably large number of experiments, or no OA may exist. For example, an OA (18, 9, 2¹3⁸, 2) may be desired, but only OA (36, 9, 2¹3⁸, 2) is currently available. In some cases, for a set of levels and factors an OA doesn't exist, e.g., OA (36, 14, 3⁶4⁸, 2). In such conditions nearly orthogonal arrays (NOA) can be used. NOA (36, 14, 3⁶4⁸, 2) and NOA (18, 9, 2¹3⁸, 2) designs are possible. The number of runs can also be decreased by the use of nearly orthogonal designs. In these designs, columns will not be perfectly

orthogonal, which will (as for optimal designs) result in correlation of the regressed coefficients.

NOA have been used by Wang and Wu (1992) and Nguyen (1996). An algorithm developed using A-optimality is used in this research to develop OA/NOA. This algorithm is flexible for constructing various mixed-level designs and outperforms existing algorithms in speed and efficiency (Xu, 2002).

3.5.1. Benefits of OA/NOA

1. Many factors can be studied and conclusions are valid over the entire region spanned by the control factors.
2. Large savings in the experimental effort result from decreasing number of simulations.
3. Analysis is easy as columns are orthogonal (OA only).
4. OA is an adaptive design. Changing the values of the factors can refine OA design, columns (factors) can be eliminated, and new columns can be added. Previous runs are efficiently re-used.
5. Wide ranges of factors and levels can be used, especially with NOA.
6. Excellent screening designs can be generated by two level supersaturated designs.
7. OA designs can be called adaptive designs. Quadratic and interaction effects can be evaluated by augmenting another design with same levels and factors as the linear effects design.

3.5.2. Algorithm for OA/NOA

If no orthogonal array exists or the number of runs is to be decreased as explained above, then nearly orthogonal array can be used. NOA partially confounds the factors.

The NOA/OA algorithm used in this research is based on J_2 criterion (Xu, 2002). In this algorithm all levels in the design appear equally often in the array, leading to a balanced design.

3.6. LATIN HYPERCUBES AND THE HAMMERSLEY SEQUENCE

Latin hypercube and Hammersley sequences are additional alternative designs. Much of the work in analysis of computer experiments (e.g., the classic Sacks *et al.*, 1989) used Latin hypercubes as a design for obtaining response surfaces. They are not as efficient as OA/NOA but the ease of computation and opportunity to use many levels are appealing. Sensitivity and uncertainty analysis can be done at the same time. Coupled with hyperkriging, the results could be flexible and efficient.

3.6.1. Latin Hypercube Sampling (LHS)

If it is not clear which factors are important (where mixed level OA/NOA designs cannot work) and if few parameters control the response, then Latin hypercube sampling may be used. In LHS, each variable or constraint is divided equally over the range. If n runs are planned and if the range of a parameter is R then the whole parameter space is divided into R/n equal parts and bifurcation is done for all the parameters (McKay *et al.*, 1979). This is a random sampling method, unlike OA/NOA, optimal designs, or factorials. The notation followed is LHS (n, k) where n is the number of runs and k is the number of variables. An example of LHS (5, 4) is shown below in Table 3.3

LHS (McKay *et al.*, 1979) gives a sample data set for a group of uncorrelated variables in which the univariate characteristics of the original data are reproduced almost exactly. LHS is a stratified sampling technique where the random variable distributions

are divided into equal probability intervals. A probability is randomly selected from within each interval, inverted and combined with other variables randomly.

Table 3.3 Latin Hypercube LHS (5, 4)

Runs	Factors			
	1	2	3	4
1	3	2	1	0
2	1	3	4	2
3	0	4	2	3
4	2	1	0	4
5	4	0	3	1

LHS is a frequently used sampling technique for kriging. As in classical designs, LHS does not specify the relation between input and output by a function as in polynomial regression. Instead, the design points are selected to cover the whole k -dimensional space. It assumes a prior distribution of the input variables like Monte Carlo sampling and selects the sample by inverting those distributions.

The CDF or PDF range of a variable is divided in to n intervals. As the CDF y -axis is uniform, each interval of n will have equal probability. One value is taken from each interval in a random manner and inverted to the respective distribution. These n values from each parameter will be combined again with other parameters randomly to generate the LHS sample.

To understand the LHS technique, consider an example where it is desired to generate a LHS of size $n = 5$ with two input variables. Let us assume that first random

variable X_1 has a normal distribution with a mean value of μ and σ^2 . The intervals for $n = 5$ should satisfy the condition:

$$P(0 \leq X_1 \leq A) = P(A < X_1 \leq B) = P(B < X_1 \leq C) = P(C < X_1 \leq D) = P(D < X_1 \leq 1) = 0.2$$

where A , B , C , and D are the boundaries of the intervals; one would use an inverse Gaussian function to compute X_i from P_i . Thus the five intervals correspond to 20% probability each. Also assume that the second variable, X_2 , has a uniform distribution. As above, the whole probability range is divided into five equiprobable regions. The design can be obtained by picking specific values of X_1 and X_2 in each of the five intervals in a random manner. Because the elements are selected in equal probability regions, the selection will reflect the density across the interval. For example, the third interval for X_1 will have a higher probability density than intervals abutting A or D because the density in the normal distribution is high around the mean value. Next, the selected values of X_1 and X_2 are paired to form the design with five runs and two variables.

A Latin hypercube can be thought of as an orthogonal array. OA of same size $n \times k$ with $\lambda = 1$ gives $n = s^t$. If s is equal to n then $t = 1$ means that the strength is equal to one and index is one. In $n \times k$ Latin hypercube there will be n levels in each column and index is also one, so a Latin hypercube is an OA of strength one. The strength equals one in Latin hypercubes signifies that it fills the one-dimensional space efficiently but not in a high dimensional hyperspace, which an OA can fill efficiently.

3.6.2. Hammersley Sequence Sampling (HSS)

A Hammersley sequence is a low-discrepancy sequence. Low-discrepancy sequences of n sample points fill n dimensional space more uniformly than uncorrelated random points (MathWorld, 2005). Both uniform random generators and quasi-random sequences produce n -tuples on a uniform random scale, but Hammersley sequences are quasirandom; the design point m (where $1 \leq m \leq n$) is conditioned on the $m-1$ points and total sample points n . A uniform random generator over the interval $(0, 1)$ produces values within the subintervals with equal probability. For example in the subintervals $(0, 1/2)$ and $(1/2, 1)$, if n values are generated in the first half by chance, there is a probability of half that next $n+1^{\text{th}}$ point selected could be in the first half. This happens because the $(n+1)^{\text{th}}$ point does not know the location of n points. In quasirandom sequences, the samples are low-discrepancy so that the points generated in a highly ordered manner, but not on a regular lattice and still possessing some desirable properties of random sequences in terms of estimating expected values and variance.

Hammersley sequences have better dispersion properties but the number of points to be generated must be specified in advance. Thus, if a Hammersley sequence of length 100 is generated for analysis, and accuracy is then found unacceptable, then more samples must be generated. In such a case, the analyst must discard current values and start over. This same limitation applies to stratified schemes such as Latin hypercube sampling. Low discrepancy (quasirandom) sequences are used in numerical integration, simulation and optimization. They are like uniformly distributed random numbers without statistical independence, and are designed to give more uniformity in

multidimensional space. Therefore they are often more efficient than random numbers in multidimensional Monte Carlo or modified Monte Carlo methods

Quasirandom Hammersley sequences require fewer runs than Latin hypercube or Monte Carlo sampling to attain similar levels of accuracy in estimating expectations and variances. Correlations built between HSS columns are close to zero but in Latin hypercubes significant correlations commonly exist between the columns, degrading sampling efficiency. As a result, the Hammersley sampling method can span the k -dimensional space with a relatively small but still representative sample (Kalagnanam and Diwekar, 1997).

Any nonnegative integer n can be expanded using a prime base p :

$$n = a_0 + a_1 p + a_2 p^2 + \dots + a_r p^r$$

where each a_i is an integer in $[0, p-1]$.

Now define a function Φ_p of n by

$$\phi_p(n) = \frac{a_0}{p} + \frac{a_1}{p^2} + \frac{a_2}{p^3} + \dots + \frac{a_r}{p^{r+1}}$$

Hammersley points in a k -dimensional space can be given by

$$\vec{x}_k(n) = \left(\frac{n}{N}, \phi_{R_1}(n), \phi_{R_2}(n), \dots, \phi_{R_{k-1}}(n) \right)^T$$

where $k-1$ prime numbers are used for k factors

3.7. INDUCING CORRELATION STRUCTURE INTO THE DESIGNS

Iman and Conover (1982) used rank correlations to preserve the correlation between variables. If \mathbf{R} is the design having some correlation \mathbf{T} and \mathbf{C} is the desired

correlation matrix then a new design \mathbf{R}^* can be generated by rearranging the column elements in \mathbf{R} to obtain a correlation matrix close to the desired correlation matrix \mathbf{C} .

As \mathbf{C} is positive definite and symmetric, it may be decomposed by Cholesky Decomposition as $\mathbf{C} = \mathbf{P}\mathbf{P}^T$ (\mathbf{P}^T is the transpose of \mathbf{P}). Then the matrix $\mathbf{R}^* = \mathbf{R}\mathbf{P}^T$ has the desired correlation. The input matrix \mathbf{R} may not have correlation matrix $\mathbf{T} = \mathbf{I}$ (orthogonal and thus equal to the identity matrix). In such cases more manipulations on \mathbf{T} could help to get zero correlation for uncorrelated variables. The advantages of this algorithm include:

1. The variance and mean for all the variables are the same as HSS or LHS design as the algorithm is not changing the deviates but only changing the combinations.
2. Zero correlation between the variables can be approximately achieved.
3. It is very easy to implement and code.

This algorithm is also used in the LU-simulation method (Deutsch and Journal, 1998) and has been used in recent research (Kurniawan, 2005). However, factors are modeled as uncorrelated in this thesis.

3.8. NUMBER OF SAMPLES OR RUNS

Performance of the designs or sampling techniques depends on the sample size. The larger the sample the more accurate the response model will be. But beyond some threshold number of runs, the efficiency of the design does not increase significantly. Under such conditions, it is better to run the design with fewer runs. The general question is: "How many runs are ideal for a given case?". The answer depends on how complex the system is and how expensive the runs are. As the complexity of the system increases, the sample data set should be denser. For time-consuming problems, it is better to use

fewer samples. For designs, a second order polynomial has $m = (k + 1)(k + 2)/2$ coefficients to estimate and the number of experiments should be at least m . Giunta *et al.*, (1994) found that for a reasonably accurate regression, a useful empirical guideline is about $1.5m$ runs for 5-10 variables, $3m$ runs for 10-20 variables, and $4.5m$ runs for a 20-30 variables. These guidelines help choose appropriate designs for reservoir studies with various numbers of regressors.

For Latin Hypercubes, Chen and Simpson (2000) suggested $3k$ runs for expensive experiments, scarce data set, $10k$ for moderately expensive experiments, and $3(k + 1)(k + 2)/2$ for large data sets typical of low-cost experiments. These relations are reasonable and could be crosschecked by sample points as in Sacks *et al.* (1989). These relations would be very useful as a part of a software interface for sampling.

A particular design for same number of runs can be chosen using the D-Optimality criterion as explained before. The set of points that maximize the $|\mathbf{X}^T \mathbf{X}|$, is the set of points that minimizes the maximum prediction variance and also the set of points that minimizes the variance of the parameters estimated. Low discrepancy samples (for example, Hammersley sequences) might be used as starting points for optimization of such designs.

CHAPTER 4. COMPARISON OF DESIGNS

In this chapter the uniformity and space filling characteristics of different designs are compared and analyzed for choosing better designs in the reservoir applications. First, the properties of Latin hypercubes and orthogonal arrays are compared using two-factor examples. Second, D-Optimality calculations compare OA and LHS designs. A water-drive gas reservoir problem is selected to compare orthogonal arrays and a full factorial to quantify possible computational savings. Finally, the convergence properties of different sampling techniques (Monte Carlo, Latin hypercube, and Hammersley sequence) are compared.

4.1. COMPARING ORTHOGONAL ARRAYS AND LATIN HYPERCUBES

Space-filling designs with fewer points that accurately describe system responses are the goal of model building. A unit square where two variables are represented by two axes compares OA and LHS space filling properties. An orthogonal array with parameters OA (9, 2, 3, 2) *i.e.*, two variables and nine runs and Latin hypercube LHS (9, 2) are shown graphically (Figure 4.1; Sandor and Andras, 2004). The criterion for OA design is to distribute the nine points as far apart as possible so that design covers the factor space. Response models could be computed using the samples selected by the two methods. However, there is no point in the lower middle in the LHS (9, 2). Therefore, if the function in this area has values that are very different from the average, then the estimate of the response in that zone will have a large error. This deficiency can

be avoided if the sample fills the space more uniformly as the OA does (Space filling characteristics are shown in Figure 4.1).

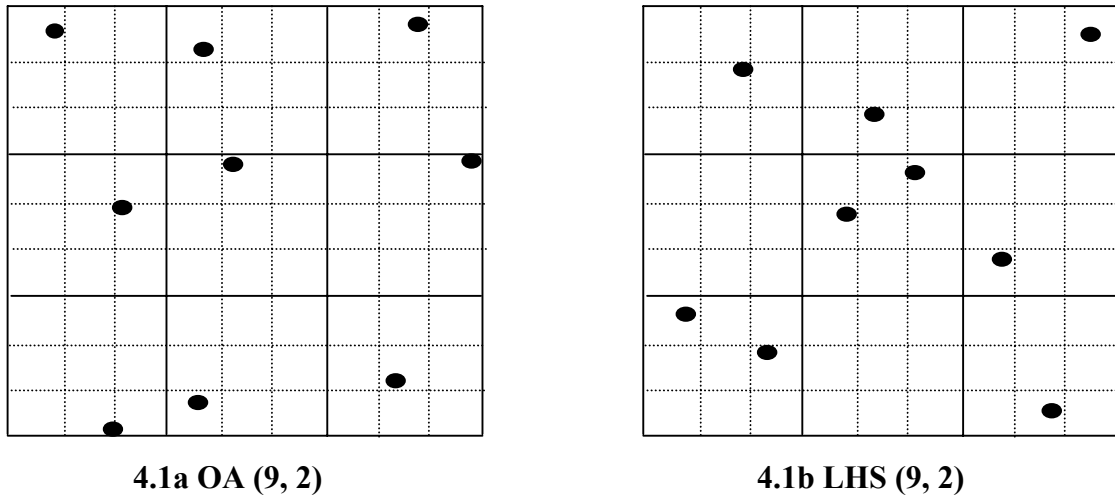


Figure 4.1 Sample Points on a Unit Square

Orthogonal arrays and Latin hypercubes are compared for linear and quadratic models. D-Optimality criterion is used in generating D-optimal designs (maximizing $|\mathbf{X}^T\mathbf{X}|$ or minimize $|\mathbf{X}^T\mathbf{X}|^{-1}$). The determinant of the information matrix $\mathbf{X}^T\mathbf{X}$ is computed and inverted; the design with lower D-Optimality is better. D-Optimality minimizes the maximum prediction variance (Chapter 3). OAD is the D-optimality of the orthogonal array and LHD is the D-optimality of the Latin hypercube (Table 4.1). Two models are considered for comparison, linear model with linear effects for the two level designs and quadratic model with interactions and quadratic effects for higher level designs. Orthogonal arrays are taken from an online library (Sloane, 2005) and Latin hypercubes are generated as explained in the chapter three. In each table row, the designs have the same number of runs. In all the cases, the OA D-optimality is considerably less (*i.e.*,

better) than the LHS D-Optimality. This indicates that OA designs will be more accurate than the LHS designs.

Table 4.1 D-optimality Test Results for OA and LHS

Model			$y = \beta_0 + \sum \beta_1 X_1$		
Two levels and strength 2			Two levels and strength 3		
OA Size	OAD	LHD	OA Size	OAD	LHD
4.3.2.2	0.25	5.520593	8.4.2.3	0.007813	1.2269445
8.5.2.2	0.003906	3.274702	16.8.2.3	9.54E-07	0.0815949
8.7.2.2	0.000977	51.07244	24.12.2.3	1.91E-11	0.000244
12.11.2.2	4.7E-07	211.0781	32.16.2.3	1.11E-16	1.793E-07
16.15.2.2	5.82E-11	148.3284	40.20.2.3	2.5E-22	2.73E-10
20.19.2.2	2.62E-15	32.08318	48.24.2.3	2.62E-28	5.369E-14
			56.28.2.3	1.45E-34	1.129E-17
Model			$y = \beta_0 + \sum \beta_1 X_1 + \sum \beta_{11} X_{11}^2$		
Three and strength 2			More than 3 level of Fixed Strength		
OA Size	OAD	LHD	OA Size	OAD	LHD
9.4.3.2	5.36E-06	0.991747	16.5.4.2	1.86E-14	3.644E-10
18.7.3.2	9.46E-14	7.7E-06	32.9.4.2	3.39E-30	4.872E-22
27.13.3.2	1.36E-28	1.02E-09	64.21.4.2	1.99E-80	1.841E-60
36.13.3.2	5.76E-32	3.06E-17	64.6.4.3	8.67E-25	7.58E-21
			25.6.5.2	2.18E-23	3.088E-19
			50.11.5.2	4.94E-48	2.793E-40
			49.8.7.2	6.56E-43	6.03E-39
			98.15.7.2	1.13E-87	2.303E-80
			64.9.8.2	2E-53	2.147E-49
			81.10.9.2	2.19E-64	1.034E-60
			162.19.9.2	1.1E-131	1.6E-124
			121.12.11.2	1.68E-87	8.514E-84

4.2. COMPARING ORTHOGONAL ARRAYS AND FULL FACTORIAL DESIGNS

In this section, the use of OA for response surface modeling is investigated. A full factorial design having five parameters (three four level and two three level) Box-Tidwell

powers, and response (recovery factor) is taken from Armenta *et al.* (2003) an OA for this model is validated. The full factorial design has 576 simulation runs and response is the gas recovery (G_p). The factors shown below are initial reservoir pressure (psia), permeability (md), permeability ratio (vertical / horizontal), aquifer support (gas reservoir size / aquifer size), and perforation ratio (perforated length / total gas thickness).

$$\text{Reservoir Pressure Factor : } P_{iD} = \frac{(P_i - P_{i \min})}{(P_{i \max} - P_{i \min})}$$

$$\text{Permeability Factor : } K_{hD} = \frac{(\ln(K_h) - \ln(K_{h \min}))}{(\ln(K_{h \max}) - \ln(K_{h \min}))}$$

$$\text{Permeability Ratio Factor : } K_{zD} = \frac{(\ln(K_z) - \ln(K_{z \min}))}{(\ln(K_{z \max}) - \ln(K_{z \min}))}$$

$$\text{Aquifer Support Factor : } V_{aD} = \frac{(V_a^{-1/2} - V_{a \min}^{-1/2})}{(V_{a \max}^{-1/2} - V_{a \min}^{-1/2})}$$

$$\text{Perforation Ratio Factor : } H_{pD} = \frac{(H_p - H_{p \min})}{(H_{p \max} - H_{p \min})}$$

Scaling makes the factors approximately linearly correlated to gas recovery where P_i is the pressure, K_h is the permeability, K_z is the anisotropy, V_a is the aquifer volume, and H_p is the completion length. The regression model fit may include linear, quadratic, and interactions of these terms.

The powers on the factors are estimated from Box-Tidwell transforms which linearize the relation between the factors and the response (e.g., the number -0.5 on the factor V_a). If the powers are below zero (as for the aquifer size), a hyperbolic relation is implied and power approximately equal to zero implies a logarithmic transformation is implied. Polynomial regression has been used to fit a model relating inputs and output using full factorial.

$$\begin{aligned}
G_p = & 58.78 + 33.65 P_{iD} + 55.47 K_{hD} - 16.57 K_{zD} - 19.86 V_{aD} - 44.66 H_{pD} \\
& - 11.82 P_{iD}^2 - 33.04 K_{hD}^2 + 2.44 K_{zD}^2 + 0.70 V_{aD}^2 + 23.35 H_{pD}^2 - \\
& 26.70 P_{iD} K_{hD} + 3.16 P_{iD} K_{zD} - 2.92 P_{iD} V_{aD} + 12.82 P_{iD} H_{pD} + 17.82 K_{hD} K_{zD} \\
& + 8.37 K_{hD} V_{aD} + 13.05 K_{hD} H_{pD} - 12.86 K_{zD} V_{aD} + 0.73 K_{zD} H_{pD} + 10.29 V_{aD} H_{pD}
\end{aligned}$$

This is complex model with many significant interactions and quadratic terms; the R^2 is 0.8724 and R^2_{Adj} is 0.8678. Equations for R^2 and R^2_{Adj} are

$$R^2 = \frac{SS_R}{SS_T}$$

$$R^2_{Adj} = 1 - \frac{SS_E / (n - p)}{SS_T / (n - 1)}$$

The total sum of square SS_T can be divided in to sum of squares due to model and sum of square due to residuals.

$$SS_T = SS_R + SS_E$$

R^2 is the reduction in the variability of the response because of the regression model.

High R^2 alone does not mean that the model is good. Increasing the number of coefficients of freedom always increases R^2 . So, another statistic called R^2_{Adj} is used. This statistic will not always increase as number of coefficients increase and also it could decrease if unnecessary terms are used in the design. R^2_{Adj} is a good measure in comparing the fit of the model. Both the SS_E and SS_R are independent and distributed as χ^2 . The F statistic compares model to residual variance,

$$F = \frac{SS_R / (k)}{SS_E / (n - k - 1)}$$

The F statistic for this model is 189.7 and $Pr(F)$ is 2.2e-16, which says that the model is highly significant – the probability of such a low proportion of variance being left after the model is very small.

A NOA (36, 5, 3² 4³, 2) is used to compare the orthogonal array design with the factorial result discussed above. The chosen NOA design (Chapter 3) has 36 instead of 576 simulation runs used for the full factorial design. Scaling is as for the factorial design and the equation fitted is

$$\begin{aligned}
 G_p = & 44.42 + 39.00 P_{iD} + 63.38 K_{hD} - 16.84 K_{zD} - 1.69 V_{aD} - 41.60 H_{pD} \\
 & - 6.38 P_{iD}^2 - 33.39 K_{hD}^2 + 6.37 K_{zD}^2 - 2.90 V_{aD}^2 + 15.91 H_{pD}^2 - \\
 & 28.47 P_{iD} K_{hD} - 7.03 P_{iD} K_{zD} - 11.96 P_{iD} V_{aD} + 19.19 P_{iD} H_{pD} + 18.50 K_{hD} K_{zD} \\
 & + 2.47 K_{hD} V_{aD} + 4.04 K_{hD} H_{pD} - 18.60 K_{zD} V_{aD} + 12.62 K_{zD} H_{pD} + 12.59 V_{aD} H_{pD}
 \end{aligned}$$

R^2 is 0.9127 and R^2_{Adj} is 0.7962. R^2_{Adj} is lower because degrees of freedom available are smaller for the NOA (fewer runs), even though the number of coefficients is similar to the full factorial. The F statistic for this model is 7.8363 and $Pr(F)$ is 9.402e-05; the model is still highly significant.

4.2.1 F Test to Compare Two Variances

An F -test compares the two models predicted by factorial design and NOA. The number of degrees of freedom for the NOA is only 15 as 21 terms are evaluated by regression and only 36 runs were used. In the factorial design, 555 degrees of freedom are available. The results of the test are

Table 4.2 *F* Test Results for OA and Full Factorial Models

$F: 0.586$
$p\text{-value}: 0.092$
95 percent Confidence Interval: 0.243 1.089

An *F*-test compares the variances of two models to determine whether the models are drawn from the same population (both the models are statistically similar or not). The 95 percent confidence interval is the upper and lower critical values (rejection regions) of the *F* distribution with 15 and 555 degrees of freedom for 95% significance (Table 4.2). As the ratio of variances ($F = 0.586$) is within the interval ($0.243 < F < 1.089$), the null hypothesis that models are similar is accepted. This implies that orthogonal array with 36 runs is sufficient to analyze the system instead of a factorial design with 576 runs; the 576-run factorial model is not significantly different from the 36-run NOA model. Put more conservatively, the difference in variability captured by the much more expensive full factorial is difficult to justify statistically. This example demonstrates the enormous potential savings attainable using NOA designs.

4.3. COMPARING DIFFERENT SAMPLING TECHNIQUES

The uniformity characteristics of a sampling technique or a traditional experimental design are clearly important, as discussed above. In this section, Monte Carlo, Latin hypercube, and Hammersley sampling techniques are compared (Figure 4.2). All the methods sampled 100 points. These graphs show the space filling properties of these techniques. The 100 Hammersley points are spread evenly, almost on a lattice in a unit square. This uniformity on the lattice makes it an attractive method for high dimensional

sampling. The Latin hypercube (which is designed to perform well in single dimension) randomly combines with second variable to get a two-dimensional design. As a result, it is not as efficient in two dimensions and this can reasonably be extrapolated from a 2- to a k -dimensional space. Monte Carlo, which is unstratified Latin hypercube sampling, is more erratic and leaves more sampling space open than the other two and is the least favored sampling technique (although the easiest to interpret statistically, because of its randomness).

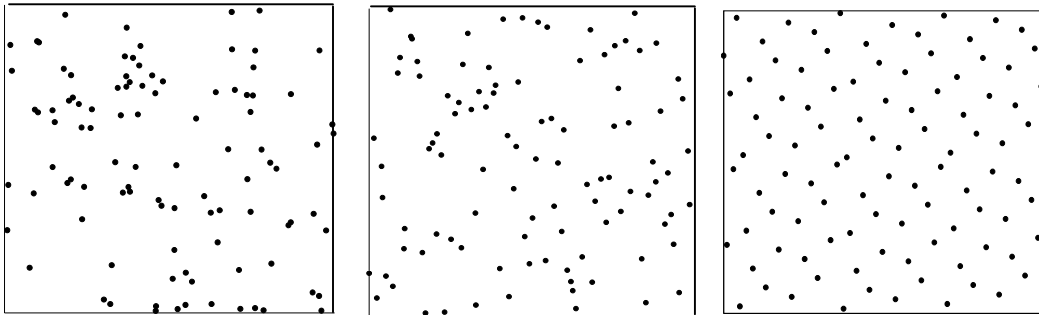


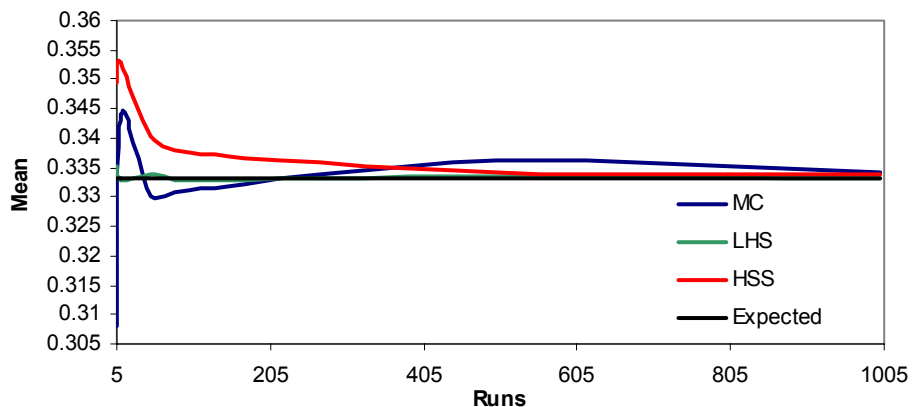
Figure 4.2 Sample Points (100) on a Unit Square for Monte Carlo, LHS and HSS

HSS and LHS reduce significantly the number of points required for representative sample statistics for high-dimensional problems such as reservoir modeling. That is, they reproduce the univariate characteristics of the distributions with fewer runs and also include any correlations (if they are any) between factors with fewer sample points.

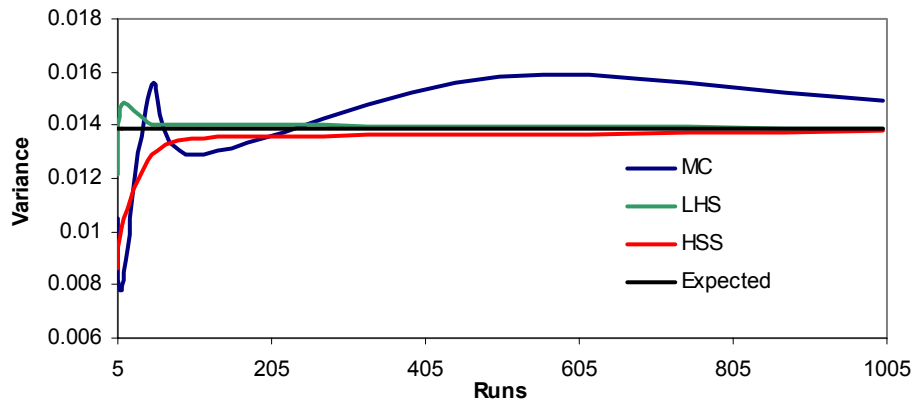
Numerical experiments compared sample means, standard deviations, and skewness obtained by a series of independent samples of Monte Carlo, LHS, and HSS, selecting 5, 10, 50, 100, 500, and 1000 runs (Figure 4.3). The input distribution is a beta distribution. HSS and LHS are compared with Monte Carlo for different distributions including uniform, triangular, normal, lognormal, and beta distributions. LHS and HSS

attain the required characteristics in fewer runs than Monte Carlo. Another important characteristic is that as the number of runs increases HSS and LHS converge to the actual value monotonically, but in Monte Carlo the estimated property oscillates around the expected value, making it hard to detect convergence. For the cases with only 30 or 50 sample points, the HSS and LHS more consistently generate means and standard deviations closer to the true mean and standard deviation compared with Monte Carlo sampling.

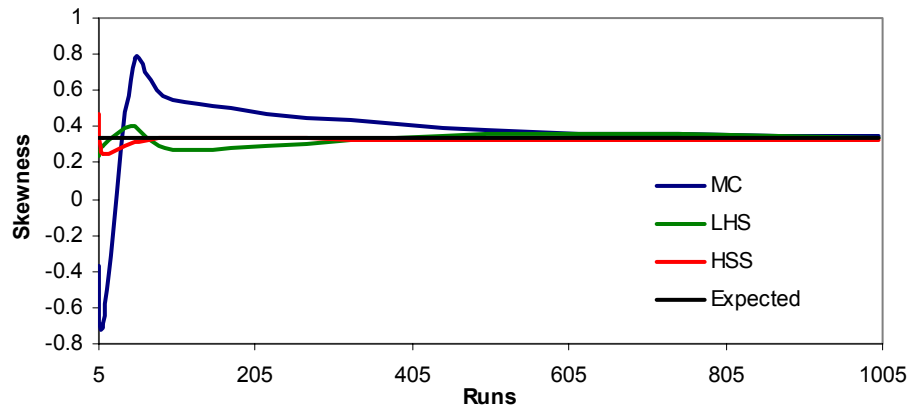
Latin hypercubes are very efficient in filling single dimensional space but Hammersley sequence uses discrepancy in placing n points in k -dimensional hypercube efficiently. Multi-dimensional properties are very important and the results of two dimensional space filling characteristics (Figure 4.2) signifies that Hammersley set of points fills the hyperspace unlike Monte Carlo and Latin hypercube. Kalganum and Diwekar (1997) provide a comparison of the performance of the HSS sampling to that of LHS and Monte Carlo. It was found that the HSS technique is at least 3–100 times faster than LHS and Monte Carlo techniques.



4.3a Convergence of mean for different techniques as number of samples increase



4.3b Convergence of variance for different techniques as number of samples increase



4.3c Convergence of skewness for different techniques as number of samples increase

Figure 4.3 Convergence Properties of Monte Carlo, LHS and HSS

CHAPTER 5. APPLICATION OF RESPONSE SURFACES

Response surface and experimental design methods are frequently used for sensitivity study of complex reservoir systems. In this study, these methods are applied to a single-well water-drive gas reservoir with a radial geometry. Factors that could influence the responses are considered and parameter ranges are selected depending on the uncertainty of each factor (Chapter 2). Fourteen factors are considered for this study and they are transformed using coding functions (Section 4.2). Cumulative gas recovery and breakthrough time are the responses examined. A quadratic regression model is compared with a kriged model.

Extensive investigation has been conducted on gas reservoirs under water influx using material balance calculations and simulation studies (McMullan and Bassiouni, 2000) to improve gas recovery by decreasing the potential gas trapping of advancing waterfront. McMullan *et al.* (2000) suggest that outrunning water influx by fast gas depletion can decrease gas trapping of advancing waterfront. Experimental and simulation studies (Valjak *et al.*, 2001) to use water drive energy to maintain gas rates with high pressures are being considered in low pressure gas reservoirs. Acaro *et al.* (1987) suggest that coproduction of the strong aquifers attached to gas reservoirs can minimize gas trapping.

Because high water cuts in gas production are economically unfavorable, operators may try to avoid high water production by partially completing the wells. However, partial completions result in productivity losses caused by Darcy and, at high gas flowrates, non-Darcy pressure drops (Lee *et al.*, 1987; Allam *et al.*, 1981). The

above-mentioned studies indicate a complex relationship between gas recovery, and aquifer strength, reservoir properties, and completion length. Therefore, it is important to understand the relationship for improved gas recoveries. This study analyzes water-drive gas reservoirs with water coning problems.

5.1. SIMULATION MODEL

Analytical methods for multiphase flow and water coning problems are generally limited and difficult to apply. Therefore, numerical reservoir simulators are preferred over analytical methods for modeling these reservoirs.

A radial model with water drive mechanism is used to study the performance of the orthogonal array design mentioned in chapter 4. The model consists of 26 radial grid rings and 110 layers with no angular variation (*i.e.*, an r - z *vis-à-vis* r - z - θ geometry).

5.1.1. Reservoir Geometry and Properties

In order to capture the uncertainty of reservoir geometric properties, reservoir dimensions (radius and thickness) have to be varied in reservoir simulation models. However, gas zone length, h , is kept constant at 100 feet for all simulation runs in this study (Figure 5.1). Other parameters such as porosity of gas zone, fluid densities, and reservoir temperature are also constant. The aquifer volume is varied by increasing the pore volume of the aquifer. Tubing performance is integrated with reservoir performance so that tubing head pressure and tubing diameter can be varied and optimized.

The model thickness is expanded and the expanded zone contains only water. A sketch of the gas-water system is illustrated in Figure 5.1.

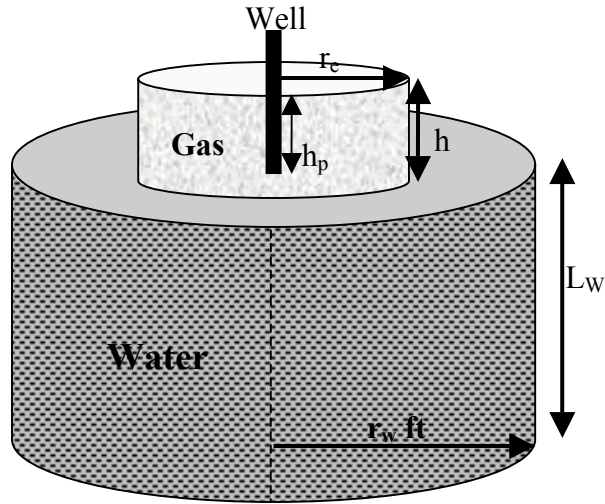


Figure 5.1 Sketch of the Radial Reservoir Model

The depth of the reservoir in the model is 5000 ft and was not varied. The ground surface temperature is 60°F with a temperature gradient of 1.2°F per 100 feet. The reservoir temperature is 120°F. The porosity of the gas zone is 25 percent. Other parameters like irreducible water, residual gas saturation, and vertical permeability to the horizontal permeability ratio were varied (Table 5.1).

The gas and water properties were estimated using correlations. The gas specific gravity is 0.65 with no CO₂, H₂S or N₂. The gas viscosity and gas deviation factor are from Armenta *et al.* (2003). The water specific gravity is 1.0 and the water viscosity is estimated using correlations programmed by McMullan (2000).

The gas-water relative permeability curves are for water-wet system and the capillary pressures are assumed to be negligible due to high contrasts in gas and water densities and the relatively high permeability considered in this study; high permeability implies large pore size and low capillary pressure. Gas deviation factor, gas viscosity, and rock properties are from Armenta *et al.* (2003).

Because a radial model is used for this study, only one central producing well is considered. Well production constrained by a constant tubing head pressure and terminated when rate reaches 1 MMscf/day. The tubing head pressure was related to the bottomhole pressure using Gray method (Eclipse Reference Manual, 2000).

5.1.2. Parameters under Consideration

Table 5.1 shows the fourteen factors, *viz.*, eleven geologic and three engineering variables. The engineering factors (tubing head pressure, tubing diameter, and completion length) are varied to optimize the production behavior. Most of the factors nonlinearly influence the response. Apart from three engineering factors, which are to be optimized to increase the reservoir performance, the other factors are uncertain factors. Factors chosen are in different areas of the reservoir simulation model: reservoir geometry, reservoir properties, well data, and relative permeability curves.

The factor ranges are selected based on practical factor distributions (Table 5.1). If the factors are more uncertain then a larger range is considered because the effect of those factors at extreme values should not be missed. A three-level and four-level mixed design is used to ensure adequate resolution of the influential factors (Section 2.1). Higher levels are chosen for the factors that are expected to affect the response most and are that are expected to have nonlinear effects. Factor values are scaled, coded, and spread depending on the factor statistical distribution. The list of factors chosen for water-drive gas reservoir analysis and the levels, symbols, and ranges are shown in Table 5.1.

Table 5.1. Factors considered in the Gas Reservoir Problem

Symbol	Parameters	Levels	Values of Levels			
P_i	Initial Pressure (psia)	3	2000	2500	3000	
K_h	Horizontal Permeability (md)	3	1	10	100	
S_{wc}	Connate Water Saturation	3	0.1	0.3	0.5	
S_{gc}	Critical Gas Saturation	3	0.1	0.2	0.3	
K_{rg}	Gas End Point	3	0.6	0.8	1.0	
K_{rw}	Water End Point	3	0.2	0.3	0.4	
N_w	Water Corey Exponent	3	2	3	4	
N_g	Gas Corey Exponent	3	1	2	3	
B	Non-Darcy Coefficient	4	0.1	1	10	100
V_a	Aquifer Size (ft)	4	100	600	1100	1600
K_z	Anisotropy Ratio	4	0.1	0.3	0.6	1
H_p	Completion Length Ratio	4	0.1	0.4	0.7	1
P_{tf}	Tubing Head Pressure $((p_i - p_{tf})/p_i)$	4	0.4	0.5	0.6	0.7
D_t	Tubing Diameter (in)	4	1.5	2.0	3.0	4.0

1. Initial Pressure

Uncertainty in the reservoir pressure is often encountered during the initial stages of exportation where data is very limited. Varying the initial reservoir pressure in the simulation model from 1500 psia to 3000 psia captures this uncertainty; alternatively, this could be viewed as a variable but near-certain property (that is, observable in the sense of White and Royer, 2003).

2. Permeability

A reservoir is composed of material formed by different sedimentation processes followed by chemical and biological diagenesis and structural adjustments. So, heterogeneity in the reservoir is common and may occur at scales from sand ripples (inch scale) to salt encroachment that traps the reservoir (mile scale).

In the current example, three different homogeneous permeability values quantify the uncertainty in reservoir permeability. In general, geostatistical realizations differing significantly can be taken as levels for an experimental design (Li and White, 2003). Geostatistical models (*e.g.*, conditional simulations) were not considered in this study.

3. Aquifer Strength

Aquifers are one of the least known elements of the reservoir simulation as few wells are intentionally drilled to investigate their properties. Material balance equations can be used to estimate the influence of aquifer size and productivity on production performance. However, uncertainty in aquifer models is propagated study by taking it as a parameter in the design. Four aquifer sizes are taken as four levels in the design.

4. Anisotropy Ratio

Anisotropy ratio is the ratio of vertical to horizontal permeability. Reservoir rocks can exhibit small anisotropy ratios because of heterogeneities. Anisotropy ranges from zero for laminated sands, like Lake Delta, to one for massive swamp deposits in the same delta sequence (Coleman, 1966). Four levels are considered for sensitivity analysis.

5. Non-Darcy Coefficient

Inertial and turbulent flow in the reservoir causes non-Darcy effects. These increase the pressure drop quadratically as the flow rate increases; in Darcy flow, the

pressure drop increases linearly with flow rate. Inertial effects (ρv^2 , Eqn. 5.1) influence on pressure drop in the porous media is given by the *Forchheimer* equation (Dake, 2001). In this equation, another proportional constant roughly analogous to μ/k , β , is incorporated to scale the increase in pressure drop. The final equation to consider the Non-Darcy pressure drop is:

$$\frac{\partial p}{\partial r} = \left(\frac{\mu}{k} \right) u + \beta \rho u^2 \dots\dots\dots(5.1)$$

Many empirical models have been proposed for modeling the β factor. Many of these models consider permeability, porosity, and tortuosity as the dominating terms controlling β .

Liquid saturation affects β (Armenta *et al.*, 2003) and most frequently used correlation for β is Frederick and Graves (1994) correlation.

$$\beta = \frac{7.89 \cdot 10^{10}}{k^{1.6} \phi^{0.404}}$$

The term β depends on many factors and considerable uncertainty is involved in its estimation. Thus, it is useful to know the significance of this term on production prediction. In this study, a standard correlation is used and then four log-distributed multipliers for β are used and effects are examined (four levels).

5. Relative Permeabilities

The shapes of the relative permeability curves are approximated by using Corey equations. However, relative permeability is a function of wettability, pore geometry, fluid distribution and saturation history (Craig *et al.*, 1981). Many of these factors are

hard to characterize -- especially wettability of a heterogeneous reservoir. Heterogeneous reservoirs often have large absolute permeability variations, which leads to changes in end point effective permeability causing changes in the relative permeability curves.

In practice, Corey exponents and end points are initially obtained from measured data or correlations (Craig, 1981). Here, reasonable values have been chosen, and a range is considered to examine sensitivities of different terms in the Corey equation on production performance.

$$K_{rwr} = K_{rw} \left(\frac{S_w - S_{wc}}{1 - S_{wc} - S_{gc}} \right)^{N_w}$$

$$K_{rg} = K_{rg} \left(\frac{S_g - S_{gr}}{1 - S_{gr} - S_{wr}} \right)^{N_g}$$

6. Completion Length

Decreasing the completion length may delay water coning in water-drive reservoirs. Partial completions over the reservoir thickness will always have higher pressure drops or lower rates than full completions. However, producing at high rates in reservoirs with high vertical permeability results in water coning. Therefore, completion length is a decision factor in asset development. Optimizing the reservoir performance using experimental design helps determine completion length.

7. Tubing Head Pressure and Tubing Diameter

The tubing head pressure (p_{th}) is related to the bottom hole pressure (p_{wf}) by Cullender-Smith method. Tables are generated by using commercial software (Eclipse

Reference Manual, 1997). Tubing diameter is varied for the table sets. Tubing diameter is related to the pressure drop by

$$\Delta p \propto \frac{q^2}{D^4}$$

The fourth root of the diameter should be proportional to the pressure and this relation is taken to determine the level values of the tubing diameter.

5.2. RESPONSE MODEL BUILDING

Complexity of the responses is expected to require a second order models.

$$y = \beta_0 + \sum_{i=1}^k \beta_i x_i + \sum_{i=1}^k \beta_{ii} x_i^2 + \sum_{i=1}^k \sum_{j>i} \beta_{ij} x_i x_j$$

This model includes one mean effect term, fourteen main effects for fourteen variables, fourteen quadratic terms for fourteen terms and $91(k(k-1)/2)$ interaction terms. As *NOA* design considered here has 36 runs, the degrees of freedom will allow us only to estimate 36 terms in the above equation. The main effects and quadratic terms are considered to generate the regression surfaces; two-term and higher interactions will be confounded with the estimated coefficients (section 3.2).

A full two-level factorial design for fourteen factors will have 2^{14} (16,384) runs. The design used here is a mixed design with three and four factor levels. A full factorial for such a design would be even more expensive with number of runs mounting to $3^8 \times 4^6$ that would be 26,873,856 runs. Even the fractional factorial should be fractionated hugely to decrease the number of runs, giving very complex confounding and probably undesirable design properties. Orthogonal arrays and nearly orthogonal arrays reduce the number of runs but confounding pattern also may be complex too. A 36 run *NOA* is chosen.

5.3. AUTOMATION

Once the design is selected, the reservoir models are built. Each row in the design array with k columns (k parameters) can be used as data points in the preparation of data deck for each simulation run (White and Royer, 2003). Any program that handles strings efficiently can be used to prepare the data decks. Programs create the data decks and included files corresponding to 36 design points. The coded variables in the design array are converted (using the scaling discussed above) to parameters with units which simulator reads. A commercial simulator is used (CMG Technologies Launcher, 2002). After all 36 data files are constructed; decks are run using batch files. Recovery factor and break through time are extracted as responses from the summary files.

5.4. FITTING RESPONSE SURFACE MODELS USING REGRESSION

Multiple linear regression fits response surface models for the fourteen factors and various responses. Normal score transformation is done on the responses for better fit. Main effects and quadratic effects are obtained (29 regression coefficients). The responses include gas recovery and break through time.

5.4.1. Gas Recovery

The derived second order model with out interaction terms for the recovery factor is

$$\begin{aligned} G_{pD} = & -0.53 + 0.21P_{iD} + 0.52K_{hD} - .34S_{wcD} - 0.02S_{gcD} + 0.08K_{rgD} + 0.06K_{rwD} + 0.06N_{wD} \\ & - 0.18N_{gD} + 0.04P_{fD} + 0.06D_{iD} + 0.08V_{aD} - 0.28K_{zD} - 0.22H_{pD} - 0.40B_D + 0.11P_{iD}^2 \\ & + 0.05K_{hD}^2 + 0.20S_{wcD}^2 - 0.02S_{gcD}^2 + 0.06K_{rgD}^2 + 0.01K_{rwD}^2 + 0.27N_{wD}^2 - 0.04N_{gD}^2 \\ & + 0.14P_{fD}^2 - 0.18D_{iD}^2 + 0.04V_{aD}^2 + 0.23K_{zD}^2 + 0.04H_{pD}^2 - 0.36B_D^2 \end{aligned}$$

where $G_{pD} = G_p / G$.

The coefficients of the factors determine the significance of that factor as the factors are scaled. The response surface gives the sensitivities of the factors. The steeper the surfaces, the more sensitive the corresponding factors to the response. Quadratic coefficients are smaller relative to the main effect coefficients. For gas recovery the most important factors are permeability, initial water saturation, and non-Darcy correction factors as shown in the response surface plot (Figure 5.2). The statistics for this model are:

Table 5.2 Regression Statistics for the Gas Recovery
Regression Statistics

R^2	0.9464
R^2_{adj}	0.7322
F	4.418 on 28 and 7 DF
p	0.0243

Table 5.3 Analysis of Variance for the Gas Recovery
Analysis of Variance

	Df	Sum Sq	Mean Sq	F value	Pr(>F)
X	28	32.893	1.175	4.4184	0.02436
Residuals	7	1.861	0.266		

R^2 is high for this model but as the degrees of freedom is small for the model (number of coefficients interpreted are high) R^2_{adj} is not very high. But the high F value and low p -value indicate the model is significant; most variance is explained by the

model. In the research, models are fit with low degrees of freedom; it is possible that R^2_{adj} would be small. But the mean square prediction errors and F statistics indicate a good fit.

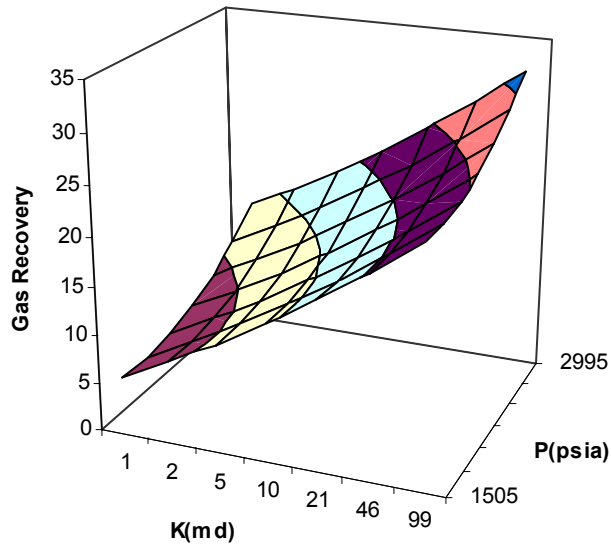


Figure 5.2 Gas Recovery Sensitivity to Pressure and Permeability

5.4.2. Water Breakthrough

Breakthrough (B_T) time is defined as the time (days) in the production history when the water production is above two stb/day. In a few of the simulations (34 out of the 36 total simulations considered) the breakthrough never occurs, so a higher rate of 100,000 stb/day is used and the normal score transform of the data is used to regress on the factors. Q_{max} was set constant for all simulations. These improve the regression statistics. The derived model is

$$\begin{aligned}
B_r = & -1.01 - 0.14P_{iD} - 0.47K_{hD} - 0.07S_{wcD} + 0.02S_{gcD} - 0.01K_{rgD} + 0.07K_{rwD} + 0.07N_{wD} \\
& - 0.001N_{gD} - 0.14P_{fD} - 0.10D_{iD} + 0.12V_{aD} - 0.25K_{zD} - 0.58H_{pD} + 0.06B_D + 0.16P_{iD}^2 \\
& + 0.27K_{hD}^2 + 0.23S_{wcD}^2 - 0.11S_{gcD}^2 + 0.06K_{rgD}^2 + 0.22K_{rwD}^2 - 0.02N_{wD}^2 - 0.04N_{gD}^2 \\
& - 0.28P_{fD}^2 + 0.12D_{iD}^2 + 0.05V_{aD}^2 + 0.06K_{zD}^2 - 0.04H_{pD}^2 - 0.13B_D^2
\end{aligned}$$

As could be expected permeability and completion length dominates the response. This also validates the procedure followed to predict and optimize the completion length and other responses.

Table 5.4 Regression Statistics for the Breakthrough Time
Regression Statistics

R ²	0.9527
R ² _{adj}	0.7636
F	5.038 on 28 and 7 DF
p	0.0168

Table 5.5 Analysis of Variance for the Breakthrough Time
Analysis of Variance

	DF	Sum Sq	Mean Sq	F	Pr(>F)
X	28	33.110	1.183	5.0376	0.01680
Residuals	7	1.643	0.235		

5.5. KRIGING SURFACES

There are many types of kriging but interpolation is done in this research by using simple kriging. The basics are already explained in chapter 3. Let $\mathbf{x} = (\mathbf{x}^1 \dots \mathbf{x}^{36})^T$ denote the vector of inputs and y denote a response (say, gas recovery). Then it is known that

$$y(\mathbf{x}) = f(\mathbf{x}) + z(\mathbf{x})$$

where $f(\mathbf{x})$ is assumed to be a constant and $z(\mathbf{x})$ is a random function with mean zero, variance σ_z^2 and correlation $R(\mathbf{x}^i, \mathbf{x}^j)$ between two z values at input vector \mathbf{x}^i and \mathbf{x}^j (Sacks *et al.*, 1989). Most important in this model is to build the correlation function $R(\mathbf{x}^i, \mathbf{x}^j)$.

$$R(\mathbf{x}^i, \mathbf{x}^j) = \exp \left[- \sum_{k=1}^{36} \theta_k (x_k^i - x_k^j)^2 \right]$$

where θ is the uncorrelated parameters used to fit the model, and the x_k^i and x_k^j are the k^{th} components of sample points \mathbf{x}^i and \mathbf{x}^j as explained in chapter 2. This assumes that the stochastic process has a Gaussian correlation model and the system is isotropic which means all θ_k are identical. The value of θ^* (isotropic) is estimated using a maximum likelihood method where the difference between the predicted and estimated $R(\mathbf{x}^i, \mathbf{x}^j)$ is minimized for all i and j . For a semivariogram $\gamma(\mathbf{h})$ where $\mathbf{h} = \mathbf{x}^i - \mathbf{x}^j$ is a k -dimensional distance separation vector between the two points \mathbf{x}^i and \mathbf{x}^j is

$$\gamma(|\mathbf{h}|) = c \left\{ 1 - \exp \left[- (\boldsymbol{\theta} \cdot \mathbf{h})^2 \right] \right\}$$

and $\gamma(|\mathbf{h}|)$ is also defined as half of the average square distance between two response values approximately separated by a vector \mathbf{h} . For an isotropic variogram,

$$\gamma(|\mathbf{h}|) = c \left\{ 1 - \exp \left[- \left(\frac{|\mathbf{h}|}{a} \right)^2 \right] \right\}$$

where a is the range which corresponds to the inverse of the norm of the parameter vector $\boldsymbol{\theta}$. The experimental semivariogram, computed from simulation results (in this study) is:

$$\gamma(|\mathbf{h}|) = \frac{1}{2N(\mathbf{h})} \sum_{i=1}^{N(\mathbf{h})} \{y(\mathbf{x}^i) - y(\mathbf{x}^j)\}^2$$

where $N(\mathbf{h})$ is the number of data pairs. Let S be the average of $\{y(\mathbf{x}^i) - y(\mathbf{x}^j)\}^2$, then the plot between estimated values of S^2 and $|\mathbf{h}|$ is as shown below in Figure 5.3

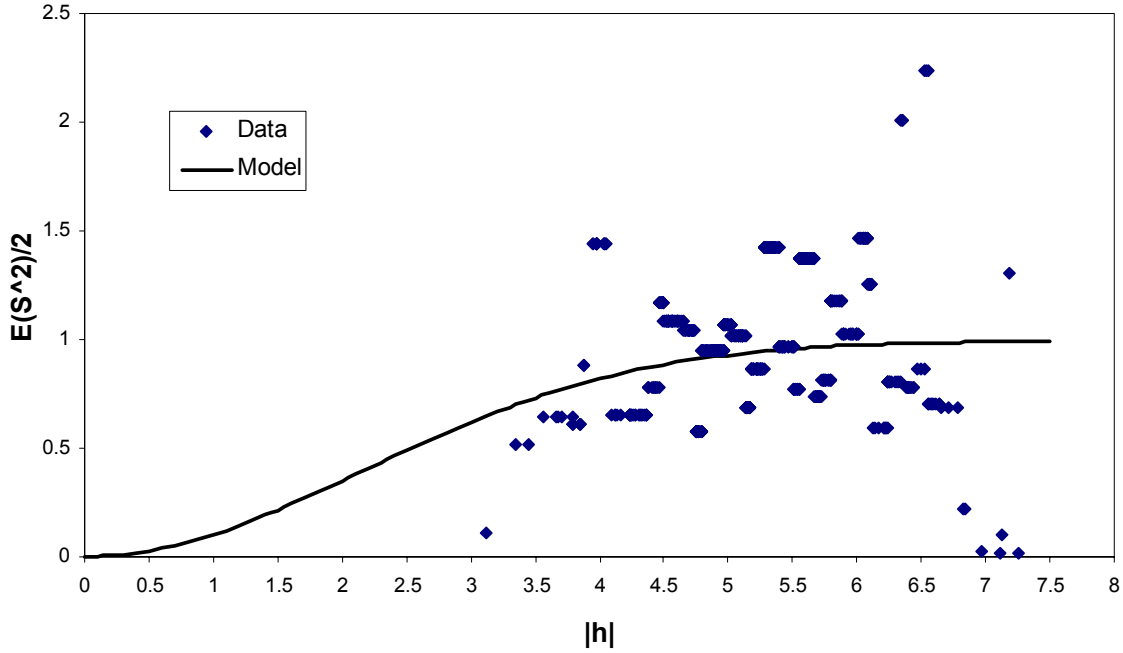


Figure 5.3 Estimation of Correlation Function Parameters

Minimizing the sum of errors by using maximum likelihood would give the value of kriging parameters θ . In the current work, because of the small number of runs (36) and the high dimensionality of the problem (14), a simple isotropic model ($\theta_i = \theta^*, \forall i$).

From the fitted model, a best linear unbiased predictor can be constructed by

$$\hat{y} = \hat{\beta}_0 + \mathbf{r}^T(\mathbf{x})\mathbf{R}^{-1}(\mathbf{y} - \hat{\beta}_0 \mathbf{1})$$

where $\mathbf{r}(\mathbf{x})$ is an $n \times 1$ vector of correlations with unknown point i and the design points given by $R(\mathbf{x}^i, \mathbf{x})$. \mathbf{R} is the correlation matrix of the design points and $\hat{\beta}_0$ is the least square estimate of β_0 . Prediction and visualization are done with the DACE (2002) software in MATLAB. Figure 5.4 is a surface generated by the kriging model. The variables are permeability and the connate water saturation and the response is the gas recovery. It is showing that as connate water saturation increases the recovery will be low. The change in concavity along the S_w axis shows that the kriging method can

produce surface shapes that quadratic models cannot; in addition, the estimation error at all sample data is identically zero.

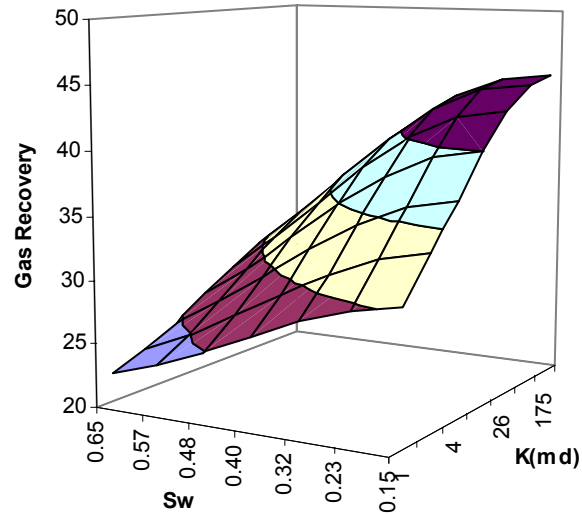


Figure 5.4 Gas Recovery Sensitivity to Pressure and Permeability

5.6. COMPARISON BETWEEN POLYNOMIAL REGRESSION AND KRIGING

Validating of the models built by regression and kriging is done using bootstrapping and jackknifing, respectively. Both this statistical methods estimate the standard error for a simple mean.

5.6.1. Bootstrapping the Polynomial Model

The notion of bootstrapping depends on the bootstrap sample (Efron and Tibshirani, 1993). A bootstrap sample $\mathbf{x}^* = (x_1^*, x_2^*, x_3^*, \dots, x_n^*)$ is obtained by randomly sampling with replacement the actual data set of $x_1, x_2, x_3, \dots, x_n$ N times. N should be large enough to stabilize the statistic in concern. In the current studies, $x_1, x_2, x_3, \dots, x_n$ are the 36 responses from the simulator (i.e, n is equal to 36). Thus the bootstrap sample $(x_1^*, x_2^*, x_3^*, \dots, x_n^*)$ consists of members of the original data set, some elements appearing more than once and other not sampled. For each bootstrap sample \mathbf{x}_i^* there

exists the sample mean let it be $S(\mathbf{x}_i^*)$ which could be evaluated for all the bootstrap samples. Then the bootstrap estimate of standard error is

$$\hat{SE} = \left\{ \sum_{i=1}^N \frac{[S(x_i^*) - \alpha]^2}{(N-1)} \right\}^{\frac{1}{2}}$$

Where $\alpha = \sum_{i=1}^N S(x_i^*)/N$ and the above estimate is standard error. That is because, for a

set of data points $x_1, x_2, x_3, \dots, x_n$ and $\bar{x} = \sum_{i=1}^n x_i/n$ the standard error is $(s^2/n)^{\frac{1}{2}}$ and it is

close to \hat{SE}

where
$$s^2 = \sum_{i=1}^n \frac{(x_i - \bar{x})^2}{(n-1)}$$

The bootstrap method is applied to the residuals by taking 10,000 bootstrap samples. The set of $x_1, x_2, x_3, \dots, x_n$ are the 36 residuals of the regression. The standard error is **0.038**.

5.6.2. Jackknifing the Kriging Model

Jackknifing is an approximation method to bootstrapping for predicting the standard error. Jackknifing is done by removing a sample point or design point from the data set of $\mathbf{x}_1, \mathbf{x}_2, \mathbf{x}_3, \dots, \mathbf{x}_n$ to get the cross validation sample. The responses for those data points are $y_1, y_2, y_3, \dots, y_n$; from the sample, predict the kriging value for the removed design point. For the j^{th} data point removal

$$\tilde{y}_j = y_j - \hat{y}_j$$

where y_j is the original j^{th} response with out removing any data points and \hat{y}_j is the kriging estimate after removing that data point from the data set. The estimate of standard error is

$$\hat{SE} = \left(\sum_{j=1}^n \left[\tilde{y}_j - \overline{\tilde{y}} \right]^2 / n (n - 1) \right)^{\frac{1}{2}}$$

where $\overline{\tilde{y}} = \sum_{j=1}^n \frac{\tilde{y}_j}{n}$. Jackknifing the kriging model of the gas recovery estimates

the standard error **0.14**.

5.6.3. Comparison of Errors

The standard errors show that predictions using regression (0.038) are much smaller than kriging (0.14). This implies that regression for this model is better for sensitivity analysis and optimization. These results are obtained partly because the number of degrees of freedom for kriging is smaller, especially with the assumption of isotropy in the fourteen dimensions. Using more degrees of freedom in the kriging model (for example, by computing an anisotropic parameter vector, θ , for the correlation function rather than using a single isotropic value θ_i) can improve kriging accuracy (Sacks *et al.*, 1989; Chen and Simpson, 2000; Kleijnen and Beers, 2004)

CHAPTER 6. RESERVOIR APPLICATIONS

Applications for the response surfaces derived in the gas coning simulation study are discussed in this chapter. Response surfaces can approximate production performance more quickly and easily than numerical reservoir simulators. Response surfaces can be used instead of more expensive numerical models and are applied in this chapter for

1. optimization of engineering factors
2. expected production of reserves under uncertainty
3. estimating sensitivity

Effectiveness of the response surfaces generated by a design determines the robustness of the risk analysis and optimum solution. So, the 36 run NOA (Chapter 5) is compared with different NOA designs with more runs.

6.1. UNCERTAINTY ANALYSIS

Factor uncertainty influences the reliability of the response functions and the optimization process in the reservoir simulation (Chapters 2, 3). Uncertainty analysis estimates the mean response and standard deviation. As discussed in chapter 5, uncertainty is caused by (Narayanan, 1999)

1. measurement errors in factors like permeability and relative permeability
2. upscaling errors in estimating of effective properties
3. estimated reservoir properties at biased and very few reservoir locations
4. approximations and simplifications in the model building

6.1.1 Robustness of Uncertainty Estimate

Traditionally uncertainty assessment in a model is done using Monte Carlo simulation. This analysis translates the specified uncertainty in factors to uncertainty in

responses by randomly varying factors and computing responses (Chapter 2). Monte Carlo analysis with fine scale flow simulation would be prohibitively expensive, as the number of Monte Carlo samples required is around 10^4 to 10^5 for a moderately complex problem. The use of an efficient Hammersley sequence and Latin hypercubes for such fine scale simulations is discussed in chapter 3.

A Hammersley sequence of 1000 factor combinations (or realizations) is considered. Responses are computed using polynomial response models, which are trivial in expense (calculate in a desktop spreadsheet). The simulations use response surfaces from NOA designs with 36, 72, and 108 runs. The 108-run NOA design is generated by adding the 36-run and 72-run NOA designs because orthogonal arrays are adaptive designs (adding two carefully selected designs generates a third design with good statistical properties). Response surfaces using these different runs are compared (section 6.3). The possible values of the uncertain factors are specified as probability distributions. The factor distributions may be estimated using historical and analog data and knowledge about the physics controlling the factor (e.g., depositional and diagenetic processes controlling permeability). Inferred distributions may be uniform, normal, lognormal, and triangular. To evaluate responses in a spreadsheet, the relations between factors and responses are specified as polynomials computed from linear regression. The stochastic draws many factor combinations using Hammersley sequences; those combinations are used as an input to the polynomial response models.

These simulations reveal the range of possible outcomes for the response (here, gas recovery) and the probability of occurrence. In this study, the mean and standard

deviation for NOA designs with 36, 72, and 108 runs are estimated (Table 6.1; Figure 6.1).

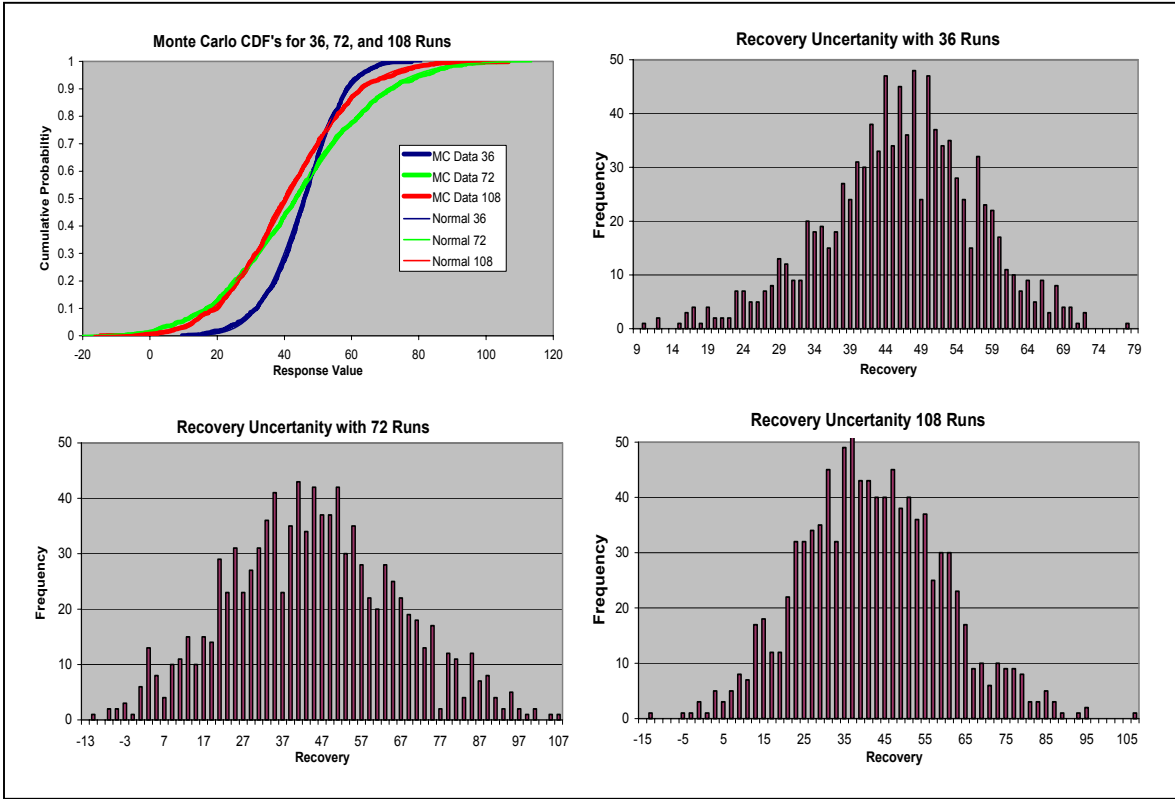


Figure 6.1 Monte Carlo Simulation Distributions

6.1.2 Response Surfaces for Uncertainty Estimation

Uncertainty analysis uses response surfaces generated from an NOA(72, 14, $3^6 4^8$, 2) with 72 runs and an NOA(108, 14, $3^6 4^8$, 2) with 108 runs. These models are compared with analysis done by response surfaces of an NOA(36, 14, $3^6 4^8$, 2) with 36 runs (Chapter 5). Response surfaces from the 72-run NOA contain main and quadratic effects as in 36-run model and also include interactions of engineering factors (tubing head pressure, tubing diameter, and completion length). Response surfaces generated by the NOA with 108 runs contain the effects of 72-run design plus the interactions of dominating factors

as determined from earlier, smaller designs (aquifer size, connate water saturation, and water relative permeability). Table 6.1 shows that response models from all designs predict similar mean gas recovery. The 36-run design variance estimate is less than the 108-run variance estimate, but the mean gas recovery estimates from all designs are statistically indistinguishable at the 0.05 confidence level. However, the difference in standard deviation indicates that extrema are estimated much differently, particularly between the 36-point and more densely sampled designs.

Table 6.1 Monte Carlo Simulation Statistics

Design	Mean	Standard Deviation
NOA(36, 14, 3648, 2)	45.52	10.74
NOA(72, 14, 3648, 2)	43.76	21.23
NOA(108, 14, 3648, 2)	40.96	17.50

6.2. OPTIMIZATION OF ENGINEERING FACTORS

Response surfaces can be used with conventional optimization techniques. Although the factors are nonlinearly influencing the response, polynomial function (response surfaces) optimization is fast and accurate. Optimization of engineering factors is done with realistic factor uncertainty to improve the gas recovery. The objective function is the response surface for gas recovery obtained by regressing 36-run NOA. The optimal solution precision depends on the accuracy of the response surface, and the response surface accuracy depends on the choice of design and design parameters like levels and number of runs. In this section the choice of 36-run in generating response

surfaces and doing optimization is justified by comparing with the optima obtained by 72 and 108-run NOA designs.

6.2.1 Formulation of the Objective Function

The optimization of gas recovery under uncertainty maximizes the expected value of this response by changing the engineering factors (White and Royer, 2003).

$$\bar{f}(\mathbf{c}) = \int_{\Omega} f(\mathbf{c}, \boldsymbol{\omega}) d\Pi(\boldsymbol{\omega}) \dots\dots\dots (6.1)$$

where $\bar{f}(\mathbf{c})$ is the expected value of the gas recovery, which is a function of only engineering factors \mathbf{c} , obtained by integrating over the uncertain factors $\boldsymbol{\omega}$. The vectors $\boldsymbol{\omega}$ and \mathbf{c} are subvectors of \mathbf{x} . $d\Pi(\boldsymbol{\omega})$ is the joint probability function of all the uncertain factors. This formulation (Eqn. 6.1) recognizes that reservoir properties such as k_v / k_h are rarely precise; by averaging over all uncertain factors, it incorporates uncertainty into the objective function and includes nonlinearity in all factor-response relationships.

6.2.2 Optimization Method and Results

If Equation 6.1 is solved analytically, $\bar{f}(\mathbf{c})$ is a second order equation without any interactions as no interaction terms are assessed in the 36 run NOA design. The uncertain factors influence the objective function \bar{f} if the uncertain factors are interacting with the engineering factors. When integrated over uncertain factors, the coefficients of the main effects influence the optimization process (caused by interacting factors) but average mean effect change (caused by non-interacting factors) cannot influence the optimum values. As the quadratic effects are estimated in 36-run design, \bar{f} for this response surface would be a parabola in space \mathbf{c} , and may have an optimum within the

factor space (not possible for a linear model). In this study, \bar{f} is obtained by numerically integrating f over the uncertain factor space using an HSS sample and optimization of engineering factors is done using a nonlinear regression method (White and Royer, 2003; Excel, 2003). Response surfaces from 72-run and 108-run NOA having interaction terms are used similarly to do the optimization (Table 6.2).

Table 6.2 Gas Recovery Optimum Values

Design	$h_{pD}=(h/h_t)$	p_{tf} (psia)	D_t (in)
NOA(36, 14, 3648, 2)	1.0	$0.3p_i$	2.1
NOA(72, 14, 3648, 2)	1.0	$0.38p_i$	1.5
NOA(108, 14, 3648, 2)	1.0	$0.3p_i$	2.0

The optima of 36-run and 108-run are close, but the 72-run deviates from both the models. Although the 72-run design selects smaller tubing and higher tubing pressure, all designs have optima around the full completion length, average tubing diameter around 2 in, and low tubing head pressure. The results should be further analyzed by using another design with engineering factors and few important uncertainty factors (aquifer size, connate water saturation, and water relative permeability) in the optimal region. The high initial ranges for the factors chosen for this study and number of factors make it nearly impossible for designs to precisely locate optima.

6.2.3 Discussion of Optimization

These results show that response surfaces generated by the 36-run NOA is an efficient design for optimization, as optimum estimated by this design are very close to

the results for the NOA with 108 runs. Tubing head pressure is scaled to the initial reservoir pressure, so the optimum value is a function of p_i . The change in the objective function with the three engineering factors for this design is shown in Figure 6.2. Completion length and tubing head pressure are going to the boundaries (Figure 6.2). These results show that a gas well with coning and uncertain reservoir properties should have completions through the pay zone and the tubing head pressure kept as low as possible to produce as much gas as possible to maximize gas recovery.

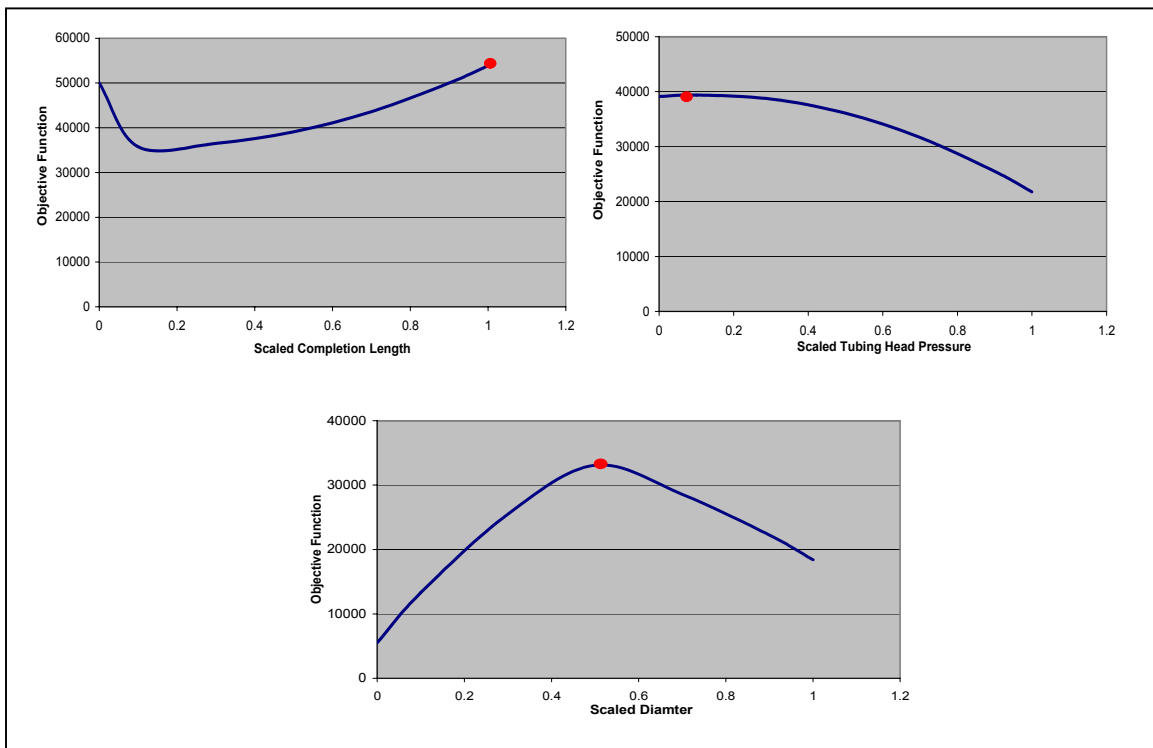


Figure 6.2 Objective Function Sensitivity to the Engineering Factors

The limit for the head pressure could be facilities restriction and sand production. As these are not inputs to the simulation, the limits of the ranges should be chosen carefully to model those effects. Higher tubing diameter is optimized to a diameter where recovery will be high. Higher diameters give loading problems especially under water

encroachment and lower diameters may be giving lower recovery. Further optimization analysis would include the three major uncertain factors, tubing diameter ranging from 1.5 to 2.5 inch, and head pressure with high $((p_i - p_g) / p_i)$. The results are different if water handling, compression, and tubing costs are included in the objective function (Section 6.3).

Factor distributions also influence the optimization and uncertainty analysis significantly. Reasonable distribution models are essential for meaningful results. For example, an alternative uncertainty evaluation uses all factor distributions shifted to left (using triangular distributions). Using response surfaces generated by the 108- run NOA and resampling from these factor distributions, the uncertainty and optimization results could change. Average recovery increases from 41 percent to 52 percent and standard deviation increases from 17.7 to 26.9. The optimal tubing diameter changes from 2 inch to 2.3 in (other two controllable factors had the same optima as previously).

Such analyses of factor distribution effects can be used to justify the resources for additional data. For example, in current research aquifer properties are influencing the uncertainty and optimization significantly, so more resources could be shifted for acquiring that data.

6.3. STATISTICAL COMPARISON OF DESIGNS

Uncertainty and optimization analyses (sections 6.1 and 6.2) show that a 36-run NOA provides efficient uncertainty analysis and optimization. Similar results are obtained using 36 runs compared with larger designs of 72 and 108 runs. In this section statistical analysis of the residuals from the regression models uses F - and t -tests. NOA designs with 108, 36, and 72 runs are compared. The factors considered in regression for

all the factors are similar. The F -test compares the model residuals and the t -test compares the means of the residuals generated by different response surfaces. The F -test (Table 6.3) shows that 36-run NOA residuals are different than the 108-run NOA. As the number of residuals and expected variances are different for the two models, variances are pooled (Montgomery, 2001) and the t -test shows that the means are obtained from the same distribution.

Table 6.3 Comparison of NOA(36, 14, 3⁶4⁸, 2) and NOA(108, 14, 3⁶4⁸, 2)
T-Test

Method	Variances	DF	t Value	Pr > t
Pooled	Equal	142	0.00	1.0000
Satterthwaite	Unequal	125	0.00	1.0000

F-Test				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	107	35	4.38	<.0001

Similar tests compare the residuals generated by regression for 72-run and 108-run designs. The results show that both the variances and means are obtained from the same sample population (Table 6.4). Although the models are indistinguishable according to these univariate tests, the differences in optimization results indicate that the models are different for some applications. The differences in optimization results could be caused by subtle differences in how the residuals are distributed in the factor space; such differences are not assessed by t - or F -tests.

Table 6.4 Comparison of NOA(72, 14, 3⁶4⁸, 2) and NOA(108, 14, 3⁶4⁸, 2)
T-Test

Method	Variances	DF	t Value	Pr > t
Pooled	Equal	178	0.00	1.0000
Satterthwaite	Unequal	168	0.00	1.0000

F-Test				
Method	Num DF	Den DF	F Value	Pr > F
Folded F	107	71	1.40	.128

6.3. ECONOMIC ANALYSIS

In gas and oil development economic factors like discount factor and gas price have a significant influence on the viability of the project. These factors determine the rate of return on investment and also provide reserves depletion strategy. Economic factors should be considered which optimizing the production parameters and so more important than cumulative gas recovery is the net present value (NPV), which should be optimized. The cost of tubing, water disposal cost (C_w), gas price (C_g) and gas pumping costs could change the optimum operating conditions.

Considering water disposal costs may penalize the option of full completion of reservoir. Full completion would increase the cumulative gas recovery but under increased water production, which could be a significant cost, may decrease NPV. While optimizing the NPV the completion length is chosen where gas recover is high (C_g is a factor and discount rate) and also cost for water treatment is low (C_w is a factor and discount rate). Similarly, tubing cost depends on both length and diameter of the tubing could change the optimum in Table 6.2. Tubing cost is not influenced by discount rate as

it is an initial cost of investment. Other factors influencing the results could be the pumping costs, compression costs, and sand control investment.

In this section NPV optimization is evaluated with factors like water treatment costs and discount ratio under multifactor uncertainty. The objective function is the response surface for NPV obtained by regressing 36-run NOA. Water treatment costs and annual discount rate are assumed to be 20 cents/bbl and 10 percent, respectively. Gas price is taken as \$4/MCF. Inflation in gas and water treatment is assumed to be three percent. Tubing cost is included, but considering tubing costs alone for the prescribed depth does not significantly influence the decisions as the costs are not large (less than 0.1%) compared with net revenue. If the initial investment required for drilling and completion for the destined depth were included, then tubing diameter might make significantly influence NPV. Similarly, compression costs may change the optimum p_{tf} estimates but this factor is not considered.

Table 6.5 Net Present Value Optima

Design	$h_p(h/h_t)$	p_{tf} (psia)	D_t (in)
NOA(36, 14, 3648, 2)	0.57	$0.3p_i$	2.3

Engineering optima values for NPV (Table 6.5) differ from results optimizing gas recovery only (Table 6.2). Optimum tubing head pressure is not changed and tubing diameter changed slightly.

CHAPTER 7. CONCLUSIONS AND RECOMMENDATIONS

Experimental designs help examine reservoir behavior if it is affected by many factors over wide ranges. The ranges of the factors, levels of the factors, and number of runs are chosen depending on model complexity and nature of dependencies between factors and responses.

Uncertainty can be assessed efficiently using experimental designs. Orthogonal arrays and nearly orthogonal arrays are more efficient than full factorials, Latin hypercubes, and other partial factorial designs. They significantly decrease the number of runs compared to traditional designs like central composite and Box Behnken designs. Uncertainty analysis using Monte Carlo sampling can be improved and better variance and mean predictions of the responses can be obtained with less number of runs by using Hammersley sampling. Modified Monte Carlo, Latin hypercube sampling can be efficient with fewer factors but Hammersley is better in high-dimensional problems.

Response surfaces analyze sensitivity of production responses to factor variations. Polynomial models or high-dimensional kriging can be used to create response models. Bootstrapping and jackknifing provide error estimates for these models. In the current study, polynomial models had lower errors than kriged response models. Kriged estimates may be improved by not assuming anisotropy in the factor space or kriging residuals to a low order polynomial.

Response surfaces for the gas coning simulation model with a 36-run NOA show that the response surfaces can relate the responses like NPV or gas recovery to the production parameters accurately. Optimization, sensitivity, and uncertainty studies can be done more quickly and easily with proxy response models than with numerical

reservoir simulators. In some applications, response surfaces can be used instead of more expensive numerical models.

REFERENCES

- Acaro, D.P., Bassiouni, Z. A.: "The Technical and Economic Feasibility of Enhanced Gas Recovery in the Eugene Island Field by Use of the Co-production Technique", JPT 585-590 (1987)
- Allam, M., Crichlow, H. B., Soliman, Y.: "Numerical Simulation of Heterogeneous Fractured Gas Reservoir Systems with Turbulence and Closure Stress Effects," SPE 10119 (1981)
- Armenta, M., White. C. D., Wojtanowicz. A. K.: "Completion Length Optimization in Gas Wells", Canadian International Petroleum Conference 192 (2003)
- Aslett, R., Buck, R, J., Duvall, S, G.: "Circuit Optimization via Sequential Computer Experiments: Design of an Output Buffer", Applied Statistics (1998), 47, 31-48
- Box, G. E. P., Behnken D. W.: "Some New Three Level Designs for the Study of Quantitative Variables," Technometrics (1960), 2, 455-476
- Box, G. E. P., Draper, N. R.: "Empirical Model Building and Response Surfaces", New York: Wiley (1987)
- Box, G.E.P and Hunter, J.S.: "The 2(k-p) Fractional Factorial Designs", part I, Technometrics (1961), 3, 311-351
- Box, G. E. P., Hunter, W. G., Hunter, J. S.: "Statistics for Experimenters: An Introduction to Design, Data Analysis, and Model Building" John Wiley & Sons (1978)
- CMG Technologies Launcher, Version 2002.1, Revision 4, Computer Modeling Group, 1978-2002
- Coleman, J. M.: "Ecological Changes in a Massive Freshwater Clay Sequence", Gulf Coast Association Geological Society Transactions (1966), 14, 159-174
- Craig, F.F.: "The Reservoir Engineering Aspects of Water-Flooding", Dallas: AIME, 1981
- Cullick, A. S., Heath, D., Narayanan, K.: "Optimizing Multiple-Field Scheduling and Production Strategy with Reducing Risk", SPE 84239 (2003)
- Dake, L. P.: "Fundamentals of Reservoir Engineering" Elsevier Science (2001)
- Deutsch, C. V., Journel, A. G.: "GSLIB: Geostatistical Software Library and User's Guide", New York: Oxford (1998)

Efron, B., Tibshirani, R. J., "An Introduction to the Bootstrap", New York: Chapman & Hall (1993)

Frederick, D.C., Graves, R.M.: "New Correlations to Predict Non-Darcy Flow Coefficients at Immobile and Mobile Water Saturation", SPE 28451 (1994)

Giunta, A. A., Watson, L. T.: "A Comparison of Approximation Modeling Techniques: Polynomial versus Interpolating Models", AIAA (1998), 4758

Giunta, A. A., Wojtkiewicz, S. F., Eldred, M.S.: "Overview of Modern Design of Experiments Methods for Computer Simulations" AIAA (2003), 0649

Goovaerts, P.: "Geostatistics for Natural Resources Evaluation", New York: Oxford (1997)

Gregory D. W., Kelly H. J.: "A User's Guide to LHS: Sandia's Latin Hypercube Sampling Software"

Halton, J.H.: "On the Efficiency of Certain Quasi-Random Sequences of Points in Evaluating Multi-Dimensional Integrals", Numerische Mathematik (1960), 2, 84–90

Hedayat, A. S., Pesotan, H.: "Two-Level Factorial Designs for Main Effects and Selected Two-Factor Interactions", Statistical Sinica (1992), 453-464

Hedayat, A. S., Sloane, N.J.A., John Stufken.: "Orthogonal Arrays: Theory and Applications", Springer Series, New York (1999)

Iman, R. L., Conover, W. J.: "A Distribution- Free Approach to Inducing Rank Correlation among Input Variables", Communications in Statistics (1982), B11 (3), 311-334

Informatics and Mathematical Modeling, 2002, Design and Analysis of Computer Experiments, Version 2.0

Jin. R., Chen. W., Simpson, W. T.: "Comparative Studies of Metamodeling Techniques under Multiple Modeling Criteria", AIAA (2000), 4801

Kalagnanam, J. R., Diwekar, U. M.: "An Efficient Sampling Technique for Off-Line Quality Control", Technometrics (1997), 39, 308 – 319

Kleijnen, J. P. C., Van Beers, W. C. M.: "Kriging Interpolation in Simulation: A Survey", Winter Simulation Conference (2004)

Kohler. J. R., Owen. A.B.: "Computer Experiments", Handbook of Statistics, Elsevier Science, New York (1996)

- Kurniwaran, B.: "Shaly Sand Interpretation Using CEC-Dependent Petrophysical Parameters", PhD Dissertation, Louisiana State University, 2005
- Landa. J.L., Guyaguler. B.: "A Methodology for History Matching and the Assessment of Uncertainties Associated with Flow Prediction", SPE 84465(2003)
- Lee, R. L., Logan, R. W., Tek, M. R.: "The Effect of Turbulence on Flow of Natural Gas through Porous Media," SPE 14295 (1987)
- Li. H., White, C. D.: "Geostatistical Models for Shales in Distributary Channel Point Bars (Ferron Sandstone, Utah): From Ground-Penetrating radar data to Three-Dimensional Flow Modeling", AAPG (2003) 87, 12, 1851-1868
- McKay. M.D., Conover. W.J., Beckman. R.J.: "A Comparison of Three Methods for Selecting Values of Output Variables in the Analysis of Output from a Computer Code" Technometrics (1979), 21, 239-245
- McMullan, J. H., Bassiouni, Z.A.: "Optimization of Gas-Well Completion and Production Practices", SPE 58983 (2000)
- Microsoft Excel 2003, Microsoft Professional Office Edition 2003
- Montgomery, D.C.: "Design and Analysis of Experiments", 5th ed., Wiley, New York (2001)
- Montgomery, D. C.: "Introduction to Linear Regression Analysis", Wiley, New York (2001)
- Myers, R. H., Montgomery, D.C.: "Response Surface Methodology: Process and Product Optimization Using Designed Experiments", New York: Wiley (1995)
- Narayanan, K.: "Applications for Response Surfaces in Reservoir Engineering", MS Thesis, University of Texas, 1999
- Narayanan, K., Cullick, A. S., Bennett, M.: "Better Field Development Decisions from Multi-Scenario, Interdependent Reservoir, Well, and Facility Simulations", SPE 79703 (2003)
- "NIST/SEMATECH e-Handbook of Statistical Methods", <http://www.itl.nist.gov/div898/handbook/>, Jan 2005
- Nguyen, N. K.: "A Note on the Construction of Near-Orthogonal Arrays with Mixed Levels and Economic Run Size", Technometrics (1996), 38, 279-283
- Owen, A. B.: "Orthogonal Arrays for Computer Experiments, Integration and Visualization", Statistical Sinica (1992), 439-452

Peake, W.T., Abadah, M., Skander, L.: “Uncertainty Assessment using Experimental Design: Minagish Oolite Reservoir” SPE 91820 (2005)

Peng, C.Y., and Gupta, R.: “Experimental Design and Analysis Methods in Multiple Deterministic Modeling for Quantifying Hydrocarbon In-Place Probability Distribution Curve”, SPE 87002 (2004)

Sacks, J., Schiller, S.B., Welch, W.J.: “Designs for Computer Experiments”, Technometrics (1989) 31:41-47.

Sacks, J., Welch, W. J., Mitchell, T. J., Wynn, H. P.: “Design and Analysis of Computer Experiments”, Statistical Science (1989), 4(4), 409-435

Sandor, Z. and Andras, P.: “Alternative Sampling Methods for Estimating Multivariate Normal Probabilities”, Journal of Econometrics (2004), 120, 2, 207-234

Schlumberger Technology Co., 1997, Eclipse 100 Reference Manual, unpaginated

Simpson, W. T., Mauery, M. T., Korte, J.J., Mistree, F.: “Comparison of Response Surface and Kriging Models for Multidisciplinary Design Optimization”, ASME (1997)

Sloane, N. J.: “A Library of Orthogonal Arrays”, <http://www.research.att.com/~njas/oadir/>, Jan 2005

Sobol, I.M.: “On the Distribution of Points in a Cube and the Approximate Evaluation of Integrals”, Computational Mathematics and Mathematical Physics (1967), 7, 86–112.

Valjak, M., Novakovic, D., Bassiouni, Z. A.: “Physical and Economic Feasibility of Waterflooding of Low-Pressure Gas Reservoirs”, SPE 69651 (2001)

Wang, F., White, C. D.: “Designed Simulation for a Detailed 3D Turbidite Reservoir Model”, SPE 75515 (2002)

Wang, J. C., Wu, C. F. J.: “Nearly Orthogonal Arrays with Mixed Levels and Small Runs”, Technometrics (1992), 34, 409-422

Weisstein, Eric W.: “Quasirandom Sequence.” MathWorld -- A Wolfram Web Resource, <http://mathworld.wolfram.com/QuasirandomSequence.html>

White, C.D., Royer, S.A.: “Experimental Design as a Framework for Reservoir Studies” SPE 79676 (2003)

White, C. D., Wills, B. J., Narayanan, K., Dutton, S. P.: “Identifying Controls on Reservoir Behaviour using Designed Simulations”, SPE 62971 (2000)

Xu, H.: “An Algorithm for Constructing Orthogonal and Nearly-Orthogonal Arrays with Mixed Levels and Small Runs”, *Technometrics* (2002), 44, 356-368

VITA

Subhash Kalla was born in Madhavaram, India, on August 12, 1981, the son of Gangadhara R. Kalla and Laxmi V. Kalla. After completing his school studies at Aditya Junior College, Kakinada, India, he entered Regional Engineering College, and obtained a bachelor's degree with distinction in chemical engineering in 2002. In August 2002, he joined the Graduate School of Louisiana State University, Baton Rouge, United States, in The Craft and Hawkins Department of Petroleum Engineering. The degree of Master of Science in Petroleum Engineering will be conferred in August 2005.