THE UNIVERSITY OF TULSA

THE GRADUATE SCHOOL

STOCHASTIC CONDITIONAL SIMULATION FOR DESCRIPTION OF RESERVOIR PROPERTIES

by

Godofredo Pérez

A dissertation submitted in partial fulfillment of
the requirements for the degree of Doctor of Philosophy
in the Discipline of Petroleum Engineering
The Graduate School
The University of Tulsa
1991

THE UNIVERSITY OF TULSA

THE GRADUATE SCHOOL

STOCHASTIC CONDITIONAL SIMULATION FOR
DESCRIPTION OF RESERVOIR PROPERTIES

by

Godofredo Pérez

A DISSERTATION

APPROVED FOR THE DISCIPLINE OF

PETROLEUM ENGINEERING

By Dissertation Committee

Balmohan A Kelkar, Chairperson

Peyton Cook

Edgar Kaya

Guerrero

ABSTRACT

Pérez, Godofredo (Doctor of Philosophy in Petroleum Engineering)

Stochastic Conditional Simulation for Description of Reservoir Properties

(227 pp. - Chapter V)

Directed by Dr. Mohan Kelkar

(338 words)

In this work, a stochastic conditional simulation method founded on the simulated annealing optimization technique is developed to generate three-dimensional distributions of a property. This method allows one to specify various conditioning data, and univariate and spatial statistical attributes in the distributions of a property. The conditional simulation method is robust, and reproduces all the specifications and constraints with remarkable accuracy. A new testing procedure is introduced to evaluate the quality of the simulations and is used to establish the simulated annealing parameters required for high quality and small computation cost simulations.

Field studies in sandstone and carbonate environments indicate that descriptions of reservoir properties generated by the stochastic conditional simulation method closely resemble the observed distributions provided that adequate models are used to represent the spatial correlation structure of the properties. Yet, in most field cases, the spatial correlation structure maybe the greatest source of uncertainty in the descriptions. Assuming that same amount of information is available, the conditional simulation method based on simulated annealing is compared with other conditional simulation methods. The quality of the descriptions generated by the method developed in this work is similar or superior than for other conditional simulation methods tested. For similar degrees of quality, the conditional simulation method of choice should be the one that yields the smallest computation cost.

A field study is conducted in a carbonate reservoir using horizontal and vertical well log data to evaluate the statistical attributes of the inter-vertical-well distribution of reservoir properties. Although the evaluations suggest that the porosity correlation range in the inter-well region is more than two times longer than in the vertical direction, other statistical attributes in the horizontal and vertical directions are not significantly different. Guidelines are developed to assess information about the inter-well spatial correlation of reservoir properties from vertical well data. Conditional simulations of porosity generated in this reservoir highlight the role of the spatial correlation models on simulated inter-well distributions.

ACKNOWLEDGEMENTS

I thank Dr. Mohan Kelkar for his advice, enthusiasm and guidance throughout this investigation. I also thank Dr. Peyton Cook, Dr. E.T. Guerrero and Dr. Ekrem Kasap for providing suggestions and reviewing this work.

I am grateful for the financial support provided by the Tulsa University Reservoir Exploitation Projects (TUPREP), the Department of Energy (DOE) and the University of Tulsa during my studies.

I dedicate this work to my parents Víctor and María. The support and encouragement they provided me all my life has been essential for the completion of this work.

TABLE OF CONTENTS

TABLE OF CONTENTS (Continued)

TABLE OF CONTENTS (Continued)

LIST OF TABLES

LIST OF TABLES (Continued)

LIST OF FIGURES

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

LIST OF FIGURES (Continued)

CHAPTER I

INTRODUCTION

Description of reservoir properties is an important aspect of several engineering evaluations, including assessment of hydrocarbon reserves, selection of infill well locations, and prediction of displacement processes performance. As a result, description of reservoir properties can have a direct impact in the economic decisions and benefits of a reservoir exploitation project. This view has been a major motivation in the petroleum industry to develop new methodologies to improve reservoir description, as indicated by recent investigations reported in two reservoir characterization conferences.[1,2]

The nature of the problems in the description of reservoir properties is related to availability of sample data and geologic complexities. The sample data in most reservoirs are available at vertical wells and represent only a small fraction of the reservoir. Such sample data include properties derived from cores, logs and well test analysis. Other measurements and sources of information which are valuable for reservoir description include geophysical data, geologic interpretations, evaluations of outcrop analogs, and production history. In addition to the scarcity of sample data, the description process must account for geologic complexities, such as characteristics of the depositional environment, which may be reflected as spatial relationships in the distribution of a property. These relationships include spatial variability and anisotropy observed in the sample data of reservoir properties.

Stochastic conditional simulation is a geostatistical method to generate descriptions of reservoir properties which use the available scarce data and the spatial relationships outlined above. The most important elements of the method are noted within the terms stochastic, conditional and simulation. This method is a stochastic approach because reservoir properties are represented by random variables or stochastic processes. Descriptions of properties generated by this method are conditional since the available data are honored at the sampled locations. And last,

the method simulates or predicts several equiprobable descriptions of the actual distribution of a property in the reservoir.

Although, in this investigation, the stochastic conditional simulation approach is developed exclusively for the generation of reservoir properties, such as porosity and permeability, in general, this approach can be used to develop descriptions of other geologic features. Haldorsen and Damsleth[3] review the state of the art in stochastic modeling of different geologic features, such as sand bodies, shales, facies, fractures and faults. An extensive description of these stochastic modeling methods is given by Haldorsen et al.[4] Recently, Damsleth et al.[5] proposed a hybrid approach which combines two stochastic modeling methods to generate facies descriptions and distributions of reservoir properties within the facies. The stochastic conditional simulation method described in this dissertation utilizes geostatistical relations to represent the spatial correlation structures of properties. Any large scale geologic features that might be apparent in the descriptions generated by this method are a consequence of the information embedded in the statistics of the sample data.

Geostatistical methods for the spatial description of properties were originally introduced for mining applications. The theoretical aspects of geostatistics are given in details by Journel and Huijbregts[6] and summarized by Journel.[7] Extensive examples on the application of geostatistical methods are described by Isaaks and Srivastava.[8] In addition, geostatistical methods have been used in hydrology to describe groundwater flow parameters.[9] In petroleum engineering, one of the first applications of geostatistics to describe reservoir properties has been presented by Da Costa e Silva.[10]

The objective of geostatistical methods is to generate a distribution of a property which satisfies a specified model for the spatial correlation. Most geostatistical methods, such as ordinary kriging, indicator kriging and others described by Journel and Huijbregts,[6] use a function known as the semi-variogram to represent the spatial correlation of a variable. The semi-variogram allows one to quantify the variability of a property as a function of distance and direction and it is used extensively throughout this investigation. The definition of the semi-variogram, γ, of

a variable V is given by[6]

$$\gamma(\vec{h}) = \frac{1}{2}\mathrm{E}\left[V(\vec{\omega}_i + \vec{h}) - V(\vec{\omega}_i)]^2\right], \tag{1.1}$$

where, $\vec{h}$ is the lag vector, $\vec{\omega}$ represents the locations of the variable V in the region considered and the symbol E denotes the expectation operator. Perhaps, the better known covariance function, $COV(\vec{h})$, is related to the semi-variogram by[6]

$$COV(\vec{h}) = COV(0) - \gamma(\vec{h}). \tag{1.2}$$

The estimator of the semi-variogram is[6]

$$\hat{\gamma}(\vec{h}) = \frac{1}{2N_p(\vec{h})} \sum_{i=1}^{N_p(\vec{h})} \left[V(\vec{\omega}_i + \vec{h}) - V(\vec{\omega}_i)\right]^2, \tag{1.3}$$

where, $N_p(\vec{h})$ are the number of pairs of the variable V separated by a distance or lag $|\vec{h}|$ along the direction of the vector $\vec{h}$. For example, in a well log of a property with n samples uniformly spaced by 1 foot, the number of pairs in the semi-variogram are $N_p(1 \text{ foot}) = n - 1$, $N_p(2 \text{ feet}) = n - 2$, ...

Usually, the sample semi-variograms are presented as a plot of $|\vec{h}|$ versus $\gamma(\vec{h})$ for one or more directions. For a correlated variable, the magnitude of the semi-variogram is small for short lags and increases as the lag increases until it reaches an approximate constant value denoted as the sill, as illustrated in Figure 1.1. Normally, the sill is close to the sample variance. The correlation range denotes the lag at which the semi-variogram reaches the sill and for lags greater than the correlation range, the variable is uncorrelated. The spatial correlation models used in geostatistical methods are mathematical expressions fitted to the sample semi-variogram plot. Several types of spatial correlation models or mathematical expressions are available which satisfy the condition of positive definiteness required for kriging applications[6] and a few of these models are described in the following chapters.

One of the first geostatistical conditional simulation method based on kriging is described by Journel and Huijbregts.[6] The simulated distribution of a property

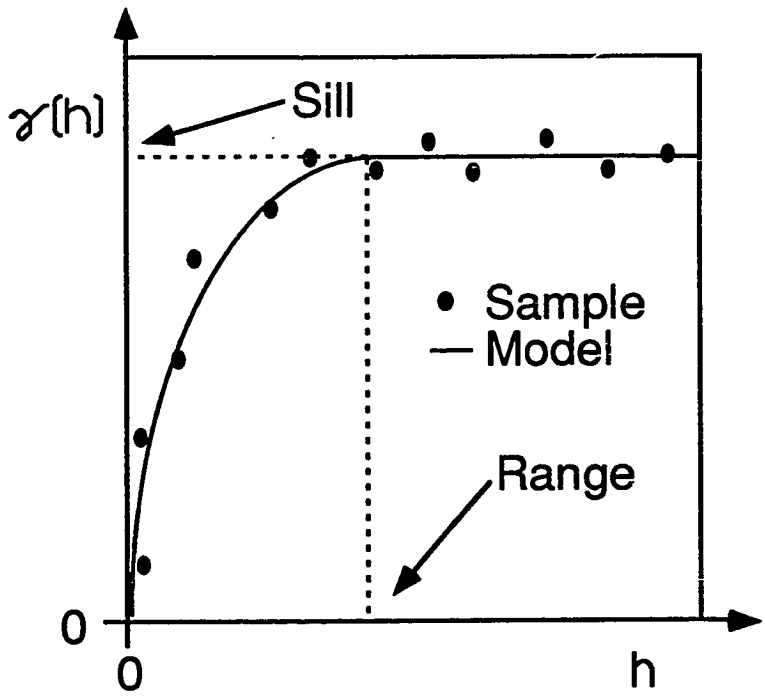

Figure 1.1 - Schematic of sample and model semi-variograms

generated by this method consists of the sum of the kriging estimates and a distribution of unconditional random components. These random components must have the same spatial correlation structure as the kriging estimates. The turning bands method[11] can be used to generate these unconditional random distributions with a prescribed spatial correlation model. Farmer[12] describes several other methods to generate unconditional distributions. Recently, Journel and Alabert[13] introduced the sequential indicator simulation method based on indicator kriging. According to Journel and Alabert[13] an advantage of the sequential indicator simulation method is that it accounts for the spatial connectivity of extreme values of a property, such as permeability, which can have an important effect on the presence of preferential flow paths in reservoirs. Journel and Gómez-Hernández[14] presented a field study using the sequential indicator simulation method to generate sand-shale sequences in an oil reservoir. The sequential indicator simulation method is described and used in the Carbonate Field section of Chapter III.

In addition to the conventional spatial correlation models developed in geostatistical applications, the concept of statistical fractals can be used to represent the spatial correlation of a property. Statistical fractals are stochastic processes which closely resemble the records of several natural phenomena.[15,16] The spatial correlation models of statistical fractal processes are characterized by a parameter known as the fractal dimension and these models are described in more detail in the Fractal Models section of Chapter III. Hewett[17] proposed the use of statistical fractals to analyze and represent spatial correlation of porosity logs in vertical wells. Hewett[17] developed a stochastic conditional simulation method to generate distributions of porosity in a vertical reservoir section between two conditioning wells based on assumptions about the spatial correlation structure of porosity in the inter-well region. This stochastic conditional simulation method[17] is described and used in the Sandstone Field Section in Chapter III. Furthermore, Hewett[17] presented a hybrid simulation approach to extend the use of the conditional simulations of properties in vertical reservoir sections to predict the production performance of large fields. A great interest in Hewett's approach arose when several field studies[18–22] showed

that this approach closely predicts the production histories. Other recent investigations presented by Aasum et al.[23] and Crane and Tubman[24] used statistical fractals to study the performance of miscible flooding.

The major objectives of the research reported in this dissertation are:

(1) Develop a stochastic conditional simulation method to generate three-dimensional descriptions of reservoir properties which overcomes the restrictions of currently available methods.

(2) Establish the validity of a stochastic conditional simulation method to represent the distributions of properties observed in the field.

(3) Evaluate the nature of the distribution and spatial correlation of reservoir properties in the inter-well regions, which is critical for stochastic conditional simulation applications, using horizontal and vertical well data.

The Stochastic Conditional Simulation chapter presents the development of a stochastic conditional simulation method based on the simulated annealing technique. A general description of the simulated annealing technique, and the basis and motivations to select this technique are given in the Background section of Chapter II. Also, Chapter II presents the optimization of the conditional simulation method. Chapter III reports the field validations of the stochastic conditional simulation method developed in Chapter II to generate descriptions of reservoir properties in a sandstone and a carbonate field and presents comparisons with other conditional simulation methods. In Chapter IV, log data from a horizontal well and several closely spaced vertical wells in a carbonate reservoir are used to evaluate the distribution of reservoir properties in the inter-well region and to derive guidelines to assess the spatial correlation of a property in the inter-well regions from vertical well data. In addition, Chapter IV describes the effects of different spatial correlation models for the inter-well regions on the distributions of properties generated by stochastic conditional simulation. Finally, the conclusions of this dissertation are summarized in Chapter V.

CHAPTER II

STOCHASTIC CONDITIONAL SIMULATION

This chapter describes the development of a stochastic conditional simulation method to generate multi-dimensional distributions of a random variable with specified univariate and spatial statistical attributes. As discussed in the Chapter I, this method is useful to generate descriptions of properties because it accounts for heterogeneities, such as spatial variability and anisotropy, observed in the sample data of reservoirs.

The basis of the stochastic conditional simulation method described in this chapter is the simulated annealing optimization technique. Therefore, the first section of this chapter provides a background and a description of simulated annealing. In the following sections, the three-dimensional conditional simulation problem is formulated and all the definitions and steps of the simulation method are described. Then, a procedure is designed to evaluate the performance of the conditional simulation method. This procedure is used to derive the simulation parameters, referred as the annealing schedule, required for computationally efficient simulations. Additional numerical experiments are conducted to examine the effect of different degrees of information about the statistical specifications on the magnitude of the uncertainties of the simulated distributions. Finally, potential extensions of the stochastic conditional simulation method are briefly discussed.

Background

Farmer[25] introduced the "Numerical Rocks" conditional simulation method based on simulated annealing to generate patterns of discrete rock types representing different lithologies. The statistical properties of the rock types imposed in the patterns generated by this conditional simulation method are a single-point histogram, a two-point histogram and a correlation function. Farmer[25] defines

two-point histograms as the density distribution of sample pairs as a function of separation distance and the direction between pairs. The correlation function is calculated using the two-point histogram functions. The main difference between the conditional simulation method described in this chapter and Farmer's[25] method is that semi-variograms are used to represent the spatial correlation of a continuous variable instead of two-point histograms and correlation functions. Additional differences between these methods pertain to the annealing schedule parameters and are discussed in the Simulation Parameters and Annealing Schedule sections.

Simulated annealing is a simple, yet general and flexible, optimization technique which has been applied to problems in diverse disciplines. According to Aarts and Korst,[26] the simulated annealing optimization method was introduced independently by Kirkpatrick et al.[27] and Cerny.[28] A thorough literature review, which includes 292 articles about developments and implementations of simulated annealing optimization, is given by Collins et al.[29] Some areas of application of simulated annealing include placement of computer components, wiring of electronic components and image processing. Several investigations reported by Collins et al.[29] used the traveling salesman problem (calculation of the shortest distance required to visit several sites) as a benchmark for different implementations of simulated annealing.

Simulated annealing is founded in concepts from statistical mechanics and the annealing process. Therefore, the implementation of simulated annealing requires the definition of three components of a system which are analogous to the energy, the temperature and the interactions of a molecular system. The objective function represents the energy of a system and it is defined as the function to be minimized (or maximized). The control parameter represents the temperature of a system which is an independent parameter and not necessarily related to any other parameter of the problem. The interchange mechanism corresponds to the molecular interactions and it usually consists of a finite set of perturbations to the independent variables which produce a change of the objective function. The following sections provide more specific definitions of these components for the conditional simulation problem studied in this chapter.

The goal of the simulated annealing method is to determine the configuration of the independent variables that yields a global minimum of the objective function. The simulation process consists of performing a specified number of changes to the independent variables with the interchange mechanism and for each change, the objective function is evaluated. A change of the variables is accepted and retained if it reduces the magnitude of the objective function. If a change increases the magnitude of the objective function, it is accepted or rejected and discarded according to the probability function proposed by Metropolis et al.[30] This probability function, introduced by Metropolis et al.[30] to study the statistical mechanics of molecular systems, is a Boltzmann distribution which depends on the energy and the temperature of the system. At high temperatures, when the energy and the degree of disorder of the system is large, there is a high probability of accepting a change which results in an increase in the energy. As the temperature decreases, the energy decreases and the system (molecules) becomes more ordered, and at this stage, there is a small probability of accepting changes which result in an increase of the energy of the system. This probabilistic feature of simulated annealing allows the system to escape from solutions which correspond to local minima of the objective function. Other algorithms, such as the iterative improvement or the greedy algorithm, which accept only the changes resulting in a decrease of the energy of the system may become easily trapped in a local minimum solution.

The procedure to reduce the temperature or control parameter and specify the number of iterations or changes required at each temperature is known as the annealing schedule. Annealing is the process of carefully reducing the temperature of a metal, from an initial high temperature (high energy state near the melting point) to a low temperature (low energy state near the freezing point), so that at the final state the metal can achieve a high strength. If the temperature is reduced too fast, a metal will cool to an unstable state (analogous to a local minimum). At each temperature, the number of iterations or changes should be large enough so that the energy achieves an equilibrium state. The temperature is reduced until a stopping criteria is satisfied. Stopping criteria can be when the energy becomes

small or until further changes at subsequent temperatures have negligible effect on the changes in the energy function.

Several annealing schedules have been proposed in the simulated annealing literature. The review by Collins et al.[29] reports nine temperature reduction functions, eleven number of iterations schemes and five stopping criteria. Most annealing schedules can be divided into two major groups on the basis of some common features. In one group, the number of total iterations per step is constant and the size of the step or decrement of the control parameter changes in successive steps. In the second group, the number of total iterations per step is adjusted between steps and the size of the steps is fixed. Theoretical developments by Aarts and Korst[26] and Geman and Geman[31] using Markov random fields theory demonstrate the convergence properties of simulated annealing and derive annealing schedules. Aarts and Korst[26] noted that these theoretical annealing schedules lead to an exponential-time execution of the algorithm and showed that other approximate annealing schedules provide near optimal solutions for a significantly smaller execution time. The annealing schedule used in the conditional simulation method developed in this chapter is described in the Annealing Schedule section.

Problem Formulation

The objective of the stochastic conditional simulation method described in this chapter is to generate a distribution of a random variable with specified statistical properties in a three-dimensional rectangular grid. The statistical properties and constraints specified in this conditional simulation method are:

(1) conditioning data,
(2) distribution function and
(3) semi-variogram models for several directions.

The simulation variable is defined in a three-dimensional rectangular region discretized using N_x, N_y and N_z grid points along the x, y and z directions, respectively. The coordinates of a point, (x_i, y_j, z_k), in this rectangular region are given by

$$x_i = O_x + i\Delta_x, \tag{2.1}$$

$$y_j = O_y + j\Delta_y \tag{2.2}$$

and

$$z_k = O_z + k\Delta_z, \tag{2.3}$$

where, O_x, O_y and O_z are the coordinates of the origins of the simulation region and Δ_x, Δ_y and Δ_z are the spacing between grid points in the x, y and z directions, respectively. In Equations 2.1 through 2.3, $1 \leq i \leq N_x$, $1 \leq j \leq N_y$ and $1 \leq k \leq N_z$. The total number of simulation points, N_s, is

$$N_s = N_x N_y N_z. \tag{2.4}$$

For convenience in the presentation of later derivations, the coordinates of the location vector of a grid point, $\vec{\omega}_l = (x_i, y_j, z_k)$, are defined with a single index, l, as follows

$$k = 1 + \text{INT}\left[\frac{l-1}{N_x N_y}\right], \tag{2.5}$$

$$j = 1 + \text{INT}\left[\frac{l-(k-1)N_x N_y - 1}{N_x}\right] \tag{2.6}$$

and

$$i = l - (k-1)N_x N_Y - (j-1)N_x, \tag{2.7}$$

where, $l = 1, \ldots, N_s$ and the symbol INT refers to the integer part of the number inside the brackets. Equations 2.5 through 2.7 allow a sequential arrangement of the coordinates, such that the x, y and z coordinates cycle first, second and third, respectively. Hence, the coordinates of the location vector are $\vec{\omega}_1 = (x_1, y_1, z_1)$, ..., $\vec{\omega}_{N_x} = (x_{N_x}, y_1, z_1)$, $\vec{\omega}_{N_x+1} = (x_1, y_2, z_1)$, ..., $\vec{\omega}_{2N_x} = (x_{N_x}, y_2, z_1)$, ... , ..., $\vec{\omega}_{N_s} = (x_{N_x}, y_{N_y}, z_{N_z})$. The simulation variable is $V(\vec{\omega}_l)$, for $l = 1, \ldots, N_s$

The conditioning data values are assigned to the simulation variable at specified locations. The conditioning data remain fixed at the specified locations throughout the conditional simulation. The constraints corresponding to a number N_c of conditioning data can be expressed as

$$V(\vec{\omega}_{c,l}) = V_c(\vec{\omega}_{c,l}), \tag{2.8}$$

where, V_c are the values of the conditioning data and $\vec{\omega}_{c,l}$ for $l = 1, \ldots, N_c$ are the set of locations corresponding to the conditioning data. Usually, the conditioning data are a set of measurements available at sampled locations, such as log or core measurements of the property to be simulated.

The cumulative distribution function represents the probability that the variable is smaller than a given value. The discrete cumulative distribution function, $F(V_{f,k})$, of the simulation variable, for an upper class limit value equal to $V_{f,k}$, is given by

$$F(V_{f,k}) = \frac{1}{N_s} \sum_{l=1}^{N_s} i(\vec{\omega}_l; V_{f,k}), \tag{2.9}$$

where, $k = 1, \ldots, N_f$ and N_f is the number of classes used to discretize the cumulative distribution function. The indicator function, i, in Equation 2.9 is

$$i(\vec{\omega}_l; V_{f,k}) = \begin{cases} 1, & V(\vec{\omega}_l) \leq V_{f,k} \\ 0, & V(\vec{\omega}_l) > V_{f,k} \end{cases}. \tag{2.10}$$

The specified discrete cumulative distribution function, F_o, for the simulation variable requires the following constraints

$$F(V_{f,k}) = F_o(V_{f,k}), \tag{2.11}$$

for $k = 1, \ldots, N_f$. The cumulative distribution function must be a strictly increasing function, or equivalently, $0 < F_o(V_{f,1}) < F_o(V_{f,2}) \ldots < F_o(V_{f,N_f}) = 1$. In practice, the specified discrete cumulative distribution function can be calculated from the conditioning data or may be available from other sources of information.

As explained in Chapter I, the semi-variogram is used in geostatistics to quantify the spatial correlation of a variable as a function of separation distance and direction. The sample semi-variogram, γ_s, of the simulation variable, V, is given by

$$\gamma_s(\vec{h}_{l,i}) = \frac{1}{2N_p(\vec{h}_{l,i})} \sum_{j=1}^{N_p(\vec{h}_{l,i})} \left[V(\vec{\omega}_j + \vec{h}_{l,i}) - V(\vec{\omega}_j) \right]^2, \tag{2.12}$$

where, $\vec{h}_{l,i} = (h_{x,l,i}, h_{y,l,i}, h_{z,l,i})$ is the lag distance vector and $N_p(\vec{h}_{l,i})$ is the number of pairs within the simulation grid corresponding to the lag $\vec{h}_{l,i}$. For the lag distance

vector in Equation 2.12, the index i refers to a direction and the index l refers to the lag distances considered for a given direction. The number of semi-variogram directions considered in the simulations is N_d and for each direction i the number of semi-variogram lags is $N_{h,i}$. Hence, the range of the indices for the lag distance vector in Equation 2.12 are $l = 1, \ldots, N_{h,i}$ and $i = 1, \ldots, N_d$. The semi-variogram models specified in the conditional simulations are denoted by γ_o and require that

$$\gamma_s(\vec{h}_{l,i}) = \gamma_o(\vec{h}_{l,i}), \tag{2.13}$$

for $l = 1, \ldots, N_{h,i}$ and $i = 1, \ldots, N_d$. Usually, the semi-variogram models along different directions are derived by fitting one of the conventional models through the sample semi-variogram of the available data of the simulation variable or it may be necessary to infer a model based on external information if the available data is undersampled along a specific direction. The mathematical expressions for the semi-variogram models widely used in geostatistics and the fractal methods are given in the description of the implementation of the conditional simulation method in Appendix A.

It should be noted that there are two major differences in the role of the semi-variograms in other geostatistical methods, such as kriging, and the conditional simulation method described in this chapter. First, kriging methods require semi-variogram models which satisfy the conditions of positive definiteness[6] while the semi-variogram models specified in Equation 2.13 are not restricted to models which satisfy this condition. Second, kriging methods require a semi-variogram model along all the directions between the location where a variable is estimated or simulated and the locations of the surrounding sample data. In the conditional simulation method discussed in this chapter, the semi-variogram models are reproduced only along the directions specified in Equation 2.13, and further assumptions about the spatial correlation structure for other directions are not needed.

All the constraints and specifications required in a conditional simulation of a variable are given in Equations 2.8, 2.11 and 2.13. The total number of constraints, N_o, for the conditional simulation is

$$N_o = N_c + N_f + \sum_{i=1}^{N_d} N_{h,i}. \tag{2.14}$$

The simulation procedure to obtain distributions of the variable V which satisfy these N_o specifications is described in the Initial Distribution and the Simulated Annealing sections.

Initial Distribution

The first step of the conditional simulation method is to generate an initial distribution of the simulation variable which satisfies the conditioning data constraints and cumulative distribution function specifications given by Equations 2.8 and 2.11. This first step can be readily accomplished using the transformation method[32] to generate a distribution of a variable with a specified cumulative distribution function.

The procedure to generate the initial distribution consists of visiting all the points of the simulation grid, one at a time. If the location of a point is equal to the location of a conditioning data point ($\vec{\omega}_{c,i}, i = 1,\ldots,N_c$), then the variable is assigned the specified conditioning data value, as required in Equation 2.8. Otherwise, the variable is assigned a random value from the specified cumulative distribution function.

The transformation method is used to assign a random value from a specified distribution function and it is illustrated in Figure 2.1. A random number, R, in the range of (0,1] is drawn from a uniform random generator[33] and the variable is assigned a value according to one of the following two alternatives

$$V(\vec{\omega}_l) = V_{f,k-1} + (V_{f,k} - V_{f,k-1})\frac{R - F_o(V_{f,k-1})}{F_o(V_{f,k}) - F_o(V_{f,k-1})} \tag{2.15}$$

for a uniform subclass distribution, or

$$V(\vec{\omega}_l) = \frac{V_{f,k-1} + V_{f,k}}{2} \tag{2.16}$$

for a discrete distribution, for $l = 1,\ldots,N_s$. In Equations 2.15 and 2.16, the index k is such that $F_o(V_{f,k-1}) \leq R < F_o(V_{f,k})$ and for the case when $k = 1$ then $F_o(V_{f,0}) = 0$ and $V_{f,0}$ is set equal to the specified minimum value of the distribution. Equation 2.15 assumes that the distribution between the lower and upper limits of

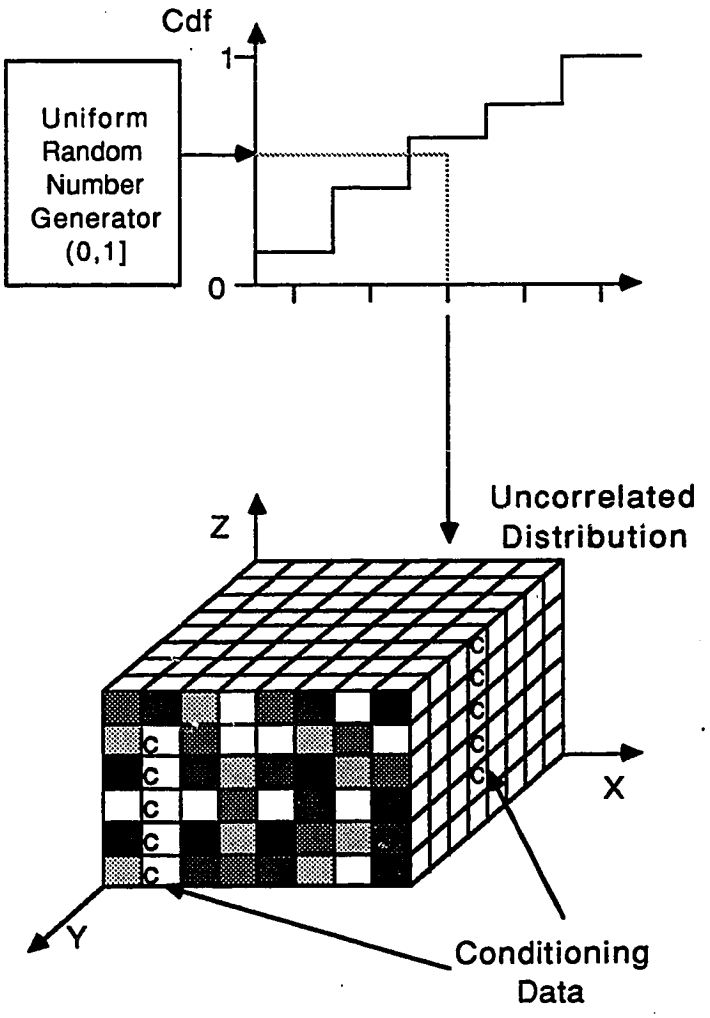

Figure 2.1 - Schematic of the transformation method to generate an initial distribution

a class is a uniform distribution and Equation 2.16 is used to assign discrete values equal to the class marks.

The transformation method described in this section yields an uncorrelated distribution of the simulation variable. The results in the Reproduction of Statistics section, described later, show that the distributions generated with this method are in close agreement with the specified distributions. Minor deviations from the specified distribution function might occur if the distribution of the conditioning data is different than the specified distribution. However, these deviations will seldom happen, since in most cases the specified distribution function is derived from the same sample data used as conditioning data. The sample semi-variograms of the initial distribution of the simulated variable for all directions are close to the sample variance of the distribution.

Simulated Annealing

The objective of the simulated annealing method is to rearrange the initial distribution generated in the previous section until the semi-variogram specifications given by equation 2.13 are satisfied. The rearrangements of the spatial distribution of the variable performed by the simulation process will not affect the specifications imposed by the initial distribution (Equations 2.8 and 2.11). Therefore, upon the completion of the simulation process described in this section, the simulated variable will fulfill all the required specifications and constraints.

The basic elements of simulated annealing are defined and described in the Objective Function, Interchange Mechanism and Metropolis Condition sections. Then, these elements are integrated in the Simulation Process section.

Objective Function

The objective function is the function to be minimized with the simulated annealing method and it is analogous to the energy of the system in the context of statistical mechanics. The objective function is defined as a function of the difference between the sample semi-variograms of the simulation variable and the

required semi-variogram models. Therefore, by minimizing this objective function, the requirements in Equation 2.13 can be satisfied.

The objective function, E, is defined as

$$E^k = \frac{1}{E_0} \sqrt{\sum_{i=1}^{N_d} \sum_{l=1}^{N_{h,i}} \left[\frac{\gamma_s^k(\vec{h}_{l,i}) - \gamma_o(\vec{h}_{l,i})}{\gamma_o(\vec{h}_{l,i})} \right]^2}, \tag{2.17}$$

where, the index k refers to the iteration number within a step and E_0 is a normalizing constant equal to

$$E_0 = \sqrt{\sum_{i=1}^{N_d} \sum_{l=1}^{N_{h,i}} \left[\frac{\gamma_s^0(\vec{h}_{l,i}) - \gamma_o(\vec{h}_{l,i})}{\gamma_o(\vec{h}_{l,i})} \right]^2}. \tag{2.18}$$

In Equation 2.18, $\gamma_s^0(\vec{h}_{l,i})$ are the semi-variograms of the simulation variable at initial conditions calculated from the initial distribution of the variable. The purpose of the normalizing constant used in Equation 2.17 is to assign a value of one to the objective function at the initial conditions for any distribution or specifications.

The advantage of the objective function defined in Equation 2.17 over other possible definitions is the ability to assign a different weight to each semi-variogram term in the function. It can be noted in Equation 2.17 that differences between the sample semi-variogram of the variable and the specified model $(\gamma_s(\vec{h}) - \gamma_o(\vec{h}))$ are assigned greater weights $(1/\gamma_o(\vec{h}))$ at smaller lags $(|\vec{h}|)$ because the magnitude of the semi-variogram model decreases as the lag becomes smaller. This feature is useful because it amplifies the effect of the semi-variograms for small lags in the objective function, which are the terms with greater contribution to the magnitude of the energy function.

Interchange Mechanism

The interchange mechanism are perturbations to the distribution of the simulation variable resulting in changes in the objective function. There are several different ways to design an interchange mechanism. However, for efficiency reasons, the interchange mechanism should be simple, so that the computational costs required to update the objective function for one interchange is as small as possible.

The interchange mechanism implemented in the conditional simulation method consists of exchanging or swapping the values of the variable at two locations selected at random in the simulation region. The exchange locations within the simulation region are selected by drawing at random two integer numbers, I_1 and I_2, between one and the number of simulated grid points (N_s) as follows

$$I_1 = 1 + \text{INT}(N_p R_1) \tag{2.19}$$

and

$$I_2 = 1 + \text{INT}(N_p R_2), \tag{2.20}$$

where, R_1 and R_2 are two random numbers from a uniform distribution in the range (0,1] and the symbol INT indicates the integer part of a number.

In addition, prior to exchanging the values of the two selected random points, the following three conditions must be satisfied:

(1)

$$\vec{\omega}_{I_1} \neq \vec{\omega}_{c,l}, \tag{2.21}$$

(2)

$$\vec{\omega}_{I_2} \neq \vec{\omega}_{c,l} \tag{2.22}$$

for $l = 1, \ldots, N_c$ and

(3)

$$|V(\vec{\omega}_{I_1}) - V(\vec{\omega}_{I_2})| > \Delta V_{f,min}. \tag{2.23}$$

In Equation 2.23, $\Delta V_{f,min}$ is a constant equal to the smallest inter-class size of the distribution function and it is given by

$$\Delta V_{f,min} = \text{MIN}[V_{f,i} - V_{f,i-1}] \tag{2.24}$$

for $i = 1, \ldots, N_f$ and the symbol MIN refers to the minimum value. The first two conditions (Equations 2.21 and 2.22) require that the selected random points do not coincide with the locations of the conditioning data. If one of these two conditions are not satisfied, then new random points are drawn according to Equations 2.19 and

2.20, as required. The third condition (Equation 2.23) does not allow to exchange the selected points when the magnitude of the difference of the variable at these points is smaller than the smallest magnitude of the difference between consecutive class limits. If this third condition is not satisfied, one of the locations of the random points is selected again with Equation 2.19 or 2.20. This last condition is not necessary to satisfy the simulation specifications; but it eliminates further computational efforts required by unnecessary exchanges which have none or small effect on the objective function.

The actual interchange of the values of the simulation variable at the selected random locations is performed later in the simulation process. Nevertheless, it is possible to calculate the semi-variograms and the corresponding objective function that will result if the exchange is performed. During the simulation process many iterations of the above exchanges will be performed; therefore, it is not feasible to recalculate the sample semi-variograms using Equation 2.12. Instead, the new values of the semi-variograms of the distribution of the simulated variable, γ'_s, that result due to the exchange of the random points is calculated by

$$\gamma'_s(\vec{h}_{l,i}) = \gamma^k_s(\vec{h}_{l,i}) + \lambda^k(\vec{h}_{l,i}), \tag{2.25}$$

where, $\lambda^k(\vec{h}_{l,i})$ is a semi-variogram correction term. If the two random points are interchanged, then the new objective function, E', can be calculated by inserting Equation 2.25 into Equation 2.17, or

$$E' = \frac{1}{E_0} \sqrt{\sum_{i=1}^{N_d} \sum_{l=1}^{N_{h,i}} \left[\frac{\gamma^k_s(\vec{h}_{l,i}) + \lambda^k(\vec{h}_{l,i}) - \gamma_o(\vec{h}_{l,i})}{\gamma_o(\vec{h}_{l,i})} \right]^2}. \tag{2.26}$$

In Equations 2.25 and 2.26, $l = 1, \ldots, N_{h,i}$ and $i = 1, \ldots, N_d$, and the index k refers to an iteration within a step.

During the simulation process, most of the computational effort is dedicated to the calculation of the correction terms $\lambda^k(\vec{h})$ in Equation 2.25. Accordingly, in practice, it is vital to calculate these correction terms in as small number of operations as possible. The formulas developed to calculate the correction terms in Equation 2.25 are described in Appendix B.

Metropolis Condition

The Metropolis condition is a Boltzmann distribution introduced by Metropolis et al.[30] to study the statistical mechanics of molecular systems. Metropolis et al.[30] used this distribution to calculate the probability of moving a molecule in a lattice as a function of the change of energy resulting from this move and the temperature of the system. In the simulated annealing method, the Metropolis condition is used to calculate the probability of the transition between two states of the objective function at a given value of a control parameter analogous to the temperature of the system.

The Metropolis condition is given by

$$P(\Delta E^k, T^r) = \begin{cases} 1, & \Delta E^k \leq 0 \\ \exp\left(\frac{-\Delta E^k}{T^r}\right), & \Delta E^k > 0, \end{cases} \tag{2.27}$$

where, the index k refers to an iteration within a step represented by the index r. In Equation 2.27, the change of the objective function, ΔE^k, due to one interchange is

$$\Delta E^k = E' - E^k, \tag{2.28}$$

where, the terms on the right hand side are given by Equations 2.17 and 2.26. The implementation of the probability function given by Equation 2.27 is discussed in the Simulation Process section.

The control parameter, T^r, in Equation 2.27 is calculated according to the following equation

$$T^r = \alpha T^{r-1}, \tag{2.29}$$

where, the index r refers to a step number equal to $r = 2, 3, \ldots$. For the first step, $r = 1$ in Equation 2.29, the control parameter is set equal to the initial value described in the Initial Control Parameter section. The parameter α in Equation 2.29 is the convergence rate factor and it is a constant value in the range $0 < \alpha < 1$ throughout the simulation. The effect of the convergence rate factor on

the computational efficiency of simulated annealing is discussed in the Optimum Schedule section.

As noted in the Background section, several ways to calculate the control parameter have been proposed in different investigations. For the method discussed in this chapter, other forms of Equation 2.29 were considered, such as the equation proposed by Aarts and Korst[26] which uses an additional parameter referred as the specific heat of the system. However, preliminary tests indicated no advantages of other schedules over the simple equation given by Equation 2.29.

The Metropolis condition allows to perform exchanges of two values of the simulated variable, even if a change increases the magnitude of the objective function. The probability of approving an exchange which results in an increase of the magnitude of the objective function decreases as the magnitude of the control parameter is decreased from step to step according to Equation 2.29. Figure 2.2 illustrates this effect of the control parameter on the Metropolis condition. For a fixed value of the change of objective function, the probability of accepting an exchange decreases as the control parameter becomes smaller.

Simulation Process

The simulation process consists of a series of steps and within each step several iterations are performed. At each step, the value of the control parameter is reduced gradually according to Equation 2.29. Then, within each step, several iterations are performed using the interchange mechanism described in the Interchange Mechanism section.

In addition to the specifications and constraints for the simulation variable (Equations 2.8, 2.11 and 2.13), the simulation process of simulated annealing requires other parameters. The methods designed to evaluate two parameters, the initial control parameter (T^0) and the maximum number of total iterations per step (M_t), are described in the Simulation Parameters section. The remaining parameters, referred as the annealing schedule, are the convergence rate factor (α), the maximum number of accepted iterations per step (M_a), the tolerance for the

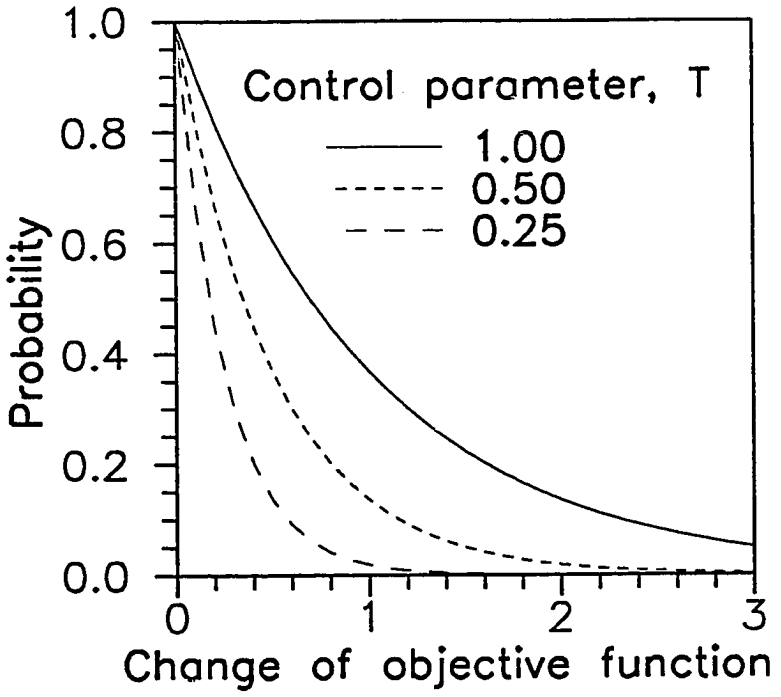

Figure 2.2 - Probability function ($P(\Delta E)$) for the Metropolis condition

objective function (ϵ_v) and the tolerance for the acceptance ratio (ϵ_a). These parameters have an important effect on the efficiency of the simulation and the quality of the results. A detailed analysis of these parameters is conducted in the Annealing Schedule section.

The simulation process is summarized in the flow chart shown in Figure 2.3. The following steps explain the details of this process.

(1) Calculate the initial distribution of the simulation variable, $V(\vec{\omega}_j)$ for $j = 1, \ldots N_s$, using Equation 2.15 or 2.16.

(2) Calculate the normalizing constant of the objective function, E_0, using Equation 2.18.

(3) Calculate the initial value of the control parameter, T^0, using Equation 2.40. The control parameter for the first step is set equal to the initial value, $T^1 = T^0$.

(4) For the first step, set the maximum number of total iterations per step equal to the maximum number of accepted iterations per step, $M_t^1 = M_a$.

(5) Select two random locations in the simulation grid, I_1 and I_2, using the interchange mechanism given by Equations 2.19 and 2.20.

(6) Calculate the sample semi-variograms, $\gamma_s'(\vec{h}_{l,i})$ for $l = 1, \ldots, N_{h,i}$ and $i = 1, \ldots, N_d$, using Equation 2.25 and the objective function, E', using Equation 2.26 that will result if the exchange is accepted by the Metropolis condition.

(7) Check if the Metropolis condition accepts the interchange proposed in Step (5). This test consists of comparing a random number, R, from a uniform distribution in the range (0,1] to the probability, $P(\Delta E^k, T^r)$, calculated from Equation 2.27. Note that if $\Delta E^k \leq 0$ the exchange is always accepted, according to the definition given by Equation 2.27. Depending on the results of the test in this step, sub-step (i) or (ii) is followed.

(i) If $R \leq P(\Delta E^k, T^r)$, the interchange proposed in Step (5) is accepted. Then, the distribution of the simulated variable is updated by the following changes. Using the following temporary variables, $W_1 = V(\vec{\omega}_{I_1})$ and $W_2 = V(\vec{\omega}_{I_2})$, then the exchange of values of the simulated variable is

$$V(\vec{\omega}_{I_1}) = W_2 \tag{2.30}$$

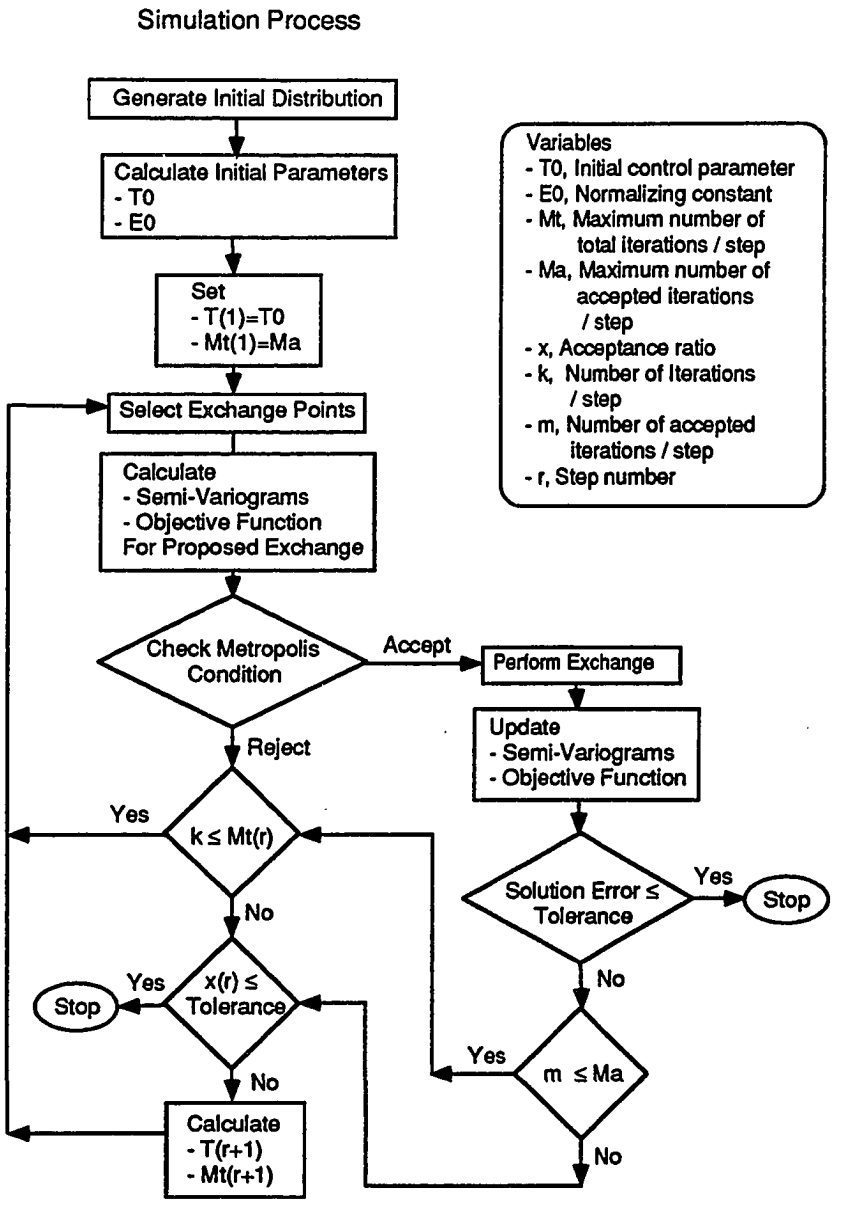

Figure 2.3 - Flow chart of the simulation process

and

$$V(\vec{\omega}_{I_2}) = W_1. \tag{2.31}$$

The update of the sample semi-variograms is given by

$$\gamma_s^{k+1}(\vec{h}_{l,i}) = \gamma_s'(\vec{h}_{l,i}), \tag{2.32}$$

for $l = 1, \ldots, N_{h,i}$ and $i = 1, \ldots, N_d$ and the update of the objective function is

$$E^{k+1} = E'. \tag{2.33}$$

Two additional conditions are tested in this sub-step.

(a) If the tolerance condition for the objective function,

$$\frac{E^k E_0}{\sum_{i=1}^{N_d} N_{h,i}} \leq \epsilon_v, \tag{2.34}$$

is satisfied then the simulation is stopped. The condition given by Equation 2.34 requires that root mean square average of the semi-variograms differences terms is less than a specified tolerance (ϵ_v).

(b) If the total number of accepted iterations in the current step exceed the maximum number of accepted iterations per step, M_a, then proceed to Step (9).

(ii) If $R > P(\Delta E^k, T^r)$, then the interchange proposed in Step (5) is not accepted. In this case, it is not necessary to update the simulation variable, the semi-variogram or the objective function since these will remain the same for the next iteration or equivalently

$$\gamma_s^{k+1}(\vec{h}_{l,i}) = \gamma_s^k(\vec{h}_{l,i}), \tag{2.35}$$

for $l = 1, \ldots, N_{h,i}$ and $i = 1, \ldots, N_d$,

$$E^{k+1} = E^k. \tag{2.36}$$

(8) Check if the number of total iterations k in the step r does not exceed the maximum number of total iterations per step, M_t^r. If the condition given by

$$k \leq M_t^r, \tag{2.37}$$

is satisfied, then proceed to Step (5).

(9) Check the acceptance ratio tolerance for the current step. The acceptance ratio is the fraction of the total iterations in a step which were accepted by the Metropolis condition. (Step (7), a). The acceptance ratio, χ^r, is given by

$$\chi^r = \frac{m}{k}, \tag{2.38}$$

where, m is the number of accepted iterations in the step r. Note, that for any step in the simulation $m \leq M_a$ and $M_a \leq k \leq M_t^r$. If the acceptance ratio condition given by

$$\chi^s \leq \epsilon_a \tag{2.39}$$

is satisfied, then the simulation is stopped.

(10) Calculate the control parameter, T^{r+1}, for the next step using Equation 2.29.

(11) Calculate the maximum number of iterations for the next step, M_t^{r+1}, using Equation 2.46.

(12) Proceed to Step (5).

The simulation process terminates when one of the conditions given by Equation 2.34 in Step (5,i,a) or Equation 2.39 in Step (9) is satisfied. Depending on the relative size of the specified tolerances, ϵ_v and ϵ_a, either condition may be satisfied first. The condition given in Equation 2.39 stops the simulation when many iterations are performed in a step with minimal changes in the objective function. Strictly, only the condition given by Equation 2.34 is sufficient; however, in some cases the condition given by Equation 2.39 allows to save the additional computational effort that will be required for only a marginal improvement in the final solution.

Simulation Parameters

This section describes the methods to calculate the value of the initial control parameter and the maximum number of total iterations per step required in the simulation process described in the Simulation Process section. The approach used

in other applications of simulated annealing is to specify these two parameters in the annealing schedule. In that approach, prior information of the value of these parameters can be obtained experimentally; However, in the proposed approach, these simulation parameters are calculated internally. Therefore, the methods discussed in this section are superior, because prior information or experimentation is not required.

Initial Control Parameter

The performance efficiency of simulated annealing is sensitive to the value of the initial control parameter. The effects of the control parameter can be deduced from the Metropolis condition given by Equation 2.27. If the magnitude of the initial control parameter is excessively high, then most of the iterations proposed by the interchange mechanism will be accepted for the first several steps. During these steps, the distribution of the simulation variable remains essentially uncorrelated because interchanges resulting in either increase or decrease of the objective function have about the same probability to be accepted. Hence, these steps will be unproductive since the initial distribution of the simulated variable, as defined in the method developed in this chapter, is already uncorrelated. If the magnitude of the initial control parameter is too small, then most of the proposed iterations resulting in an increase of the objective function will be rejected. In this case, there is the possibility that the simulated distribution will become trapped in a local minimum rather than a global minimum, as it is the case in other algorithms such as the iterative improvement or greedy algorithm mentioned in the Background section.

A method of calculating the value of the control parameter at initial conditions has been derived by Aarts and Korst.[26] The procedure for calculating the initial control parameter consists of generating $M_{initial}$ iterations cycles with the interchange mechanism given by Equations 2.19 and 2.20 and for each iteration, calculate the objective function, $E^k_{initial}$, using Equation 2.26 for E'. One iteration cycle is equal to the number of iterations equal to the number of grid points (N_s) and the reasons to used this unit is discussed in the Annealing Schedule section.

The corresponding change in the objective function, $\Delta E^k_{initial}$, for each iteration is given by

$$\Delta E^k_{initial,} = E^k_{initial} - E^0, \tag{2.40}$$

for $k = 1, \ldots, M_{initial}N_s$, where, E^0 is the initial objective function. Note that $E^0 = 1$ according to the definition given in Equation 2.17. From the $M_{initial}N_s$ iterations, there is a number of m_1 iterations with $\Delta E^k_{initial} \leq 0$ and a number of m_2 iterations with $\Delta E^k_{initial} > 0$. Denoting the m_2 positive changes in the objective function as $\Delta E^+_{initial}$, the mean of the positive changes in the objective function, $\overline{\Delta E}^+_{initial}$, is calculated as

$$\overline{\Delta E}^+_{initial} = \frac{1}{m_2} \sum_{i=1}^{m_2} \Delta E^+_{initial}. \tag{2.41}$$

The initial control parameter, T^0, is given by[26]

$$T^0 = \frac{\overline{\Delta E}^+_{initial}}{\ln \left[\frac{m_2}{\chi_{initial} m_2 - (1-\chi_{initial}) m_1} \right]}, \tag{2.42}$$

where, $\chi_{initial}$ is the initial acceptance ratio.

Even though, the above procedure introduces two new parameters ($M_{initial}$ and $\chi_{initial}$), these parameters can be readily specified or set equal to a constant value for any simulation. The initial acceptance ratio should a value near one (e.g., 0.98 or 0.99) but not equal to one in order to allow an acceptance probability near one only during the first step. The number of iterations should be sufficiently large to allow for a stable estimate of the mean change of the objective function in Equation 2.41. This requirement can be satisfied by comparing the estimates from Equation 2.41 over a successive number of iterations and stopping the procedure when the estimate stabilizes. An example of the effect of the number of iterations on the estimates of the initial control parameter for four realizations of a conditional simulation is illustrated in Figure 2.4 for a wide range of $M_{initial}$ cycles and $\chi_{initial}$ =0.99. The specifications of these conditional simulations (Table 2.2) are discussed in more details in the Reproduction of Statistics section. The results in Figure 2.4 show

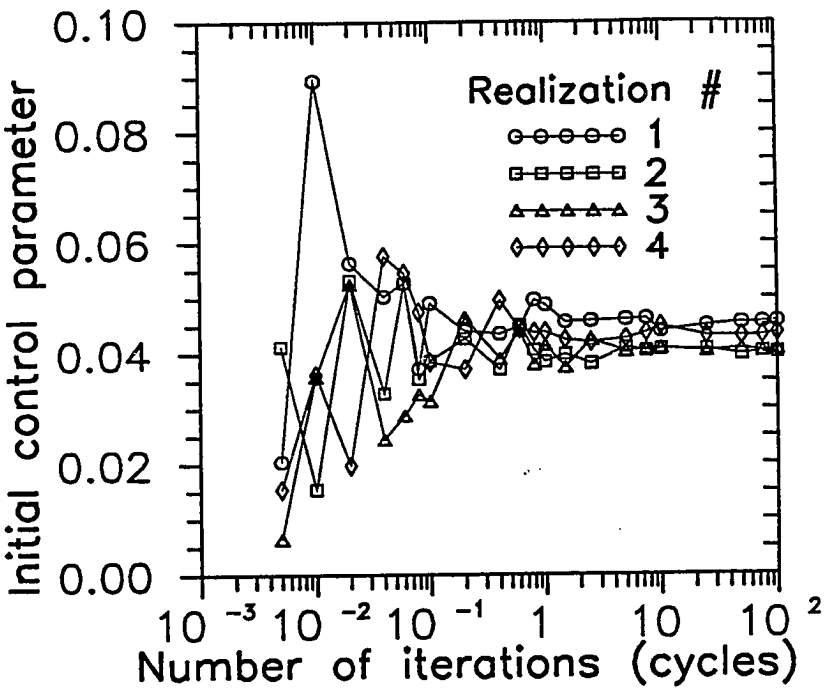

Figure 2.4 - Estimated initial control parameter (T^0) for different number of iterations ($M_{initial}$) and realizations

that the initial control parameter estimated by Equation 2.42 attains a stable value for $M_0 > 0.20$ cycles.

Maximum Number of Total Iterations

The maximum number of total iterations provides a practical upper limit for the number of iterations conducted in one step of the simulation procedure described in the Simulation Process section. Convergence or equilibrium conditions are achieved in a step when the number of iterations accepted by the Metropolis condition reaches the maximum number of accepted iterations specified in the annealing schedule described in the Annealing Schedule section. Therefore, the actual number of iterations conducted in each step is between the maximum number of accepted iterations and the maximum number of total iterations. For the first few steps, the actual number of iterations is close to the maximum number of accepted iterations and for the last few steps, the actual number of iterations approaches the maximum total number of iterations.

The method developed to calculate the maximum number of total iterations per step consists of predicting an estimate of the acceptance ratio for the subsequent step. For several simulations with a large constant value of the maximum number of total iterations per step, different relations between the acceptance ratio and the control parameter were evaluated. Results of these evaluations indicated that a practical method to predict the acceptance ratio is to perform a logarithmic extrapolation of the acceptance ratio from the two previous steps. This logarithmic extrapolation provides the following relation

$$\frac{\ln(\chi^{r+1}/\chi^{r})}{\ln(T^{r+1}/T^{r})} = \frac{\ln(\chi^{r}/\chi^{r-1})}{\ln(T^{r}/T^{r-1})}, \tag{2.43}$$

where, $r-1$, r and $r+1$ are the previous, the current and the next steps, respectively. From Equation 2.29, it can be noted that $T^{r+1}/T^{r} = T^{r}/T^{r-1} = \alpha$, then Equation 2.43 becomes

$$\chi^{r+1} = \frac{(\chi^{r})^2}{\chi^{r-1}}, \tag{2.44}$$

for $r = 1, 2, \ldots$. For $r = 1$ in Equation 2.44, $\chi^0 = 1$. The relation between the number of total iterations and the number of accepted iterations for any step r is

$$M_t^{r+1} = \frac{M_a}{\chi^{r+1}}. \tag{2.45}$$

Inserting the predicted acceptance from Equation 2.44 into Equation 2.45 yields

$$M_t^{r+1} = M_a \frac{\chi^{r-1}}{(\chi^r)^2}, \tag{2.46}$$

for $r = 1, 2, \ldots$. For the first step M_t is set equal to M_a as indicated in Step (4) of the Simulation Process section.

Numerical experiments were performed to compare the simulations using Equation 2.46 and simulations where M_t^r=40 $\times M_a$= 300 cycles, for $r = 1, 2, \ldots$. The reason to set the maximum number of total iterations equal to a large constant value is to ensure that the specified maximum number of accepted iterations is reached at each step. The simulation specifications for these experiments (Table 2.2) are discussed in more detail in the Reproduction of Statistics section. The simulation experiments reported in this section were conducted with three values for the convergence rate factor, equal to $\alpha = 0.1$, 0.5 and 0.9. For all simulations the maximum number of accepted iterations per step is $M_a = 7.5$ cycles.

Results of the numerical experiments indicate a close agreement between the simulations performed with the method given by Equation 2.46 and with the maximum number of total iterations per step equal to a large constant value. These simulations reached the specified constant value of $M_t = 300$ cycles only in the last simulation step. The values of the acceptance ratio and the objective function throughout conditional simulations using the method given by Equation 2.46 are compared to the simulations for $M_t = 300$ cycles in Figures 2.5 through 2.7 for $\alpha =$ 0.10, 0.50 and 0.90. These results indicate a close agreement between the simulation results of the two methods. For the small convergence rate factor ($\alpha = 0.10$), the slight difference in the results arises because the predictions of the acceptance ratio using Equation 2.46 are made over large increments of the control parameter.

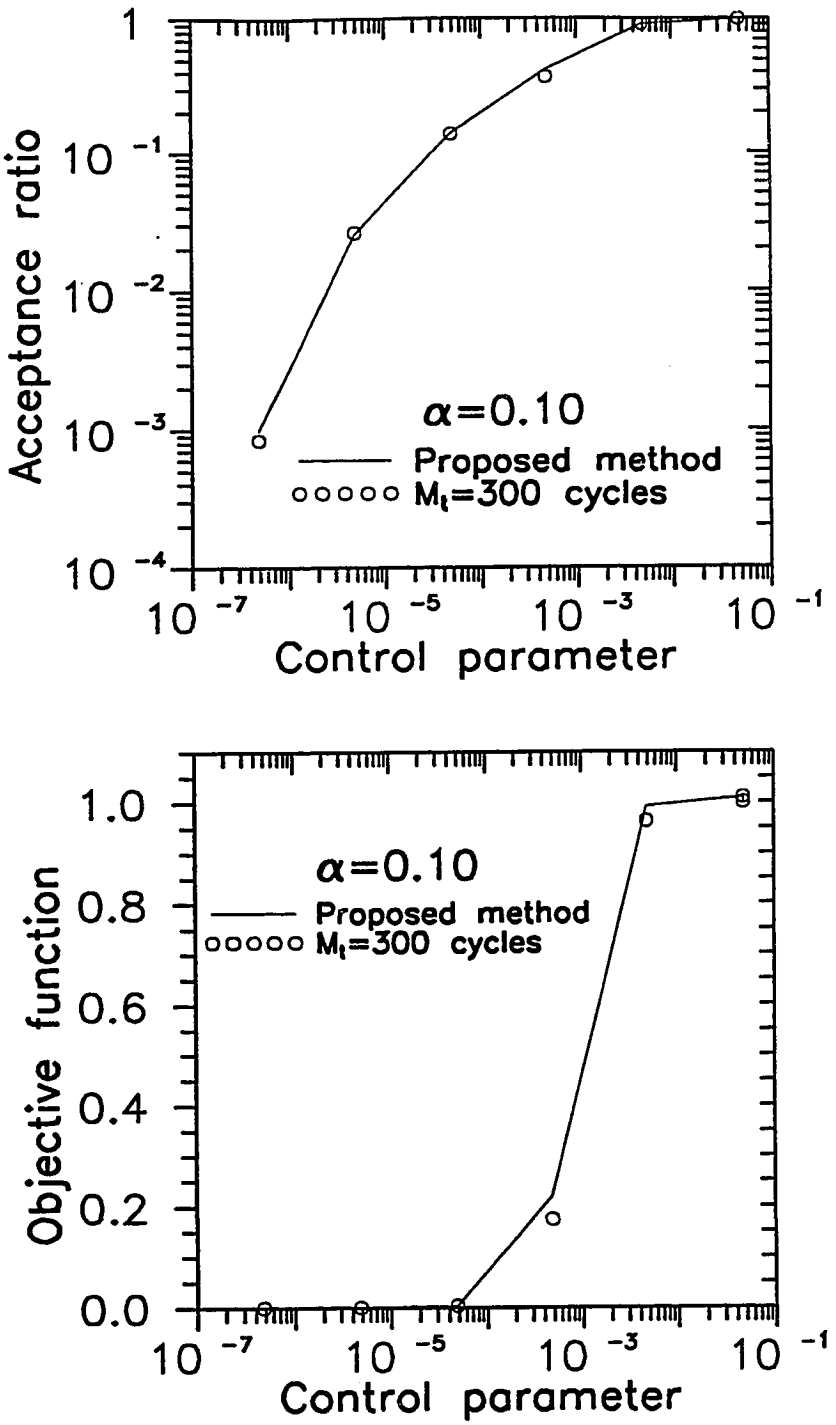

Figure 2.5 - Acceptance ratio and objective function throughout a conditional simulations with M_t calculated by the proposed method (Equation 2.46) and M_t=300 cycles for α=0.10

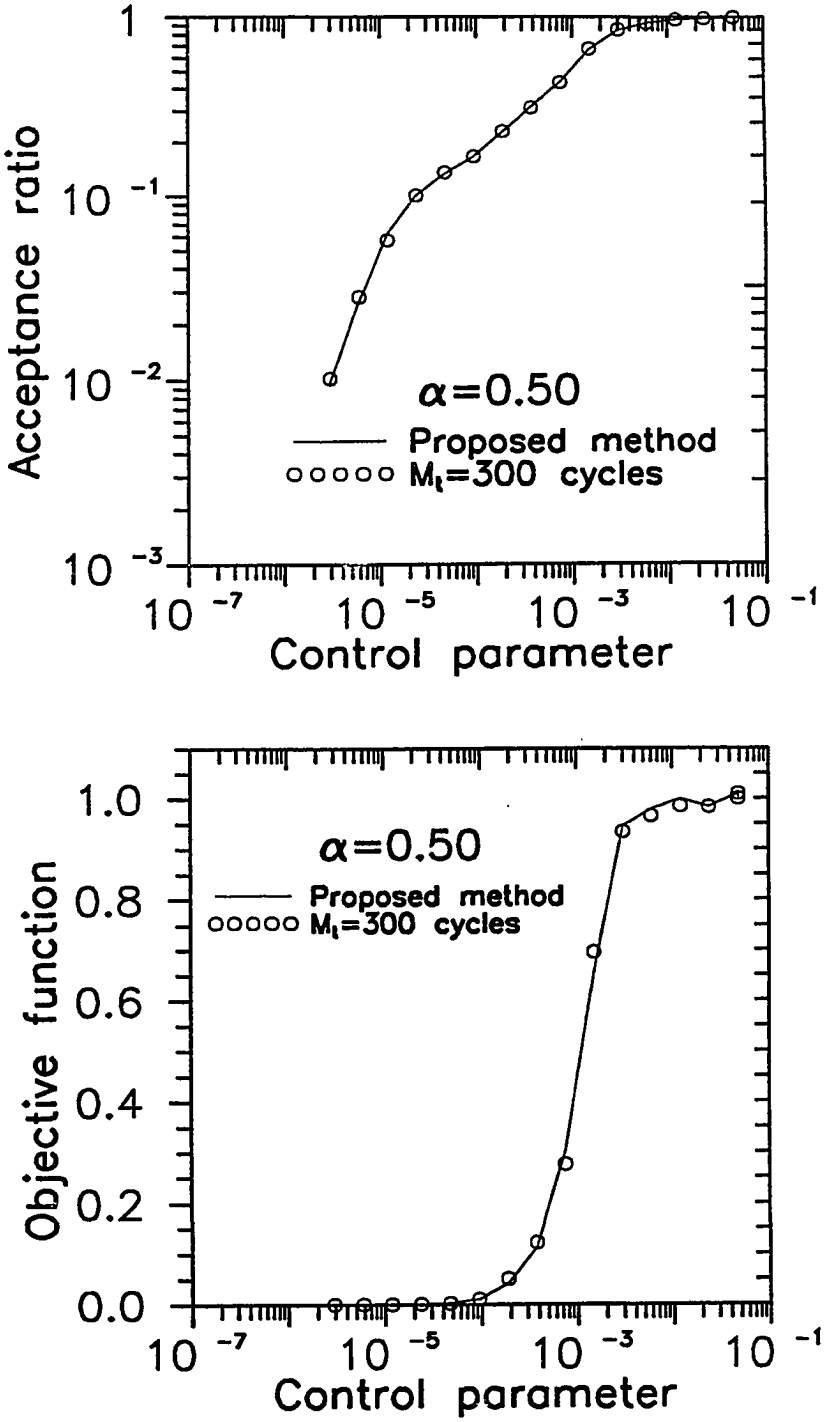

Figure 2.6 - Acceptance ratio and objective function throughout a conditional simulations with M_t calculated by the proposed method (Equation 2.46) and M_t=300 cycles for α=0.50

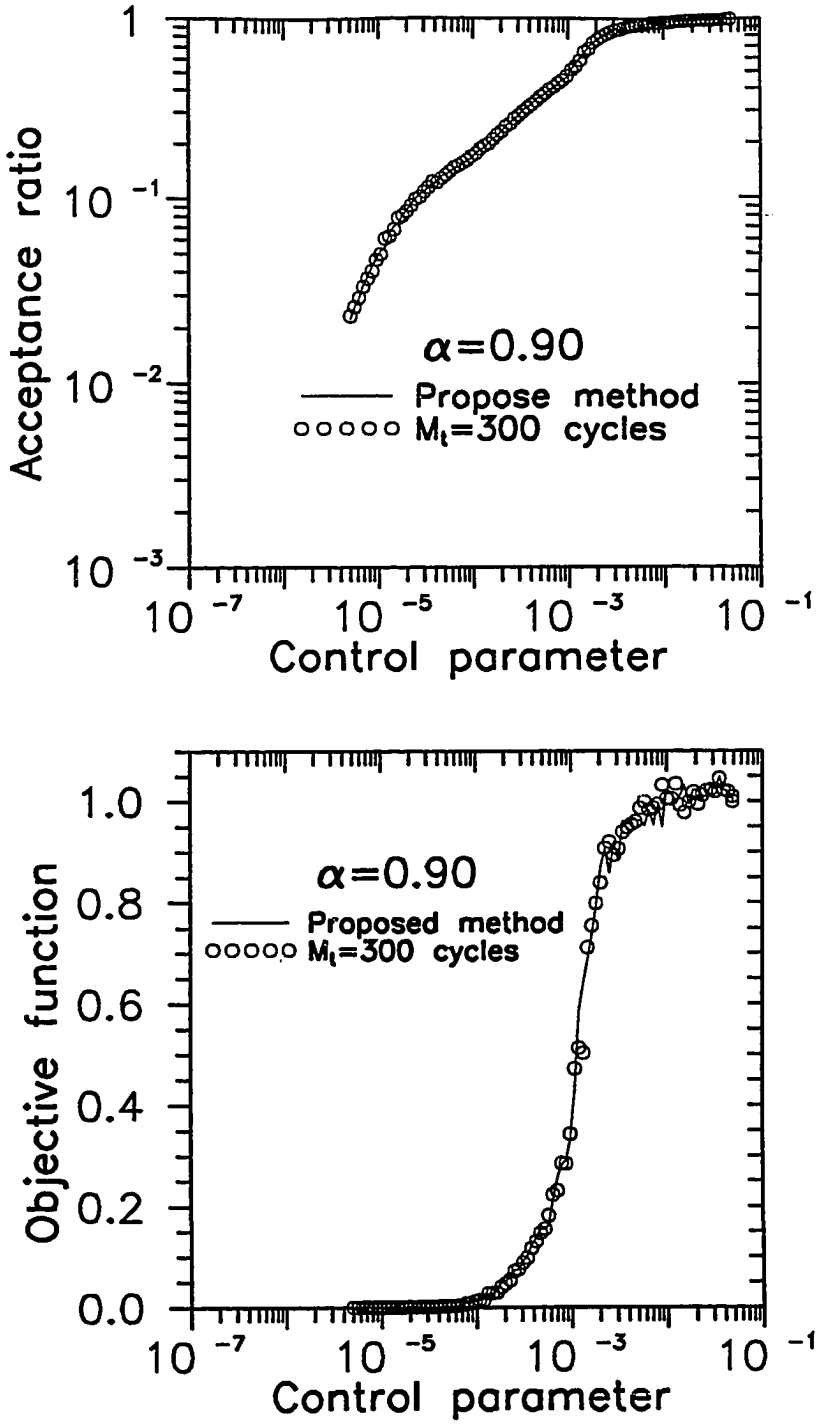

Figure 2.7 - Acceptance ratio and objective function throughout a conditional simulations with M_t calculated by the proposed method (Equation 2.46) and M_t=300 cycles for α=0.90

Annealing Schedule

The annealing schedule refers to the set of parameters that must be specified in the simulated annealing method. In the method developed in this chapter, the four parameters of the annealing schedule are the convergence rate factor (α), the maximum number of accepted iterations per step (M_a), the tolerance for the solution (ϵ_v) and the tolerance for the acceptance ratio (ϵ_a). The objective of this section is to evaluate the performance of conditional simulations for different annealing schedules and to derive guidelines to select the parameters of the annealing schedule needed for efficient simulations.

The convergence rate factor controls the magnitude of the decrements of the control parameter from step to step in the simulation process. As indicated by Equation 2.29, small convergence rate factors, close to zero, yield large decrements of the control parameter, while large convergence rate factors, close to one, yield small decrements of the control parameter. The maximum number of accepted iterations per step used in Step (7,i,b) in the Simulation Process section is the criterion to assume equilibrium conditions of the objective function in each step. The tolerance for the solution stops the simulation when the attributes of the simulated variable are close to the set of specified properties. The tolerance for the acceptance ratio stops the simulation when changes in the objective function are insignificant due to only few accepted iterations during a step.

The convergence rate factor is the only independent parameter of the system and in principle it should be the only parameter required in the annealing schedule. However, in most implementations of simulated annealing, including the method described in this chapter, a robust procedure is not available to identify the onset of equilibrium of the objective function at each step of the simulation. The tolerance for the solution and for the acceptance ratio are practical stopping criteria which are required in most numerical methods.

The efficiency of a conditional simulation can be ranked by the quality of the solution and the computational effort. The combination of the four parameters of the annealing schedule which yield high quality solutions for the smallest computational effort is the optimal annealing schedule. Unfortunately, for the conditional

simulation problem considered in this chapter, it is difficult to measure the quality of the solutions because the problem inherently entangles uncertainties. Furthermore, a close convergence to all the constraints and specifications of the simulation is not sufficient to measure the quality of the solution. In fact, the evaluation of the results of several conditional simulations indicates that all of the constraints and specifications are closely reproduced for a wide range of values of the parameters in the annealing schedule but the computational effort for the simulations varies significantly for different annealing schedule parameters.

In this section, an approach is developed to measure the quality of the solutions from the conditional simulations. This approach consists of estimating a certainty index between a known master distribution of a variable and the corresponding simulated distributions. The certainty index is defined in this section and is used to measure the quality of the simulations. The master distribution is the known distribution of a variable and it is used to derive the constraints and specifications used in the simulations.

Master Distribution

The master distribution is a two-dimensional synthetic distribution of a variable. The value of the variable is known everywhere in the master distribution. The function of the master distribution is to use it as a standard to measure the quality of the results from conditional simulations designed to reproduce this master distribution. The advantage of using a master distribution is that uncertainties involved in the evaluation of the statistical properties of the variable can be eliminated. The exhaustive statistical properties, such as the distribution function and the semi-variograms, required as specifications in the simulation are available since the value of the variable is known at every grid point.

The master distribution has been generated with a sequential Gaussian simulation method.[34] The specifications used in this simulation method allowed to generate a master distribution with four regions of alternating low and high mean values. This type of distribution resembles a layered structure which is found in

several reservoir engineering applications. Also, this layered structure serves as a severe test for a conditional simulation method due to the nonstationarities resulting from the marked changes of mean values among the layers. The master distribution consists of 1300 grid points, and there are 65 grid points in the x or lateral direction and 20 grid points in the z or vertical direction. Other specifications used in the sequential Gaussian simulation are not discussed because only the sample statistics derived from the master distribution are required for the conditional simulation performed in this section. In fact, the origin of the master distribution is not important and other conditional simulation methods can be used to generate this distribution.

A gray scale map of the master distribution is shown in Figure 2.8. The white colors in the gray scale map correspond to the minimum value, the black colors correspond to the maximum value and different gray tones in between white and black are assigned by linear interpolation of the corresponding value. The univariate statistical parameters of the master distribution are summarized in Table 2.1. The histogram of the master distribution is shown in Figure 2.9 and it is close to a uniform distribution. Actually, this distribution is a mixture of a low mean distribution with the highest frequency near the value of 9.0 and a high mean distribution with the highest frequency near the value of 29.0. These low and high mean distributions correspond to the layers with a high and a low magnitude, but the overall distribution appears to be a uniform due to the gradual transition of values in between the distributions. The exhaustive cumulative distribution function of the master distribution is shown in Figure 2.10.

The exhaustive sample semi-variograms of the master distribution along four directions are shown by the circles in Figures 2.11 and 2.12 and the corresponding models fitted to these semi-variograms are shown by the solid lines. The parameters of the semi-variogram models are given in Table 2.2 for the simulation specifications used in Reproduction of Statistics section. The semi-variogram along the horizontal or x direction (Figure 2.11) is almost an order of magnitude smaller than the semi-variograms along the other directions. Also, the correlation range in the x direction

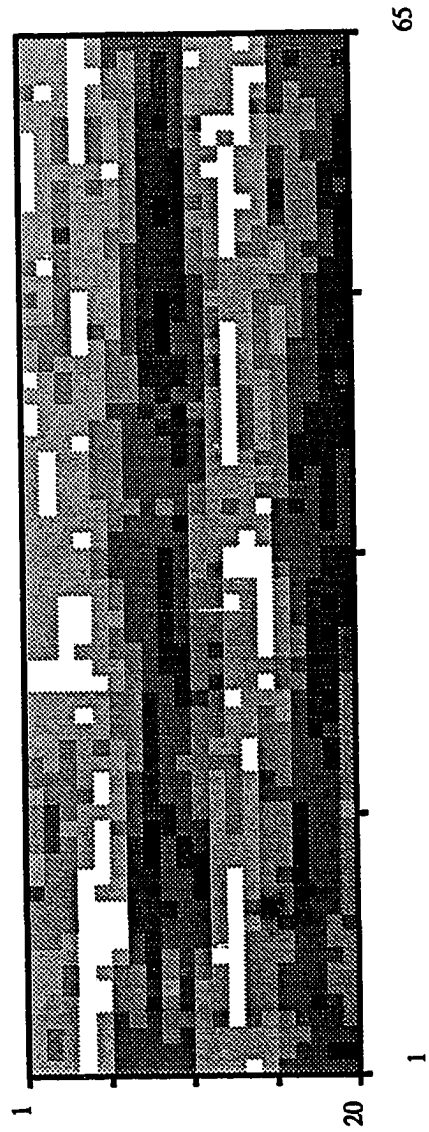

Figure 2.8 - Map of the master distribution

Table 2.1

Univariate statistics of the master distribution

Parameter	
Number	1300
Mean	16.3933
Variance	96.5527
Standard Deviation	9.8261
Coefficient of Variation	0.5994
Minimum	0.0
Maximum	40.2470
Median	15.7100
1^{st} quartile	8.0000
3^{rd} quartile	24.6030

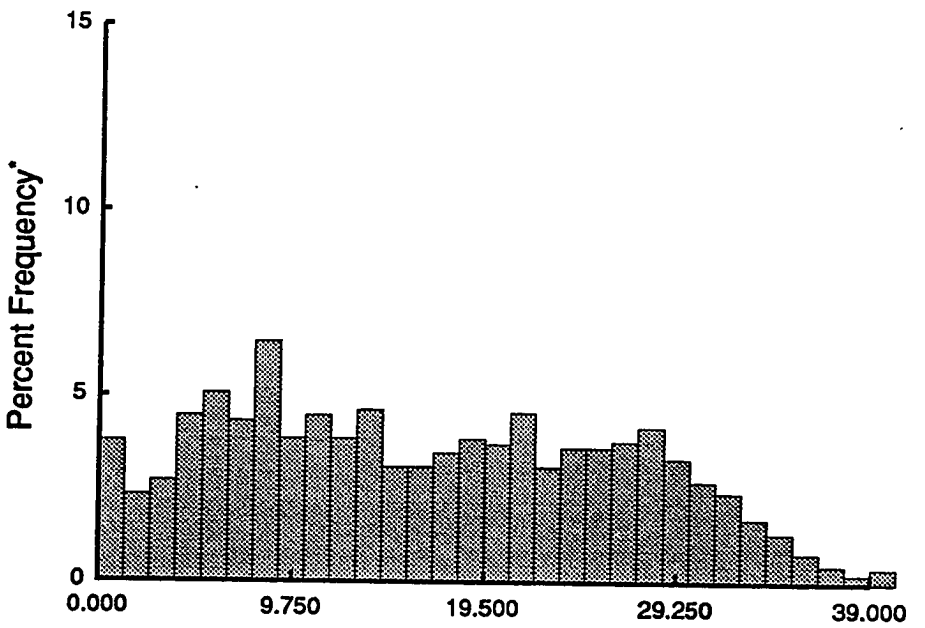

Figure 2.9 - Histogram of the master distribution

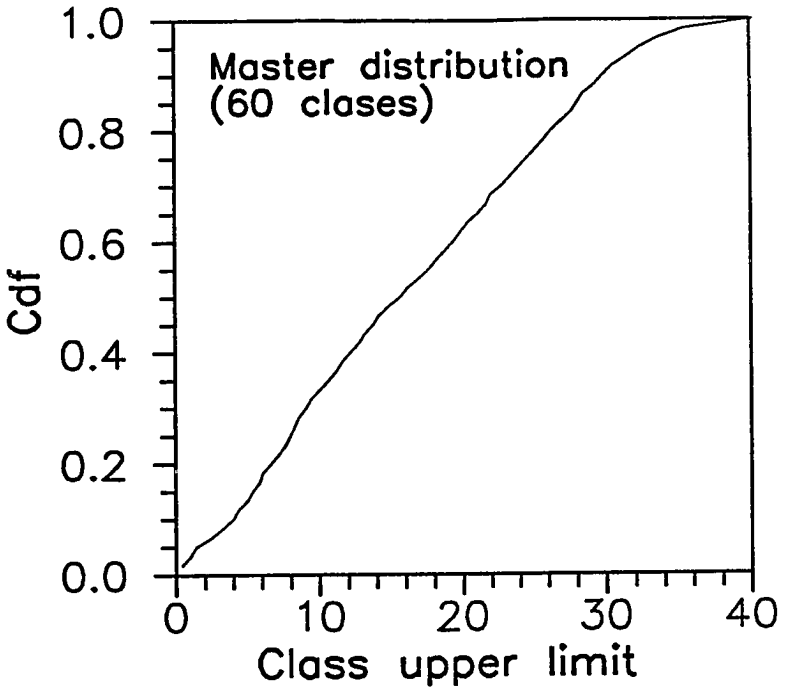

Figure 2.10 - Exhaustive cumulative distribution function (Cdf) of the master distribution

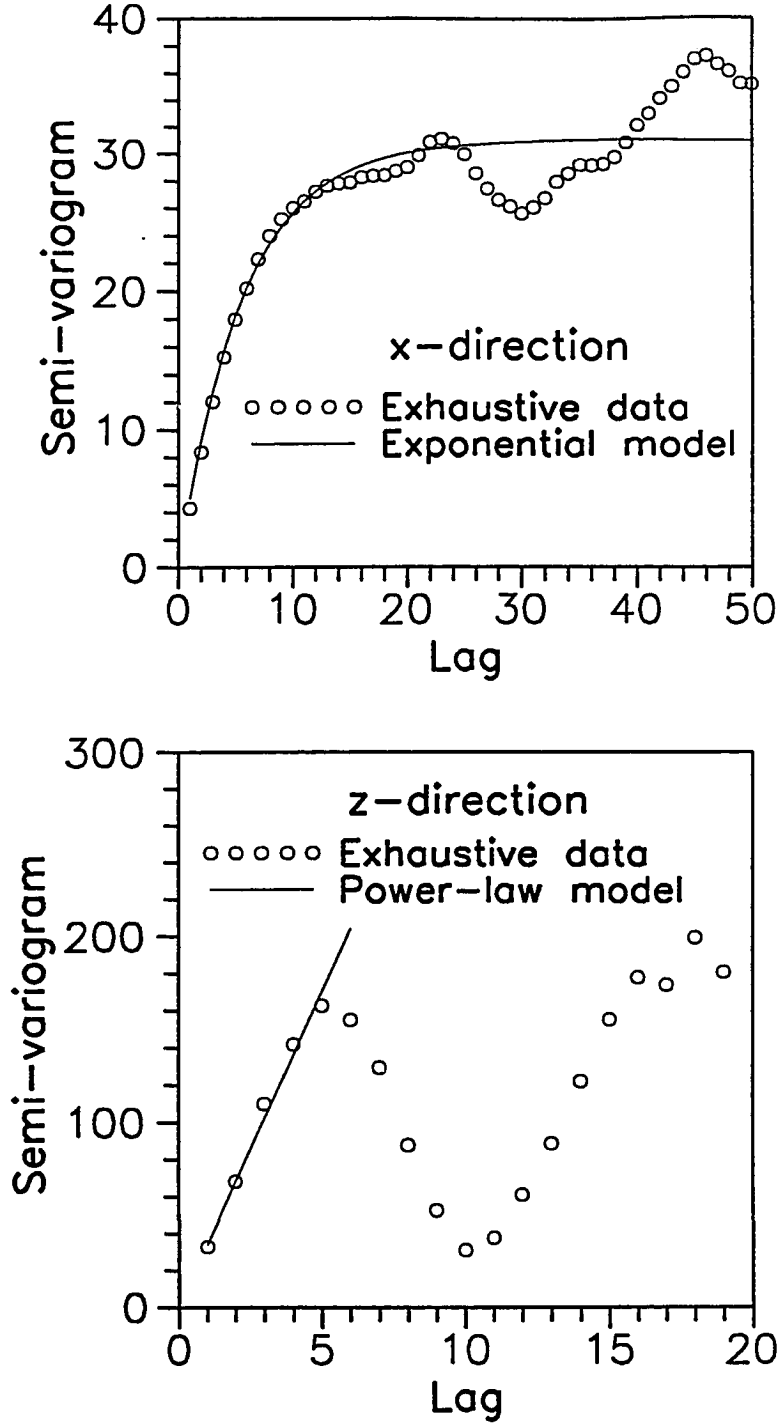

Figure 2.11 - Exhaustive semi-variograms of the master distribution and models for the x and z directions

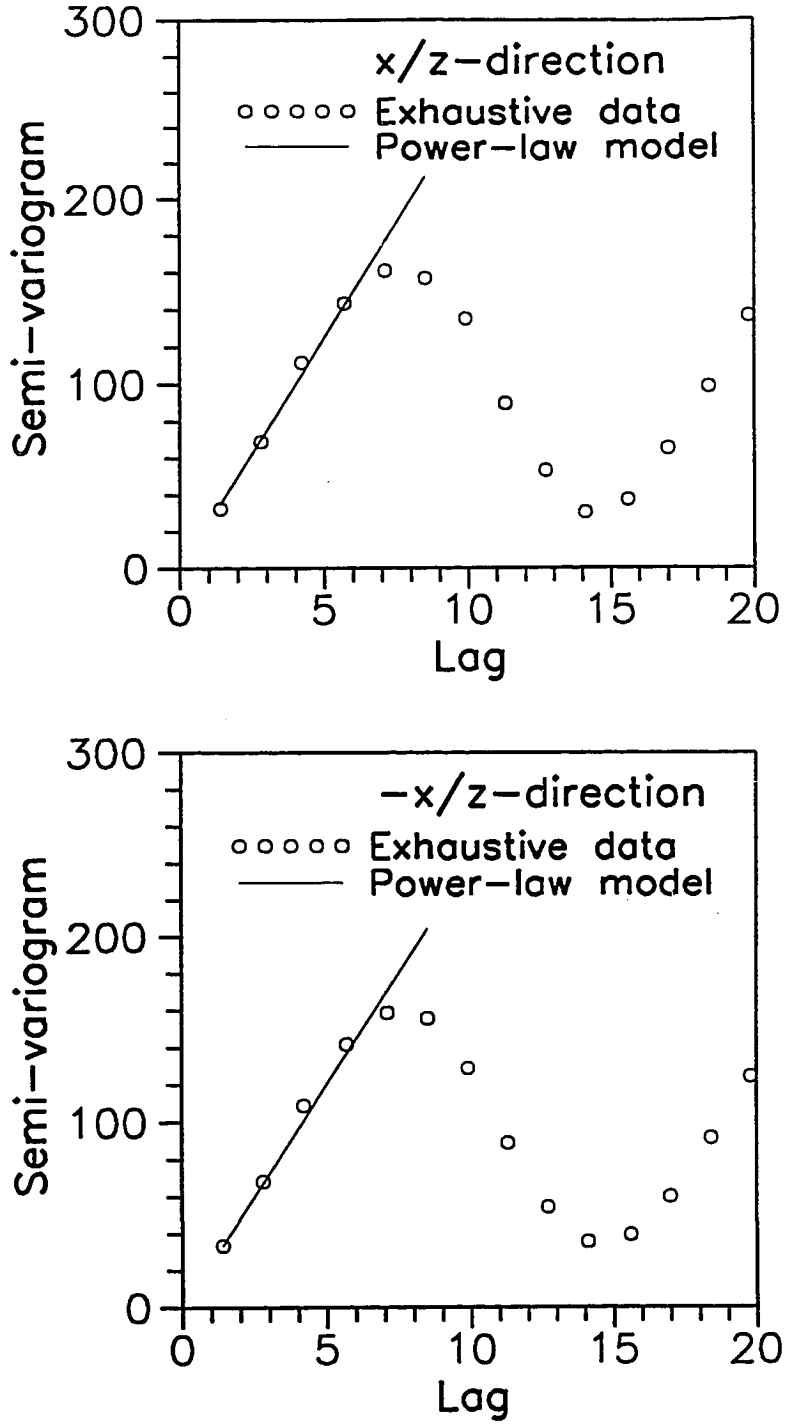

Figure 2.12 - Exhaustive semi-variograms of the master distribution and models for the x/z and -x/z directions

Table 2.2

Base specifications for conditional simulations

of the master distribution

Grid Geometry			
Direction	x	y	z
Number	65	1	20
Spacing	1.0	1.0	1.0
Conditioning Data			
Number		60	
Distribution Function			
Number of classes		60	
Subclass distribution		Uniform	
Semi-variogram Models			
x direction		$\gamma(h) = 31.0\,[1 - \exp(-3h/17.0)]$	
z direction		$\gamma(h) = 34.0\,h, h \leq 5$	
x/z direction		$\gamma(h) = 25.0\,h, h \leq 7$	
-x/z direction		$\gamma(h) = 24.0\,h, h \leq 7$	

is longer than the correlation ranges in the other directions. These differences of the semi-variograms are due to the small variability and greater continuity as a result of the layered structure of the master distribution. The hole effects[6] in the semi-variogram along the vertical or z direction (Figure 2.11), and for the 45° or x/z and -45° or -x/z directions (Figure 2.12) are also due to the layered structure of the master distribution. For the z direction the semi-variogram reaches a minimum value at a lag equal to 10 units which corresponds to two times the approximate thickness of each layer. In the x/z and -x/z directions, the layers appear to be thicker and the distance at which the semi-variograms reach a minimum value increases to a lag equal to 14 units. However, the observed hole effects will be ignored and only the models shown in Figures 2.11 and 2.12 up to a maximum lag equal to the approximate layer thickness will be used for the z, x/z and -x/z directions in the conditional simulations.

Reproduction of Statistics

Several numerical experiments were conducted to investigate the effect of the annealing schedule parameters on the reproduction of the specified statistics and on the computational performance of simulated annealing. The specifications for the conditional simulations were derived from the master distribution.

The simulation region for the conditional simulations consists of 1300 grid points and it has the same geometric configuration as the master distribution. The conditioning data, cumulative distribution function, and semi-variogram models for all the conditional simulations described in this section are the same and are summarized in Table 2.2. The conditioning data consists of 60 values of the master distribution. These conditioning data are located in three sections along the z direction corresponding to the first, the middle and the last x locations. The cumulative distribution function and the semi-variogram models are shown in Figures 2.10 through 2.12.

The parameters of the annealing schedule used in the numerical experiments consisted of convergence rate factors (α) between 0.10 and 0.90 and maximum

number of accepted iterations per step (M_a) between 0.50 and 30.0 cycles. All the iteration parameters used in this section and the following sections are given in units of cycles. One cycle equals to the number of iterations equal to the number of grid points of the simulation region. The cycle is a convenient unit because the number of combinations of the simulated variable or the solution space is related to the number of the simulated grid points. For most conditional simulations, the solution tolerance (ϵ_v) is equal to 1×10^{-4} and the acceptance ratio tolerance (ϵ_a) is equal to 2.5×10^{-2}. These tolerances were varied in a few conditional simulations described later.

The initial distribution of the variable for all the conditional simulations described in this section is the same (i.e., same seed used to initialized the random number generator). The sample semi-variograms of the initial distribution for the four directions considered in the simulations are shown in Figures 2.13 and 2.14 and indicate that the initial distribution is uncorrelated. All the simulations, for different annealing schedule parameters, were capable of reproducing the specified statistical properties (Table 2.2). Comparisons between the simulated and the specified statistical attributes for the simulation case with α =0.50 and M_a = 5.0 cycles are shown in Figures 2.15 through 2.17. For this case, the agreement is close between the specified and simulated cumulative distribution function (Figure 2.15) and semi-variograms (Figures 2.16 and 2.17). For the simulation cases with other annealing schedules, the agreement between the specified and simulated statistical attributes is as close as for the case compared in Figures 2.15 through 2.17. However, the computational effort (total number of iterations throughout the simulation) varies significantly for the different cases. For example, the computational effort required for the case with α = 0.50 and M_a = 5.0 cycles is about ten times greater than for case with α = 0.10 and M_a = 0.50 cycles. For these and other cases discussed later, it is observed that the reproduction of the specified statistical properties appears to be independent of the annealing schedule parameters.

In addition to the simulations with different annealing schedules, a conditional simulation was conducted using the greedy algorithm for the specifications shown

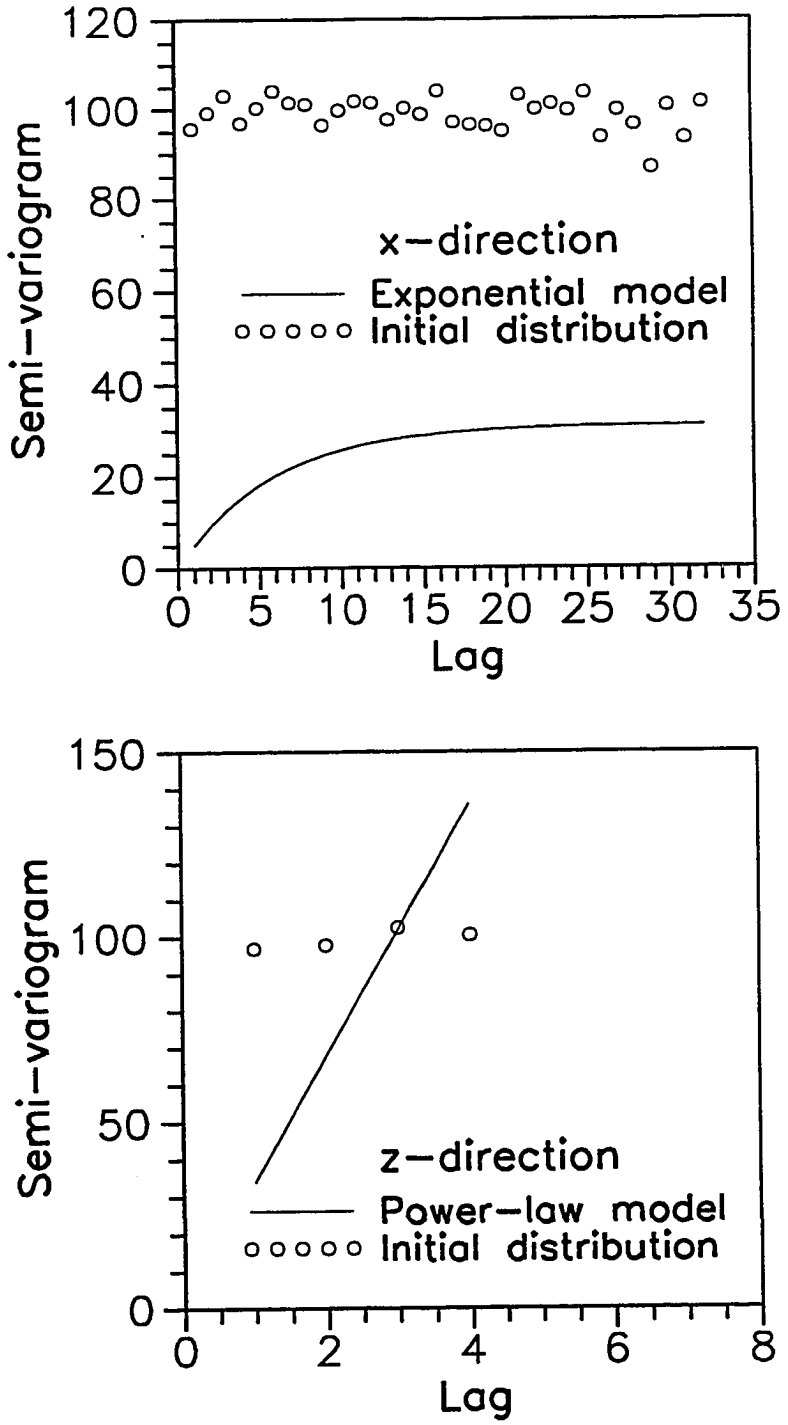

Figure 2.13 - Semi-variograms for the initial distribution of the conditional simulations for the x and z directions

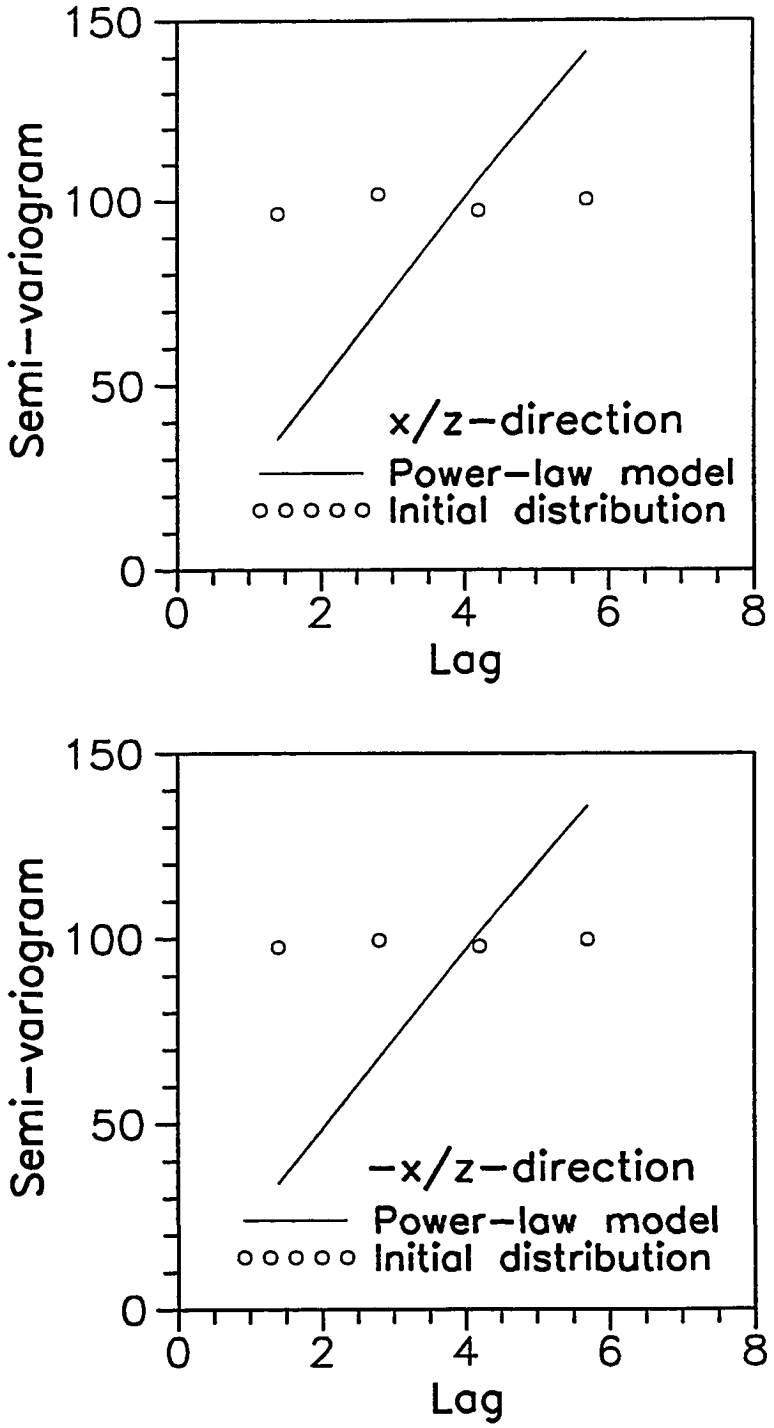

Figure 2.14 - Semi-variograms for the initial distribution of the conditional simulations for the x/z and -x/z directions

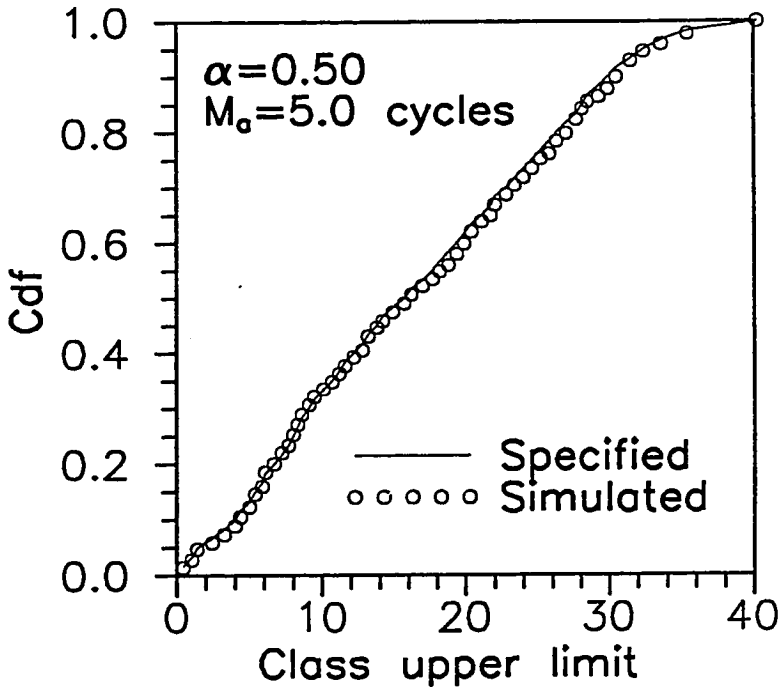

Figure 2.15 - Specified and simulated cumulative distribution function (Cdf)

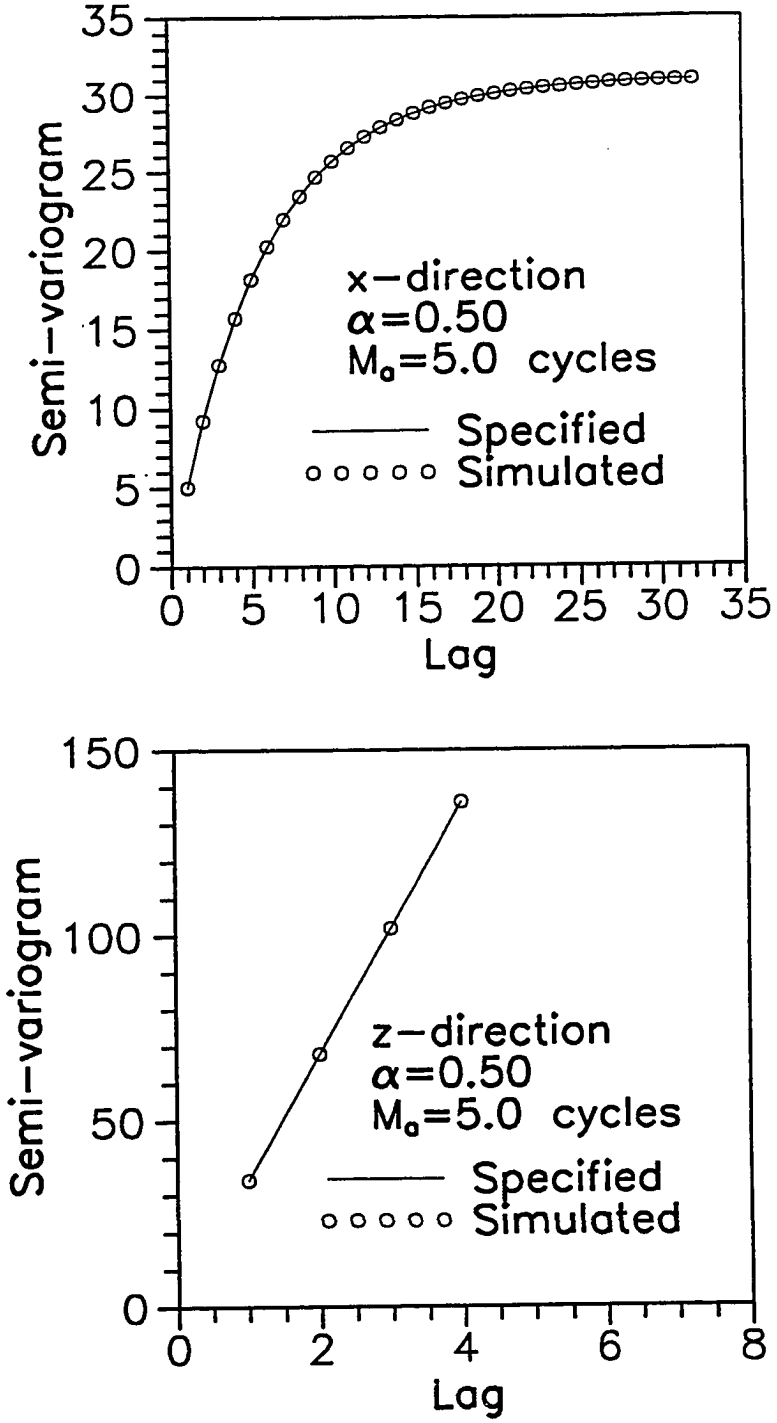

Figure 2.16 - Specified and simulated semi-variograms for the x and z directions

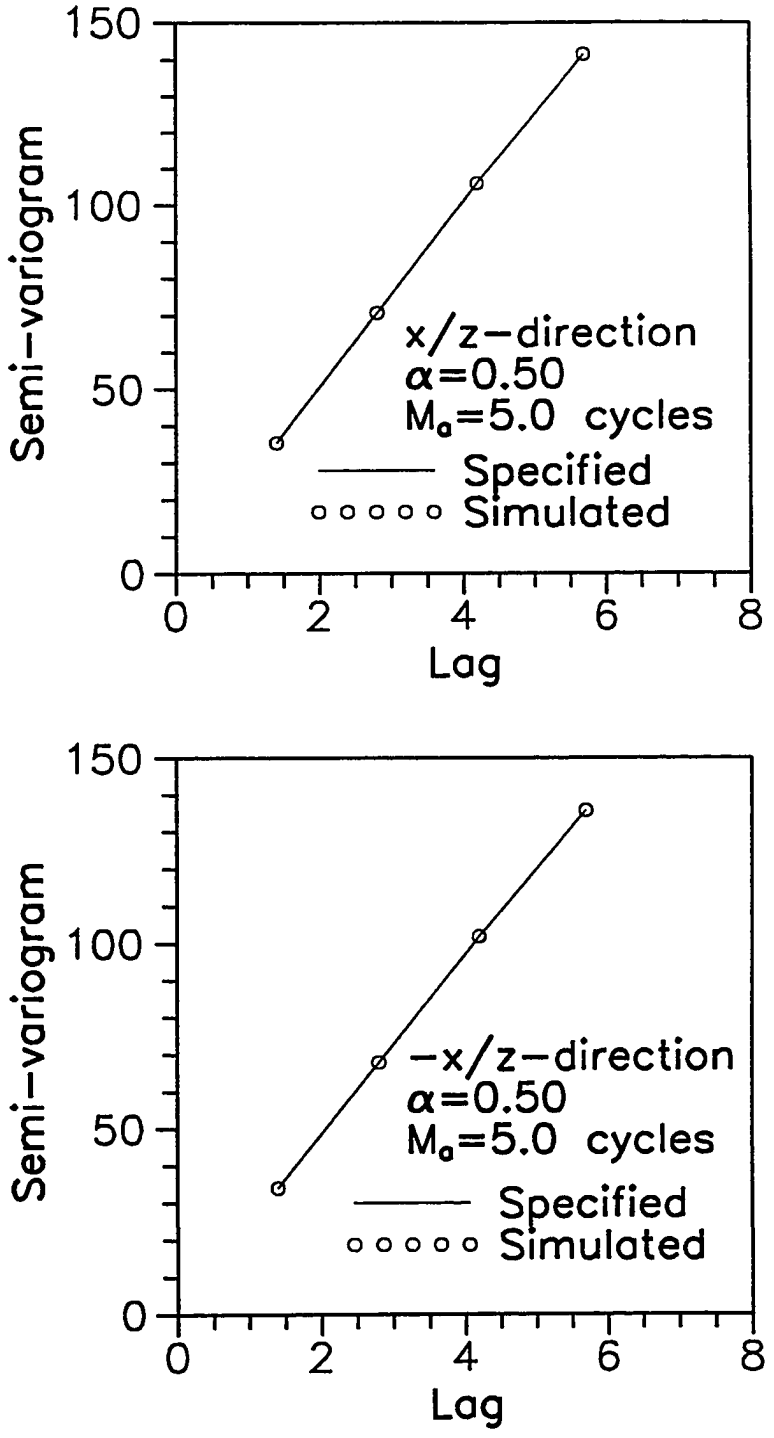

Figure 2.17 - Specified and simulated semi-variograms for the x/z and -x/z directions

in Table 2.2. The greedy algorithm is a limiting case of simulated annealing because it does not use the Metropolis condition and, therefore, it does not require an annealing schedule. Instead, it accepts only the iterations which reduce the magnitude of the objective function. The greedy algorithm corresponds to the limit when the control parameter approaches zero and the exponential probability function of the Metropolis condition in Equation 2.27 becomes equal to zero. Therefore, the convergence rate factor and the maximum number of accepted iterations per step are meaningless in the greedy algorithm. For the conditional simulation conducted with the greedy algorithm, the specified and the simulated statistical attributes are in close agreement and the comparison between these attributes is similar to the simulation case shown in Figures 2.15 through 2.17. The computational effort for the greedy algorithm (50.3 cycles) is smaller than for the simulated annealing conditional simulations. The computational cost of simulated annealing approaches the greedy algorithm computational cost as α and M_a tend to zero.

Optimum Schedule

The optimum schedule is defined as the set of parameters of the annealing schedule which yield the lowest computational cost simulations without sacrificing the quality of the solution. An optimum schedule will be needed for practical application of conditional simulations consisting of a large number of grid points, potentially in the order of millions of grid points.

The computational cost of a conditional simulation is measured by the total number of iterations required to achieve the final solution. The computational cost of a simulation increases as the convergence rate factor increases and as the maximum number of accepted iterations per step increases. The computational costs of the conditional simulations of the master distribution described in this section are shown in Figure 2.18. These results indicate that for a constant convergence rate factor, the total number of iterations of a simulation increases as a linear function of the maximum number of accepted iterations per step. The results of Figure 2.18 are redrawn in Figure 2.19 using a semi-logarithmic scale to illustrate that for a

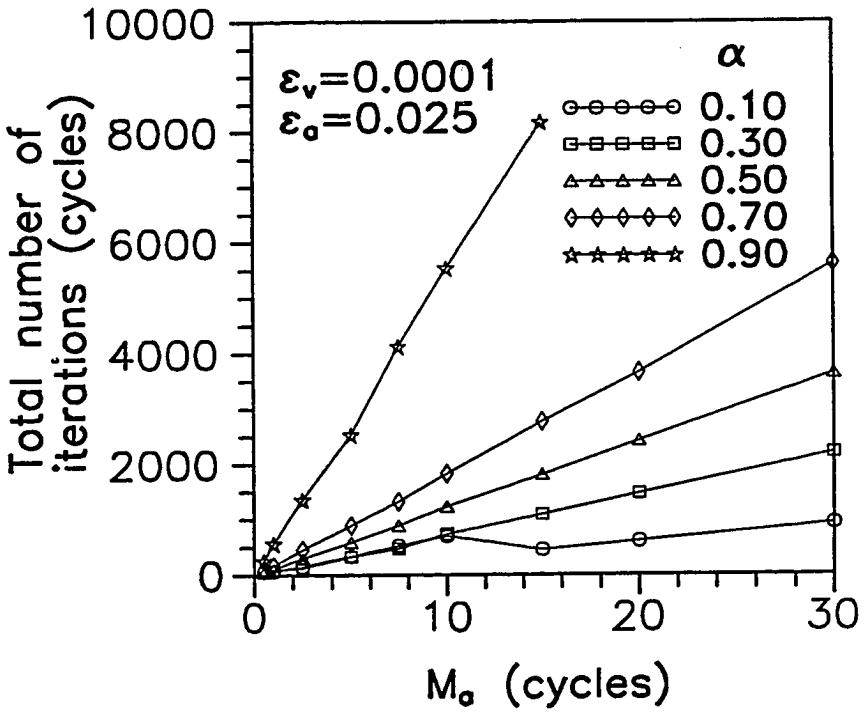

Figure 2.18 - Computational cost of conditional simulations as a function of number of accepted iterations per step (M_a) for different convergence rate factors (α)

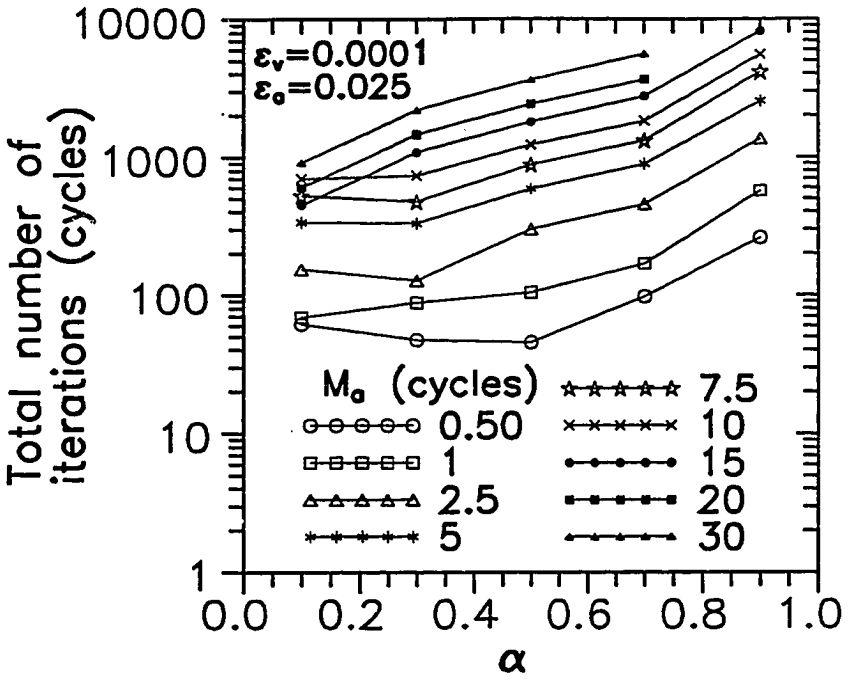

Figure 2.19 - Computational cost of conditional simulations as a function of convergence rate factor (α) for different numbers of accepted iterations per step (M_a)

constant number of maximum accepted iterations per step, the total number of iterations in a simulation increases as an exponential function of the convergence rate factor.

Results of the numerical experiments indicate that the conditional simulations for different values of the parameters of the annealing schedule reproduced closely the specified statistical properties. Based on these results, it appears that reproduction of the specified statistical properties is not enough to measure the quality of a conditional simulation. Otherwise, the greedy algorithm can be judged to be the best algorithm. The problem, however, with this conclusion is the inability to discern the local from the global minimum simply based on matching the specified statistical attributes. In order to distinguish these minima, it is necessary to evaluate other fundamental differences which must exist among the simulations for different parameters. One approach to measure the quality of the simulation can be to visually compare the simulated distribution to the master distribution using, for example, gray or color scale maps. However, this approach is too subjective. Since the problem is inherently stochastic, a better approach to estimate the quality of a simulation is to use a coefficient to measure the certainty of the simulated distribution.

The certainty coefficient, C, between a simulated variable, V_s, and the true value, V_t, is defined using the linear correlation coefficient[32] and it is given by

$$C = \frac{\sum_{i=1}^{N_s}[V_s(\vec{\omega}_i) - \overline{V}_s][V_t(\vec{\omega}_i) - \overline{V}_t]}{\sqrt{\sum_{i=1}^{N_s}[V_s(\vec{\omega}_i) - \overline{V}_s]^2 \sum_{i=1}^{N_s}[V_t(\vec{\omega}_i) - \overline{V}_t]^2}}, \tag{2.47}$$

where,

$$\overline{V}_s = \frac{1}{N_s}\sum_{i=1}^{N_s} V_s(\vec{\omega}_i) \tag{2.48}$$

and

$$\overline{V}_t = \frac{1}{N_s}\sum_{i=1}^{N_s} V_t(\vec{\omega}_i). \tag{2.49}$$

The certainty coefficient varies between 1 and -1. The extreme values of the certainty coefficient indicate the following relations between the true and the simulated

distributions: (1) for $C = 1$ there is a perfect positive linear relation, (2) for $C = 0$ there is no correlation and (3) $C = -1$ there is a perfect negative correlation. Also, the certainty coefficient can be interpreted as the slope of the least-squares regression line between the standard scores transform of the true and simulated distributions. If the simulated distribution is a perfect prediction of the true distribution, then the certainty coefficient is equal to one. As the quality of this prediction degrades the certainty coefficient decreases and it becomes equal to zero when there is no correlation between the simulated distribution and the true distribution. The certainty coefficient becomes negative when the values predicted by the simulated distribution and the corresponding true values consistently have opposite magnitudes.

The certainty coefficients between the simulated distributions and the master distribution for the different annealing schedules and the greedy algorithm are shown in Figure 2.20. These results reveal important differences between simulated annealing and the greedy algorithm and among the simulations with different annealing schedule parameters. The certainty coefficient of the greedy algorithm distribution is equal to 0.3232 and it is significantly smaller than the lowest certainty coefficient in a simulated annealing distribution. The simulated distributions for $\alpha = 0.50$ and $M_a = 5.0$ cycles and for the greedy algorithm are shown in the gray scale maps in Figures 2.21 and 2.22, respectively. It is clear from these results that the distribution for the greedy algorithm (Figure 2.22) has a different appearance than the true master distribution (Figure 2.8), while the distribution for simulated annealing (Figure 2.21) closely resembles the master distribution.

For different annealing schedule parameters, the certainty coefficient decreases as the convergence rate factor and the maximum number of accepted iterations per step decrease. For small values of the convergence rate factor and the maximum number of accepted iterations per step, the large fluctuations of the certainty coefficient indicate that the number of iterations at each step is too small to allow the equilibrium of the objective function for the large steps corresponding to a small convergence rate factor. As the maximum number of accepted iterations per step

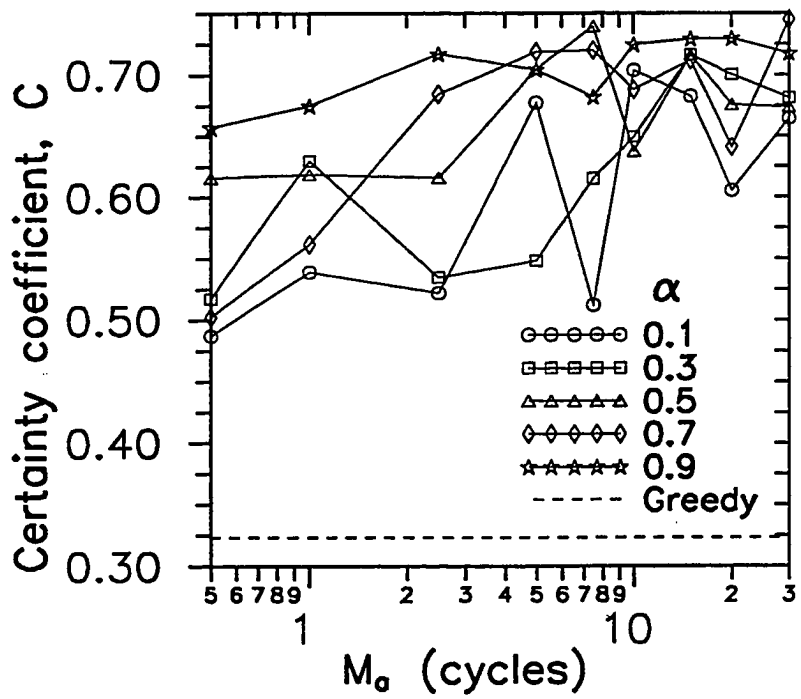

Figure 2.20 - Certainty coefficient of conditional simulations as a function of number of accepted iterations per step (M_a) for different convergence rate factors (α)

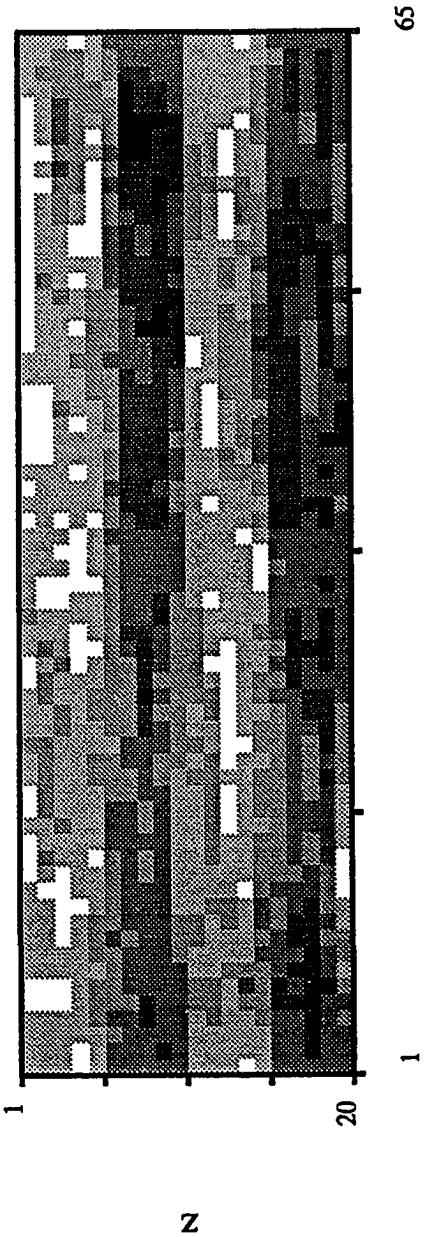

Figure 2.21 - Map of conditional simulation of the master distribution for α=0.50 and M_a=5.0 cycles

Figure 2.22 - Map of conditional simulation of the master distribution for the greedy algorithm

increases the certainty coefficient tends to reach a stable value and for the largest maximum number of accepted iterations per step shown in Figure 2.20, $M_a = 30.0$ cycles, the mean certainty coefficient for the different convergence rate factors is equal to 0.695.

The criterion of high quality solutions for the selection of the optimum parameters of the annealing schedule can be fulfilled by maximizing the certainty coefficient of the simulations. For each of the convergence rate factors considered in the numerical experiments, the optimum value of the maximum number of accepted iterations per step has been selected from the results in Figure 2.20 as the smallest value with a certainty coefficient equal to 0.70. These optimal annealing schedules are shown in Table 2.3 along with the corresponding total number of iterations in the simulation or computational cost. The optimum annealing schedules indicate that as the convergence rate factor increases the required maximum number of accepted iterations per step decreases, except for the case of $\alpha = 0.30$. As the convergence rate factor increases, Equation 2.29 indicates that the size of the control parameter steps decreases and it can be anticipated that a smaller number of iterations is required for the objective function to reach equilibrium at each step. The exception observed for the smaller convergence rate factors, $\alpha = 0.10$ and 0.30 may be due to the instabilities resulting from the large steps for these small convergence rates. Table 2.3 indicates that the optimum annealing schedule with the lowest computational cost corresponds to a convergence rate factor equal to 0.50.

The tolerance parameters in the annealing schedule for the numerical experiments reported in Figures 2.18 through 2.19 are fixed. The effect of the acceptance ratio tolerance on the computational cost is shown in Figure 2.23 for conditional simulations with $\alpha = 0.50$. As the acceptance ratio tolerance decreases, the computational cost increases. Further evaluation of these simulations indicates that when the acceptance ratio tolerance becomes greater than 0.10 (10 % of the iterations are accepted in the last step) the quality of the solution or certainty coefficient decreases. The solution tolerance used in the numerical experiments, $\epsilon_v = 1 \times 10^{-4}$, is significantly smaller than the tolerance required to get high quality solutions. In

Table 2.3

Optimum annealing schedule parameters for conditional simulations of the master distribution

Convergence Rate Factor, α	M_a for $C = 0.70$ (cycles)	Computation Cost (cycles)
0.10	10.0	697.482
0.30	13.8	1000.587
0.50	4.9	576.409
0.70	3.6	644.868
0.90	1.9	1039.115

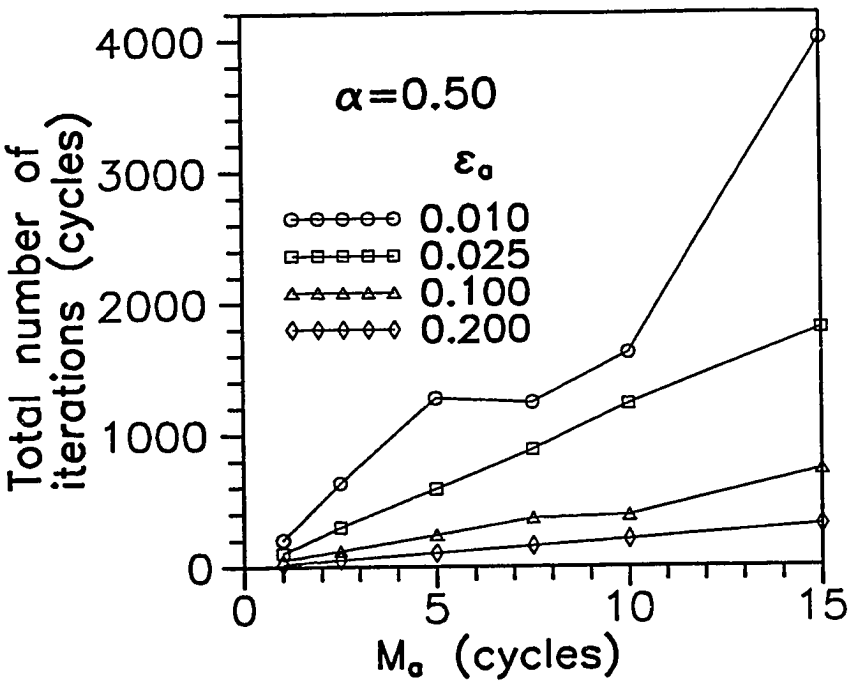

Figure 2.23 - Computation cost of conditional simulations as a function of number of accepted iterations per step (M_a) for different acceptance ratio tolerances (ϵ_a)

other simulations with a tolerance $\epsilon_v = 2\times10^{-3}$ the statistical specifications are closely reproduced and the magnitude of the certainty coefficients are close to the simulations with a tolerance $\epsilon_v = 1\times10^{-4}$. An approach to specify these tolerances of the annealing schedule is to set the acceptance ratio tolerance to a small value, such as $\epsilon_a = 2.5\times10^{-2}$, and to set the solution tolerance to a value which provides a close reproduction of the specified statistics, such as $\epsilon_v = 2\times10^{-3}$.

The analysis of the simulations discussed in this section indicates that the value of the parameters of an optimum annealing schedule are $\alpha = 0.50$, $M_a = 5.0$ cycles, $\epsilon_a = 2.5\times10^{-2}$ and $\epsilon_v = 2\times10^{-3}$.

Quantifying Uncertainty

The objective of this section is to quantify the uncertainty of simulated distributions for various types of constraints and specifications used in conditional simulations of the master distribution. In practice, the information required to accesses uncertainty is seldom available. Therefore, a major contribution of this section is to provide an order of magnitude of the uncertainty of the predicted distributions relative to the amount and quality of information available.

In the previous section, the statistical specifications were fixed and equal to the parameters derived from the exhaustive master distribution. In this section, the conditioning data, the distribution function and the semi-variogram models derived from the master distribution are systematically altered to observe the effects on the simulated distributions. Also, different equiprobable realizations of the simulated distributions are evaluated. The base specifications for the simulations are the same as the ones used in the previous section (Table 2.2). Except for the changes noted throughout the discussion, the remaining specifications of the simulations described in this section are the same as the base specifications. For all the conditional simulations discussed in this section, the parameters of the annealing schedule are $\alpha = 0.50$, $M_a = 5.0$ cycles, $\epsilon_a = 2.5\times10^{-2}$ and $\epsilon_v = 1\times10^{-4}$.

Equiprobable Realizations

Different equiprobable realizations were generated by varying the seed of the random number generator. For each seed, a different initial distribution for the simulations is generated. In the previous section, all the conditional simulations used the same seed and have the same initial distribution.

The univariate statistics of the certainty coefficient and the computational cost for ten realizations of the simulated distribution are given in Table 2.4. For the ten generated realizations, the certainty coefficient does not change significantly and indicates that the quality of different realizations of the solutions is similar. However, for the same realizations, the computational cost of the simulations varies more than the certainty coefficient and have range of about 200 cycles.

In an attempt to quantify the differences or variations among the realizations of the simulated distribution, certainty coefficients between the realizations were compared. The purpose of this comparison is to determine if the different realizations reproduce the same characteristics of the master distribution or if each realization reproduces characteristics of the master distribution different from other realizations. This implies that if two realizations have similar certainty coefficients with respect to the master distribution and the realizations are similar, then the certainty coefficient between the realizations can be greater than the individual certainty coefficients with respect to the master distribution. The approach to calculate the certainty coefficient between realizations consists of specifying one realization as the true distribution and calculating the certainty coefficient between this realization and each of the remaining nine realizations. This approach allowed to calculate 45 certainty coefficients for different possible combinations among the ten realizations. The univariate statistics of the certainty coefficients between the realizations is shown in the third column of Table 2.4. The mean certainty coefficient between the realizations is lower than mean certainty coefficient with respect to the master distribution (first column of Table 2.4). Also, the variance and the range of the certainty coefficients between the realizations are significantly greater than for the certainty coefficients with respect to the master distribution. None of the certainty

Table 2.4

Univariate statistics of certainty coefficients
for different realizations of conditional
simulations of the master distribution

Parameter	Simulation Certainty Coefficient	Computation Cost (cycles)	Among Simulations Certainty Coefficient
Number	10	10	45
Mean	0.7042	491.057	0.6581
Variance	0.0003	5040.348	0.0016
Standard Deviation	0.0160	70.995	0.0405
Coefficient of Variation	0.0227	0.1446	0.0615
Minimum	0.6777	410.126	0.5679
Maximum	0.7297	603.248	0.7418
Median	0.7061	484.029	0.6532
1^{st} quartile	0.6878	425.212	0.6294
3^{st} quartile	0.7135	530.034	0.6927

coefficients between the realizations approaches a value near one, implying that any pair of the realizations is not close to be identical. On the other extreme, the minimum certainty coefficient between the realizations is significantly greater than zero, indicating that at least some major characteristics of the realizations are similar.

Conditioning Data

The conditioning data used in the simulations of the master distribution in the previous section consisted of 60 values arranged in three vertical segments. In this section, the number of conditioning data values is changed from zero (unconditional) to 120 values or close to 10 % of the number of grid points in the simulation region. For all cases, the locations of conditioning data is arranged in vertical segments, where each segment consists of 20 values. For the case corresponding to one vertical segment or 20 conditioning data values, this segment is assigned to the middle of the lateral or x direction and for the case of two vertical segments or 40 conditioning data values, these segments are assigned to the edges of the lateral direction. For the remaining cases, the vertical segments of conditioning data are uniformly distributed along the lateral direction of the simulation region.

The certainty coefficients of the simulated distribution for different number of conditioning data values are shown in Figure 2.24. Surprisingly, the simulated distribution with no conditioning data has a certainty coefficient greater than the certainty coefficients for the cases with 20 and 40 conditioning data. For the remaining cases, the certainty coefficient increases as the number of conditioning data values increases.

In order to investigate the effect of the arrangement of the conditioning data, an additional simulation has been conducted with the conditioning data arranged in horizontal segments instead of vertical segments. In this case, the conditioning data consists of 80 sample values arranged in four horizontal segments, where each segment consists of 20 values. The horizontal segments are located in the middle of the simulation region and are uniformly distributed along the vertical direction. The certainty coefficient for this case is shown by the solid point in Figure 2.24 and

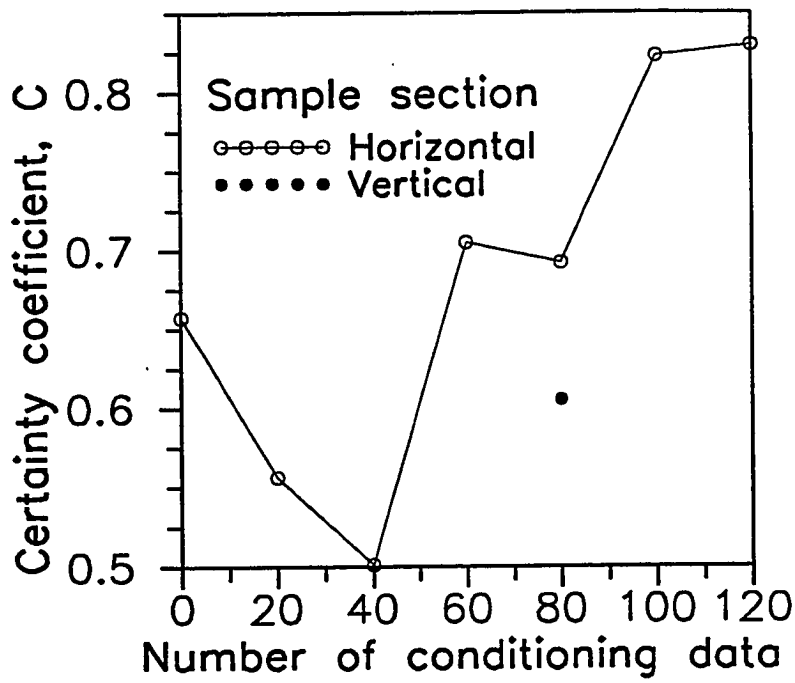

Figure 2.24 - Certainty coefficient of conditional simulations for different number of conditioning data

indicates that this is slightly lower than the certainty coefficient corresponding to the 80 conditioning values arranged in vertical segments.

Distribution Functions

The cumulative distribution functions used in the conditional simulations were derived from the master distribution and only the number of classes varies among the different distribution functions. The number of classes for the distribution functions ranges from 5 to 80 classes and for each case the within class distribution is assumed to be a uniform distribution. The certainty coefficients of the simulated distributions with different number of distribution classes are shown in Figure 2.25. These results indicate that the certainty coefficients for the different number of classes are not significantly different, except for the case of 25 classes which reaches an apparent minimum value. In these simulations, the number of classes of the distribution function should not have a significant effect on the simulated distributions because all the simulations use a uniform within class distribution which is a good approximation to the exhaustive distribution function of the master distribution which is also near a uniform distribution.

Semi-Variograms

Conditional simulations were conducted to investigate the effect of different number of semi-variogram models on the simulated distributions. In each simulation case, the semi-variogram models are specified for different directions and the models are the same as the ones given in Table 2.2. Only in one simulation discussed below, a parameter of a semi-variogram model has been modified.

The certainty coefficients of the simulated distributions with different semi-variograms specifications are given in Table 2.5. The certainty coefficients of the simulated distributions generated with one semi-variogram model for the x direction and two models for x and z directions are close to the certainty coefficient for the case of four models. For the case with only one semi-variogram in the z direction the certainty coefficient is almost equal to zero. Including the semi-variogram models

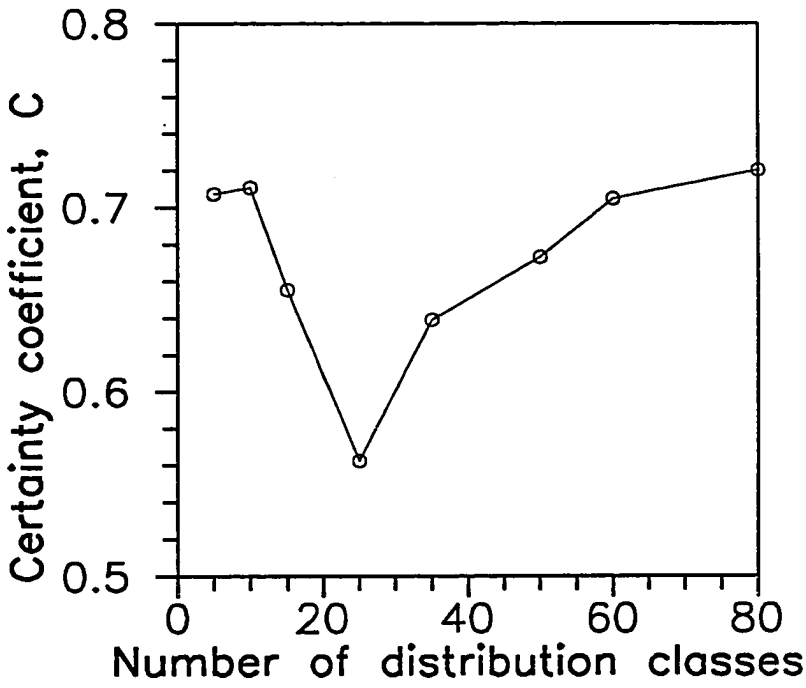

Figure 2.25 - Certainty coefficient of conditional simulations for different number of classes for the cumulative distribution function

Table 2.5

Certainty coefficient for conditional simulations of the master distribution for semi-variogram models along different directions

Simulated Directions	Certainty Coefficient
x	0.6558
z	0.0074
x and z	0.7288
x/z and -x/z	0.2831
z, x/z and -x/z	0.3010
x (sill=17), z, x/z and -x/z	0.7047
x (sill=97), z, x/z and -x/z	0.4062

in the diagonal directions yields a slight improvement of the certainty coefficient over the case with only a model in the z direction. These results indicate that the semi-variogram model in the direction of greater continuity or small variability has a significant effect on the quality of the predicted distributions.

Ironically, in practice, reliable information about the directions which are suspected to have greater continuity is seldom available. For example, if the conditioning data used for the simulations is assumed to be the available sample data then the sill of the semi-variogram in the direction of greater continuity calculated from the exhaustive master distribution cannot be estimated from the sample data. In order to test the importance of these practical considerations, in one simulation case, the sill of the semi-variogram model in the x direction is set to a false value equal to the variance of the conditioning data which is about three times greater than the sill found from the exhaustive distribution. The results in Table 2.5 indicate the certainty coefficient for the case with the false sill is smaller than for the case with the sill from the exhaustive distribution but it is greater than for the cases where the semi-variogram model in the x direction is ignored. This simulation case shows that it is better to make use of approximate or even poor information about the direction of greater continuity than ignoring this direction in the conditional simulations. Some guidelines to obtain information about the spatial correlation in the direction of greater continuity are discussed in Chapter IV for properties in a carbonate reservoir.

Potential Extensions

The conditional simulation method described in this chapter is fundamentally different from other conventional simulation techniques, such as the family of kriging techniques.[6] The simulated annealing conditional simulation method does not invoke or requires the conditions of stationarity. Thus, any realistic distribution or semi-variogram functions can be generated with the simulated annealing conditional simulation method. However, at this point, the differences or advantages of simulated annealing over conventional conditional simulation methods have not been established.

The generality and flexibility of the simulated annealing method opens the door to several potential extensions to incorporate more constraints or features into the simulation distributions. For example, one extension is to specify different distribution functions in different sub-regions of the simulation region. This constraint will be useful to incorporate information obtained from well test analysis in the description of a property, since this information provides a measure of the effective mean value of a property for a sub-region of the reservoir significantly greater than for the sub-region sampled by logging tools. For this extension, it will be necessary to include the distribution function of each sub-region as an additional term in the objective function. Another potential extension is to simulate simultaneously the distribution of two variables. The advantage of this extension is that a local relationship between two reservoir properties can be incorporated in the objective function, in addition to the variogram models of each variable.

CHAPTER III

FIELD VALIDATION

In this chapter, descriptions of reservoir properties are generated for a sandstone and a carbonate field using the stochastic conditional simulation method developed in Chapter II. In each field, a validation procedure is used to evaluate the reliability of the simulated distributions to represent the observed distributions of reservoir properties. A major difference between these field conditional simulations and the simulations presented in Chapter II is the additional degree of uncertainty in the evaluation of the statistical attributes of a reservoir property due to the scarcity of sample data in the field cases. The intent of selecting the two fields of distinct geologic origins is to expand the scope of the validation procedure for cases where the spatial distributions of reservoir properties are significantly different.

The validation procedure used in this chapter can be described in four steps. The first step consists of selecting a simulation region in the field which includes all the wells available for the study. The wells in the simulation region are treated as either conditioning wells or test wells. In the second step, the statistical properties of the simulation variable are estimated using the sample data from the conditioning wells. In the third step, conditional simulations of the property are generated using the statistical properties from the second step and the sample data at the conditioning wells as simulation specifications. In the last step, the values of the simulated variable at the locations of the test wells are compared to the sample data of the test wells.

The analysis of the spatial correlation of reservoir properties in this chapter is based on statistical fractal models. A brief introduction to fractal models is provided in the first section of this chapter. Then, for each field, the results of the steps in the validation procedure are discussed. In addition, the results of other conditional simulation methods are compared to the conditional simulation method developed in Chapter II for each field case.

Fractal Models

Statistical fractal models are useful to represent spatial records and time series of natural phenomena. Hewett[17] proposed an approach to model the spatial variations of reservoir properties based on the concept of statistical fractals. The two statistical fractal processes used in this chapter are the fractional Brownian motion (fBm) and the fractional Gaussian noise (fGn). A Comprehensive description and theoretical developments of statistical fractal models are given by Mandelbrot and Van Ness.[35] More general reviews on the application of fractals to model porous media are given by Williams and Dawe[36] and Sahimi and Yortos.[37]

The spatial correlation model of a fBm process is a power-law relation given by[35]

$$\gamma(h) = V_H h^{2H}. \tag{3.1}$$

fGn is approximately the derivative of fBm. An expression for the spatial correlation model of a fGn process has been derived by Mandelbrot and Van Ness[35] and Mandelbrot,[38] and it is given by

$$\gamma(h) = \frac{1}{2} V_H \delta^{2H-2} \left[2 - (\frac{|h|}{\delta} + 1)^{2H} + 2|\frac{h}{\delta}|^{2H} - (\frac{|h|}{\delta} - 1)^{2H} \right]. \tag{3.2}$$

The intermittency exponent, H, in Equations 3.1 and 3.2 is related to the fractal dimension and it varies between zero and one. Different ranges of H represent different types of spatial correlation. For $0 < H < 0.50$, the correlation is anti-persistent (high values tend to be followed by low values and vice versa) and for $0.50 < H < 1$, the correlation is persistent (successive values tend to be of similar magnitudes). For the case where $H = 0.50$, Equation 3.1 represents the semi-variogram of a Brownian motion or a random walk and Equation 3.2 becomes a constant value corresponding to a Gaussian noise or a white noise. In Equation 3.2, the factor $V_H \delta^{2H-2}$ is the variance of a fGn process. The parameter δ in Equation 3.2 is a smoothing factor which is required to differentiate fBm in order to arrive at the spatial correlation model for fGn.

In theory, the correlation ranges of fBm and fGn processes are infinite (except for a fGn process with $H = 0.50$). The semi-variogram models of fBm and fGn in Equations 3.1 and 3.2 are shown in Figure 3.1 for different intermittency exponents. The semi-variograms of fBm for intermittency exponents greater than 0.50 increase without bounds as the lag increases, while for fGn the semi-variograms tend to reach a constant value. The semi-variograms of fBm for small intermittency exponents (close to zero) are similar to the semi-variograms of fGn with large intermittency exponents (close to one). These similarities are illustrated in Figure 3.2 for the semi-variograms of fBm for $0.05 \leq H \leq 0.50$ and of fGn for $0.50 \leq H \leq 0.95$.

The smoothing factor, δ, in Equation 3.2 accounts for the influence of the size sample on the spatial correlation structure.[17,35] An example of this sample size effect on the semi-variogram of a fGn process is illustrated in Figure 3.3. In this example, the semi-variograms of fGn are compared for two cases with sample spacings equals to 1 and 5 units and smoothing factors equal to corresponding sample spacing. For both cases, the intermittency exponent and the factor V_H have the same value. Figure 3.3 shows that for the same lag, the magnitude of the semi-variogram is smaller for the larger sample spacing. This observation indicates that the semi-variogram of fGn accounts for the expected reduction on the variability of the processes as the sample spacing increases.

Several methods are available to estimate the intermittency exponent of fBm and fGn processes. For a fBm process, the power-law relation given by Equation 3.1 can be used to estimate the intermittency exponent directly from the slope of a plot of sample semi-variogram versus lags in logarithmic coordinates. Hewett[17] describes another method to evaluate the intermittency exponent of fBm using an expression for the spectral density representation of the semi-variogram given by Equation 3.1. The methods to evaluate the intermittency exponent of fGn include the Re-scaled range (R/S) analysis, spectral density and the grading methods. An extensive description of these methods is given by Hewett.[17] The R/S analysis is based on a scaling expression developed by Hurst[39] which relates the sequential range of the cumulative departures from the mean, lag and intermittency exponent.

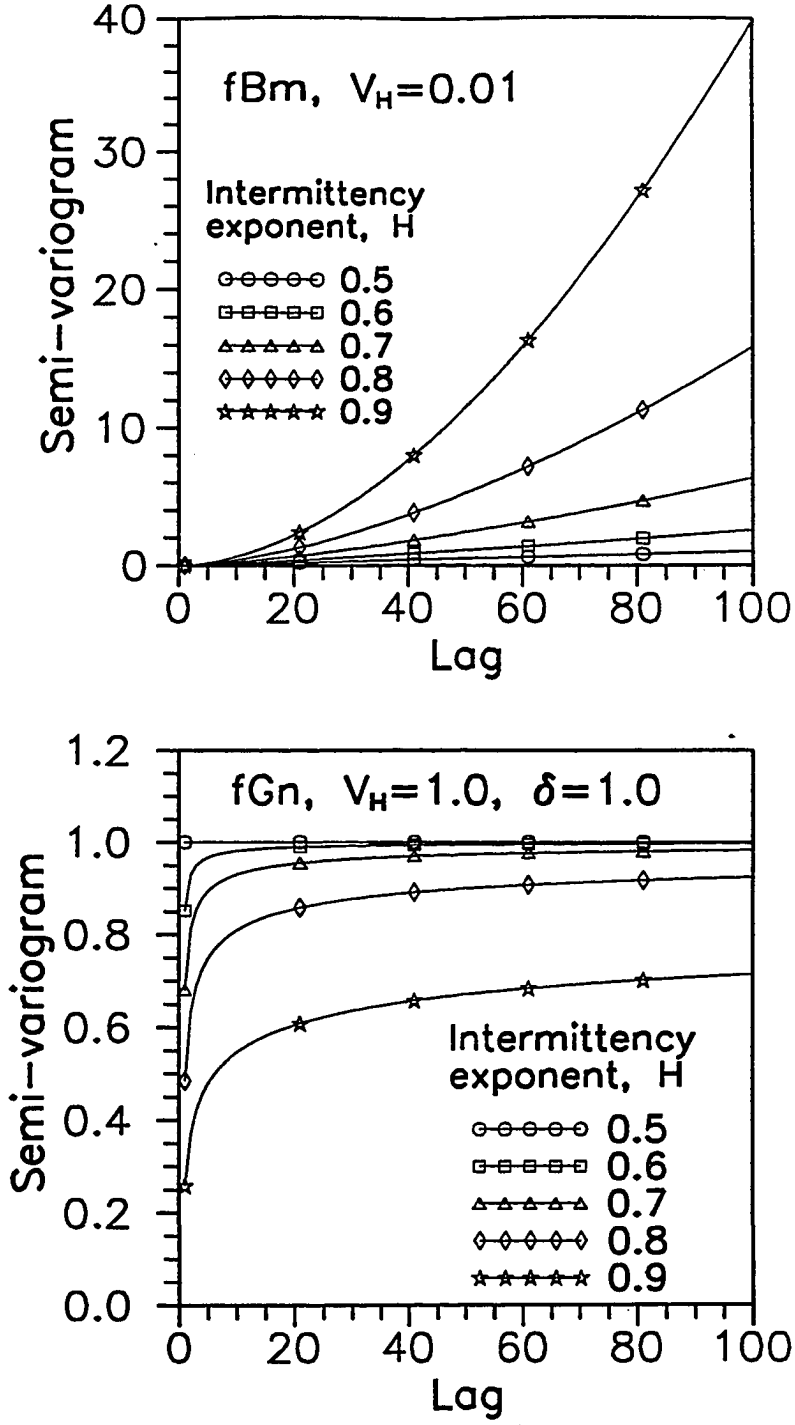

Figure 3.1 - Semi-variogram models of fBm for V_H=0.01 and of fGn for V_H=1.0 and δ=1.0

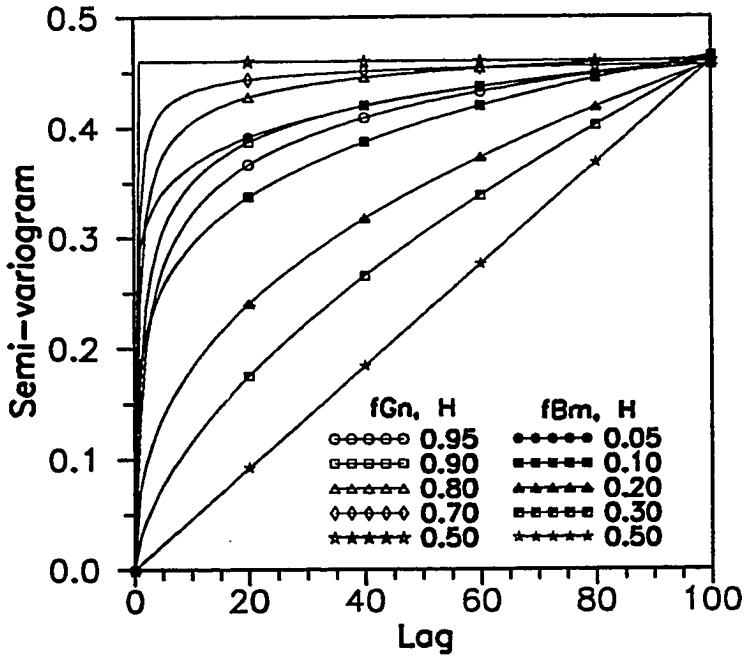

Figure 3.2 - Comparison of semi-variogram models of fBm for $0.05 \leq H \leq 0.50$ and of fGn for $0.50 \leq H \leq 0.95$ and $\delta = 1.0$

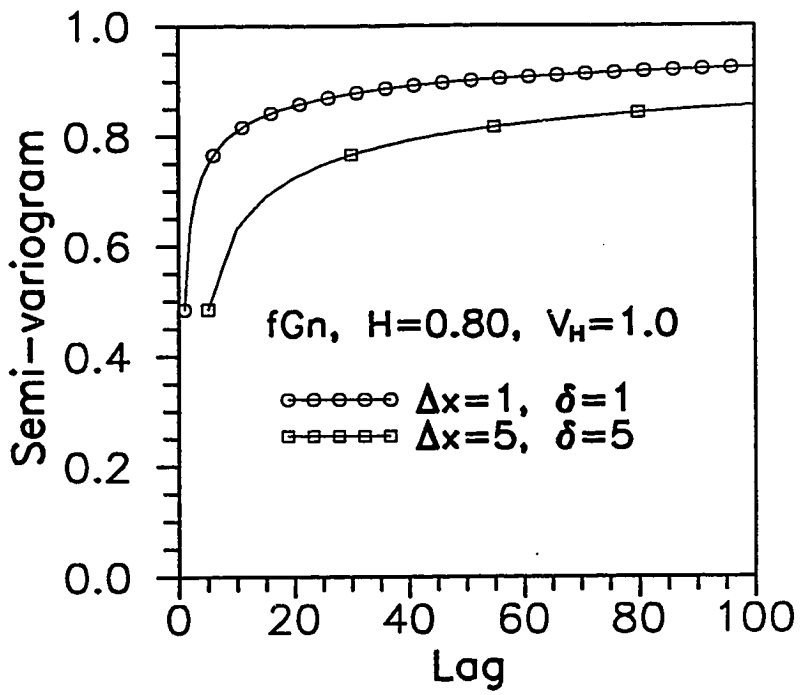

Figure 3.3 - Comparison of semi-variogram models of fGn for different sample spacings and size

In the spectral density method the intermittency exponent is calculated from an approximation of the spectral density of the semi-variogram of fGn. The grading methods consist of applying the methods developed to analyze fBm, such as the semi-variogram and the spectral density methods, to a graded fGn process. The graded sequence is the cumulative value or the discrete integral of a fGn process. Another method to estimate the intermittency exponent of a graded fGn sequence is the box-counting method described by Feder.[40] The box-counting method is based on the scaling expression for deterministic fractals which relates the number of boxes required to cover the trace of the process, the size of the boxes and the intermittency exponent. According to Hewett,[17] an advantage of the grading methods is that a graded sequence is more regular and may be easier to evaluate than a more erratic fGn process.

Sandstone Field

The sandstone oil field is part of the Robinson lease located in Muskogee county, Oklahoma. The lease covers a surface area of 160 acres. Subsequent to the discovery of the field in the 1930's, several wells were drilled during different primary development programs. Since 1985, the field production is supported by waterflooding.

This section provides a brief geologic description of the sandstone field. Then, the steps of the validation procedure for the conditional simulations of porosity are described in details. Finally, conditional simulations are generated with a method based on statistical fractals known as the successive random additions method and these results are compared to the conditional simulation method developed in Chapter II.

Geology

The sandstone field produces from a group of sands known as the Dutcher sands. Three sands can be identified in the Gamma-ray logs and these are denoted as A, B and C sands starting from the top of the reservoir. These sands are separated by thin shales. Core data indicate that the B sand has the highest rock quality

(average porosity and permeability equal to 19 % and 100 millidarcys, respectively) and it is the most productive sand.

The top of the reservoir is sealed by an impermeable limestone streak observed in the Gamma-ray logs throughout the field. The top of the reservoir dips in the southwest direction with a small angle of about 1 degree. Laterally, the reservoir is bounded by a combination of structural and stratigraphic traps. In the west side of the lease, the reservoir ends in a sealing fault defined by dry wells, as indicated in Figure 3.4. Towards the east side of the lease, the quality of the sands gradually diminishes.

Simulation Region

The simulation region in the sandstone field contains the four wells shown in a map of the lease in Figure 3.4. The conditioning wells for the conditional simulations are Well 2-6A and 2-10 and the test wells are Well 2-8A and 2-11. The simulation region is a vertical reservoir section between the conditioning wells. The length of this vertical reservoir section, corresponding to the distance between Wells 2-6A and 2-10, is 3030 feet and the thickness is 95 feet. This vertical reservoir section includes the three Dutcher sands. In this simulation region, the tops of the reservoir have been aligned to a common reference depth; thus, the small dip angle of the reservoir is ignored.

The sample data for the conditioning and test wells are porosity values derived from density logs at one foot spacing. For each well, there are 95 sample data points. The porosity logs of the conditioning wells are shown in Figure 3.5. The two high porosity regions around 2120 and 2150 feet in Well 2-6A (Figure 3.5) correspond to the B and the C sands, respectively and the A sand is directly above the B sand but it has lower porosities. In Well 2-10 (Figure 3.5), due to the poor quality of the reservoir on the east side of the lease, the three sands are not easily distinguishable on the basis of porosity as in the case of Well 2-6A.

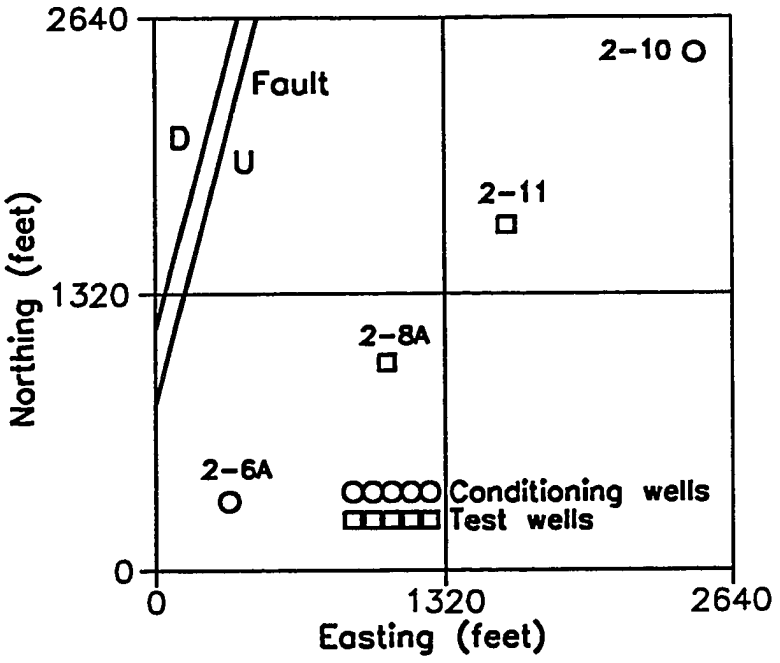

Figure 3.4 - Location of wells in the sandstone field

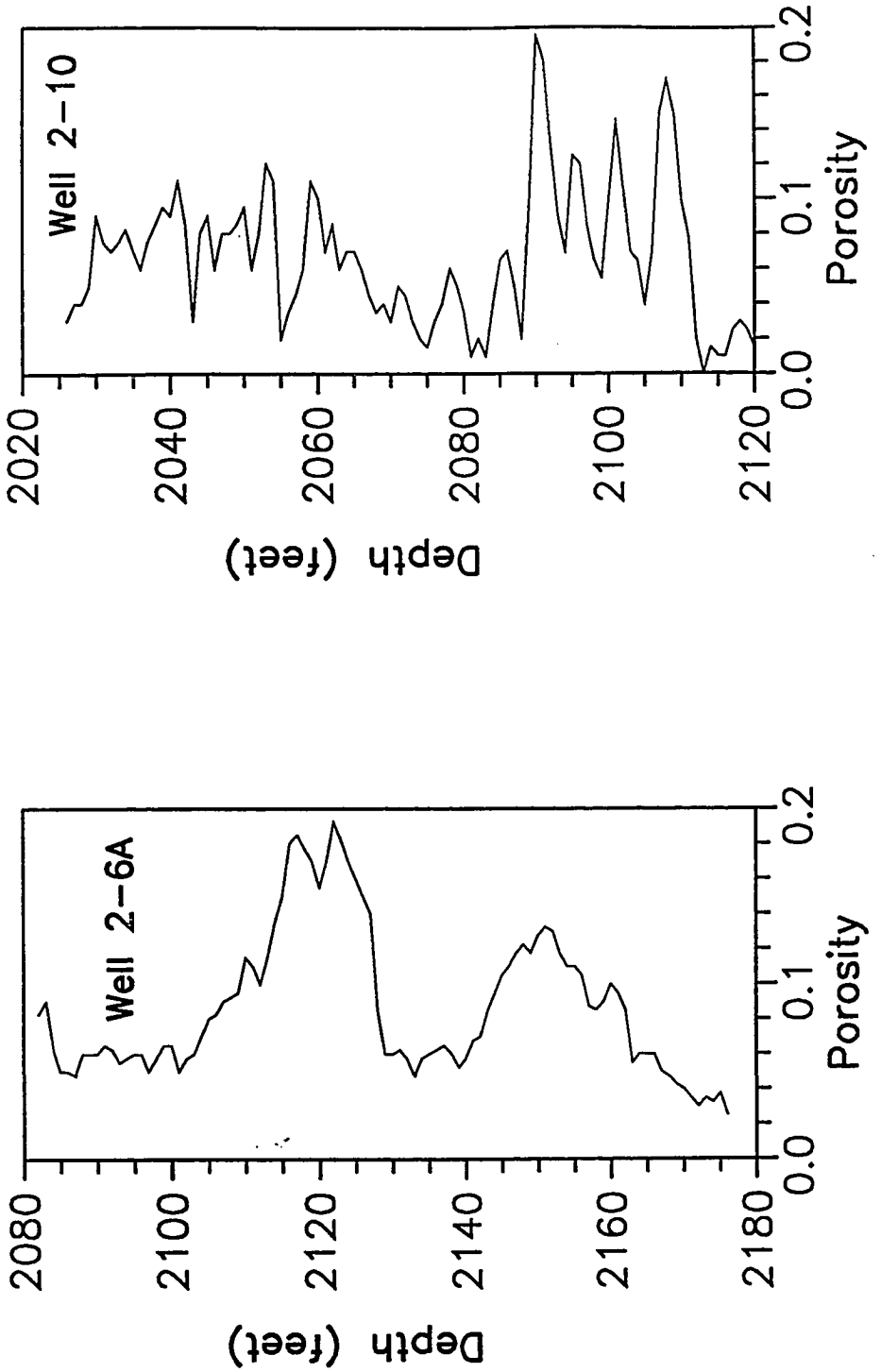

Figure 3.5 - Porosity logs of the conditioning wells

Univariate Statistics

The univariate statistics for the porosity sample data from the conditioning wells are summarized in Table 3.1. Overall, the univariate statistics of porosity for the two conditioning wells are similar. Although, the mean and variance of porosity for Well 2-6A are higher than for Well 2-10, these differences are not significant. The histogram for the porosity sample data from the conditioning wells is shown in Figure 3.6. This histogram is slightly skewed to the right. The cumulative distribution function specified in the conditional simulations is calculated from the histogram shown in Figure 3.6.

Spatial Statistics

In this section, a model of the spatial correlation of porosity is developed from the conditioning data using statistical fractals. For this field case, the available conditioning data allows a reliable evaluation of the spatial statistics only along the vertical direction. For the horizontal direction, two alternatives to infer the spatial correlation model required in the simulations are described later in the Conditional Simulations and the Successive Random Additions sections.

The intermittency exponents of the porosity logs from the conditioning wells are evaluated with the R/S analysis and the box counting methods. The results of the R/S analysis are shown in Figure 3.7 for Wells 2-6A and 2-10. For both conditioning wells, the slopes of the R/S versus lag plots are greater than one and indicate that the spatial correlation can be represented with fBm models. The intermittency exponents of the fBm models from the R/S results are 0.52 and 0.04 for the logs of Wells 2-6A and 2-10, respectively. The plots used to evaluate the intermittency exponent with the box counting method for the logs of the conditioning wells are shown in Figure 3.8. The box-counting method indicates a fGn type of behavior with intermittency exponents equal to 0.93 and 0.92 for the logs of Wells 2-6A and 2-10, respectively.

The results of the spatial statistics analysis from the R/S and the box counting methods appear to be contradictory; However, these discrepancies can be explained

Table 3.1

Univariate statistics of porosity (fraction) from the density logs of the conditioning wells in the sandstone field

Parameter	Well 2-6A	Well 2-10	Wells 2-6A and 2-10
Number	95	95	190
Mean	0.0869	0.0673	0.0771
Variance	0.00172	0.00158	0.0017
Standard Deviation	0.0415	0.0398	0.0417
Coefficient of Variation	0.4774	0.5922	0.5415
Minimum	0.025	0.0	0.0
Maximum	0.1925	0.1950	0.1950
Median	0.070	0.0650	0.0688
1^{st} quartile	0.0575	0.0350	0.0500
3^{rd} quartile	0.110	0.0862	0.0975

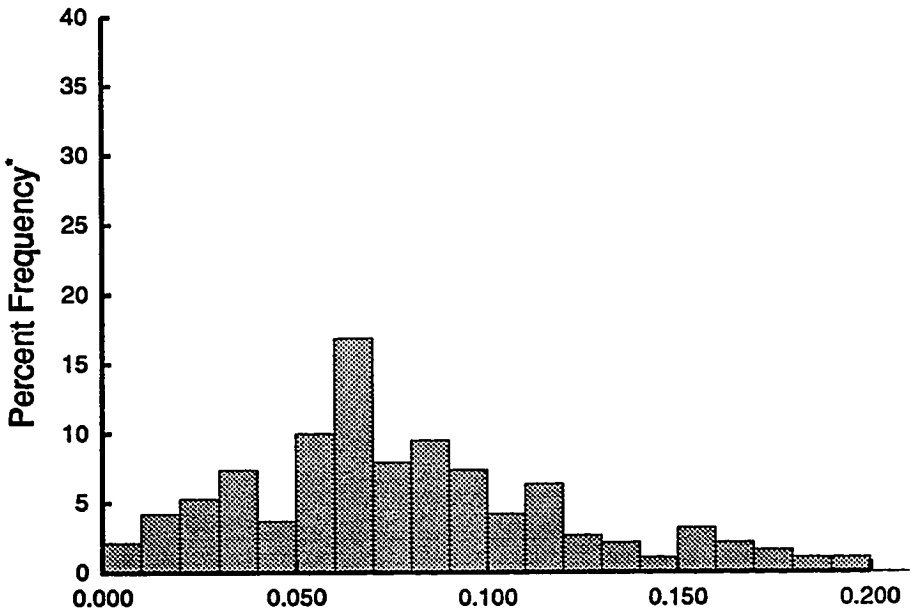

Figure 3.6 - Histogram for porosity logs of the conditioning wells

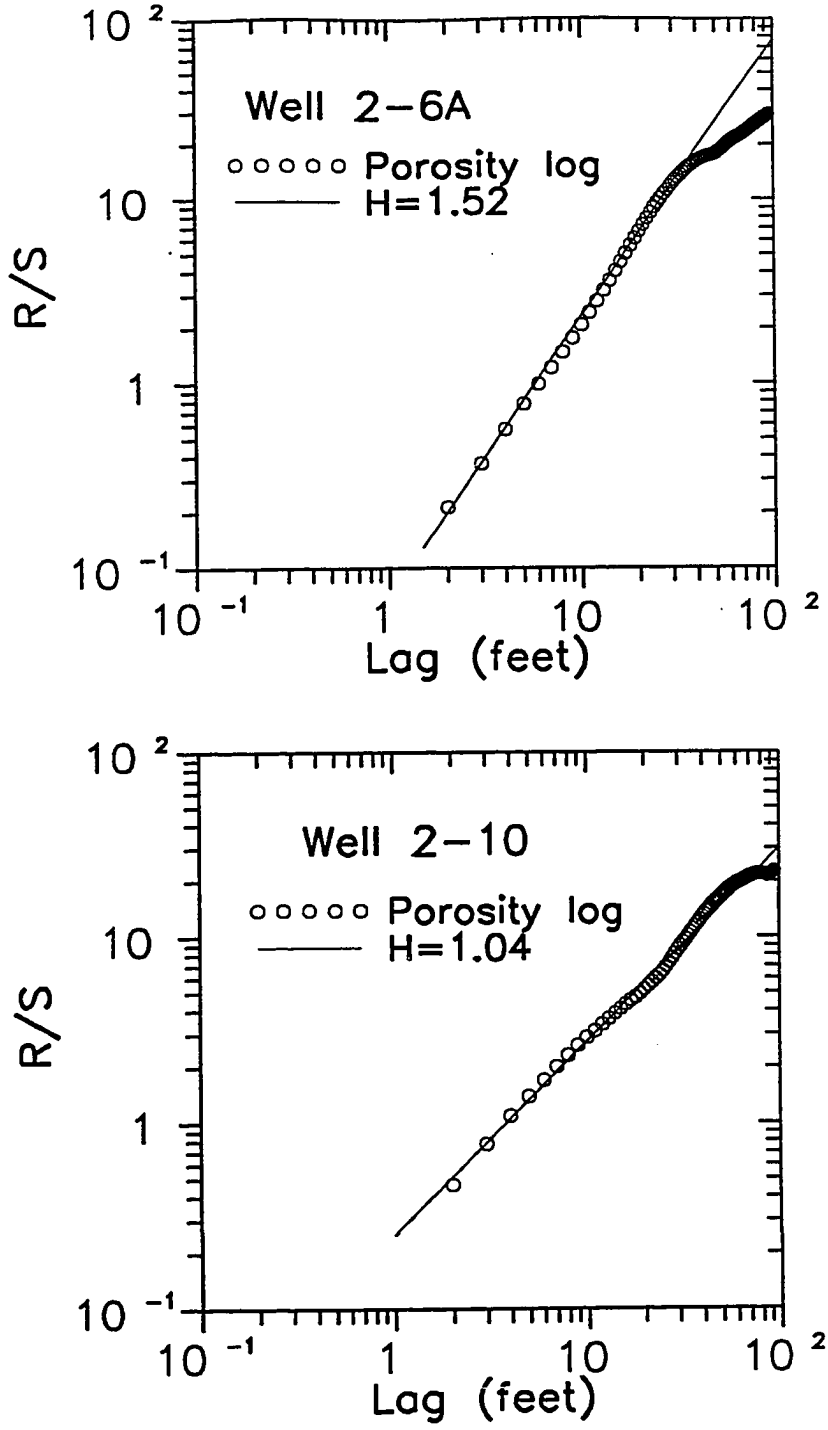

Figure 3.7 - R/S analysis for porosity logs of the conditioning wells

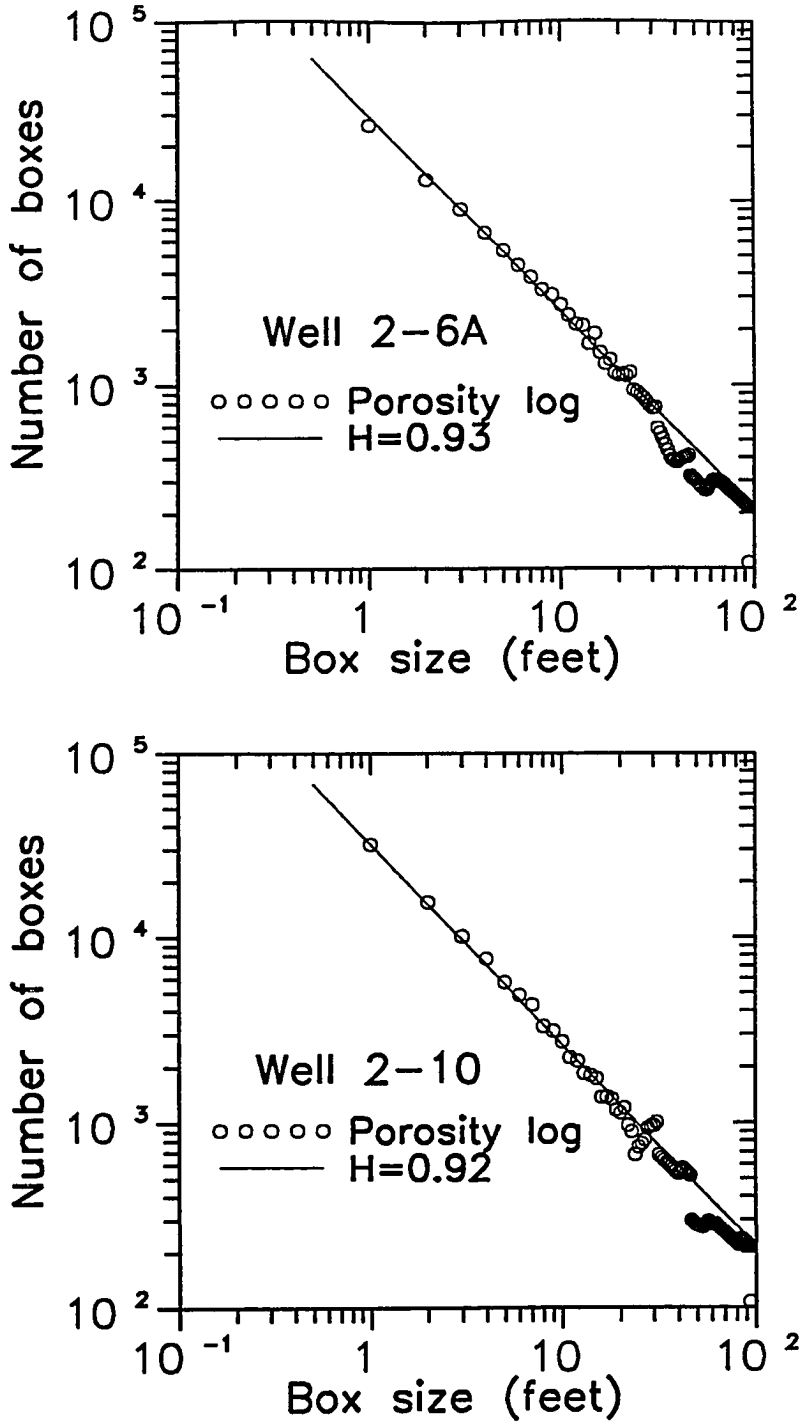

Figure 3.8 - Box counting method plot for porosity logs of the conditioning wells

for both wells. For Well 2-6A, it can be observed in Figure 3.5 that the trace of the porosity log has the appearance of a Brownian motion. Figure 3.9 shows the sample semi-variogram of the the standardized porosity log (zero mean and unit variance) of the conditioning wells. As indicated in Figure 3.9, the semi-variogram for Well 2-6A follows the expected linear relation in logarithmic coordinates for a fBm model (Equation 3.1) and the intermittency exponent calculated from this semi-variogram plot, $H = 0.68$ is also in agreement with the one found in the R/S method. Since it is clear that the log of Well 2-6A can be represented with a fBm process, then the box-counting method is not applicable because it is used as a grading method. For Well 2-10, the intermittency exponents derived from the R/S and the box counting methods are close to one for a fGn model and close to zero for a fBm model. In this range of intermittency exponent values, as explained in the Fractal Models section, it may be difficult to discern between the fBm and fGn processes.

The analysis of the spatial statistics for the individual logs is useful because it provides direct evidence of the presence of lateral variations of porosity as a result of the observed changes in the type of vertical correlation at different lateral locations. However, a conditional simulation method requires a single and unified semi-variogram model along a direction. Hence, the semi-variogram model for the vertical direction in the conditional simulations is derived from the single sample semi-variogram which includes the data of the two conditioning wells. The sample semi-variogram of the two conditioning wells and the fitted fGn model are shown in Figure 3.9. The parameters of this fGn model (Equation 3.2) are $H = 0.92$, $V_H = 0.0055$ and $\delta = 3$ feet. A comparison of the structure of the semi-variograms for the individual conditioning wells and for the two wells in Figure 3.9, indicates that linear trend observed for Well 2-6A is weaker than for the two wells. The hole effect observed for Well 2-6A in Figure 3.9 at a lag around 30 feet is still present in the semi-variogram of the two wells and it is an indication of the layered nature of this reservoir.

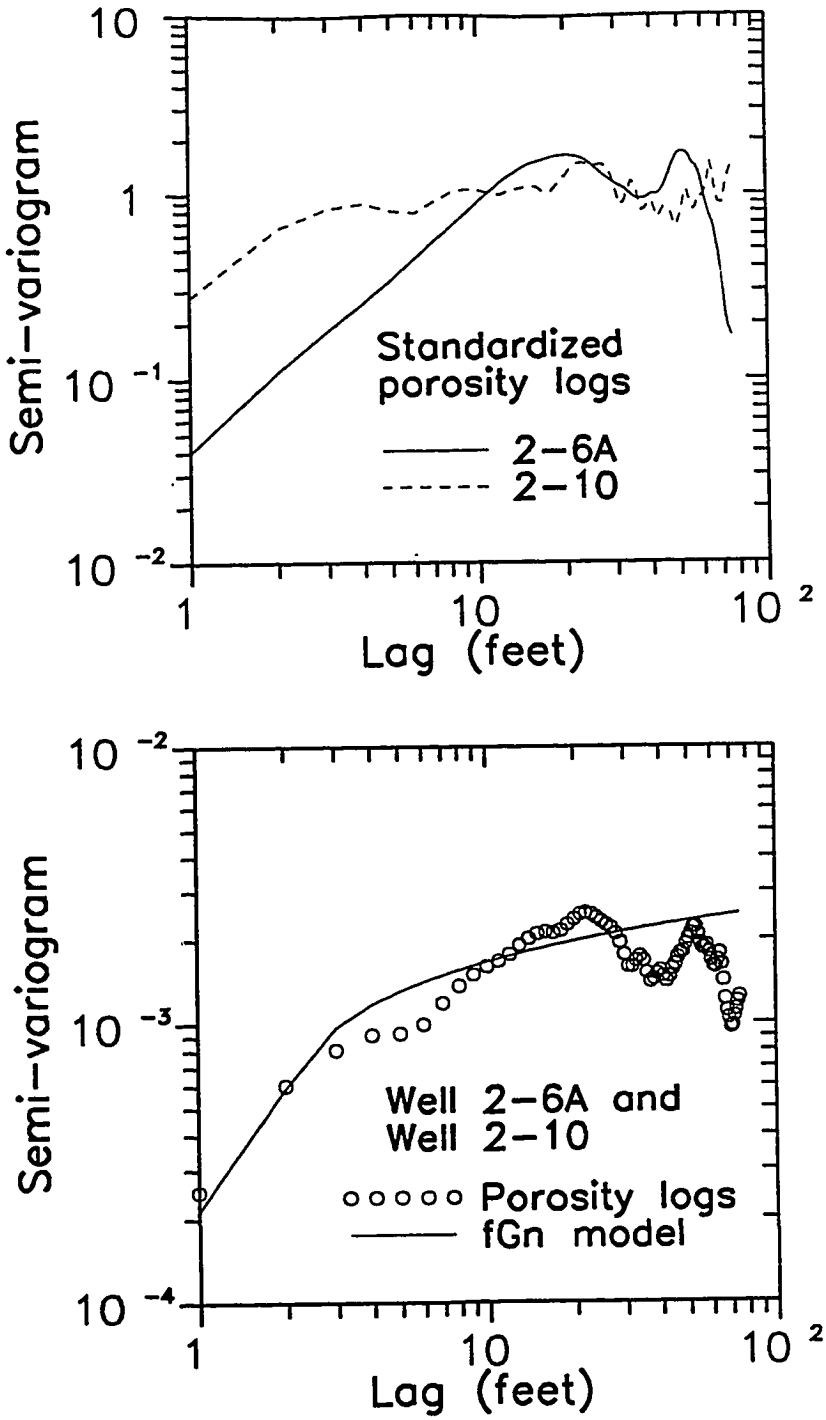

Figure 3.9 - Semi-variograms for porosity logs of the conditioning wells

Conditional Simulations

This section corresponds to the last step of the validation procedure and presents the results of the conditional simulations of the porosity variable using the method developed in Chapter II. The specifications for the conditional simulations for the vertical reservoir section described in the Simulation Region section are summarized in Table 3.2. The number of points used to discretize the vertical direction has been selected so as to include all the available conditioning data in the simulations. The conditioning data and the distribution function were described in the Simulation Region and the Univariate Statistics sections, respectively, and the parameters of the annealing schedule were selected according to the guidelines established in the Annealing Schedule section in Chapter II.

In this section, the semi-variogram model for the horizontal direction is set equal to the fGn model derived for the vertical direction in the Spatial Statistics section, except for the differences in the parameter δ of Equation 3.1 which depends on the grid spacing. As explained in the Fractal Models section (Figure 3.3), the magnitude of the variability for this horizontal semi-variogram model is smaller than for the model in the vertical direction due to the longer grid spacing in the horizontal direction. This choice of the semi-variogram model for the horizontal direction has been made due to the lack of information to evaluate a semi-variogram model in this direction. The analysis in the Quantifying Uncertainty section in Chapter II shows that semi-variogram model in the direction of greater continuity has a significant effect on the conditional simulations. Therefore, other semi-variogram model to represent the spatial correlation in the horizontal direction is attempted in the Successive Random Additions section.

A realization of the conditional simulation of the distribution of porosity is shown in the gray scale map in Figure 3.10. These map shows that there are regions of low and high porosity magnitudes which are more continuous in the horizontal than in the vertical direction. However, these regions of high porosity continuity are not sufficiently long in the lateral direction to indicate the presence of a layered structure throughout the simulation region.

Table 3.2

Specifications for conditional simulations of porosity in the vertical section between the conditioning wells in the sandstone field

Grid Geometry			
Direction	x	y	z
Spacing	47.34	1.0	1.0
Grid points	65	1	95
Total points	6175		
Conditioning Data			
Source		Density logs of Wells 2-6A and 2-10	
Number		190	
Distribution Function			
Source		Conditioning data	
Number of classes		15	
Subclass distribution		Uniform	
Semi-variogram Models			
x direction		Equation 3.2, $H = 0.92$, $V_H = 0.0055$, $\delta = 3\times1$ feet	
z direction		Equation 3.2, $H = 0.92$, $V_H = 0.0055$, $\delta = 3\times47.34$ feet	
Annealing Schedule			
α		0.50	
M_a		5.0 cycles	
ϵ_v		1.0×10^{-4}	
ϵ_a		2.5×10^{-2}	

Figure 3.10 - Map of porosity distribution from conditional simulation with fGn models

Comparisons of the actual porosity logs of the test wells and the simulated porosities are shown in Figures 3.11 and 3.12 for two realizations of the conditional simulations. These comparisons show that simulated porosity logs appear to resemble the principal characteristics of the test well logs with only few major deviations. However, a more qualitative evaluation between the actual logs and the corresponding simulated porosities, using the certainty coefficient defined in Chapter II (Equation 2.47), indicates that the certainty coefficients are nearly equal to zero for these conditional simulations. There are two explanations for the low magnitudes of these certainty coefficients compared to the certainty coefficients for the simulations of a synthetic distribution in the Annealing Schedule section in Chapter II. First, the low magnitude of the certainty coefficient for the simulations in this section reflect the uncertainties about the distribution function and, specially, the semi-variogram models in addition to the inherent uncertainties. Second, the number of available test or actual values is small relative to the total number of grid points in the simulation, while in the simulations in Chapter II, the test values for the master distribution are available for every simulation grid point. This implies that the tests in this chapter are more severe since only few poor predictions can have a significant effect in the certainty coefficient because these few predictions can represent a large fraction of the test data.

Successive Random Additions

Successive random additions is a stochastic interpolation method proposed by Hewett[17] for conditional simulation of a property in a vertical reservoir section between two conditioning wells. This method consists of performing a one dimensional stochastic interpolation between the values at the conditioning wells at each vertical level. The successive random additions method reproduces only semi-variogram models of fBm (Equation 3.1) in the horizontal direction with a specified intermittency exponent. The hypothesis of Hewett's[17] approach is that the intermittency exponent required for the semi-variogram of fBm in the horizontal direction can be set equal to the intermittency exponent of a fGn semi-variogram model calculated

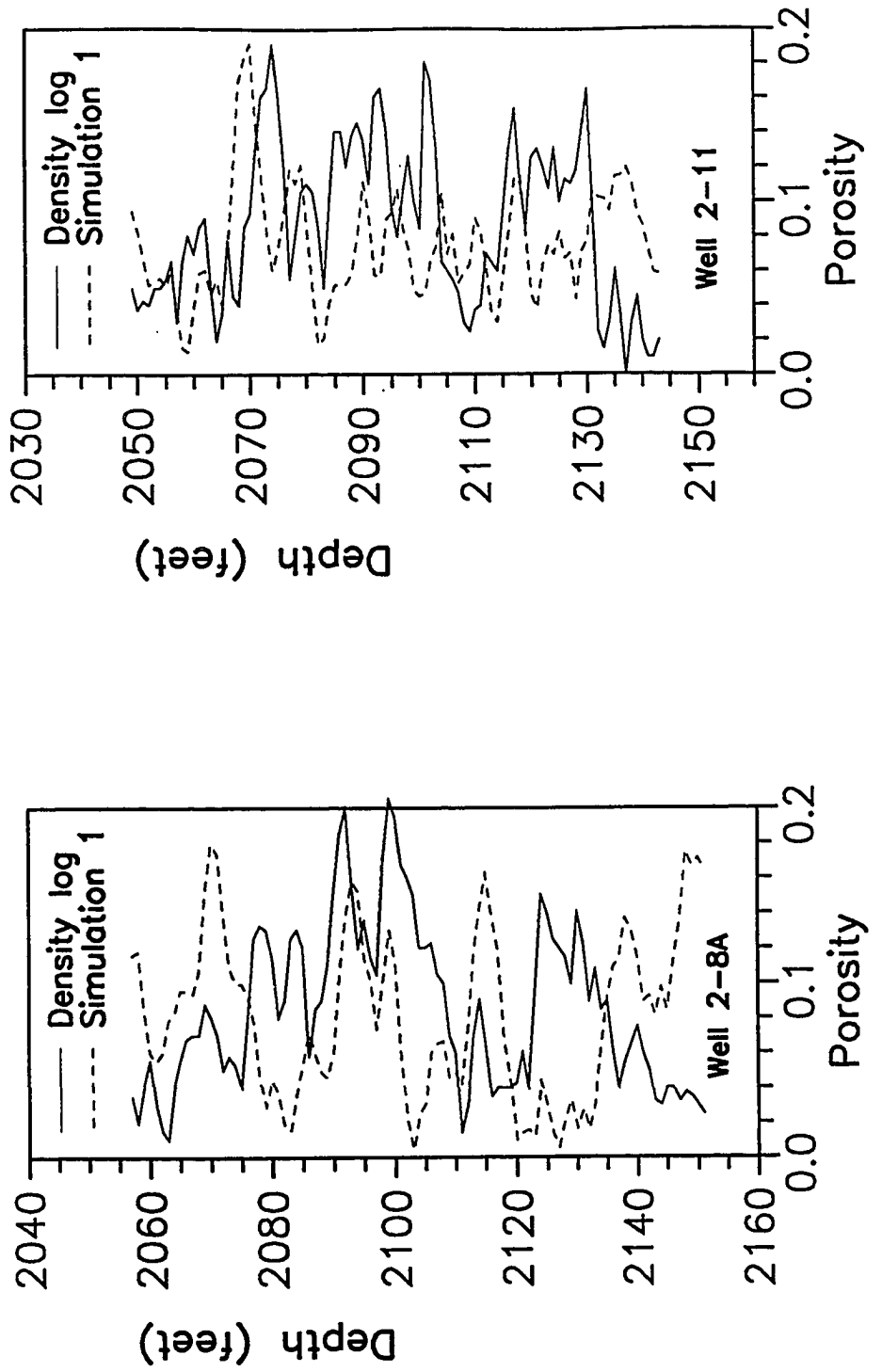

Figure 3.11 - Comparison of porosity logs of the test wells and first realization of conditional simulation with fGn models

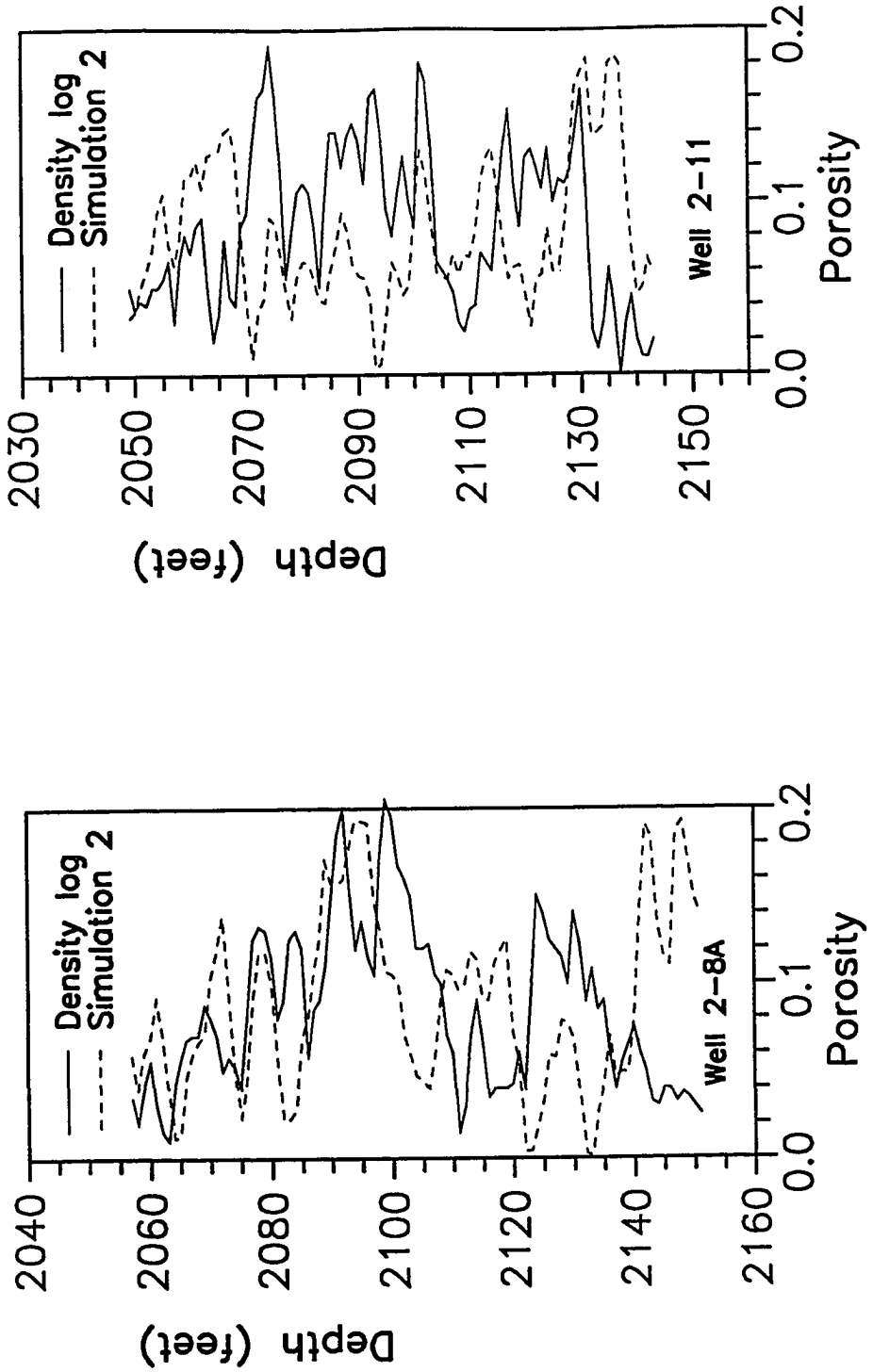

Figure 3.12 - Comparison of porosity logs of the test wells and second realization of conditional simulation with fGn models

from the sample data in the vertical direction. For example, if the analysis of a vertical log yields an intermittency exponent equal to 0.90, corresponding to the model shown in Figure 3.1, then the semi-variogram model for the horizontal direction should be similar to the model of fBm for $H = 0.90$ also shown in Figure 3.1. This example indicates that the semi-variograms for the vertical and the horizontal directions, in Hewett's hypothesis, are significantly different. Furthermore, for the horizontal direction, the semi-variogram does not reach a sill within the simulation region.

The successive random additions method has been used to generate a distribution of porosity in the vertical reservoir section described in the Simulation Region section. The intermittency exponent used for the semi-variogram model of fBm in the horizontal direction is 0.92. The distribution of porosity generated with successive random additions is shown in the gray scale map of Figure 3.13. The continuous black regions in this distribution correspond to the B and the C sands which have high porosities. The porosity distribution in Figure 3.13 has a lateral variability significantly smaller than the variability of the distribution shown in Figure 3.10 for simulated annealing. This small variability or large continuity occurs because in this case the successive random additions are almost identical to a linear interpolation. In fact, the case when $H = 1.0$ in the semi-variogram model of fBm, the successive random additions method is identical to a linear interpolation. In addition, Aasum et al.[23] showed that the distributions generated by successive random additions are insensitive to the value of the intermittency exponent for the range of $0.70 < H \leq 1.0$, if the variance of the difference between the values at the two conditioning wells is small.

Comparisons between the porosity logs of the test wells and the distribution simulated by the successive random additions method are shown in Figure 3.14. These comparisons indicate that there is a close agreement between the logs of the test wells and the simulations. The variability of the simulated logs appears to be less than the variability observed in the test wells. The certainty coefficient for these simulated logs is 0.53 and it is greater than the certainty coefficient of the simulated annealing distribution in the described in the Conditional Simulation section.

Figure 3.13 - Map of porosity distribution from successive random additions method

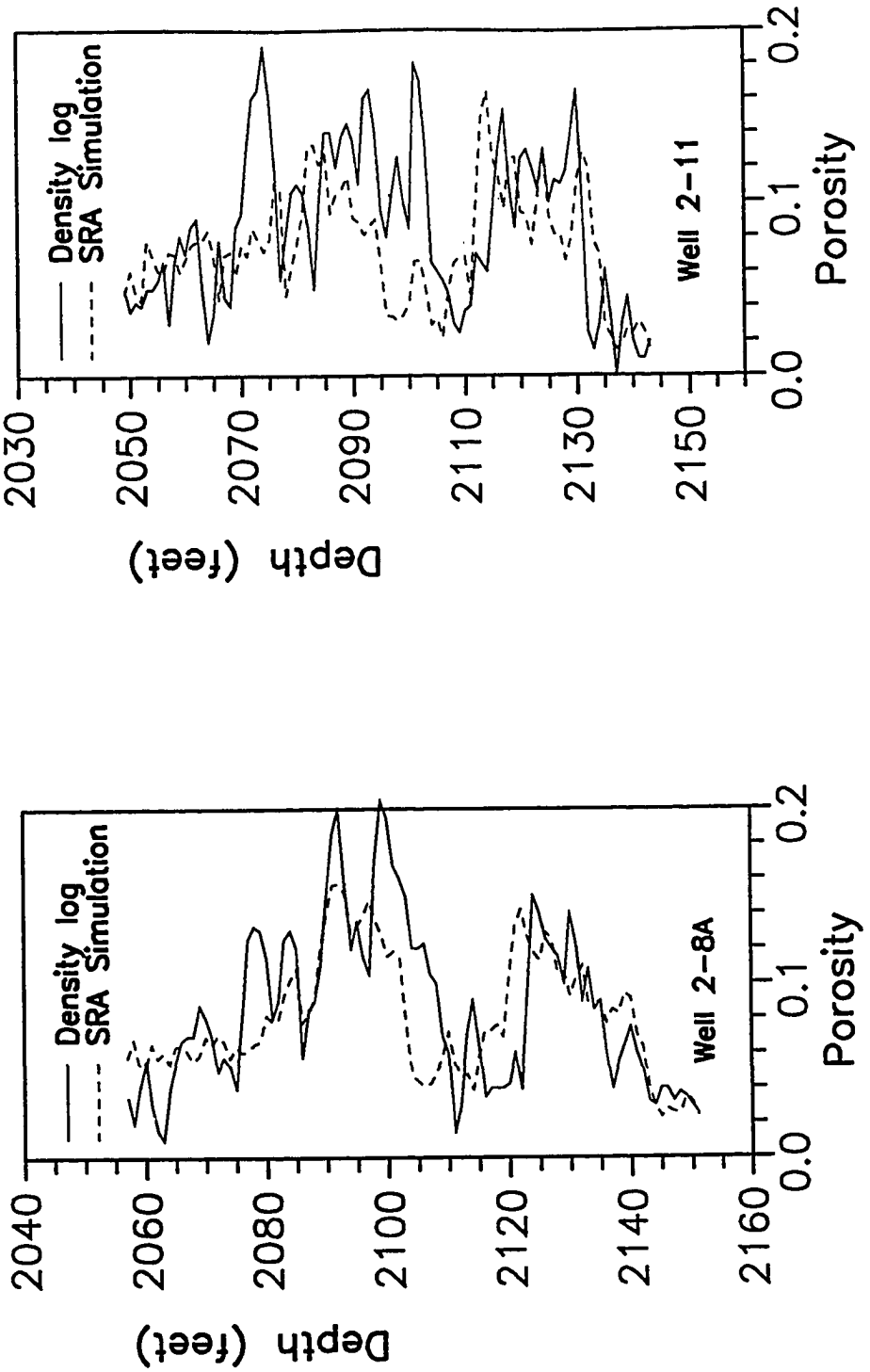

Figure 3.14 - Comparison of porosity logs of the test wells and conditional simulation from successive random additions method

The conditional simulations for the simulated annealing method were repeated for the specifications of Table 3.2 but with a semi-variogram model of fBm in the horizontal direction similar to the model used in the successive random additions simulations instead of the model of fGn. The parameters of this semi-variogram model (Equation 3.1) are $H = 0.92$ and $V_H = 1.032 \times 10^{-9}$. The value of factor V_H has been calculated from Equation 3.1 by setting the semi-variogram equal to the variance of the difference of the porosity values ($\gamma(h) = 0.002627$) of the conditioning wells for a lag equal to the separation between the conditioning wells ($h = 3030$ feet). Incidentally, an additional evidence that the porosity is still correlated at a lag distance corresponding to the separation between the conditioning wells is that variance of the difference of the conditioning wells (0.002627) is smaller than the sum of the variances of the porosity logs (0.003309) of the individual conditioning wells.

The distribution of porosity for the simulated annealing conditional simulation using the fBm model is shown in the gray scale map in Figure 3.15. The regions of high porosity values in this distribution are similar to the distribution of successive random additions in Figure 3.13. However, the variability of the simulated annealing distribution is greater because a semi-variogram model in the vertical direction is also specified. The comparisons of the logs of the test wells and the simulated logs are shown in Figure 3.16. These results indicate that the agreement between the logs of test wells and the simulations is closer than for the simulations with fGn models in the horizontal direction shown in Figures 3.11 and 3.12. The certainty coefficient for these simulations is 0.21 and it is smaller than for the successive random additions simulations.

An additional reason for the high variability observed in the results from simulated annealing is that the simulated semi-variogram model in the horizontal direction differs slightly from the specified model of fBm. In order to test if increasing the number of steps in simulated annealing allows a closer convergence to the specified semi-variogram model, the acceptance ratio tolerance has been reduced to $\epsilon_a = 1.0 \times 10^{-3}$ from $\epsilon_a = 2.5 \times 10^{-2}$. Comparisons of the semi-variograms of the simulated

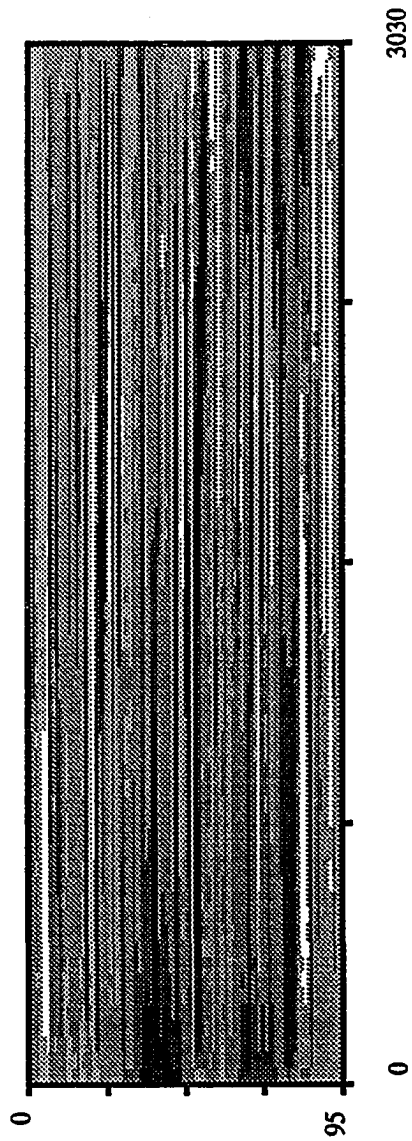

Figure 3.15 - Map of porosity distribution from simulated annealing method with horizontal fBm model

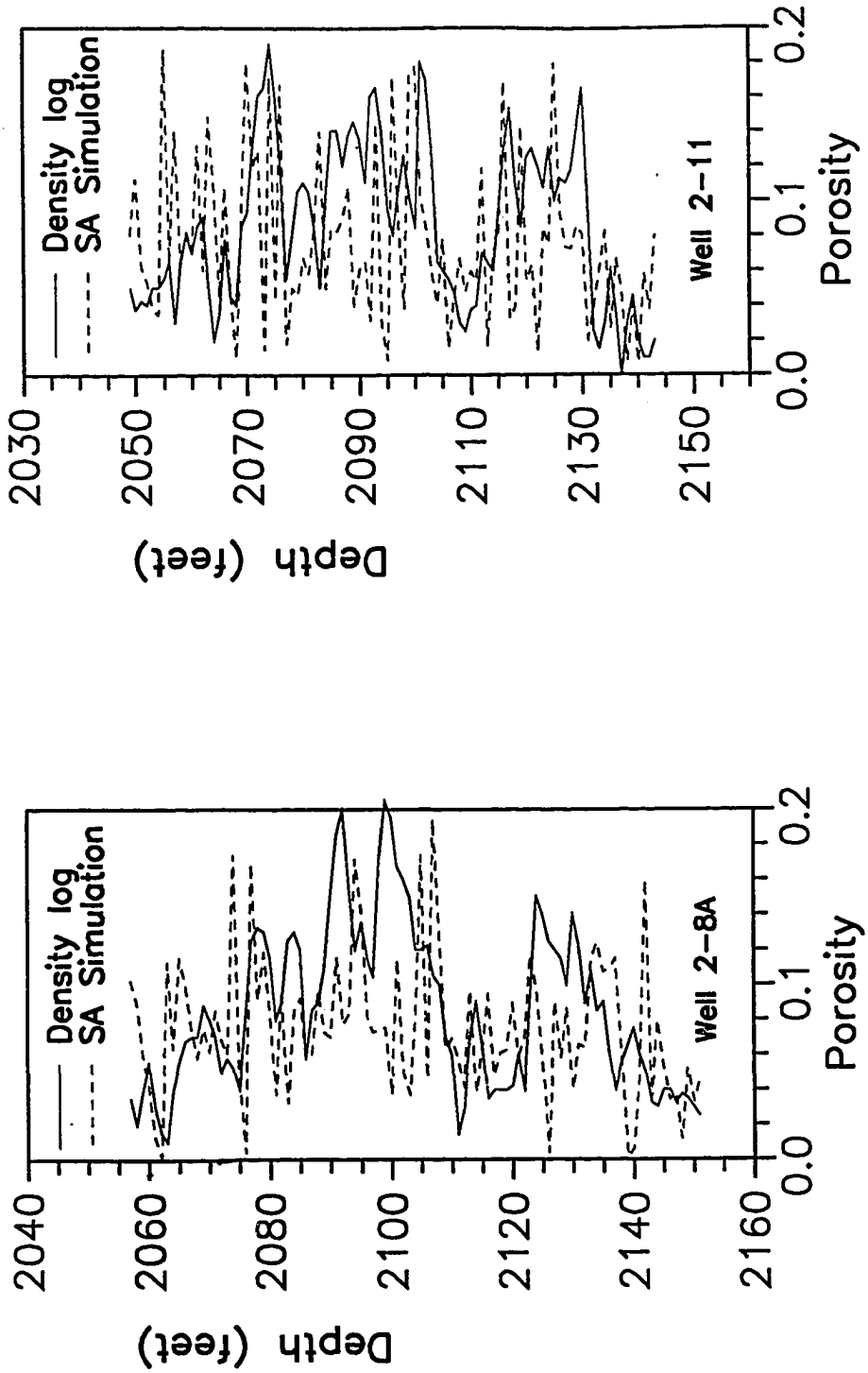

Figure 3.16 - Comparison of porosity logs of the test wells and conditional simulation from simulated annealing method with horizontal fBm model

distributions and the specified model are shown in Figure 3.17. For the smaller acceptance ratio tolerance, the semi-variogram of the simulated distribution is in a close agreement with the specified model (Figure 3.17). Also, it can be noted that the semi-variogram of the initial distribution (close to the sample variance) is several orders of magnitude greater than the specified model for the small lags and it is a remarkable task for the simulated annealing method to be able to reproduce such a strict constraint. The certainty coefficient between this simulation with a smaller acceptance ratio tolerance and the data of the test wells is 0.40 and it is close to the successive random additions simulation. However, the computational effort required for the simulation with the smaller acceptance ratio tolerance (1.0×10^{-3}) is about ten times greater than for the simulation with the larger acceptance tolerance (2.5×10^{-2}).

The analysis in this section indicates that the spatial correlation of porosity in the horizontal direction of the sandstone field is well represented by the semi-variogram model of fBm which is close to a linear interpolation. This lateral correlation is so strong that it masks the influence of the vertical correlation. Under these specific conditions, successive random additions is a preferable conditional simulation method because it requires only a small computational cost relative to the simulated annealing method. However, the conditions observed in the sandstone field may be rare even for reservoirs with a layered structure, such as the case discussed in the Master Distribution section in Chapter II. The conditional simulation method based on simulated annealing is more general since it can represent diverse types of spatial correlation models. Even though, the distributions generated by simulated annealing with a fGn model in the horizontal direction have low certainty coefficients due to the reasons explained in the Conditional Simulations section, the overall agreement with the data of the test wells (Figures 3.11 through 3.12) is acceptable and the few deviations reflect the different uncertainties, especially with respect to the semi-variogram model in the horizontal direction.

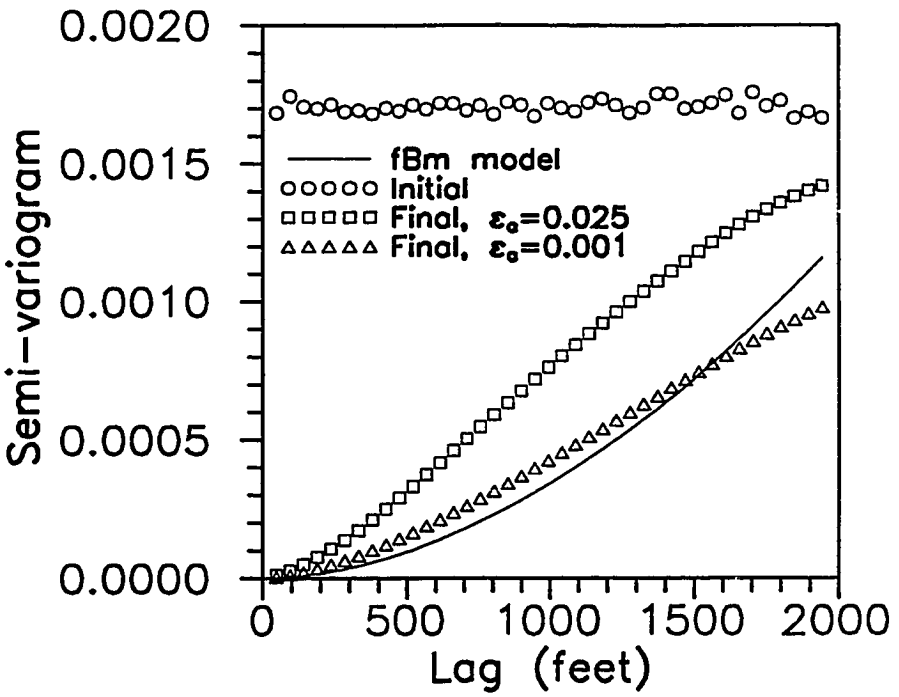

Figure 3.17 - Specified and simulated semi-variograms for conditional simulations from simulated annealing method with horizontal fBm model

Carbonate Field

The carbonate field is an oil field located in west Texas and covers approximately 1950 surface acres. The natural production mechanism in the carbonate field is solution-gas-drive. The primary production life of the field was 10 years and was followed by 17 years of waterflooding with peripheral and staggered line patterns. The current tertiary production mechanism in the field is water alternating miscible gas flooding with a 20 acres well spacing. The carbonate field has 32 production wells and 33 water and gas injection wells.

This section follows the same organization as the Sandstone Field section. The validation procedure is used to test three-dimensional distributions of a porosity variable generated with the conditional simulation method developed in Chapter II. Then, this conditional simulation method is compared to the sequential indicator simulation[13] method for two-dimensional conditional simulations of porosity in a vertical reservoir section.

Geology

The carbonate field produces from the San Andres formation. The lithology of the predominant facie in this formation is dolomite. Several heterogeneities as a result of the dolomitization process have been identified in cores from the carbonate field, including burrows, anhydride cementation and chert nodules. The depositional environments of the San Andres formation in the carbonate field are shallow restricted marine and intertidal environments.

The producing formation in the carbonate field has an anticline structure and it is located at an average subsea depth of 3700 feet. The maximum gross thickness of this formation is approximately 150 feet at the center of the anticline. The top and the bottom seals of this formation are a dense dolomite and an inactive aquifer, respectively.

Simulation Region

The simulation region is located near the center of the carbonate field, where the formation reaches the maximum thickness. The locations of the conditioning

(Wells 8, 11, 42 and 46) and test (Wells 36, 39 and 40) wells within the simulation region used in the three-dimensional conditional simulations are shown in the map in Figure 3.18. The dimensions of this simulation region along the easting, northing and vertical directions are 1219, 2068 and 149 feet, respectively. For the two-dimensional conditional simulations, a subset of the simulation region shown in Figure 3.18 is used, as explained in the Sequential Indicator Simulations section.

The porosity logs of the conditioning wells are shown in Figures 3.19 and 3.20. The porosity values in these logs have been derived from the traveling times from the acoustic logs using a matrix and fluid traveling times equal to 43.5 and 189.0 μseconds/feet, respectively. Although, the acoustic logs are available at 0.50 feet spacing, the porosity logs used for the conditioning wells are at a 1 foot spacing. For Well 46, Figure 3.20 shows that there is an irregularity in the porosity log between 5200 and 5225 feet and it may be due to measurement errors of the acoustic logging tool (cycle skipping).

Univariate Statistics

The univariate statistics of the porosity logs for the conditioning wells are summarized in Table 3.3. A point to note in Table 3.3 is that the variances of the porosity logs for Wells 8 and 46, located in the north part of the simulation region, are higher than the variances for the for Wells 11 and 42. A similarity among the porosity logs of the conditioning wells is the region of a high porosity mean between 5175 and 5200 feet, as indicated in Figures 3.19 and 3.20.

A histogram of the porosity logs for all the conditioning wells is shown in Figure 3.21 and indicates that the distribution of the sample data is close to a Normal distribution.

Spatial Statistics

The spatial statistics of the porosity logs for the conditional wells are analyzed with statistical fractal models. The intermittency exponents have been estimated with the R/S analysis and the box counting method and the results are summarized in Table 3.4. The plots used to evaluate these intermittency exponents for

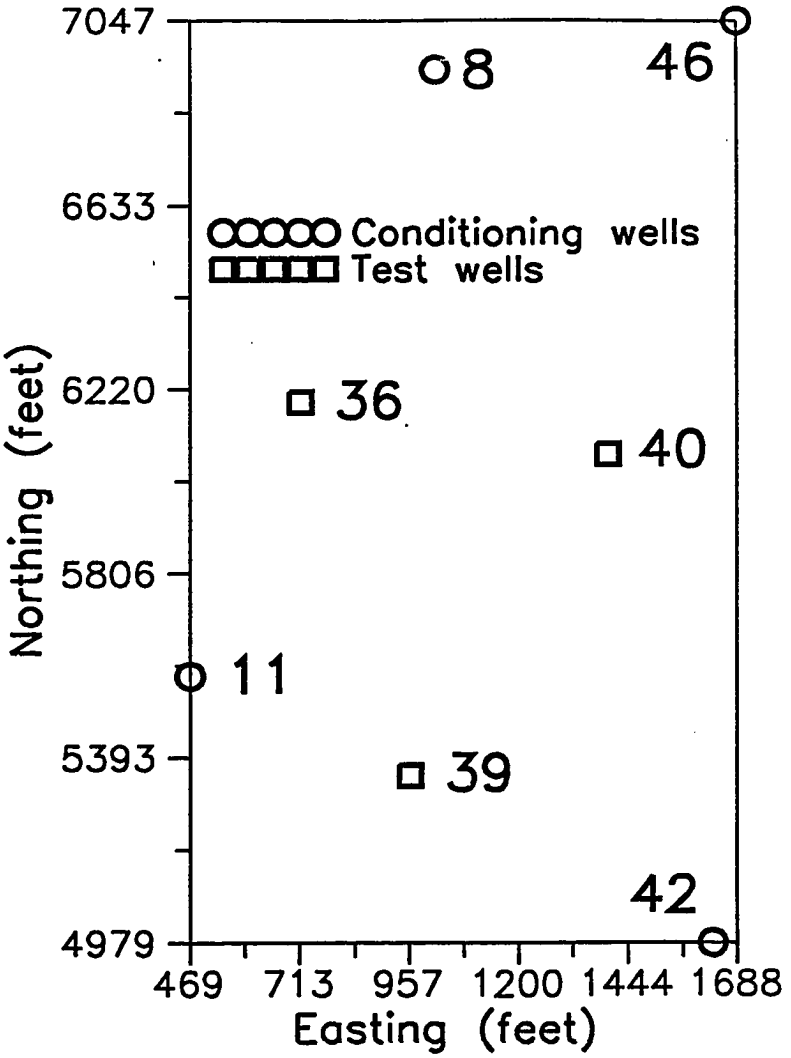

Figure 3.18 - Location of wells in the carbonate field

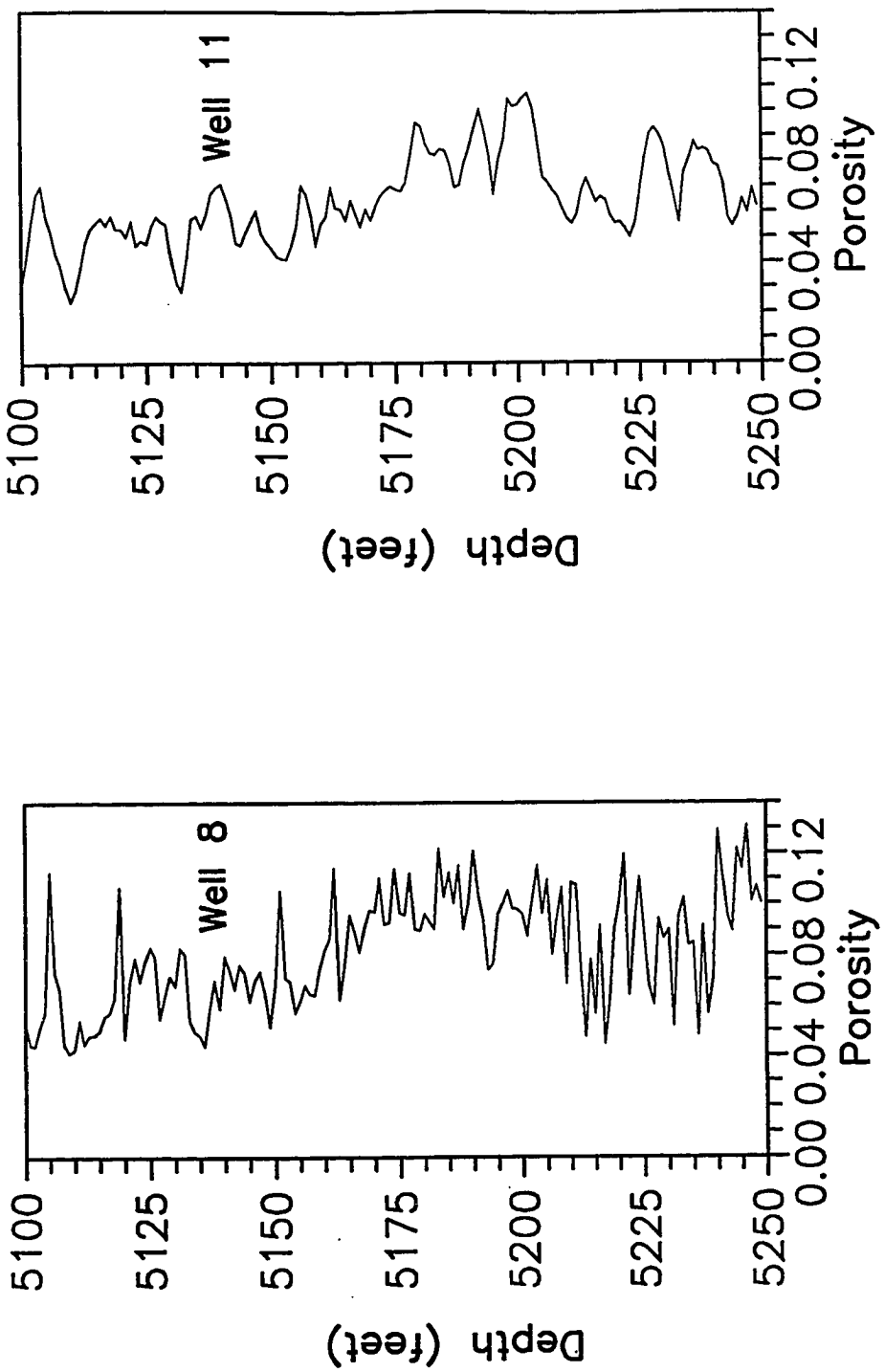

Figure 3.19 - Porosity logs of Wells 8 and 11

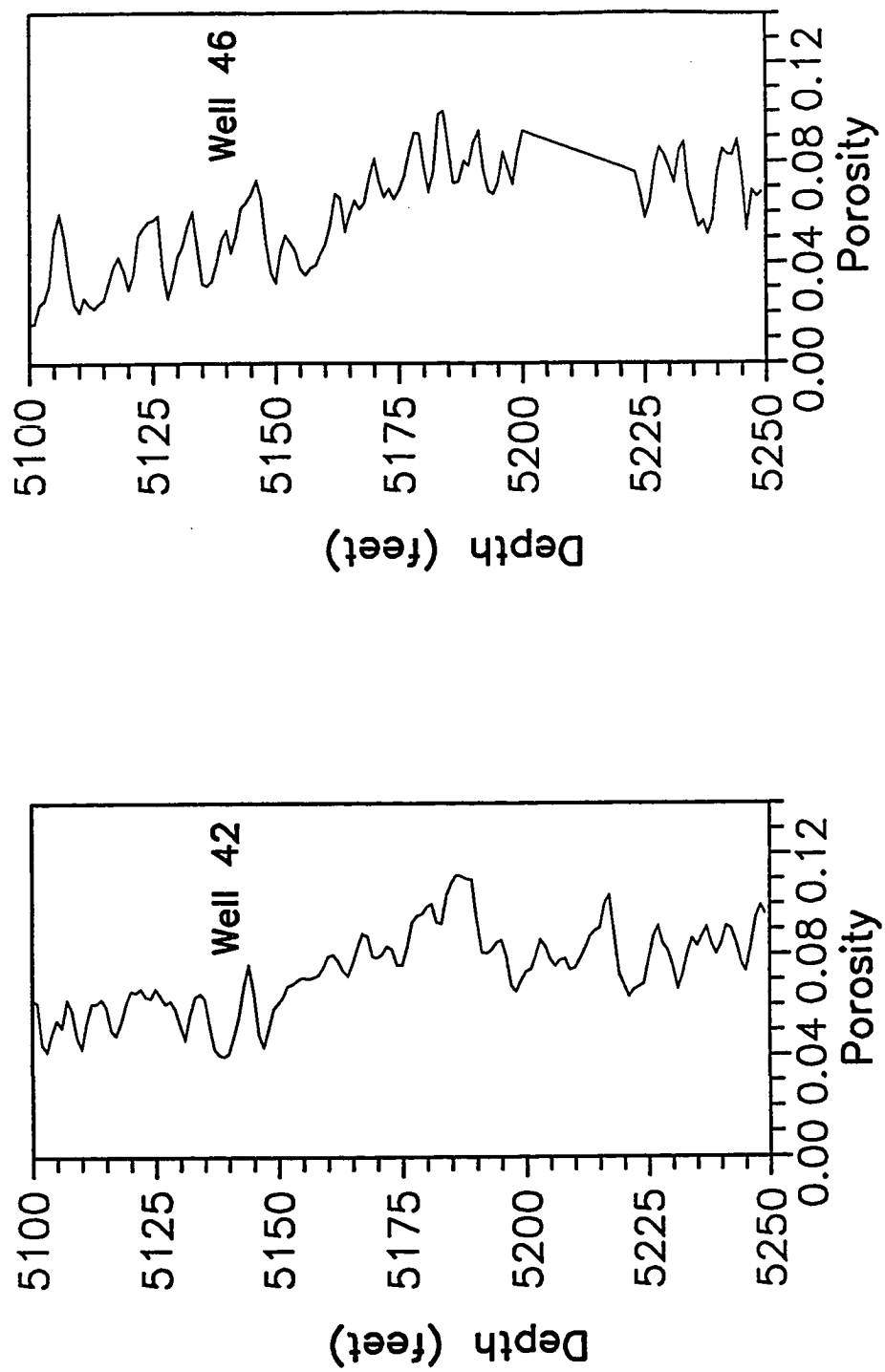

Figure 3.20 - Porosity logs of Wells 42 and 46

Table 3.3

Univariate statistics of porosity (fraction) from the acoustic logs of the conditioning wells in the carbonate field

Parameter	8	11	42	46	All
Number	150	150	150	150	600
Mean	.0812	.0646	.0734	.0620	.0703
Variance	.00049	.00029	.00027	.00045	.00043
Standard Deviation	.0221	.0172	.0165	.0212	.0208
Coefficient of Variation	.2722	.2668	.2249	.3417	.2959
Minimum	.0412	.0245	.0395	.0155	.0155
Maximum	.1309	.1069	.1116	.1001	.1309
Median	.0816	.0615	.0744	.0658	.0695
1^{st} quartile	.0632	.0544	.0620	.0453	.0557
3^{rd} quartile	.0972	.0741	.0843	.0797	.0846

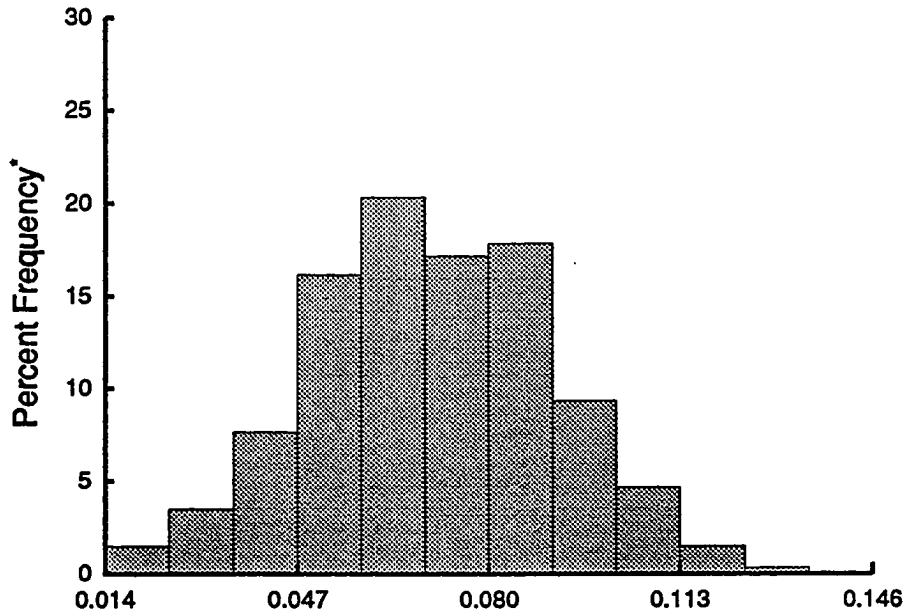

Figure 3.21 - Histogram for porosity logs of the conditioning wells

Table 3.4

Intermittency exponents for the porosity logs of the conditioning wells evaluated with the R/S analysis and the box counting method

Well	R/S	Box Counting
8	0.95	0.92
11	0.98	0.96
42	0.98	0.97
46	0.96	0.97

Well 11 are shown in Figures 3.22 and 3.23 for the R/S analysis and the box counting method, respectively. The results of both methods indicate that the spatial correlation of the porosity logs can be represented by fGn models with intermittency exponents close to one. Except for Well 8, the intermittency exponent of the porosity logs is greater than 0.95. For the four porosity logs, the mean of the intermittency exponents calculated by the R/S analysis is 0.97 and by the box counting method is 0.96.

The semi-variogram along the vertical direction for all the porosity sample data of the conditioning wells and a fGn model are shown in Figure 3.24. The parameters of this fGn model are $H = 0.97$, $V_H = 0.001$ and $\delta = 1$ feet. This fGn model is specified for the vertical direction in the three-dimensional conditional simulations. The semi-variogram models specified for the x and y directions in the conditional simulations are similar to the fGn model for the vertical direction, except that the smoothing factor δ is adjusted according to the grid spacings in the lateral directions. These models simply allow to account for the expected smaller variability as a result of the larger grid spacings in the lateral directions. The significance of using these models for the lateral directions in the conditional simulations is investigated in Chapter IV.

Conditional Simulations

The specifications for the three-dimensional conditional simulations of porosity are given in Table 3.5. The simulation region has been discretized in the vertical direction with 150 grid points in order to include in the simulations all the sample data available at the conditioning wells. For the lateral directions, the grid spacing has been restricted to approximately 50 feet because it yields a total number of grid points close to the practical limitations of the available computational facilities. The conditioning data consist of the porosity logs shown in Figures 3.19 and 3.20. The distribution function and semi-variogram models used in the conditional simulations are described in the Univariate Statistics and the Spatial Statistics sections, respectively. The parameters of the annealing schedule have been selected based on the analysis in the Annealing Schedule section in Chapter II.

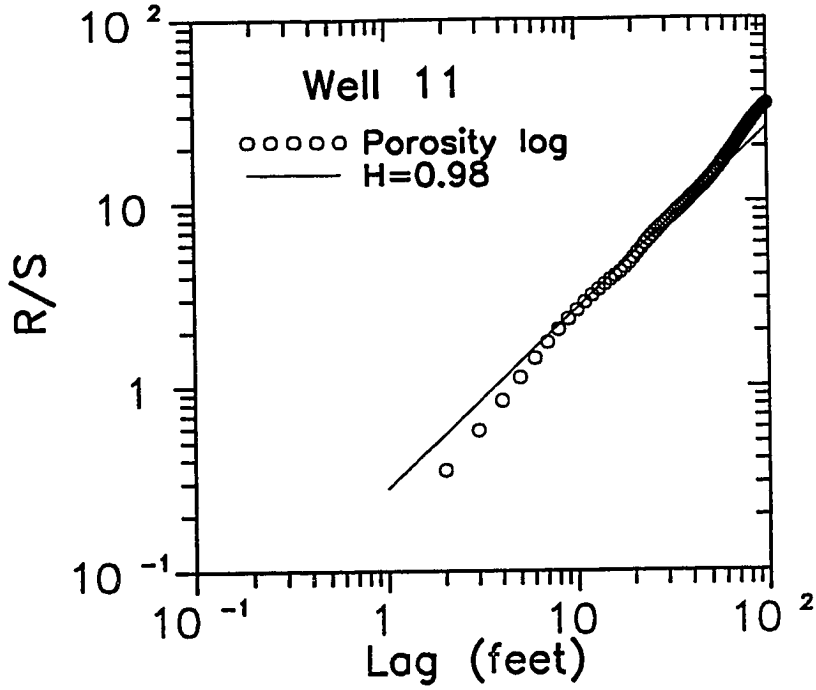

Figure 3.22 - R/S analysis for the porosity log of Well 11

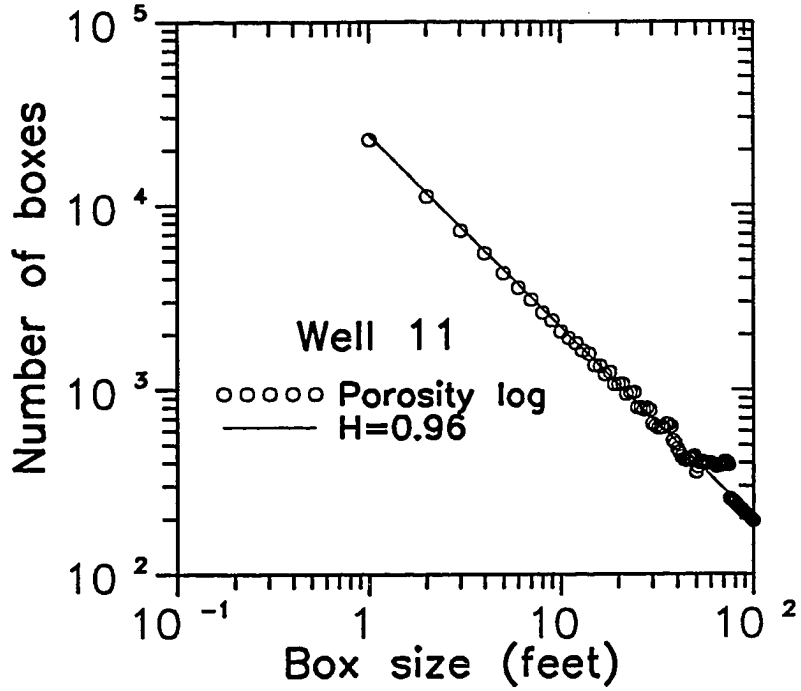

Figure 3.23 - Box counting method plot for the porosity log of Well 11

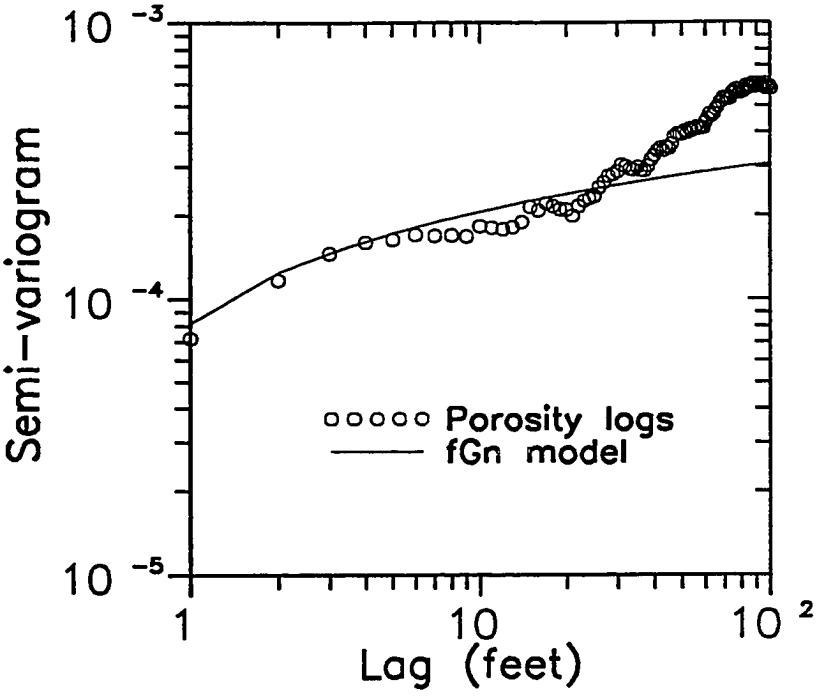

Figure 3.24 - Semi-variogram for the porosity logs of the conditioning wells and fGn model with H=0.97, V_H=0.001 and δ=1.0

Table 3.5

Specifications for three-dimensional conditional simulations of porosity in the carbonate field

Grid Geometry			
Direction	x	y	z
Spacing	50.792	48.093	1.0
Grid points	25	44	150
Total points	165000		
Conditioning Data			
Source		Acoustic logs of Wells 8, 11, 42 and 46	
Number		600	
Distribution Function			
Source		Conditioning data	
Number of classes		14	
Subclass distribution		Uniform	
Semi-variogram Models			
x direction		Equation 3.2, $H = 0.97$, $V_H = 0.001$, $\delta = 1 \times 50.792$ feet	
y direction		Equation 3.2, $H = 0.97$, $V_H = 0.001$, $\delta = 1 \times 48.093$ feet	
z direction		Equation 3.2, $H = 0.97$, $V_H = 0.001$, $\delta = 1 \times 1.0$ feet	
Annealing Schedule			
α		0.50	
M_a		5.0 cycles	
ϵ_v		1.0×10^{-4}	
ϵ_a		2.5×10^{-2}	

The distribution of porosity for one realization of the three-dimensional conditional simulations is illustrated with gray scale maps in Figures 3.25 through 3.27 for x/z, y/z and x/y sections. In these figures, the two-dimensional sections of porosity are uniformly spaced along the corresponding third dimension in the simulation region. Figures 3.25 and 3.26 indicate that porosity is high in the middle and bottom regions of the vertical x/z and y/z sections and that there are a few regions of low porosity in the top of these sections. The maps of the horizontal x/y sections in Figure 3.27 also show that there is a trend of increasing porosity values with increasing depth. For each horizontal section there are regions with connected high porosity values, but there are no apparent similarities in the spatial distributions of porosity among the x/y sections at different depths.

Comparisons between the porosity logs of the test wells and the simulated porosity distributions, used for the last step in the validation procedure, are shown in Figures 3.28 through 3.30 for two realizations of the conditional simulations. The overall agreement between the porosity logs and the simulated distributions is close for the three test wells. There are only a few noticeable departures of the simulated distributions from the porosity logs in the two realizations for Well 39 around 5110 feet and 5245 feet and in the second realization for Well 40 around 5240 feet. For Well 36, the predicted porosity distributions are remarkably close to the porosity log.

Certainty coefficients have been calculated for the two realizations of the porosity distribution at Wells 39 and 40, shown in Figures 3.29 and 3.30. Unfortunately, the sample data at Well 36 cannot be used to calculate the certainty coefficients because this porosity log is not available at the regular one foot spacing used for the conditional simulations; instead there are 137 sample data values irregularly spaced in this 150 feet reservoir interval. The certainty coefficients of the first and the second realizations of the porosity distributions are 0.28 and 0.54. The fact that these certainty coefficients are significantly greater than the ones found for the conditional simulations in Sandstone Field section provides an indication that the fGn models are more appropriate to represent the lateral spatial correlation of porosity in the carbonate field than in the sandstone field.

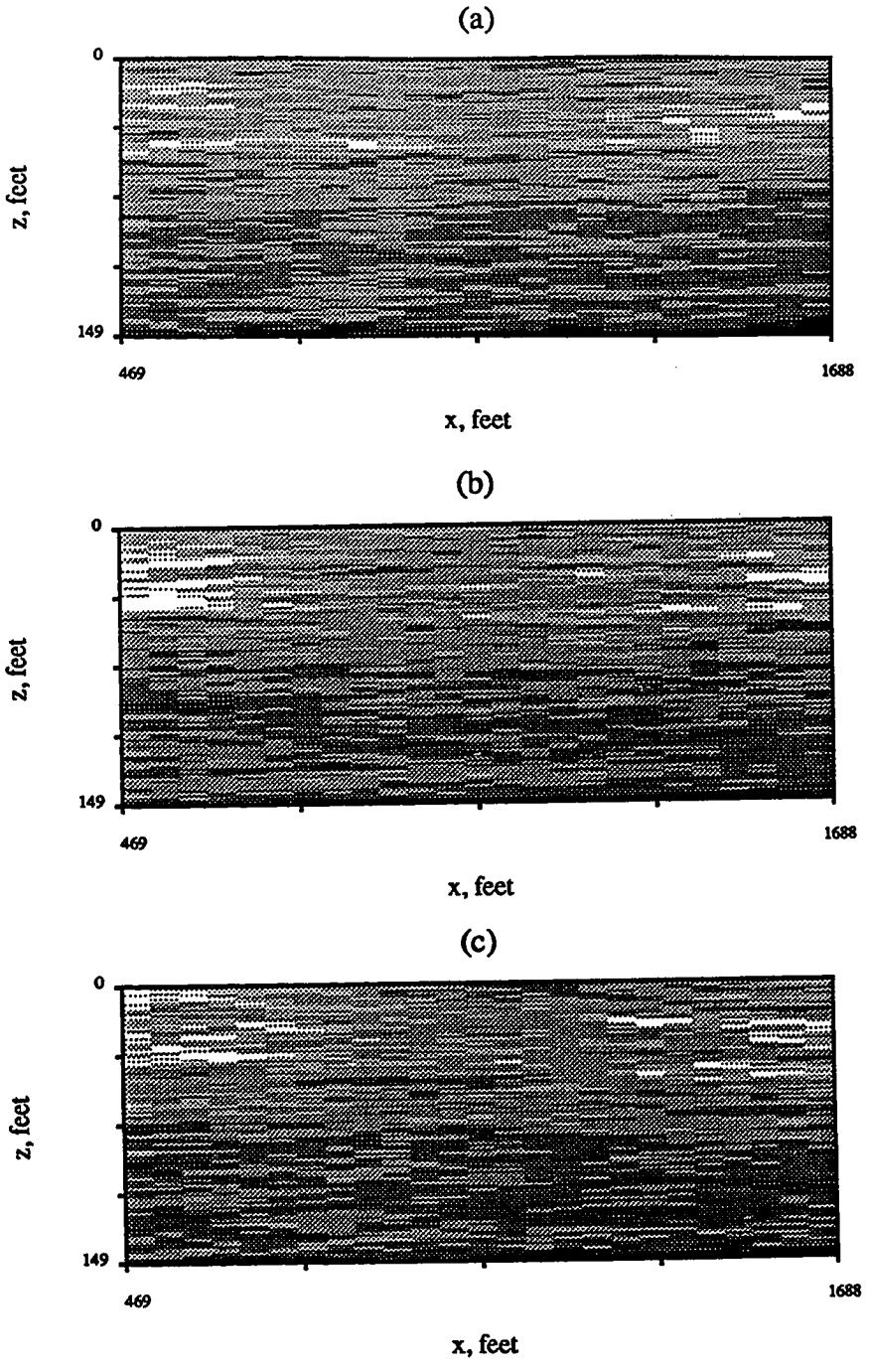

Figure 3.25 - Maps of porosity vertical sections (x/z) from three-dimensional conditional simulations at (a) y=5460 feet, (b) y=5989 feet and (c) 6518 feet

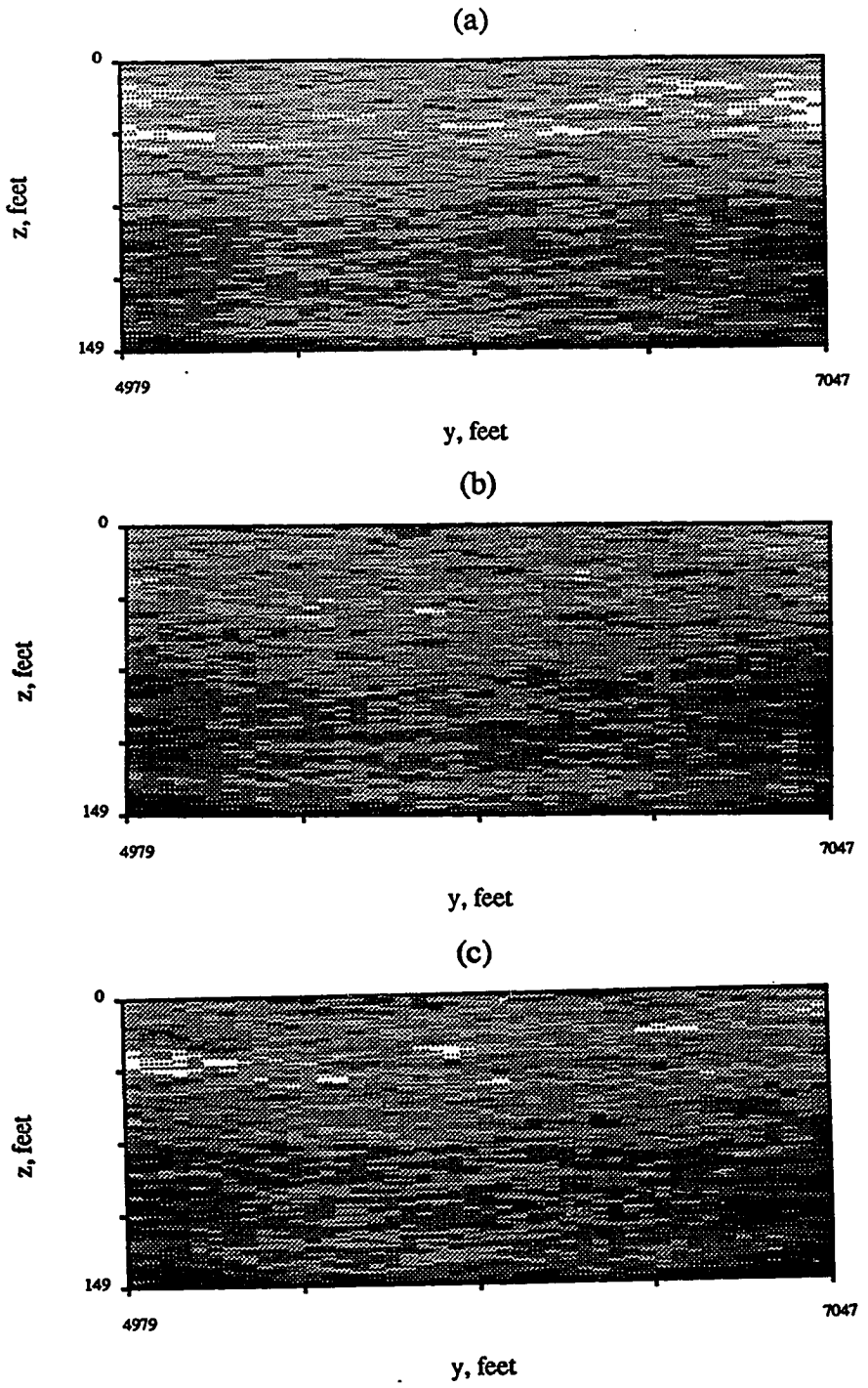

Figure 3.26 - Maps of porosity vertical sections (y/z) from three-dimensional conditional simulations at (a) x=723 feet, (b) x=1028 feet and (c) x=1332 feet

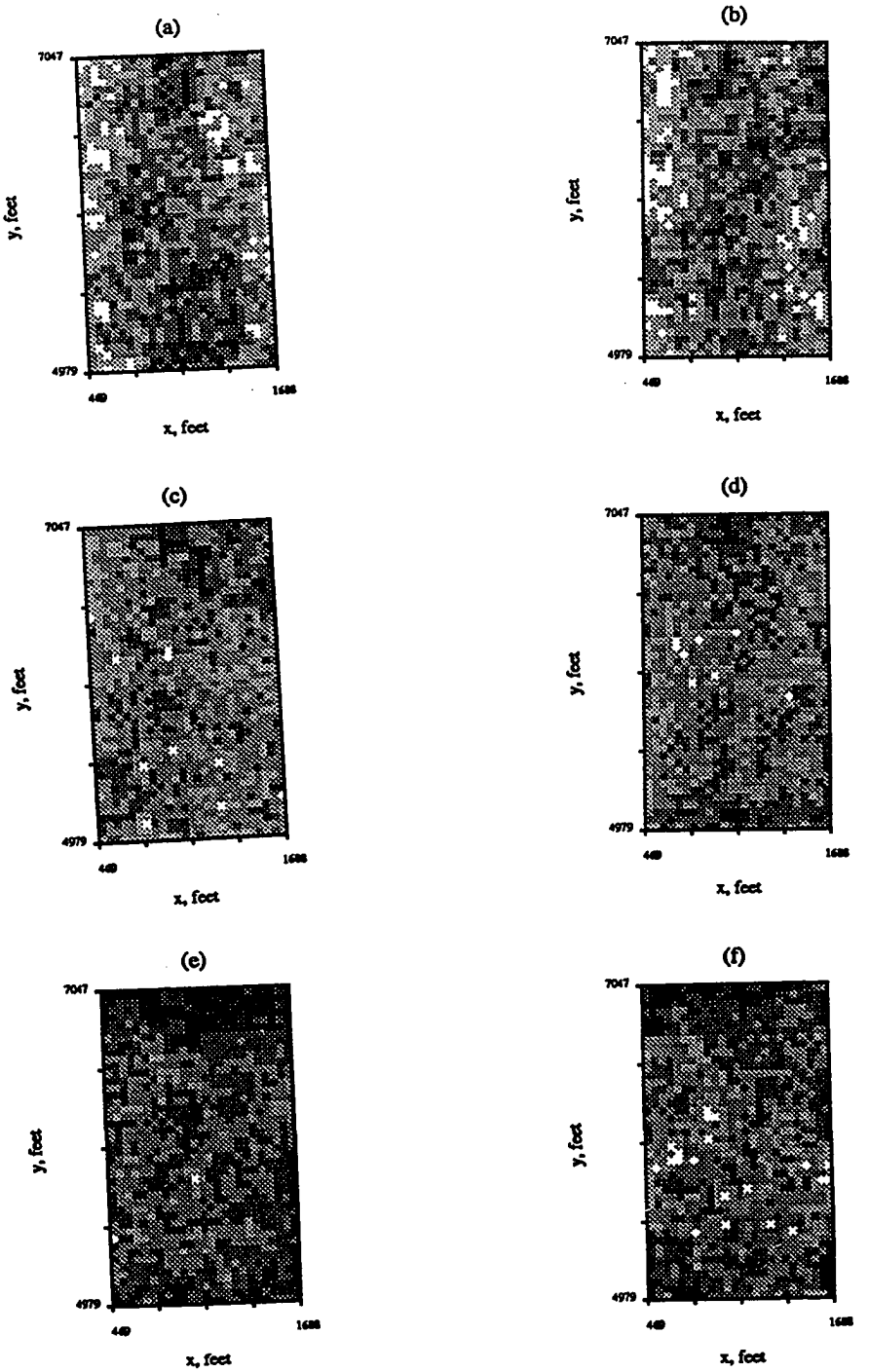

Figure 3.27 - Maps of porosity horizontal sections (x/y) from three-dimensional conditional simulations at (a) z=5120 feet, (b) z=5141 feet, (c) z=5162 feet, (d) z= 5183 feet, (e) z=5204 feet and (f) z=5225 feet

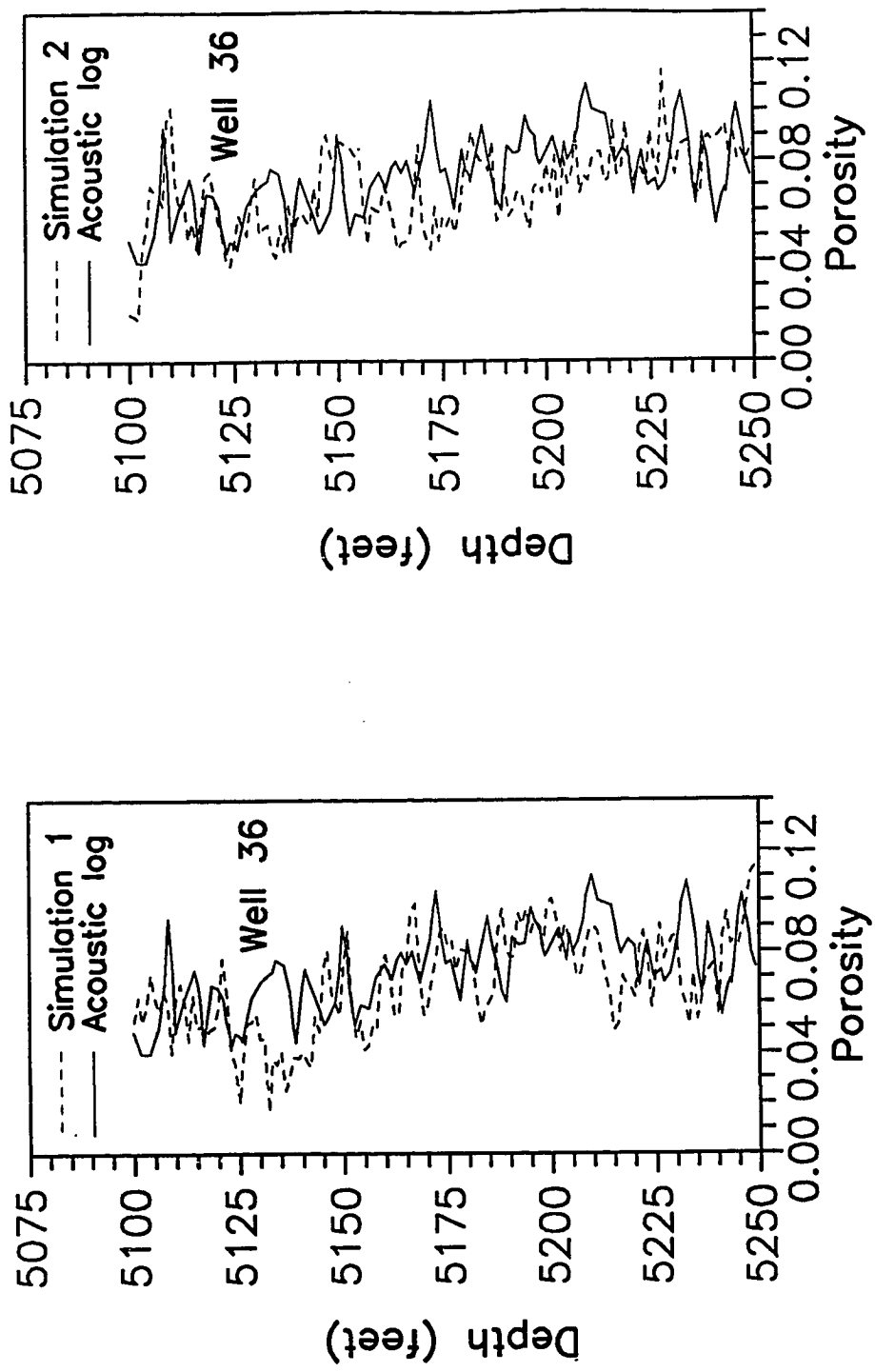

Figure 3.28 - Comparison of porosity log from test Well 36 and two realizations of three-dimensional conditional simulations

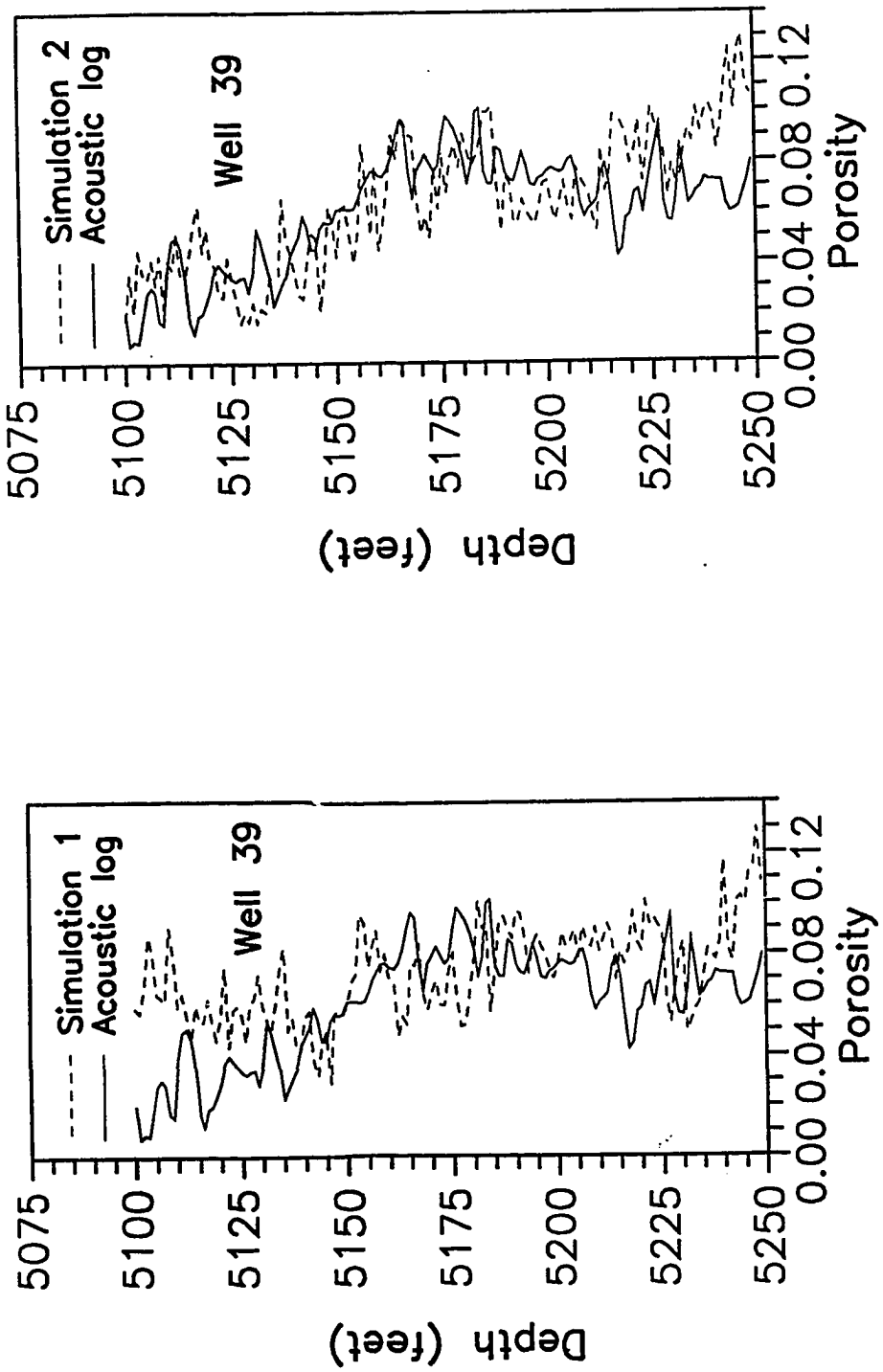

Figure 3.29 - Comparison of porosity log from test Well 39 and two realizations of three-dimensional conditional simulations

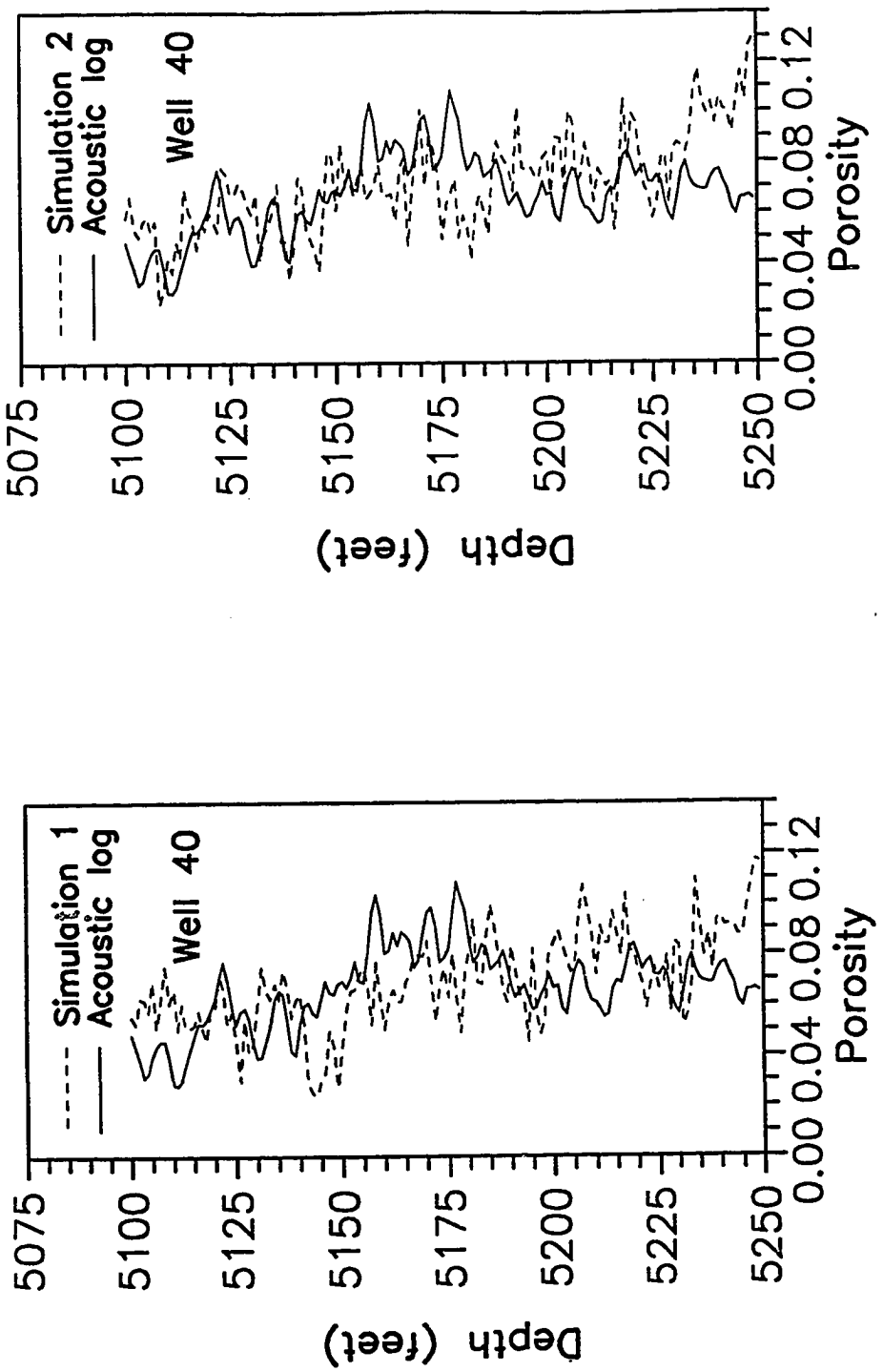

Figure 3.30 - Comparison of porosity log from test Well 40 and two realizations of three-dimensional conditional simulations

Sequential Indicator Simulations

Sequential indicator simulation is a conditional simulation method introduced by Journel and Alabert.[13] This method consists of estimating the posterior probability distribution for specified threshold values of a variable using indicator kriging. The estimated posterior distribution is used to calculate the simulated value using Monte Carlo sampling. The estimations are performed at each point in the simulation region in a sequential mode which follows a random path. At each stage of the simulation, all the values estimated at the previously selected random locations are used to condition the estimation at the present location. The computer program used to generate the conditional simulations in this section with the sequential indicator simulation method has been developed by Gómez-Hernández and Srivastava.[41]

The simulation region in this section is a subset of the region shown in Figure 3.18 and consists of a vertical reservoir section between Wells 11 and 42. In this simulation region, the conditioning wells are Wells 11 and 42 and the test well is Well 39. The specifications for the conditional simulations are given in Table 3.6. The vertical direction of the simulation region has been discretized with 150 grid points in order to maximize the use of the sample data available at the conditioning wells. The distribution function and the semi-variogram models in the vertical direction for the conditional simulation have been derived from the porosity log data of the conditioning wells. The semi-variogram models for the sequential indicator simulations are specified for one threshold equal to the median of the sample data. The parameters of the annealing schedule have the same values as the parameters used for the simulations throughout this chapter.

In order to use the same type of semi-variogram models for the two conditional simulation methods compared in this section, exponential models have been selected. Figure 3.31 shows the semi-variograms along the vertical direction for the sample porosity data of Wells 11 and 42 and for the indicator transform at the median threshold (0.0677) of these sample data. The exponential models shown in

Table 3.6

Specifications for two-dimensional conditional simulations of porosity in the vertical section between Wells 11 and 42 in the carbonate field

Grid Geometry			
Direction	x	y	z
Spacing	13.1818	1.0	1.0
Grid points	100	1	150
Total points	15000		
Conditioning Data			
Source		Acoustic logs of Wells 11 and 42	
Number		300	
Distribution Function			
Source		Conditioning data	
Number of classes		8	
Subclass distribution		Uniform	
Semi-variogram Models			
Simulated Annealing			
x direction		$\gamma(h) = 0.00030[1 - \exp(-3h/320.0)]$	
z direction		$\gamma(h) = 0.00030[1 - \exp(-3h/32.0)]$	
Sequential Indicator Simulation			
x direction		$\gamma(h) = 0.25[1 - \exp(-3h/300.0)]$	
z direction		$\gamma(h) = 0.25[1 - \exp(-3h/30.0)]$	
Annealing Schedule			
α		0.50	
M_a		5.0 cycles	
ϵ_v		1.0×10^{-4}	
ϵ_a		2.5×10^{-2}	

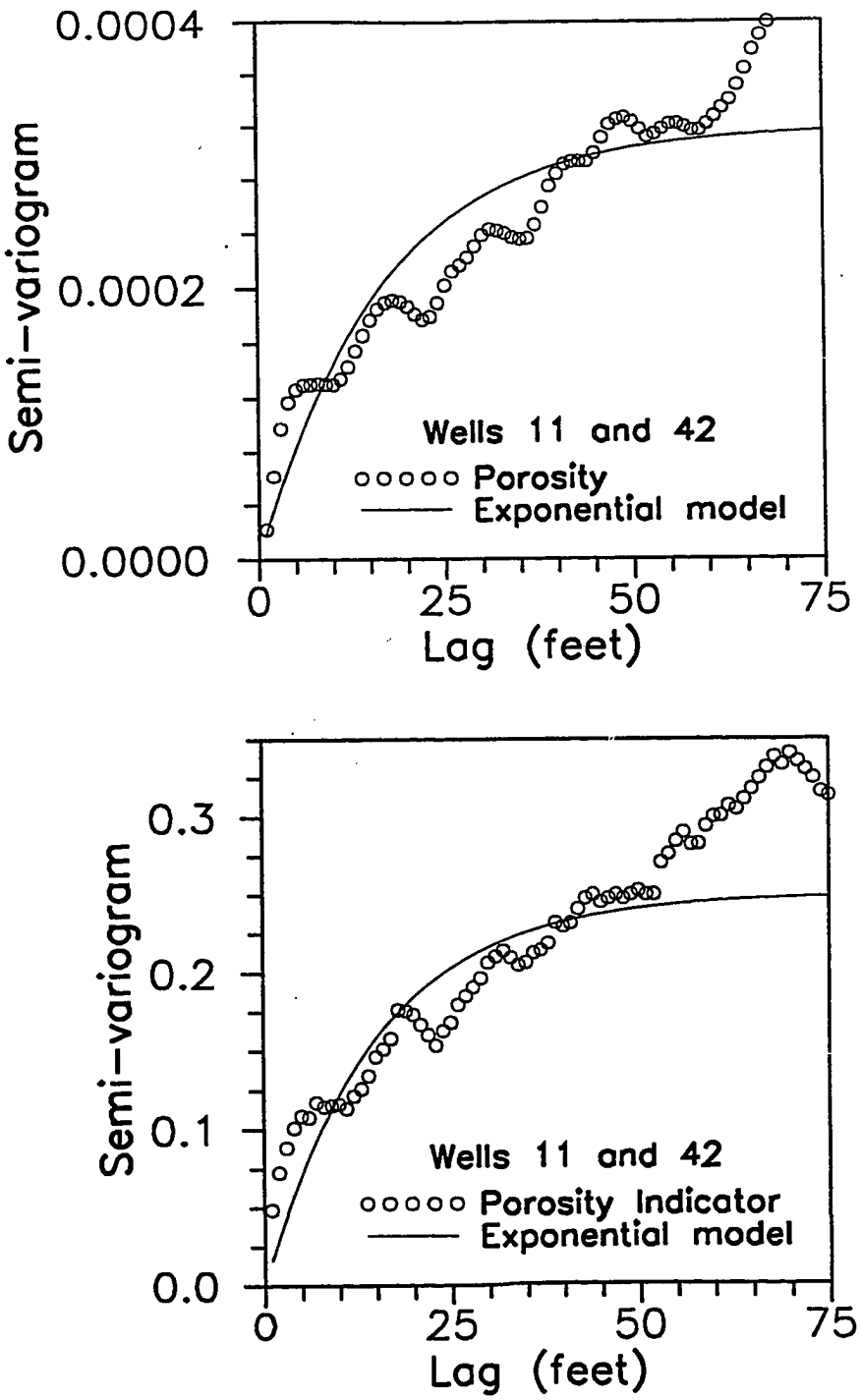

Figure 3.31 - Semi-variograms of porosity logs and indicator transform (median threshold) of porosity logs from Wells 11 and 42

Figure 3.31 are given in Table 3.6. The semi-variogram models for the horizontal direction of the conditional simulations are similar to the models in the vertical direction, except that the practical correlation ranges are set equal to ten times the correlation ranges of the vertical direction, as indicated in Table 3.6. This assumption about the horizontal direction correlation range is a practical approximation when no information is available to evaluate spatial correlation along this direction and the validity of this assumption is discussed in Chapter IV.

The porosity distributions of three realizations generated with the sequential indicator simulation method and the simulated annealing method are shown by the maps in Figures 3.32 and 3.33, respectively. These results indicate that the porosity distributions for both conditional simulation methods have large continuous regions of high porosities in the middle and lower parts of simulation region. The overall smooth appearance of the porosity distributions for the sequential indicator simulations occurs because only one threshold is used in the simulations and the simulated values reflect the underlying smooth characteristics of indicator kriging.

The comparisons between the porosity log of the test well and the simulated distributions for the sequential indicator simulation method are shown in Figures 3.34 and 3.35 and for the simulated annealing method are shown in Figures 3.36 and 3.37. Except for the first realization of sequential indicator method (Figure 3.34), the other simulations in Figures 3.34 and 3.35 provide a reasonably close prediction of the porosity log of Well 39. The certainty coefficients of the results shown in Figures 3.34 through 3.37 are summarized in Table 3.7. For these three realizations of the porosity distribution, the mean certainty coefficient is greater for the simulated annealing method than for the sequential indicator simulation method but the range of certainty coefficients is greater for the sequential indicator simulation method. The results of this limited number of realizations suggest that the simulated annealing method yields a high level of certainty and the sequential indicator simulation method provides a wider range of equiprobable distributions. It should be noted that the results of the sequential indicator simulation method may improve if the indicator semi-variograms are specified in the simulations for more than one threshold.

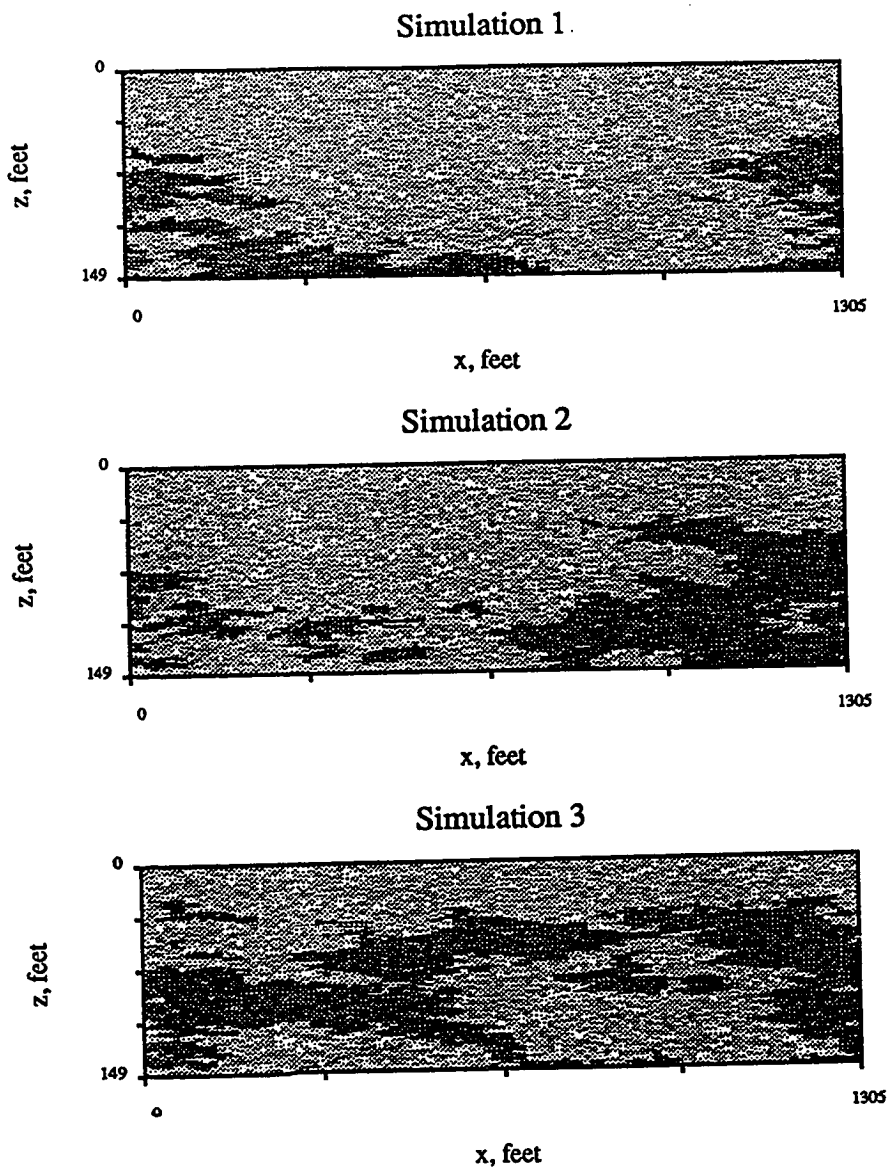

Figure 3.32 - Maps of porosity for three realizations from the sequential indicator simulation (SIS) method

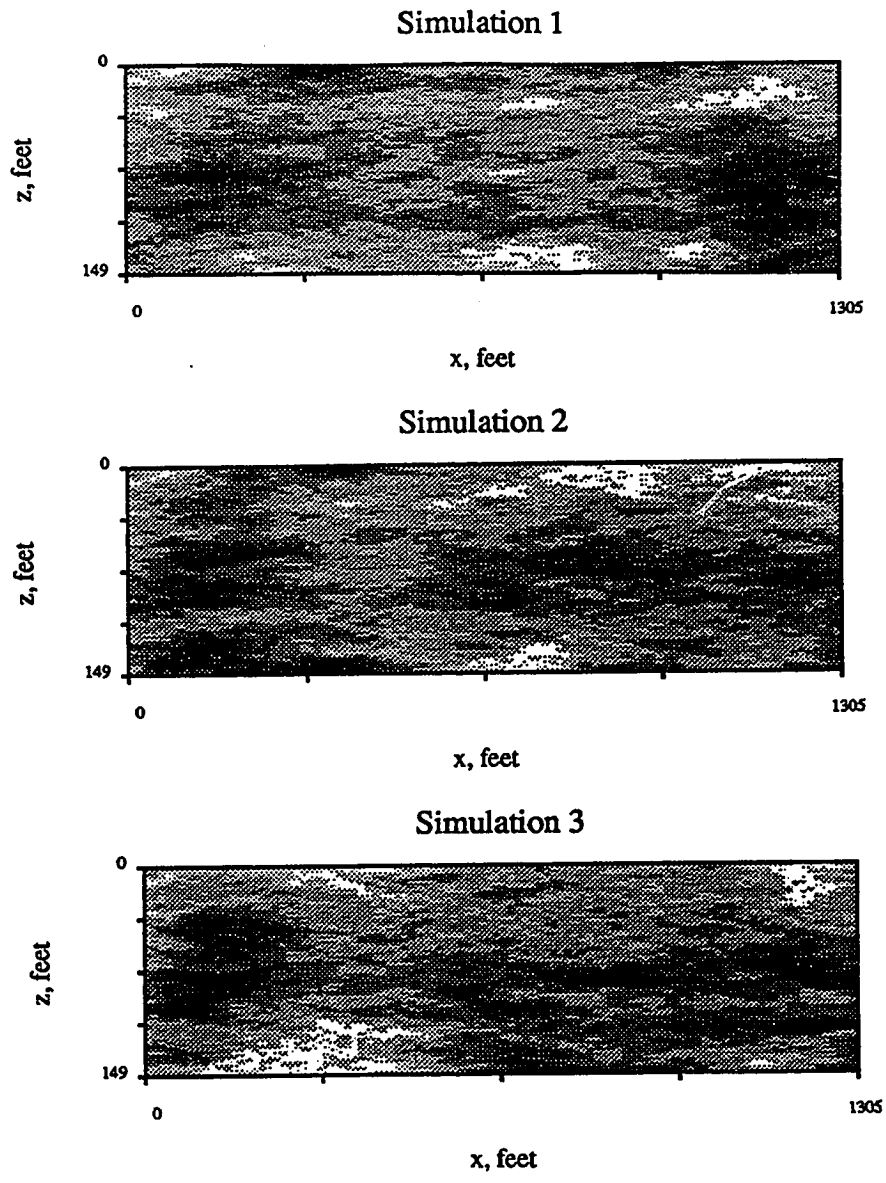

Figure 3.33 - Maps of porosity for three realizations from the simulated annealing (SA) method

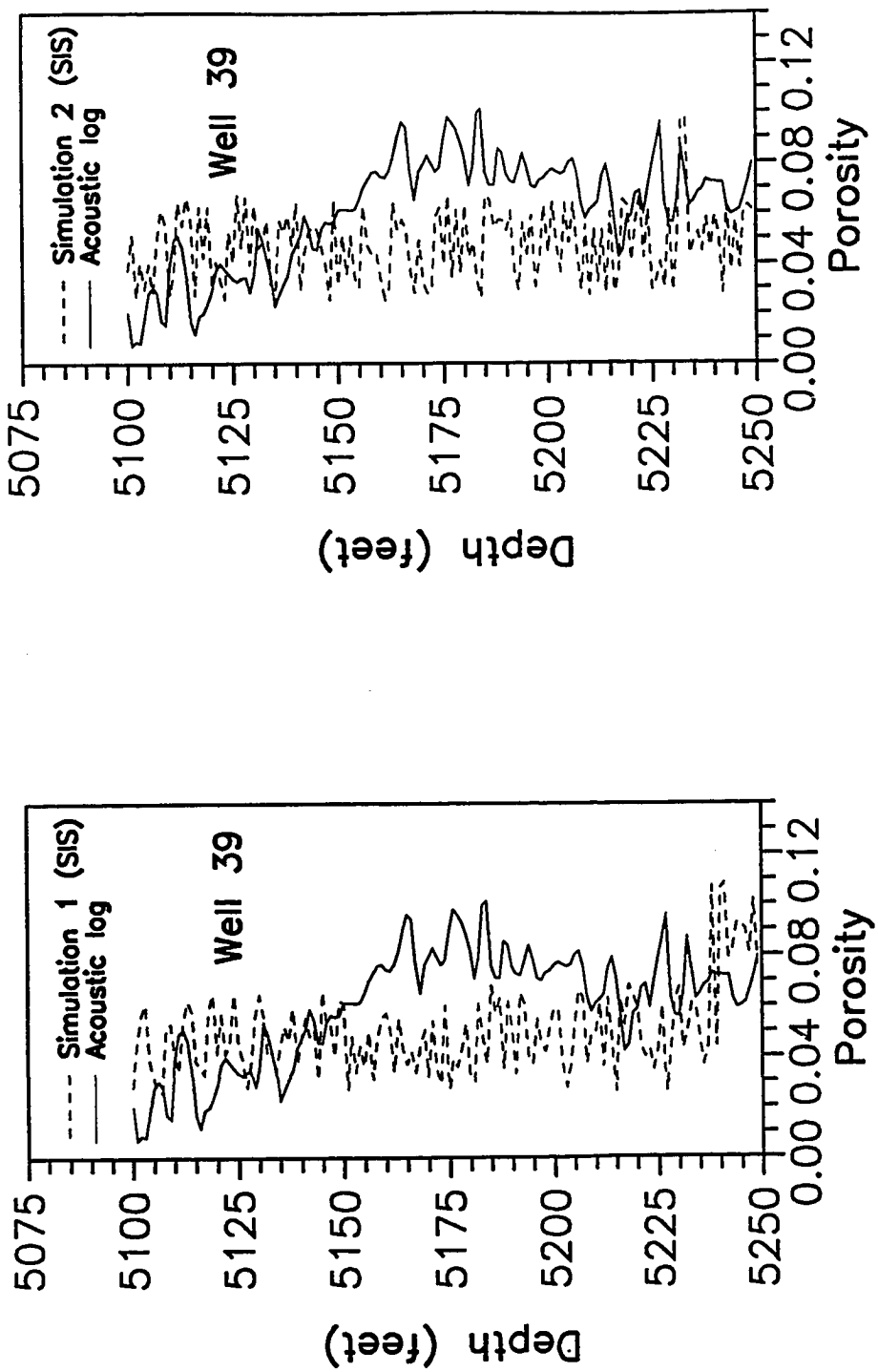

Figure 3.34 - Comparison of porosity log from test Well 39 and first and second realizations from the sequential indicator simulation (SIS) method

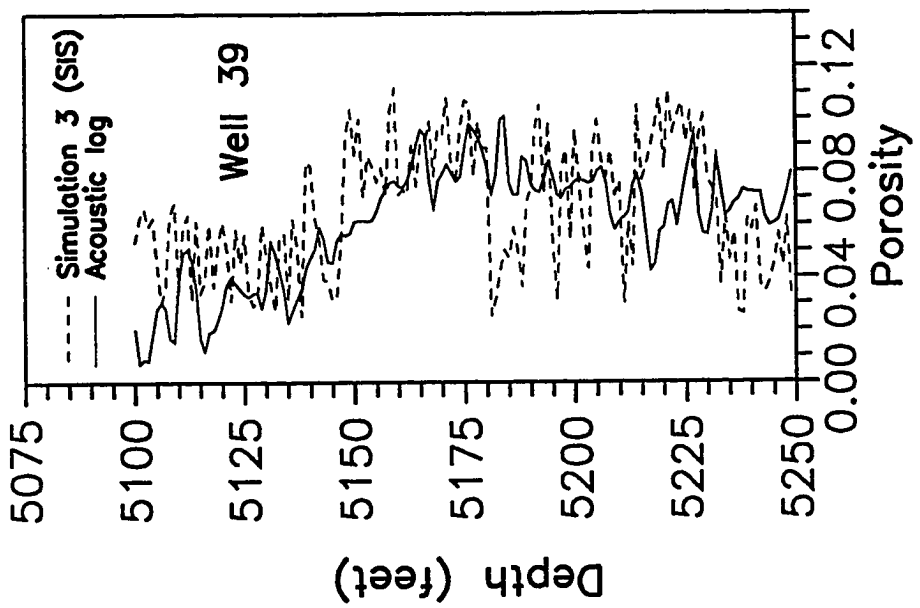

Figure 3.35 - Comparison of porosity log from test Well 39 and third realization from the sequential indicator simulation (SIS) method

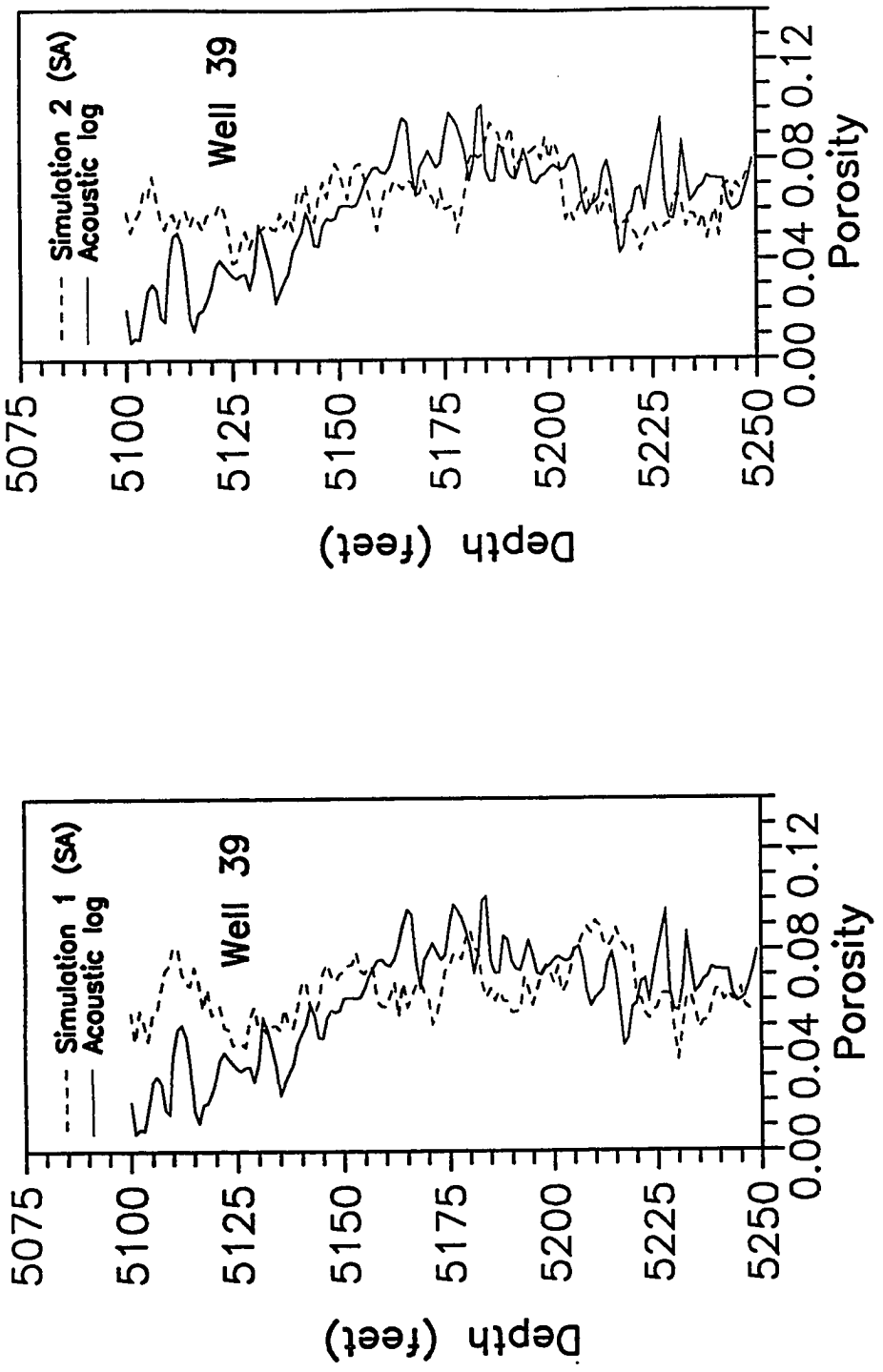

Figure 3.36 - Comparison of porosity log from test Well 39 and first and second realizations from the simulated annealing (SA) method

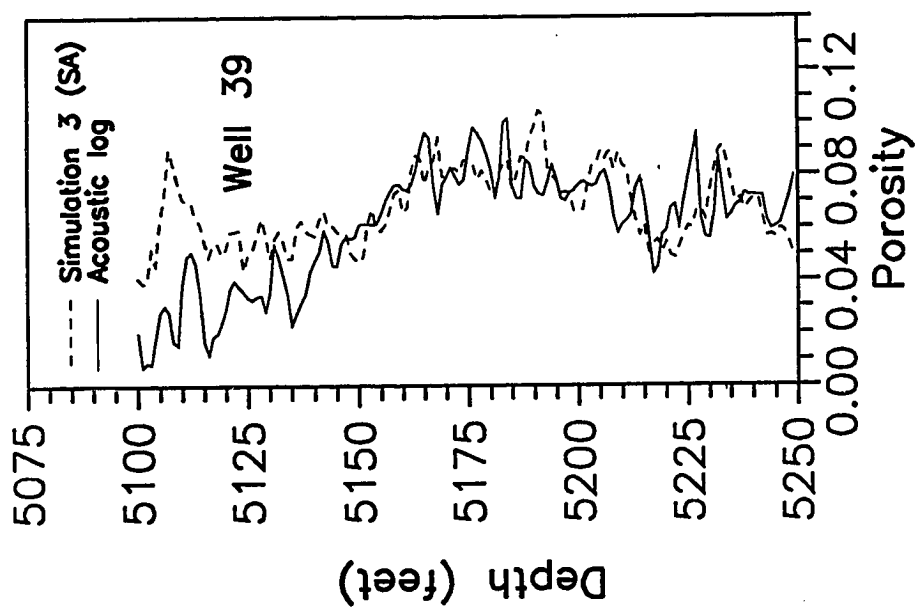

Figure 3.37 - Comparison of porosity log from test Well 39 and third realization from the simulated annealing (SA) method

Table 3.7

Certainty coefficients for porosity log of the test well and different realizations of distributions generated by simulated annealing and sequential indicator simulation methods

Realization Number or Statistic	Simulated Annealing	Sequential Indicator Simulation
1	0.364	0.0956
2	0.475	0.142
3	0.624	0.432
Mean	0.488	0.223
Range	0.260	0.336

The evaluation of the semi-variograms for the simulated distributions indicates that the sequential indicator simulation method does not reproduce the specified models as close as the simulated annealing method. Comparisons between the sample semi-variograms of a median indicator distribution generated with the sequential indicator method and the specified exponential models are shown in Figure 3.38, for the horizontal and vertical directions. The comparisons in Figure 3.38 indicate that the magnitudes of the sample semi-variograms of the simulated indicator distribution are smaller than the specified models and supports the observation that the distributions generated with the sequential indicator simulation method have a smooth appearance. For the simulated annealing method, Figure 3.39 shows that the semi-variograms of the simulated distributions are in close agreement with the specified models.

In this carbonate field two- and three-dimensional conditional simulations were conducted with different models for the horizontal spatial correlation of porosity. The results show that the quality of the tests for the two-dimensional simulations is slightly better than for the three-dimensional simulations. This observation is an indication that the exponential model used in the two-dimensional simulations provide a better representation of the horizontal spatial correlation of porosity than the fGn model used in the three-dimensional simulations. An explanation for these results is that the exponential model properly represents the small support volume of the logs of the test wells. The fGn model accounts for the large support volume used in the lateral direction of the simulations which is significantly greater than the support volume of the logs. This effect of the support volume in the spatial correlation models is investigated in the Conditional Simulations section of Chapter IV.

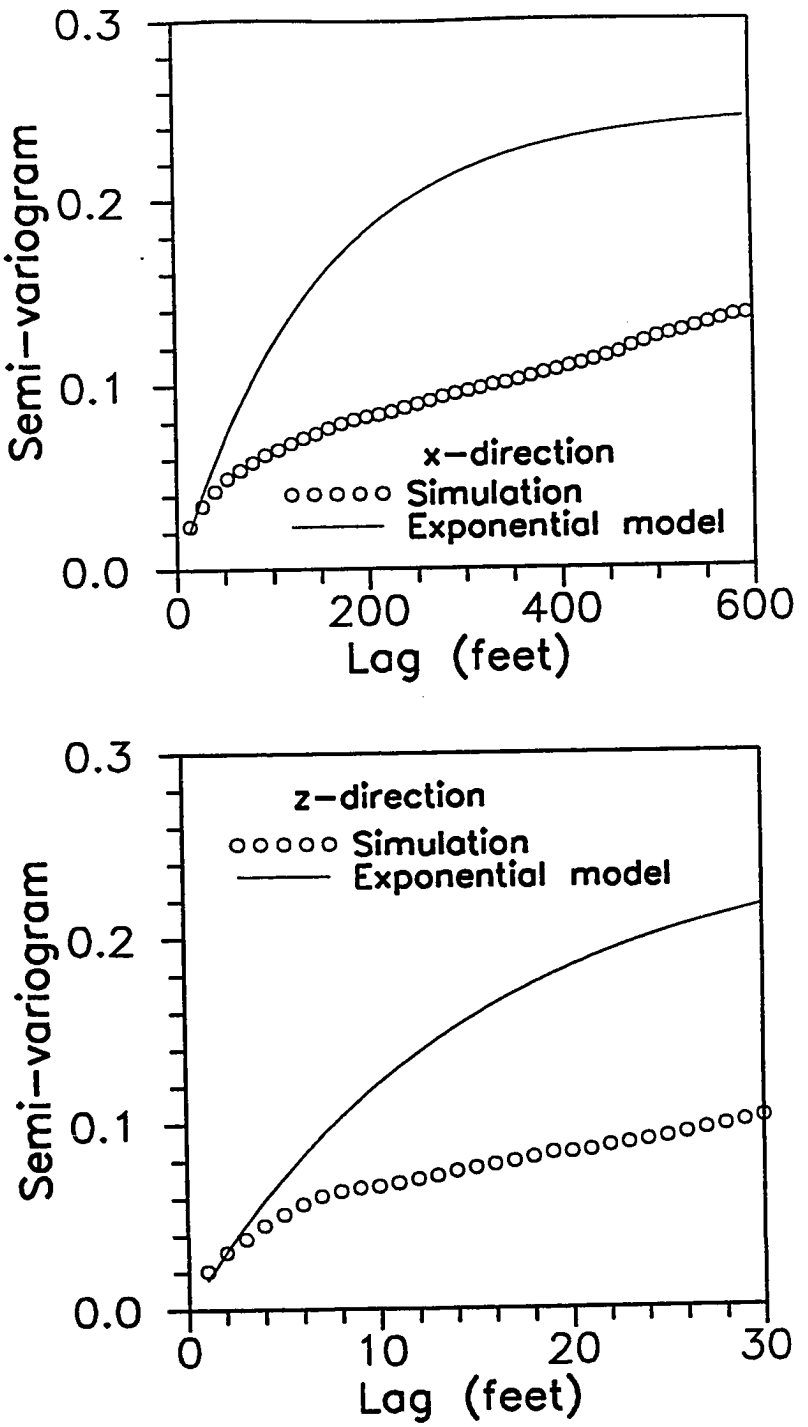

Figure 3.38 - Specified and simulated semi-variograms of porosity median indicator transform distribution generated by the sequential indicator simulation method for the x and z directions

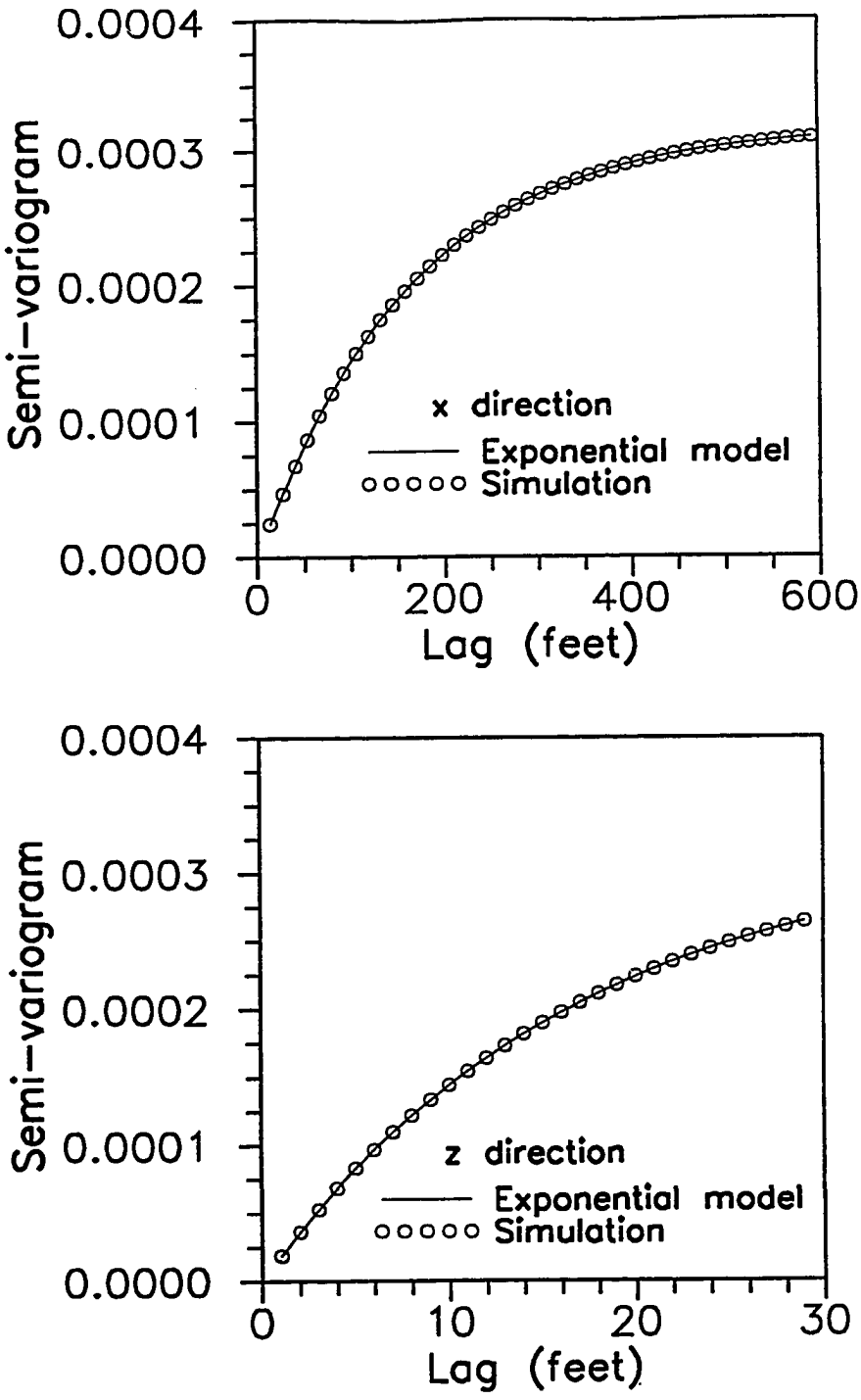

Figure 3.39 - Specified and simulated semi-variograms of porosity distribution generated by the simulated annealing method for the x and z directions

CHAPTER IV

DISTRIBUTIONS OF RESERVOIR PROPERTIES

This chapter presents a comprehensive geostatistical evaluation of the distribution of porosity in a carbonate reservoir using logs of a horizontal well and several closely spaced vertical wells. In many field cases, such as the field validations in Chapter III, a significant part of the uncertainty in the descriptions of reservoir properties can be due to the limited information about the nature of the spatial distribution of the properties in the inter-well regions. However, as the results of the analyses in Chapter II indicate, the certainty levels of the descriptions increase if adequate information about the spatial distribution is supplied in the conditional simulations. Therefore, these results are a further incentive to investigate the statistical character of the distribution of reservoir properties. The major contribution of this chapter is to provide guidelines to assess information about, rarely available, small scale variability of reservoir properties in inter-well regions, from the widely available measurements at vertical wells.

The first section of this chapter contains a background about the needs and problems in the spatial description of properties for reservoir engineering applications and reviews the most recent investigations pertinent to the evaluation of spatial statistics of reservoir properties. The Study Region section provides a brief description of the geology and the sources of well data for the carbonate field discussed in this chapter. The following sections contain the univariate and spatial statistical analyses of porosity and highlight the comparisons among the statistical attributes of porosity for the horizontal and the vertical well logs. The analysis is extended to investigate the effect of the sample volume size on the spatial statistics. Finally, three-dimensional porosity distributions generated with the conditional simulation method developed in Chapter II using different spatial correlation models derived from the vertical well data are evaluated, and compared to the observed porosity distributions along the horizontal direction.

Background

Proper quantification of the spatial distribution of properties and associated uncertainties is important for most reservoir engineering applications. For example, descriptions of reservoir properties are needed for oil-in-place evaluation, displacement mechanisms models and infill wells selection, among other engineering applications. As a result, a major goal of recent reservoir characterization practices has been to develop methods to describe the spatial distribution of properties which incorporate the observed heterogeneity and variability due to the geologic complexities. Geostatistical approaches allow to quantify the spatial variability of a property and to generate descriptions at the unsampled inter-well regions which conform with the observed statistical attributes of a property.

In several field cases, porosity logs (Neutron, Density and Acoustic) are the only source available to estimate the distributions of reservoir properties. A conventional approach to develop descriptions of properties is to divide the reservoir into layers of constant properties and/or to include lateral variability by interpolation of "average" values from the well logs. The assumptions behind this conventional approach can result in significant errors because it ignores the geologic complexities, the inherent heterogeneity of reservoir properties, and the uncertainties associated with the relatively small amount and irregular density of the sample data.

Distribution of properties, such as porosity and permeability, are affected by depositional and post-depositional geologic processes and must possess certain spatial correlation or continuity in order for the hydrocarbon accumulations to be productive reservoirs. This nature of the geologic processes controls the spatial continuity of properties along different directions in the reservoir.

Information about the magnitude and variation of properties at small scales in the inter-well regions of reservoirs is limited because sources of direct measurements, such as horizontal wells and seismic, are not commonly available. Only a few studies, based on data from outcrops and horizontal wells, provide analyses and comparisons of the statistics and the spatial correlation of properties along different

directions in reservoirs. Measurements of rock properties in outcrops may provide a viable alternative to overcome the difficulties of sampling reservoirs along different directions.

The work of Smith[42] includes a statistical analysis of core measurements of porosity, permeability, rock compressibility and grain size in a stratified, unconsolidated sandstone outcrop. The sampling configurations in this study include transects parallel and perpendicular to the direction of stratification. The observations of Smith[42] indicate that the means of porosity and permeability in the transects parallel and perpendicular to the direction of stratification are close. However, the standard deviations of porosity and permeability for the perpendicular transect are about two times greater than for the parallel transect. The spatial correlation analysis in Smith[42] investigation indicates that the autocorrelation functions of porosity and permeability are similar for the parallel transect, but these have different character for the perpendicular transect. Even though the character of the autocorrelation functions are different for the parallel and the perpendicular transects, the extent of the correlation is of the same order of magnitude for these two transects.

Goggin et al.[43] presented a geologic description and a statistical analysis of permeability for an eolian sandstone outcrop. The permeability sample data in this study consisted of core calibrated minipermeater measurements in five concentric grids on the face of the outcrop. The dimensions of the grids ranged from a few feet to hundreds of feet. The observations of Goggin et al.[43] indicate that permeability distribution varies for different stratification types in this eolian outcrop and the coefficient of variation of permeability follows an irregular trend as a function of the area of the concentric grids. The correlation ranges of permeability for the horizontal and vertical directions observed in the spatial analysis are significantly different for the larger and the smaller sampling grids. The ratios of the horizontal to vertical correlation ranges for the largest and the intermediate grids are 1.5 and 3.0, respectively, while, for the smallest grid this ratio is 1/6. Goggin et al.[43] explain that these differences are a result of the distinct character of the heterogeneities for the different types of stratification.

Kittridge et al.[44] conducted a study to evaluate the statistical properties of permeability in an outcrop of a carbonate formation. In addition, the observations from the outcrop were compared to statistical analyses of core data of a field producing from the same formation as the outcrop. The permeability sample data for the outcrop consisted of minipermeater measurements in a large grid (100 by 80 feet), six small grids within the large grid and a vertical transect. The univariate statistical analysis of permeability indicates that mean from the outcrop data differs by several orders of magnitude from the field measurements but the coefficients of variation are similar for these two sources. On the spatial analysis of permeability, Kittridge et al.[44] found that the correlation range of the vertical transect is in a close agreement with the average correlation range of the core data from several wells in the field. For the field data, the correlation range of permeability and porosity are similar.

More recently, Crane and Tubman[24] evaluated the spatial correlation of density, neutron and sonic logs of horizontal and vertical wells with fractal models in a carbonate and a sandstone reservoir. In both reservoirs, Crane and Tubman[24] used models of fractional Gaussian noise processes to represent the spatial correlation of the logs from horizontal and vertical wells. For the carbonate reservoir, the intermittency exponent of the logs for three horizontal wells range between 0.85 and 0.93, and for four vertical wells range between 0.88 and 0.89. For the sandstone reservoir, the average intermittency exponents of the logs for five horizontal wells is 0.90 and for seven vertical wells is 0.83. A conclusion of this investigation[24] is that the depositional and post-depositional environments appear to influence the differences or similarities between the intermittency exponents of the horizontal and vertical well logs in a similar manner. According to Crane and Tubman,[24] for the uniform marine depositional environment of the carbonate reservoir, the values of the horizontal and vertical intermittency exponents are similar, while for the more complex braided-stream environment of the sandstone reservoir, the horizontal and vertical intermittency exponents are different.

Study Region

This section provides a brief geologic description of the carbonate reservoir studied in this chapter. Then, the relative locations of the wells within the study region and the available well data are described.

Geology

The data set used in this investigation consists of porosity logs from eight wells in a tight carbonate reservoir. The predominant lithology of the reservoir is dolomite. The oil producing formation in this reservoir is very complex and heterogeneous due to the post-depositional processes, such as cementation and several re-dolomitization cycles. Other heterogeneities observed in the cores include cement filled Burrows and Vugs. As a result, the formation is a very tight rock (low porosity and permeability) and includes non-porous intervals.

The producing formation has an average thickness of fifty feet and it is overlayed by a dense shale (the oil source rock). At the bottom of the producing formation, a tight shale streak is present in many wells which probably prevented the migration of oil to a high porosity region located immediately below.[45]

Well Data Set

The horizontal well and the surrounding vertical wells used in this study are within a surface area of approximately one square mile. Figure 4.1 shows a map of the location of all the wells. The relative position and length of the horizontal well log with respect to the logs of the vertical wells can be seen in the vertical profile shown in Figure 4.2. The horizontal section of the horizontal well (dotted line in Figure 4.2) is 1481 feet long and it has been selected by ensuring that the angle of inclination at each point does not deviate by more than five degrees from a horizontal reference. The vertical wells logs vary in length from 365 feet (Well S5) to 721 feet (Well S2) and these logs include reservoir regions above and below the oil production formation.

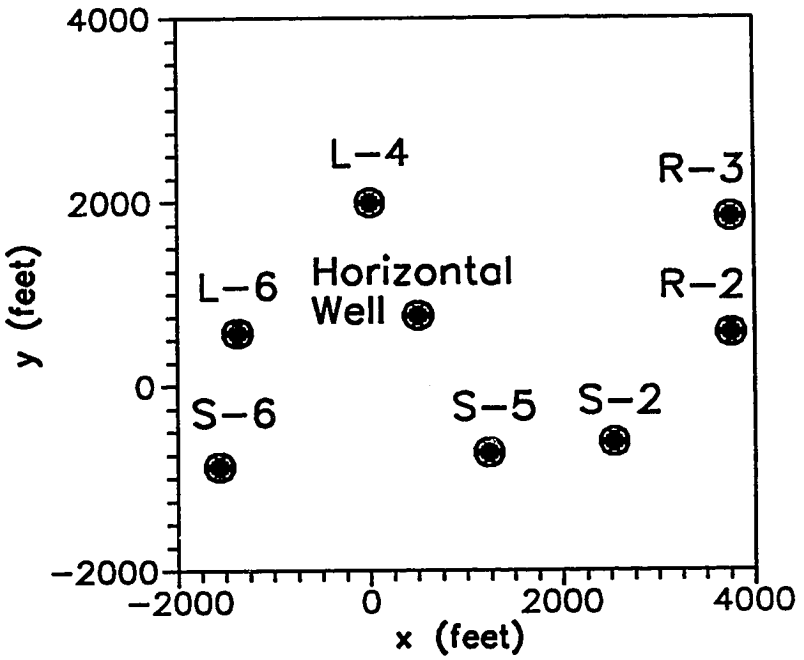

Figure 4.1 - Map of well locations in the study region

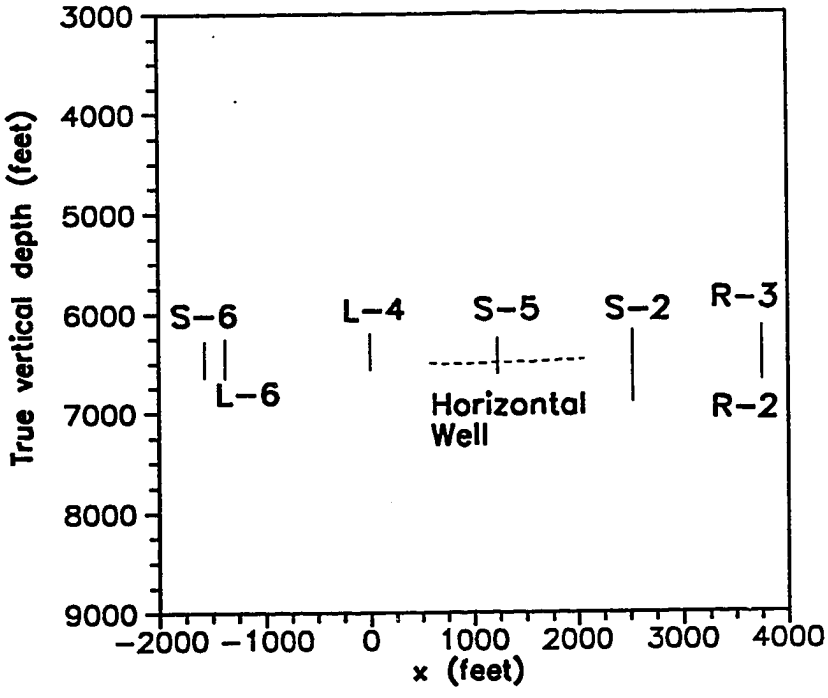

Figure 4.2 - Vertical profile of wells in the study region

The porosity logs for all wells have been derived from Neutron and Density logs at a one foot spacing. These porosity values include the conventional log corrections, such as lithology and environmental corrections.

Analysis of Univariate Statistics

The univariate statistical parameters used in the analysis of the log sample data are the mean, the variance, the coefficient of variation (standard deviation divided by mean), the range, the median, and the quartiles. Table 4.1 summarizes the parameters estimated from the porosity logs of the horizontal well, all combined vertical wells and the individual vertical wells. Figure 4.3 shows the porosity log of the horizontal well and Figures 4.4 through 4.7 show the porosity logs of the vertical wells. In order to facilitate the comparison of the relative porosity variability among the wells, all the logs in Figures 4.3 through 4.7 are standardized to zero mean and unit variance.

Overall mean porosity of all the well logs is low because of the tight nature of the reservoir. The mean porosity of all the vertical well is greater than the mean porosity of the horizontal well. The mean porosity of only two vertical wells (Wells L6 and S5) is lower than the mean porosity of the horizontal well. The mean porosity of the horizontal well is low because it intercepts a large fraction of non-porous regions (zero porosity measurements) as indicated in Figure 4.3 and as reflected in the low values of the first quartile and the median. The first quartile of porosity for all the vertical wells is higher than for the horizontal well.

The porosity variance of the log from the horizontal well is smaller than the variance of all the combined data of the vertical wells. However, the variance of the horizontal well is within the range of variances of individual vertical well logs. Referring to Figure 4.1, it can be noted that the three wells (Wells L4, L6, and S6) with a lower variance are on the west side of the horizontal well, while the three wells (Wells R2, R3 and S2) with a higher variance are in the east side of the horizontal well. The remaining well, Well S5, has a porosity variance smaller than the variance of the horizontal well and it is located between the groups of wells with

Table 4.1

Univariate statistics of porosity logs

Well name	Number	Mean (%)	Variance ($\%^2$)	Coefficient of variation	Range* (%)	Median (%)	1^{st} quartile (%)	3^{rd} quartile (%)
Horizontal	1481	1.389	3.518	1.350	9.780	0.380	0.000	2.470
All vertical	3183	1.752	5.392	1.326	17.74	1.060	0.360	2.180
L4	371	1.470	1.771	0.905	8.000	1.240	0.540	2.010
L6	401	1.243	2.008	1.140	9.930	0.860	0.360	1.510
R2	521	1.905	3.791	1.022	11.19	1.390	0.620	2.540
R3	434	2.057	4.962	1.083	10.68	1.555	0.470	2.690
S2	721	2.222	13.02	1.624	17.74	0.860	0.190	2.230
S5	365	1.350	2.155	1.087	7.170	0.880	0.350	1.810
S6	370	1.491	2.694	1.101	10.02	1.040	0.310	2.200

* The range is also equal to the maximum porosity since the minimum is equal to zero for all the wells

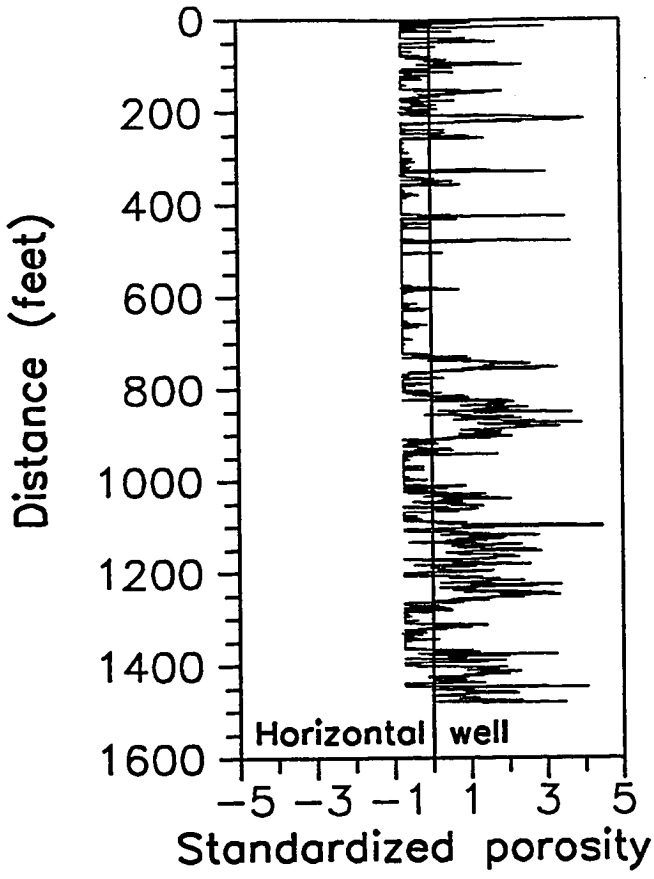

Figure 4.3 - Standardized porosity log of horizontal well (zero mean and unit variance)

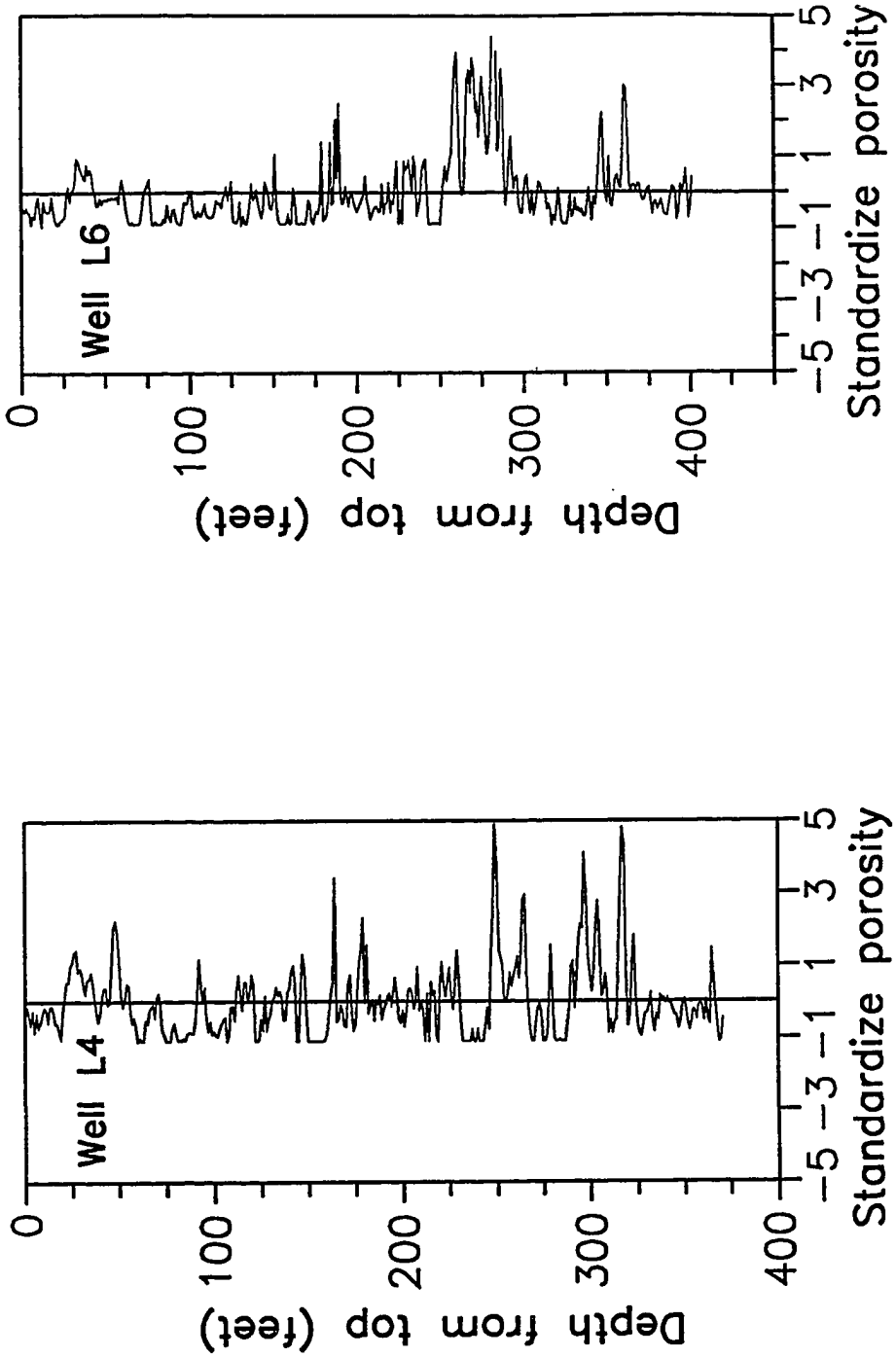

Figure 4.4 - Standardized porosity logs of Wells L4 and L6 (zero mean and unit variance)

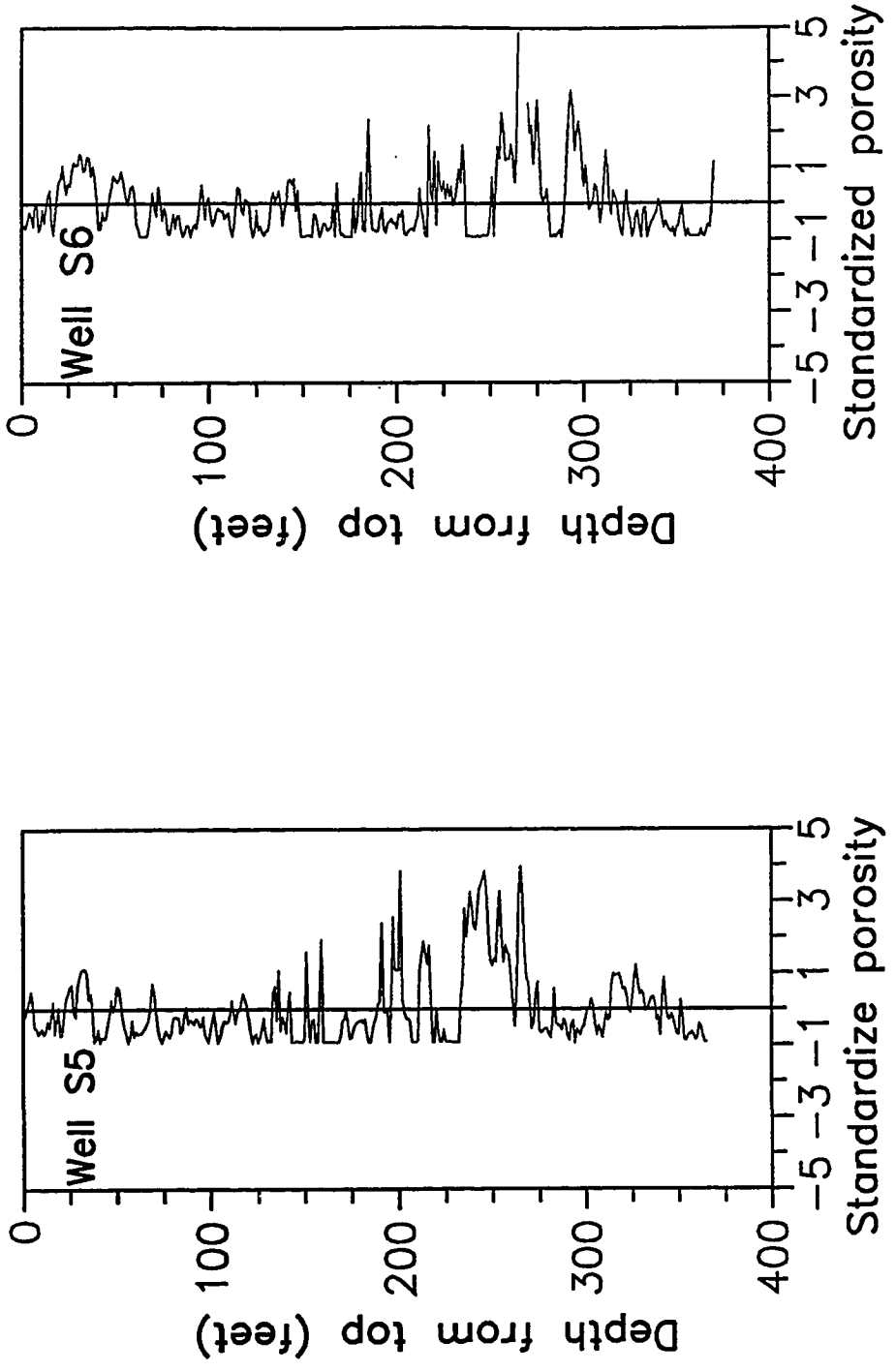

Figure 4.5 - Standardized porosity logs of Wells S5 and S6 (zero mean and unit variance)

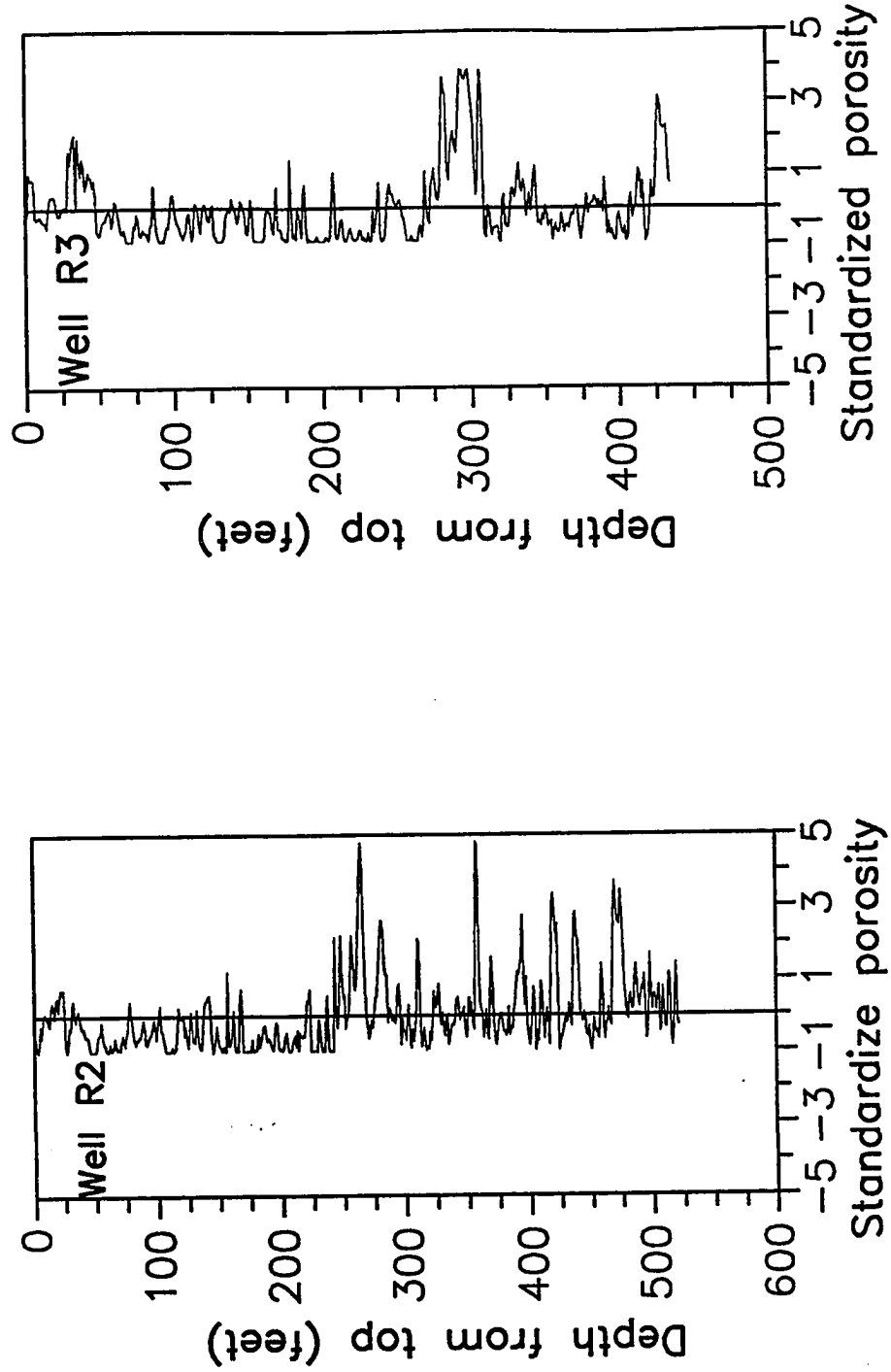

Figure 4.6 - Standardized porosity logs of Wells R2 and R3 (zero mean and unit variance)

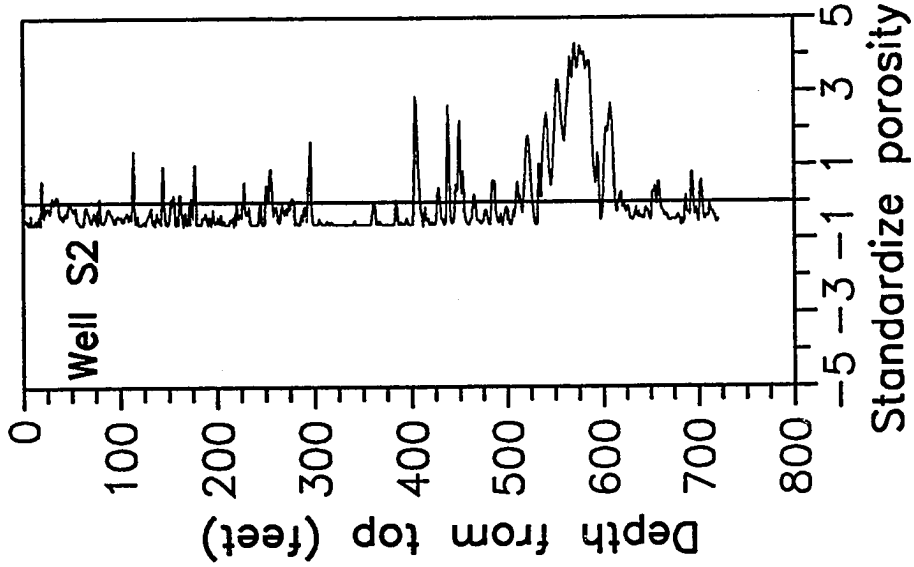

Figure 4.7 - Standardized porosity log of Well S2 (zero mean and unit variance)

low and high porosity variance. This apparent regionalized trend of the porosity variance is evaluated further in the Horizontal Subsets section.

The coefficients of variation of the porosity logs from the horizontal well and all the vertical wells logs are close in magnitude. For the horizontal well, the coefficient of variation is slightly higher (greater than one) due to the low magnitude of the mean porosity. Except for Well S2, the coefficients of variation of individual vertical well logs are close to one. The coefficient of variation is high for Well S2 because of a small region in the lower part of the formation with abnormally high porosities, as indicated in Figure 4.7.

The histograms of porosity for the horizontal well log is shown in Figure 4.8 and the histogram for all the data of the vertical wells is shown in Figure 4.9. It can be noted that the distributions shown in Figures 4.8 and 4.9 are skewed to the right due to the large fraction of low porosity measurements. For the horizontal well, the high frequency (about 43 %) of the first class is due to the large number of zero porosity values. The small frequency of the last porosity class for the vertical well logs is due to the high porosity region of Well S2 (with a maximum porosity of 17.74 %), which is also responsible for the irregularities in the coefficient of variation, as explained above.

Analysis of Spatial Statistics

Spatial correlation and variability of the porosity logs from the horizontal and the vertical wells are evaluated with fractal models and the semi-variogram analysis. The intermittency exponents of the fractal models are evaluated with the re-scaled range analysis described in the Fractal Models section of Chapter III. The semi-variogram analysis is mainly used in this section to evaluate the correlation range or the distance over which the property is correlated. The semi-variogram analysis is extended to investigate different regions along the horizontal well and the effect of the sample volume on the spatial statistics of porosity.

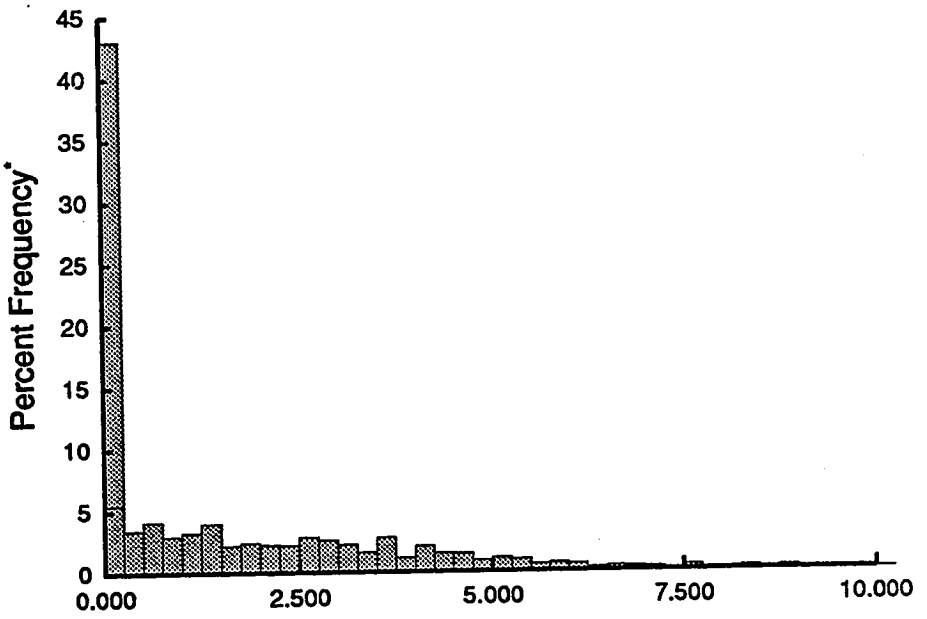

Figure 4.8 - Histogram for porosity log of horizontal well

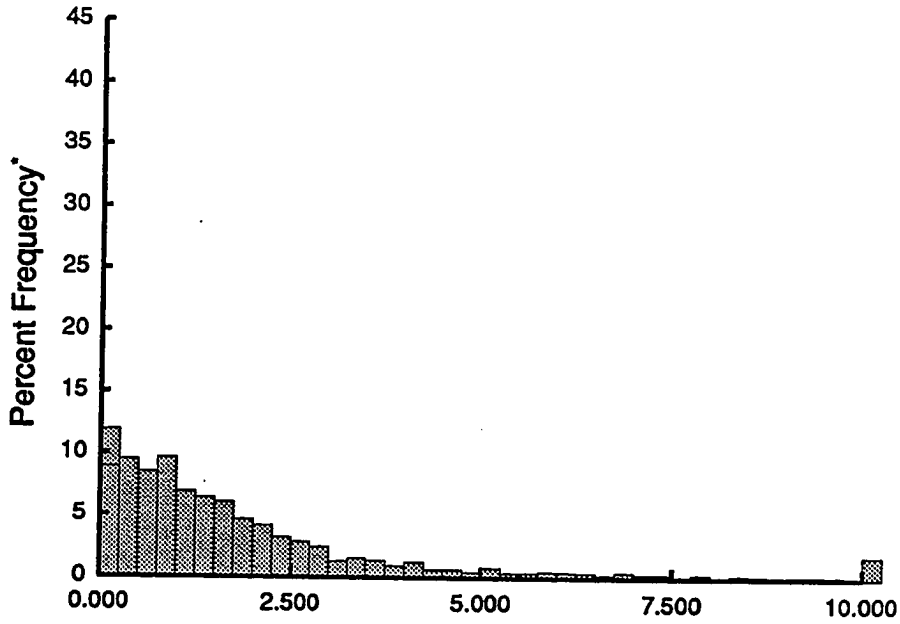

Figure 4.9 - Histogram for porosity logs of vertical wells

Fractal Analysis

The results of the R/S analysis of the porosity logs for all the wells are summarized in Table 4.2. These results are the slopes of the line fitted to the plots of R/S versus lag in log-log coordinates, as shown in Figure 4.10 for the horizontal well log and for the log of Well L4. The horizontal spatial correlation can be represented with a power-law semi-variogram model of a fBm process (Equation 3.1). In the case of a fBm process, the slope of the R/S plot is approximately equal to one plus the intermittency exponent; Thus, the intermittency exponent for the horizontal well log is equal to 0.073. Indeed, a plot in logarithmic coordinates of the sample semi-variogram versus lag for the horizontal well log follows a well defined linear trend, as illustrated in the Semi-Variogram Analysis section. The intermittency exponent calculated from the semi-variogram plot of the horizontal well log is equal to 0.13 and it is in a close agreement with the intermittency exponent calculated in the R/S analysis. For all the vertical wells, Table 4.2 shows that the R/S slope is smaller than one. This implies that the spatial correlation of the vertical wells logs can be represented with the models of fGn processes.

In the case of the horizontal well, a fBm intermittency exponent smaller than 0.5 indicates an anti-persistent behavior or that deviations from the mean are likely to be followed by deviations of the opposite sign. This behavior occurs because the horizontal well intercepts many times non-porous regions (zero porosity) resulting in many fluctuations of opposite sign around the mean line, as indicated in Figure 4.3. This characteristic of the formation may be a reason why a fBm model is obtained instead of the fGn models observed by Crane and Tubman[24] in the porosity logs of three horizontal wells in a carbonate reservoir. The intermittency exponents of the vertical well logs are in the range of the values observed by Hewett[17] and Crane and Tubman[24] for carbonate and sandstone reservoirs.

Semi-Variogram Analysis

The semi-variograms of the horizontal and vertical wells porosity logs are compared in Figures 4.11 through 4.14. The apparent correlation ranges calculated from

Table 4.2

Spatial statistics of porosity logs

Well name	R/S slope*	Apparent range (feet)	Apparent range / horizontal range
Horizontal	1.073	143	1.000
All Vertical	–	56	0.392
L4	0.845	9	0.0621
L6	0.965	29	0.200
R2	0.778	40	0.276
R3	0.928	32	0.221
S2	0.970	59	0.407
S5	0.928	26	0.179
S6	0.868	17	0.117

* If the R/S slope is less than one, then this slope is equal to the intermittency exponent of a fGn model. Otherwise, if the R/S slope is greater than one, then the slope minus one is equal to the intermittency exponent of a fBm model.

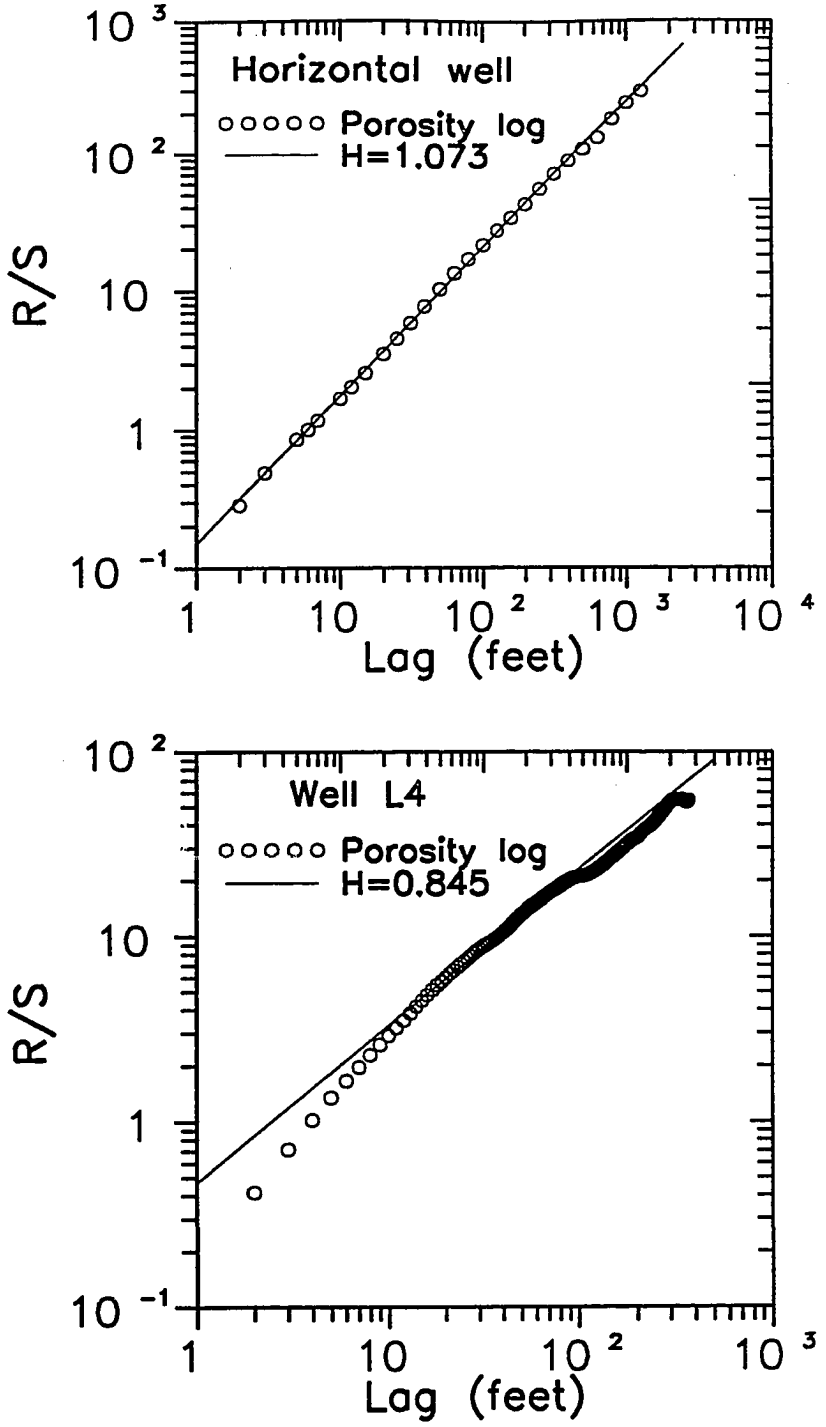

Figure 4.10 - R/S analysis for porosity log of horizontal well and Well L4

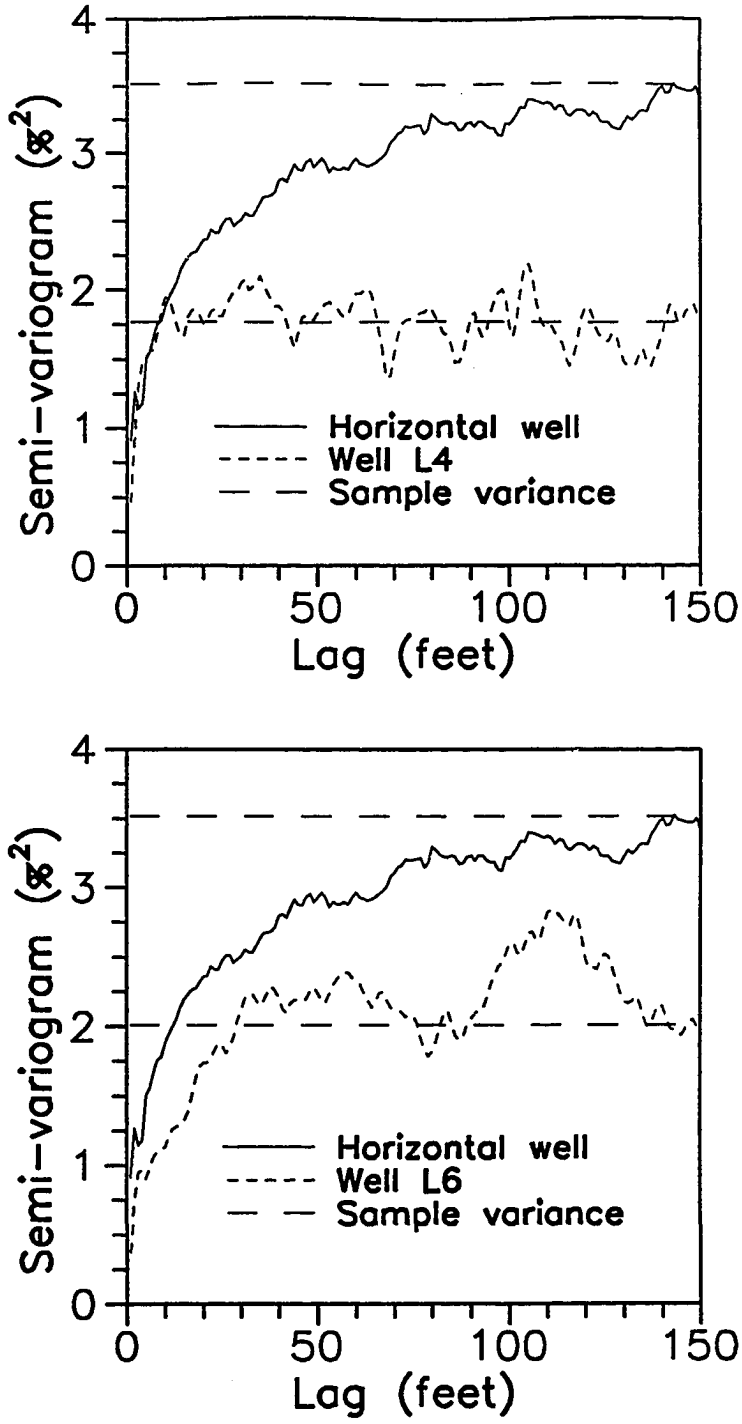

Figure 4.11 - Comparison of semi-variograms of porosity log of horizontal well and of Wells L4 and L6

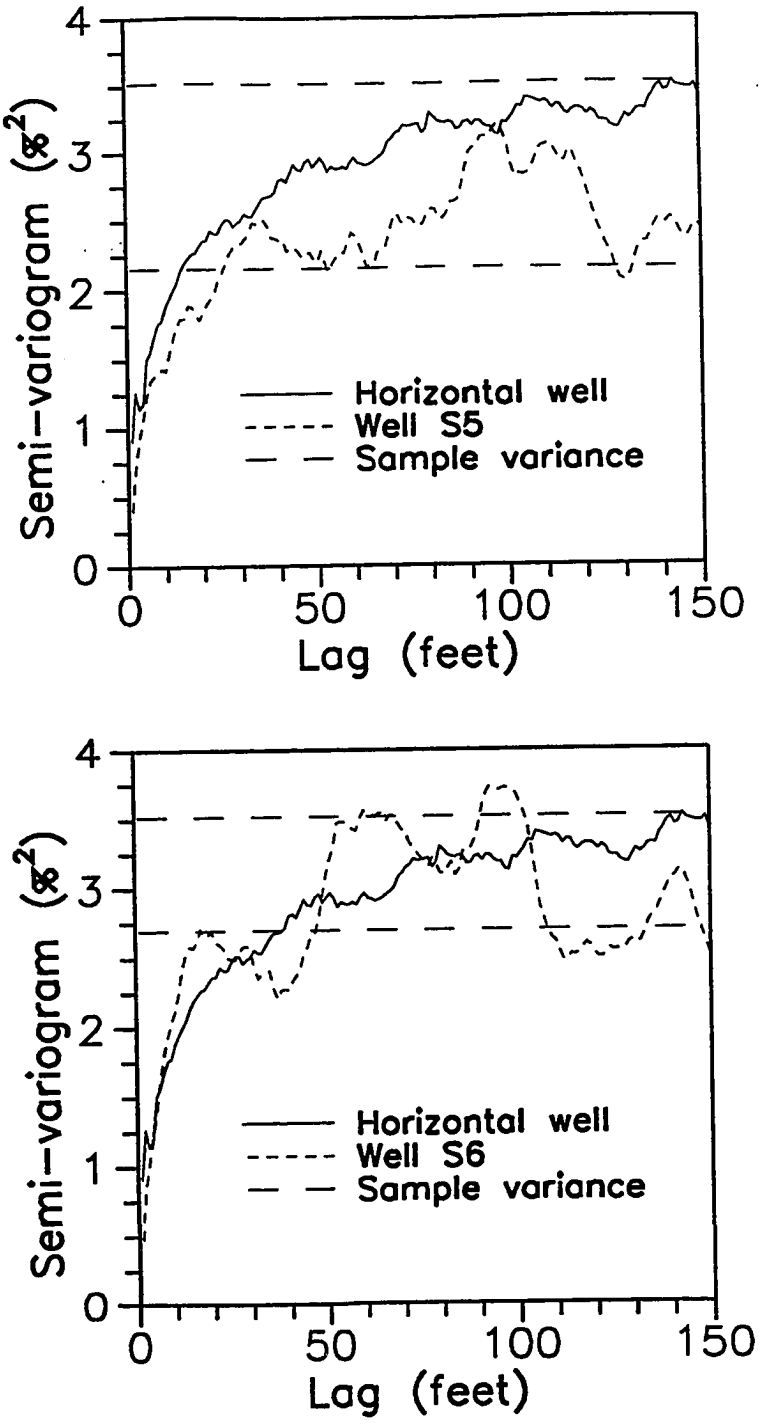

Figure 4.12 - Comparison of semi-variograms of porosity log of horizontal well and of Wells S5 and S6

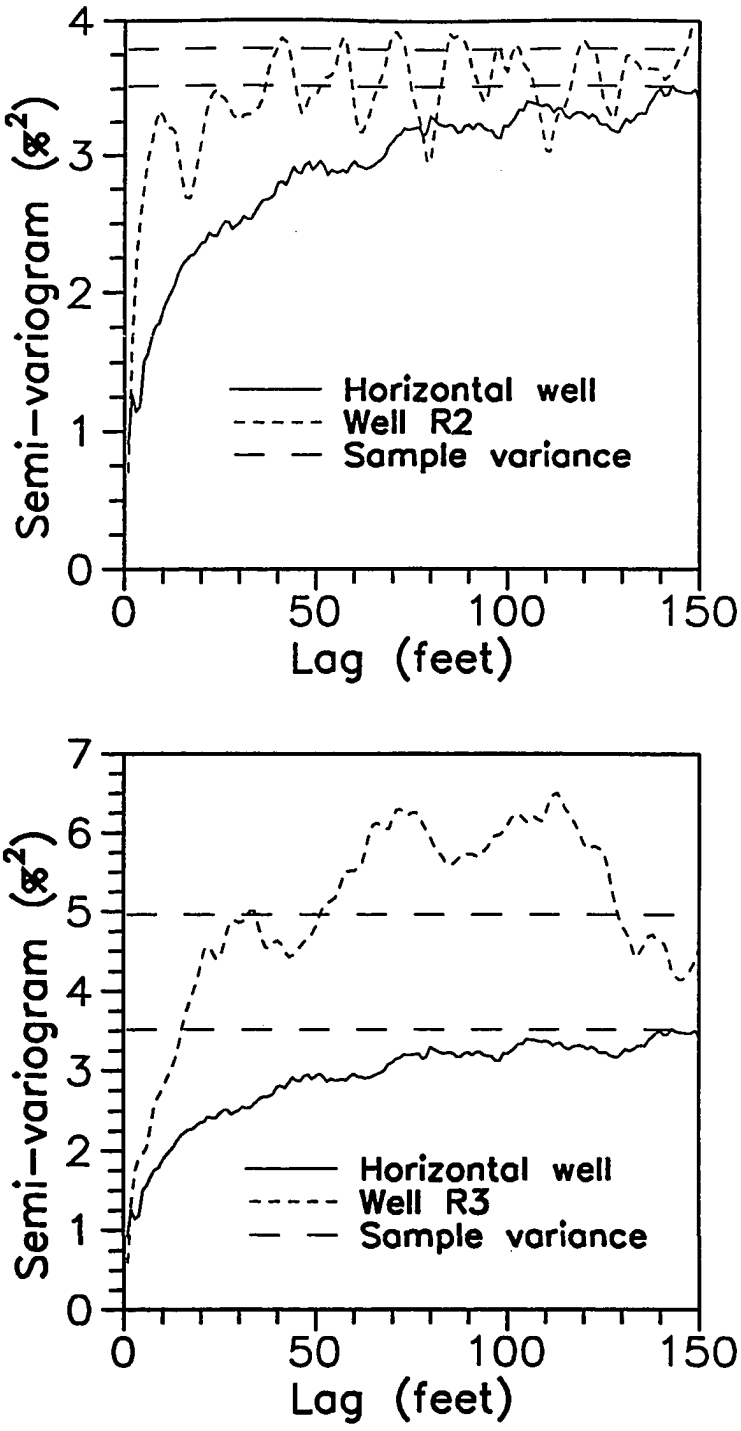

Figure 4.13 - Comparison of semi-variograms of porosity log of horizontal well and of Wells R2 and R3

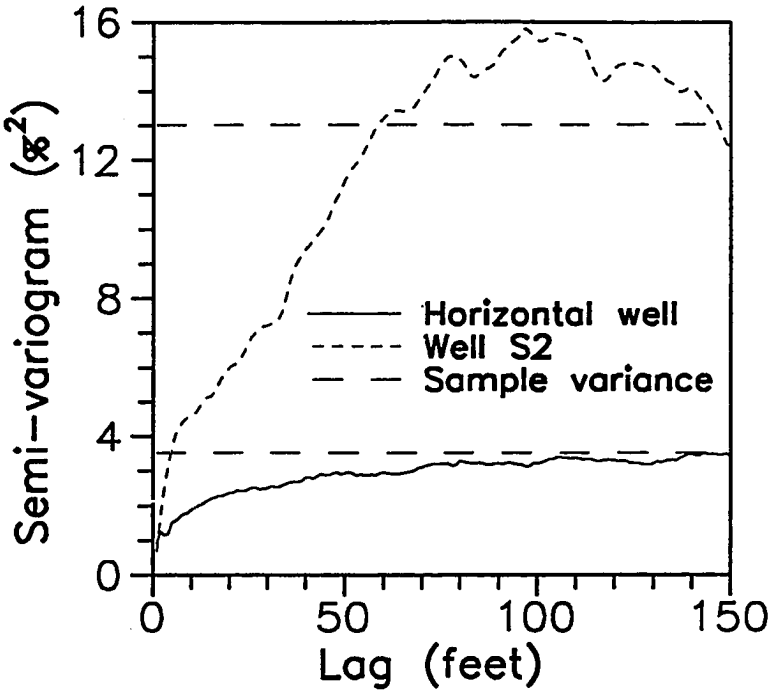

Figure 4.14 - Comparison of semi-variograms of porosity log of horizontal well and of Well S2

these semi-variograms are given in Table 4.2. Here, the apparent correlation range is defined as the smallest lag at which the semi-variogram becomes greater or equal to the variance of the samples. This definition of the apparent correlation range, depending on the behavior of the semi-variogram, may be slightly different than a correlation range calculated by fitting a model. However, the apparent correlation range has the advantage that it is independent of a model and it is not subject to the bias involved in fitting a model.

The apparent correlation range of the horizontal well log is more than two times longer than any of ranges of the vertical wells logs. The vertical wells with porosity variance greater than the horizontal well (Figures 4.13 and 4.14) have an apparent correlation range longer than the vertical wells of lower porosity variance (Figure 4.11 and 4.12). The anisotropy ratio (vertical divided by horizontal apparent correlation range) of the vertical wells is between 0.0621 (Well L4) and 0.407 (Well S2).

The shape of the semi-variogram is a measure of the degree of variability of the spatial distribution of a property. The extreme shapes correspond to a completely uncorrelated distribution (constant semi-variogram with magnitude approximately equal to the sample variance) and a distribution consisting of a constant value (semi-variogram equal to zero). In order to compare the shape of the semi-variograms from Figures 4.11 through 4.14, the sample semi-variograms have been normalized by dividing the lag by the apparent correlation range and the semi-variogram by the sample variance of each porosity log. These normalized semi-variograms for the horizontal and vertical wells logs are shown in Figure 4.15. The vertical wells with a low porosity variance have similar semi-variogram shapes and are slightly more continuous than the shape of the horizontal well semi-variogram. For the vertical wells of high porosity variance, only Well R2 has a semi-variogram shape which indicates more relative variability than the horizontal well semi-variogram; the other two wells (Wells R3 and S2) have semi-variogram shapes similar to the low porosity variance wells.

The intermittency exponents calculated in the Fractal Analysis section is another measure of the degree of relative spatial variability of a variable around the

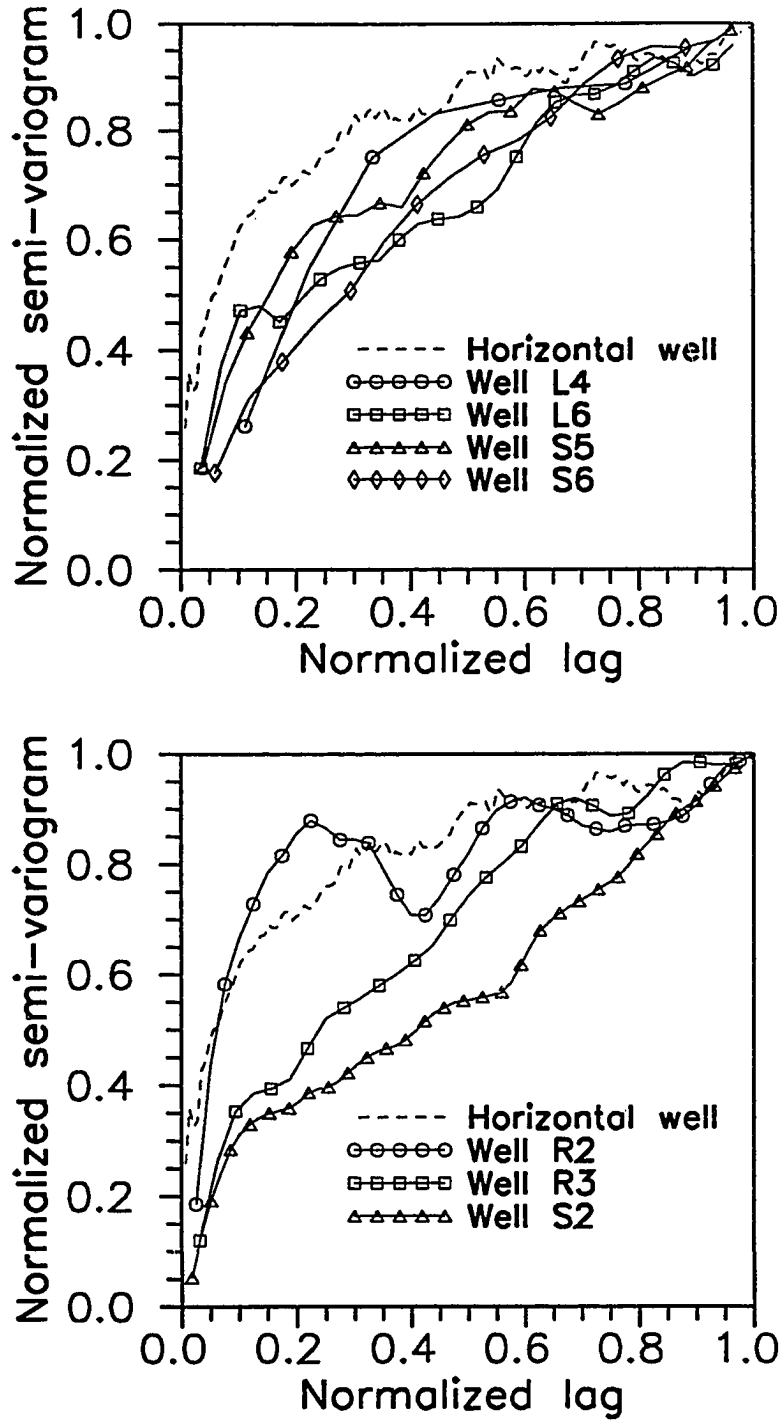

Figure 4.15 - Comparison of normalized semi-variograms of porosity logs of horizontal and vertical wells

mean. The results of the Fractal Analysis section indicate that the porosity logs of the horizontal and the vertical wells resemble different fractal processes. However, it is not surprising that the shape of the normalized semi-variograms for the horizontal and vertical wells in Figure 4.15 are similar because the shape of the semi-variograms of fBm and fGn can be similar, as explained in the Fractal Models section in Chapter III, for the range of intermittency exponents considered (Table 4.2).

The porosity semi-variogram of the horizontal well log is compared in Figure 4.16 to the semi-variograms along the vertical and horizontal directions derived from the sample data of all the vertical wells logs. A comparison of the semi-variograms for the logs of the horizontal well and all the vertical wells indicates that there is a greater spatial variability (except for the first lag) in the vertical than in the horizontal directions. The semi-variogram in the horizontal direction can be computed from the vertical wells logs only for a few large lags corresponding to the separation distance among the vertical wells. As shown in Figure 4.16, for the configuration of the vertical wells in this study region, the shortest horizontal lag is 1200 feet and the longest horizontal lag is 5900 feet. The magnitude of these semi-variogram values at the smaller lags is close to the porosity variance of the horizontal well log while for the larger lags it is close to the variance of the sample data of all the vertical wells. The values of this semi-variogram along the horizontal direction show a large scatter and there are too few points to be able to infer a correlation structure.

The observations from Figure 4.16 can be used to derive practical guidelines to assess information about the actual correlation range for the horizontal direction from the vertical well data. Based on the semi-variogram of the vertical well data in the horizontal direction, it can only be inferred that the actual correlation range for the horizontal direction is smaller than the minimum distance between the vertical wells. For the reservoir studied in this chapter, it is known that the actual correlation range in the horizontal direction is longer than the correlation range along the vertical direction. Generally, it can be expected that the correlation range in the

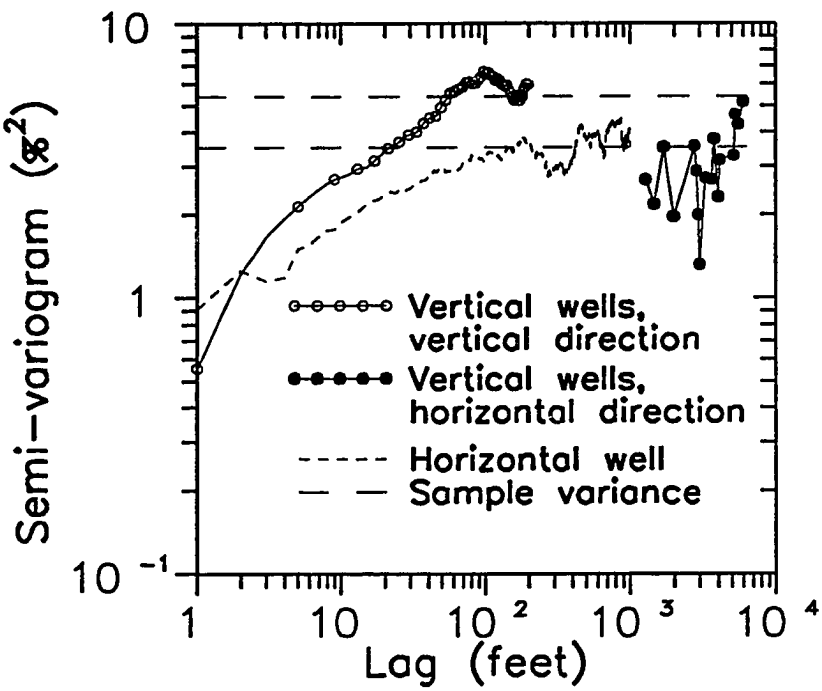

Figure 4.16 - Comparison of semi-variograms of porosity logs of horizontal well and of vertical wells for the vertical and horizontal directions

horizontal direction is longer than the range in the vertical direction. Therefore, the correlation ranges calculated from the vertical well data for the horizontal and vertical directions provide upper and lower limits for the actual horizontal correlation range, respectively. For the carbonate reservoir considered in this chapter, these upper limit and lower limits for the horizontal correlation range are 1200 and 56 feet, respectively. Since, the range of these limits can be large, in practice the sensitivity of the conditional simulations or estimations to the horizontal correlation range within these limits can be included as an additional source of uncertainty.

Even though, the nature of the geologic setting is not part of the analyses in this section, additional research might show that it is possible to narrow even further the limits of the spatial statistical parameters by incorporating information about the depositional environment. For example, in the sandstone field described in Chapter III, a horizontal correlation range equal to the separation distance between the wells (3030 feet) which is significantly longer than the vertical correlation range (17 feet) was adequate to represent the lateral spatial correlation of porosity while in the carbonate reservoir discussed in this chapter, the magnitude of the horizontal correlation range is closer to the vertical correlation range of all the porosity data from the vertical wells than to the separation distance among the vertical wells.

Evaluation of Horizontal Subsets

In this section, the horizontal well log is divided into subsets to investigate the changes on the statistical parameters describing porosity in different regions of the formation. These changes are expected because of the lateral trend of increasing porosity variance from the west to the east region of the formation, as noted in the Analysis of Univariate Statistics section and which coincides with the direction of the horizontal well. If the statistical parameters for these subsets of the horizontal well follow the trend observed in the vertical wells then it provides an indication of the vertical and lateral coverage of post-depositional effects on reservoir properties.

The four subsets of the horizontal well log considered are equal to 350 feet each. These subsets are denoted as Subset 1 through 4, where Subset 1 is in west section

of the horizontal well and Subset 4 is the east section. The estimated values of the statistical parameters for each subset are summarized in Table 4.3. As suspected, the porosity variances of the subsets in the western region of the formation (Subset 1 and 2) are lower than the variances of the subsets in the eastern region (Subset 3 and 4). The apparent correlation ranges of the western region subsets are significantly smaller than the correlation ranges of the eastern region subsets. However, the apparent ranges of the two subsets in the west side and the two subsets in the east side are close. Despite these differences in the correlation ranges, the normalized semi-variograms in Figure 4.17 show that the shapes of the semi-variograms of all the subsets are similar.

The porosity semi-variograms of the subsets from the horizontal well log are compared in Figures 4.18 and 4.19 to the semi-variograms of five vertical well logs. The semi-variograms for Well S6 (Figure 4.18) and for Subset 1 are closer than any of the other semi-variogram of the vertical wells and the horizontal subsets. The semi-variograms for Wells L6 and S5 (Figure 4.18) are similar to the semi-variogram of Subset 1. For the wells with a porosity variance greater than the variance of the horizontal well log, the semi-variograms are not similar to any of the semi-variograms of the subsets of the horizontal well, as shown in Figure 4.19 for Well R2. These observations indicate that horizontal and vertical spatial correlations are similar in the west section of the study region. However, in the east section, similarities between horizontal and vertical spatial correlations are not observed.

Sample Volume Effect

The effect of the size of sample volumes on the statistics of porosity is investigated using the horizontal well log. In this section, support volume refers to the volume of rock around the well sampled by the measurements of the logging tool. For example, Neutron logs in vertical wells have a depth of investigation into the formation of about 1 to 2 feet[46] and a vertical resolution of about one foot. The practical implication of the effect of the support volume in the statistics of a property is that the volumes considered in most engineering calculations, such as grid

Table 4.3

Univariate and spatial statistics for

subsets of porosity log of the horizontal well

Subset number	Mean (%)	Variance ($\%^2$)	Apparent range (feet)
1 (West)	0.950	2.580	12
2	0.226	0.666	8
3	1.826	4.096	37
4 (East)	2.172	3.910	40

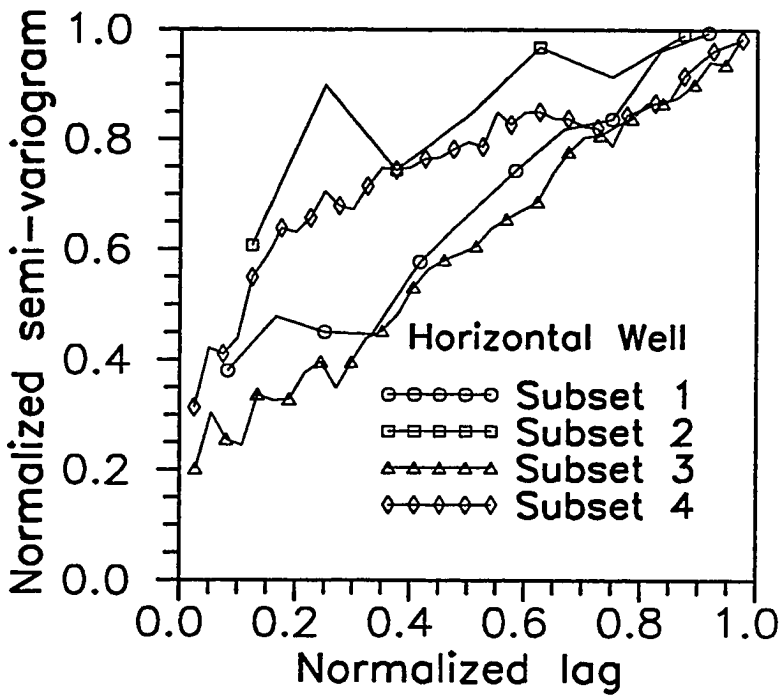

Figure 4.17 - Comparison of normalized semi-variograms of four subsets of porosity log of horizontal well

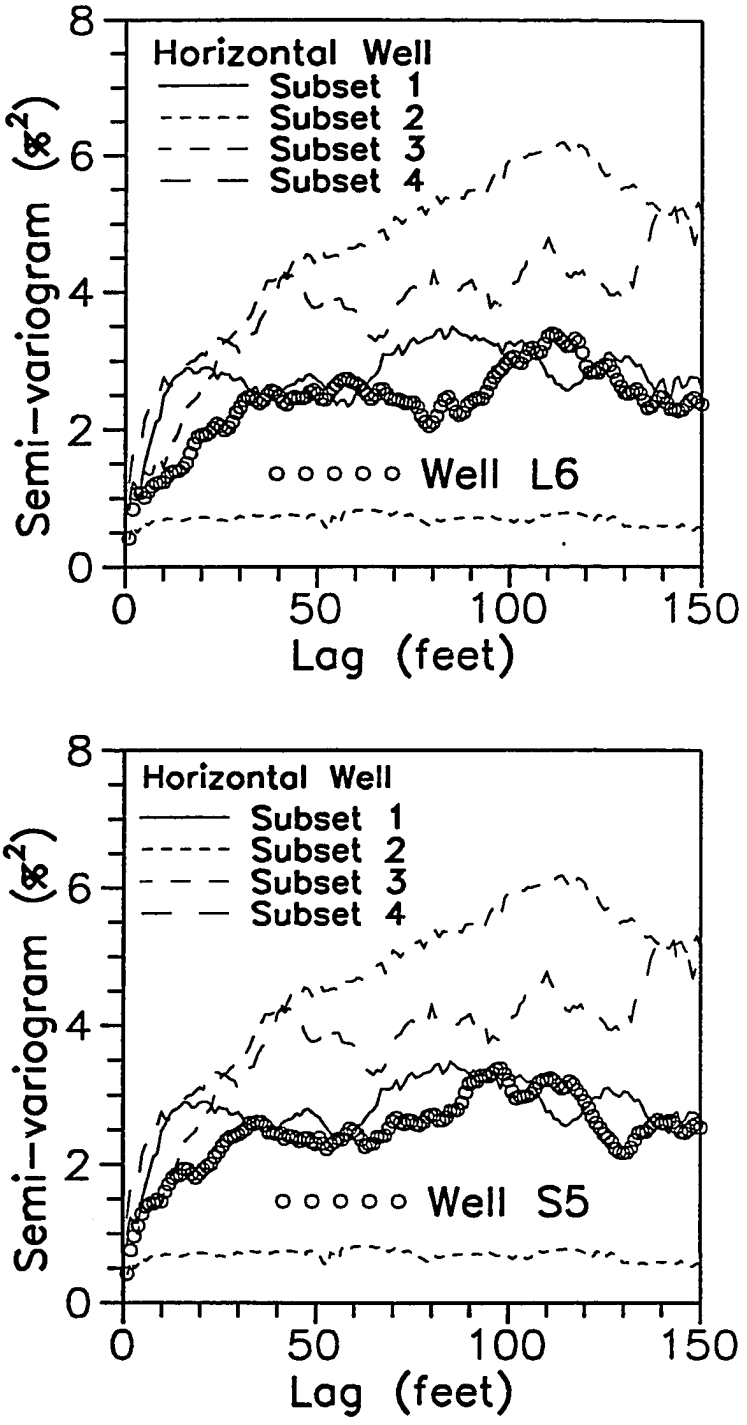

Figure 4.18 - Comparison of semi-variograms of four subsets of porosity logs of horizontal well and Wells L6 and S5

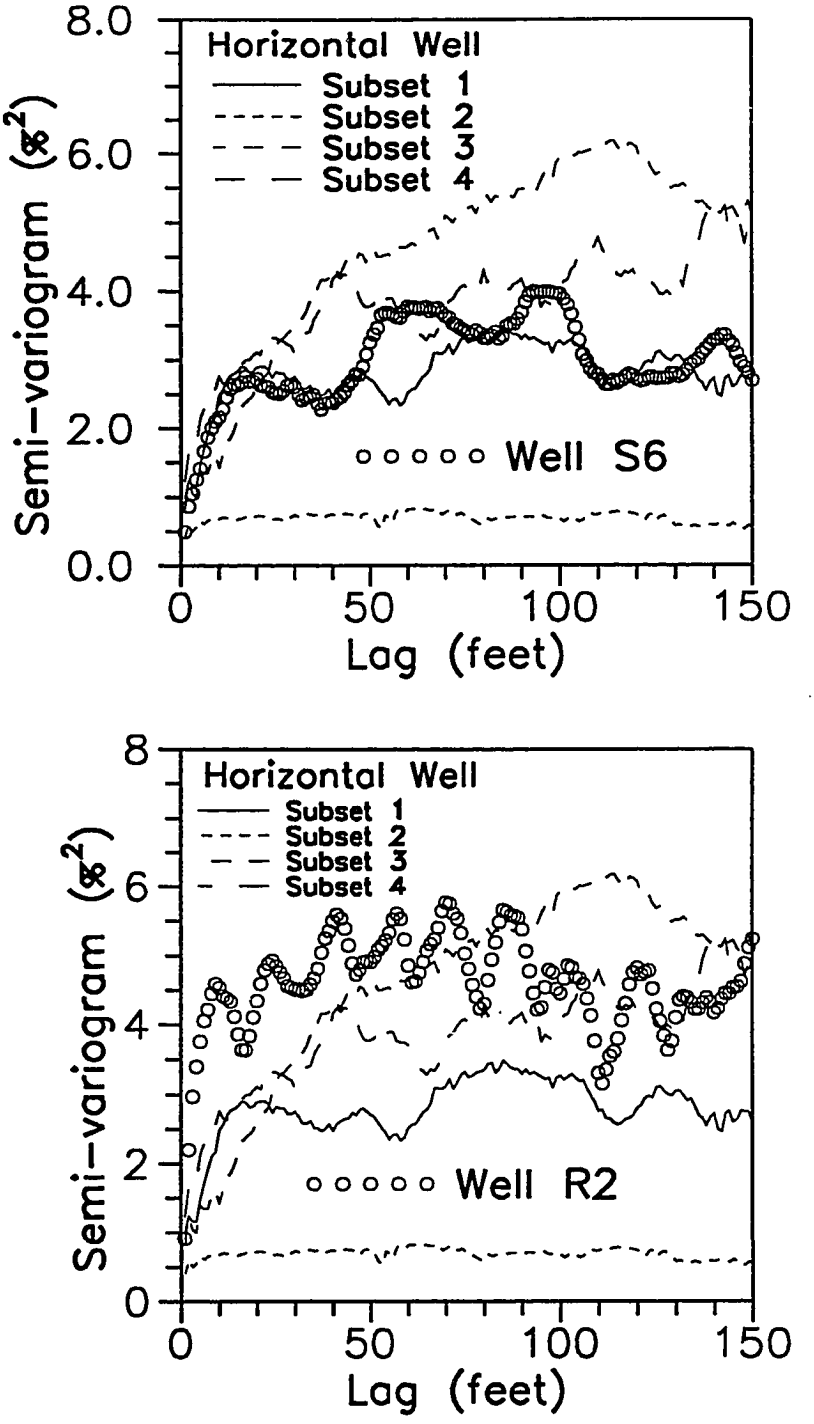

Figure 4.19 - Comparison of semi-variograms of four subsets of porosity logs of horizontal well and Wells S6 and R2

blocks of reservoir simulators, are significantly greater than the volumes sampled by the logging tools.

The support volume of porosity from the horizontal well log has been increased by averaging consecutive measurements and assigning this average porosity to the middle location of these measurements. In this case, the support volume per unit area of the plane normal to the direction of the well is equal to the length of the log measurements that were used in computing the average porosity. Figure 4.20 shows the porosity logs of the horizontal well for five support volumes per unit area ranging from 1 to 50 feet. Most of the high porosity regions observed in the original 1 foot log are preserved in the 4 feet volume per unit area log but almost disappeared (or average out) in the 50 feet volume per unit area log.

The semi-variograms of the horizontal well logs for different support volumes are compared in Figure 4.21 and the values of the statistical parameters of these logs are summarized in Table 4.4. These semi-variograms indicate that as the support volume increases the porosity variance becomes smaller and the degree of continuity increases. (The porosity means in Table 4.4 vary slightly because some samples at the ends of the logs were not included in the calculations.) The apparent correlation ranges do not change significantly for the semi-variograms of the logs for different support volumes considered, except for the 50 feet volume per unit area log. For the large support volumes, the observed trend of increasing correlation range may be due to the bias in the estimates of the semi-variograms as a result of the reduction of the number of samples in the logs. For example, the logs with a support volume per unit area equal to 25 and 50 feet have only 59 and 29 sample data points, respectively.

Conditional Simulations

In this section, three-dimensional distributions of porosity are simulated in the study region of the carbonate field using the conditional simulation method described in Chapter II. The objective of these simulations is to evaluate the simulated horizontal porosity distributions for different spatial correlation models in the

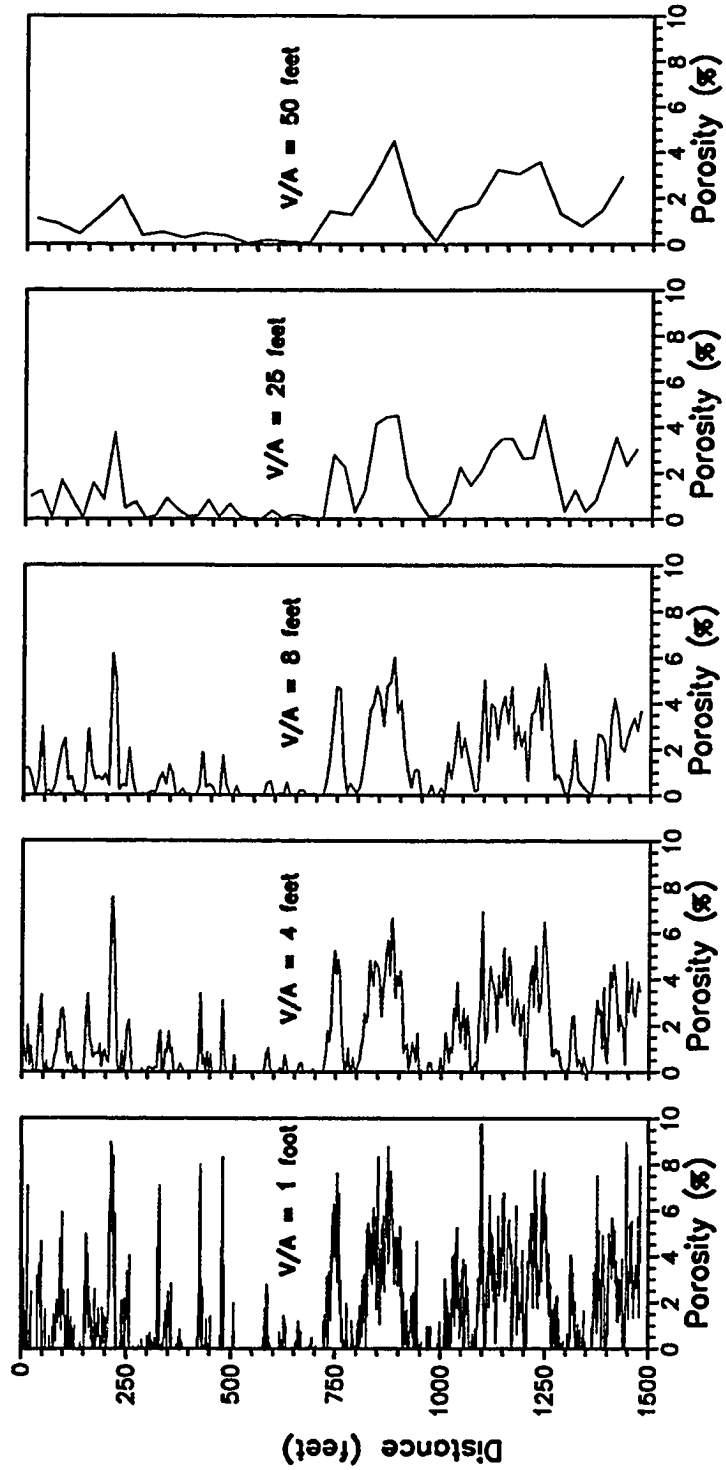

Figure 4.20 - Porosity log of horizontal well for different support volume

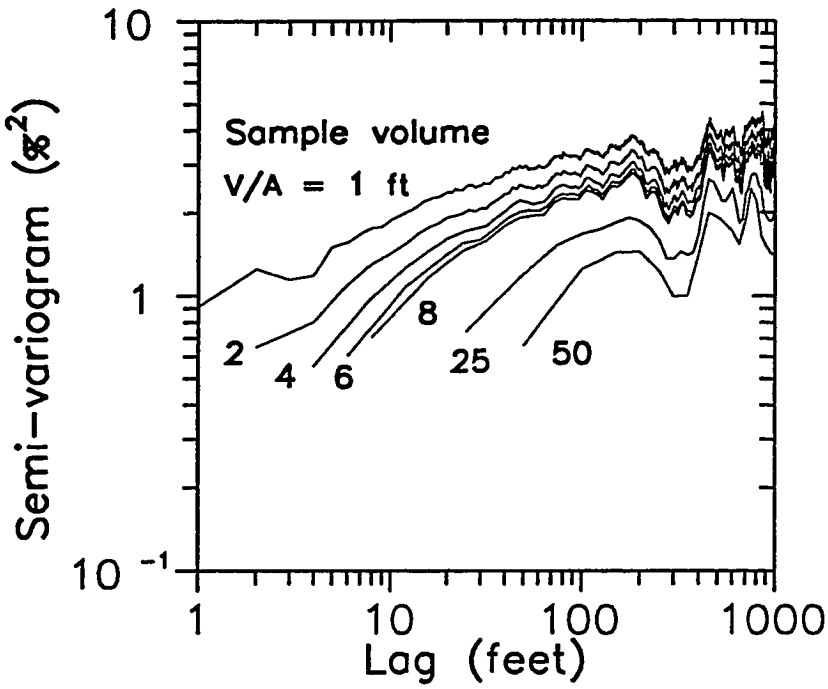

Figure 4.21 - Semi-variograms of porosity log of horizontal well for different support volumes

Table 4.4

Univariate and spatial statistics of porosity log
of the horizontal well for different support volumes

Volume / Area (feet)	Mean (%)	Variance ($\%^2$)	Apparent range (feet)
1	1.389	3.518	143
2	1.388	3.052	144
4	1.388	2.758	164
6	1.382	2.593	144
8	1.389	2.518	160
25	1.386	1.869	175
50	1.360	1.474	450

horizontal direction which might be used when the available information consists of porosity logs along the vertical direction.

Simulation Specifications

The conditional simulation region includes the horizontal well and portions of all the vertical wells shown in Figure 4.1 for the study region considered in this chapter. The specifications for the simulations are given in Table 4.5. The conditioning data consist of 50 porosity values at the location of each of the seven vertical wells. The cumulative distribution function is calculated from the conditioning data. The porosity conditioning data have a mean equal to 2.040 % and a standard deviation equal to 2.037 %. As noted in the Analysis of Univariate Statistics section, the distribution of porosity from the vertical well logs underestimates the large proportion of low porosity values observed in the distribution of the horizontal well log. The semi-variogram models specified for different conditional simulation cases are described in the next section. The values of the annealing schedule parameters (α and M_a) shown in Table 4.5 are slightly different than the optimum values derived in Chapter II because the simulations in this section were performed just before the optimum values were available. However, this slight difference does not affect the quality of the simulation results but perhaps will yield a higher computational cost, as indicated by the results of the Annealing Schedule section in Chapter II.

Spatial Correlation Models

In this section, semi-variogram models for four conditional simulation cases are developed using the porosity log data of the vertical wells. The effects of the different sets of semi-variogram models on the simulated porosity distributions and the reasons for selecting these models are discussed in the following section.

The semi-variogram models developed for the horizontal and vertical directions are fGn and exponential models. The fGn and exponential models fitted to the sample semi-variogram of the log data for the vertical wells are shown by the solid lines in Figure 4.22. The parameters of the vertical fGn model (Equation 3.2) are

Table 4.5

Specifications for three-dimensional conditional simulations of porosity in the study region of the carbonate reservoir

Grid Geometry			
Direction	x	y	z
Spacing	50.0	50.0	1.0
Grid points	108	59	50
Total points	318600		
Conditioning Data			
Source		Porosity logs of Wells L4, L6, R2, R3, S2, S5 and S6	
Number		350	
Distribution Function			
Source		Conditioning data	
Number of classes		15	
Subclass distribution		Uniform	
Annealing Schedule			
α		0.65	
M_a		7.0 cycles	
ϵ_v		1.0×10^{-3}	
ϵ_a		2.5×10^{-2}	

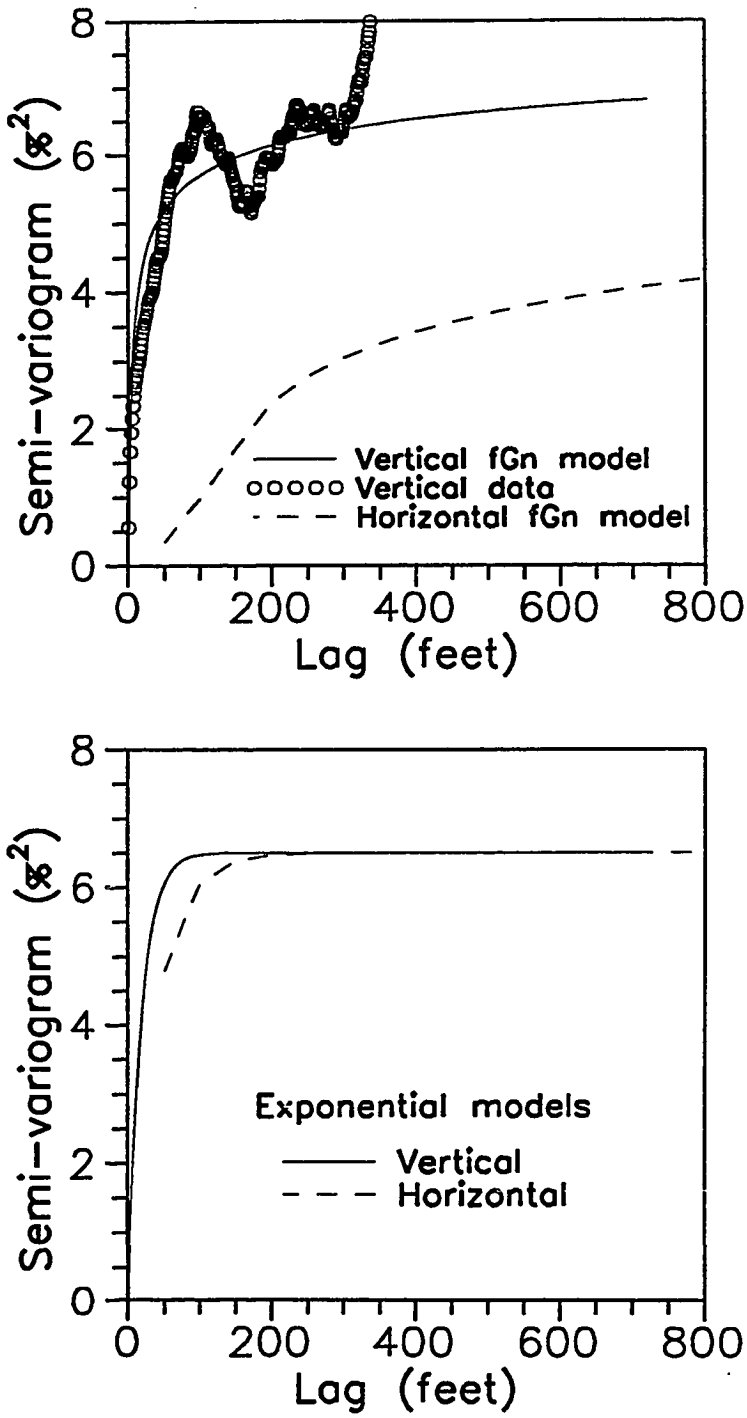

Figure 4.22 - Semi-variogram of vertical wells and models for conditional simulations

$H = 0.897$, $V_H = 12.0$ and $\delta = 4.0$ feet. The intermittency exponent, H, is the mean value of the intermittency exponents of the porosity logs of the vertical wells (Table 4.2). The vertical exponential model has a sill equal to 6.5 and a practical correlation range equal to 56 feet. For these models, the value of the sill is slightly greater than the porosity variance of the vertical wells due to the influence of the large variance of Well S2 (Table 4.1). Both, the fGn and the exponential models in Figure 4.22 provide a close fit to the vertical sample data. However, the exponential model is primarily used because it allows to explicitly specify a correlation range.

Four cases with different semi-variogram models for the vertical and horizontal directions are considered in the conditional simulations. Case A is an uncorrelated simulation, where the semi-variograms are a constant value and equal to the variance of the distribution. For Case B, a semi-variogram model is specified only for the vertical direction and it is equal to the fGn model shown by the solid line in Figure 4.22. In Case C, the fGn semi-variogram models shown in Figure 4.22 are specified for the vertical and horizontal directions. The fGn model for the horizontal directions, shown by the dashed line in Figure 4.22, has similar parameters as the model for the vertical direction, except that the smoothing factor is adjusted to $\delta = 4 \times 50.0$ feet, in order to account for the different grid spacing in the lateral directions. For Case D, the exponential models shown in Figure 4.22 are specified for the vertical and horizontal directions. The practical correlation range of the exponential model for the horizontal direction, shown by dashed the line in Figure 4.22, is set equal to two times the range of the model for the vertical direction and it is equal to 112 feet.

Simulated Distributions

The results of the three-dimensional conditional simulations analyzed in this section are horizontal porosity sections for each of the four simulation cases described in the previous section. These horizontal sections are simulated porosity values along the x direction of the simulation grid at a location near to the horizontal well. The length and spacing of the simulated horizontal sections are different

from the porosity log of the horizontal well. The total length of the horizontal sections is 5350 feet and it is approximately equal to the size of x direction of the study region shown in Figure 4.1. The spacing of simulated values in the horizontal sections is 50 feet.

The simulated horizontal porosity sections for Cases A and B are shown in Figure 4.23 and it can be noted that for both cases the sections appear to be uncorrelated. The sample semi-variograms shown in Figure 4.24 correspond to the sections in Figure 4.23 and confirm the uncorrelated character of these sections. These results show that a poorly constrained simulation such as Case B which does not include explicit information about the horizontal spatial correlation yields horizontal distributions similar to an uncorrelated distribution such as Case A. Even though, the porosity section for Case A is expected to be uncorrelated, for Case B, a horizontal correlation does not develop due to the implicit influence of the conditioning data at vertical locations.

The result of the conditional simulation for Case C and the horizontal porosity log for a support volume equal to 50 feet per unit area are shown in Figure 4.25. The porosity log of the horizontal well shown in Figure 4.25 is the same log shown in Figure 4.20 but it plotted with the same scale used for the simulated section. Figure 4.25 indicates that the simulated horizontal porosity section for Case C resembles the horizontal well porosity log of a large support volume. The sample semi-variograms shown in Figure 4.26 correspond to the porosity sections of Figure 4.25 and indicate that there is a reasonably close agreement between the spatial correlation of the simulated section and the horizontal well porosity log. Therefore, the horizontal fGn model with a smoothing parameter which accounts for the grid spacing provides an effective representation of the lateral spatial correlation for large support volumes.

In Case C, two points about the sample semi-variograms in Figure 4.26 should be noted. First, the semi-variogram of the simulated horizontal porosity section differs slightly from the specified horizontal fGn model (Figure 4.22) because the conditional simulation method developed in Chapter II does not simulate the semi-variogram of a single section but the average semi-variogram of all the sections in

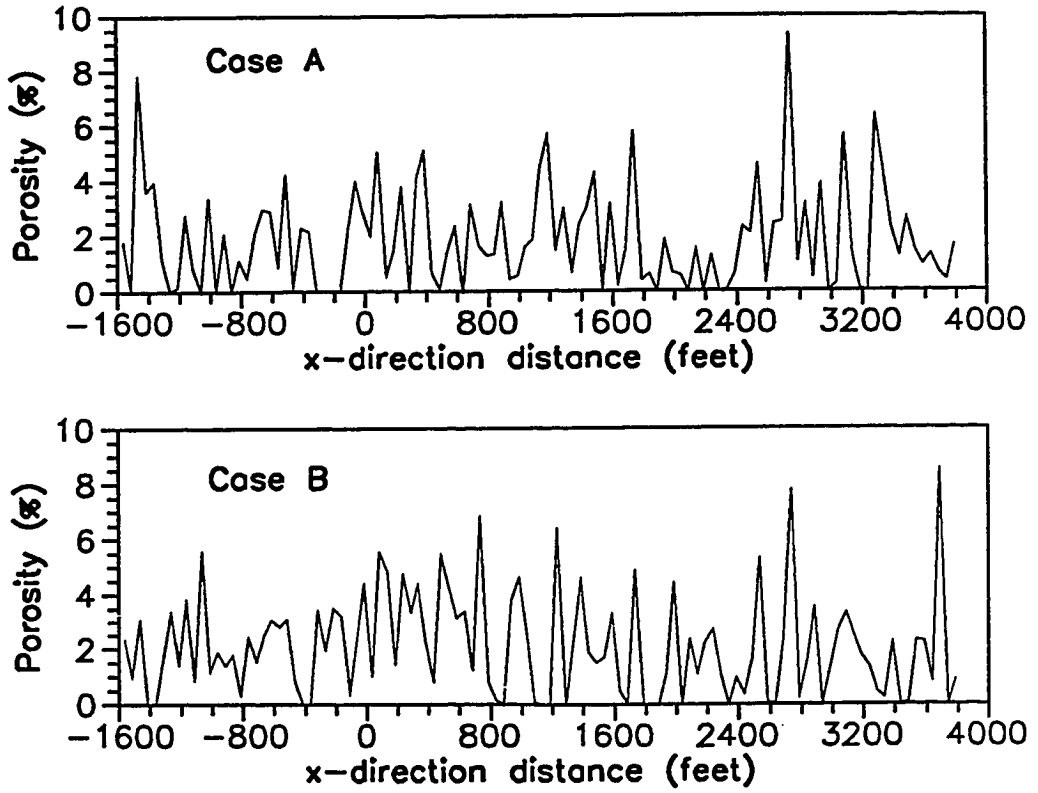

Figure 4.23 - Horizontal porosity sections for conditional simulation Cases A and B

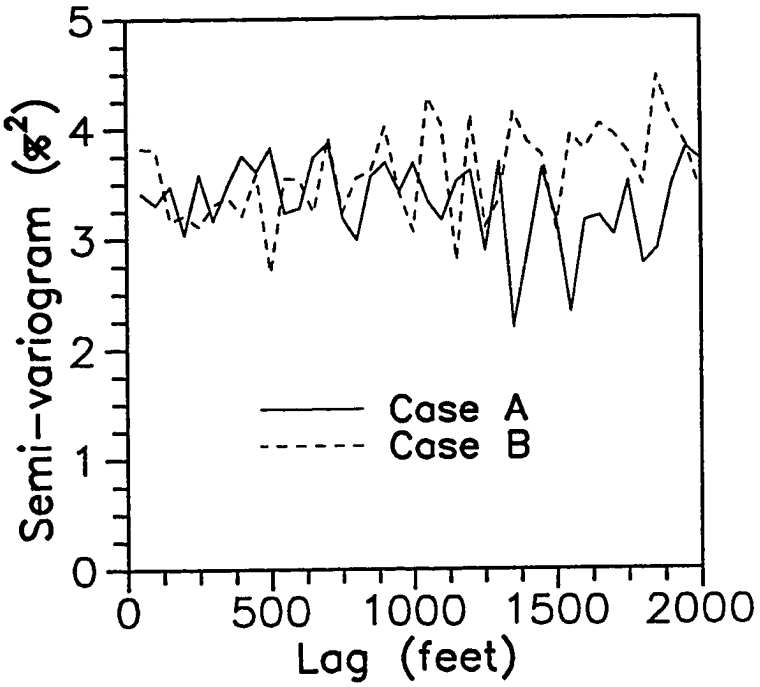

Figure 4.24 - Semi-variograms of horizontal porosity sections for conditional simulation Cases A and B

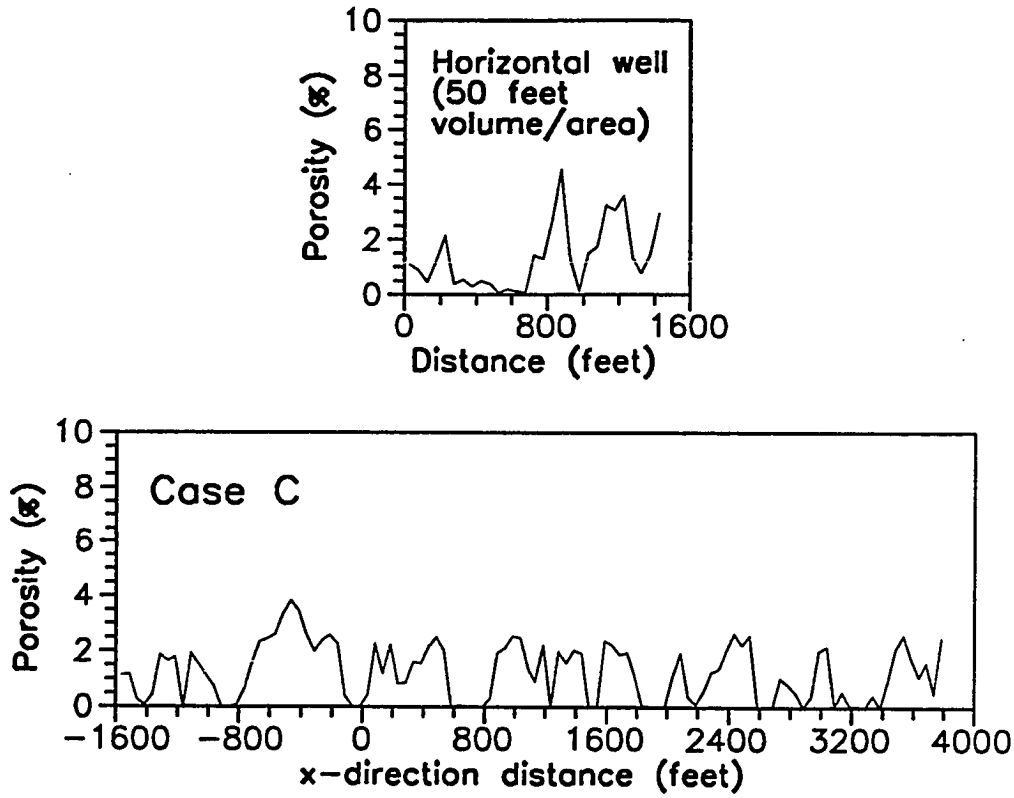

Figure 4.25 - Horizontal porosity section for conditional simulation Case C and horizontal well log with a 50 feet support volume per unit area

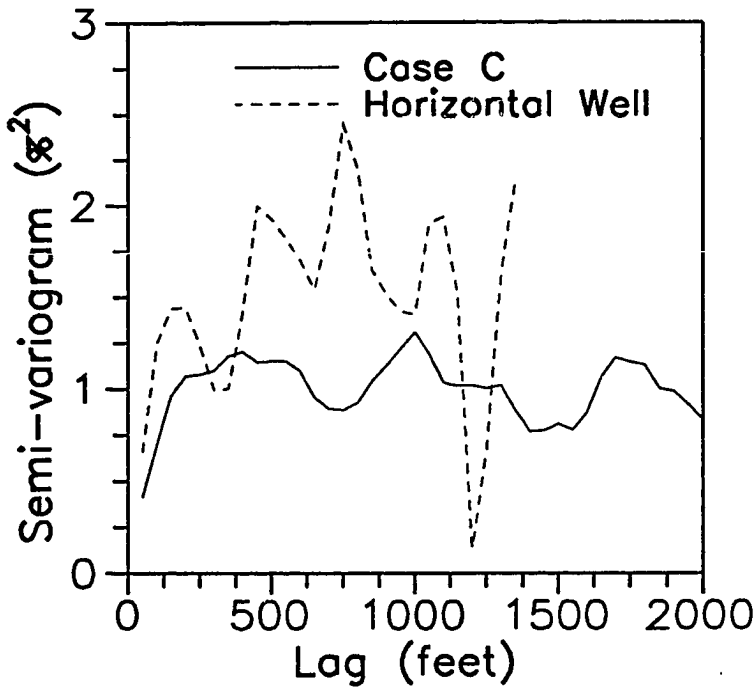

Figure 4.26 - Semi-variograms of horizontal porosity sections for conditional simulation Case C and of horizontal well log with a 50 feet support volume per unit area

a specified direction. Second, the estimation of the sample semi-variogram of the horizontal well porosity log for the 50 feet support volume may not be accurate because it contains only 29 sample data points.

The simulated horizontal porosity section for Case D and a sample of the horizontal well porosity log with values spaced by 50 feet are shown in Figure 4.27. The appearance of the simulated section for Case D is similar to the horizontal well sampled at a spacing equal to the grid spacing specified for the lateral directions in the conditional simulations. These porosity sections are comparable because the exponential semi-variogram model for the horizontal direction used in Case D is similar to the sample semi-variogram of the horizontal well porosity log at one foot spacing. The sample semi-variograms of the porosity sections in Figure 4.27 are shown in Figure 4.28. These results indicate that the horizontal spatial correlation of the simulated section and the horizontal well log are of the same magnitude and the large fluctuations in the sample semi-variograms are due to the reasons discussed above for Case C.

Even though, the objective of this section is not to attempt a close match between the simulations and the actual data as in the field validations of Chapter III, the results emphasize the importance of using spatial correlation models which properly account for the correlation range and the scale.

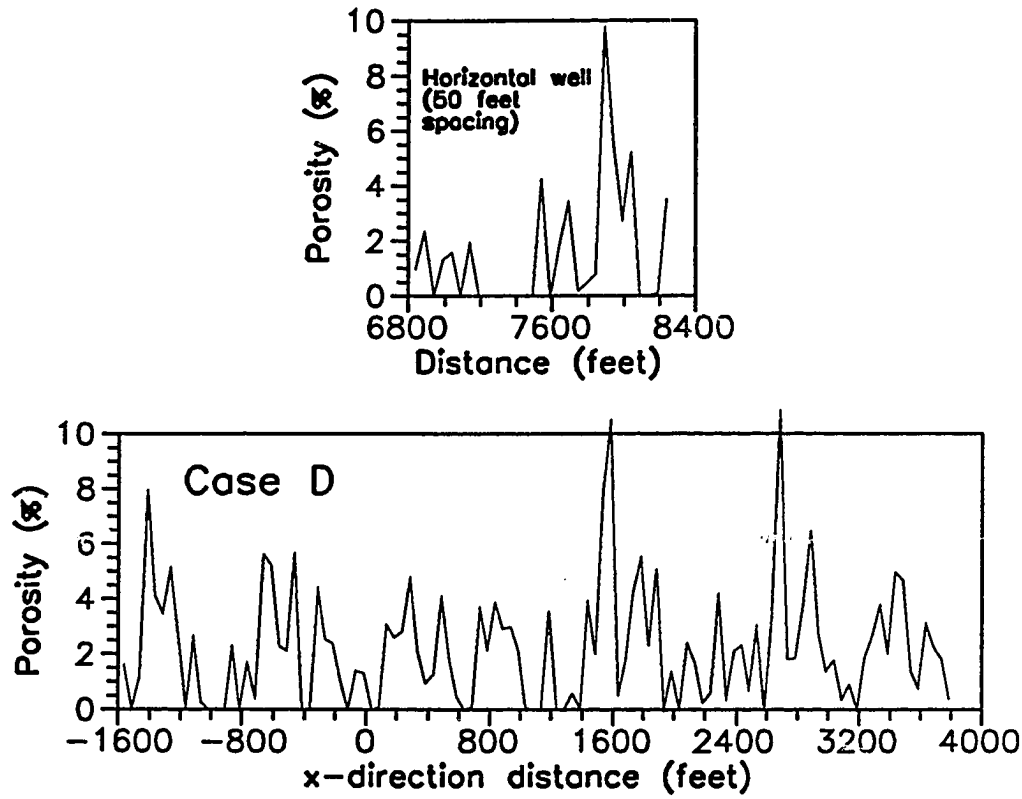

Figure 4.27 - Horizontal porosity section for conditional simulation Case D and horizontal well log sampled at a 50 feet spacing

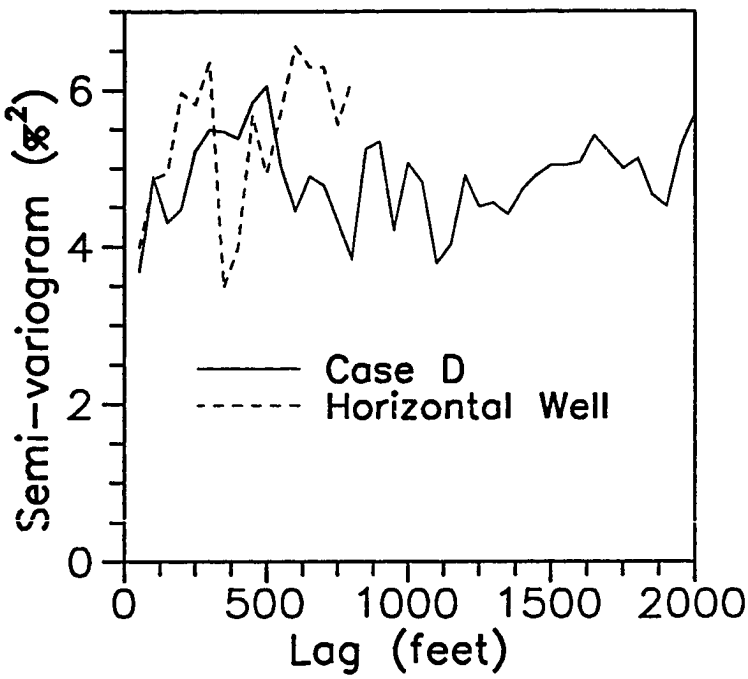

Figure 4.28 - Semi-variograms of horizontal porosity section for conditional simulation Case D and of horizontal well log sampled at a 50 feet spacing

CHAPTER V

CONCLUSIONS

This chapter summarizes the contributions and significant observations of the developments presented in the preceding chapters of this dissertation. At the end, the conclusions derived from these developments are enumerated.

Chapter II presented the development of a stochastic conditional simulation method based on the simulated annealing optimization technique to generate equiprobable correlated distributions of a multi-dimensional variable. The constraints reproduced by this conditional simulation method are a set of conditioning data, a cumulative distribution function and semi-variogram models in multiple directions. Although the original basis for using simulated annealing to generate correlated distributions was developed by Farmer,[25] several enhancements to the original method are included in the method discussed in Chapter II. In order to reduce the number of simulation parameters required in the simulated annealing technique, procedures are introduced to calculate two parameters within the conditional simulation method. Furthermore, a numerical approach is developed to derive the values of the additional simulation parameters, referred as the annealing schedule, required for optimal conditional simulations. Another benefit of this numerical approach is that the certainty coefficient introduced in Chapter II allows to quantify the quality of the simulated distributions. The applicability of these developments is not restricted to the conditional simulation method described in Chapter II; these can be useful for the implementation of simulated annealing in other disciplines.

In Chapter III, distributions of porosity are generated in a sandstone and a carbonate field using the conditional simulation method developed in Chapter II. A validation procedure is developed to test the reliability of the simulated distributions to represent the observed distributions of reservoir properties. The results of the field validations indicate that the simulated distributions are a close representation

of the observed distributions provided that the semi-variogram models specified in the simulations are adequate to represent the spatial correlation of the reservoir properties. For the sandstone and the carbonate fields, the quality of the results obtained with the conditional simulation method developed in Chapter II is of the same magnitude or better than the results obtained with the other conditional simulation methods tested.

The results of Chapters II and III suggested that the effectiveness of conditional simulation to predict the observed distributions of a reservoir property is a strong function of the semi-variograms models used to represent the spatial correlation of the property along the major directions. However, a practical problem in most petroleum reservoirs is the irregular availability of the sample data necessary to evaluate the spatial correlation of a property because the data are scarce in the areal directions and abundant in the vertical direction. Therefore, in Chapter IV, log data from a horizontal and several vertical wells in a carbonate reservoir are used to assess the nature of the small-scale inter-well spatial distribution of porosity. Based on extensive analyses and comparisons of the univariate and spatial statistical attributes of the porosity logs, guidelines are developed to infer the horizontal spatial correlation from the vertical well log data in this particular carbonate reservoir. The evaluation of three-dimensional simulated porosity distributions in the carbonate reservoir to represent the horizontal spatial correlations, using different models, indicates that the size of the support volume should be included in the conditional simulations procedures.

The conclusions and observations derived from the preceding chapters of this dissertation are listed below.

(1) The conditional simulation method developed in Chapter II using the simulated annealing technique is robust. For most distributions generated with this conditional simulation method, the statistical properties of the simulated distributions are in excellent agreement with the simulation specifications.

(2) The approach proposed by Aarts and Korst,[26] to calculate the control parameter at initial conditions in simulated annealing, provides adequate estimate of

the initial control parameter for the conditional simulation method developed in Chapter II.

(3) The method developed to estimate the number of total iterations per step required in simulated annealing eliminates the need to specify this parameter in the annealing schedule. This method yields simulation results which are in close agreement with simulations where the number of total iterations per step is set equal to a large constant number.

(4) A parameter denoted as the certainty coefficient is defined as the correlation coefficient between the actual and simulated distributions. This parameter allows to measure the quality of simulated distributions.

(5) The certainty coefficient and the testing approach based on a known master distribution introduced in Chapter II have been useful in optimizing the annealing schedule parameters. This approach can be extended to evaluate other conditional simulation methods.

(6) For the conditional simulation method developed in Chapter II, a close reproduction of all simulation specifications is required, but it is not a sufficient condition to ensure the highest possible quality of the simulations.

(7) The quality of a simulated distribution is significantly affected by the parameters of the annealing schedule. The quality of a distribution increases as the convergence rate factor and the maximum number of accepted iterations per step increase.

(8) The numerical experiments conducted in Chapter II to evaluate the quality and the computational cost of conditional simulations for different annealing schedules, indicate that a convergence rate factor equal to 0.50 and a maximum number of accepted iterations per step equal to 5.0 cycles yield optimum results.

(9) The quality of distributions generated with the simulated annealing technique is superior to the greedy (or iterative improvement) technique that does not use the Metropolis conditions.[30]

(10) Further numerical experiments in Chapter II indicate that the spatial correlation model for the direction of greater continuity has the greatest effect on the quality of the simulated distributions.

(11) The conditional simulation method developed in Chapter II is flexible and offers the potential for several practical extensions to enhance the descriptions of reservoir properties.

(12) The validation procedure developed in Chapter III provides a viable approach to test the reliability of reservoir properties descriptions generated by a conditional simulation in field cases.

(13) The validation of two-dimensional simulated porosity distributions for the sandstone field in Chapter III indicates that a model of a fractional Brownian motion (fBm) process is a better representation of the unknown horizontal spatial correlation than the model of a fractional Gaussian noise (fGn) process derived from the vertical log data. These results suggest that the porosity variability in the horizontal direction is significantly smaller than the variability observed in the vertical well logs.

(14) For the particular conditions present in the sandstone field of Chapter III, the successive random additions method is computationally more efficient than the conditional simulation method of Chapter II for the generation of two-dimensional distributions in a vertical reservoir section between two conditioning wells.

(15) For the carbonate field in Chapter III, the validation of three-dimensional simulated porosity distributions indicates a close agreement with the actual distributions when a fGn model similar to the model derived from the vertical well data is used to represent the inaccessible horizontal spatial correlations in the conditional simulations. This horizontal fGn model accounted for the large grid spacing in the lateral directions.

(16) The validation of two-dimensional simulated porosity distributions for the carbonate reservoir of Chapter III indicates that three equi-probable realizations generated using the conditional simulation method of Chapter II are slightly closer to the distributions observed in a test well than the realizations generated using the sequential indicator simulation method.

(17) In the carbonate reservoir of Chapter IV, the univariate analysis of porosity logs for a horizontal well and several surrounding vertical wells shows that the

porosity variance of the horizontal log is slightly smaller than the variance of all the vertical well logs but it is within the variance of individual vertical well logs. The spatial analysis indicates that the porosity correlation range of the horizontal well log is 2.5 times greater than the vertical correlation range of all the vertical well logs.

(18) The fractal analysis of the porosity logs in the carbonate field of Chapter IV indicates that the vertical well logs follow a fGn process with intermittency exponents ranging between 0.78 and 0.97, while the horizontal well log follows a fBm process with an intermittency exponent close to 0.10.

(19) The observations from the spatial analysis of porosity in Chapter IV indicate that the sample data from vertical wells can provide upper and lower limits for the actual horizontal correlation range.

(20) The three-dimensional conditional simulations of porosity for the carbonate field of Chapter IV show that simulations with information only about the vertical spatial correlation are poorly constrained and yield uncorrelated horizontal porosity sections.

(21) The evaluations and comparisons in the carbonate field of Chapter IV between the horizontal porosity log and horizontal sections from three-dimensional conditional simulations indicate that different models for the horizontal spatial correlation can account for the sample volume size.

NOMENCLATURE

Symbol

a	=	correlation range for semi-variogram model
C	=	certainty coefficient
C_o	=	nugget
COV	=	covariance function
$d_{1,2}$	=	parameter for semi-variogram correction term
$D_{1,2}$	=	parameter for semi-variogram correction term
E	=	objective function
E'	=	objective function after an exchange
E_0	=	normalizing constant for objective function
$E_{initial}$	=	objective function for initial conditions
F	=	cumulative distribution function
F_o	=	specified cumulative distribution function in simulation
$\vec{h}$	=	lag distance vector for semi-variograms
h_1	=	normalized lag for fGn models
h_{max}	=	maximum lag for a specified direction
H	=	intermittency exponent
i	=	indicator function
i_s	=	search increment
I_1	=	first random integer for exchange mechanism
I_2	=	second random integer for exchange mechanism
L	=	dimension of simulation region
m_1	=	number of iterations for $\Delta E_{initial} \leq 0$
m_2	=	number of iterations for $\Delta E_{initial} > 0$
M_a	=	maximum number of accepted iterations per step
$M_{initial}$	=	number of iterations for evaluation of initial control parameter
M_t	=	maximum number of total iterations per step

N_c = number of conditioning data

N_d = number of semi-variogram directions

N_f = number of classes for distribution function

N_h = number of semi-variogram lag distances

N_o = number of constraints for a conditional simulation

N_p = number of pairs in a semi-variogram

N_s = total number of simulation grid points

N_x = number of grid points in the x direction

N_y = number of grid points in the y direction

N_z = number of grid points in the z direction

O_x = x coordinate of origin of simulation grid system

O_y = y coordinate of origin of simulation grid system

O_z = z coordinate of origin of simulation grid system

P = probability function for Metropolis condition

R = random number from a uniform distribution (0,1]

r_s = search radius fraction

S = sill for semi-variogram model

T = control parameter

T^0 = initial control parameter

V = simulation variable

V_c = conditioning data value

V_f = upper class limit of distribution function

V_H = scaling factor for fGn and fBm semi-variogram models

W_1 = temporary variable for simulation process

W_2 = temporary variable for simulation process

x = x coordinate of a grid point in simulation region

y = y coordinate of a grid point in simulation region

z = z coordinate of a grid point in simulation region

α = convergence rate factor

δ = smoothing parameter for fGn semi-variogram model

Δ_x = x direction spacing in simulation grid system

Δ_y = y direction spacing in simulation grid system

Δ_z = z direction spacing in simulation grid system

ΔE = change in objective function for an exchange

$\Delta E_{initial}$ = change in objective function for initial conditions

$\overline{\Delta E}^{+}_{initial}$ = mean positive change in objective function for initial conditions

$\Delta V_{f,min}$ = minimum inter-class size

ϵ_a = tolerance for acceptance ratio

ϵ_v = tolerance for objective function

γ = semi-variogram

$\hat{\gamma}$ = estimator of semi-variogram

γ_o = specified semi-variogram model for a conditional simulation

γ_s = sample semi-variogram of simulation variable

γ'_s = sample semi-variogram of simulation variable after an exchange

χ = acceptance ratio

$\chi_{initial}$ = acceptance ratio for initial conditions

λ = correction terms for sample semi-variograms

$\lambda^{+}_{I_1}$ = component of semi-variogram correction term

$\lambda^{-}_{I_1}$ = component of semi-variogram correction term

$\lambda^{+}_{I_2}$ = component of semi-variogram correction term

$\lambda^{-}_{I_2}$ = component of semi-variogram correction term

$\vec{\omega}$ = coordinates of grid points in the simulation region

$\vec{\omega}_c$ = coordinates of the conditioning data

REFERENCES

1. Lake, L. W. and Carroll, H. B. Jr. (Editors): *Reservoir Characterization*, Academic Press, Inc. (1986) Orlando, FL.
2. Lake, L.W., Carroll, H.B. Jr. and Wesson, T.C. (Editors): *Reservoir Characterization II*, Academic Press, Inc. (1991) San Diego, CA.
3. Haldorsen, H.H. and Damsleth, E.: "Stochastic Modeling," Journal of Petroleum Technology (April 1990) 404-12.
4. Haldorsen, H.H., Brand, P.J. and Macdonald, C.J.: "Review of the Stochastic Nature of Reservoirs," *Mathematics in Oil Production*, S. Edwards and P. King (editors), Oxford Science Publications, Clarendon Press, Oxford (1988) 109-209.
5. Damsleth, E., Tjolsen, C.B., Omre, K.H. and Haldorsen, H.H.: "A Two-Stage Stochastic Model Applied to a North Sea Reservoir," paper SPE 20605 presented at the 1990 SPE Annual Technical Conference and Exhibition, New Orleans, LA, Sept. 23-26.
6. Journel, A.G. and Huijbregts, C.J.: *Mining Geostatistics*, Academic Press, New York (1978).
7. Journel, A.G.: *Fundamentals of Geostatistics in Five Lessons*, American Geophysical Union, Washington, D.C. (1989).
8. Isaaks, E.H. and Srivastava, R.M.: *Applied Geostatistics*, Oxford University Press, New York (1989).
9. Delhomme, J.P.: "Spatial Variability and Uncertainty in Groundwater Flow Parameters: A Geostatistical Approach," Water Resources Research (April 1979) 269-80.
10. Da Costa e Silva, A.J.: "A New Approach to the Characterization of Reservoir Heterogeneity Based on the Geomathematical Model and Kriging Technique," paper SPE 14275 presented at the 1985 Annual Technical Conference and Exhibition, Las Vegas, Sept. 22-25.

11. Mantoglou, A. and Wilson, J.L.: "The Turning Bands Method for Simulation of Random Fields Using Line Generation by a Spectral Method," Water Resources Research (Oct. 1982) 1379-94.

12. Farmer, C.L.: "The Generation of Stochastic Fields of Reservoir Parameters With Specified Geostatistical Distributions," paper presented at the 1987 IMA on The Mathematics of Oil Production conference, Robinson College, Cambridge, July 6-7.

13. Journel, A.G. and Alabert, F.G.: "New Method for Reservoir Mapping," Journal of Petroleum Technology (Feb. 1990) 212-18.

14. Journel, A.G. and Gómez-Hernández, J.J.: "Stochastic Imaging of the Wilmington Clastic Sequence," paper SPE 19857 presented at the 1989 Annual Technical Conference and Exhibition, San Antonio, Oct. 8-11.

15. Mandelbrot, B.B. and Wallis, J.R.: "Some Long-Run Properties of Geophysical Records," Water Resources Research (April 1969) 321-40.

16. Burrough, P.A.: "Fractal Dimensions of Landscapes and Other Environmental Data," Nature (Nov. 19, 1981) 240-42.

17. Hewett, T.A.: "Fractal Distribution of Reservoir Heterogeneity and Their Influence on Fluid Transport," paper SPE 15386 presented at the 1986 Annual Technical Conference and Exhibition, New Orleans, Oct. 5-8.

18. Tang, R.W., Behrens, R.A. and Emanuel, A.S. "Reservoir Studies With Geostatistics To Forecast Performance," SPE Reservoir Engineering Journal (May 1991) 253-258.

19. Payne, D.V., Edwards, K.A. and Emanuel, A.S.: " Examples of Reservoir Simulation Studies Utilizing Geostatistical Models of Reservoir Heterogeneity," Reservoir Characterization II, Lake, L.W., Carroll, H.B. Jr. and Wesson, T.C. (Editors), Academic Press, Inc., San Diego, CA (1991) 497-523.

20. Hewett, T.A. and Behrens, R.A.: "Conditional Simulation of Reservoir Heterogeneity With Fractals," SPE Formation Evaluation Journal (Sept. 1990) 217-225.

21. Emanuel, A.S., Alameda, G.K., Behrens, R.A. and Hewett, T.A.: "Reservoir Performance Prediction Methods Based on Fractal Geostatistics," SPE Reservoir Engineering Journal (Aug. 1989) 311-318.

22. Mattews, J.L., Emanuel, A.S. and Edwards, K.A.: "A Modeling Study of the Mitsue Stage 1 Miscible Flood Using Fractal Geometries," paper SPE 18327 presented at the 1988 Annual Technical Conference and Exhibition, Houston, Oct. 2-5.

23. Aasum, Y., Kelkar, M.G. and Gupta, S.P.: "An Application of Geostatistics and Fractal Geometry for Reservoir Characterization," SPE Formation Evaluation Journal (March 1991) 11-19.

24. Crane, S.D. and Tubman, K.M.: "Reservoir Variability and Modeling With Fractals," paper SPE 20606 presented at the 1990 Annual Technical Conference and Exhibition, New Orleans, Sept. 23-26.

25. Farmer, C.L.: "Numerical Rocks, The Mathematical Generation of Reservoir Geology," paper presented at the 1989 Joint IMA/SPE European Conference, Cambridge University, July 25-27.

26. Aarts, E. and Korst, J.: *Simulated Annealing and Boltzmann Machines: A Stochastic Approach to Combinatorial Optimization and Neural Computing*, John Wiley & Sons Ltd, Chichester (1989).

27. Kirkpatrick, S., Gelatt, C.D. Jr. and Vecchi, M.P.: "Optimization by Simulated Annealing," Science (May 13, 1983) 671-80.

28. Cerny, V.: "Thermodynamical Approach to the Traveling Salesman Problem: An Efficient Simulation Algorithm," Journal of Optimization Theory and Applications (1985) vol. 45, no. 1, 41-51.

29. Collins, N.E., Eglese, R.W. and Golden, B.L.: "Simulated Annealing – An Annotated Bibliography," *Simulated Annealing (SA) & Optimization: Modern Algorithms with VLSI, Optimal Design, and Missile Defense Applications*, M.E. Johnson (editor), American Sciences Press, New York (1988) 209-307.

30. Metropolis, N., Rosenbluth, A.W., Rosenbluth, M.N., Teller, A.G. and Teller, E.: "Equation of State Calculations by Fast Computing Machines," The Journal of Chemical Physics (June 1953) 1087-92.

31. Geman, S. and Geman, D.: "Stochastic Relaxation, Gibbs Distributions, and the Bayesian Restoration of Images," IEEE Transactions on Pattern Analysis and Machine Intelligence, vol. PAMI-, no. 6 (Nov. 1984) 721-41.

32. Press, W. H., Flannery, B. P., Teukolsky, S. A. and Vetterling, W. T.: *Numerical Recipes: The Art of Scientific Computing*, Cambridge University Press, Cambridge (1986).

33. Wikramaratna, R.S.: "ACORN- A New Method for Generating Sequences of Uniformly Distributed Pseudo-random Numbers," Journal of Computational Physics (1989) 83, 16-31.

34. Gómez-Hernández, J.J.: "A Stochastic Approach to the Simulation of Block Conductivity Fields Conditioned upon Data Measured at a Smaller Scale," Ph D Thesis, Branner Earth Sciences Library, Stanford University, (Oct. 1990).

35. Mandelbrot, B.B. and Van Ness, J.W.: "Fractional Brownian Motions, Fractional Noises and Applications," SIAM Review (Oct. 1968) 422-37.

36. Williams, J.K. and Dawe, R.A.: "Fractals - An Overview of Potential Applications to Transport in Porous Media," Transport in Porous Media, Reidel Publishing Company, vol. 1 (1986) 201-9.

37. Sahimi, M. and Yortsos, Y.C.: "Applications of Fractal Geometry to Porous Media: A Review," paper SPE 20476 presented at the 1990 Annual Technical Conference and Exhibition, New Orleans, Sept. 23-26.

38. Mandelbrot, B.B.: "A Fast Fractional Gaussian Noise Generator," Water Resources Research (June 1971) 543-53.

39. Hurst, H.E.: "Long-Term Storage Capacity of Reservoirs," Trans. Amer. Soc. Civil Eng., vol. 116 (1951) 770.

40. Feder, J.: *Fractals*, Plenum Press, New York (1988).

41. Gómez-Hernández, J. J. and Srivastava, M.R.: "isim3d: A 3-dimensional Multiple Indicator Simulation Program," Computer and Geosciences (1990), vol. 16, no. 4, 396-440.

42. Smith, L.: "Spatial Variability of Flow Parameters in a Stratified Sand," Mathematical Geology (1981) V. 13, No. 1.

43. Goggin, D.J., Chandler, M.A., Kocurek, G. and Lake, L.W.: "Patterns of Permeability in Eolian Deposits: Page Sandstone (Jurassic), Northeastern Arizona," SPE Formation Evaluation Journal (June 1988) 297-306.

44. Kittridge, M.G., Lake, L.W., Lucia, F.J. and Fogg G.E.: "Outcrop/Subsurface Comparisons of Heterogeneity in the San Andres Formation," SPE Formation Evaluation Journal (Sept. 1990) 233-40.

45. Deines, T.: Personal Communication (Feb. 1991).

46. Schlumberger: *Log Interpretation, Volume I - Principles*, Schlumberger Limited, New York (1972).

47. Greenkorn, R.A. and Johnson, C.R.: "Variation of a Natural Sandstone Reservoir Element: An Objective Analysis of Core Measurements," paper SPE 1577-G presented at the 1960 Annual Fall meeting, Denver, Oct. 2-5.

48. Abramowitz, M. and Stegun, I.A.: "Handbook of Mathematical Functions," Dover Publications (1970) New York.

APPENDIX A

IMPLEMENTATION OF CONDITIONAL SIMULATION METHOD

This appendix describes the capabilities and options of the computer program developed for the stochastic conditional simulation method described in Chapter II. The program generates distributions of a variable with all the constraints and specifications described in Chapter II in a one-, two- or three-dimensional rectangular grid system. This computer program is written in ANSI FORTRAN 77 and the required input data is specified in a template data file. The first section of this appendix provides a description of the input data requirements for the computer program. The second section describes an example of a two-dimensional conditional simulation to illustrate the input and output capabilities of the computer program.

Input Requirements

Input data required by the conditional simulation program, such as grid geometry and semi-variogram models, are specified in the template data file. In this file, the data is typed only at fields (or spaces) indicated to the right of the symbols " $->$ " and in a few cases the spaces below the symbols " $-----$ ". Figure A-1 shows a template data file for one of the examples discussed later in the report. Two optional data files to input the conditioning data and a cumulative distribution function are specified in this template data file.

The template data file consists of six parts: Grid Configuration, Conditioning Data, Distribution Function, Semi-Variogram Models, Simulation Parameters and Output Files. This section describes the input data for each part of the template data file.

Grid Configuration

The following data defines the geometry of a rectangular simulation grid system.

- Number of grid points along the x, y and z directions

These numbers must be integers and equal to or greater than one. For 1- or 2-dimensional simulation systems, the number of grid points along the direction(s)

```
+-----------------------------------------------------------+
|        Conditional Correlated Distributions               |
|          by Simulated Annealing Program                   |
|     Simulation Specifications Template Data File          |
+-----------------------------------------------------------+
|
| Grid Configuration ------------------------------------+
|                          Number of grid points         |
|                        x          y          z         |
|                     ---------- ---------- ----------+
|                        31         31         1
|
|                          Origin of coordinates
|                        x          y          z
|                     ---------- ---------- ----------+
|                       0.0        0.0        1.0
|
|                               Grid spacing
|                        x          y          z
|                     ---------- ---------- ----------+
|                       5.0        5.0        1.0
|
| Conditioning Data -------------------------------------+
|                                    Number of points   ->25
|                                    = 0 No conditioning
|                                              File name ->ch2d.dat
|
| Distribution Function ---------------------------------+
|                  Source for Cdf                        ->2
|                  = Source 1, Sample conditioning data
|                  =         2, Input
|
|                                 Cdf number of classes ->20
|                  Type of subclass distribution         ->2
|                  = Type 1, Discrete (class mark)
|                  =         2, Continous (linear)
|
|                File name for Cdf (only for Source 2) ->nr20.dis
| First class lower boundary (minimum, only for Source 2) ->100.0
|
| Semi-Variogram Models ---------------------------------+
|                                    Number of directions ->2
|
|  #      Direction  cosine          Search     Number of
|       x        y        z      radius incr.  structures
|  -- -------- -------- -------- ------ -----  ----------+
|   1       1.       0.       0.    .50     1       1
|   2       0.       1.       0.    .50     1       1
|   3
|   4
|   5
|   6
|   7
|   8
|   9
|                                                Nugget ->0.0
|
|                  Type
|                  = 1 Spherical
|                  = 2 Exponential
|                  = 3 Gaussian
|                  = 4 fBm ()
|                  = 5 fGn ()
|
|                #  Type     Sill          Range     Delta
|                           (Scalng)        (H)      (fGn)
|                -- ---- ------------ ------------ ------+
|                 1    2       30000.          20.    1.
|                 2    2       30000.          60.    1.
|                 3
|                 4
|                 5
|                 6
|                 7
|                 8
|                 9
|
| Simulation Parameters ---------------------------------+
|                                                  Seed ->87586
|                                Convergence rate factor ->0.50
|          Maximum number of accepted iterations cycles ->5.0
|                                    Solution tolerance ->0.0001
|                               Acceptance ratio tolerance ->0.0250
|
| Output Files ------------------------------------------+
|                  File name for summary of simulation ->dsch2ds3.spe
|                                  simulated variable  ->dsch2ds3.out
|                      Output mode                     ->1
|                      = Mode 1, Do not overwrite files
|                      =      2, Overwrite files
|
+-----------------------------------------------------------+
```

Figure A-1 - Template data file with specifications for conditional simulation example

not included (e.g., z-direction for a x-y simulation) must be equal to one. The total number of grid points of the system is the product of the number of grid points along the three directions.

• Origin of coordinates for the x, y and z axis

The three origins must be specified even for 1- or 2-dimensional simulation systems.

• Grid spacing along the x, y and z directions

These are the distances between two consecutive grid points along the x, y and z directions. The three grid spacings (non-zero numbers) must be specified even for 1- or 2-dimensional simulation systems.

Note that the data specified in this part and in the other parts (locations of conditioning data and correlation range of the semi-variogram models) must be in the same units. The dimensions of the simulation grid system (L_x, L_y and L_z) along the x, y and z directions are

$$L_x = (N_x - 1)|\Delta_x|, \qquad (A-1)$$

$$L_y = (N_y - 1)|\Delta_y|, \qquad (A-2)$$

and

$$L_z = (N_z - 1)|\Delta_z|, \qquad (A-3)$$

respectively. Where, N_x, N_y and N_z are the number of grid points and Δ_x, Δ_y and Δ_z are the grid spacings along x, y and z directions, respectively.

Conditioning Data

The following data specifies the conditioning data, i.e., values of the variable at given locations which are fixed throughout the simulation.

• Number of conditioning data points

This number must be equal to the number of conditioning data points present in the file specified below or must be equal to zero for unconditional simulations.

• Name of file containing conditioning data

This file consists of four columns of data. The first three columns are the x, y and z coordinates of the conditioning data point and the fourth column is the value of the variable. These x, y and z coordinates must be in the same units used in the Grid Configuration part and inside the boundaries of the simulation grid.

Note that the coordinates of the conditioning data are transformed to discrete grid point indexes (i, j and k) according to

$$i = \frac{(x - O_x)}{\Delta_x} + 1, \qquad (A-4)$$

$$j = \frac{(y - O_y)}{\Delta_y} + 1 \qquad (A-5)$$

and

$$k = \frac{(z - O_z)}{\Delta_z} + 1 \qquad (A-6)$$

for the x, y and z directions, respectively. Where, x, y and z are the coordinates of a conditioning data point, O_x, O_y and O_z are the origins of the simulation grid and Δ_x, Δ_y and Δ_z are the grid spacings along the x, y and z directions, respectively. Obviously, i, j and k should be between 1 and N_x, N_y and N_z, respectively, for every conditioning data point.

Distribution Function

The following data defines the discrete cumulative density function (Cdf) of the variable to be simulated.

• Source for Cdf

The cumulative distribution function (Cdf) can be calculated by the program from the sample Cdf of the conditioning data (Source 1) or it can be specified in a file (Source 2). If the Cdf is specified in a file, then the file name and the lower boundary of the first class must be specified (see data below).

• Cdf number of classes

The number of classes (N_f) discretizes the Cdf into classes with inter-quantile sizes equal to $1/N_f$. The i^{th} quantile of the discrete Cdf is equal to i/N_f, for i=1,...N_f.

• Type of subclass distribution

Within each Cdf class the distribution can be approximated by a single number equal to the class mark (Type 1) or by linear interpolation of the class boundaries (Type 2). The class mark is taken as the mean of the lower and upper boundaries of a class. Note that lower boundary of the first class equals to the minimum of the simulated variable and the upper boundary of the last class equals to the maximum. The within class linear interpolation is equivalent to a uniform distribution within a class.

• File name for Cdf

This file is required if the Source 2 option is specified above. Each line of this file contains the value of the upper class boundary of each Cdf class, starting with the first class and arranged in a strictly increasing order. Therefore, the value in the last line will be the maximum of the distribution. Note that the number of lines in this file must be equal to the number of Cdf classes (N_f) specified above.

• First class lower boundary

This value is the minimum of the distribution only when the Cdf is specified in a file (Source 2 option).

Semi-Variogram Models

The following data specifies the semi-variogram models to be simulated along several directions.

• Number of directions

Specifies the number of different directions along which a semi-variogram needs to be simulated. For each direction, the remaining parameters in this section must be input. If the number of directions is equal to zero, the remaining parameters in this section are ignored and the simulated variable will be uncorrelated.

• Direction cosines

These parameters are the cosines of the angles between a vector pointing in the direction in which the semi-variogram is to be simulated and each major axis (x, y

and z). Note that the sum of the squares of the x, y and z direction cosines must be equal to one. For example, to specify directions which coincide with the positive x-, y- and z-axis, the x, y and z direction cosines are (1,0 0), (0,1,0) and (0,0,1), respectively.

• Search radius

The search radius is defined as the lag distance fraction of the maximum lag along the specified direction. This parameter must be between zero and one. For example, a semi-variogram evaluated along the x-direction up to a lag distance h_x, has a search radius (r_s) equal to

$$r_s = \frac{h_x}{h_{max}}, \qquad (A-7)$$

where, h_{max} is the maximum lag distance equal to

$$h_{max} = (N_x - 1)|\Delta_x|. \qquad (A-8)$$

• Search increment

The search increment is an index to calculate successive lags within the search radius. For most simulations this parameter should be equal to one. For example, the i^{th} lag (h_i) for a semi-variogram along the x-direction with a search increment equal to i_s is

$$h_i = i\ i_s|\Delta_x|. \qquad (A-9)$$

• Number of structures

The number of structures specifies the number of semi-variogram models to be nested along a direction. Currently, the number of structures must be equal to one or two. For each direction, the semi-variogram model (γ) is calculated as

$$\gamma(h) = C_o + \gamma_1(h) + \gamma_2(h), \qquad (A-10)$$

where, h is the lag distance, C_o is the nugget, and γ_1 and γ_2 are semi-variogram models of the types specified below for the two structures. If only one structure is specified then $\gamma_2(h) = 0$ in Equation A-10.

• Nugget

The nugget must be equal to or greater than zero and it is included in the semi-variogram model of each direction.

• Type

The type of semi-variogram models must be specified in the same order as the directions specified above. If there are two structures along a direction, then the type of each structure must be specified on consecutive lines. The five types of semi-variogram models available are:

(1) Spherical,

$$\gamma(h) = 0, \quad h = 0,$$
$$\gamma(h) = S[\frac{3}{2}(\frac{h}{a}) - \frac{1}{2}(\frac{h}{a})^3], \quad 0 < h < a \qquad (A-11)$$
$$\gamma(h) = S, \quad h \geq a,$$

(2) Exponential,

$$\gamma(h) = S[1 - \exp(-3\frac{h}{a})], \qquad (A-12)$$

(3) Gaussian,

$$\gamma(h) = S[1 - \exp(-3(\frac{h}{a})^2)], \qquad (A-13)$$

(4) Fractional Brownian motion (fBm),

$$\gamma(h) = V_H h^{2H} \qquad (A-14)$$

and (5) Fractional Gaussian noise (fGn),

$$\gamma(h) = \frac{1}{2}V_H \delta^{2H-2}[2 - (\frac{|h_1|}{\delta} + 1)^{2H} + 2|\frac{h_1}{\delta}|^{2H} - (\frac{|h_1|}{\delta} - 1)^{2H}]. \qquad (A-15)$$

Where, $\gamma(h)$ is the semi-variogram for a lag distance h, S is the sill minus the nugget, a is the correlation range or practical range (i.e., h at which 95% of the sill is reached for the exponential and Gaussian models), V_H is the scaling factor for fBm and fGn models, H is the intermittency exponent, h_1 is the lag divided by the distance between two consecutive points along the direction considered (e.g.,

for a semi-variogram direction which coincides with the x axis, this distance is Δ_x) and δ is the smoothing factor for fGn models. Note, h_1 is used in the fGn models instead of h so that δ can be expressed as the number of points used for smoothing, independently of the grid spacing.

• Sill, s (or Scaling Factor, V_H)

This parameter is the sill minus the nugget (S) for spherical, exponential, Gaussian models. For fBm and fGn models this parameter is a scaling factor (V_H), as defined in Equations A-14 and A-15.

• Range, a (or Intermittency Exponent, H)

This parameter is the correlation range or practical range (a) for spherical, exponential and Gaussian models and it is the intermittency exponent (H) for fBm and fGn models. Note that the range must be in the same units as the data of the Grid Configuration.

• Delta, δ (only for fGn Models)

This parameter is the smoothing factor (δ) for fGn models. The data field corresponding to δ must contain a numerical value for any model, but it is ignored for models other than a fGn model.

It should be noted that simulations become computationally more demanding as the number of directions and the search radius increase.

Simulation Parameters

Simulated annealing is an iterative stochastic method consisting of a discrete number of steps. At each step many iterations are conducted until equilibrium is achieved. Steps and iterations continue until one of the convergence criteria is satisfied.

The following data defines the simulation parameters, referred as the annealing schedule.

• Seed

This seed initializes the uniform random number generator.[33] Different seeds can be used to generate different realizations of the simulated variable. The seed must be an integer number in the range of the computer accuracy.

• Convergence rate factor

This factor controls the size of the simulation steps. The convergence rate factor must be a number between zero and one. A factor close to zero gives large steps, and close to one gives small steps. Depending on the number of iterations per step, too large steps may yield poor simulations and too small steps may take long running times. The results in Chapter II indicate that a convergence rate factor equal 0.50 provides optimum simulations when the maximum number of accepted iterations per step is 5.0 cycles.

• Maximum number of accepted iterations per step

This parameter controls the number of accepted iterations at each step of the simulation, required to assume equilibrium conditions. This parameter must be specified in units of cycles. One cycle is defined as the number of iterations equal to the number of grid points in the system ($N_x N_y N_z$). The results of Chapter II indicate that a maximum number of accepted iterations per step equal to 5.0 cycles provides optimum simulations when the convergence rate factor is equal 0.50.

• Solution tolerance

The solution tolerance specifies the value of the root mean square of the difference between the sample and model semi-variograms normalized by the model semi-variogram below which convergence is assumed and the simulation is stopped. This tolerance should be a number close to zero (e.g., 1×10^{-4}).

• Acceptance ratio tolerance

The acceptance ratio tolerance allows to stop a simulation when the fraction of accepted iterations per step becomes smaller than this tolerance. This tolerance should be a small number, such as 2.5×10^{-2} (i.e., the simulation stops when less than 2.5 % of the iterations in a step are accepted).

Note that simulations become computationally more demanding as the value of the convergence rate factor and the maximum number of accepted iterations per step increases.

Output Files

The following data specifies the files to store the output from the simulation.

• File name for summary of simulation

This file includes the following information: input data specifications for the simulation, univariate statistics of the conditioning data and simulated variable, model and sample semi-variogram of the simulated variable before and after the simulation and statistics of the simulation (e.g., total number of cycles).

• File name for simulated variable

This file contains the values of the simulated variable. Each line of the file contains one value of the variable, and it is arranged with the x, y and z coordinates cycling first, second and third, respectively. Starting with the first line of the file, the grid point index coordinates of the values in each line are: (1,1,1), (2,1,1) ... (N_x,1,1), (1,2,1), (2,2,1) ... (N_x,2,1) ... ... (1,N_y,N_z) ... (N_x,N_y,N_z). The simulation grid coordinates corresponding to a point with indices (i,j,k) can be obtained by solving for (x,y,z) from Equations A-4 through A-6.

• Output mode

The output Mode 1 option is used to avoid overwriting an existing file with one of the files specified above. If the output Mode 2 option is specified then existing files will be overwritten.

Conditional Simulation Example

This section describes the input and results of 2-dimensional conditional simulations generated with the computer program described in the previous section. The objective is to simulate distributions of permeability in the Chandler field test site[47]

for different anisotropy ratios. The template data file with the input specifications for one simulation is shown in Figure A-1.

The dimensions of the simulation region are 150 by 150 feet and it has been discretized with a 5 by 5 feet grid spacing (31 by 31 grid points). The locations of the 25 conditioning wells are shown in Figure A-2. The conditioning data, given in Table A-1, are average permeabilities from cores along each well. The mean, the standard deviation, the minimum and the maximum of the conditioning data are 381.24 md, 181.105 md, 100 md and 714 md, respectively.

The specified cumulative distribution function (Cdf) is the discrete Normal Cdf given in Table A-2 consisting of 20 classes, a mean and a variance equal to those of the conditioning data values. The limits of the Cdf are calculated by multiplying the limits of a Normal Cdf of zero mean and unit variance (second column of Table A-2 obtained from Reference 48) by the required standard deviation and adding the required mean. Figure A-3 compares the specified Cdf and the Cdf of the variable for one simulation and indicates a close agreement.

Semi-variograms are specified for the x- and y-directions with exponential models of sills equal to 30,000 md^2. The correlation ranges for the x- and y-directions for three anisotropy ratios (range in the x-direction divided by range in the y-direction) are 60 and 20 feet (anisotropy ratio equal to 3), 20 and 60 feet (anisotropy ratio equal to 1/3) and 60 and 60 feet (anisotropy ratio equal to 1), respectively. These models and the sample semi-variogram of the variable for the simulation with an anisotropy ratio equal to 3 are are shown in Figure A-4. The semi-variograms for the other anisotropy ratios are also in close agreement.

The simulations of permeability for the three anisotropy ratios are shown in the gray scale maps (colors are scaled between white and black linearly between the minimum and maximum values, respectively) in Figure A-5. The three simulations have identical cumulative distribution functions (and identical initial distributions at the beginning of the simulation). The effect of anisotropy is clearly indicated by the horizontal and vertical stripes in Figures A-5(a) and A-5(b), respectively. In the isotropic simulation, Figure A-5(c), the continuity of low and high values are influenced by underlying continuities in the conditioning data.

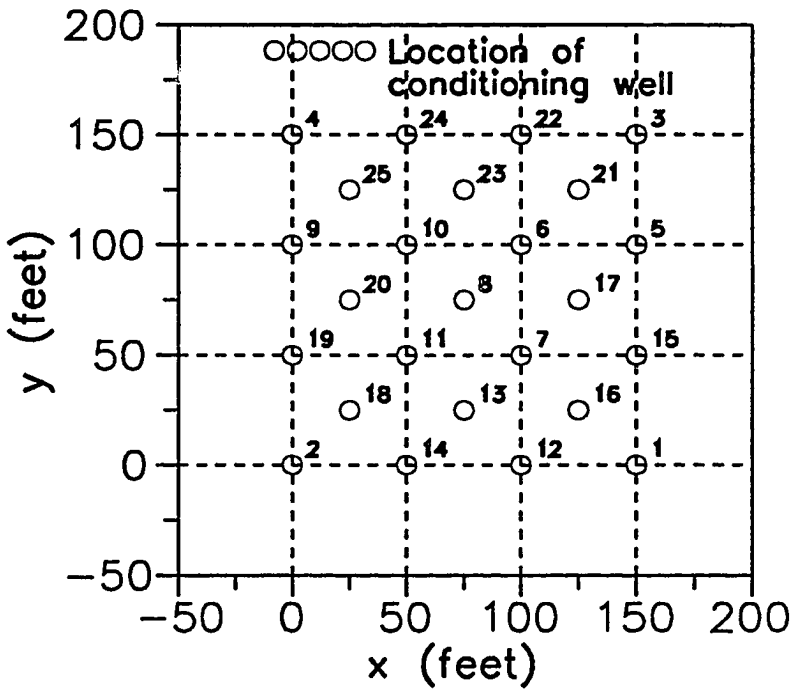

Figure A-2 - Locations of the conditioning data (the simulation region is the area between 0 and 150 feet for the x and y directions)

Table A-1

Permeability conditioning data for

two-dimensional simulation example

Well #	x (feet)	y (feet)	Permeability (md)	Well #	x (feet)	y (feet)	Permeability (md)
1	150	0	331	14	50	0	100
2	0	0	315	15	150	50	373
3	150	150	132	16	125	25	460
4	0	150	233	17	125	75	714
5	150	100	122	18	25	25	168
6	100	100	368	19	0	50	456
7	100	50	424	20	25	75	364
8	75	75	700	21	125	125	657
9	0	100	249	22	100	150	553
10	50	100	592	23	75	125	292
11	50	50	205	24	50	150	465
12	100	0	598	25	25	125	252
13	75	25	408				

Table A-2

Normal cumulative distribution function for

two-dimensional simulation example

Cumulative Distribution Function	Limit Mean=0 Std. Dev.=1	Limit (md) Mean=381.24 md Std. Dev.=181.105 md
0.05	-1.64485	100.000*
0.10	-1.28155	149.145
0.15	-1.03643	193.537
0.20	-0.84162	228.818
0.25	-0.67449	259.086
0.30	-0.52440	286.269
0.35	-0.38532	311.457
0.40	-0.25335	335.357
0.45	-0.12566	358.482
0.50	0.00000	381.240
0.55	0.12566	403.998
0.60	0.25335	427.123
0.65	0.38532	451.023
0.70	0.52440	476.211
0.75	0.67449	503.394
0.80	0.84162	533.662
0.85	1.03643	568.943
0.90	1.28155	613.335
0.95	1.64485	679.131
1.00**	3.00000	714.000*

* Limits adjusted to honor minimum and maximum of the conditioning data

** Approximate value, actual value equals 0.99865

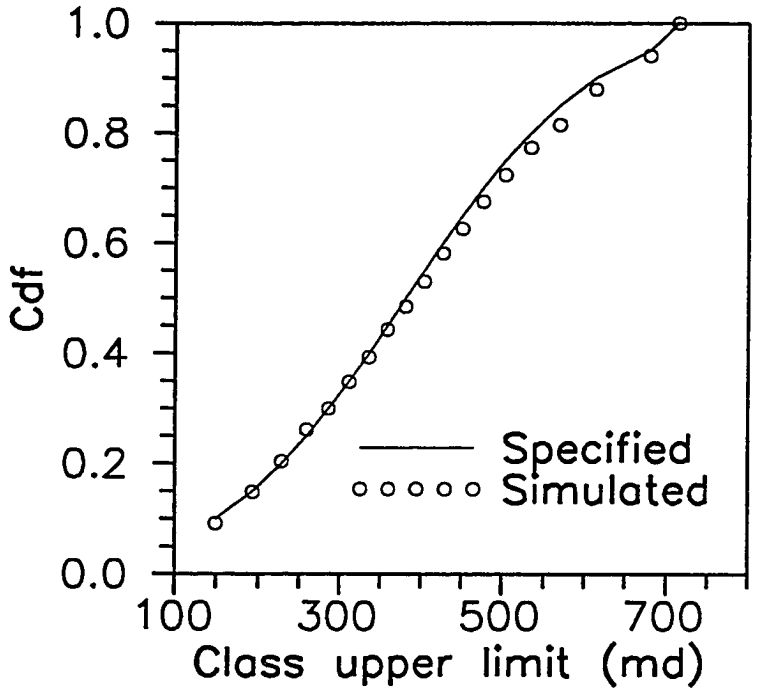

Figure A-3 - Specified and simulated cumulative distribution function (Cdf) for conditional simulation example

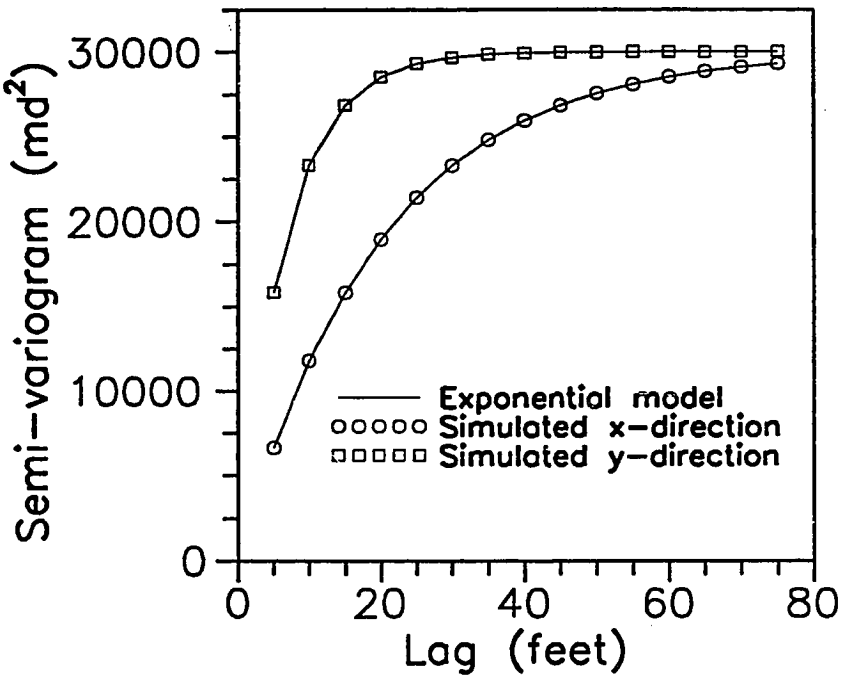

Figure A-4 - Specified and simulated semi-variograms for conditional simulation example with anisotropy ratio equal to 3 (a_x/a_y=3)

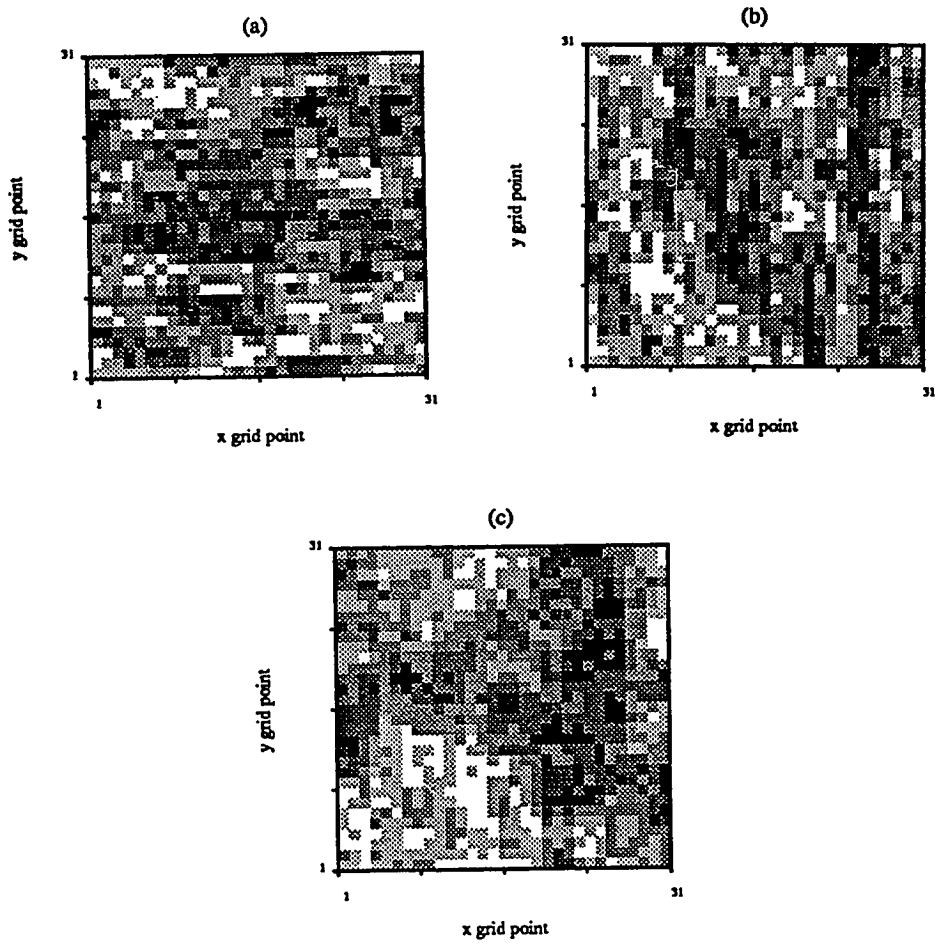

Figure A-5 - Maps of permeability for conditional simulations with anisotropy ratios equal to (a) a_x/a_y=3, (b) a_x/a_y=1/3 and (c) a_x/a_y=1

The output file containing the simulation specifications, univariate statistics of the conditioning data and the simulated variable, the initial and final sample semi-variograms of the simulated variable and the statistics of the simulation are shown in Figure A-6.

```
-------------------------------------------------------------
3-D M-D Conditional Simulated Annealing - Specifications
-------------------------------------------------------------

................................................
Input/Output Files
................
Parameters specification file name ...= sani.dat
Simulated distribution             ...= dsch2ds2.o
................................................
................................................................
Grid Configuration
................
                                      Direction
                                  x          y          z
                              .......... .......... ..........
Number of grid points ...=          31         31          1
Origin ..................=   .0000E+00   .0000E+00   .1000E+01
Grid spacing ............=   .5000E+01   .5000E+01   .1000E+01
System size .............=   .1500E+03   .1500E+03   .0000E+00

Total number of grid points ...........=         961
................................................................
................................................
Conditioning Data
................
Input file with conditioning data ...= ch2d.dat
Number of conditioning points .......=         25
................................................
................................................
Distribution Function
..................
Source for Cdf ..................=     Input
Type of subclass distribution ...= Continous
Cdf number of classes ...........=        20
File name for Cdf ...............= nr20.dis
Input Cdf
.........
 N     L. Bound      U. Bound    Class Mark       Cdf
 .     ........      ........    ..........       ...
 1     .100E+03      .100E+03      .100E+03       .05000
 2     .100E+03      .149E+03      .125E+03       .10000
 3     .149E+03      .194E+03      .171E+03       .15000
 4     .194E+03      .229E+03      .211E+03       .20000
 5     .229E+03      .259E+03      .244E+03       .25000
 6     .259E+03      .286E+03      .273E+03       .30000
 7     .286E+03      .311E+03      .299E+03       .35000
 8     .311E+03      .335E+03      .323E+03       .40000
 9     .335E+03      .358E+03      .347E+03       .45000
10     .358E+03      .381E+03      .370E+03       .50000
11     .381E+03      .404E+03      .393E+03       .55000
12     .404E+03      .427E+03      .416E+03       .60000
13     .427E+03      .451E+03      .439E+03       .65000
14     .451E+03      .476E+03      .464E+03       .70000
15     .476E+03      .503E+03      .490E+03       .75000
16     .503E+03      .534E+03      .519E+03       .80000
17     .534E+03      .569E+03      .551E+03       .85000
18     .569E+03      .613E+03      .591E+03       .90000
19     .613E+03      .679E+03      .646E+03       .95000
20     .679E+03      .714E+03      .697E+03      1.00000
.........................................
.................................................................
Variogram Models
```

Figure A-6 - Computer program output for conditional simulation example

```
................
Number of search directions ...=          2
Nugget ......................=          .00

#   Direction  cosine      Search       Type       Sill       Range    Delta
    x       y      z    radius inc.             (Scalng)      (H)     (fGn)
.. ...... ...... ...... ...... .... ........ ......... ......... .....
1 1.0000  .0000  .0000  .5000    1 Expntial  .300E+05  .600E+02  1.00
2  .0000 1.0000  .0000  .5000    1 Expntial  .300E+05  .600E+02  1.00
..........................................................................
.............................................................
Simulation Schedule
.................
Seed .........................................=      87586
Convergence rate factor ......................=  .5000E+00
System tolerance .............................=  .1000E-03
Acceptance ratio tolerance ...................=     .02500
Max. # of accepted iteration cycles per step ...=   5.0000
.............................................................
..........................................................
Current Storage Limits
.....................
Maximum number of simulation grid points ......=     6200
                  conditioning data points ....=      200
                  distribution classes ........=      100
                  searching directions ........=        9
                  lags per search direction ...=      100
..........................................................
.....................................................
Univariate Statistics of Conditioning Data
...................................
Number .....................=            25
Mean .......................=   .38124E+03
Variance ...................=   .32799E+05
Standard deviation .........=   .18110E+03
Coefficient of variation ...=   .47504E+00
Minimum ....................=   .10000E+03
Maximum ....................=   .71400E+03
Median .....................=   .36800E+03
1st quartile ...............=   .23300E+03
3rd         ...............=   .46000E+03
........................................................
.....................................................
Univariate Statistics of Initial Pattern
...................................
Number .....................=           961
Mean .......................=   .39084E+03
Variance ...................=   .29005E+05
Standard deviation .........=   .17031E+03
Coefficient of variation ...=   .43575E+00
Minimum ....................=   .10000E+03
Maximum ....................=   .71400E+03
Median .....................=   .38874E+03
1st quartile ...............=   .25317E+03
3rd         ...............=   .51963E+03
Sample Cdf
.........
 N     L. Bound      U. Bound    Class Mark       Cdf
 .     ........      ........    ..........       ...
 1     .100E+03      .100E+03     .100E+03       .00000
 2     .100E+03      .149E+03     .125E+03       .09157
 3     .149E+03      .194E+03     .171E+03       .14880
```

Figure A-6 - Computer program output for conditional simulation example (cont'd)

```
  4     .194E+03      .229E+03      .211E+03       .20395
  5     .229E+03      .259E+03      .244E+03       .26119
  6     .259E+03      .286E+03      .273E+03       .29969
  7     .286E+03      .311E+03      .299E+03       .34860
  8     .311E+03      .335E+03      .323E+03       .39334
  9     .335E+03      .358E+03      .347E+03       .44329
 10     .358E+03      .381E+03      .370E+03       .48491
 11     .381E+03      .404E+03      .393E+03       .53070
 12     .404E+03      .427E+03      .416E+03       .58169
 13     .427E+03      .451E+03      .439E+03       .62643
 14     .451E+03      .476E+03      .464E+03       .67534
 15     .476E+03      .503E+03      .490E+03       .72320
 16     .503E+03      .534E+03      .519E+03       .77315
 17     .534E+03      .569E+03      .551E+03       .81478
 18     .569E+03      .613E+03      .591E+03       .87929
 19     .613E+03      .679E+03      .646E+03       .94069
 20     .679E+03      .714E+03      .697E+03      1.00000
 ..................................................

 ..............................................................
 Initial Pattern Variography
 ...........................

 Search Direction #:  1
 x-index ...=    1 -> x-angle ...=    .00 deg.
 y-      ...=    0 -> y-      ...=  90.00 deg.
 z-      ...=    0 -> z-      ...=  90.00 deg.

              Error Function = (Sample/Model  - 1)**2

#          Lag              Semi-variogram          # of     Error
    Relative  Actual       Model       Sample      Pairs    Function
... ........ ........ ............ ............ ........ ............
  1        1      5.0   .66360E+04   .27953E+05      930   .10319E+02
  2        2     10.0   .11804E+05   .29535E+05      899   .22562E+01
  3        3     15.0   .15829E+05   .28689E+05      868   .66006E+00
  4        4     20.0   .18964E+05   .29714E+05      837   .32138E+00
  5        5     25.0   .21405E+05   .28744E+05      806   .11758E+00
  6        6     30.0   .23306E+05   .29155E+05      775   .62980E-01
  7        7     35.0   .24787E+05   .31998E+05      744   .84646E-01
  8        8     40.0   .25940E+05   .29999E+05      713   .24489E-01
  9        9     45.0   .26838E+05   .30368E+05      682   .17302E-01
 10       10     50.0   .27537E+05   .29394E+05      651   .45442E-02
 11       11     55.0   .28082E+05   .27984E+05      620   .12158E-04
 12       12     60.0   .28506E+05   .29110E+05      589   .44776E-03
 13       13     65.0   .28837E+05   .29426E+05      558   .41745E-03
 14       14     70.0   .29094E+05   .28537E+05      527   .36708E-03
 15       15     75.0   .29294E+05   .29957E+05      496   .51171E-03

 Search Direction #:  2
 x-index ...=    0 -> x-angle ...=  90.00 deg.
 y-      ...=    1 -> y-      ...=    .00 deg.
 z-      ...=    0 -> z-      ...=  90.00 deg.

              Error Function = (Sample/Model  - 1)**2

#          Lag              Semi-variogram          # of     Error
    Relative  Actual       Model       Sample      Pairs    Function
... ........ ........ ............ ............ ........ ............
 16        1      5.0   .66360E+04   .28953E+05      930   .11310E+02
 17        2     10.0   .11804E+05   .29603E+05      899   .22738E+01
 18        3     15.0   .15829E+05   .28908E+05      868   .68274E+00
```

Figure A-6 - Computer program output for conditional simulation example (cont'd)

```
 19       4       20.0    .18964E+05    .28534E+05       837    .25468E+00
 20       5       25.0    .21405E+05    .28981E+05       806    .12527E+00
 21       6       30.0    .23306E+05    .29505E+05       775    .70752E-01
 22       7       35.0    .24787E+05    .29544E+05       744    .36832E-01
 23       8       40.0    .25940E+05    .29594E+05       713    .19839E-01
 24       9       45.0    .26838E+05    .28880E+05       682    .57876E-02
 25      10       50.0    .27537E+05    .27779E+05       651    .77173E-04
 26      11       55.0    .28082E+05    .28710E+05       620    .49917E-03
 27      12       60.0    .28506E+05    .29640E+05       589    .15805E-02
 28      13       65.0    .28837E+05    .27644E+05       558    .17095E-02
 29      14       70.0    .29094E+05    .28458E+05       527    .47733E-03
 30      15       75.0    .29294E+05    .29302E+05       496    .68620E-07

     ..................................................
     Simulation Statistics
     ......................
```

Tem Lev	Temp	Total # of Cycls	Accptc Ratio	St #	Energy Change	Energy Mean	Energy Std. Dev.
0	.831E-01	.000	1.000000	0	.1000E+01	.0000E+00	.1000E+01	.0000E+00
1	.831E-01	5.000	.992716	1	.1031E+01	.3055E-01	.1001E+01	.1972E-01
2	.416E-01	10.073	.983385	2	.1033E+01	.2403E-02	.9947E+00	.2248E-01
3	.208E-01	15.205	.963909	3	.1008E+01	-.2482E-01	.9860E+00	.2562E-01
4	.104E-01	20.496	.925467	4	.9272E+00	-.8091E-01	.9599E+00	.2059E-01
5	.520E-02	26.123	.862770	5	.8840E+00	-.4325E-01	.9007E+00	.1828E-01
6	.260E-02	32.338	.709526	6	.6701E+00	-.2139E+00	.7271E+00	.4569E-01
7	.130E-02	40.906	.485305	7	.2374E+00	-.4327E+00	.3477E+00	.9722E-01
8	.649E-03	55.278	.347911	7	.7164E-01	-.1657E+00	.1025E+00	.3140E-01
9	.325E-03	74.105	.265572	7	.1538E-01	-.5626E-01	.1776E-01	.1061E-01
10	.162E-03	98.769	.180871	8	.5955E-02	-.9427E-02	.5121E-02	.1264E-02
11	.812E-04	139.358	.110547	9	.2552E-02	-.3402E-02	.2373E-02	.5127E-03
12	.406E-04	213.359	.060817	10	.1459E-02	-.1094E-02	.1262E-02	.2620E-03
13	.203E-04	362.799	.031592	11	.5647E-03	-.8939E-03	.6262E-03	.1213E-03
14	.101E-04	562.798	.015905	12	.2576E-03	-.3071E-03	.3263E-03	.5195E-04

```
  * Acceptance Ratio = Fraction of accepted iterations
                       of total iterations per step (level)
  * St # = Number of steps reaching iterations limit

     Initial normalized energy ........=   .10000E+01
     Final                    ........=   .25762E-03
     Initial r.m.s. error .............=   .97731E+00
     Final                .............=   .25178E-03
     Total number of          swaps ...=      540849
                    accepted        ...=       62051
                    rejected        ...=      478798
     Number of temperature levels .....=          14
     Initial temperature ..............=   .83129E-01
     Final               ..............=   .10148E-04

     ..............................................................
     Final Pattern Variography
     .........................

     Search Direction #:  1
     x-index ...=    1 -> x-angle ...=    .00 deg.
     y-      ...=    0 -> y-      ...=  90.00 deg.
     z-      ...=    0 -> z-      ...=  90.00 deg.

               Error Function = (Sample/Model  - 1)**2
```

Figure A-6 - Computer program output for conditional simulation example (cont'd)

#	Lag		Semi-variogram		# of	Error
	Relative	Actual	Model	Sample	Pairs	Function
1	1	5.0	.66360E+04	.66372E+04	930	.32578E-07
2	2	10.0	.11804E+05	.11804E+05	899	.20722E-09
3	3	15.0	.15829E+05	.15830E+05	868	.30693E-08
4	4	20.0	.18964E+05	.18968E+05	837	.52798E-07
5	5	25.0	.21405E+05	.21411E+05	806	.93383E-07
6	6	30.0	.23306E+05	.23302E+05	775	.27644E-07
7	7	35.0	.24787E+05	.24784E+05	744	.15975E-07
8	8	40.0	.25940E+05	.25948E+05	713	.92027E-07
9	9	45.0	.26838E+05	.26834E+05	682	.19321E-07
10	10	50.0	.27537E+05	.27551E+05	651	.23923E-06
11	11	55.0	.28082E+05	.28093E+05	620	.14873E-06
12	12	60.0	.28506E+05	.28503E+05	589	.18034E-07
13	13	65.0	.28837E+05	.28831E+05	558	.43688E-07
14	14	70.0	.29094E+05	.29093E+05	527	.17887E-08
15	15	75.0	.29294E+05	.29297E+05	496	.62420E-08

Search Direction #: 2
x-index ...= 0 -> x-angle ...= 90.00 deg.
y- ...= 1 -> y- ...= .00 deg.
z- ...= 0 -> z- ...= 90.00 deg.

Error Function = (Sample/Model - 1)**2

#	Lag		Semi-variogram		# of	Error
	Relative	Actual	Model	Sample	Pairs	Function
16	1	5.0	.66360E+04	.66393E+04	930	.25680E-06
17	2	10.0	.11804E+05	.11799E+05	899	.16845E-06
18	3	15.0	.15829E+05	.15832E+05	868	.31922E-07
19	4	20.0	.18964E+05	.18961E+05	837	.18764E-07
20	5	25.0	.21405E+05	.21406E+05	806	.35939E-08
21	6	30.0	.23306E+05	.23297E+05	775	.14412E-06
22	7	35.0	.24787E+05	.24785E+05	744	.52439E-08
23	8	40.0	.25940E+05	.25944E+05	713	.21337E-07
24	9	45.0	.26838E+05	.26839E+05	682	.11061E-08
25	10	50.0	.27537E+05	.27552E+05	651	.28342E-06
26	11	55.0	.28082E+05	.28075E+05	620	.56612E-07
27	12	60.0	.28506E+05	.28505E+05	589	.98471E-09
28	13	65.0	.28837E+05	.28831E+05	558	.43292E-07
29	14	70.0	.29094E+05	.29090E+05	527	.18425E-07
30	15	75.0	.29294E+05	.29288E+05	496	.52878E-07

Figure A-6 - Computer program output for conditional simulation example (cont'd)

APPENDIX B

CORRECTION FORMULA FOR SEMI-VARIOGRAM

The objective of this appendix is to present the formulas used to calculate the correction terms required to account for the exchange of the values of two points of the simulation grid in the semi-variograms. The equation used to apply these correction terms is given by Equation 2.25 in the Interchange Mechanism section of Chapter II. In this appendix, the correction terms for the semi-variograms are derived for an arbitrary lag vector, $\vec{h}_a$. For this lag, Equation 2.25 can be written as

$$\gamma_s'(\vec{h}_a) = \gamma_s(\vec{h}_a) + \lambda(\vec{h}_a), \qquad (B-1)$$

where, γ_s' is the new semi-variogram if the two points are exchanged, γ_s is the current semi-variogram and λ is the correction term derived in this appendix. The location of the two exchange points is denoted by $\vec{\omega}_{I_1}$ and $\vec{\omega}_{I_2}$, where $1 \leq I_1 \leq N_s$ and $1 \leq I_2 \leq N_s$. The actual coordinates in the simulation grid of the points denoted by I_1 and I_2 can be calculated by Equations 2.1 through 2.3 and Equations 2.5 through 2.7.

The procedure to calculate the correction term (λ) consists of removing the effect of the value at one exchange point and adding the effect of the value of the other exchange point in the semi-variogram. This procedure is applied for each exchange point. For the first interchange point, I_1, the values required to correct the semi-variogram can be grouped into two individual components given by

$$\lambda_{I_1}^{+} = -[V(\vec{\omega}_{I_1}) - V(\vec{\omega}_{I_1} + \vec{h}_a)]^2 + [V(\vec{\omega}_{I_2}) - V(\vec{\omega}_{I_1} + \vec{h}_a)]^2 \qquad (B-2)$$

and

$$\lambda_{I_1}^{-} = -[V(\vec{\omega}_{I_1}) - V(\vec{\omega}_{I_1} - \vec{h}_a)]^2 + [V(\vec{\omega}_{I_2}) - V(\vec{\omega}_{I_1} - \vec{h}_a)]^2, \qquad (B-3)$$

where, V is the simulation variable. The positive and negative components, $\lambda_{I_1}^{+}$ and $\lambda_{I_1}^{-}$, in Equations B-2 and B-3 correspond to the two possible pair points around a lag $\vec{h}_a$. Equations B-2 and B-3 can be rearranged as

$$\lambda_{I_1}^{+} = V(\vec{\omega}_{I_2})^2 - V(\vec{\omega}_{I_1})^2 + 2V(\vec{\omega}_{I_1} + \vec{h}_a)[V(\vec{\omega}_{I_1}) - V(\vec{\omega}_{I_2})] \qquad (B-4)$$

and

$$\lambda_{I_1}^{-} = V(\vec{\omega}_{I_2})^2 - V(\vec{\omega}_{I_1})^2 + 2V(\vec{\omega}_{I_1} - \vec{h}_a)[V(\vec{\omega}_{I_1}) - V(\vec{\omega}_{I_2})], \qquad (B-5)$$

respectively. A more compact representation of Equations B-4 and B-5 is given by

$$\lambda_{I_1}^{+} = d_{1,2}V(\vec{\omega}_{I_1} + \vec{h}_a) - D_{1,2} \qquad (B-6)$$

and

$$\lambda_{I_1}^{-} = d_{1,2}V(\vec{\omega}_{I_1} - \vec{h}_a) - D_{1,2}, \qquad (B-7)$$

where, $d_{1,2}$ and $D_{1,2}$ are

$$d_{1,2} = 2[V(\vec{\omega}_{I_1}) - V(\vec{\omega}_{I_2})] \qquad (B-8)$$

and

$$D_{1,2} = [V(\vec{\omega}_{I_1}) + V(\vec{\omega}_{I_2})][V(\vec{\omega}_{I_1}) - V(\vec{\omega}_{I_2})]. \qquad (B-9)$$

For the second interchange point, I_2, a derivation similar to the one used to arrive at Equations B-6 and B-7 yields the following correction components

$$\lambda_{I_2}^{+} = -d_{1,2}V(\vec{\omega}_{I_2} + \vec{h}_a) + D_{1,2} \qquad (B-10)$$

and

$$\lambda_{I_2}^{-} = -d_{1,2}V(\vec{\omega}_{I_2} - \vec{h}_a) + D_{1,2}, \qquad (B-11)$$

Note that the components of the correction term given by Equations B-6, B-7, B-10 and B-11 are not simplified and combined into a single compact term, because for each component, it is necessary to check that the points $(\vec{\omega}_{I_1} + \vec{h}_a)$, $(\vec{\omega}_{I_1} - \vec{h}_a)$, $(\vec{\omega}_{I_2} + \vec{h}_a)$ and $(\vec{\omega}_{I_2} - \vec{h}_a)$ are inside the simulation region. If one of the points falls outside the simulation region, then the corresponding correction component is equal to zero.

The correction term in Equation B-1 is given by

$$\lambda(\vec{h}_a) = \frac{1}{2N_p(\vec{h}_a)}\left[\lambda_{I_1}^{+} + \lambda_{I_1}^{-} + \lambda_{I_2}^{+} + \lambda_{I_2}^{-}\right], \qquad (B-12)$$

where, N_p is the number of pairs for the lag $\vec{h}_a$. The number of pairs for each lag vector required in the correction term in Equation B-12 is known because this number is calculated when the sample semi-variogram of the initial distribution is computed.