GEOSTATISTICS OF POROUS MEDIA
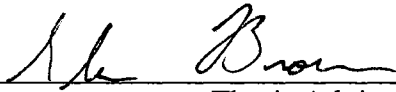
TOMOGRAPHY IMAGES   .

By

JASON R. VOGEL

Bachelor of Science
University of Nebraska
Lincoln, Nebraska
1995

Master of Science
Texas A&M University
College Station, Texas
1997

Submitted to the Faculty of the
Graduate College of the
Oklahoma State University
in partial fulfillment of
the requirement for
the Degree of
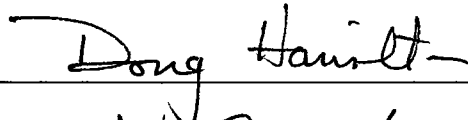DOCTOR OF PHILOSOPHY
December, 2001

GEOSTATISTICS OF POROUS MEDIA
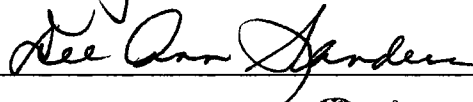
TOMOGRAPHY IMAGES

Thesis Approved:

_____
Thesis Advisor

_____

_____

_____

_____
Dean of the Graduate College

PREFACE

Mathematical models are used extensively to predict contamination and design cleanup of polluted aquifers. Such computer methods are generally sophisticated and theoretically accurate. The difficulty in applying these models, however, comes from the complexity of the subsurface environment. All the differential equations in the world are of no use if the modeler does not know the properties of the aquifer. The ultimate goal of this research is to start the development of methods that allow practitioners to analyze the complex properties of the subsurface from small cores using computer tomography (CT) and to then project those properties to predict the characteristics of a larger aquifer. Those projections can then be used as input in mathematical flow and transport models. A significant contribution toward completion of this goal is described within. This is accomplished while addressing many of the basic theoretical questions and concerns associated with this type of analysis, which heretofore had gone unanswered.

How big is big enough? When are objects larger than they appear? These questions are addressed in the context of rock and soil materials in Chapter 2. A method is developed to determine the size of sample needed for a particular problem. This methodology also results in a method of rock and soil classification based upon uniformity of size of the material. Image resolution is also shown to affect the size of an object that is detected so that objects may not be as large as they appear.

Are the objects that are seen really there? Soil and rock could be thought of as a mass of relatively large and small objects that are clumped together. Chapter 3 explores this question and develops a way to detect actual size of features within soil and rock materials, as opposed to current methods that basically match curves to assign size to the objects.

Are all objects created equal? Can we take measurements on something small and make something big? Chapter 4 shows that all objects within a soil and rock sample are not created equal. Methods are developed to determine the size and amount of features of different materials within the soil. Using this information, a larger projection of the soil or rock materials can be generated from a small sample. These projections have the potential to be used in mathematical models for pollution prediction and cleanup.

A contemporary publication format is followed in this dissertation. Chapter 1 introduces the problem and objectives of the entire dissertation along with providing recommendations for future research. Chapters 2, 3, and 4 are structured as individual refereed publications. The reference format follows that of *Water Resources Research*, to which much of this research will be directed. Finally, additional information and the Fortran computer codes used for these analyses are provided in the appendices.

I have been blessed to work for and with a talented pair of individuals in Dr. C. T. Haan and Dr. G. O. Brown during my time at Oklahoma State. I thank them for serving as excellent advisors, teachers, supervisors, mentors, and friends over the past five years. Also thanks to my other committee members, Drs. D. W. Hamilton and D. A. Sanders for their valuable contributions to this research. I also wish to acknowledge Sandia National

Laboratory and Dr. J. E. Gazin for providing the Culebra dolomite and concrete cores, respectively, used in much of these studies.

So many individuals are vital for the completion of an endeavor such as this. Being the first in your immediate family to complete college does not happen without a great deal of support. Uprooting oneself and moving several hundred miles away from home has also definitely strengthened my appreciation of family. To my parents, thank you for the gift of life and the determination and work ethic which you have instilled in me. May we continue to learn every day. Traci, thanks for all your love and support, the occasional late-night phone calls, and for always being there. Last but not least, thank you Anthony for being my "bud" and my best friend. I dedicate this to you all. Thanks to Mrs. Vollbrecht for seeing something in me a long time ago and starting me down the path of knowledge. Lifelong friendships are also one of the great benefits of my studies and have been forged from a variety of arenas. Thanks to Haggard, Duane, Derek, Van, Jenny, Boland, the Achafoa Chapter of Mortar Board, and my AGS brothers for continuing inspiration and support.

Most of all I would like to thank God for His love and forgiveness. He has definitely carried me when I needed it most.

Finally, I also dedicate this dissertation to my grandpa, Leonard Vogel, who with an eighth-grade education is still the smartest man I have ever met. It is the greatest gift that I can give to you to carry on your name as Dr. Vogel.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

NOMENCLATURE

a	Correlation length
a_x	Correlation length in the x direction
a_y	Correlation length in the y direction
a_z	Correlation length in the z direction
C	Calibration coefficient for CT images
Cov[...]	Covariance
CT	Computerized Tomography
D	Difference scores for Wilcoxon matched-pairs signed-ranks test
E[...]	Expected value
$f_m^*\left(\dfrac{\rho_m}{\rho_{max}}\right)$	True mixed voxel distribution
h	lag distance
$h_{\alpha\beta}$	lag distance from point x_α to point x_β
L{...}	Law, or distribution of a random variable
N(0,1)	Standard normal distribution
$N(h)$	Number of pairs for semivariogram calculation
$N_j(h)$	Number of pairs in bin j for conditional semivariogram calculation
r^2	Coefficient of determination
REV	Relative Elementary Volume

REV_{min}	Minimum volume of the REV region
R_N	Random value from standard normal distribution
s^2	Sample variance
SEM	Scanning Electron Micrograph
SEV	Standard Element Volume
$U(h)$	Standardized distribution of semivariogram values
Var[...]	Variance
V_i	Volume
V_{max}	Maximum value of the REV region
V_{min}	Minimum value of the REV region
w_i	Weighting function for combination semivariogram model
WIPP	Waste Isolation Pilot Plant
x_i	Data value at the tail for semivariogram calculation
y'	Random value from the beta distribution
y_i	Data value at the head for semivariogram calculation
$Z(x,V_i)$	Random function with volume V_i and a centroid at x
$Z(x)$	Intrinsically stationary random function
α	Fitting coefficient for the beta distribution
β	Fitting coefficient for the beta distribution
Γ	Gamma function
$\gamma(h)$	Semivariogram value
$\gamma_i(h)$	Individual semivariogram value in combination semivariogram
Λ	Scale disparity factor

$\mu(x,y,z)$	Attenuation at point (x, y, z)
$\rho_b(x,y,z)$	Bulk density at point (x, y, z)
ρ_{max}	Largest bulk density in a beta distribution
σ	Population standard deviation
$\Sigma R+$	Sum positive scores for Wilcoxon matched-pairs signed-ranks test
$\Sigma R+$	Sum negative scores for Wilcoxon matched-pairs signed-ranks test

CHAPTER ONE

INTRODUCTION

1.1 Stochastic Subsurface Modeling and Geostatistics

Stochastic modeling of subsurface flow and transport has been a subject of intensive research over the past twenty-five years. The word "stochastic" derives from the Greek word *stochos*, for a target stake, reflecting the uncertainty of the aim of an archer. Likewise, stochastic modeling considers the uncertainty and natural variability associated with hydrogeologic materials. In principle, this approach recognizes that hydrogeologic variables are affected by uncertainty and regards them as random (Dagan, 1997).

Geostatistics is a term first coined by the French statistician G. Matheron of the Ecole des Mines Supériur, Paris, France in the 1950s, and deals with the estimation of variables that depend on spatial location and exhibit a stochastic spatial structure. Classical geostatistics was developed for applications in mining applications and are not necessarily adequate for groundwater studies. Data availability and the scale of variation which influences flow and transport in subsurface hydrology is orders of magnitude smaller than in mining applications. Moreover, variables such as hydraulic conductivity and transmissivity can represent average values over the entire thickness of an aquifer and are subject to measurement and interpretation errors. For this reason, hydrologists have concentrated on methods for parameter estimation for spatial characterization of the

1

subsurface at these smaller scales of interest. Geostatistical parameter estimation involves quantification of the spatial correlation function (semivariogram), and analysis of the structure through trend detection, transformations, semivariogram modeling, and histogram analysis. The development of geostatistical parameter estimation methods has in turn led to simulation techniques that allow for equiprobable realizations of a variable for use in hydrologic flow and transport models (Calvete, 1997). Examples of simulation techniques include conditional simulation and simulated annealing (Goovaerts, 1997).

1.2 The Problem

The application of geostatistically estimated parameters to subsurface hydrologic modeling faces a number of obstacles. Structural analysis is highly subjective. The decision on what is actual statistical structure of a variable is arguable. There are also no geostatistical tests to check for stationarity of the variable (Calvete, 1997). Gelhar (1993) discusses the importance of the scale associated with the heterogeneity of a variable in a subsurface flow system. The scale of heterogeneity of media in which dispersion and diffusion are the dominant transport pathways will be different than that of media in which advection or macropore flow is the dominant transport pathway. In addition, the relatively small-scale processes are extremely variable in time and space, making it difficult to use local measurements directly and extrapolate them to large scale behavior. Finally, distinguishing individual structure characteristics of more than two phases within a media is difficult with current geostatistical methods.

1.3 Objectives

The overall objective of this dissertation is to develop geostatistical methods for more reliable estimation of spatial parameters associated with porous media using computerized tomography (CT) images. CT images provide high-quality, high-quantity data sets that are used to develop geostatistical methods for more reliable estimation of the spatial parameters associated with porous media. These parameters can then in the future be applied in simulation models for improved prediction of flow and transport in subsurface media.

Specific objectives of this study are:

1. Compare the geostatistical properties of similar cores at two different resolutions (CHAPTER TWO).

2. Determine the disparity of scale required for ergodicity of a sample through the use of relative elementary volume (REV) concepts (CHAPTER TWO).

3. Determine the effect of resolution on correlation lengths by addressing the modifiable areal unit problem (CHAPTER TWO).

4. Determine the maximum reliable lag on a semivariogram by analyzing the variance of the semivariogram families at each lag (CHAPTER THREE).

5. Automatically detect actual structure by analyzing the distribution of semivariogram families (CHAPTER THREE).

6. Determine structure size of components of different bulk density ranges using a conditional semivariogram (CHAPTER FOUR).

7. Create a stochastic representation of a porous media utilizing the conditional semivariogram and phantom images (CHAPTER FOUR).

3

1.4 Overall Conclusions

This research makes a significant contribution towards the ultimate goal of allowing practitioners to analyze small cores and to then create a large representation of an aquifer with similar spatial properties. Each of the results are original and without precedent in the literature. Likewise, each may be directly applied toward the goal of creating reasonable synthetic input data for computer modeling. This will provide an improvement over the arbitrary practices used today. Thus, the prediction of groundwater contamination at a site or the development of remediation strategies for environmental cleanup will be greatly enhanced.

1.5 Recommendations for Future Research

Many implications of these studies to future research are apparent. Specifically, the scale disparity factor can be explored and developed as a way to classify the heterogeneity of porous media. It is hypothesized that similar types of soil will have similar scale disparity factors. Furthermore, the inherent range can be investigated at scales larger and smaller than those addressed here by using different types of imaging such as scanning electron microscopes and different porous media. Additional analysis should be carried out on uniformly packed beads of a single size to confirm that predicted and measured correlation lengths are in agreement. Analysis should also be completed with conditional semivariogram families to determine if automatic correlation length detection is possible on conditional semivariograms. Simulated annealing or another

simulation technique can be performed to see how using the stochastic representation

from conditional semivariograms as an initial image compares to using a random image,

both in flow and transport results and computational time. Flow and transport studies on

stochastic representations can be also completed and compared to field data to determine

possible effects of changing scales on model results.

The methodology developed here has the potential to be applied to image analysis of a

variety of materials such as concrete, fruits and vegetables, and wood products. Possible

results of these analyses could include new ways of characterizing, classifying, and/or

sorting of these materials. Application to remotely sensed images could also be used to

characterize structure of variables on the earth's surface and detect large-scale changes in

global variables.

1.6 References

Calvete, F. J. S., Application of geostatistics in subsurface hydrology, in *Subsurface Flow and Transport: A Stochastic Approach*, ed. By G. Dagan and S. P. Neuman, Cambridge University Press, Cambridge, UK, 1997.

Dagan, G., Stochastic modeling of flow and transport: the broad perspective, in *Subsurface Flow and Transport: A Stochastic Approach*, ed. By G. Dagan and S. P. Neuman, Cambridge University Press, Cambridge, UK, 1997.

Gelhar, L. W., *Stochastic Subsurface Hydrology*, Prentice-Hall, Inc., Englewood Cliff, NJ, 1993.

Goovaerts, P., *Geostatistics for Natural Resources Evaluation*, Oxford University Press, New York, 1997.

CHAPTER TWO

GEOSTATISTICS OF COMPUTERIZED TOMOGRAPHY IMAGES OF THE CULEBRA DOLOMITE

2.1 Abstract

The three-dimensional geostatistical properties of samples of Culebra Dolomite have been determined at two different resolutions using computer tomography images of gamma ray attenuation. By fitting models to semivariograms, small-scale and large-scale correlation lengths are determined for four samples. Different geostatistical properties were found for adjacent cores at both resolutions. Relative elementary volume concepts are related to the stationarity of the sample. A scale disparity factor is defined and is used to determine sample size required for intrinsic stationarity with a specified correlation length. This allows for comparison of geostatistical measures and representative elementary volumes. The modifiable areal unit problem is also addressed. By changing resolution, a range of correlation lengths can be determined for the same sample. Comparison of voxel volume to the correlation length of a single sample at different resolutions reveals a linear scaling effect. Using this relationship, the inherent range is defined. This is the range approached as the voxel size goes to zero. This analysis shows the resolution and sample size should be a function of the scale of the intended data use.

2.2 Introduction

Computerized tomography (CT) is being increasingly used in porous media research as a nondestructive tool for measurement of interior properties of solid cores. *Petrovic et al.* [1982], *Crestana et al,* [1985], *Tollner and Verma* [1989], *Phogat et al.* [1991], *Brown et al.,* [1993], *Hsieh et al.* [1998b], and others have met with increasing success in producing images of the spatial distribution of solids, fluids, and voids within porous media columns. These images have been used to analyze many porous media phenomena such as permeability, pore space distribution, connectivity, and relative elementary volume [*Coskun and Wardlaw,* 1993; *Grevers and de Jong,* 1994; *Ioannidis et al.,* 1999; *Brown et al.,* 2000].

CT images characterize porous media content at a scale smaller than is possible with most physical sampling methods. This data provides a unique opportunity to use geostatistics, specifically the semivariogram, to determine the spatial structure and anisotropy of porous media at these smaller scales. *Schafmeister-Spierling and Pekdeger* [1989] have shown that dispersion can be modeled more accurately when small-scale spatial variability is considered. *Flühler et al.* [2001] also shows that mixing processes in the millimeter scale exert a pronounced effect on flow and transport behavior of water and solutes. For porous media in which macropore dispersion is a significant pathway of contaminant transport, such as Culebra Dolomite, this small-scale geostatistical characterization of the rock may yield improved aquifer characterization.

The large amount of spatial data provided by CT images also allows for investigation of the effect of resolution and scale on correlation lengths. *Gelhar* [1986] states that "spatial correlation scales should not be viewed in any absolute sense, but rather will

depend on the problem at hand." *Openshaw and Taylor* [1979] also show that aggregating fixed observations on small areas in various ways can produce a large range of correlation coefficients.

This paper uses CT images to characterize the geostatistical properties and anisotropy of Culebra Dolomite structure at core scales. With this information, the geostatistical properties of adjacent cores and cores at two different scales are compared. Through relative elementary volume concepts, the disparity scale required for stationarity of spatial data is quantified. Finally, the modifiable areal unit problem is addressed and the scaling factor associated with increasing scanning resolution is determined.

2.3 Theory

2.3.1 Semivariograms

The semivariogram measures the average degree of dissimilarity between an unsampled value and a nearby data value [*Deutsch and Journel*, 1998]. The traditional semivariogram is defined as half of the average squared difference between two data points approximately separated by a vector *h*:

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} (x_i - y_i)^2 \qquad (2.1)$$

where $\gamma(h)$ is half of the average squared difference; $N(h)$ is the number of pairs; x_i is the data value at the start, or tail; and y_i is the data value at the end, or head.

A sill and range characterize the semivariogram. The sill is a plateau in the semivariogram values that corresponds to the variance of the sampled data. If the semivariogram is standardized, the $\gamma(h)$ value is divided by the variance and the sill value will be equal to one. The range, or correlation length, is the separation distance at which the semivariogram reaches a plateau. A non-zero intercept on the semivariogram may exist and is termed the nugget effect. It may result from sampling error or variability at scales less than the smallest sampling interval [*Cressie*, 1991]. Another possible semivariogram characteristic is the hole effect, which is described as a semivariogram that peaks and dips, suggesting that at greater distances samples are more related. The hole effect may be seen as the rough spacing between adjacent lenses or bedding planes.

Semivariograms are fit by a number of analytical relationships to ensure that the semivariograms model is positive definite. The exponential model is often utilized by researchers in stochastic hydrology [*Woodbury and Sudicky*, 1991]. The model is given by

$$\gamma(h) = 1 - e^{\left(-\frac{3h}{a}\right)} \tag{2.2}$$

where a is the range and h is the lag distance. The exponential model reaches its sill asymptotically, with the practical range a defined as that distance at which the variogram value is 95% of the sill. Another common semivariogram model is the spherical model,

$$\gamma(h) = \begin{cases} 1.5\frac{h}{a} - 0.5\left(\frac{h}{a}\right)^3 & \text{if } h < a, \\ 1 & \text{elsewhere.} \end{cases} \tag{2.3}$$

The spherical model has a linear behavior near the origin, flattens out at larger distances, and reaches the sill at a [*Isaaks and Srivastava*, 1989].

Multiple, or nested, sills in the semivariogram may be associated with physical phenomena occurring at different scales. Detection of ranges nested within larger ranges requires sample spacing shorter than the minimum detected range. In general, nested sills are only considered when they can be associated with physical phenomena [*Solie et al.*, 1999]. Nested sills can be modeled by using a linear combination of positive definite semivariogram models with positive coefficients. This property results in the family of positive definite models

$$\gamma(h) = \sum_{i=1}^{n} |w_i| \gamma_i(h) \tag{4}$$

which are positive definite as long as the n individual models are all positive definite. A weighting function, w_i, is defined for each individual model subject to $\sum_{i=1}^{n} w_i = 1$ [*Isaaks and Srivastava*, 1989].

Interpretation of 3-D semivariograms of geologic data will yield information on subsurface anisotropy. Geometric anisotropy exists when the directional semivariograms show the same shape and sill, but different range values. An anisotropy factor may be used to characterize the geometric anisotropy and is defined as the ratio of the ranges in the major and minor direction. Another type of anisotropy is zonal anisotropy, where both the sill and range change with direction. A lower horizontal sill represents anisotropy due to geologic strata. The distance between the horizontal (lower) and vertical (upper) sills represents the variability between different geologic strata and may be related to laterally continuous sedimentary features. When the vertical semivariogram

reaches a lower sill than the horizontal semivariogram, zonal anisotropy is present and is due to areal trends. Geometric and both types of zonal anisotropy typically exist simultaneously in the field and complicate analysis [*Kupfersburger and Deutsch*, 1999; *Goggin et al.*, 1988].

Recent studies have also shown that the size of the sampled area (or volume) causes a scaling effect on semivariogram correlation lengths. The size of the sampled area influences the correlation scale in the sense that the sampled area might actually be too small to capture a long-range structure [*Kupfersburger and Deutsch*, 1999]. *Gelhar* [1986] and *Schafmeister and Pekdeger* [1993] show that correlation length increases with increasing scale. Thus, it is not an absolute property of the subsurface, but rather related to the size of the investigated area. Larger scales generally lead to an increased anisotropy ratio. Comparison of semivariograms should therefore be done with caution if similarly sized sampling and investigated areas are not used.

2.3.2 Geostatistics and the REV

Conditions required for geostatistical interpretation can be related to the relative elementary volume (REV) region. *Brown et al.* [2000] uses REV concepts to evaluate CT images. The REV is defined as the volume range, $V_{min} < V_0 < V_{max}$, where

$$\left. \frac{\partial Z(x, V_i)}{\partial V} \right|_{V_i = V_0} = 0, \tag{2.5}$$

where $Z(x, V_i)$ is the value of a hydrologic property measured on a volume V_i with a centroid at x, with a given domain $\Re$. *Bear and Bachmat* [1990] state that a volume size V_i that falls within the REV region can be treated as an ergodic stationary random

11

function in $\mathfrak{R}$. A stationary random function is said to be ergodic if any statistical characteristic of the function, taken over a sufficiently large volume in a single realization, is an unbiased and consistent estimate of the same characteristic over the entire data set. Functions that are ergodic for the mean and covariance are a subset of second-order stationary functions. By definition, a function is second-order stationary in $\mathfrak{R}$ if

$$E[Z(x,V_i)] = \text{constant} \qquad (2.6)$$

and

$$Cov[Z(x+h,V_i), Z(x,V_i)] = f(h) \qquad (2.7)$$

where $E[Z(x,V_i)]$ is the expected value of Z, and $Cov[Z(x+h),Z(x)]$ is the covariance of two points and is a function of the separation distance, h, only, for all $x \in \mathfrak{R}$. Then,

$$Var(Z(x)) = f(0) = \text{constant}. \qquad (2.8)$$

In a strict sense, (2.6) may be met by a linear function in space. However, as discussed by Brown et al. [2000], the limited scanning domains possible in CT preclude random, stationary volume averaging procedures. Thus, in the practical sense, any V_0 within the REV range will correspond to the volume required for ergodicity. This ensures a complete sample for the realization and makes it possible to make inferences from the semivariogram.

2.3.3 Scale Disparity

Gelhar [1986] states that "...to be meaningful and estimable, the correlation scale must be small compared with the scale of the problem. Essentially, a disparity in scale is

required." To quantify this scale disparity, Λ, I define the standard element volume (SEV) as the volume represented by the product of the correlation length in the x, y and z directions. By considering a REV_{min} with the same proportional dimensions as the SEV such that

$$REV_{min} = (a_x \Lambda)(a_y \Lambda)(a_z \Lambda) \qquad (2.9)$$

where a_x, a_y, and a_z are the correlation length in the x, y, and z direction, respectively. The scale disparity factor, Λ, is defined as

$$\Lambda = \sqrt[3]{\frac{REV_{min}}{SEV}}. \qquad (2.10)$$

Clausnitzer and Hopmans [1999] show that CT scanning of uniformly packed beads produces an REV_{min} with side length about two times the bead diameter, d. That implies a REV volume of $8d^3$. A semivariogram of that packing would produce a correlation length equal to d, and thus a Λ of 2. This is probably the minimum Λ that can be expected in porous media. Λ can be seen as the ratio of the characteristic length of the porous media's features, (in that case d) to the geometric mean length of the REV_{min}.

REV analysis is based on volume averaging concepts, which are traditionally scalar in nature and based upon a cubic volume. Conversely, semivariograms are fundamentally based on vector analysis. In highly anisotropic porous media, therefore, a REV dimensionally proportional to the SEV would give lower V_{min} values than V_{min} calculated using traditional volume averaging concepts.

2.4 Materials and Methods

2.4.1 Culebra Dolomite

Core samples were taken from the Culebra Dolomite Member of the Rustler Formation from drill holes near the Waste Isolation Pilot Plant (WIPP) site in southeastern New Mexico. The WIPP is a U.S. Department of Energy research and development facility designed to demonstrate safe disposal of transuranic radioactive waste resulting from the United States' defense programs. These samples were obtained because the Culebra Dolomite is the most transmissive confined unit above the waste repository and therefore is considered the most likely transport path by which radionuclides could be transported to the accessible environment over time spans of interest to regulatory agencies [*Lappin et al.*, 1989].

The Culebra is a finely crystalline, vuggy dolomite $(CaMg(CO_3)_2)$ which is often argillaceous and fractured. Kelley and Saulnier, 1990, report that the mean horizontal and vertical hydraulic conductivity for the Culebra Dolomite at this site is 6.2×10^{-9} m/s and 5.1×10^{-9} m/s, respectively. These estimates, however, exhibit a six order-of-magnitude variation in the vicinity of the sampling site. Density frequency analysis by *Hsieh et al.* [1998b] divides Culebra Dolomite into four components: solid dolomite, gypsum infills, silty dolomite and mixed regimes (dolomite, gypsum, and/or voids), and voids.

2.4.2 Sample Description

Three horizontal cores were imaged and analyzed. C2611A had a diameter of 147 mm and length of 450 mm. The barrel size for sample collection was selected to so that the core was representative with respect to heterogeneity [*Lucero et al.*, 1997]. C2611A had been used in numerous actinide transport experiments and was scanned saturated for operational reasons. The other two samples (C1AV and C2AV) were scanned dry and had diameters of 37 mm and lengths of 52 mm. These two samples were subsamples of a larger core that visually appeared to have different compositions.

2.4.3 Computerized Tomography

Each core was scanned in a three dimensional grid using the custom pencil-beam gamma ray CT scanner of *Brown et al.* [1993]. The x, y, and z axes on the horizontal cores are defined as shown in Figure 2.1. Cores C1AV and C2AV were imaged at 71 uniformly spaced slices, while C2611A was imaged at 150 positions, not all of which were used here. The z direction voxel spacing was double the spacing in the x and y direction for these cores. The images from C1AV and C2AV have also been used for analyses in *Hsieh et al.* [1998b] and *Brown et al.* [2000].

2.4.4   Semivariogram Analysis

Deutsch and Journel [1998] provides a package of FORTRAN 77 geostatistical programs that was used here. The GAM code was recompiled using Visual Fortran 6.6 to increase the maximum number of data points that could be analyzed. Semivariogram

15

generation was then completed for gamma ray attenuation of core samples using the recompiled GAM program.

2.5 Results and Discussion

2.5.1   CT Measurements

CT scanning parameters are listed in Table 2.1. For the analysis here, core C2611a has been separated into two different sections to illustrate differences in slices 30 to 90 and 90 to 150 of the same core. Core samples are cylinders, but for ease of computation, semivariograms are completed on data in rectangular prisms in the center of the cylinders. The bulk density in Table 2.1 is calculated by

$$\rho_b(x,y,z) = C\mu(x,y,z) \tag{2.11}$$

where C is the calibration factor and $\mu(x,y,z)$ is the point attenuation value [*Luo and Wells*, 1992]. *Brown et al.* [1999] reported a calibration coefficient of 147.8 g mL^{-1} mm^{-1} for core C2611A. The calibration coefficient for cores C1AV and C2AV is 127 g mL^{-1} mm^{-1}.

Cores C1av and C2av are the smaller resolution cores (0.10 mm^3 voxel spacing) and the two samples from core C2611a are larger resolution (6.8 mm^3 voxel spacing). The two samples from core C2611a exhibit larger bulk density means due to their being scanned wet, but smaller standard deviations than cores C1av and C2av. Core C1av has the smallest mean dry bulk density and highest standard deviation. As a demonstration of the variability in the cores, the mean bulk density along the z axis of each sample is shown in Figure 2.2.

Table 2.2 uses four structural components and the statistical segregation threshold of *Hsieh et al.* [1998a] to differentiate components of the bulk density frequency distribution in each of the cores. The two samples from core C2611A indicate predominantly silty and solid dolomite. Core C1AV is similar to samples C2611A, except there is a detectable percentage of solid gypsum. Core C2AV is different from the other cores, being mostly solid dolomite with a smaller percentage of silty dolomite areas. Core C2AV also has the largest percentage of solid gypsum areas.

2.5.2   Geostatistics

A reconstruction of a x-y slice from each of the cores and standardized semivariograms for each direction are shown in Figures 3 and 4, and geostatistical properties of each are listed in Table 2.3. For the two larger resolution samples (C2611a) the small-scale correlation lengths ranged from 3.3 to 17.3 mm and the large-scale correlation lengths ranged from 24.0 to 52.0 mm. For the two smaller resolution samples (C1AV and C2AV) the small-scale correlation lengths ranged from 1.6 to 3.1 mm and the large-scale correlation lengths ranged from 13.7 to 28.7 mm. Two semivariograms were best fit with exponential semivariogram models. The rest of the semivariograms were best fit with a combination model, indicating the presence of nested sills. Best fit models were chosen by maximizing the coefficient of determination, r^2, between predicted and measured values. The best fit semivariogram models all show a r^2 value of 0.98 or greater from $h=0$ to the lag indicated, denoting good fit for all the semivariograms. The best fit model was not determined for the entire semivariogram because as the semivariogram lag approached the far end of the sample, the semivariogram often begins

17

to drift up of down. This is a result of a smaller number of comparisons in the semivariogram mean squared difference near the edges of the sample and is not necessarily considered representative of the geostatistical properties of the media. The best fit model was also not determined for the entire semivariogram if a hole effect was exhibited. None of the semivariograms for the cores shown in Figure 2.3 and 2.4 and described in Table 2.3 have been modeled with a nugget. This indicates that at these resolutions, the spatial variability in the cores is fully represented in the semivariogram. Large scale correlation lengths in the x and z direction of core C2AV have values approached the sample length. This may be indicative of a large-scale trend in the mean [*Gelhar*, 1986]. The C2AV curve in Figure 2.2 may indicate a slight trend in the mean along the z direction. Core C2611a (30-90) exhibits a hole effect in the y direction. This cannot be modeled by nested sills, but is estimated from the semivariogram and listed in Table 2.3.

The three-dimensional anisotropy ratios and anisotropy in the large and small scale are characterized in Table 2.4. The anisotropy characterizations are based on the definitions of *Kupfersberger and Deutsch* [1999] for well-log analysis. Anisotropy ratios have been defined to describe the geometric anisotropy of a sample (ie. sill values are equal). Therefore, the ratios listed in Table 2.4 for cores where geometric anisotropy does not exist, are not true anisotropy ratios, but simply a ratio of the range values in the two directions of interest. Because nested ranges have been modeled on these cores, a small scale and large scale anisotropy ratio has been calculated corresponding to the small scale and large scale structure within the core.

Three anisotropy characterizations are provided for each core to demonstrate if the samples are isotropic or geometrically anisotropic in the x-z plane (two horizontal directions) as expected. Core C2611a (90-150) and C2AV exhibit these anisotropies, while C2611a (30-90) and C1AV exhibit zonal anisotropy in the x-z plane. The zonal anisotropy is not further defined in this instance because neither direction is vertical. Core C2611a (30-90) was zonal because of different sill values between the x- and z-directions. The x direction of core C1AV was modeled by a basic model and did not exhibit a large-scale structure. Nearly all of the x-y and z-y (horizontal-vertical) anisotropy characterizations are zonal due to strata, with the exception of C1AV in the x-y direction, which is zonal due to areal trends, and core C2AV, which exhibits isotropy and geometric anisotropy.

2.5.3 Geostatistical Comparisons of Adjacent Cores

Two adjacent sections of core C2611a have been analyzed for geostatistical properties with a resolution of 1.5 mm per voxel. For these cores, range values differ in all three primary directions. A nested sill is indicated in the y direction of core c2611a, slice 90-150, while a hole effect is shown in the slice 30-90 core. The hole effect phenomenon may be representative of strata or horizontal bedding planes in the core and this difference is likely a result of a change in stratification between the two core subsections. These results indicate that different geostatistical properties can be easily identified for adjacent cores at this scanning resolution.

Cores C1AV and C2AV were extracted from visually different sections of a larger core and were located within 70 mm of each other. The main structural difference

between these two samples is the absence of a nested sill in the x-direction on core C1AV. A higher percentage of silty dolomite in C1AV (Table 2.2) could be partially responsible for this result. If there is more silty dolomite (mixed voxels), there conversely will be smaller continuous regions of solid dolomite. This would account for the absence of a larger-scale range in this core, and may also explain the anisotropy differences exhibited by these two cores. As with the larger resolution, these results indicate that at this resolution samples obtained next to each other can exhibit clearly different geostatistical properties.

2.5.4   Geostatistical Comparisons at Different Resolutions

In general, when comparing the geostatistical properties of smaller resolution cores (C1AV and C2AV) to the larger resolution cores (C2611a), the two resolutions show similar range values in the vertical direction when comparing the nested ranges in the large resolution to the ranges indicated in the smaller resolution. In the horizontal, the ranges are larger when comparing equal scales of the large resolution to the small resolution. However, the smaller resolution cores have larger ranges than the nested ranges in the larger resolution. This could be a result of changing horizontal geologic structure of the Culebra Dolomite, comparison of semivariograms from differing sample sizes and resolutions, and/or axes of anisotropy that may not be exactly horizontal. The larger range values associated with the larger scales are not detected in the smaller resolution cores because of the smaller sample size used in the smaller resolution analysis. Nested range values on the smaller scale cores are not detectable in the larger resolution because the values approach the resolution limit in the CT images.

20

Comparison of the small-scale and large-scale anisotropy ratios on these cores shows the ratio increasing at increasing scales. The results reinforce the results of *Gelhar* [1986] and *Schafmeister and Pekdeger* [1993] that increasing scales show increasing ranges and anisotropy ratios.

2.5.5 Scale Disparity Quantification

Table 2.3 lists the small-scale correlation lengths for cores C1AV and C2AV ranging from 1.6 mm to 3.1 mm. SEV volumes were thus 9.4 mm^3 and 8.5 mm^3 for these two cores, respectively. Brown et al., 2000, reports a REV_{min} of 1000 mm^3 and 2200 mm^3 for cores C1AV and C2AV, respectively. Therefore, Λ is 4.7 and 6.4 for cores C1AV and C2AV, respectively. Since the REV_{min} corresponds to the volume required for ergodicity, the sample length should be at least five or six times larger than the correlation lengths to detect the spatial structure of the Culebra Dolomite at this small scale. The analyzed sample volume is also larger than the minimum REV for both samples.

The SEV can be viewed as the characteristic size of the density features (pore, solid dolomite, or silty regions) within the core. Conversely, the REV_{min} is the minimum volume required to obtain a statistically significant average of all these dominant features. The Λ measured for the cores thus imply that Λ^3 density features are within V_{min}, or 104 and 262 features for C1AV and C2AV, respectively.

21

2.5.6   Modifiable Areal Unit Problem and the Inherent Range

Comparison of CT images at different resolutions also lends itself to an analysis of the modifiable areal unit problem. *Cressie* [1991] discusses the situation where averaging, or aggregating, of regionalized variables leads to larger sample correlation lengths. *Openshaw and Taylor* [1979] show that by averaging fixed observations on small areas in various ways, a large range of correlation coefficients can be produced.

Gamma ray CT is similar to the modifiable areal unit problem, as the image attenuation value is simply the average attenuation through the area encompassed by the voxel. Therefore, semivariograms were calculated for CT images of core C2AV with voxels modified by increasing voxel lengths by 2x, 3x, 4x, and 5x, and averaging the attenuation values within these new volumes. This increased the voxel volume by a factor of 8, 27, 64, and 125, respectively. The results of these exercises are shown in Figure 2.5 and Table 2.5.

As the voxel size increases, the correlation length increases in all directions. All of these semivariograms become more linear as voxel size increases. Larger resolution semivariograms would likely be best fit with a nugget. Figure 2.6 summarizes the scaling effect caused by increasing voxel volume. A linear trend is indicated between the correlation length and voxel volume. I propose that the intercept of the trend line with the y-axis (range) in this graph could be said to represent the *inherent range*, or the range that is unaffected by the scaling effect of increased scanning resolution. The slope of the trend line indicates the scaling effect caused by increasing voxel volume. Figure 2.6 also demonstrates that multiple inherent ranges may exist in a particular direction.

2.6 Summary and Conclusions

Geostatistics has been used to determine the spatial structure and anisotropy of Culebra Dolomite at very small scales. This analysis was possible because of characterization of the rock from CT images utilizing gamma ray attenuation. Horizontal range values for the four samples ranges from 1.6 to 17.3 mm for small-scale structure and from 13.7 to 52.0 mm for large-scale structure. The smaller-resolution analysis did not detect the largest-scale structure because the total sample size was smaller than the range values associated with this structure. Likewise, the larger-resolution analysis did not detect the smallest-scale structure because the correlation lengths approached the resolution limit.

Results suggest that cores sampled adjacent to each other can show different geostatistical properties at these resolutions. In general, correlation lengths were similar at different resolutions in the vertical direction. These vertical direction results could be a result of fairly uniform stratification in the Culebra Dolomite. However, since four different samples are being compared, this is not conclusive enough evidence to discount a scaling effect in the vertical direction. Correlation lengths in the horizontal, however, changed with resolution and sample size. Possible reasons for these changes include comparing geostatistical properties for differ size samples with different resolutions, changing horizontal geologic structure of the Culebra Dolomite, and/or axes of anisotropy that may not be exactly horizontal.

Relative elementary volume (REV) concepts have been shown to be analogous to the ergodic property of second-order stationarity. Using the REV and standard element volume (SEV), a scale disparity factor has been quantified to determine sample size

necessary to determine a meaningful and estimable semivariogram with specified correlation lengths.

The modifiable areal unit problem has also been addressed with this data. By changing resolution, one can estimate a range of correlation lengths. A linear relationship between voxel volume and correlation length has been determined for all directions. The y-intercept of the linear relationship has been defined as the inherent range. This is the range unaffected by the scaling affect of increasing resolution. Multiple inherent ranges can exist in a given direction. Therefore, based on these results, sample size and resolution selection for future core analysis should be a function of the scale of interest for the intended data use.

2.7 References

Bear, J., and Y. Bachmat, *Introduction to Modeling of Transport Phenomena in Porous Media*, Kluwer, Norwell, MA, pp. 16-28, 1990.

Brown, G. O., H. T. Hsieh, and D. A. Lucero, *Analysis of Culebra Dolomite B Core by Gamma Ray CT*, unpublished research report submitted to Sandia National Laboratories, Albuquerque, NM, under contract AT-8739, Oklahoma State University, 1999.

Brown, G. O., H. T. Hsieh, and D. A. Lucero, Evaluation of laboratory dolomite core size using representative elementary volume concepts, *Water Resourc. Res.*, 36(5), 1199-1207, 2000.

Brown, G. O., M. L. Stone, and J. E. Gazin, Accuracy of Gamma Ray Computerized Tomography in Porous Media, *Water Resourc. Res.*, 29(2), 479-486, 1993.

Chiles, J.-P., and P. Delfiner, *Geostatistics, Modeling Spatial Uncertainty*, John Wiley and Sons, Inc., New York, p. 17, 1999.

Clausnitzer, V and J. W. Hopmans, Determination of phase-volume fractions from tomographic measurements in two-phase systems, *Water Resources Res.*, 22(6), 577-584, 1999.

Coskun, S. B., and N. C. Wardlaw, Estimation of Permeability from Image Analysis of Reservoir Sandstones, *J. of Pet. Sci. Eng.*, 10, 1-16, 1993.

Cressie, N. A. C., *Statistics for Spatial Data*, John Wiley and Sons, NY, 1991.

Crestana, S., S. Mascarenhas, and R. S. Pozzi-Mucelli, Static and dynamic three-dimensional studies of water in soil using computed tomography scanning, *Soil Science*, 140, 326-332, 1985.

Deutsch, C.V., and A.G. Journel, *GSLIB: Geostatistical Software Library and User's Guide*, Oxford University Press, 2nd edition, New York, 1998.

Flühler, H., N. Ursino, M. Bundt, U. Zimmerman, and C. Stamm, The Preferential Flow Syndrome -- A Buzzword or a Scientific Problem, in *Preferential Flow, Water Movement and Chemical Transport in the Environment*, Proceedings of the 2nd International Symposium, American Society of Agricultural Engineers, Honolulu, HA, pp. 21-24, 2001.

Gelhar, L.W., Stochastic subsurface hydrology from theory to applications, Water Resour. Res., 22(9), 1358-1458, 1986.

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

Grevers, M. C. J., and E. de Jong, Evaluation of Soil-Pore Continuity Using Geostatistical Analysis on Macroporosity in Serial Sections Obtained by Computed Tomography Scanning, in *Tomography of Soil-Water-Root Processes*, SSSA Special Publication no. 36, Soil Science Society of America, Madison, WI, pp. 73-86, 1994.

Hsieh, H. -T., G. O. Brown, and M. L. Stone, Quantification of Porous Media Using Computerized Tomography and a Statistical Segregation Threshold, *Trans. ASAE*, 41(6), 1697-1706, 1998a.

Hsieh, H. -T., G. O. Brown, M. L. Stone, and D. A. Lucero, Measurement of porous media component content and heterogeneity using gamma ray tomography, *Water Resourc. Res.*, 34(3), 365-372, 1998b.

Ioannidis, M. A., I. Chatzis, and M. J. Kwiecien, Computer Enhanced Core Analysis for Petrophysical Properties, *J. Can. Pet. Tech.*, 38(3), 18-24, 1999.

Isaaks, E.H., and R.M. Srivastava, *An Introduction to Applied Geostatistics*, Oxford University Press, New York, 1989.

Kelley, V.A., and G. J. Saulnier, Jr., Core Analyses for Selected Samples from the Culebra Dolomite at the Waste Isolation Pilot Plant Site, *Contractor Report SAND90-7011*, Sandia National Laboratories, Albuquerque, NM, 1990.

Kupfersberger, H., and C.V. Deutsch, Methodology for Integrating Analog Geologic Data in 3-D Variogram Modeling, *AAPG Bulletin*, 83(8), 1262-1278, 1999.

Lappin, A. R., R. L. Hunter, D. P. Garber, and P. B. Davies, (Editors), Systems Analysis, Long-Term Radionuclide Transport, and Dose Assessments, Waste Isolation Pilot Plant (WIPP), Southeastern New Mexico, *SAND89-0462*, Sandia National Laboratories, Albuquerque, NM, 1989.

Lucero, D. A., G. O. Brown, and C. E. Heath, Laboratory Column Experiments for Radionuclide Adsorption Studies of the Culebra Dolomite Member of the Rustler Formation, *SAND97-1763*, Sandia National Laboratories, Albuquerque, NM, 1997.

Luo, X., and L. G. Wells, Evaluation of gamma ray attenuation for measuring soil bulk density, 1, Laboratory investigation, *Trans. ASAE.*, 35(1), 17-26, 1992.

Openshaw, S., and P. J. Taylor, A Million or So Correlation Coefficients: Three Experiments on the Modifiable Areal Unit Problem, in *Statistical Applications in the Spatial Sciences*, N. Wrigley, ed. Pion, London, pp. 127-144, 1979.

Petrovic, A. M., J. E. Siebert, and P. E. Rieke, Soil bulk density analysis in three dimensions by computed tomography scanning, *Soil Sci. Soc. Am. J.*, 46(3), 445-450, 1982.

Phogat, V. K., L. A. G. Aylmore, and R. D. Schuller, Simultaneous measurement of the spatial distribution of soil water content and dry bulk density, *Soil Sci. Soc. Am. J.*, 55, 908-915, 1991.

Schafmeister, M. -Th. and A. Pekdeger, Spatial structure of hydraulic conductivity in various porous media -- problems and experiences, in *Geostatistics Troia 1992,* A. Soares (ed), Kluwer, Quantitative Geology and Geostatistics Series, 1, pp. 733-744, 1993.

Schafmeister-Spierling, M. -Th. and A. Pekdeger, Influence of Spatial Variability of Aquifer Properties on Groundwater Flow and Dispersion, in *Contaminant Transport in Groundwater*, Dobus and Kinzelbach (eds.), Balkema, Rotterdam, The Netherlands, pp. 215-220, 1989.

Solie, J.B., W.R. Raun, and M.L. Stone, Submeter Spatial Variability of Selected Soil and Bermudagrass Production Variables, *Soil Sci. Soc. Am. J.*, 63(6), 1724-1733, 1999.

Tollner, E. W., and B. P. Verma, X-ray CT of quantifying water content at points within a soil body, *Trans. ASAE*, 32, 901-905, 1989.

Woodbury, A. D., and E. A. Sudicky, The Geostatistical Characteristics of the Borden Aquifer, *Water Resour. Res.*, 27(4), 533-546, 1991.

Table 2.1. Summary of gamma ray attenuation statistical properties of Culebra Dolomite samples.

Sample Name	Sample Dimensions, X x Y x Z, (Voxels)	Voxel Length (mm) [1]	Attenuation (mm^{-1})		Bulk Density (g/ml)	
			Mean	St. Dev.	Mean	St. Dev.
C2611a, slice 30-90	61 x 61 x 81 (301,401)	1.5	0.0174	0.0013	2.57	0.19
C2611a, slice 90-150	61 x 61 x 81 (301,401)	1.5	0.0171	0.0013	2.53	0.19
C1av	71 x 71 x 46 (231,886)	0.37	0.0183	0.0030	2.32	0.38
C2av	71 x 71 x 46 (231,886)	0.37	0.0192	0.0024	2.44	0.30

[1] z direction voxel spacing is double the x and y direction voxel spacing; x and y direction is listed.

Table 2.2. Bulk density frequency distributions of Culebra Dolomite samples.

Sample Name	% Voids	% Silty Dolomite	% Solid Gypsum	% Solid Dolomite
C2611a, slice 30-90	<1	38	<1	62
C2611a, slice 90-150	<1	46	<1	54
C1av	<1	30	4	66
C2av	<1	9	9	82

Table 2.3. Properties of standardized semivariograms for the culebra dolomite samples. [1]

Sample Name and Direction		Semivariogram Model	Range (mm)	Sill, w_i	r^2, mm [3]	
C2611a, 30-90	x (nested) x	Combination of Two Spherical	6.6 52.0	0.49 0.47	0.992	to 90
	y y (hole)	Exponential Visually estimated	17.3 60	1.02 0.84	0.997	to 18
	z (nested) z	Combination of Two Exponential	4.8 52.0	0.41 0.44	0.997	to 72
C2611a, 90-150	x (nested) x	Combination of Spherical and Exponential	3.5 35.0	0.37 0.52	0.999	to 63
	y (nested) y	Combination of Spherical and Exponential	24.0 7.4	0.40 0.72	0.995	to 35
	z (nested) z	Combination of Spherical and Exponential	3.3 37.4	0.43 0.51	0.997	to 66
C1av	x	Exponential	3.1	0.86	0.981	to 16.3
	y (nested) y	Combination of Spherical and Exponential	1.9 13.7	0.58 0.48	0.997	to 14.1
	z (nested) z	Combination of Two Spherical	1.6 20.1	0.12 0.81	0.983	to 15.5
C2av	x (nested) x	Combination of Spherical and Exponential	2.3 25.9 [2]	0.83 0.22	0.991	to 13.0
	y (nested) y	Combination of Spherical and Exponential	9.1 2.3	0.17 0.85	0.992	to 15.9
	z (nested) z	Combination of Spherical and Exponential	28.7 [2] 1.6	0.16 0.83	0.993	to 14.8

[1] None of the cores were modeled with a nugget

[2] Range approaching sample width

[3] In this column, the coefficient of determination and the largest lag value used for calculation of the coefficient of determination is reported.

Table 2.4. Anisotropy of Culebra Dolomite samples.

Sample Name	Direction	Anisotropy Ratio		Anisotropy	
		Small Scale	Large Scale	Small Scale	Large Scale
C2611a, slice 30-90	x-y:	0.38	0.87 [1]	Zonal due to Strata	na [2]
	z-y:	0.28	0.87 [1]	Zonal due to Strata	na
	x-z:	1.38	1.00	Zonal [3]	Zonal [3]
C2611a, slice 90-150	x-y:	0.47	1.46	Zonal due to Strata	Zonal due to Strata
	z-y:	0.45	1.56	Zonal due to Strata	Zonal due to Strata
	x-z:	1.06	0.94	Isotropic	Isotropic
C1av	x-y:	1.63	na	Zonal Areal Trends	na
	z-y:	0.84	1.47	Zonal due to Strata	Zonal due to Strata
	x-z:	1.94	na	Zonal [3]	na
C2av	x-y:	1.00	2.84	Isotropic	Geometric
	z-y:	0.65	3.15	Geometric	Geometric
	x-z:	1.53	0.90 [4]	Geometric	Isotropic [4]

[1] The range of the hole effect is used to calculate the large scale x-y and z-y anisotropy ratio for this sample.

[2] na = not available

[3] Further characterization of zonal anisotropy is not included in this characterization because neither direction is vertical.

[4] Range approaching sample width

Table 2.5. Properties of standardized semivariograms for the culebra dolomite core C2AV using different aggregations based upon voxel volume to address the modifiable area unit problem. The initial sample spacing was 0.37 mm in the x and z directions and 0.74 mm in the z direction.[1]

Resolution, Volume, Direction		Semivariogram Model	Range (mm)	Sill, w_i	r^2, mm [3]	
1x (0.101 mm³)	x (nested) x	Combination of Spherical and Exponential	2.3 25.9 [2]	0.83 0.22	0.991	to 13.0
	y (nested) y	Combination of Spherical and Exponential	9.1 2.3	0.17 0.85	0.992	to 15.9
	z (nested) z	Combination of Spherical and Exponential	28.7 [2] 1.6	0.16 0.83	0.993	to 14.8
2x (0.810 mm³)	x (nested) x	Combination of Spherical and Exponential	2.3 21.2 [2]	0.46 0.58	0.994	to 13.3
	y (nested) y	Combination of Two Spherical	2.3 9.1	0.64 0.40	0.996	to 14.1
	z (nested) z	Combination of Two Spherical	2.4 25.9	0.69 0.32	0.992	to 14.8
3x (2.74 mm³)	x (nested) x	Combination of Spherical and Exponential	3.1 25.9 [2]	0.47 0.54	0.964	to 15.5
	y (nested) y	Combination of Spherical and Exponential	3.1 9.7	0.60 0.47	0.998	to 14.4
	z (nested) z	Combination of Two Spherical	3.1 25.9 [2]	0.53 0.41	0.997	to 11.1
4x (6.48 mm³)	x (nested) x	Combination of Two Spherical	3.9 22.6 [2]	0.46 0.51	0.999	to 14.8
	y (nested) y	Combination of Spherical and Exponential	4.6 10.5	0.29 0.83	0.997	to 13.3
	z (nested) z	Combination of Two Spherical	3.7 25.9 [2]	0.53 0.41	0.978	to 14.8
5x (12.66 mm³)	x (nested) x	Combination of Two Spherical	4.3 20.2 [2]	0.36 0.56	0.999	to 13.0
	y (nested) y	Combination of Two Spherical	5.6 12.6	0.48 0.63	0.998	to 14.8
	z [4]	--	--	--	--	--

[1] None of the cores were modeled with a nugget

[2] Range approaching sample width

[3] In this column, the coefficient of determination and the largest lag value used for calculation of the coefficient of determination is reported.

[4] Not enough points in the semivariogram to fit a meaningful model at this lag distance.

32
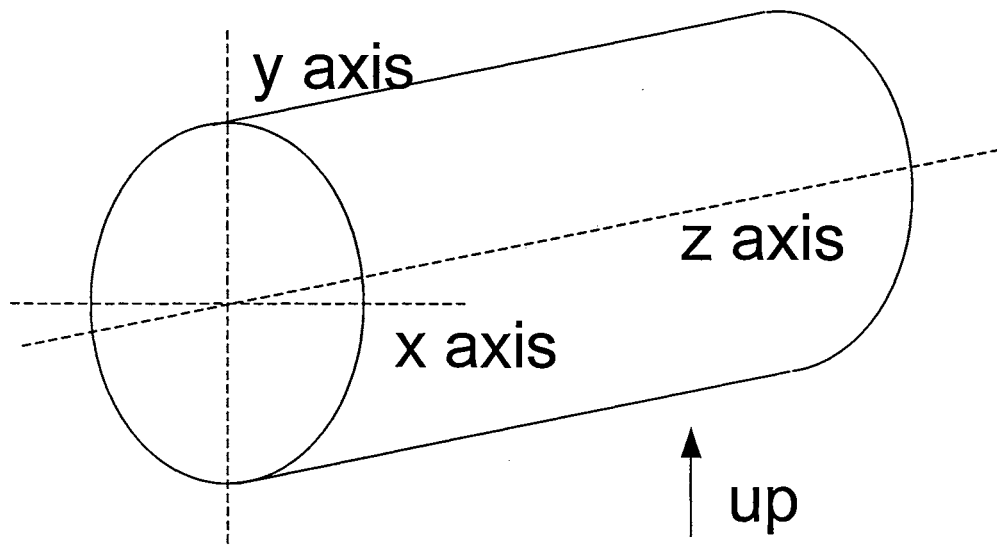
Figure 2.1. Orientation of Culebra Dolomite cores.
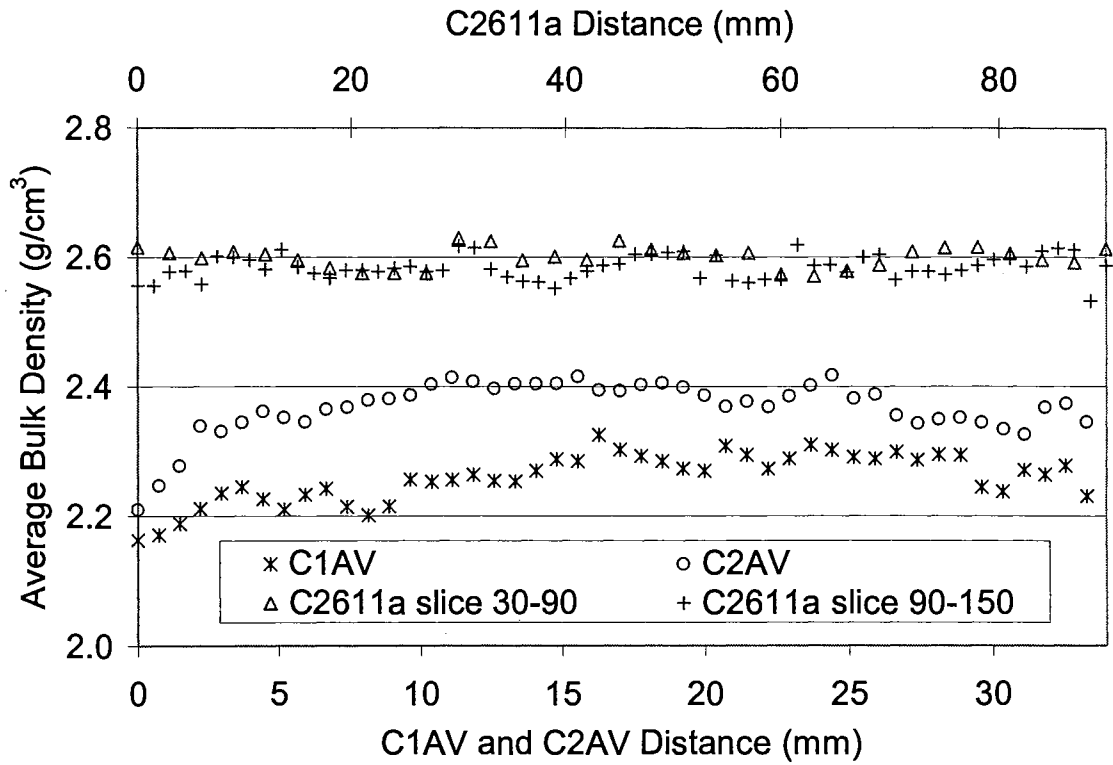
Figure 2.2. Individual slice average bulk density along each core.

c2611a, 30-90
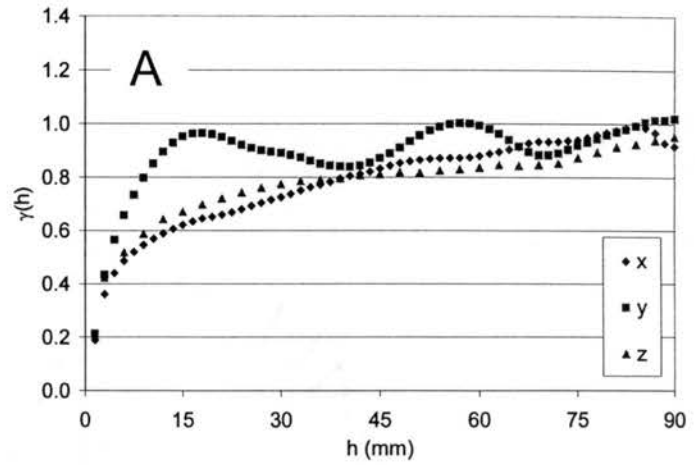
c2611a, 90-150
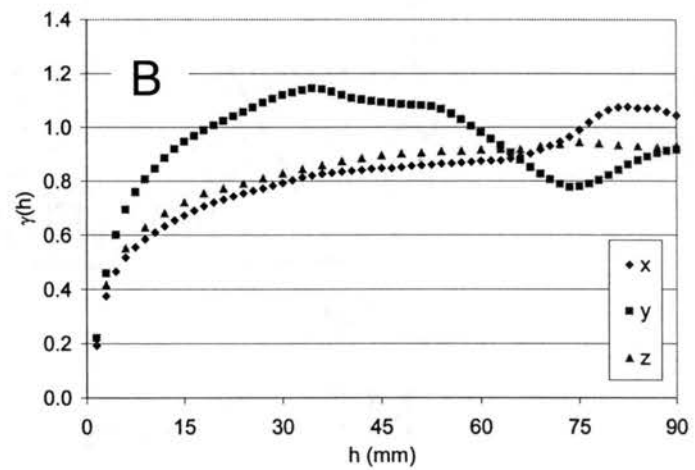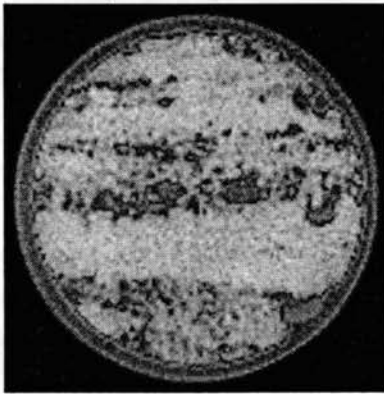
Figure 2.3. Computer tomography images and semivariograms of (A) core C2611a (slice 30-90) and (B) core c2611a (slice 90-150).

vpx26-c1av (slice 35)
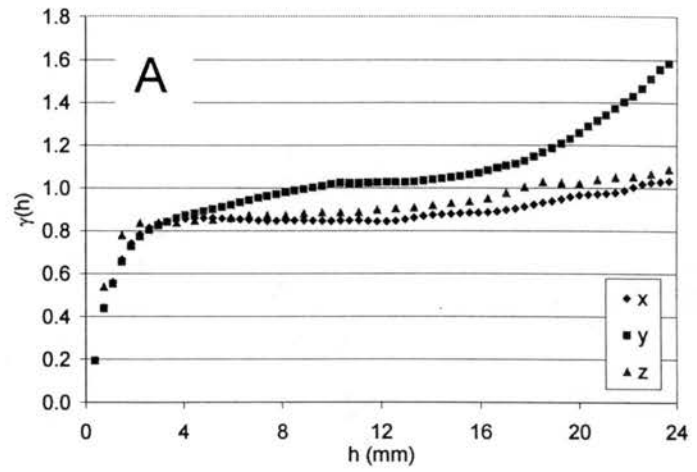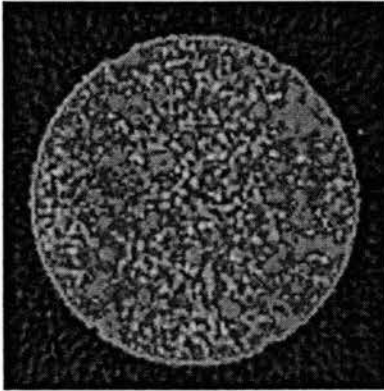
vpx26-c2av
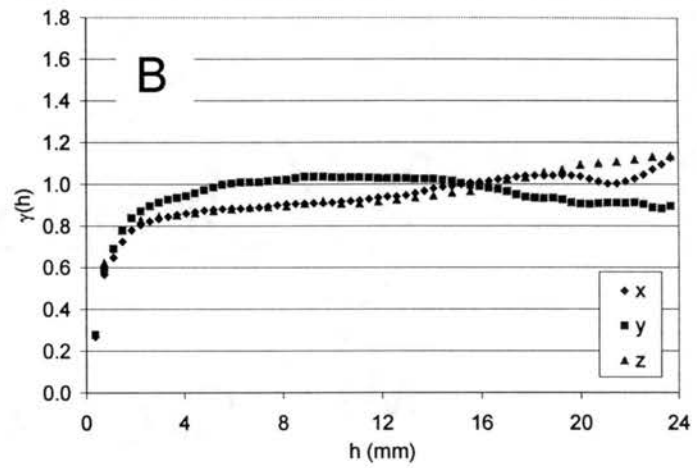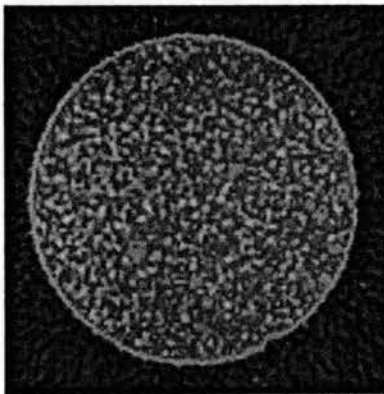
Figure 2.4. Computer tomography images and semivariograms of (A) core C1AV and
(B) core C2AV.
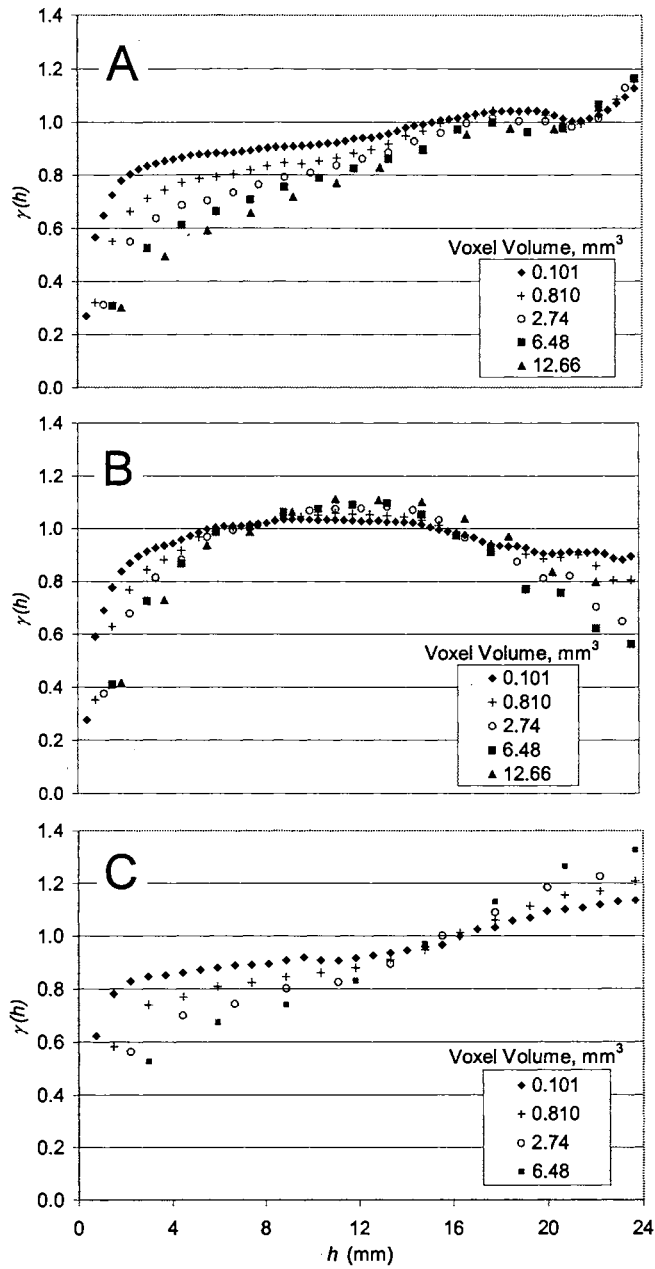
Figure 2.5. Semivariograms of computer tomography images for core C2AV with different voxel volumes in (A) the x direction, (B) the y direction, and (C) the z direction.

37
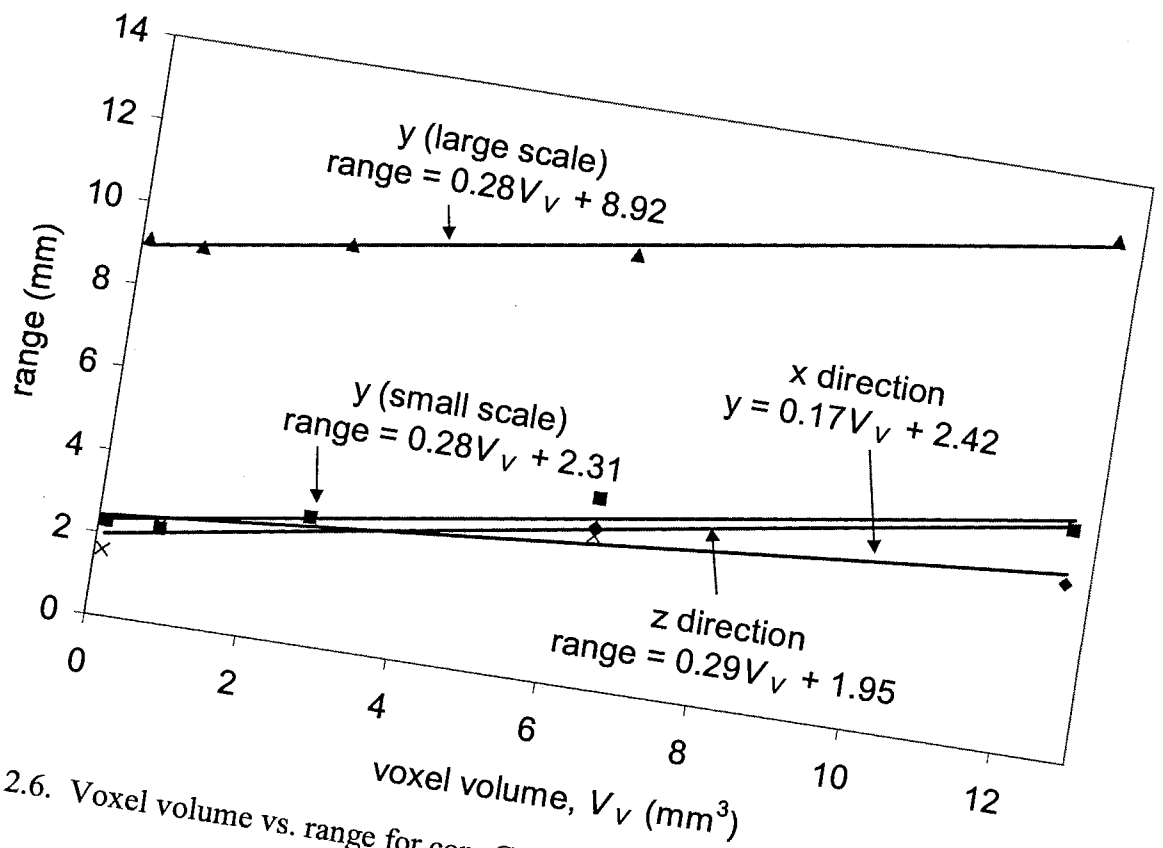
Figure 2.6. Voxel volume vs. range for core C2AV.

CHAPTER THREE

CHARACTERIZATION OF POROUS MEDIA CT IMAGES USING SEMIVARIOGRAM FAMILIES

3.1 Abstract

A semivariogram family is the collection of directional slice semivariograms from a three-dimensional sample. Analysis using semivariogram families of computerized tomography (CT) images of Culebra dolomite provides improved methods for determining the spatial correlation structure of the bulk density of porous media. The largest reliable lag in a semivariogram is quantified by analyzing the variance of the semivariogram values within the semivariogram family at each lag. This improves on previous methods for determining the largest reliable lag that rely on rules of thumb. In addition, the distribution of the semivariogram values within a semivariogram family at each lag is analyzed to automatically determine the correlation length (range) values for the porous media. Such a procedure provides range values that are representative of actual structure, as opposed to values determined from best-fitting empirical semivariogram models that may or may not be representative of the actual morphology.

3.2 Introduction

The use of geostatistics is an important tool for characterizing spatial properties of porous media. Applications have included investigations of the spatial variability of subsurface hydraulic conductivity [*Schafmeister-Speirling and Pekdeger*, 1989; *Woodbury and Sudicky*, 1991], soil geochemical properties [*Miesch*, 1975; *Solie et al.*, 1999], rock fractures [*Tavchandjian et al.*, 1993], soil temperature [*Yates et al.*, 1988], permeability [*Goggin et al.*, 1988], and soil water content [*Greenholtz et al.*, 1988; *Bárdossy and Lehmann*, 1998]. Geostatistical image analysis also allows for micro-scale non-destructive characterization of the spatial variability of porous media bulk density, which can in turn be used for a variety of predictive transport models. *Grevers and de Jong* [1994] used x-ray computer tomography (CT) images to analyze soil macroporosity and ground-penetrating radar images were analyzed by *Tercier et al.* [2000] to compare the correlation structure of different depositional environments. *Ioannidis et al.* [1999] used geostatistical analysis of scanning electron micrographs (SEM) to evaluate various transport and capillary properties of rock samples and gamma ray CT has been used to quantify the spatial variability of soil properties in soil columns [*Rasiah and Aylmore*, 1998] and rock cores [*Vogel and Brown*, 2001].

All of these applications depend on reliable determination of the spatial structure of the variable of interest. *Gelhar* [1986] discusses "...the key role that the spatial correlation structure of the input processes plays in the behavior of all the stochastic solutions..." of subsurface hydrologic models. In addition, Gelhar states that "...there is a need to find better methods to determine the input spatial correlation structure of hydraulic parameters. This should include the use of prior information on geologic and

geomorphic conditions as well as improved methods of statistical analysis to estimate these parameters." This need is reiterated by *Gelhar* [1993] and *Calvete* [1997].

Two questions addressed here concerning the reliability of spatial characterizations are: What is the maximum reliable lag on a semivariogram that should be used, and which correlation lengths determined from best fit empirical models are representative of actual structural features? In this paper the large data set provided by gamma ray CT analysis of Culebra Dolomite cores is utilized to introduce semivariogram families and show how they can be used (1) to identify the maximum lag for reliable semivariogram estimates, and (2) to automatically detect correlation lengths (ranges) representative of actual structure for semivariograms derived from computer tomography images of porous media. Combined, these two methods greatly improve the statistical analysis for estimating the spatial structure of porous media properties.

3.3 Geostatistical Theory

3.3.1 Semivariograms

The semivariogram is one of the most important tools in geostatistics. For an intrinsically stationary random function $Z(x)$, the semivariogram is the plot of the semivariogram value, $\gamma(h)$, vs. lag distance, h. The semivariogram value for each lag distance is represented by

$$\gamma(h) = \frac{1}{2} \text{Var}[Z(x) - Z(x+h)],$$ (3.1)

for each lag distance. Because $\gamma(h)$ cannot be calculated directly, an unbiased estimator for the semivariogram of a random function is used and is given by

$$\hat{\gamma}(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [Z(x_i + h) - Z(x_i)]^2 \qquad (3.2)$$

where $N(h)$ is the number of pairs of variables at h distance apart [*Olea*, 1999].

A sill and range characterize the semivariogram. The sill is a plateau in the semivariogram values that corresponds to the variance of the sampled data. If the semivariogram is standardized, the $\gamma(h)$ value is divided by the variance and the sill value will be equal to one. The range, or correlation length, is the distance at which the semivariogram reaches a plateau. One possible semivariogram characteristic is the hole effect, which is described as a semivariogram that peaks and then dips, suggesting that at greater distances samples are more related. The hole effect may be seen as the rough spacing between adjacent lenses or bedding planes. Another characteristic similar to the hole effect is pseudocycling. This is the apparent periodic cycling or oscillation of the magnitude of the semivariogram over distance and is common in minerals. However, unless they can be attributed to a physical phenomenon, these changes should be considered random even though they appear periodic [Solie, 1999].

Semivariograms are fit by a number of analytical relationships to ensure that the semivariograms model is positive definite. One of the most common semivariogram fittings used in hydrology is the exponential model [*Woodbury and Sudicky*, 1991]. The model is given by

$$\gamma(h) = 1 - e^{\left(-\frac{3h}{a}\right)}. \qquad (3.3)$$

The exponential model reaches its sill asymptotically, with the practical range a defined as that distance at which the variogram value is 95% of the sill [*Isaaks and Srivastava*,

1989]. *Agterberg* [1970] has shown that a continuous random variable in three-dimensional space has an exponential autocorrelation function if it is subject to a property analogous to the Markov property in time series analysis.

Multiple, or nested, sills in the semivariogram may be associated with physical phenomena occurring at different scales. Detection of ranges nested within larger ranges requires sample spacing shorter than the minimum detected range. As with pseudocycling, nested sills are only considered when they can be associated with physical phenomena [*Solie et al.*, 1999]. Nested sills can be modeled by using a linear combination of positive definite semivariogram models with positive coefficients. This property results in the family of models

$$\gamma(h) = \sum_{i=1}^{n} |w_i| \gamma_i(h) \tag{3.4}$$

which are positive definite as long as the n individual models are all positive definite. A weighting function, w_i, is defined for each individual model subject to $\sum_{i=1}^{n} w_i = 1$ [*Isaaks and Srivastava*, 1989].

3.3.2 Semivariogram Families

Chilès and Delfiner [1999] describe the semivariogram cloud as introduced by *Gandin* [1963] and used by *Chauvet* [1982] for meteorological fields. The semivariogram cloud is a plot of all sample pairs (α, β) showing $h_{\alpha\beta} = |x_\beta - x_\alpha|$ on the x-axis and the half squared increment, $\frac{1}{2}(Z(x_\beta) - Z(x_\alpha))^2$, on the y-axis. An example of a semivariogram

43

cloud as demonstrated by *Vogel and Haan* [2001] is shown in Figure 3.1a. I introduce the semivariogram family as an extension of the semivariogram cloud for use with three-dimensional CT data. A semivariogram family is a plot of the semivariograms for each CT image slice. It is similar to the semivariogram cloud, except that the semivariogram curve for each slice is plotted. Each point that makes up a semivariogram family "member" therefore represents the mean of all the half squared increments at that lag distance in that slice, as opposed to the point difference represented by the semivariogram cloud. A hypothetical semivariogram family is shown in Figure 3.1b.

3.3.3 Semivariogram Estimate Reliability

Besides rules of thumb, there are currently no tests to determine the reliability of semivariogram estimations at large lags [*Olea*, 1999]. Because the accuracy of estimates is proportional to the number of pairs, at greater lags (where there are fewer pairs) the estimates are less reliable. A common practice is to limit semivariogram estimation to lags with a minimum of 30 pairs [*Journal and Huijbregts*, 1978]. Another practical guideline is to limit the lag of the experimental semivariogram to half the extreme distance in the sampling domain for the direction to be analyzed [*Journal and Juijbregts*, 1978; *Clark*, 1979].

I determine the largest reliable lag for semivariogram estimation by analyzing the variance of the semivariogram families at each lag distance. A breakpoint exists at which the variance (s^2) starts to increase at a larger rate from that point to the end of the sample than before that point. This could be said to be the point where the variance goes from being a result of structural differences to being a result of too few sample pairs. For

relatively homogeneous media, this point can be determined by breaking the lag distance, h, versus s^2 plot into two parts and minimizing the sum of the squared differences between the actual s^2 values and a linear approximation of the two parts. However, as demonstrated by the results, this method may not be adequate when analyzing samples with stratification or other irregularities. In these cases the breakpoint is determined visually from the plot.

3.3.4 Automated Range Detection

The correlation length, or range, a, is currently estimated by best-fitting semivariogram models to the estimated semivariogram. However, a method to determine correlation length directly from the data would be more desirable than this empirical curve-fitting method. Estimated semivariogram values from equation 3.2 are distributed by a chi-square distribution [*Cressie*, 1991]. *Davis and Borgman* [1982] show that the distribution approaches a normal distribution as the number of data pairs goes to infinity for a random, second-order stationary function:

$$L\left\{\frac{\left[\hat{\gamma}(h)-\gamma(h)\right]}{\sigma\left[\hat{\gamma}(h)\right]}\right\} \rightarrow N(0,1)\,\text{as}\,n \rightarrow \infty, \tag{3.5}$$

where $L(X)$ denotes the law, or distribution of a random variable, $\hat{\gamma}(h)$ is the estimated semivariogram value, σ is the standard deviation, n is the number of data points, and $N(0,1)$ is the standard normal distribution.

Using this property on semivariogram families of CT images, which provide hundreds of thousands of data points, the expected distribution of

$$U(h) = \frac{\left[\hat{\gamma}(h)-\gamma(h)\right]}{\sigma\left[\hat{\gamma}(h)\right]} \tag{3.6}$$

at each lag distance h, would be the standard normal. However, if there is significant structure corresponding to correlation length h, the resulting distribution will be a multi-modal and not normally distributed. Similarly, the distribution of the $\hat{\gamma}(h)$ values at a particular lag distance can be tested for normality since the $\gamma(h)$ and $\sigma[\hat{\gamma}(h)]$ values can be considered constant at each lag distance. Therefore, by testing the normality of the $\hat{\gamma}(h)$ values at each lag I determine if there is structure occurring at that lag. In other words, the range of the semivariogram can be detected based on the distribution of the semivariogram family, instead of by empirical curve-fitting of semivariogram models.

To determine normality of the $\hat{\gamma}(h)$ distributions, the Shapiro-Wilk W test [*Shapiro and Wilk*, 1965], the Anderson-Darling test [*Anderson and Darling*, 1954], the Martinez-Iglewicz test [*Martinez and Iglewicz*, 1981], and the D'Agostino skewness test [*D'Agostino et al.*, 1990] have all been employed with a five percent decision criteria. A decision to reject normality from any one of these tests is considered sufficient to determine non-normality.

3.4 Materials and Methods

3.4.1 Sample Description

Three horizontal cores were collected from the Culebra Dolomite Member of the Rustler Formation from the Waste Isolation Pilot Plant (WIPP) site in southeastern New Mexico and are listed in Table 3.1. Core C2611A had a diameter of 147 mm and length of 450 mm. The barrel size for sample collection was selected to so that the core was representative with respect to heterogeneity [*Lucero et al.*, 1997]. *Vogel and Brown*

[2001] show that the two halves of this sample have different anisotropy and therefore it was analyzed in two sections (slices 30-90 and slices 90-150). C2611A had been used in numerous actinide transport experiments and was scanned saturated for operational reasons. The other two samples (C1AV and C2AV) had diameters of 37 mm and lengths of 52 mm and were subsamples of a larger core that visually appeared to have different compositions and differing geostatistical properties. C2AV contains a higher percentage of solid dolomite than C1AV [*Vogel and Brown*, 2001]. *Hseih et al.* [1998] and *Brown et al.* [2000] have also completed analyses of these core samples.

3.4.2 Computerized Tomography Imaging

A three-dimensional grid using the custom pencil-beam gamma ray CT scanner of *Brown et al.* [1993] has been used to scan the four samples. While much slower than x-ray CT, gamma CT does not suffer from beam hardening and photon scattering that distort x-ray images and would make geostatistical analysis problematic. The x and z axes on these samples are defined as horizontal, while the y axis is vertical. Cores C1AV and C2AV were imaged at 71 uniformly spaced slices with a voxel size of 0.37 mm by 0.37 mm by 0.75 mm, while core C2611A was imaged at 150 uniformly spaced slices, not all of which were used here, with voxel size of 1.5 mm by 1.5 mm by 3.0 mm. The z direction voxel spacing was double the spacing in the x and y direction.

3.4.3 Semivariogram Analysis

The FORTRAN 77 package of geostatical programs of *Deutsch and Journel* [1998] was used for these analyses. However, the GAM code was recompiled using Visual Fortran 6.6 to increase the maximum number of data points that could be analyzed. TableCurve 2D was used for best-fitting semivariogram models [*TableCurve 2D*, 1996] and NCSS 2000 was used for the normality tests.

3.5 Results and Discussion

3.5.1 Semivariogram Families

Semivariogram families have been plotted in the x and y direction for the four core samples (Figures 2 and 3). The number of members in each family corresponds to the number of voxels in the z direction as shown in Table 3.1. Analysis was not completed in the z direction along the core axis because of the initial slice groupings of the data, but would be possible with further data manipulation to group the slices perpendicular to the x or y direction. These samples do not exhibit a nugget effect in either the x or y direction at these resolutions. Sample C2611A (slice 30-90) (Figure 3.2b) exhibits pseudocycling in the y direction. A hole effect is shown for sample C2611A (slice 90-150) (Figure 3.2d) in the y direction. Sample C2611A (slice 90-150) in the x direction (Figure 3.2c) and C2AV in the y direction (Figure 3.3d) show slices that may be considered outliers. Further analysis of these slice semivariograms, however, indicates that these semivariograms are for sequential slices. Therefore, these outlier

48

semivariograms are a result of actual sample structure and not a result of operator error during scanning of that slice.

3.5.2 Semivariogram Estimate Reliability

In practice, rules of thumb have been generally applied to determine the largest lag on the semivariogram that constitutes a reliable estimate and therefore should be used for modeling and characterization purposes. The largest lag of reliable semivariogram estimates can also be determined from plots of the variance of the semivariogram families. Figures 3.4 and 3.5 show the sample variance (s^2) values of the semivariogram families at each lag and the break point where the variance goes from being a result of sample structure to being a result of too few lags and being at the extreme edges of the sample. The breakpoint is determined by minimizing the sum of the squared differences between the actual s^2 values and a linear approximation of the two parts. In directions where the plots are irregular (Figures 4b, 4d, and 5c), however, the break point cannot be selected by optimizing the linear models, but rather must be chosen by visually detecting the break point.

Table 3.2 summarizes these results. The percentage of reliable semivariogram estimates for these samples ranges from 54% to 78% and is generally larger in the x direction than the y direction. This is likely a result of greater heterogeneity from stratification in the vertical direction. As a check, the sum of the percentage of reliable estimates plus the percentage of pairs needed for a reliable estimate should be slightly greater than 100%. This check is not equal to exactly 100% because the maximum number of pairs will be slightly smaller than the total number of data points compared.

These results show more reliable lags than the rule of thumb that considers half of the extreme distance in the sampling domain for that direction as reliable. If the rule of thumb that 30 pairs were used as an estimate of reliable values, all semivariogram values from these CT images would be considered reliable.

3.5.3 Automated Range Detection

By analyzing the distribution of the $\hat{\gamma}(h)$ values at each lag, range values for each Culebra dolomite sample have been determined. Complete results of the normality tests at each lag for sample C2611A (slice 30-90) are shown in Figure 3.6. Of note are the groupings of non-normal semivariogram family distributions. These groupings could be said to represent elliptical structure within the porous media. *Jupp et al.* [1989] and *Swan and Garratt* [1995] discuss the use of ellipsoids to represent structure in geostatistical image analysis, and *Agterberg* [1970] describes a three-dimensional autocorrelation function as a hypersurface with an ellipsoid contour. As an example, a directional semivariogram of an ellipsoid with a major axis diameter of 11 lags would result in a grouping of correlation lengths similar to those exhibited in Figure 3.6, lag = 6 to lag = 11. The directional diameter of this ellipsoid corresponds to the largest value in this range. Single non-normal semivariogram family distributions may be a result of directional cracks or an anomaly of the normality tests at that lag. Examples of the normality plots for a normal and a non-normal lag distance from C2AV in the y direction are shown in Figure 3.7.

Each of the composite estimated semivariograms has been fitted with a combination of exponential semivariograms models using the detected range values. Table 3.3 lists the

ranges detected from analysis of semivariogram families, the number of consecutive non-normal semivariogram family distributions contributing the detected range, the sill values (w_i) and the R^2 value between the semivariogram model and the estimated semivariogram. All semivariograms have been fit with an R^2 value of at least 0.94. The number of consecutive non-normal semivariogram family distributions is generally proportional to the w_i values to the extent that a large number of consecutive non-normal distributions corresponds to larger w_i values. The exceptions to this are at small lag distances ($h < 4$ lags) and in the x direction of C2611A (30-90) at $h = 40.5$ mm. In this last case, the detected range at $h = 22.5$ mm may represent a slight hole and offset the next effect of additional structure at $h = 40.5$ mm in the best fit semivariogram model. Hole effects were also detected using this method for C2611A (30-90) in the y direction and C2611A (90-150) in the y direction.

The normality tests do not always detect ranges at small lag distances ($h < 4$ lags). Non-normality is probably not an indicator of a correlation at $h = 1$ lag, as demonstrated by the $w_i = 0.00$ values at that lag for samples C2611A (30-90) in the y direction, C2611A (90-150) in the y direction, and C2AV in the y direction. These results occur because of dependency due to the structure of the porous media and due to artifacts created by the back-projection reconstruction algorithm and filter function used for CT images (Brown et al., 1993). To offset the lack of automatic detection at these small lags, an additional exponential model is best fit to the estimated semivariogram with a maximum value equal to four lags. This method detected small-scale nested sills in samples C2611A (30-90) in the x direction, C1AV in the x direction, and C2AV in the x and y directions.

3.5.4 Semivariogram Comparisons

By determining the extent of reliable estimates on a semivariogram, one is tempted to compare semivariograms from two samples with statistical tests to that point to determine if the samples possess similar geostatistical characteristics. *Snedecor and Cochran* [1967] describe a method for comparing paired samples which may be applicable if the paired samples are normally distributed. However, as shown earlier in this paper, the sample pairs are not normally distributed at lag distances corresponding to the correlation lengths. A nonparametric test such as the Wilcoxon matched-pairs signed-ranks test as described by *Sheskin* [2000] could therefore be attempted. This test is completed by ranking the absolute values of the difference scores $\left(|D|\right)$ and summing the ranks of the positive scores $\left(\sum R+\right)$ and negative rank scores $\left(\sum R-\right)$ from $h = 0$ to the largest lag distance of a reliable semivariogram estimate. The smallest of these two values is then assigned to the Wilcoxon T test statistic and compared to the critical T value. However, this method fails as well since it discounts obvious trends if one semivariogram contains all values slightly less than the first (Figure 3.8a), while identifying two semivariograms that cross at the midpoint as similar even though the magnitude of the differences is large (Figure 3.8b). Comparisons of structure from semivariogram analysis should therefore continue to follow the classical procedure of comparing ranges, sills, and anisotropy.

3.6 Summary and Conclusions

Semivariogram families have been introduced and shown to provide improved methods for determining the spatial correlation structure of porous media. By analyzing the variance of the semivariogram values within the semivariogram family at each lag, the maximum reliable lag has been determined. The maximum reliable lag could be said to be the point of the semivariogram where the variance of the semivariogram values at that lag goes from being a result of structural differences to being a result of too few sample pairs. This quantification is an improvement on past methods that rely on rules of thumb to determine the maximum reliable lag on the semivariogram.

By testing the normality of the semivariogram values at each lag within a semivariogram family, range values are automatically detected that represent actual porous media structure. Current methodology is to best-fit empirical semivariogram models that result in range values, which may or may not correspond to actual structure.

Finally, the tempting proposition of using the increased characterization of geostatistical properties to apply a statistical test to compare semivariograms from two samples is explored. Because of normality concerns and the inherent nature of nonparametric tests, however, the classical procedure of comparing ranges, sills, and anisotropy is still recommended to compare the spatial correlation structure of two different samples.

3.7 References

Agterberg, F. P., Autocorrelation Functions in Geology, in *Geostatistics*, edited by D. F. Merriam, Plenum, NY, pp. 113-142, 1970.

Anderson, T. W. and D. A. Darling, A test of goodness-of-fit, *J. Amer. Statist. Assoc*, 49, 765-769, 1954.

Brown, G. O., H. T. Hsieh, and D. A. Lucero, Evaluation of laboratory dolomite core size using representative elementary volume concepts, *Water Resourc. Res.*, 36(5), 1199-1207, 2000.

Brown, G. O., M. L. Stone, and J. E. Gazin, Accuracy of gamma ray computerized tomography in porous media, *Water Resourc. Res.*, 29(2), 479-486, 1993.

Calvete, F. J. S., Application of geostatistics in subsurface hydrology, in *Subsurface Flow and Transport: A Stochastic Approach*, ed. By G. Dagan and S. P. Neuman, Cambridge University Press, Cambridge, UK, 1997.

Chauvet, P., The variogram cloud, in *Proceedings of the 17th APCOM International Symposium*, Golden, CO, 757-764, 1982.

Chilès, J.-P., and P. Delfiner, *Geostatistics, Modeling Spatial Uncertainty*, John Wiley and Sons, Inc., New York, pp. 34-38, 1999.

Clark, I., *Practical Geostatistics*, Applied Science Publishers Ltd., London, p. 14, 1979.
Cressie, N. A. C., *Statistics for Spatial Data*, John Wiley and Sons, NY, 1991.

D'Agostino, R. B., A. Belanger, R. B. D'Agostino Jr., A suggestion for using powerful and informative tests of normality, *The American Statistician*, 44(4), 316-321, 1990.

Davis, M. D., and L. E. Borgman, A note on the asymptotic distribution of the sample variogram, *J. Intnl. Assoc. Math. Geol.*, 14(2), 189-193, 1982.

Deutsch, C.V., and A.G. Journel, *GSLIB: Geostatistical Software Library and User's Guide*, Oxford University Press, 2nd edition, New York, 1998.

Gandin, L. S., *Ob"ektivnyi analiz meterologicheskikh polei*, Gidrometeologicheskoe Izdatel'stvo, Leningrad, Translation (1965), *Objective Analysis of Meteorological Fields*, Israel Program for Scientific Translations, Jerusalem, p. 47, 1963.

Gelhar, L. W., *Stochastic Subsurface Hydrology*, Prentice-Hall, Inc., Englewood Cliff, NJ, 1993.

Gelhar, L.W., Stochastic subsurface hydrology from theory to applications, *Water Resour. Res.*, 22(9), 1358-1458, 1986.

Goggin, D.J., M.A. Chandler, G. Kocurek, and L.W. Lake, Patterns of permeability in eolian deposits: Page Sandstone (Jurassic), Northeastern Arizona, *SPE Form. Eval.*, June, 297-306, 1988.

Greenholtz, D. E., T.-C. J. Yeh, M. S. B. Nash, and P. J. Wierenga, Geostatistical analysis of soil hydrologic properties in a field plot. *J. of Cont. Hydrology*, 3, 227-250, 1988.

Grevers, M. C. J., and E. de Jong, Evaluation of Soil-Pore Continuity Using Geostatistical Analysis on Macroporosity in Serial Sections Obtained by Computed Tomography Scanning, in *Tomography of Soil-Water-Root Processes, SSSA Special Publication no. 36*, Soil Science Society of America, Madison, WI, pp. 73-86, 1994.

Hsieh, H. -T., G. O. Brown, M. L. Stone, and D. A. Lucero, Measurement of porous media component content and heterogeneity using gamma ray tomography, *Water Resourc. Res.*, 34(3), 365-372, 1998.

Ioannidis, M. A., I. Chatzis, and M. J. Kwiecien, Computer enhanced core analysis for petrophysical properties, *J. Can. Pet. Tech.*, 38(3), 18-24, 1999.

Isaaks, E. H., and R. M. Srivastava, *An Introduction to Applied Geostatistics*, Oxford University Press, New York, 1989.

Journel, A. G., and C. J. Huijbregts, *Mining Geostatistics*, Academic Press, London, p. 194, 1978.

Jupp, D. L. B., A. H. Strahler, and C. E. Woodcock, Autocorrelation and regularization in digital images II. Simple image models, *IEEE Transactions on Geoscience and Remote Sensing*, 27(3), 247-258, 1989.

Lucero, D. A., G. O. Brown, and C. E. Heath, *Laboratory Column Experiments for Radionuclide Adsorption Studies of the Culebra Dolomite Member of the Rustler Formation*, SAND97-1763, Sandia National Laboratories, Albuquerque, NM, 1997.

Martinez, J. and B. Iglewicz, A test for departure from normality based on a biweight estimator of scale, *Biometrika*, 68, 331-333. 1981.

Meisch, A. T., Variograms and variance components in geochemistry and ore evaluation, *Geologic Society of America Memoir 142*, 333-340, 1975.

Olea, R. A., *Geostatistics for Engineers and Earth Scientists*, Kluwer Academic Publishers, Boston, 67-75, 1999.

Oliver, M. A., and R. Webster, Semi-variograms for modeling the spatial pattern of landform and soil properties, *Earth Surface Processes and Landforms*, 11, 491-504, 1986.

Rasiah, V. and L. A. G. Aylmore, Computed tomography data on soil structural and hydraulic parameters assessed for spatial continuity by semivariogram geostatistics, *Aust. J. Soil Res.*, 36, 485-93, 1998.

Schafmeister-Spierling, M. -Th. and A. Pekdeger, Influence of Spatial Variability of Aquifer Properties on Groundwater Flow and Dispersion, in *Contaminant Transport in Groundwater*, Dobus and Kinzelbach (eds.), Balkema, Rotterdam, The Netherlands, pp. 215-220, 1989.

Shapiro, S. S. and M. B. Wilk, An analysis of variance test for normality, *Biometrika*, 52, 591-611, 1965.

Sheskin, D. J., *Handbook of Parametric and Nonparametric Statistical Procedures*, 2nd Ed., Chapman and Hall/CRC, Boca Raton, pp. 467-477, 2000.

Snedecor, G. W., and W. G. Cochran, *Statistical Methods*, 6th Ed., The Iowa State University Press, Ames, IA, 91-116, 1967.

Solie, J. B., W. R. Raun, and M. L. Stone, Submeter spatial variability of selected soil and bermudagrass production variables, *Soil Sci. Soc. Am. J.*, 63(6), 1724-1733, 1999.

Swan, A. R. H. and J. A. Garrett, Image analysis of petrographic textures and fabrics using semivariance, *Mineralogical Magazine*, 59(2), 189-196, 1995.

Tavchandjian, O., A. Rouleau, and D. Marcotte, Indicator Approach to Characterize Fracture Spatial Distribution in Shear Zones, in *Geostatistics Troia 1992*, A. Soares (ed), Kluwer, Quantitative Geology and Geostatistics Series, 1, pp. 965-976, 1993.

Tercier, P., R. Knight, and H. Jol, A comparison of the correlation structure in GPR images of deltaic and barrier-split depositional environments, *Geophysics*, 65(4), 1142-1153, 2000.

Vogel, J. R., and G. O. Brown, Geostatistics of computerized tomography images of the Culebra Dolomite, *Water Resourc. Res.*, in review, 2001.

Vogel, J. R., and C. T. Haan, Chapter 17: Geostatistics, in *Statistical Methods in Hydrology*, 2nd ed., C. T. Haan, Iowa State University Press, Ames, IA, *in press*, 2001.

Woodbury, A. D., and E. A. Sudicky, The geostatistical characteristics of the Borden Aquifer, *Water Resour. Res.*, 27(4), 533-546, 1991.

Yates, S. R., A. W. Warrick, A. D. Matthias, and S. Musil, Spatial variability of remotely sensed surface temperatures at a field scale, *Soil Sci. Soc. Am. J.*, 52, 40-45.

Table 3.1. Culebra Dolomite sample descriptions.

Sample Name	Sample Dimensions, X x Y x Z, (Voxels)	Voxel Size (mm)	Bulk Density (g/ml)	
			Mean	St. Dev.
C2611A, slice 30-90	61 x 81 x 61 (301,401)	1.5 x 1.5 x 3.0	2.57	0.19
C2611A, slice 90-150	61 x 81 x 61 (301,401)	1.5 x 1.5 x 3.0	2.53	0.19
C1av	71 x 71 x 46 (231,886)	0.37 x 0.37 x 0.75	2.32	0.38
C2av	71 x 71 x 46 (231,886)	0.37 x 0.37 x 0.75	2.44	0.30

Table 3.2. Summary of s^2 breakpoint data for determining maximum lag at which a reliable semivariogram can be estimated.

Sample	Direction	Break-point (mm)	% Reliable Estimates	# of Pairs in $\hat{\gamma}$ (h) at Breakpoint	Max # Pairs (Pairs at $h = 1$)	% of Pairs for Reliable Estimate
C2611A (30-90)	x	70.5	78	69,174	296,460	23
	y	64.5	54	141,398	297,680	48
C2611A (90-150)	x	66.0	73	83,997	296,460	28
	y	78.0	65	107,909	297,680	36
C1AV	x	17.4	67	78,384	228,620	34
	y	16.7	64	84,916	228,620	37
C2AV	x	18.1	70	71,852	228,620	31
	y	15.2	59	97,980	228,620	43

Table 3.3. Range and semivariogram model parameters determined from semivariogram family analysis for Culebra dolomite samples. Lag distance is 1.5 mm for core C26a and 0.37 mm for cores C1AV and C2AV.

Sample	Direction	Range, a (mm)	# of Consecutive Non-normal Lag Distributions	Sill, w_i	R^2
C2611A (30-90)	x	4.2 [1]	--	0.27	0.97
		16.5	6	0.03	
		31.5	1	0.00	
		52.5	11	0.23	
		69.0	7	0.38	
	y	1.5	1	0.00	0.99 [2]
		4.5	1	0.35	
		24.0	6	0.84	
		37.5	5	hole [2]	
		57.0	7	pseudocycling	
		64.5	2	hole	
C2611A (90-150)	x	3.0	2	0.31	0.99
		22.5	2	0.39	
		40.5	7	0.00	
		66.0	19	0.21	
	y	1.5	1	0.00	0.99 [3]
		6.0	1	0.46	
		18.0	6	0.38	
		43.5	8	0.38	
		64.5	4	hole [3]	
C1AV	x	3.1 [1]	--	0.86	0.98
		11.1	3	0.00	
	y	4.1	1	0.73	0.96
		9.3	1	0.00	
		10.0	1	0.30	
C2AV	x	1.4 [1]	--	0.58	0.94
		4.1	9	0.11	
		10.4	3	0.00	
		11.1	1	0.07	
		12.6	2	0.22	
	y	0.4	1	0.00	0.98
		1.4 [1]	--	0.53	
		5.6	12	0.32	
		10.0	2	0.13	
		11.8	3	0.06	
		13.3	3	0.00	

[1] These ranges are not detected from semivariogram families because they occur at small lags. They are determined from the best fit of the semivariogram model.

[2] R^2 calculated only for range values up to 24 mm because of pseudocycling for this sample in this direction.

[3] R^2 calculated only for range values up to 43.5 mm because of the hole effect for this sample in this direction.
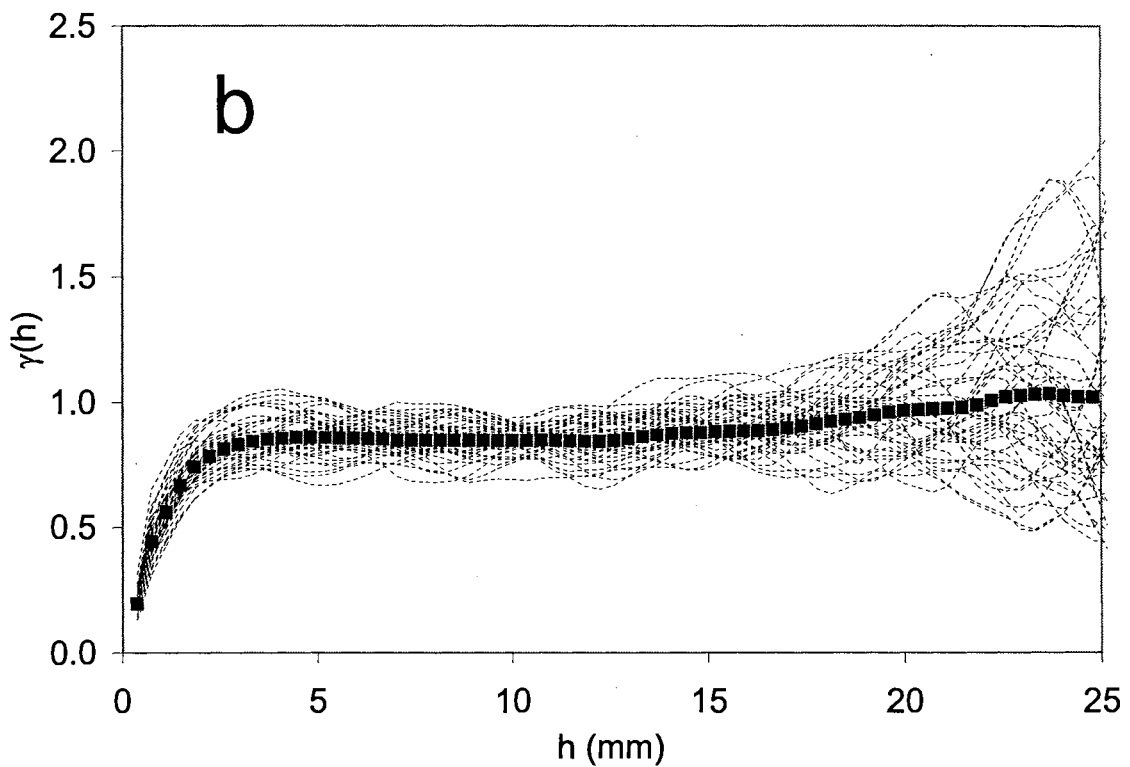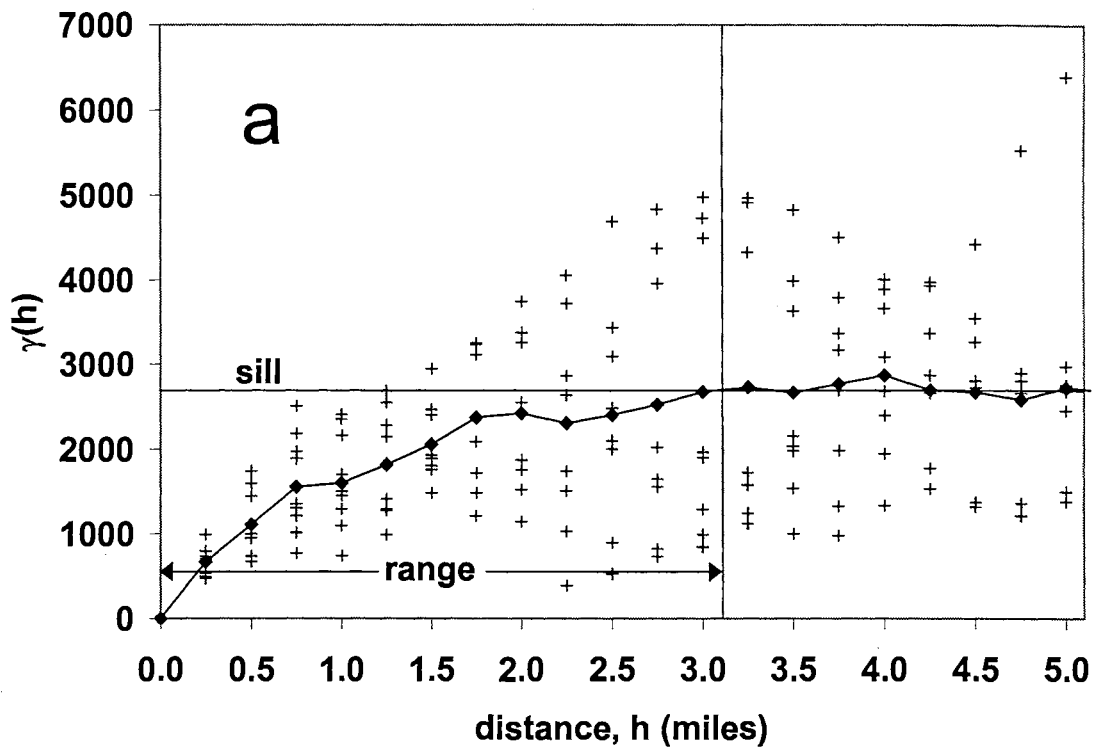
Figure 3.1. A semivariogram cloud from Vogel and Haan [2001] (a) and a semivariogram family (b). The pluses and diamonds on (a) represent the semivariogram cloud and composite semivariogram, respectively. The dashed lines and squares on (b) represent the semivariogram family members and the composite sample semivariogram, respectively.
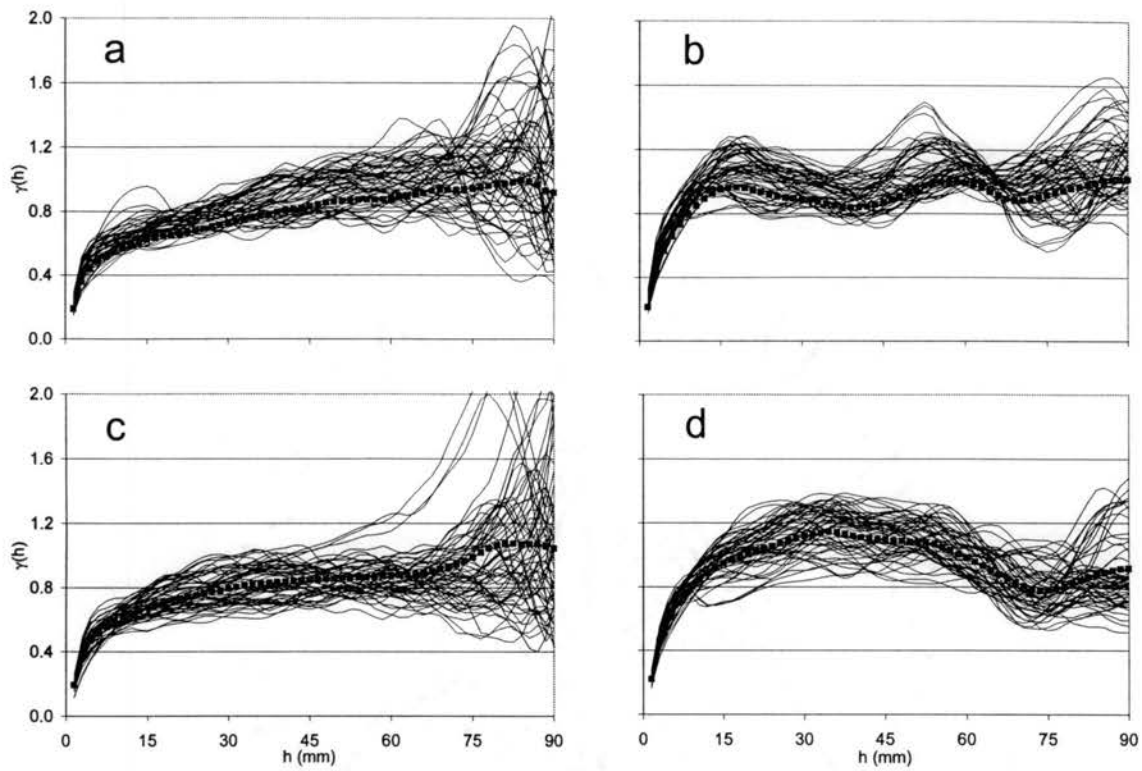
Figure 3.2. Semivariogram families (lines) and composite semivariograms (squares) for
sample C2611A for (a) slices 30-90 in the x direction, (b) slices 30-90 in the y direction,
(c) slices 90-150 in the x direction, and (d) slices 90-150 in the y direction.
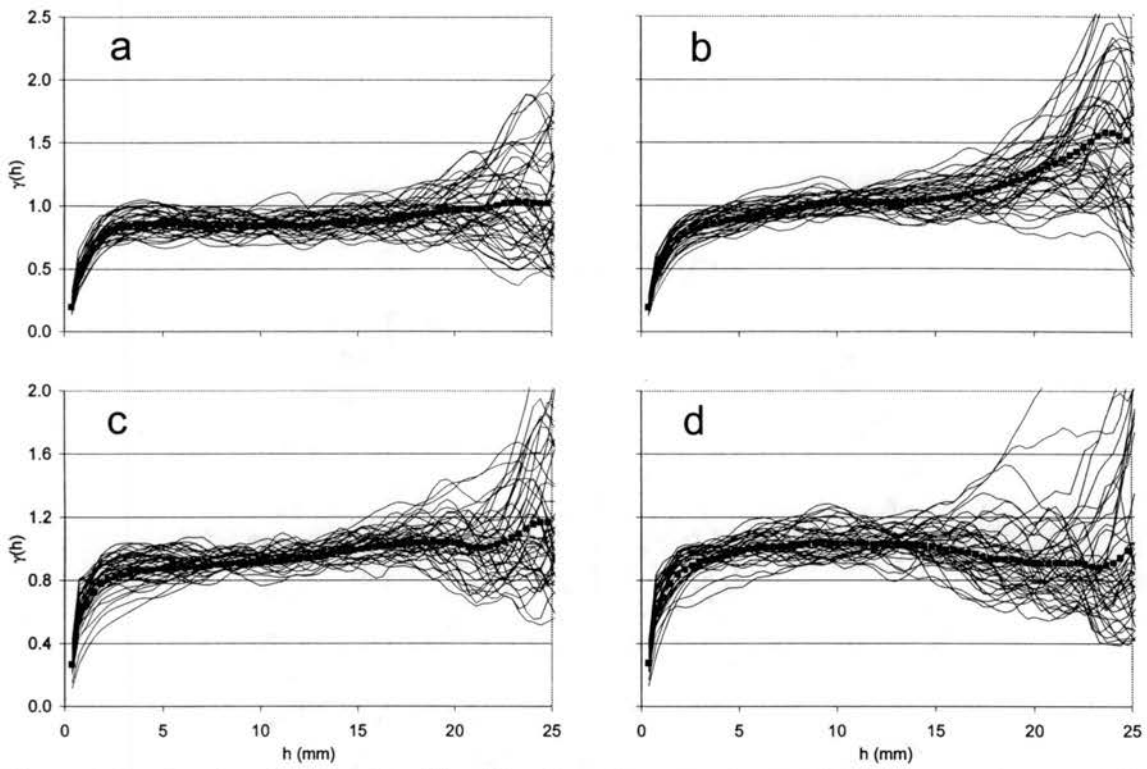
Figure 3.3. Semivariogram families (lines) and composite semivariograms (squares) for (a) C1AV in the x direction, (b) C1AV in the y direction, (c) C2AV in the x direction, and (d) C2AV in the y direction.

63
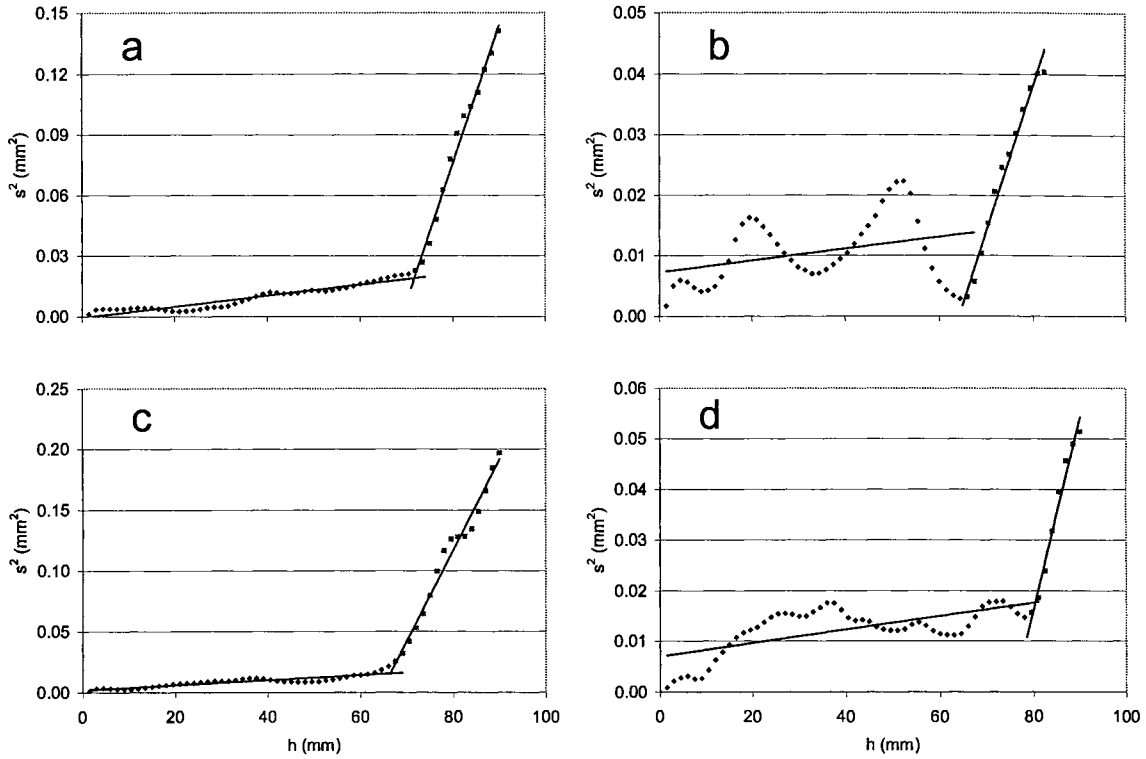
Figure 3.4. Variance of the semivariogram family distributions for core C2611A (a) slice 30-90 in the x direction, (b) slice 30-90 in the y direction, (c) slice 90-150 in the x direction, and (d) slice 90-150 in the y direction.
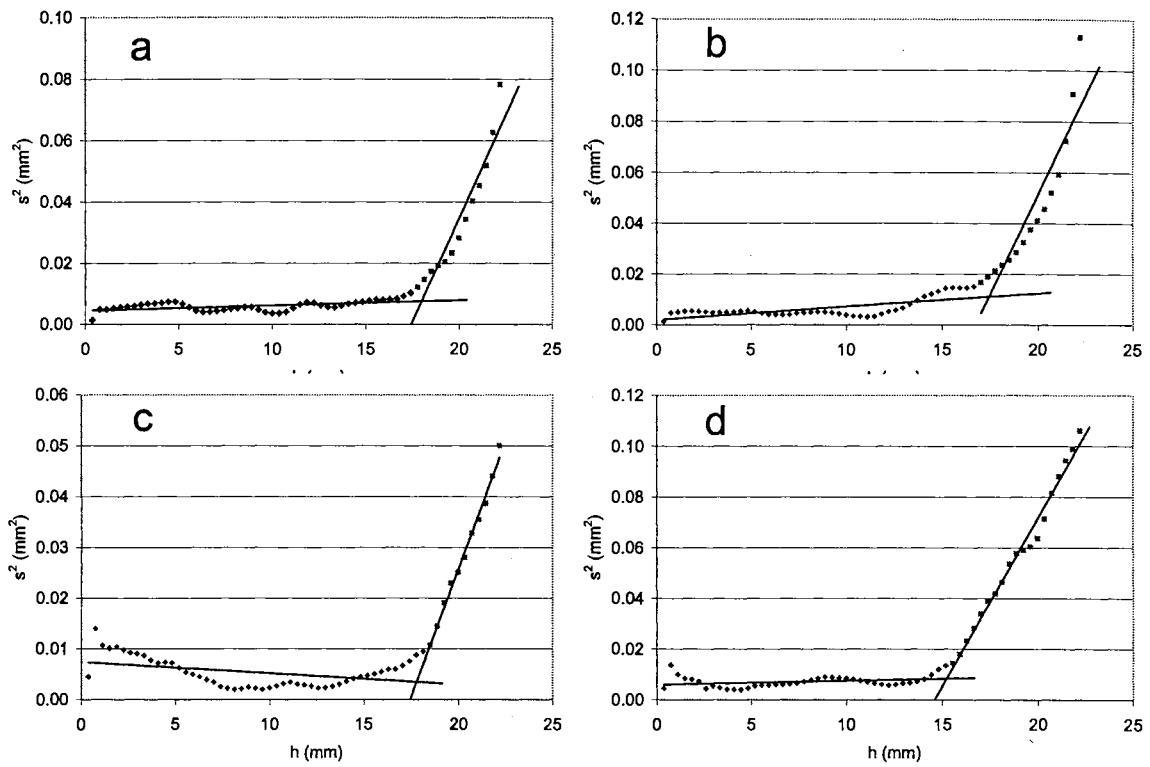
Figure 3.5. Variance of the semivariogram family distributions for (a) C1AV in the x direction, (b) C1AV in the y direction, (c) C2AV in the x direction, and (d) C2AV in the y direction.
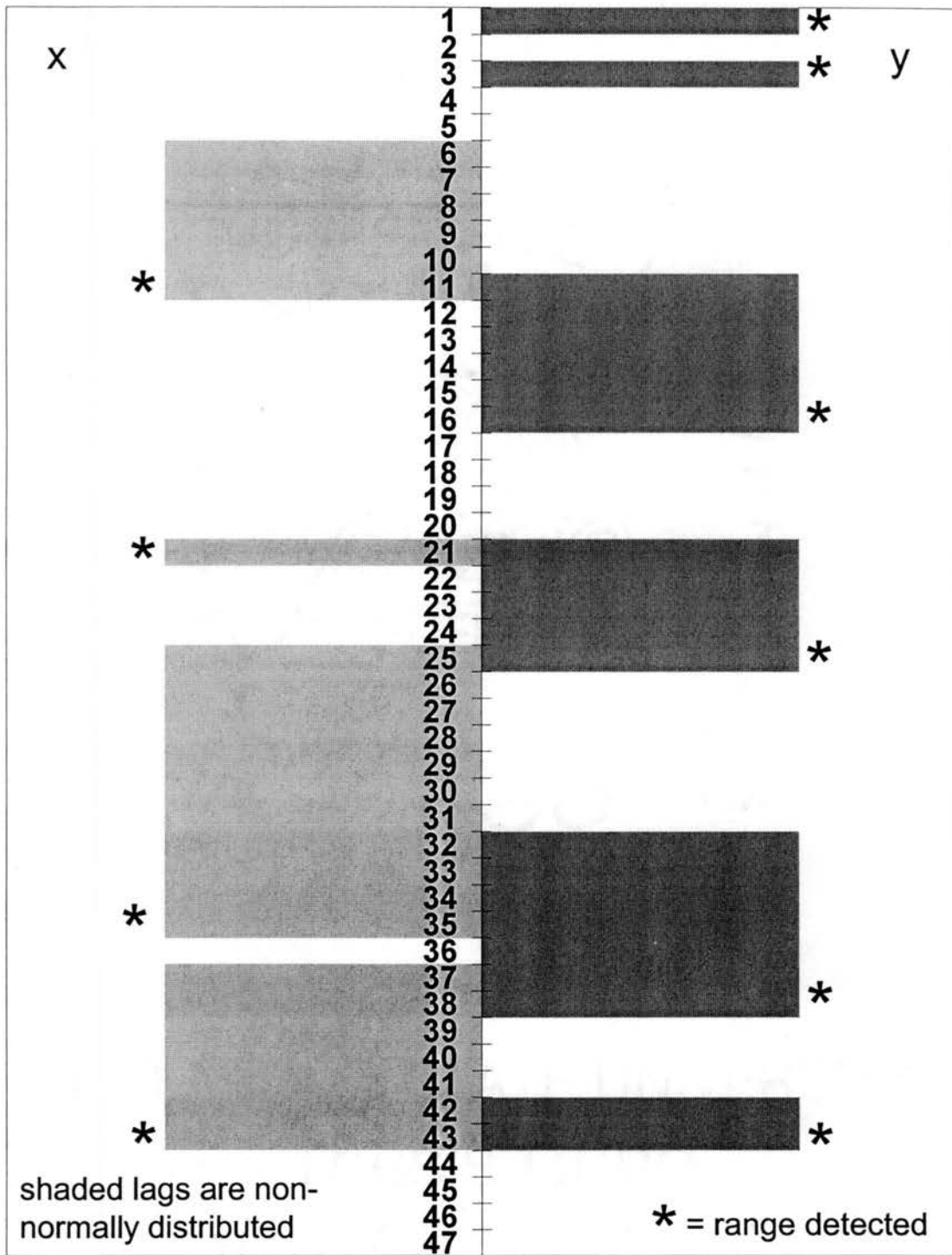
Figure 3.6. Results of normality tests for each lag of the semivariogram families of C2611A (slice 30-90). One lag is equal to 1.5 mm.
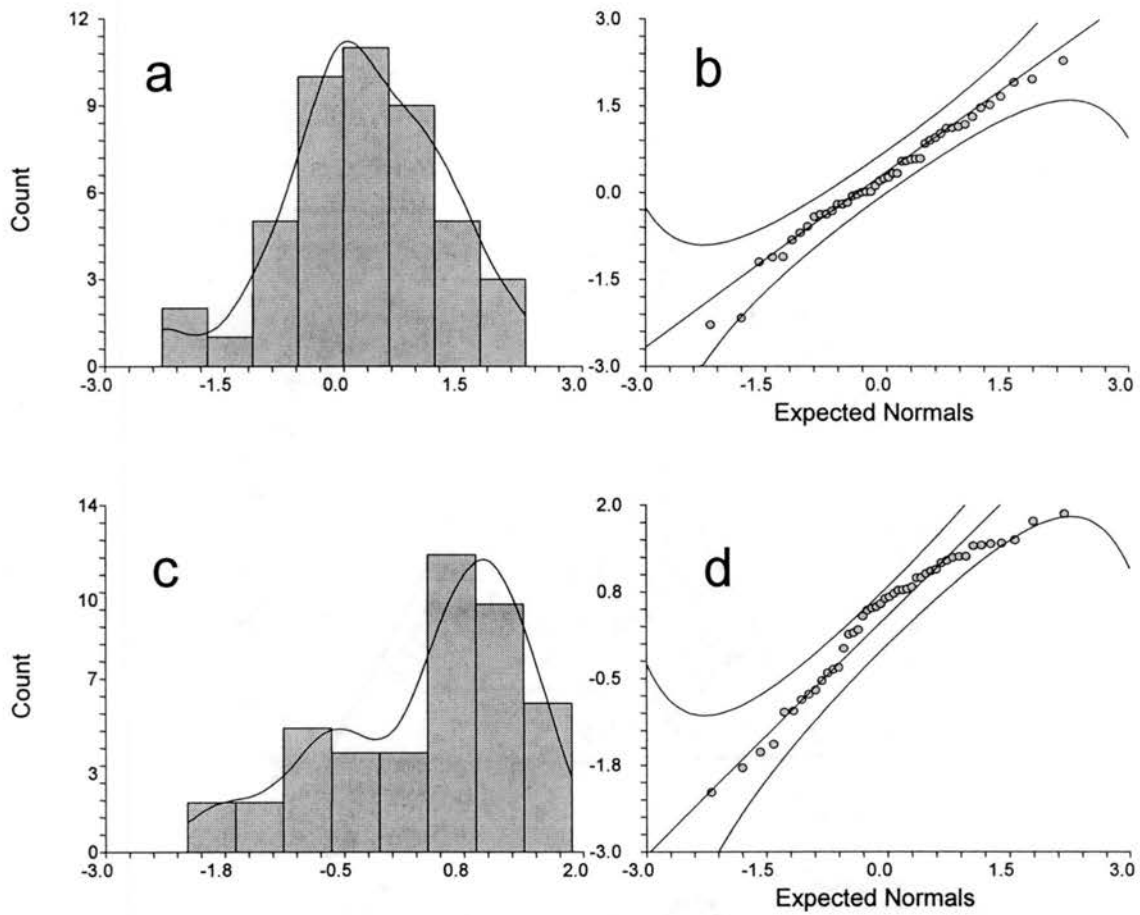
Figure 3.7. Example frequency histogram and normality plot of the $\hat{\gamma}(h)$ values for C2AV in the y direction at a lag distance that is normally distributed (a and b) and non-normally distributed (c and d).
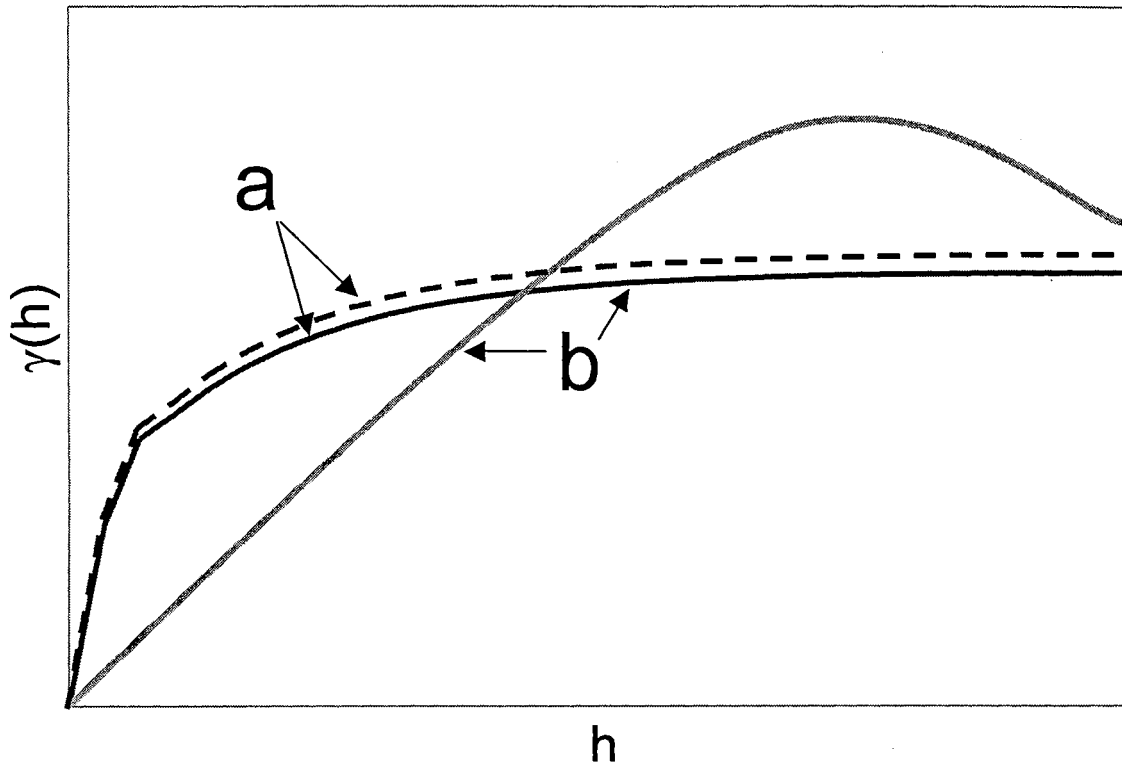
Figure 3.8. Example semivariograms for the Wilcoxon matched-pairs signed-ranks test.

CHAPTER FOUR

HETEROGENEOUS POROUS MEDIA STRUCTURE ANALYSIS WITH A CONDITIONAL SEMIVARIOGRAM AND COMPUTER TOMOGRAPHY IMAGES

4.1 Abstract

A conditional semivariogram has been introduced to characterize the structural properties of porous media features that fall within in discrete bulk density bins. In addition, artificial images, or phantoms, are used to identify semivariograms associated with several characteristic rock features. The conditional semivariogram and phantom image analysis is then combined with histograms and traditional semivariograms to create a realistic stochastic representation of the porous media.

4.2 Introduction

Soil and groundwater research is increasingly using computerized tomography (CT) images for analysis of porous media properties. *Anderson and Hopmans* [1994] describe many of these applications. Geostatistical analysis of these images provides a non-destructive method for obtaining small-scale structural properties of porous media. *Rasiah and Aylmore* [1998] and *Vogel and Brown* [2001] use CT images of gamma ray attenuation for geostatistical characterization of Culebra dolomite and soil cores, respectively. *Tercier et al.* [2000] uses ground-penetrating radar images to compare the

correlation structure of different depositional environments. Two dimensional geostatistical image analysis and three dimensional reconstruction is used by *Ioannidis et al.* [1999] to predict aquifer properties for use flow and transport models.

Geostatistical analysis can also be used for recognition of the shape the structure within images. *Jupp et al.* [1989] uses semivariogram analysis of remote-sensed images to show that overlapping disks can represent spatial structure. Semivariogram analysis of a moving window on radar images was used to classify land use [*Miranda et al.*, 1992], and ice crevace patterns are characterized using geostatistics by *Herzfeld and Zahner* [2001]. *Swan and Garratt* [1995] uses semivariance of digitized images of thin rock sections to classify texture. CT images can also be used for density frequency analysis [*Hsieh et al.*, 1998; *Clausnitzer and Hopmans*, 1999]. Indicator semivariograms have also been used for strict linear two-phase geostatistical characterization by *Grevers and DeJong* [1994] and *Tavchandjian et al.* [1993] for analysis of soil and rock fractures.

Semivariograms can be used as a significant tool to identify major features in a rock formation. By combining pattern recognition, density frequency analysis for phase differentiation, and semivariograms, the major features in the porous media can be identified to create a stochastic representation of the media. This stochastic representation could then be applied to a flow and transport model or used as an initial image for a simulation model such as simulated annealing [*Goovaerts*, 2000]. *Goovaerts* [1997] states that the initial image for simulated annealing which is realistic and meet target constraints will improve the optimization process. The stochastic representation generated using these methods that is used an initial image would be more realistic than a random image that only matches the histogram.

This paper uses semivariogram analysis of artificial images, or "phantoms", to identify dominant features of Culebra dolomite. A conditional semivariogram is also introduced to characterize geostatistical properties of different structures for three different types of heterogeneous porous media. These different structures are identified by a strict linear differentiation of the bulk density to define phase bins. Finally, by combining the phantom image analysis with the results of the conditional semivariogram, a stochastic representation of the porous media is generated incorporating the geostatistical structure in each of the phase bins.

4.3 Theory

4.3.1 Semivariograms

Semivariograms measure the degree of dissimilarity between an unsampled value and a nearby data value. Also called the structure function by Gandin (1963), Traditional semivariograms are defined as

$$\gamma(h) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} (x_i - y_i)^2 \tag{4.1}$$

where $\gamma(h)$ is half of the average squared difference; $N(h)$ is the number of pairs; x_i is the value at the start, or tail; y_i is the value at the end, or head, and h is the vector between the two data points.

A sill and range characterize the semivariogram. The sill is a plateau in the semivariogram values that corresponds to the variance of the sampled data. If the semivariogram is standardized, the $\gamma(h)$ value is divided by the variance and the sill value

will be equal to one. The range, or correlation length, is the distance at which the semivariogram reaches a plateau. A non-zero intercept on the semivariogram may exist and is termed the nugget effect. It may result from sampling error or variability at scales less than the smallest sampling interval [*Cressie*, 1991]. Another possible semivariogram characteristic is the hole effect, which is described as a semivariogram that peaks and dips, suggesting that at greater distances samples are more related. The hole effect may be seen as the rough spacing between adjacent lenses or bedding planes.

Semivariograms are fit by a number of analytical relationships to ensure that the semivariogram model is positive definite. The exponential model is often used by researchers in stochastic hydrology [*Woodbury and Sudicky*, 1991], and is given by

$$\gamma(h) = 1 - e^{\left(-\frac{3h}{a}\right)}$$ (4.2)

where a is the range and h is the lag distance. The exponential model reaches its sill asymptotically, with the practical range a defined as that distance at which the semivariogram value is 95% of the sill. Another common semivariogram model is the Gaussian model, which is given by

$$\gamma(h) = 1 - e^{\left(-\frac{3h^2}{a^2}\right)}$$ (4.3)

and is parabolic in shape. It is often representative of highly regular phenomenon such as elevation of gently undulating hills [*Goovaerts*, 1997]. The final model used is the spherical model,

$$\gamma(h) = \begin{cases} 1.5\dfrac{h}{a} - 0.5\left(\dfrac{h}{a}\right)^3 & \text{if } h < a, \\ 1 & \text{elsewhere.} \end{cases} \tag{4.4}$$

The spherical model has a linear behavior near the origin, flattens out at larger distances, and reaches the sill at a [*Isaaks and Srivastava*, 1989].

Multiple, or nested, sills in the semivariogram may be associated with physical phenomena occurring at different scales. Detection of ranges nested within larger ranges requires sample spacing shorter than the minimum detected range. In general, nested sills are only considered when they can be associated with physical phenomena [*Solie et al.*, 1999]. Nested sills can be modeled by using a linear combination of positive definite semivariogram models with positive coefficients. This property results in the family of positive definite models

$$\gamma(h) = \sum_{i=1}^{n} |w_i| \gamma_i(h) \tag{4.5}$$

which are positive definite as long as the n individual models are all positive definite. A weighting function, w_i, is defined for each individual model subject to $\sum_{i=1}^{n} w_i = 1$ [*Isaaks and Srivastava*, 1989].

4.3.2 Conditional Semivariograms

The ability to differentiate between rock features with different densities would simplify the task of interpreting the semivariogram. The indicator semivariogram allows for differentiation into two phases [*Goovaerts*, 1997]. However, I introduce a conditional

semivariogram to determine the structure of multiple rock features that fall in specified density ranges. A conditional semivariogram is defined as

$$\gamma_j(h) = \frac{1}{2N_j(h)} \sum_{i=1}^{N(h)} (x_i - y_i)^2 \qquad \text{if } x_i \in X_j \qquad (4.8)$$

where the value $(x_i - y_i)^2$ is included in $\gamma_j(h)$ if and only if the value of x_i is in the range of bin X_j, and $N_j(h)$ is the number of pairs included in the calculation.

Using this conditional semivariogram, only the spatial characteristics in the density range represented by that bin are included in the calculation. In effect, the average structure size of that "phase" can be determined. If intrinsic stationarity were assumed for the traditional semivariogram calculation, the intrinsic assumption would be appropriate within each bin for this calculation as well.

4.3.3 Phantom Images

Computer tomography uses phantom images for algorithm testing and calibration. Phantom images are artificial "scans," which are theoretically generated to see if the reconstruction algorithm can reproduce the original data. A similar approach is used here to shown how the characteristic of semivariograms are influenced by various ideal rock geometries. Semivariogram analysis of phantom images containing properties common in the Culebra dolomite formation indicates distinct shapes in the semivariogram, which can be used in interpreting the major characteristics of the rock.

Holt [1997], *Hsieh et al.* [1998], *Brown et al.* [2000], and visual inspection of Culebra dolomite cores indicate that solid dolomite; bedding planes made up of a large number of

74

vugs; vertical gypsum-filled cracks; gypsum infilling; silty dolomite outcrops, and dissolution voids are the major features of this rock formation. The distribution of bulk density of each of these features except mixed voxels is normally distributed and summarized in Table 4.1. Mixed voxels, which are primarily areas with voids smaller than the scanning resolution, are approximated by a beta distribution

$$f_m^*\left(\frac{\rho_m}{\rho_{\max}}\right) = \left[\left(\frac{\rho_m}{\rho_{\max}}\right)^{\alpha-1}\left(1-\frac{\rho_m}{\rho_{\max}}\right)^{\beta-1}\right]\frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)} \qquad 0 \le \frac{\rho_m}{\rho_{\max}} \le 1 \qquad (4.6)$$

$$f_m^*\left(\frac{\rho_m}{\rho_{\max}}\right) = 0 \qquad\qquad\qquad elsewhere$$

where $f_m^*\left(\dfrac{\rho_m}{\rho_{\max}}\right)$ is the true mixed voxel distribution, Γ is the gamma function, α and β are fitting coefficients, and ρ_{max} is the largest density in the distribution. The range of the beta distribution (Equation 4.6) for mixed voxels is from zero to 2.52, the bulk density of solid dolomite. The fitting parameter α is estimated to range from 3.30 to 7.80 and fitting parameter β from 0.41 to 1.73 for Culebra dolomite *Hsieh et al.* [1998]. Using these properties, two-dimensional, 100 x 100 pixel phantom images were constructed. Random values for the beta distribution are generated by

$$y' = \frac{\displaystyle\sum_{i=1}^{2\alpha} R_{N_i}^{\,2}}{\displaystyle\sum_{i=1}^{2\alpha} R_{N_i}^{\,2} + \sum_{i=2\alpha+1}^{2\alpha+2\beta} R_{N_i}^{\,2}} \qquad (4.7)$$

where α and β are integers, and R_{Ni} is a random value from a normal distribution with $\mu = 0$ and $\sigma = 1$ (Hahn and Shapiro, 1967). *Clausnitzer and Hopmans* [1999] describe a additional scheme for phase volume differentiation in two-phase systems which fits two

normal probability density functions and analytical expression to account for mixed voxels.

4.3.4 Semivariogram Interpretation

Knowledge of the semivariograms representative of characteristics associated with rock characteristics can aid in determining features in specific rock formations. The use of semivariograms, however, is not an absolute, as one semivariogram could be representative of numerous combinations of features. The more complex the rock, the harder it is to apply these principles. Larger resolution allows for larger scans with the same number of data points and therefore increases the possibility of larger-scale variability represented in the scan, but may not detect smaller resolution variability. Resolution selection should be based upon the scale of the problem being investigated.

4.4 MATERIALS AND METHODS

4.4.1 Sample Description

A rock core sample (C2AV) was taken from the Culebra Dolomite Member of the Rustler Formation from drill holes near the Waste Isolation Pilot Plant (WIPP) site in southeastern New Mexico. The WIPP is a U.S. Department of Energy research-and-development facility designed to demonstrate safe disposal of transuranic radioactive waste resulting from the United States' defense programs. The rock sample was obtained because the Culebra Dolomite is the most transmissive confined unit above the waste repository [*Lappin et al.,* 1989].

The large number of mixed voxels in the Culebra makes it problematic to recreate in a stochastic representation. As such, a simpler material, concrete was used. Four concrete cores (A1A, A2A, A3A, A4A) were obtained from the same bridge deck. Major features of the concrete cores include small voids, aggregates, and cement.

4.4.2 Semivariogram Analysis

Deutsch and Journel [1998] provides a package of Fortran 77 geostatistical programs. The GAM code for semivariogram analysis with a regular grid was recompiled using Visual Fortran 6.0 to increase the maximum number of data points which could be analyzed. Semivariogram generation was then completed using the recompiled GAM program. An additional Fortran code called CONVAR was developed and used for conditional semivariogram analysis.

4.5 Results And Discussion

4.5.1 Geostatistics of Phantom Images

Figures 4.1-4.5 show the phantom images and semivariograms for several characteristic properties of Culebra dolomite. The uniform phantom image in Figure 4.1a is solid dolomite. The bulk density is normally distributed with a mean and standard deviation as indicated in Table 4.1. The standardized results indicate a horizontal semivariogram with a sill and nugget at 1.0. The range is smaller than the smallest lag interval of 1 pixel.

The mean trend phantom image in Figure 4.1b represents solid dolomite with a normally distributed bulk density. The mean increases from 2.63 g/ml to 2.73 g/ml from left to right. The standardized semivariogram indicates this with an increasing $\gamma(h)$ value at larger lag distances. The nugget is approximately 0.85.

Figure 4.1c represents bedding planes found in solid dolomite. These are voids caused by layers of trapped gases in the sedimentary material. The phantom image has bedding planes in rows 19 and 45. The peaks on the semivariogram at approximately 19, 45, 55, and 81 pixels. These distances correspond to the distances to each bedding plane from the edges of the scanned area. In regards to the peak locations, the semivariogram is a mirror image at the center (50 pixels). For lags larger than 81 pixels, the semivariogram is linear because this represents comparisons of uniform regions. The semivariogram associated with bedding planes is an example of zonal anisotropy with strata, as the vertical semivariogram reaches a higher sill than the horizontal variogram because of larger average differences in the vertical direction.

Gypsum outcrops are also common in this rock formation. Phantom images of rectangular and oval-shaped gypsum outcrops have been generated and analyzed. The rectangular gypsum outcrop in Figure 4.2a extends from pixel (60, 30) to pixel (90, 80) on the 100. The x-direction semivariogram shows peaks at lags of 30 and 60 pixels, which correspond to the width of the uniform regions in that direction. The slope of the x-direction semivariogram also changes at a lag of 10 pixels, which represents the width of the smaller uniform region on the right side of the image. The y-direction semivariogram shows a peak at a lag of 50 pixels, which corresponds to height of the outcrop on the phantom image. Changes of the slope of the semivariogram are also

78

indicated at lags of 20 and 30 pixels, corresponding to the heights of the dolomite areas above and below the outcrop. Figure 4.2b represents an oval-shaped gypsum outcrop with a major axis length of 35 pixels in the y-direction and a minor axis length of 30 pixels in the x-direction and uniform dolomite surrounding. Results similar to the rectangular outcrop with regard to correlation expected ranges were found for the oval gypsum outcrop.

Figures 4.2c-4.3b represent uniform dolomite with one, two, or three small oval gypsum outcrops, respectively. The outcrops have a major axis of 17 pixels in the y-direction and 10 pixels in the x-direction. The semivariograms for all three of these representations have similar ranges of 17 and 10 pixels in the y and x-direction, respectively. For lags larger than the range in each semivariogram, the three semivariograms vary and are dependent on the physical location of the outcrops in each phantom image. The normalized nugget for the image with one small gypsum outcrop is approximately 0.8, compared to 0.7 for two gypsum outcrops and 0.6 for three outcrops. This indicates that the nugget decreases as the number of outcrops increases.

Another prominent characteristic of the Culebra dolomite is hollow voids, or vugs, spaced randomly throughout the sample. Figures 4.3c and 4.4a represent solid dolomite with one and two circle voids, respectively. The voids have a nine pixel diameter and are randomly located in the dolomite. The semivariogram for one vug (Figure 4.3c) shows an abrupt change of slope in each direction at a nine pixel lag and peaks at approximately 21 and 70 pixels. The change of slope is associated with the size of the void, and the two peaks are associated with the physical location of the vug in the matrix. The semivariogram in Figure 4.4a is for two vugs with a diameter of nine pixels. As in Figure

4.3c, there is an abrupt change of slope at 9 pixels representing the width of the voids. The y-direction semivariogram exhibits pseudocycling with peaks associated with location features of the voids. The x-direction variogram would exhibit similar behavior for two holes except for the fact that by coincidence the edges of the two holes are equidistant from the left and right edge of the image. Because of this, the effect of the second void mirrors onto the first and appears to be only one void. Both Figures 4.3c and 4.4a show a standardized nugget of approximately 0.2.

Figure 4.4b represents twenty very small circular voids with a diameter of three pixels in a uniform dolomite formation. The semivariogram clearly indicates the three pixel range associated with the voids. At lags greater than three pixels the semivariogram in each direction shows uneven pseudocycling associated with location of the voids. The standardized nugget for this phantom image is approximately 0.6.

Figures 4.4c-4.5c represent vertical gypsum-filled cracks in a uniform dolomite formation. Analysis of the first three figures reveals that differing lengths and densities of cracks provide different ranges and nuggets for the semivariograms. Figure 4.4c represents twenty randomly spaced vertical cracks 10 pixels long. This results in a y-direction semivariogram with a range of 10 pixels and a standardized nugget of approximately 0.75. The phantom image in Figure 4.5a with 20 randomly spaced vertical cracks 20 pixels long results in a y-direction semivariogram with a range of 20 pixels and a standardized nugget of approximately 0.6. Finally, Figure 4.5b shows a phantom image with 40 randomly spaced vertical cracks 10 pixels long. This results in a y-direction semivariogram with a range of 10 pixels and a standardized nugget at approximately 0.6. These results indicate that the range is the length of the cracks in that direction and that as

crack density increases the standardized nugget decreases. Figure 4.5c represents a uniform dolomite with 20 randomly spaced vertical gypsum-filled cracks 10 pixels long and 20 randomly spaced vertical gypsum-filled cracks 20 pixels long. The resulting y-direction semivariogram has a range of 20 pixels and a standardized nugget of approximately 0.5. There is also a noticeable decrease in the slope of the semivariogram at 10 pixels, corresponding to the length of the shorter cracks. This semivariogram is also more representative of the spherical and exponential models generally used for modeling semivariograms in stochastic hydrology [*Isaaks and Srivastava*, 1989; *Woodbury and Sudicky*, 1991].

4.5.2 Sample Statistics

Four two-dimensional concrete cores and one three-dimensional Culebra dolomite core have been analyzed. Characteristics of these cores are summarized in Table 4.2. The average bulk densities range from 2.35 g/ml for sample A1A to 2.42 g/ml in sample A2A.

Figures 4.6 and 4.7 show the histograms for each of the samples. Among the concrete cores, A1A (Figure 4.6a) exhibits the smallest percentage of aggregates and A2A (Figure 4.6b) exhibits the largest percentage of aggregates. The percentage of aggregates is similar for samples A3A (Figure 4.6c) and A4A (Figure 4.6d). A peak for dolomite and a mixed voxel tail is shown in the C2AV sample histogram (Figure 4.7).

Traditional semivariograms of these samples are shown in Figures 4.8 and 4.9. Figure 4.8 shows pseudocycling in all directions for the concrete cores. Results of modeling the semivariograms with a combination model are shown in Table 4.3. The four concrete

core samples were analyzed in two dimensions with a single scan. Sample A2A has larger horizontal and vertical correlation lengths than the other three concrete core samples. The C2AV sample was analyzed in three dimensions with multiple scans. Further discussion of the geostatistical characteristics of sample C2AV can be found in *Vogel and Brown* [2001].

4.5.3 Conditional Semivariogram Analysis

Conditional semivariogram analysis has been completed on the five samples. Results of the conditional semivariogram analysis are summarized in Table 4.4. Figure 4.10 shows the conditional semivariograms for a concrete sample (A1A). None of the concrete cores show greater than 2.4% voids. The size of detected voids indicates only a few voids per core. Percent aggregates range from 14.4% in A3A to 29.5% in core A2A. The smallest aggregates were detected in core A4A and the largest aggregates were shown in core A2A. The Culebra dolomite core (C2AV) indicated 92.8% solid dolomite and 7.1% mixed voxels. In general the range for the mixed voxels was larger than the range for the solid dolomite. In addition, voids in the Culebra dolomite in the x and y directions were best fit with a Gaussian semivariogram models, while the semivariograms for all other samples were best fit by spherical, exponential, or combination spherical-exponential models.

4.5.4 Traditional and Conditional Semivariogram Comparisons

Correlation lengths from traditional semivariograms can be compared with correlation lengths detected by conditional semivariograms by comparing Tables 4.3 and 4.4. In general, the traditional semivariogram correlation lengths are a weighted average of the conditional semivariogram ranges that is weighted by a combination of the sill value and the percent if points in the bin.

4.5.5 Stochastic Representation

A two-dimensional stochastic representation of the concrete core A1A has been generated. A comparison of the generated representation and the actual scan is shown n Figure 4.11. The representation is generated by placing structure of the size indicated by the conditional semivariogram of voids and aggregates on a background of cement. The generated values of voids, cement, and concrete are normally distributed with mean and standard deviation as shown in Table 4.5. The histogram showing the relative frequency distribution of the stochastic representation and actual CT data is shown in Figure 4.12. Figure 4.13 shows the semivariograms of the stochastic representation and the actual CT data. Relative frequencies are reproduced well using these assumptions for all values except between 2.5 and 2.6 g/ml, where the actual image indicates larger relative frequencies than the generated image. This may be a result of edge pixels between cement and aggregate structure that are not adequately represented in the generated image. The semivariograms shown in Figure 4.13 show similar structure, although the

hole effect in the y direction of the actual scan at approximately 15 mm is not reflected in the semivariogram model and is therefore not found in the stochastic representation.

Real and phantom images are not expected to be identical, but should show similar characteristics. The phantom image histogram closely resembles the histogram of the original CT data as it is used to determine outcrop density. Semivariograms for the real and phantom image should be similar, but not necessarily identical, especially at smaller lags. Larger lags may deviate more than smaller lags because these lags are usually more representative of location than characteristics of the rock.

Successful application of phantom images and semivariograms to identify the major features in the porous media could potentially be used to create a stochastic representation of the media which may be more representative of the true porous media structure for contaminant transport modeling than traditional finite difference modeling. This may be especially true for porous media with small hydraulic conductivities where dispersion and diffusion are important in contaminant transport and plume width.

4.6 Summary and Conclusions

Phantom images have been utilized to develop signature semivariograms of structural elements common to porous media. Knowledge of semivariograms that are associated with specific structure can be used as a tool to identify that structure within the porous media. It should be noted, however, that this is only a tool and two different structures could possess similar semivariograms.

A conditional semivariogram has been introduced to determine structure size in defined bulk density ranges, or bins. This allows for geostatistical characterization of

different rock types and voids within the porous media. As expected, the traditional

semivariogram results are a weighted average of the various conditional semivariograms.

Finally, by combining phantom images, histogram analysis, semivariogram analysis,

and conditional semivariogram analysis, a stochastic representation of porous media has

been generated. These representations have the potential to be used in flow transport

models to determine possible contamination or remediation strategies or could be used as

an initial image for a simulation technique such as simulated annealing.

4.7 References

Anderson, S. H., and J. W. Hopmans (Eds.), *Tomography of Soil-Water-Root Processes, SSSA Spec. Publ. 36*, Am. Soc. of Agron.-Soil Sci. Soc. of Am., Madison, WI, 1994.

Brown, G. O., H. T. Hsieh, and D. A. Lucero, Evaluation of laboratory dolomite core size using representative elementary volume concepts, *Water Resourc. Res.*, 36(5), 1199-1207, 2000.

Clausnitzer, V. and J. W. Hopmans, Determination of phase-volume fractions from tomographic measurements in two-phase systems, *Adv. In Water Resour.*, 22(6), 577-584, 1999.

Deutsch, C.V., and A.G. Journel, *GSLIB: Geostatistical Software Library and User's Guide*, Oxford University Press, 2nd edition, New York, 1998.

Gandin, L. S., *Ob"ektivnyi analiz meterologicheskikh polei*, Gidrometeologicheskoe Izdatel'stvo, Leningrad, Translation (1965), *Objective Analysis of Meteorological Fields*, Israel Program for Scientific Translations, Jerusalem, p. 47, 1963.

Goovaerts, P., *Geostatistics for Natural Resources Evaluation*, Oxford University Press, New York, 1997.

Goovaerts, P., Estimation or simulation of soil properties? An optimization problem with conflicting criteria, *Geoderma*, 97, 165-186, 2000.

Grevers, M. C. J., and E. de Jong, Evaluation of Soil-Pore Continuity Using Geostatistical Analysis on Macroporosity in Serial Sections Obtained by

Computed Tomography Scanning, in *Tomography of Soil-Water-Root Processes, SSSA Special Publication no. 36*, Soil Science Society of America, Madison, WI, pp. 73-86, 1994.

Hahn, G. J., and S. S. Shapiro, *Statistical Models in Engineering*, John Wiley and Sons, Inc., New York, 1967.

Herzfeld, U. C., and O. Zahner, A connectionist-geostatistical approach to automated image classification, applied to the analysis of crevasse patterns in surging ice, *Computers and Geosciences*, 27, 499-512, 2001.

Hsieh, H. -T., G. O. Brown, and M. L. Stone, Measurement of porous media component content and heterogeneity using gamma ray tomography, *Water Resour. Res.*, 34(3), 365-372, 1998.

Holt, R. M., *Conceptual Model of Transport Processes in the Culebra Dolomite Member, Rustler Formation*, Contractor Report, Sandia National Laboratories, Albuquerque, NM, 1997.

Ioannidis, M. A., I. Chatzis, and M. J. Kwiecien, Computer enhanced core analysis for petrophysical properties, *J. Can. Pet. Tech.*, 38(3), 18-24, 1999.

Isaaks, E.H., and R.M. Srivastava, *An Introduction to Applied Geostatistics*, Oxford University Press, New York, 1989.

Jupp, D. L. B., A. H. Strahler, and C. E. Woodcock, Autocorrelation and regularization in digital images II. Simple image models, *IEEE Transactions on Geoscience and Remote Sensing*, 27(3), 247-258, 1989.

Miranda, F. P., J. A. MacDonald, and J. R. Carr, Application of the semivariogram textural classifier (STC) for vegetation discrimination using SIR-B data of Borneo, *Int. J. Remote Sensing*, 13(12), 2349-2354, 1992.

Rasiah, V. and L. A. G. Aylmore, Computed tomography data on soil structural and hydraulic parameters assessed for spatial continuity by semivariogram geostatistics, *Aust. J. Soil Res.*, 36, 485-493, 1998.

Solie, J.B., W.R. Raun, and M.L. Stone, Submeter Spatial Variability of Selected Soil and Bermudagrass Production Variables, *Soil Sci. Soc. Am. J.*, 63(6), 1724-1733, 1999.

Swan, A. R. H., and J. A. Garratt, Image analysis of petrographic textures and fabrics using semivariance, *Mineralogical Mag.*, 59, 189-196, 1995.

Tavchandjian, O., A. Rouleau, and D. Marcotte, Indicator Approach to Characterize Fracture Spatial Distribution in Shear Zones, in *Geostatistics Troia 1992,* A. Soares (ed), Kluwer, Quantitative Geology and Geostatistics Series, 1, pp. 965-976, 1993.

Tercier, P., R. Knight, and H. Jol, A comparison of the correlation structure in GPR images of deltaic and barrier-split depositional environments, *Geophysics,* 65(4), 1142-1153, 2000.

Vogel, J. R., and G. O. Brown, Geostatistics of computerized tomography images of the Culebra Dolomite, *Water Resourc. Res.,* in review, 2001.

Woodbury, A. D., and E. A. Sudicky, The Geostatistical Characteristics of the Borden Aquifer, *Water Resour. Res.,* 27(4), 533-546, 1991.

Table 4.1. Mean and standard deviation of each of the major features of the Culebra dolomite soil core (Holt, 1997).

	Mean (g/ml)	Standard Deviation (g/ml)
Solid Dolomite	2.52	0.07
Gypsum	2.32	0.07
Mixed Voxels (beta distribution)	2.02	0.41
Void	0.0	-

Table 4.2. Sample characteristics.

Sample Name, Type	Sample Dimensions, X x Y x Z, (Voxels)	Voxel Size (mm)	Bulk Density (g/ml)	
			Mean	St. Dev.
A1A, concrete	93 x 44 x 1 (4092)	1.016 x 1.016 x 1.016	2.34	0.20
A2a, concrete	94 x 45 x 1 (4230)	1.016 x 1.016 x 1.016	2.42	0.26
A3A, concrete	94 x 45 x 1 (4230)	1.016 x 1.016 x 1.016	2.35	0.24
A4A, concrete	93 x 44 x 1 (4092)	1.016 x 1.016 x 1.016	2.39	0.23
C2AV, Culebra Dolomite	71 x 71 x 46 (231,886)	0.37 x 0.37 x 0.75	2.44	0.30

Table 4.3. Properties of standardized semivariograms. [1]

Sample Name and Direction		Semivariogram Model	Range (mm)	Sill, w_i	r^2, mm [3]	
A1A	x (nested) x	Combination of Two Spherical	6.5 14.3	0.58 0.45	0.973	to 40.6
	y (nested) y	Combination of Two Spherical	6.2 24.1	0.64 0.41	0.987	to 39.6
A2A	x (nested) x	Combination of Two Spherical	7.2 11.6	0.65 0.44	0.977	to 24.4
	y (nested) y	Combination of Spherical and Exponential	8.3 23.4	0.68 0.53	0.991	to 35.6
A3A	x (nested) x	Combination of Two Spherical	5.6 12.3	0.39 0.58	0.970	to 30.5
	y (nested) y	Combination of Spherical and Exponential	6.6 25.0	0.81 0.17	0.964	to 39.6
A4A	x (nested) x	Combination of Two Spherical	6.1 12.0	0.46 0.52	0.982	to 30.5
	y (nested) y	Combination of Two Spherical	6.1 25.0	0.96 0.10	0.968	to 27.4
C2av	x (nested) x	Combination of Spherical and Exponential	2.3 25.9 [2]	0.83 0.22	0.991	to 13.0
	y (nested) y	Combination of Spherical and Exponential	9.1 2.3	0.17 0.85	0.992	to 15.9
	z (nested) z	Combination of Spherical and Exponential	28.7 [2] 1.6	0.16 0.83	0.993	to 14.8

[1] None of the cores were modeled with a nugget
[2] Range approaching sample width
[3] In this column, the coefficient of determination and the largest lag value used for calculation of the coefficient of determination is reported.

Table 4.4. Properties of conditional semivariograms. [1]

Sample	Bin Range (g/ml), [% pts in range]	Direction	Model [2]	a_1 (mm)	w_1	a_2 (mm)	w_2
A1A	0.0-1.9	x	se	4.6	10.0	17.8	4.02
	[0.6%]	y	se	4.5	8.6	25.0	3.23
	1.9-2.6	x	se	6.2	0.35	25.0	0.57
	[84.3%]	y	se	4.9	0.32	25.0	0.48
	>2.6	x	s	10.9	2.02	-- [3]	--
	[15.1%]	y	s	10.0	1.36	--	--
A2A	0.0-1.9	x	se	6.4	9.9	25.0	7.0
	[2.0%]	y	se	8.1	8.4	12.1	3.3
	1.9-2.6	x	ss	4.9	0.34	16.5	0.51
	[68.5%]	y	ss	5.3	0.26	25.0	0.49
	>2.6	x	s	11.1	0.94	--	--
	[29.5%]	y	s	14.7	1.49	--	--
A3A	0.0-1.9	x	se	5.3	9.1	25.0	5.84
	[2.1%]	y	se	5.8	9.6	25.0	1.58
	1.9-2.6	x	se	8.5	0.18	12.1	0.42
	[83.5%]	y	s	5.8	0.56	--	--
	>2.6	x	s	13.2	1.46	--	--
	[14.4%]	y	s	11.5	1.82	--	--
A4A	0.0-1.9	x	se	5.9	6.1	17.5	0.18
	[2.4%]	y	s	5.5	8.8	--	--
	1.9-2.6	x	ss	4.8	0.32	15.4	0.47
	[75.7%]	y	ss	5.0	0.56	18.1	0.20
	>2.6	x	s	9.7	1.33	--	--
	[21.9%]	y	s	7.6	1.35	--	--
C2AV	0.0-0.5	x	g	2.8	25.2	--	--
	[0.1%]	y	g	2.8	26.3	--	--
		z	e	5.7	21.5	--	--
	0.5-2.00	x	se	2.2	2.38	6.5	1.19
	[7.1%]	y	se	3.4	3.64	14.7	0.44
		z	se	1.4	2.42	8.9	1.23
	>2.0	x	se	1.7	0.59	16.3	0.13
	[92.8%]	y	se	1.8	0.63	9.9	0.12
		z	se	1.4	0.64	20.0	0.11

[1] None of the semivariograms were modeled with a nugget

[2] s=spherical, se=combination spherical and exponential, ss=combination of two spherical, g=gaussian, e=exponential

[3] large scale structure not detected or hole effect

Solid Dolomite

Solid Dolomite with Mean Trend

Solid Dolomite with Vuggy Bedding Planes

▨ 0-0.25	■ 0.25-0.5	□ 0.5-0.75	□ 0.75-1	■ 1-1.25	▨ 1.25-1.5
■ 1.5-1.75	□ 1.75-2	■ 2-2.25	■ 2.25-2.5	▨ 2.5-2.75	▨ 2.75-3
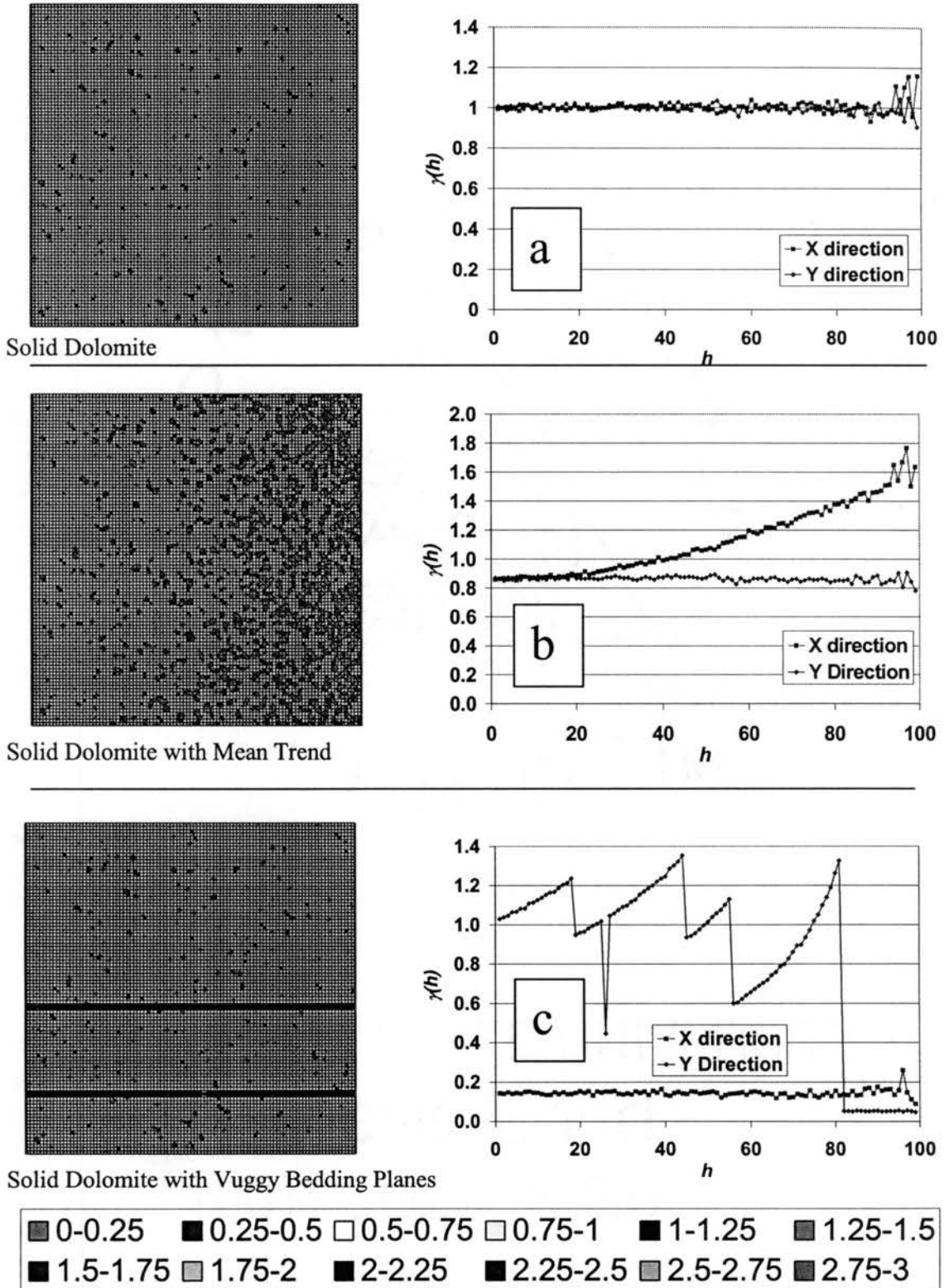
Figure 1. Phantom images and semivariograms for (A) solid dolomite, (B) solid dolomite with a mean trend, and (C) solid dolomite with vuggy bedding planes.

91
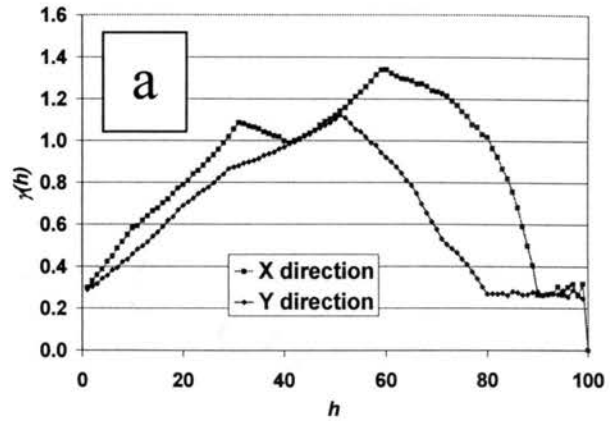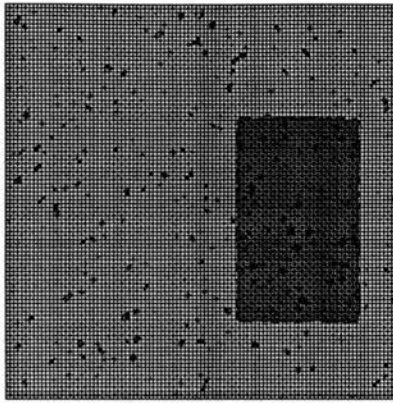
Rectangular Gypsum Outcrop
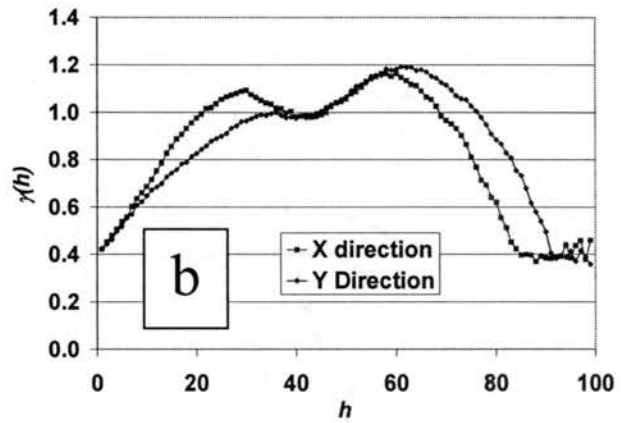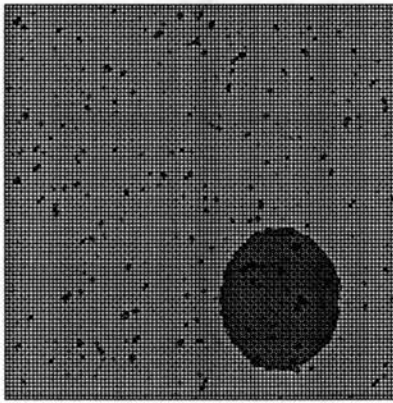
Oval Gypsum Outcrop

One Small Gypsum Outcrop

▨ 0-0.25	■ 0.25-0.5	☐ 0.5-0.75	☐ 0.75-1	■ 1-1.25	▨ 1.25-1.5
■ 1.5-1.75	▨ 1.75-2	■ 2-2.25	■ 2.25-2.5	▨ 2.5-2.75	▨ 2.75-3
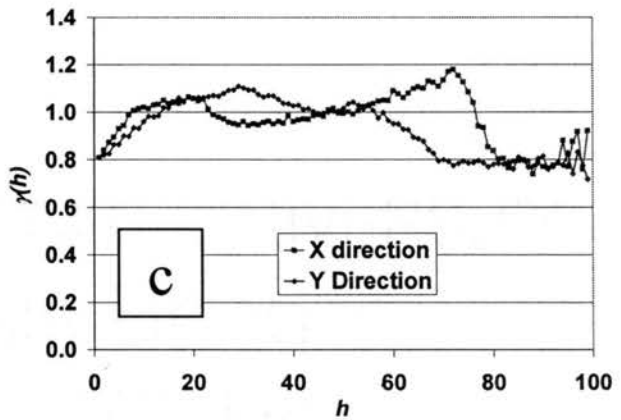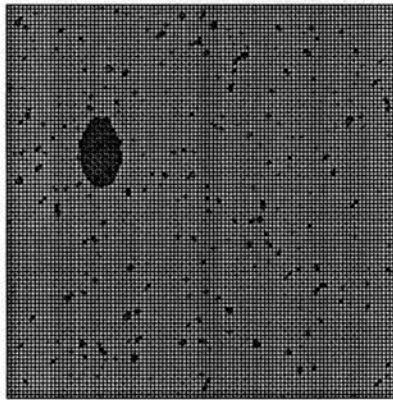
Figure 2. Phantom images and semivariograms for (A) a rectangular gypsum outcrop, (B) an oval gypsum outcrop, and (C) one small oval gypsum outcrop.
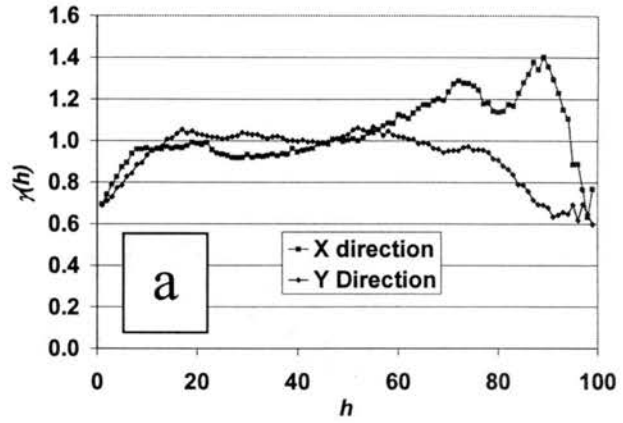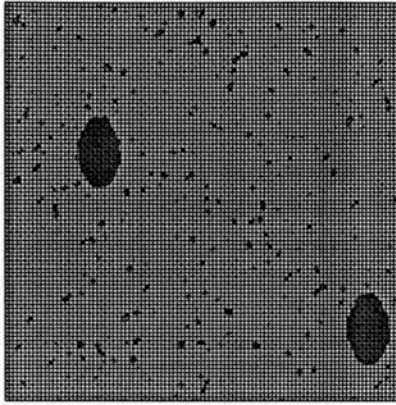
Two Small Oval Gypsum Outcrops
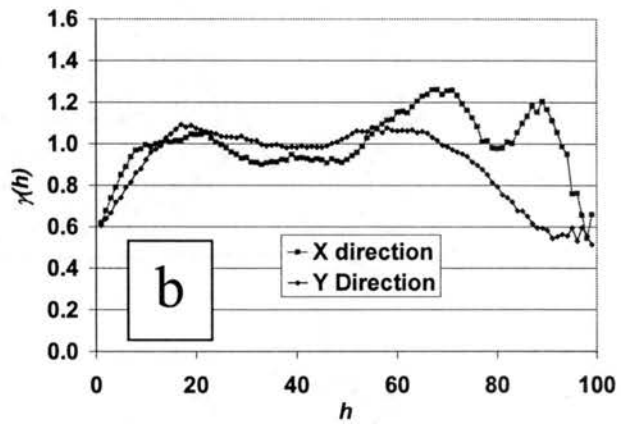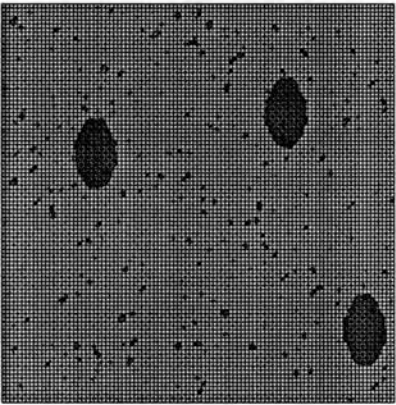
Three Small Oval Gypsum Outcrops

One Small Void

■ 0-0.25　■ 0.25-0.5 □ 0.5-0.75 □ 0.75-1　■ 1-1.25　■ 1.25-1.5
■ 1.5-1.75 □ 1.75-2　■ 2-2.25　■ 2.25-2.5 ■ 2.5-2.75 ■ 2.75-3
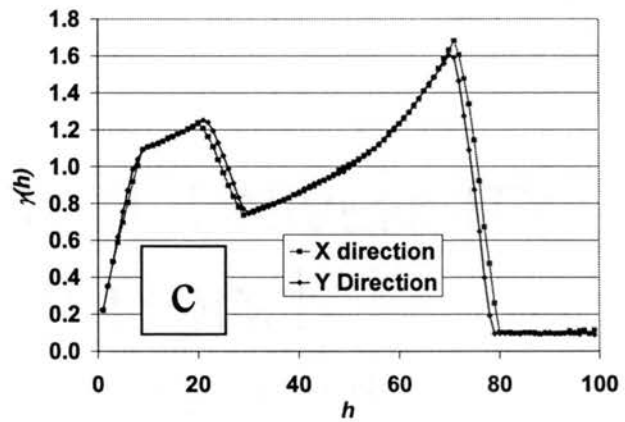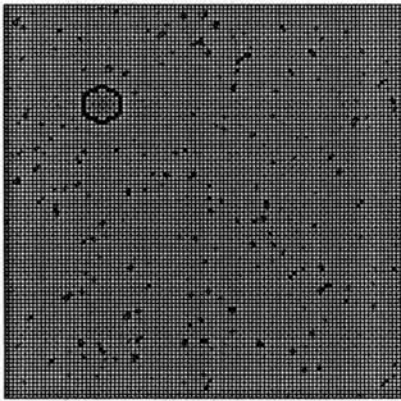
Figure 3. Phantom images and semivariograms for (A) two small gypsum outcrops, (B) three small gypsum outcrops, and (C) one small void.
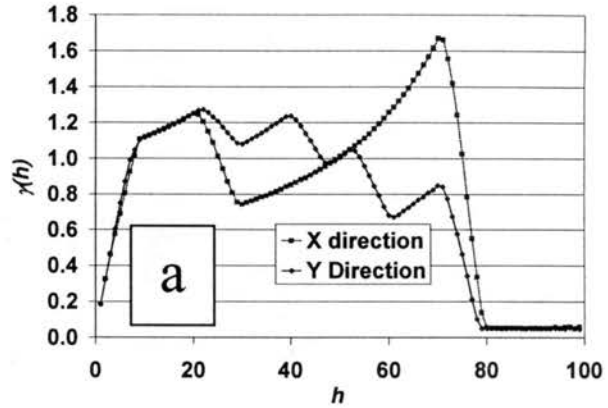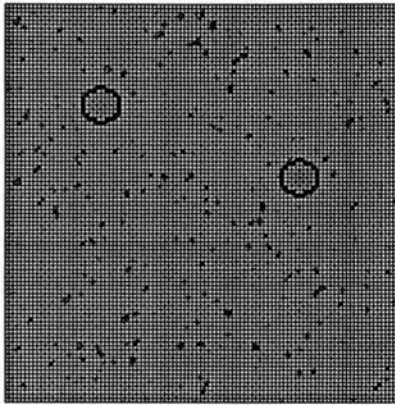
Two Small Voids
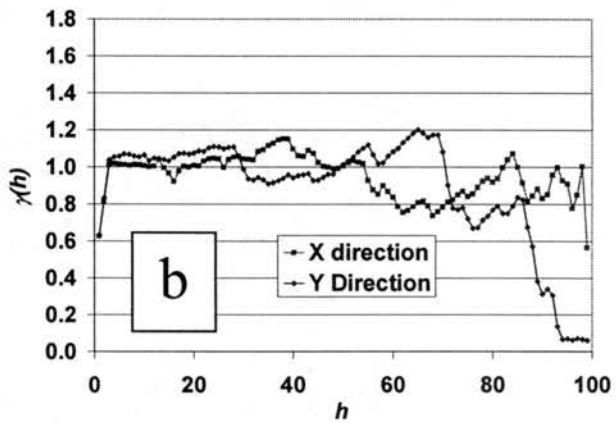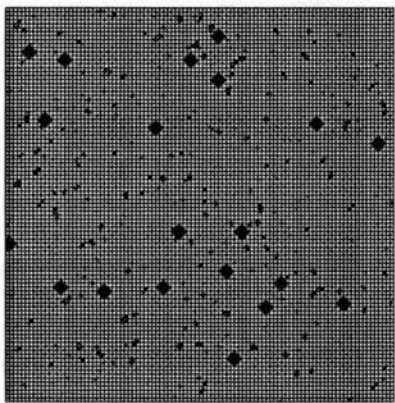
Many Very Small Voids

Twenty Gypsum-Filled Cracks (10 pixels)

| ▨ 0-0.25 | ■ 0.25-0.5 | ☐ 0.5-0.75 | ☐ 0.75-1 | ■ 1-1.25 | ▨ 1.25-1.5 |
| ■ 1.5-1.75 | ▨ 1.75-2 | ■ 2-2.25 | ■ 2.25-2.5 | ▨ 2.5-2.75 | ▨ 2.75-3 |
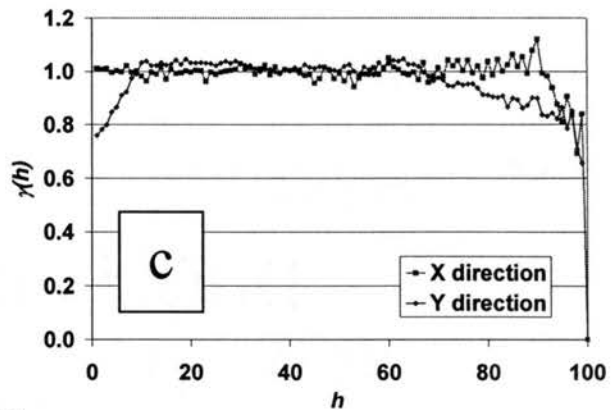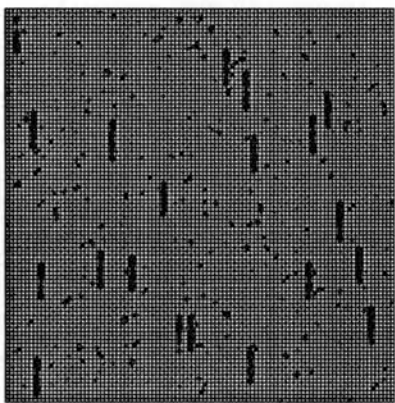
Figure 4. Phantom images and semivariograms for (A) two small voids, (B) many very small voids, and (C) twenty 10-pixel gypsum-filled cracks.
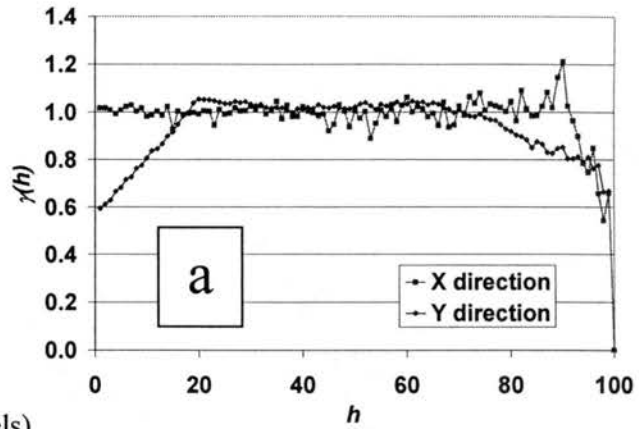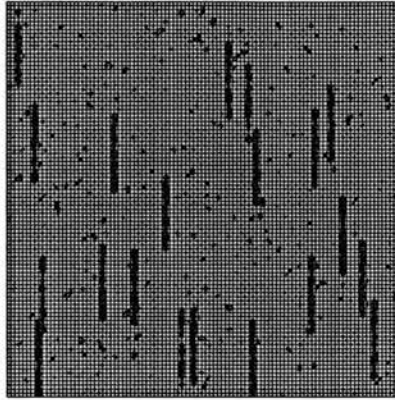
Twenty Gypsum-Filled Cracks (20 pixels)
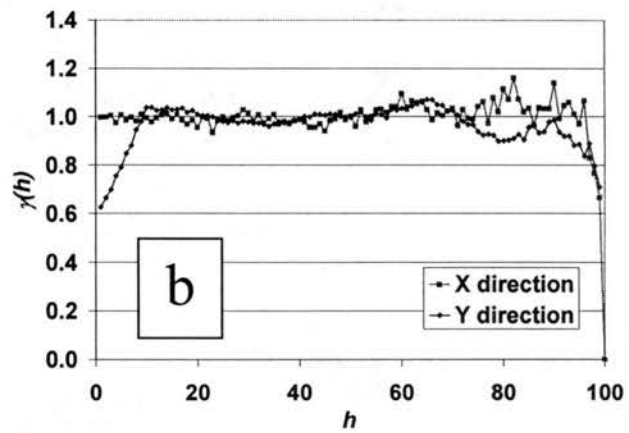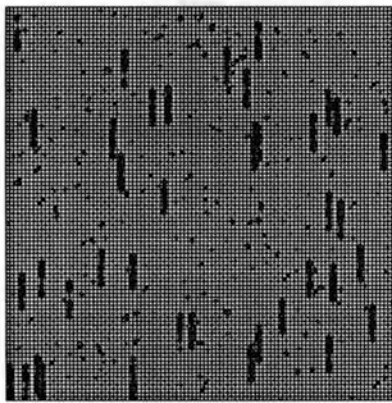
Forty Gypsum-Filled Cracks (10 pixels)
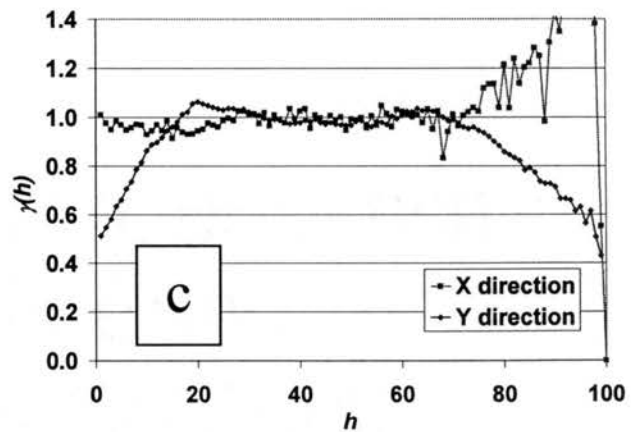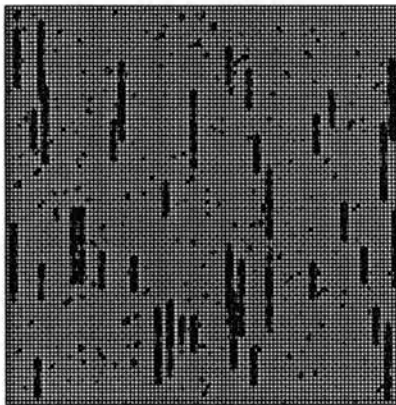
Twenty Gypsum-Filled Cracks (10 pixels) and Twenty Gypsum-Filled Cracks (20 pixels)

| ▨ 0-0.25 | ■ 0.25-0.5 | □ 0.5-0.75 | □ 0.75-1 | ■ 1-1.25 | ▨ 1.25-1.5 |
| ■ 1.5-1.75 | □ 1.75-2 | ■ 2-2.25 | ■ 2.25-2.5 | ▨ 2.5-2.75 | ▨ 2.75-3 |

Figure 5. Phantom images and semivariograms for (A) twenty 20-pixel gypsum-filled cracks, (B) forty 10-pixel gypsum-filled cracks, and (C) twenty 10-pixel gypsum-filled cracks and twenty 20-pixel gypsum-filled cracks.
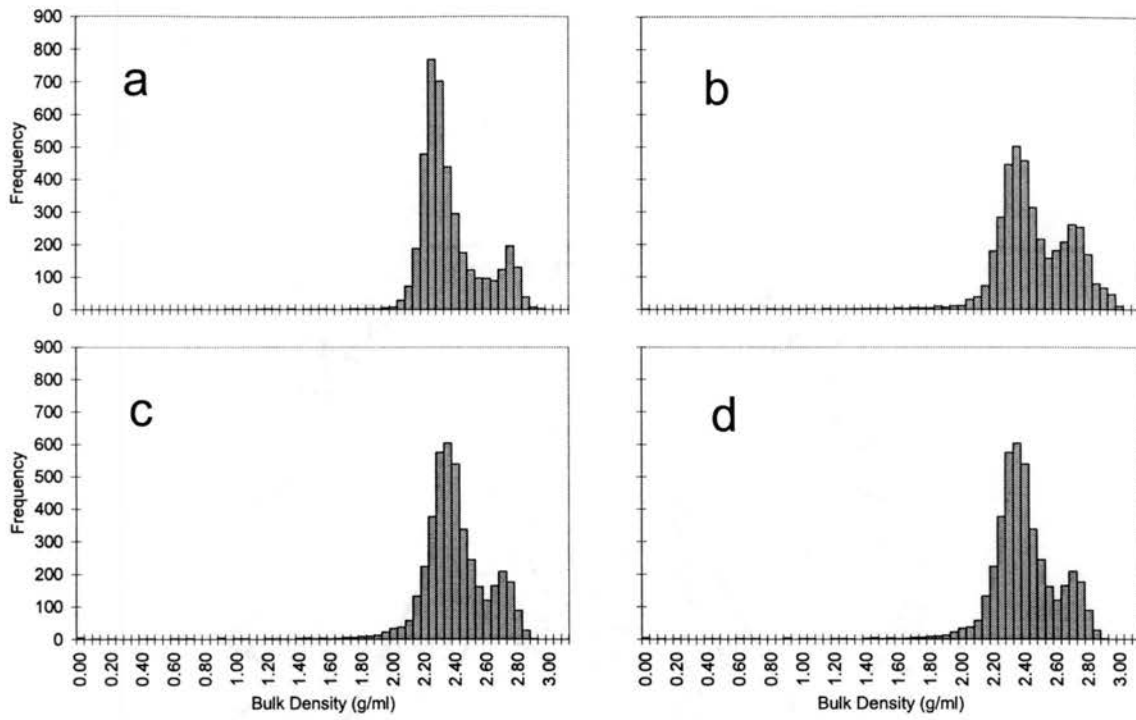
Figure 4.6. Histograms for concrete cores (a) A1A, (b) A2A, (c) A3A, and (d) A4A.
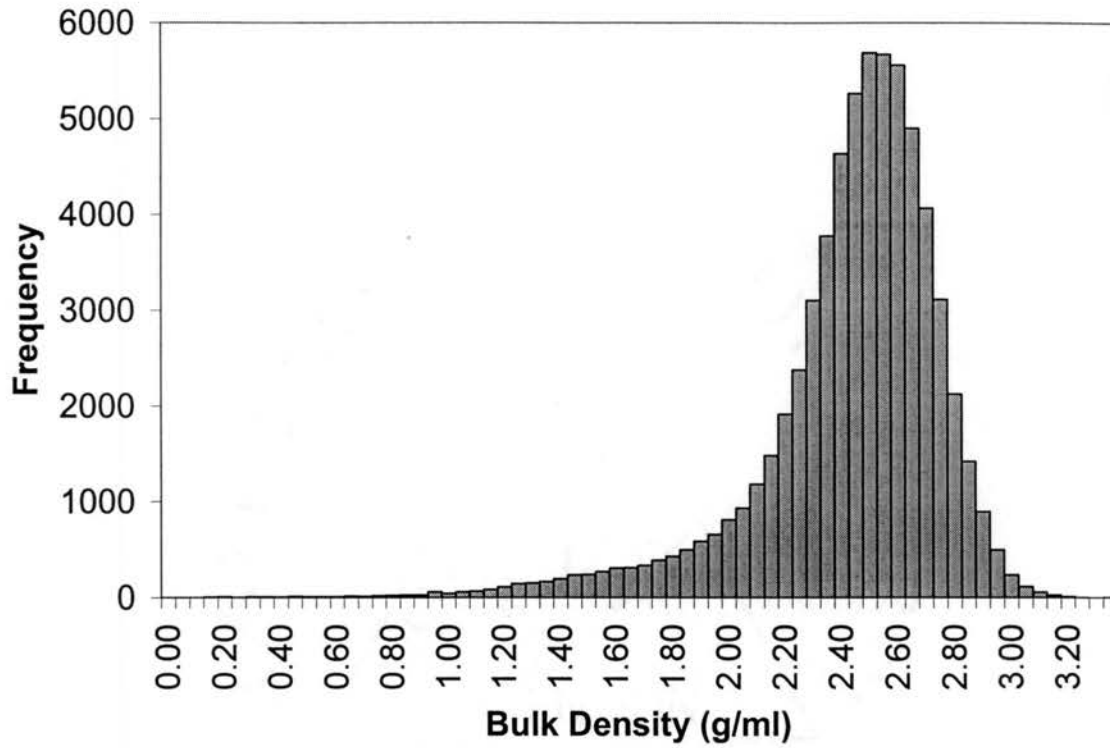
Figure 4.7. Histogram for Culebra dolomite core C2AV.
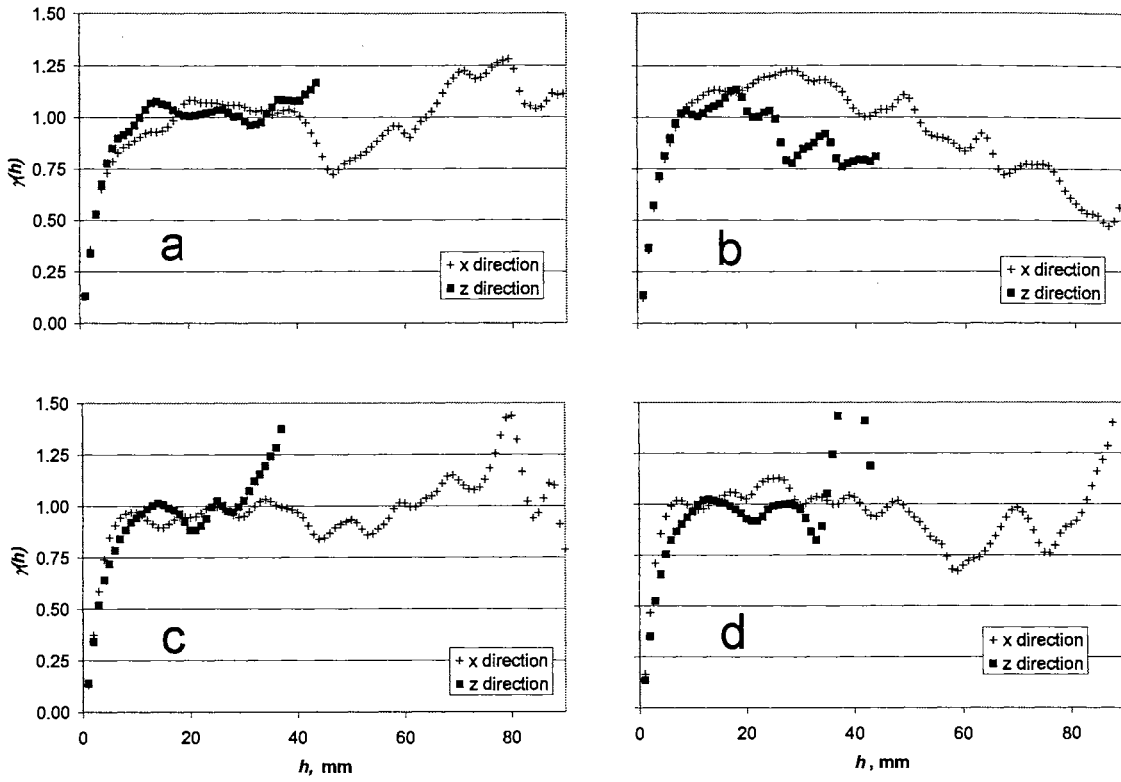
Figure 4.8. Semivariograms of concrete cores (a) A1A, (b) A2A, (c) A3A, and (d) A4A.
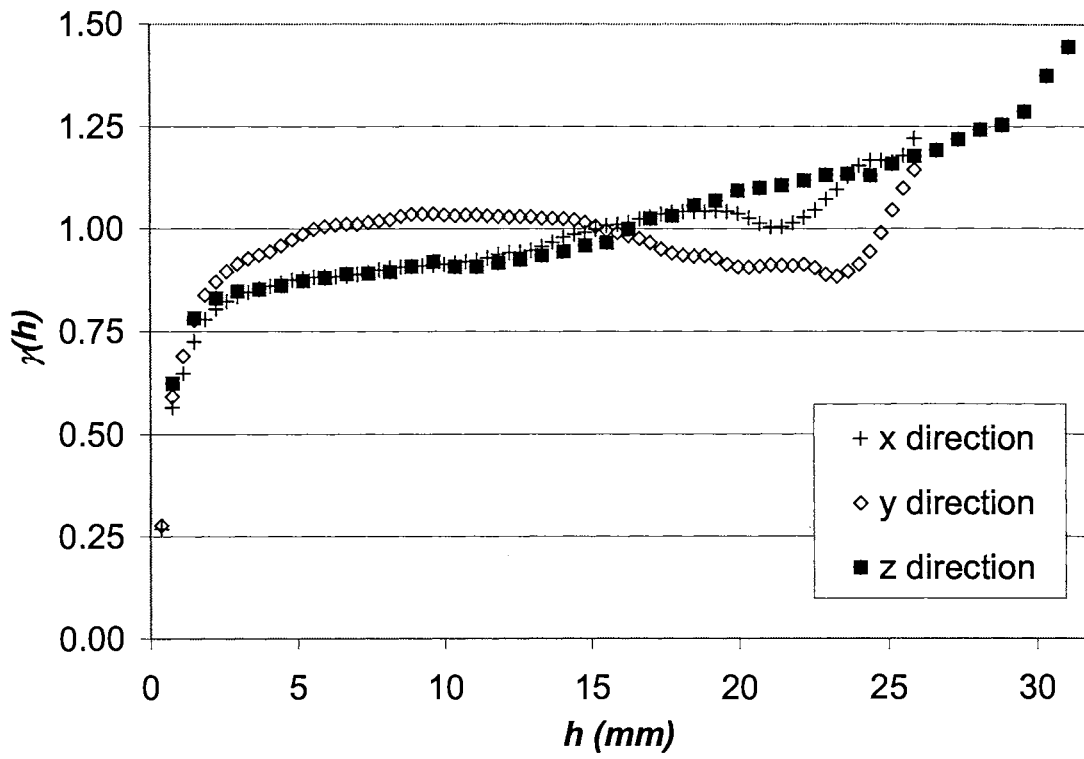
Figure 4.9. Semivariograms of the Culebra dolomite core C2AV.
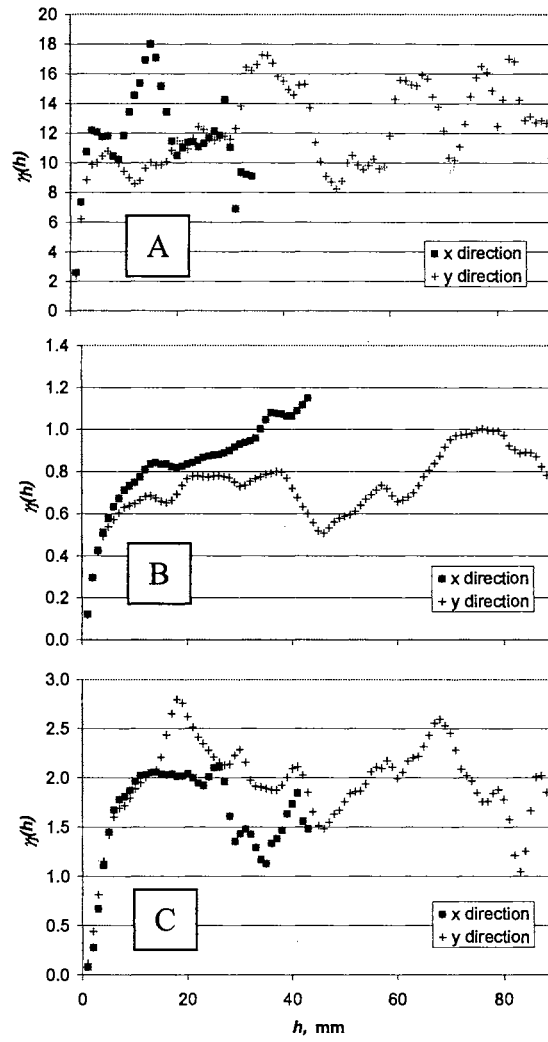
Figure 4.10. Conditional semivariograms for sample A1A for (A) air, (B) cement, and (C) aggregates.

Stochastic Representation　　　　　　　　　　CT Data

■ 2.75-3
▣ 2.5-2.75
■ 2.25-2.5
■ 2-2.25
▨ 1.75-2
■ 1.5-1.75
▨ 1.25-1.5
■ 1-1.25
□ 0.75-1
□ 0.5-0.75
■ 0.25-0.5
▨ 0-0.25
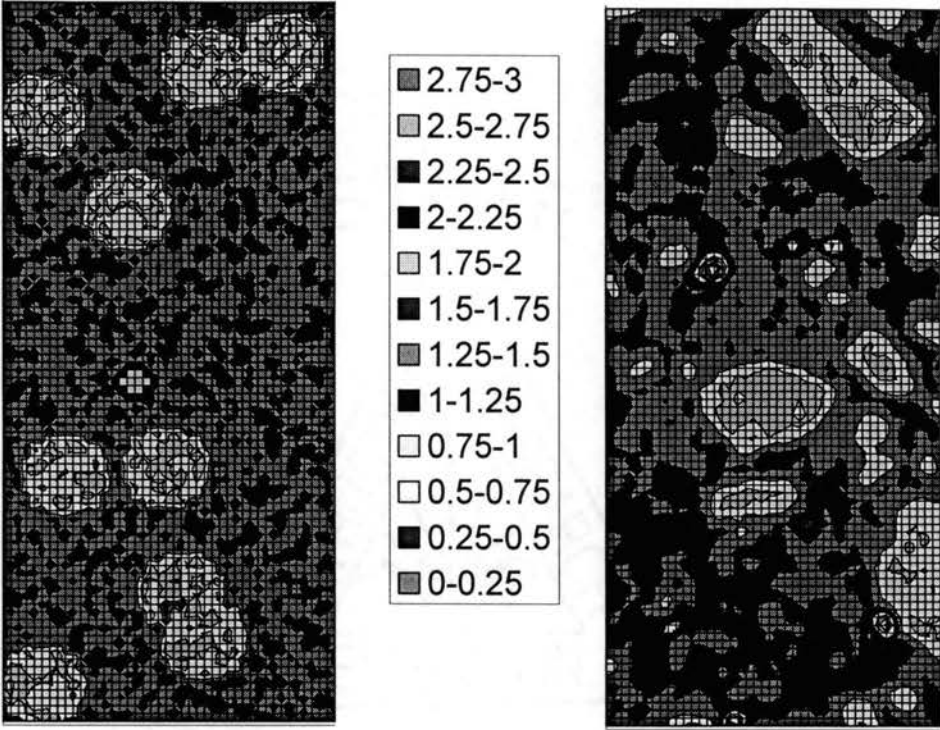
Figure 4.11. Stochastic representation and CT data of concrete core A1A.
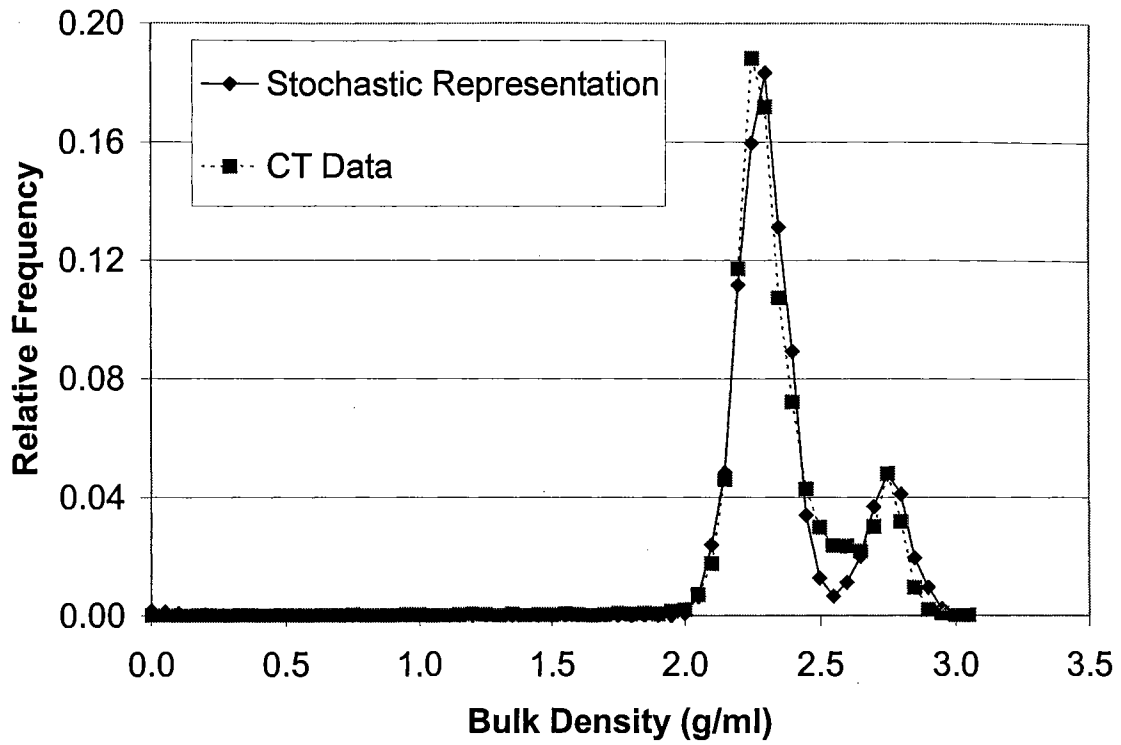
Figure 4.12. Relative histogram of a stochastic representation and the CT data for concrete core A1A.
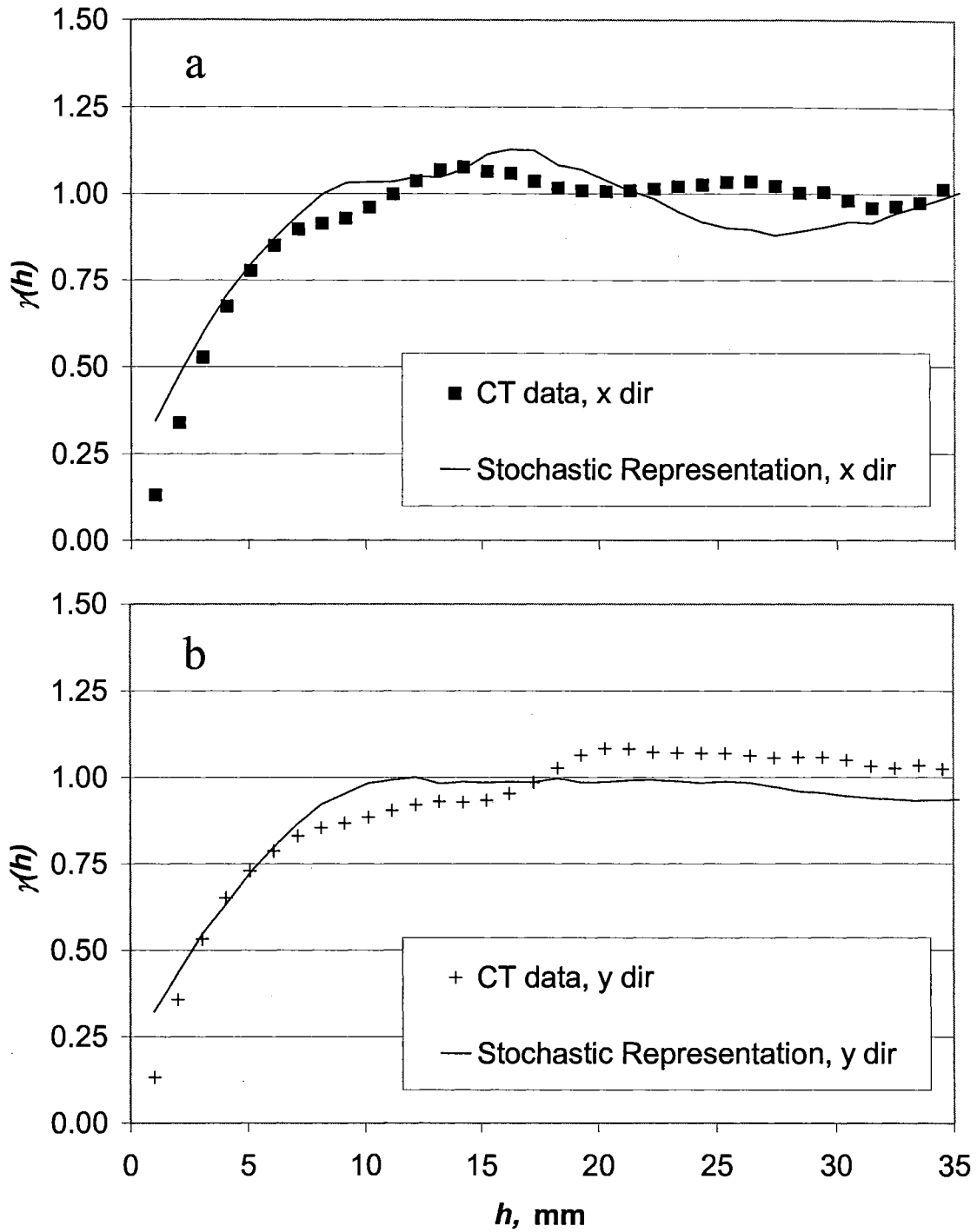
Figure 4.13. Semivariogram in the (A) x and (B) y direction of a stochastic representation and the CT data for concrete core A1A.

APPENDIX A

SEMIVARIOGRAM VALUES

Table A1. Composite semivariogram values for Culebra dolomite samples C2611A (slice 30-90) and C2611A (slice 90-150).

lag	C2611A (slice 30-90)			C2611A (slice 90-150)		
	x	y	z	x	y	z
1	0.1887	0.2123	0.4235	0.1925	0.2206	0.4168
2	0.3610	0.4345	0.5183	0.3749	0.4591	0.5517
3	0.4403	0.5653	0.5874	0.4658	0.6008	0.6284
4	0.4857	0.6574	0.6427	0.5175	0.6936	0.6809
5	0.5193	0.7341	0.6707	0.5554	0.7601	0.7219
6	0.5463	0.7978	0.6978	0.5852	0.8072	0.7549
7	0.5696	0.8511	0.7213	0.6100	0.8466	0.7725
8	0.5893	0.895	0.7429	0.6330	0.8860	0.7913
9	0.6063	0.9292	0.7604	0.6540	0.9202	0.8109
10	0.6212	0.9521	0.7743	0.6731	0.9469	0.8290
11	0.6344	0.9633	0.7853	0.6907	0.9689	0.8448
12	0.6448	0.9658	0.791	0.7071	0.9891	0.8591
13	0.6517	0.9621	0.7961	0.7210	1.0074	0.8743
14	0.6589	0.9517	0.8081	0.7321	1.0238	0.8857
15	0.6687	0.9368	0.8122	0.7428	1.0407	0.8959
16	0.6802	0.9212	0.8185	0.7540	1.0570	0.9033
17	0.6927	0.9098	0.818	0.7635	1.0740	0.9051
18	0.7043	0.9015	0.8254	0.7727	1.0917	0.9114
19	0.7153	0.8959	0.83	0.7826	1.1067	0.9125
20	0.7255	0.8916	0.8359	0.7931	1.1194	0.9172
21	0.7380	0.8844	0.8465	0.8034	1.1295	0.9197
22	0.7508	0.8742	0.8435	0.8131	1.1385	0.9192
23	0.7629	0.8627	0.847	0.8208	1.1449	0.9295
24	0.7729	0.8515	0.8534	0.8272	1.1419	0.9370
25	0.7831	0.8449	0.8736	0.8313	1.1324	0.9450
26	0.7946	0.8413	0.8943	0.8358	1.1199	0.9401
27	0.8050	0.8401	0.9144	0.8384	1.1096	0.9332
28	0.8132	0.8440	0.9248	0.8411	1.1039	0.9295
29	0.8214	0.8553	0.9388	0.8459	1.0988	0.9263
30	0.8324	0.8723	0.9533	0.8485	1.0939	0.9305
31	0.8442	0.8896	0.9697	0.8489	1.0902	0.9357
32	0.8537	0.9115	0.9866	0.8521	1.0868	0.9398
33	0.8608	0.9354	1.0067	0.8566	1.0832	0.9405
34	0.8650	0.9566	1.0194	0.8599	1.0820	0.9284
35	0.8691	0.9759	1.0284	0.8619	1.0794	0.9129
36	0.8715	0.9897	1.0319	0.8651	1.0683	0.9071
37	0.8729	0.9993	1.0564	0.8664	1.0510	0.9049
38	0.8724	1.0029	1.0635	0.8685	1.0296	0.9069
39	0.8746	1.0010	1.0752	0.8713	1.0049	0.9130
40	0.8803	0.9933	1.0635	0.8757	0.9814	0.9190
41	0.8892	0.9808	1.0593	0.8766	0.9573	0.9257
42	0.8956	0.9640	1.0723	0.8777	0.9339	0.9299
43	0.9045	0.9402	1.0806	0.8839	0.9066	0.9297
44	0.9149	0.9158	1.0859	0.8920	0.8786	0.9276
45	0.9276	0.8950	1.0950	0.9028	0.8523	0.9264
46	0.9342	0.8840	1.1004	0.9171	0.8271	0.9198
47	0.9340	0.8837	1.1059	0.9317	0.8064	0.9353
48	0.9343	0.8909	1.1086	0.9472	0.7892	0.9396
49	0.9377	0.9047	1.1131	0.9649	0.778	0.9345
50	0.9404	0.9192	1.1345	0.9897	0.7804	0.9489
51	0.9483	0.9315	1.1790	1.0188	0.7905	0.9673
52	0.9591	0.9447	1.2187	1.0456	0.8046	0.9939
53	0.9676	0.9592	1.2673	1.0653	0.8219	0.9891
54	0.9735	0.9698	1.3308	1.0744	0.8417	1.0131
55	0.9833	0.9800	1.4306	1.0757	0.8614	1.0405
56	0.9939	0.9926	1.5234	1.0712	0.8778	1.0844
57	0.9853	1.0061	1.5456	1.0708	0.8914	1.1337
58	0.9677	1.0139	1.5561	1.0710	0.9030	1.1518
59	0.9299	1.0163	1.5334	1.0574	0.9134	1.2215
60	0.9177	1.0204	1.3142	1.0443	0.9165	1.2414
61		1.0300			0.9171	
62		1.0420			0.9197	
63		1.0479			0.9315	
64		1.0402			0.9517	
65		1.0135			0.9729	
66		0.9723			0.9948	
67		0.9128			1.0188	
68		0.8446			1.0439	
69		0.7796			1.0664	
70		0.7180			1.0858	
71		0.6517			1.0988	
72		0.5880			1.0959	
73		0.5349			1.0535	
74		0.4922			0.9749	
75		0.4705			0.8604	

105

Table A2. Composite semivariogram values for Culebra dolomite samples C1AV and C2AV.

	C1AV			C2AV		
lag	x	y	z	x	y	z
1	0.1958	0.1952	0.5380	0.2684	0.2769	0.6232
2	0.4418	0.4389	0.7790	0.5652	0.5911	0.7824
3	0.5600	0.5531	0.8372	0.6484	0.6905	0.8306
4	0.6659	0.6553	0.8391	0.7247	0.7769	0.8472
5	0.7411	0.7287	0.8386	0.7794	0.8382	0.8521
6	0.7837	0.7724	0.8478	0.8044	0.8705	0.8608
7	0.8132	0.8046	0.8525	0.8228	0.8957	0.8729
8	0.8311	0.8254	0.8613	0.8347	0.9131	0.8806
9	0.8431	0.8421	0.8737	0.8445	0.9268	0.8892
10	0.8515	0.8589	0.8737	0.8536	0.9354	0.8913
11	0.8561	0.8718	0.8724	0.8599	0.9434	0.8946
12	0.8587	0.8819	0.8814	0.8669	0.9581	0.9077
13	0.8596	0.8909	0.8864	0.8747	0.9726	0.9196
14	0.8590	0.9014	0.8860	0.8784	0.9861	0.9071
15	0.8577	0.9119	0.8874	0.8811	0.9988	0.9075
16	0.8563	0.9214	0.8983	0.8833	1.0049	0.9169
17	0.8551	0.9335	0.9052	0.8829	1.0091	0.9254
18	0.8523	0.9448	0.9094	0.8849	1.0106	0.9354
19	0.8484	0.9544	0.9193	0.8886	1.0105	0.9449
20	0.8466	0.9633	0.9306	0.893	1.015	0.9596
21	0.8476	0.9714	0.9371	0.8995	1.0187	0.9661
22	0.8488	0.9805	0.9511	0.904	1.0222	0.9986
23	0.8487	0.9890	0.9775	0.9063	1.0302	1.025
24	0.8487	0.9959	1.0054	0.9074	1.0353	1.0316
25	0.8479	1.0025	1.0283	0.909	1.0357	1.0569
26	0.8468	1.0097	1.0252	0.9107	1.0361	1.0693
27	0.8469	1.0188	1.022	0.9128	1.0338	1.0931
28	0.8476	1.0248	1.039	0.9169	1.0322	1.1008
29	0.8488	1.0241	1.0507	0.9199	1.0336	1.1071
30	0.8483	1.0232	1.0541	0.9225	1.0337	1.1179
31	0.8460	1.0251	1.0679	0.9288	1.0323	1.1315
32	0.8438	1.0272	1.0893	0.937	1.0302	1.1346
33	0.8439	1.0281	1.1056	0.941	1.0288	1.1308
34	0.8469	1.0285	1.1319	0.9425	1.0297	1.1589
35	0.8539	1.0283	1.1509	0.9471	1.0289	1.1775
36	0.8616	1.0304	1.1495	0.9568	1.0258	1.1923
37	0.8693	1.0357	1.1568	0.9673	1.0258	1.2179
38	0.8741	1.0415	1.1709	0.979	1.0246	1.2433
39	0.8770	1.0461	1.2287	0.9867	1.0217	1.2548
40	0.8793	1.0501	1.3048	0.9912	1.0162	1.2865
41	0.8823	1.0558	1.3464	0.9993	1.0055	1.3755
42	0.8847	1.0631	1.3414	1.0075	0.9959	1.445
43	0.8854	1.0715	1.3437	1.0109	0.99	1.6222
44	0.8860	1.0825	1.2598	1.0153	0.9832	1.8445
45	0.8901	1.0948	1.1553	1.0235	0.9762	2.0298
46	0.8976	1.1052		1.0297	0.9662	
47	0.9040	1.1133		1.036	0.9504	
48	0.9127	1.1277		1.04	0.9394	
49	0.9237	1.1472		1.0419	0.9354	
50	0.9319	1.1668		1.0419	0.9329	
51	0.9385	1.1865		1.0416	0.9336	
52	0.9482	1.2078		1.0433	0.9271	
53	0.9594	1.2297		1.0415	0.9126	
54	0.9670	1.2574		1.0368	0.9061	
55	0.9695	1.2884		1.0259	0.9057	
56	0.9727	1.3154		1.0134	0.9083	
57	0.9750	1.3401		1.0039	0.9115	
58	0.9776	1.3722		1.005	0.9108	
59	0.9882	1.4019		1.0148	0.9104	
60	1.0054	1.4295		1.0277	0.9125	
61	1.0197	1.4655		1.0464	0.9049	
62	1.0259	1.5110		1.0717	0.8888	
63	1.0307	1.5548		1.0957	0.8836	
64	1.0341	1.5836		1.1284	0.8959	
65	1.0268	1.5805		1.155	0.9125	
66	1.0206	1.5592		1.1679	0.9438	
67	1.0211	1.5237		1.1677	0.9899	
68	1.0342	1.4796		1.1627	1.0447	
69	1.0762	1.4281		1.1791	1.0987	
70	1.1283	1.3724		1.2209	1.144	

Table A3. Composite semivariogram values for artificial porous media and concrete samples, BOT and A1A, respectively.

lag	BOT x	BOT y	BOT z	A1A x	A1A y
1	0.07058	0.06706	0.1021	0.13044	0.13291
2	0.20672	0.19652	0.24625	0.33866	0.35677
3	0.33294	0.3152	0.34982	0.52775	0.5324
4	0.43366	0.40559	0.4276	0.67451	0.65092
5	0.52345	0.4817	0.49561	0.7769	0.72926
6	0.60758	0.54952	0.54526	0.84991	0.78691
7	0.67855	0.60089	0.57713	0.89777	0.82925
8	0.73489	0.63473	0.5948	0.91376	0.85338
9	0.77738	0.65545	0.60161	0.92912	0.86817
10	0.80663	0.66677	0.60264	0.96072	0.88549
11	0.82548	0.67148	0.60152	0.99908	0.90416
12	0.83615	0.67305	0.60414	1.03682	0.9216
13	0.84199	0.67416	0.61366	1.06896	0.93011
14	0.84688	0.6792	0.62406	1.07732	0.9283
15	0.85180	0.69016	0.63087	1.0648	0.93413
16	0.85505	0.70431	0.6362	1.05887	0.95335
17	0.85746	0.71724	0.64156	1.03581	0.98581
18	0.86081	0.72793	0.64554	1.01688	1.02848
19	0.86351	0.73719	0.64771	1.00893	1.06441
20	0.86629	0.7456	0.64979	1.0063	1.08344
21	0.87347	0.75389	0.65418	1.00922	1.08246
22	0.88547	0.76278	0.661	1.01441	1.07271
23	0.90088	0.77164	0.67037	1.02097	1.07039
24	0.91979	0.78005	0.67931	1.02637	1.07042
25	0.94279	0.78773	0.68394	1.03408	1.06833
26	0.96909	0.79394	0.68709	1.03559	1.0615
27	0.99673	0.79925	0.69162	1.02284	1.05757
28	1.02316	0.80361	0.69887	1.00316	1.05882
29	1.04856	0.80591	0.70847	1.00455	1.05815
30	1.0769	0.80637	0.71753	0.98099	1.04935
31	1.11081	0.80565	0.72734	0.96004	1.03251
32	1.14374	0.80398	0.73756	0.9643	1.02606
33	1.17178	0.80033	0.74634	0.9746	1.03335
34	1.19742	0.79497	0.75188	1.01532	1.02396
35	1.2234	0.78991	0.75283	1.04806	1.01163
36	1.24713	0.78298	0.74863	1.08375	1.02021
37	1.26615	0.77274	0.74352	1.08218	1.03152
38	1.2788	0.76111	0.74178	1.08222	1.03693
39	1.28724	0.7496	0.74382	1.07771	1.02654
40	1.29618	0.74031	0.74918	1.08017	1.00451
41	1.30885	0.7346	0.75698	1.10978	0.97284
42	1.32269	0.73018	0.76699	1.13297	0.92444
43	1.33939	0.72915	0.77878	1.16597	0.87371
44	1.36281	0.73414	0.78704		0.81044
45	1.3948	0.74646	0.79238		0.74741
46	1.43179	0.76208	0.79311		0.72393
47	1.47566	0.78406	0.7912		0.7477
48	1.53261	0.81516	0.78408		0.77564
49	1.59715	0.85331	0.77763		0.78869
50	1.66238	0.8957	0.77961		0.80152
51	1.72416	0.94218	0.79043		0.81596
52	1.77653	0.99172	0.8093		0.83273
53	1.80712	1.04899	0.8347		0.85752
54	1.80988	1.118	0.87132		0.88435
55	1.80146	1.17251	0.91212		0.91025
56	1.80541		0.95689		0.93548
57	1.82027		0.99448		0.95755
58			1.0151		0.95614
59			1.03934		0.92303
60					0.90311
61					0.94133
62					0.97994
63					0.99917
64					1.02719
65					1.06804
66					1.11762
67					1.16099
68					1.18866
69					1.21764
70					1.22575
71					1.20561
72					1.18544
73					1.19277
74					1.21137
75					1.24068

107

Table A4. Composite semivariogram values for concrete samples A2A, A3A, and A4A.

lag	A2A x	A2A y	A3A x	A3A y	A4A x	A4A y
1	0.13582	0.12157	0.13888	0.13099	0.13475	0.16452
2	0.36283	0.35313	0.34127	0.37192	0.34990	0.46567
3	0.57108	0.55926	0.51814	0.58517	0.52374	0.70959
4	0.71303	0.70055	0.6397	0.74054	0.65367	0.85454
5	0.81257	0.79695	0.71831	0.84688	0.75341	0.94060
6	0.89961	0.87734	0.78573	0.90764	0.82246	0.99268
7	0.97112	0.95114	0.84087	0.94279	0.86603	1.01742
8	1.01820	1.01206	0.88346	0.96209	0.90126	1.01555
9	1.02902	1.05022	0.92020	0.96961	0.93341	0.99640
10	1.01276	1.07018	0.94259	0.96982	0.96409	0.98002
11	1.00519	1.08651	0.96368	0.95560	0.99820	0.97507
12	1.0207	1.10575	0.98497	0.93287	1.01957	0.97851
13	1.04033	1.12428	1.00137	0.91337	1.02488	0.99251
14	1.05259	1.13206	1.01352	0.89583	1.01813	1.01034
15	1.06485	1.12997	1.00632	0.89574	1.00695	1.02712
16	1.09323	1.12395	0.99062	0.91330	1.00435	1.04401
17	1.12629	1.12017	0.98157	0.93442	0.99153	1.05961
18	1.13400	1.12049	0.96179	0.94615	0.97420	1.05691
19	1.09649	1.12836	0.92418	0.94724	0.95483	1.03492
20	1.02610	1.14245	0.88349	0.94525	0.92790	1.02594
21	0.99948	1.16207	0.88340	0.94910	0.91864	1.04794
22	1.00142	1.18177	0.90452	0.96541	0.91829	1.08579
23	1.02249	1.19319	0.93873	0.98738	0.93862	1.11013
24	1.03092	1.19751	0.99419	0.99816	0.97429	1.12296
25	0.99143	1.20398	1.02519	0.99604	0.99049	1.12508
26	0.87611	1.21446	1.00234	0.98642	0.99440	1.12704
27	0.78947	1.22154	0.97571	0.97205	0.99799	1.11740
28	0.77738	1.22596	0.97103	0.95768	1.00061	1.07745
29	0.81396	1.22070	0.99628	0.94720	0.99372	1.02198
30	0.84601	1.20084	1.02731	0.95168	0.97505	0.99663
31	0.85707	1.17768	1.07311	0.97219	0.93057	1.00421
32	0.87485	1.17122	1.11997	0.99909	0.86343	1.02494
33	0.90657	1.17917	1.15421	1.02234	0.82191	1.03650
34	0.91856	1.17972	1.19521	1.03266	0.89040	1.03313
35	0.87783	1.16632	1.24229	1.02412	1.05190	1.01833
36	0.79941	1.14735	1.28398	1.00444	1.24400	0.99982
37	0.75942	1.11997	1.37304	0.99433	1.43345	1.00128
38	0.77541	1.08092	1.57193	0.98966	1.60820	1.02550
39	0.78476	1.04216	1.80824	0.98181	1.77834	1.04144
40	0.79243	1.01536	2.04803	0.96832	1.88906	1.03350
41	0.79267	1.00240	2.25914	0.94550	1.74346	1.00718
42	0.78600	1.00353	2.54161	0.90410	1.41081	0.97050
43	0.80818	1.01989	2.66412	0.86162	1.18825	0.94264
44	0.95757	1.03551	2.55629	0.83856		0.93842
45		1.03718		0.84401		0.95546
46		1.04953		0.86736		0.98482
47		1.08341		0.8942		1.01023
48		1.10719		0.91157		1.01785
49		1.09006		0.92362		0.99554
50		1.03382		0.93255		0.96117
51		0.97099		0.91914		0.93534
52		0.93056		0.88708		0.91186
53		0.91099		0.85953		0.87765
54		0.90238		0.86587		0.84036
55		0.89877		0.89066		0.82159
56		0.89082		0.91131		0.80143
57		0.87204		0.93997		0.74135
58		0.84902		0.98709		0.67995
59		0.83697		1.01608		0.67178
60		0.85173		1.01411		0.69881
61		0.89236		0.99737		0.72169
62		0.92319		0.99677		0.73248
63		0.89764		1.01506		0.73997
64		0.82172		1.03832		0.76739
65		0.74897		1.05057		0.80345
66		0.72211		1.07042		0.84339
67		0.72865		1.11084		0.88784
68		0.74780		1.14269		0.93425
69		0.76438		1.14930		0.96990
70		0.77444		1.12558		0.98354
		0.77288		1.10000		0.96125
		0.76980		1.08298		0.92684
		0.77375		1.07943		0.87242
		0.76535		1.09059		0.80796
		0.73751		1.12937		0.76112

APPENDIX B

SAMPLE METADATA

Table B1. Sample Metadata for Culebra dolomite cores.

Sample Name:	C2611a
Image Size (voxels):	120 x 120 x 162
Analyzed pixel dimensions:	x: 30, 90
	y: 30, 90
	z: 30, 150
Sample Name:	C1AV
Image Size (voxels):	120 x 120 x 54
Analyzed pixel dimensions:	x: 25, 95
	y: 25, 95
	z: 5, 50
Sample Name:	C2AV
Image Size (voxels):	120 x 120 x 54
Analyzed pixel dimensions:	x: 24, 94
	y: 25, 95
	z: 5, 50

Table B2. Sample Metadata for concrete cores.

Sample Name:	A1A
Image Size (voxels):	100 x 100
Analyzed pixel dimensions:	x: 40, 83
	y: 16, 108
Sample Name:	A2A
Image Size (voxels):	100 x 100
Analyzed pixel dimensions:	x: 37, 81
	y: 14, 107
Sample Name:	A3A
Image Size (voxels):	100 x 100
Analyzed pixel dimensions:	x: 37, 81
	y: 14, 107
Sample Name:	A4A
Image Size (voxels):	100 x 100
Analyzed pixel dimensions:	x: 38, 81
	y: 15, 107

APPENDIX C

SOURCE CODE FOR CONVAR.FOR

```
      program main
C  This is the edited gam.for source code, now called convar.for. Much of the documentation for gam.for
C  remains
C
C
C%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
C                                              %
C Copyright (C) 1996, The Board of Trustees of the Leland Stanford     %
C Junior University.  All rights reserved.                  %
C                                              %
C The programs in GSLIB are distributed in the hope that they will be  %
C useful, but WITHOUT ANY WARRANTY.  No author or distributor accepts  %
C responsibility to anyone for the consequences of using them or for   %
C whether they serve any particular purpose or work at all, unless he  %
C says so in writing.  Everyone is granted permission to copy, modify  %
C and redistribute the programs in GSLIB, but only under the condition %
C that this notice and the above copyright notice remain intact.    %
C                                              %
C%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%%
c------------------------------------------------------------------------
c
c          Variogram of Data on a Regular Grid
c          *********************************
c
c This is a template driver program for GSLIB's "gam" subroutine.  The
c input data is ordered rowwise (x cycles fastest, then y, then z) in a
c GEOEAS format file.  The User's Guide contains more details.
c
c The program is executed with no command line arguments.  The user
c will be prompted for the name of a parameter file.  The parameter
c file is described in the documentation (see the example gam.par)
c
c
c The output file will contain each directional variogram ordered by
c direction and then variogram (the directions cycle fastest then the
c variogram number).  For each variogram there will be a one line
c description and then "nlag" lines with the following:
c
c      a) lag number (increasing from 1 to nlag)
c      b) separation distance
c      c) the "variogram" value
c      d) the number of pairs for the lag
c      e) the mean of the data contributing to the tail
c      f) the mean of the data contributing to the head
c
c
c
c------------------------------------------------------------------------
      include 'gam.inc'
c
c Read the Parameter File:
c
      call readparm
c
c Call gam to compute the required variograms:
```

```
c
      call gamma
c
c Write Results:
c
      call writeout
c
c Finished:
c
      write(*,9998) VERSION
 9998 format(/' GAM Version: ',f5.3, ' Finished'/)
      stop
      end




      subroutine readparm
c-----------------------------------------------------------------------
c
c              Initialization and Read Parameters
c              ********************************
c
c The input parameters and data are read in from their files. Some quick
c error checking is performed and the statistics of all the variables
c being considered are written to standard output.
c
c
c
c-----------------------------------------------------------------------
      include 'gam.inc'
      parameter(MV=20)
      real     var(MV),cut(MXVARG)
      real*8   avg(MV),ssq(MV)
      real     vrmin(MAXVAR),vrmax(MAXVAR)
      integer  ivar(MV),num(MV),ivc(MXVARG),indflag(MXVARG)
      character datafl*40,str*40
      logical   testfl
      data      lin/1/,ncut/0/
c
c Note VERSION number:
c
      write(*,9999) VERSION
 9999 format(/' GAM Version: ',f5.3/)
c
c Get the name of the parameter file - try the default name if no input:
c
      write(*,*) 'Which parameter file do you want to use?'
      read (*,'(a40)') str
      if(str(1:1).eq.' ')str='gam.par            '
      inquire(file=str,exist=testfl)
      if(.not.testfl) then
           write(*,*) 'ERROR - the parameter file does not exist,'
           write(*,*) '        check for the file and try again  '
           write(*,*)
           if(str(1:20).eq.'gam.par            ') then
                write(*,*) '        creating a blank parameter file'
```

114

```
                call makepar
                write(*,*)
             end if
             stop
          endif
          open(lin,file=str,status='OLD')
c
c Find Start of Parameters:
c
 1      read(lin,'(a4)',end=98) str(1:4)
          if(str(1:4).ne.'STAR') go to 1
c
c Read Input Parameters:
c
          read(lin,'(a40)',err=98) datafl
          call chknam(datafl,40)
          write(*,*) ' data file = ',datafl

          read(lin,*,err=98) nvar
          write(*,*) ' number of variables = ',nvar
          if(nvar.lt.1)    stop 'nvar is too small: check parameters'
          if(nvar.gt.MAXVAR) stop 'nvar is too big: check gam.inc'
          backspace lin

          read(lin,*,err=98) j,(ivar(i),i=1,nvar)
          write(*,*) ' columns = ',(ivar(i),i=1,nvar)

          read(lin,*,err=98) tmin,tmax
          write(*,*) ' trimming limits = ',tmin,tmax

          read(lin,'(a40)',err=98) outfl
          call chknam(outfl,40)
          write(*,*) ' output file = ',outfl

          read(lin,*,err=98) isim
          write(*,*) ' grid number = ',isim

          read(lin,*,err=98) nx,xmn,xsiz
          write(*,*) ' nx,xmn,xsiz = ',nx,xmn,xsiz

          read(lin,*,err=98) ny,ymn,ysiz
          write(*,*) ' ny,ymn,ysiz = ',ny,ymn,ysiz

          read(lin,*,err=98) nz,zmn,zsiz
          write(*,*) ' nz,zmn,zsiz = ',nz,zmn,zsiz

          nxy  = nx * ny
          nxyz = nx * ny * nz
          if(nx.lt.1)      stop 'nx must be at least 1: check parameters'
          if(ny.lt.1)      stop 'ny must be at least 1: check parameters'
          if(nz.lt.1)      stop 'nz must be at least 1: check parameters'
          if(nxyz.gt.MAXXYZ) stop 'model is too big - check gam.inc'

          read(lin,*,err=98) ndir,nlag
          write(*,*) ' ndir,nlag = ',ndir,nlag
          if(ndir.lt.1)    stop 'ndir is too small: check parameters'
```

115

```fortran
      if(ndir.gt.MAXDIR) stop 'ndir is too big: check gam.inc'
      if(nlag.lt.1)      stop 'nlag is too small: check parameters'
      if(nlag.gt.MAXLAG) stop 'nlag is too big: check gam.inc'

      do i=1,ndir
            read(lin,*,err=98) ixd(i),iyd(i),izd(i)
            write(*,*) ' direction = ',ixd(i),iyd(i),izd(i)
      end do
      read(lin,*,err=98) isill
      write(*,*) ' flag to standardize sills = ',isill

      read(lin,*,err=98) nvarg
      write(*,*) ' number of variograms = ',nvarg
      if(nvarg.lt.1)     stop 'nvarg is too small: check parameters'
      if(nvarg.gt.MXVARG) stop 'nvarg is too big: check gam.inc'

      do i=1,nvarg
            read(lin,*,err=98) ivtail(i),ivhead(i),ivtype(i)
            write(*,*) ' tail,head,type = ',
     +              ivtail(i),ivhead(i),ivtype(i)
            if(ivtype(i).eq.9.or.ivtype(i).eq.10) then
                  ncut = ncut + 1
                  if(tmin.gt.0.0)stop'tmin interferes with indicators!'
                  if(tmax.le.1.0)stop'tmax interferes with indicators!'
                  if((nvar+ncut).gt.MAXVAR) then
                        write(*,*) ' Too many indicator cutoffs!'
                        write(*,*) ' use fewer or increase MAXVAR'
                        stop
                  endif
                  backspace lin
                  read(lin,*,err=98) ii,jj,kk,cut(ncut)
                  if(ivtype(i).eq.9)  indflag(ncut) = 1
                  if(ivtype(i).eq.10) indflag(ncut) = 0
                  ivc(ncut) = ivtail(i)
                  ivtail(i) = nvar + ncut
                  ivhead(i) = nvar + ncut
                  write(names(nvar+ncut),140) ncut
 140              format('Indicator ',i2)
                  write(*,*) ' indicator threshold: ',cut(ncut)
            endif
      end do
      write(*,*)
      close(lin)
c
c Check to make sure the data file exists, then either read in the
c data or write an error message and stop:
c
      inquire(file=datafl,exist=testfl)
      if(.not.testfl) then
            write(*,*) 'ERROR data file ',datafl,' does not exist!'
            stop
      endif
c
c The data file exists so open the file and read in the header
c information. Initialize the storage that will be used to summarize
c the data found in the file:
```

```
c
      open(lin,file=datafl,status='OLD')
      read(lin,'(a40)',err=99) str
      read(lin,*,err=99)      nvari
      do i=1,nvari
          read(lin,'(a40)',err=99) str
          do iv=1,nvar
              j=ivar(iv)
              if(i.eq.j) names(iv) = str(1:12)
          end do
          num(i) = 0
          avg(i) = 0.0
          ssq(i) = 0.0
      end do
c
c Read the regular grid information row wise (x cycles fastest):
c
      do is=1,isim
          do iz=1,nz
          do iy=1,ny
          do ix=1,nx
              if(is.ne.isim) then
                  read(lin,'()',err=99)
              else
                  read(lin,*,err=99) (var(i),i=1,nvari)
                  do iv=1,nvar
                      i=ivar(iv)
                      index=ix+(iy-1)*nx+(iz-1)*nxy+(iv-1)*nxyz
                      vr(index) = var(i)
                      if(var(i).ge.tmin.and.var(i).lt.tmax) then
                          num(iv) = num(iv) + 1
                          avg(iv) = avg(iv) + dble(var(i))
                          ssq(iv) = ssq(iv)
     +                            + dble(var(i)*var(i))
                      end if
                  end do
              end if
          end do
          end do
          end do
      end do
      close(lin)
c
c Compute the averages and variances as an error check for the user:
c
      do iv=1,nvar
          sills(iv) = -999.
          if(num(iv).gt.0) then
              avg(iv) = avg(iv) / dble(num(iv))
              ssq(iv) =(ssq(iv) / dble(num(iv))) - avg(iv) * avg(iv)
              sills(iv) = ssq(iv)
              write(*,*) 'Variable number ',iv
              write(*,*) ' Number   = ',num(iv)
              write(*,*) ' Average  = ',real(avg(iv))
              write(*,*) ' Variance = ',real(ssq(iv))
          endif
```

```
          end do
c
c Construct Indicator Variables if necessary:
c
      do ic=1,ncut
          iv = ivc(ic)
          jv = nvar + ic
          p0 = 0.0
          p1 = 0.0
          do ix=1,nx
          do iy=1,ny
          do iz=1,nz
              index = ix+(iy-1)*nx+(iz-1)*nxy+(iv-1)*nxyz
              jndex = ix+(iy-1)*nx+(iz-1)*nxy+(jv-1)*nxyz
              if(vr(index).lt.tmin.or.vr(index).ge.tmax) then
                  vr(jndex) = tmin - EPSLON
              else
                  if(indflag(ic).eq.1) then
                      if(vr(index).lt.cut(ic)) then
                          vr(jndex) = 0.0
                          p0 = p0 + 1.0
                      else
                          vr(jndex) = 1.0
                          p1 = p1 + 1.0
                      end if
                  else
                      vr(jndex) = 0.0
                      if(int(vr(index)+0.5).eq.int(cut(ic)+0.5))
     +                    vr(jndex) = 1.0
                  end if
              end if
          end do
          end do
          end do
          p0 = p0 / max((p1+p0),1.0)
          sills(jv) = dble(p0*(1.0-p0))
      end do
c
c Establish minimums and maximums:
c
      do i=1,MAXVAR
          vrmin(i) =  1.0e21
          vrmax(i) = -1.0e21
      end do
      do ix=1,nx
       do iy=1,ny
        do iz=1,nz
         do iv=1,nvar+ncut
          index = ix+(iy-1)*nx+(iz-1)*nxy+(iv-1)*nxyz
          if(vr(index).ge.tmin.and.vr(index).lt.tmax) then
           if(vr(index).lt.vrmin(iv)) vrmin(iv) = vr(index)
           if(vr(index).gt.vrmax(iv)) vrmax(iv) = vr(index)
          end if
         end do
        end do
       end do
      end do
```

```
      end do
c
c Check on the variogams that were requested:
c
      call check(nvarg,ivtail,ivhead,ivtype,vrmin,vrmax,names,MAXVAR)
c
c Return:
c
      return
c
c Error in an Input File Somewhere:
c
 98   stop 'ERROR in parameter file!'
 99   stop 'ERROR in data file!'
      end




      subroutine gamma
c-----------------------------------------------------------------------
c
c           Variogram of Data on a Regular Grid
c           **********************************
c
c This subroutine computes any of eight different measures of spatial
c continuity for regular spaced 3-D data.  Missing values are allowed
c and the grid need not be cubic.
c
c
c
c INPUT VARIABLES:
c
c nlag        Maximum number of lags to be calculated
c nx          Number of units in x (number of columns)
c ny          Number of units in y (number of lines)
c nz          Number of units in z (number of levels)
c ndir        Number of directions to consider
c ixd(ndir)     X (column) indicator of direction - number of grid
c               columns that must be shifted to move from a node
c               on the grid to the next nearest node on the grid
c               which lies on the directional vector
c iyd(ndir)     Y (line) indicator of direction - similar to ixd,
c               number of grid lines that must be shifted to
c               nearest node which lies on the directional vector
c izd(ndir)     Z (level) indicator of direction - similar to ixd,
c               number of grid levels that must be shifted to
c               nearest node of directional vector
c nv          The number of variables
c vr(nx*ny*nz*nv) Array of data
c tmin,tmax     Trimming limits
c isill       1=attempt to standardize, 0=do not
c sills       the sills (variances) to standardize with
c nvarg       Number of variograms to compute
c ivtail(nvarg)   Variable for the tail of the variogram
c ivhead(nvarg)   Variable for the head of the variogram
c ivtype(nvarg)   Type of variogram to compute:
```

```
c                1. semivariogram
c                2. cross-semivariogram
c                3. covariance
c                4. correlogram
c                5. general relative semivariogram
c                6. pairwise relative semivariogram
c                7. semivariogram of logarithms
c                8. madogram
c                9. indicator semivariogram: an indicator variable
c                   is constructed in the main program.
c
c OUTPUT VARIABLES:  The following arrays are ordered by direction,
c                then variogram, and finally lag, i.e.,
c                iloc = (id-1)*nvarg*nlag+(iv-1)*nlag+il
c
c  np()        Number of pairs
c  gam()         Semivariogram, covariance, correlogram,... value
c  hm()         Mean of the tail data
c  tm()         Mean of the head data
c  hv()         Variance of the tail data
c  tv()         Variance of the head data
c
c
c
c Original:  A.G. Journel              1978
c Revisions: B.E. Buxton              Apr. 1982
c----------------------------------------------------------------------
      include 'gam.inc'
c
c Initialize the summation arrays for each direction, variogram, and lag
c
      nsiz = ndir*nvarg*nlag
      if(nsiz.gt.MXDLV) then
          write(*,*) 'ERROR: available storage in gam = ',MXDLV
          write(*,*) '     requested storage       = ',nsiz
          stop
      endif
      do i=1,nsiz
          np(i)  = 0.
          gam(i) = 0.0
          hm(i)  = 0.0
          tm(i)  = 0.0
          hv(i)  = 0.0
          tv(i)  = 0.0
      end do
c
c First fix the location of a seed point on the grid (ix,iy,iz):
c
      do ix=1,nx
      do iy=1,ny
      do iz=1,nz
c
c For the fixed seed point, loop through all directions:
c
          do id=1,ndir
          ixinc = ixd(id)
```

120

```
            iyinc = iyd(id)
            izinc = izd(id)
            ix1  = ix
            iy1  = iy
            iz1  = iz
c
c For this direction, loop over all the lags:
c
            do il=1,nlag
c
c Check to be sure that the point being considered is still in the
c grid - if not, then finished with this direction:
c
            ix1 = ix1 + ixinc
            if(ix1.lt.1.or.ix1.gt.nx) go to 3
            iy1 = iy1 + iyinc
            if(iy1.lt.1.or.iy1.gt.ny) go to 3
            iz1 = iz1 + izinc
            if(iz1.lt.1.or.iz1.gt.nz) go to 3
c
c For this direction and lag, loop over all variograms:
c
            do iv=1,nvarg
             it = ivtype(iv)
c
c Get the head value, skip this value if missing:
c
            i    = ivhead(iv)
            index = ix+(iy-1)*nx+(iz-1)*nxy+(i-1)*nxyz
            vrt  = vr(index)
            if(vrt.lt.tmin.or.vrt.ge.tmax) go to 5
c
c Get the tail value, skip this value if missing:
c
            i    = ivtail(iv)
            index = ix1+(iy1-1)*nx+(iz1-1)*nxy+(i-1)*nxyz
            vrh  = vr(index)
            if(vrh.lt.tmin.or.vrh.ge.tmax) go to 5
c
c Need increment for the cross semivariogram only:
c
            if(it.eq.2) then
                i    = ivtail(iv)
                index = ix+(iy-1)*nx+(iz-1)*nxy+(i-1)*nxyz
                vrhpr = vr(index)
                if(vrhpr.lt.tmin.or.vrhpr.ge.tmax) go to 5
                i    = ivhead(iv)
                index = ix1+(iy1-1)*nx+(iz1-1)*nxy+(i-1)*nxyz
                vrtpr = vr(index)
                if(vrtpr.lt.tmin.or.vrtpr.ge.tmax) go to 5
            endif
c
c We have an acceptable pair, therefore accumulate all the statistics
c that are required for the variogram:
c
            i    = (id-1)*nvarg*nlag+(iv-1)*nlag+il
```

```fortran
          np(i)  = np(i) + 1.
          tm(i)  = tm(i) + dble(vrt)
          hm(i)  = hm(i) + dble(vrh)
c
c Choose the correct variogram type and keep relevant sums:
c
          if(it.eq.1.or.it.ge.9) then
            gam(i) = gam(i) + dble((vrh-vrt)*(vrh-vrt))
          else if(it.eq.2) then
            gam(i) = gam(i) + dble((vrhpr-vrh)*(vrt-vrtpr))
          else if(abs(it).eq.3) then
            gam(i) = gam(i) +  dble(vrh*vrt)
          else if(it.eq.4) then
            gam(i) = gam(i) +  dble(vrh*vrt)
            hv(i)  = hv(i)  +  dble(vrh*vrh)
            tv(i)  = tv(i)  +  dble(vrt*vrt)
          else if(it.eq.5) then
            gam(i) = gam(i) + dble((vrh-vrt)*(vrh-vrt))
          else if(it.eq.6) then
            if((vrt+vrh).lt.EPSLON) then
               np(i)  = np(i) - 1.
               tm(i)  = tm(i) - dble(vrt)
               hm(i)  = hm(i) - dble(vrh)
            else
               tempvar= 2.0*(vrt-vrh)/(vrt+vrh)
               gam(i) = gam(i) + dble(tempvar*tempvar)
            endif
          else if(it.eq.7) then
            if(vrt.lt.EPSLON.or.vrh.lt.EPSLON) then
               np(i)  = np(i) - 1.
               tm(i)  = tm(i) - dble(vrt)
               hm(i)  = hm(i) - dble(vrh)
            else
               tempvar= alog(vrt)-alog(vrh)
               gam(i) = gam(i) + dble(tempvar*tempvar)
            endif
          else if(it.eq.8) then
            gam(i) = gam(i) + dble(abs(vrt-vrh))
          endif
 5        continue
          end do
 4        continue
          end do
 3        continue
          end do
      end do
      end do
      end do
c
c Get average values for gam, hm, tm, hv, and tv, then compute
c the correct "variogram" measure:
c
      do id=1,ndir
      do iv=1,nvarg
      do il=1,nlag
          i = (id-1)*nvarg*nlag+(iv-1)*nlag+il
```

```fortran
      if(np(i).eq.0.) go to 6
      rnum  = np(i)
      gam(i) = gam(i) / dble(rnum)
      hm(i)  = hm(i)  / dble(rnum)
      tm(i)  = tm(i)  / dble(rnum)
      hv(i)  = hv(i)  / dble(rnum)
      tv(i)  = tv(i)  / dble(rnum)
      it     = ivtype(iv)
c
c Attempt to standardize:
c
      if(isill.eq.1) then
          if(ivtail(iv).eq.ivhead(iv)) then
              iii = ivtail(iv)
              if((it.eq.1.or.it.ge.9).and.sills(iii).gt.0.0)
     +            gam(i) = gam(i) / sills(iii)
          end if
      end if
c
c 1. report the semivariogram rather than variogram
c 2. report the cross-semivariogram rather than variogram
c 3. the covariance requires "centering"
c 4. the correlogram requires centering and normalizing
c 5. general relative requires division by lag mean
c 6. report the semi(pairwise relative variogram)
c 7. report the semi(log variogram)
c 8. report the semi(madogram)
c
      if(it.eq.1.or.it.eq.2) then
          gam(i) = 0.5 * gam(i)
      else if(abs(it).eq.3) then
          gam(i) = gam(i) - hm(i)*tm(i)
          if(it.lt.0) then
              if(sills(ivtail(iv)).lt.0.0.or.
     +           sills(ivhead(iv)).lt.0.0) then
                  gam(i) = -999.0
              else
                  variance = ( sqrt(sills(ivtail(iv)))
     +                   *  sqrt(sills(ivhead(iv))) )
                  gam(i) = variance - gam(i)
              end if
          end if
      else if(it.eq.4) then
          hv(i)  = hv(i)-hm(i)*hm(i)
          if(hv(i).le.0.0) hv(i) = 0.0
          hv(i)  = sqrt(hv(i))
          tv(i)  = tv(i)-tm(i)*tm(i)
          if(tv(i).le.0.0) tv(i) = 0.0
          tv(i)  = sqrt(tv(i))
          if((hv(i)*tv(i)).lt.EPSLON) then
              gam(i) = 0.0
          else
              gam(i) =(gam(i)-hm(i)*tm(i))/(hv(i)*tv(i))
          endif
c
c Square "hv" and "tv" so that we return the variance:
```

```
c
          hv(i)  = hv(i)*hv(i)
          tv(i)  = tv(i)*tv(i)
      else if(it.eq.5) then
          htave  = 0.5*(hm(i)+tm(i))
          htave  = htave  *  htave
          if(htave.lt.EPSLON) then
              gam(i) = 0.0
          else
              gam(i) = gam(i)/dble(htave)
          endif
      else if(it.ge.6) then
          gam(i) = 0.5 * gam(i)
      endif
6    continue
     end do
     end do
     end do
     return
     end




     subroutine writeout
c-----------------------------------------------------------------------
c
c          Write Out the Results of GAM
c          ***************************
c
c An output file will be written which contains each directional
c variogram ordered by direction and then variogram (the directions
c cycle fastest then the variogram number).  For each variogram there
c will be a one line description and then "nlag" lines with:
c
c     a) lag number (increasing from 1 to nlag)
c     b) separation distance
c     c) the "variogram" value
c     d) the number of pairs for the lag
c     e) the mean of the data contributing to the tail
c     f) the mean of the data contributing to the head
c     g) IF the correlogram - variance of tail values
c     h) IF the correlogram - variance of head values
c
c
c
c
c
c-----------------------------------------------------------------------
     include 'gam.inc'
     character title*80
     data     lout/1/
c
c Loop over all the variograms that have been computed:
c
     open(lout,file=outfl,status='UNKNOWN')
     do iv=1,nvarg
```

```fortran
c
c Construct a title that reflects the variogram type and the variables
c that were used to calculate the variogram:
c
      it = abs(ivtype(iv))
      if(it.eq. 1) title(1:24) = 'Semivariogram           '
      if(it.eq. 2) title(1:24) = 'Cross Semivariogram     '
      if(it.eq. 3) title(1:24) = 'Covariance              '
      if(it.eq. 4) title(1:24) = 'Correlogram             '
      if(it.eq. 5) title(1:24) = 'General Relative        '
      if(it.eq. 6) title(1:24) = 'Pairwise Relative       '
      if(it.eq. 7) title(1:24) = 'Variogram of Logarithms '
      if(it.eq. 8) title(1:24) = 'Semimadogram            '
      if(it.eq. 9) title(1:24) = 'Indicator 1/2 Variogram '
      if(it.eq.10) title(1:24) = 'Indicator 1/2 Variogram '
      write(title(25:62),100) names(ivtail(iv)),names(ivhead(iv))
 100  format('tail:',a12,' head:',a12)
c
c Loop over all the directions (note the direction in the title):
c
      do id=1,ndir
         write(title(62:74),101) id
 101     format('direction ',i2)
         write(lout,'(a74)') title(1:74)
c
c Compute the unit lag distance along the directional vector:
c
         dis = sqrt( max(((ixd(id)*xsiz)**2 + (iyd(id)*ysiz)**2 +
     +           (izd(id)*zsiz)**2),0.0) )
c
c Write out all the lags:
c
         do il=1,nlag
            i = (id-1)*nvarg*nlag+(iv-1)*nlag+il
            disl = real(il)*dis
            nump = int(np(i))
            if(it.eq.4) then
               write(lout,102) il,disl,gam(i),nump,
     +                  hm(i),tm(i),hv(i),tv(i)
            else
               write(lout,102) il,disl,gam(i),nump,
     +                  hm(i),tm(i)
            endif
 102        format(1x,i3,1x,f12.3,1x,f12.5,1x,i8,4(1x,f14.5))
         end do
      end do
      end do
      close(lout)
      return
      end




      subroutine check(nvarg,ivtail,ivhead,ivtype,vrmin,vrmax,
     +           names,MAXVAR)
c-----------------------------------------------------------------------
```

```fortran
c
c          Error Check and Note Variogram Types
c          ***********************************
c
c Go through each variogram type and note the type to the screen and
c report any possible errors.
c
c
c
c
c
c-----------------------------------------------------------------------
      real     vrmin(*),vrmax(*)
      integer  ivtail(*),ivhead(*),ivtype(*)
      character title*80,names(MAXVAR)*12
c
c Loop over all the variograms to be computed:
c
      write(*,*)
      do iv=1,nvarg
c
c Note the variogram type and the variables being used:
c
      it = abs(ivtype(iv))
      if(it.eq. 1) title(1:24) = 'Semivariogram         :'
      if(it.eq. 2) title(1:24) = 'Cross Semivariogram    :'
      if(it.eq. 3) title(1:24) = 'Covariance            :'
      if(it.eq. 4) title(1:24) = 'Correlogram           :'
      if(it.eq. 5) title(1:24) = 'General Relative      :'
      if(it.eq. 6) title(1:24) = 'Pairwise Relative     :'
      if(it.eq. 7) title(1:24) = 'Variogram of Logarithms:'
      if(it.eq. 8) title(1:24) = 'Semimadogram          :'
      if(it.eq. 9) title(1:24) = 'Indicator 1/2 Variogram:'
      if(it.eq.10) title(1:24) = 'Indicator 1/2 Variogram:'
      write(title(25:64),100) names(ivtail(iv)),names(ivhead(iv))
 100  format(' tail=',a12,' head=',a12)
      write(*,101) iv,title(1:64)
 101  format(' Variogram ',i2,1x,a64)
c
c Check for possible errors or inconsistencies:
c
      if(it.eq.2) then
          if(ivtail(iv).eq.ivhead(iv)) write(*,201)
 201      format(' WARNING: cross variogram with the same variable!')
      else if(it.eq.5) then
          if(ivtail(iv).ne.ivhead(iv)) write(*,501)
          if(vrmin(ivtail(iv)).lt.0.0.and.vrmax(ivtail(iv)).gt.0.0)
     +         write(*,502)
          if(vrmin(ivhead(iv)).lt.0.0.and.vrmax(ivhead(iv)).gt.0.0)
     +         write(*,502)
 501      format(' WARNING: cross general relative variogram are',
     +         ' difficult to interpret!')
 502      format(' WARNING: there are both positive and negative',
     +         ' values - lag mean could be zero!')
      else if(it.eq.6) then
          if(ivtail(iv).ne.ivhead(iv)) write(*,601)
```

126

```fortran
            if(vrmin(ivtail(iv)).lt.0.0.and.vrmax(ivtail(iv)).gt.0.0)
     +          write(*,602)
            if(vrmin(ivhead(iv)).lt.0.0.and.vrmax(ivhead(iv)).gt.0.0)
     +          write(*,602)
601         format(' WARNING: cross pairwise relative variogram are',
     +             ' difficult to interpret!')
602         format(' WARNING: there are both positive and negative',
     +             ' values - pair means could be zero!')
        else if(it.eq.7) then
            if(ivtail(iv).ne.ivhead(iv)) write(*,701)
            if(vrmin(ivtail(iv)).lt.0.0.or.vrmin(ivhead(iv)).lt.0.0)
     +       write(*,702)
701         format(' WARNING: cross logarithmic variograms may be',
     +             ' difficult to interpret!')
702         format(' WARNING: there are zero or negative',
     +             ' values - logarithm undefined!')
        else if(it.eq.8) then
            if(ivtail(iv).ne.ivhead(iv)) write(*,901)
901         format(' WARNING: cross madograms may be difficult to',
     +             ' interpret!')
        endif
c
c Loop over all variograms:
c
      end do
      return
      end




      subroutine makepar
c-----------------------------------------------------------------------
c
c                 Write a Parameter File
c                 **********************
c
c
c
c
c-----------------------------------------------------------------------
      lun = 99
      open(lun,file='gam.par',status='UNKNOWN')
      write(lun,10)
10    format('               Parameters for GAM',/,
     +       '               ****************',/,/,
     +       'START OF PARAMETERS:')

      write(lun,11)
11    format('../data/true.dat     ',
     +       '-file with data')
      write(lun,12)
12    format('2  1  2           ',
     +       '- number of variables, column numbers')
      write(lun,13)
13    format('-1.0e21    1.0e21   ',
     +       '- trimming limits')
      write(lun,14)
```

```
14  format('gam.out          ',
   +      '-file for variogram output')
    write(lun,15)
15  format('1                 ',
   +      '-grid or realization number')
    write(lun,16)
16  format('50  0.5  1.0      ',
   +      '-nx, xmn, xsiz')
    write(lun,17)
17  format('50  0.5  1.0      ',
   +      '-ny, ymn, ysiz')
    write(lun,18)
18  format(' 1  0.5  1.0      ',
   +      '-nz, zmn, zsiz')
    write(lun,19)
19  format('2 10              ',
   +      '-number of directions, number of lags')
    write(lun,20)
20  format(' 1 .0  0          ',
   +      '-ixd(1),iyd(1),izd(1)')
    write(lun,21)
21  format(' 0  1  0          ',
   +      '-ixd(2),iyd(2),izd(2)')
    write(lun,22)
22  format('1                 ',
   +      '-standardize sill? (0=no, 1=yes)')
    write(lun,23)
23  format('5                 ',
   +      '-number of variograms')
    write(lun,24)
24  format('1  1  1           ',
   +      '-tail variable, head variable, variogram type')
    write(lun,25)
25  format('1  1  3           ',
   +      '-tail variable, head variable, variogram type')
    write(lun,26)
26  format('2  2  1           ',
   +      '-tail variable, head variable, variogram type')
    write(lun,27)
27  format('2  2  3           ',
   +      '-tail variable, head variable, variogram type')
    write(lun,28)
28  format('1  1  9 2.5       ',
   +      '-tail variable, head variable, variogram type')
    write(lun,40)
40  format(//,'type 1 = traditional semivariogram',/,
   +      '     2 = traditional cross semivariogram',/,
   +      '     3 = covariance',/,
   +      '     4 = correlogram',/,
   +      '     5 = general relative semivariogram',/,
   +      '     6 = pairwise relative semivariogram',/,
   +      '     7 = semivariogram of logarithms',/,
   +      '     8 = semimadogram',/,
   +      '     9 = indicator semivariogram - continuous',/,
   +      '    10= indicator semivariogram - categorical')
```

```
      close(lun)
      return
      end


      subroutine chknam(str,len)
c----------------------------------------------------------------------
c
c               Check for a Valid File Name
c               **************************
c
c This subroutine takes the character string "str" of length "len" and
c removes all leading blanks and blanks out all characters after the
c first blank found in the string (leading blanks are removed first).
c
c
c
c----------------------------------------------------------------------
      parameter (MAXLEN=132)
      character str(MAXLEN)*1
c
c Remove leading blanks:
c
      do i=1,len-1
          if(str(i).ne.' ') then
              if(i.eq.1) go to 1
              do j=1,len-i+1
                  k = j + i - 1
                  str(j) = str(k)
              end do
              do j=len,len-i+2,-1
                  str(j) = ' '
              end do
              go to 1
          end if
      end do
 1    continue
c
c Find first blank and blank out the remaining characters:
c
      do i=1,len-1
          if(str(i).eq.' ') then
              do j=i+1,len
                  str(j) = ' '
              end do
              go to 2
          end if
      end do
 2    continue
c
c Return with modified file name:
c
      return
      end
```

129

APPENDIX D

SOURCE CODE FOR CONVAR.INC

C This is the edited gam.inc file, now called convar.inc. It is used with convar.for
C
c---
c
c   Variogram of Data on a Regular Grid
c   ********************************
c
c The following Parameters control static dimensioning within gam3:
c
c MAXXYZ   maximum number of nodes in the 3-D grid
c MAXVAR   maximum number of variables
c
c MAXDIR   maximum number of directions possible at one time
c MAXLAG   maximum number of lags at one time
c MXVARG   maximum number of variograms possible at one time
c
c MXDLV   MAXDIR*MAXLAG*MXVARG - used for dimensioning
c EPSLON   a small number to avoid dividing by zero
c
c---
c
c User Adjustable Parameters:
c
      parameter(MAXXYZ = 500000,
     +   MAXVAR = 1,
     +   MAXDIR = 5,
     +   MAXLAG = 200,
     +   MXVARG = 5)
c
c Fixed Parameters:
c
      parameter(MAXLG=MAXLAG+2,MXDLV=MAXDIR*MAXLG*MXVARG,
     +   EPSLON=1.0e-20,VERSION=2.000)
c
c Variable Declaration:
c
      real   xsiz,ysiz,zsiz,tmin,tmax,vr(MAXXYZ*MAXVAR)
      real*8   gam(MXDLV),hm(MXDLV),tm(MXDLV),hv(MXDLV),tv(MXDLV),
     +   np(MXDLV),sills(MAXVAR)
      integer nlag,nx,ny,nz,nxy,nxyz,ndir,ixd(MAXDIR),iyd(MAXDIR),
     +   izd(MAXDIR),ivtail(MXVARG),ivhead(MXVARG),

131

```
+        ivtype(MXVARG),isill,nvarg
     character outfl*40,names(MAXVAR)*12
c
c Common Blocks:
c
     common /datagv/ xsiz,ysiz,zsiz,tmin,tmax,vr,nlag,nx,ny,nz,nxy,
    +         nxyz,ndir,ixd,iyd,izd,ivtail,ivhead,ivtype,
    +         isill,nvarg

     common /parmdp/ gam,hm,tm,hv,tv,np,sills

     common /chargv/ outfl,names
```

VITA

Jason Robert Vogel

Candidate for the Degree of

Doctor of Philosophy

Thesis: GEOSTATISTICS OF POROUS MEDIA TOMOGRAPHY IMAGES

Major Field: Biosystems Engineering

Biographical:

Education: Graduated from Norfolk High School, Norfolk, Nebraska, in May 1990; received Bachelor of Science degree in Biological Systems Engineering from the University of Nebraska, Lincoln, Nebraska in May 1995; received Master of Science degree in Agricultural Engineering from Texas A&M University, College Station, Texas in August, 1997. Completed the requirements for Doctor of Philosophy degree with a major in Biosystems Engineering in December, 2001.

Experience: Employed by United States Geological Survey, Lincoln, Nebraska, as a technician from June 1994 to January 1995; employed by Texas A&M University, Department of Agricultural Engineering as a graduate research assistant May 1995 to December 1996; employed by Oklahoma State University, Department of Biosystems and Agricultural Engineering as a research engineer, January 1997 to present.

Professional Memberships: American Society of Agricultural Engineers, American Geophysical Union, American Institute of Hydrology, American Society of Civil Engineers, Alpha Epsilon, Gamma Sigma Delta