GEOSTATISTICAL ANALYSIS OF
HYDRAULIC CONDUCTIVITY
IN HETEROGENEOUS AQUIFERS

THESIS

Craig S. Biondo, Capt, USAF

AFIT/GEE/ENP/95D-01

DEPARTMENT OF THE AIR FORCE

AIR UNIVERSITY

AIR FORCE INSTITUTE OF TECHNOLOGY

Wright-Patterson Air Force Base, Ohio

DTIC QUALITY INSPECTED 1

GEOSTATISTICAL ANALYSIS OF
HYDRAULIC CONDUCTIVITY
IN HETEROGENEOUS AQUIFERS

THESIS

Craig S. Biondo, Capt, USAF

AFIT/GEE/ENP/95D-01

19960207 032

GEOSTATISTICAL ANALYSIS OF
HYDRAULIC CONDUCTIVITY
IN HETEROGENEOUS AQUIFERS

Craig S. Biondo

Captain, USAF

Approved:

_____   7 Dec 95
Lt Col Richard Hartley

_____   4 DEC 95
Maj Edward Heyse

_____   4 Dec 95
Maj David Coulliette

_____   4 Dec 95
Prof. Dan Reynolds

DISCLAIMER:

The views expressed in this thesis are those of the author and do not reflect the official policy or position of the Department of Defense or the U.S. government.

AFIT/GEEM/ENG/95D

GEOSTATISTICAL ANALYSIS OF

HYDRAULIC CONDUCTIVITY IN HETEROGENEOUS AQUIFERS

THESIS

Presented to the Faculty of the School of Engineering

of the Air Force Institute of Technology

Air University

in Partial Fulfillment of the Requirements for the

Degree of Master of Science in Engineering and Environmental Management

Craig S. Biondo, B.S.

Captain, USAF

December 1995

Table of Contents

Acknowledgments

List of Figures

List of Tables

Abstract

Observations of the spatial variability of hydraulic conductivity of the heterogeneous alluvial aquifer at Columbus Air Force Base in Mississippi are analyzed using parametric geostatistical approaches. Field studies have revealed that the heterogeneity of the aquifer hydraulic conductivity field controls the movement and dispersion of groundwater solutes (Rehfeldt, et al., 1992). Therefore, a means of quantifying spatial variability is essential for the application of flow and solute transport models to practical problems. Application of these models requires a large number of hydraulic conductivity measurements. Geostatistical analysis and kriging estimation procedures assist in providing these large numbers of values when sampling designs have provided sparse data.

The purpose of this research is to demonstrate a practical methodology for characterizing hydraulic conductivity variability in heterogeneous aquifers. Using kriging estimation procedures, provide estimations of a hydraulic conductivity field for deterministic groundwater flow.

I. Introduction

Overview

The purpose of Chapter I is to motivate the need for characterizing spatial variability of a hydraulic conductivity field with geostatistical tools. In reviewing an example "roadmap" of a typical hydrogeological effort, it is crucial to understand where these geostatistical tools and other estimation methodologies fit into the "big picture" of the modeling process. This enables the scope for the geostatistical estimation sub-process to be defined for this research.

Motivation

A majority of the U.S. population uses groundwater as its primary drinking source (Jensen, 1993). However, groundwater contamination exists in 85% of the Environmental Protection Agency's (EPA) 1208 National Priority List (NPL) sites. In addition to the NPL sites, there are 33,000 contaminated sites in the EPA's Comprehensive Environmental Response, Compensation, and Liability Act Informational System (CERCLIS). Also, there are over 1700 sites identified by the Resource Conservation and Recovery Act (RCRA) requiring cleanup (Olsen and Kavanaugh, 1993). Based on 1990 dollars, the U.S. spends an approximate average of $10 billion

annually on environmental cleanup with total cost of cleanup at $750 billion, taking 75 years (Bredehoeft, 1994).

Because of the growing economic importance of groundwater resources and the health threat posed by contamination, there exists a need to better predict groundwater flow and contaminant transport (Hess et al., 1992). Two areas of focus govern current literature: making better mathematical models and building better input maps to models - the latter being the focus area of this research.

The inability to sufficiently characterize the spatial distribution of hydraulic conductivity (K) is an obstacle to constructing reliable predictive models of groundwater flow and solute transport, to remediating groundwater contamination, and to effectively managing and protecting groundwater resources (Ritzi and Dominic, 1993). Although most researchers generally recognize these difficulties in modeling, they also recognize the reality that heterogeneity of the aquifer decisively influences the flow of groundwater and the movement and dispersion of solutes (Jussel et al., 1994). In a heterogeneous aquifer with hydraulic conductivity values differing by several orders of magnitude, physically connected paths of either high or low hydraulic conductivity values provide the largest influence on groundwater flow as opposed to the shape of a distribution of hydraulic conductivity or the proportion of extreme hydraulic conductivity values. For example, randomly disconnected small fractures (high K, low connectivity) may not generate flow

paths, but a proportion of connected hydraulic conductivity values may wholly condition flow (Journel and Alabert, 1988).

In general, most modeling efforts tend to follow the chain of events in Figure I.1. The true hydrogeological setting (box 1) is a map of the hydraulic conductivity field possessing a certain degree of heterogeneity that accounts for the real spatial distribution of sands, silts, clays, shales, or other soil types in a connected or non-connected fashion. This "map," in simplistic terms, serves as input to nature's transfer function (box 2), which individuals describe using terms such as advection, hydrodynamic dispersion, molecular diffusion, and sorption. With the hydrogeological setting serving as input to these naturally occurring transfer functions, the output is true groundwater flow and/or true contaminant transport (box 3).

Many hydrogeological researchers attempt to account for naturally occurring heterogeneities in groundwater models in order to predict the outcome of the many physical and chemical processes involved in groundwater flow and contaminant transport. Unfortunately, groundwater flow and contaminant transport take place in the inaccessible subsurface, where individuals are not privy to the true hydrogeological map. Therefore, a sampling design and execution is necessary to collect samples in an attempt to describe the actual hydrogeological setting. Sampling, whether for developing hydrogeological input maps (box 4) or for interpreting output from groundwater models for validation purposes (box 10), has the potential to introduce error which may overshadow the actual

spatial variability of the site, or heterogeneity. Sources of error include random

measurement uncertainties, measurement bias, and gross mistakes during sample

collection and laboratory analyses. Hence, sampling is a weak link in the chain of events

necessary to model the environment. Even with an assumption of no measurement error,

it is impossible to sample every point in space of the setting without significantly altering

the characteristics of the site, incurring significant cost, and consuming vast amounts of

time.
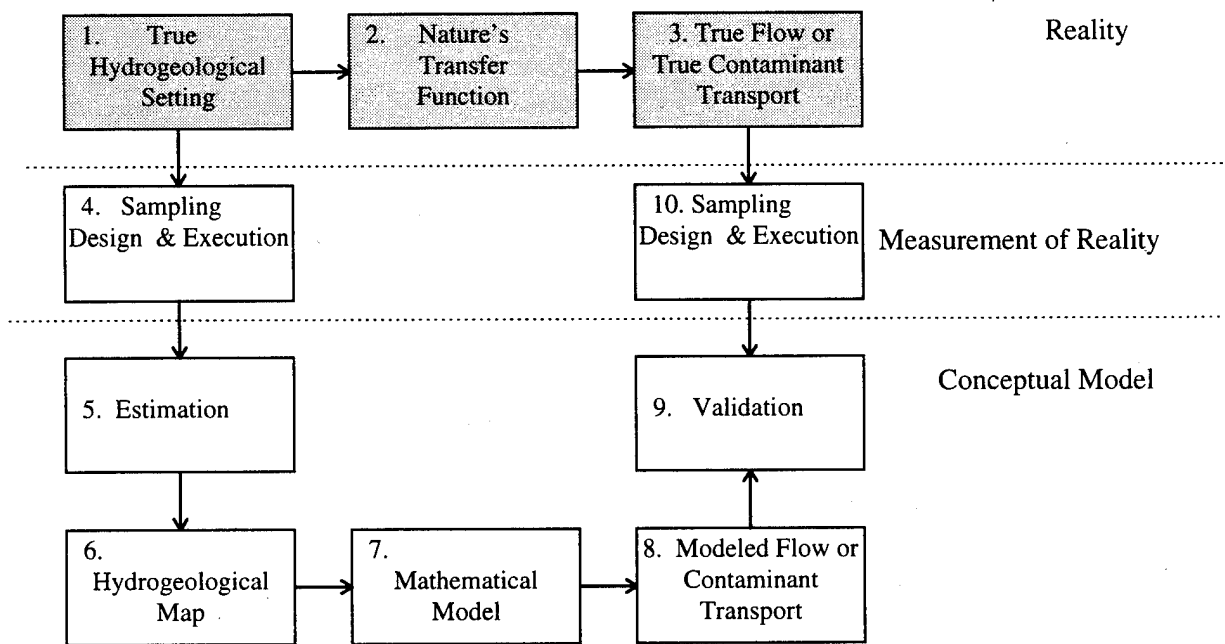
Figure I.1 Generalized chain of events for a groundwater modeling effort.

Estimation is a means of providing values at unsampled locations based on known values

at sampled locations. The correct selection of estimation techniques (box 5) improves the

ability to faithfully reproduce the natural variability and, in turn, provide a better input

map to a groundwater model. Also, because the sampling effort is demanding on a study's time and financial resources, estimation by design provides the most information possible from the samples taken.

Therefore, estimation is a crucial element in modeling the environment and is the focal point for this research. As a demonstration of the importance of selection of the best estimation technique, the Walker Lake data set can be taken as an example of an exhaustive sampling effort that produced true values, assuming no measurement error, at each of 78,000 points on a 260 x 300 rectangular grid (Isaaks and Srivastava, 1989). Desbarats and Srivastava [1991] numerically transformed the elevation variable from the Walker Lake data set into a set of pseudo-transmissivity values for simulation purposes. Figure I.2(a) is a graphical representation of the quartiles of that transformed transmissivity field, where lighter shading denotes higher hydraulic conductivity while darker shading denotes lower hydraulic conductivity. With "truth" known, studying the effect of estimation techniques of varying sophistication is more convenient. In typical sampling programs only a few samples, or a small subset of this exhaustive set, is available for use. With the limited amount of available data, the construction of a hydrogeological map using one of several estimation methodologies is essential. Among the many techniques available for use, three typify general categories: regions of constant hydraulic conductivity, inverse distance method, and kriging, each using increasingly more available information from the samples taken, respectively. A brief overview of these techniques illustrates the general approach to estimation.

The most simplistic level of estimation is constant regions of hydraulic conductivity. In certain situations, a homogeneous hydrogeological setting enables a sufficient description of the entire hydraulic conductivity field from only a few sample values. However, complex heterogeneities found in natural geological settings make it impractical to describe aquifer characteristics in this block fashion. To demonstrate the limited view of reality using this method, the site in Figure I.2(a) was divided into a uniform grid of 25 cells, each 60m x 52m. Taking 25 samples, each from the center of each cell, to represent the hydraulic conductivity field, Figure I.2(b) assigns those sample values with equal weight to every point in space bounded by each cell region, respectively. Hence, this employment of block regions of hydraulic conductivity suggests a possible, but not very realistic, hydrogeological map as input to a mathematical model.
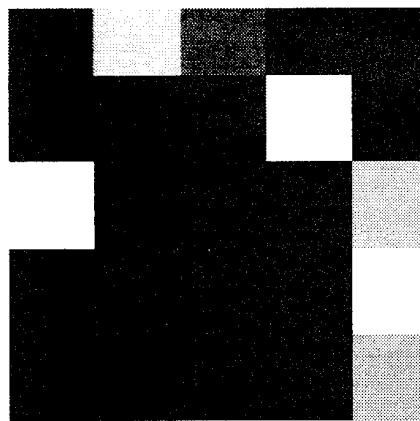
(a)   (b)

Figure I.2 Comparison of "truth" and regions of constant hydraulic conductivity.

Estimates of how the field behaves between sampled locations enhances the realism of

the hydraulic conductivity map. Block regions of constant hydraulic conductivity fail to

provide this information. In contrast, the inverse distance method is an improvement on

naively giving equal weight to all samples by giving more weight to the closest samples

and less to those that are farthest away. One obvious way to do this is to make the weight

for each sample inversely proportional to its distance, or some power of its distance, from

the point being estimated. The inverse distance estimator is:

$$v = \frac{\sum\limits_{i=1}^{n} \frac{1}{d_i^{\ p}} v_i}{\sum\limits_{i=1}^{n} \frac{1}{d_i^{\ p}}}$$

where $d_1, ..., d_n$ are the distances from each of the n known sample locations to the point

being estimated, $v_1, ..., v_n$ are the sample values, and choices for exponent p result in

different estimates (Isaaks and Srivastava, 1989). Therefore, the inverse distance method

is an attempt to estimate values at every point in space based solely on a power of the

distance between samples. Figure I.3 (b) represents the resulting quartiles of an inverse

distance methodology applied to the same 25 samples taken from Figure 1.2(a). Figure

1.3(a) is a representation of the true hydrogeological setting to allow a qualitative

comparison to Figure 1.3(b).
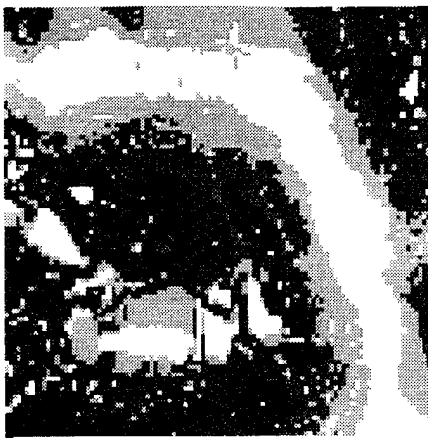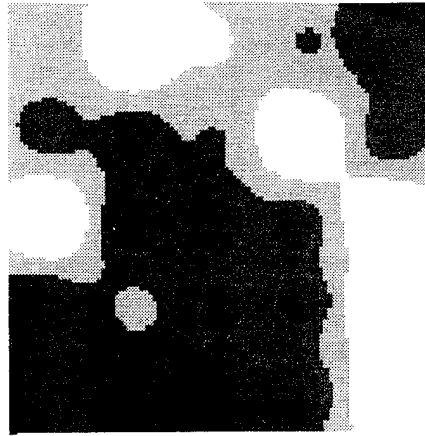
(a)   (b)

Figure I.3 Comparison of "truth" and inverse distance methodology results.

From a qualitative comparison of Figures I.3 (a) and (b), the connectivity and channeling

of areas of higher hydraulic conductivity show some similarity to the crescent shape of

extreme values in the true hydrogeological setting. Obviously, an inverse distance

estimation technique provides more information to a hydrogeological map than simply

using regions of constant hydraulic conductivity. Although this example uses regularly-

gridded data, the biggest drawback of an inverse distance method is its inability to

account directly for irregularly-gridded, closely spaced data. Therefore, an estimation

technique needs to account for this clustering of nearby samples and for their distance to

the point being estimated (Isaaks and Srivastava, 1989). Although some non-

geostatistical declustering techniques such as polygonal declustering and cell declustering

are available, estimation results from these methods force the incorporation of these

I-8

estimation results in constant regions of hydraulic conductivity when constructing the hydrogeological map.

On the other hand, a geostatistical estimation technique such as ordinary kriging meets this need to account for clustering of nearby samples and distance to the point being estimated. Chapter II contains a complete explanation of ordinary kriging. In summary, however, ordinary kriging uses a best linear unbiased estimator to characterize this spatial variability. Ordinary kriging is linear because its estimates are weighted linear combinations of the available data. It is unbiased because it forces the mean residual, or error, to zero. Finally, it is best because it aims to minimize the variance of the errors. Although the other estimation techniques shown are also linear and unbiased, the distinguishing feature of ordinary kriging is its aim to minimize error variance (Isaaks and Srivastava, 1989). Also, because ordinary kriging uses statistical distance rather than Euclidean distance in estimation, it provides more information to a hydrogeological map than an inverse distance estimation technique or regions of constant hydraulic conductivity. Following the same example, Figure I.4(b) is the resulting hydrogeological map of quartiles from an ordinary kriging estimation using the same 25 samples taken from Figure 1.2(a). Figure 1.4(a) is a representation of the true hydrogeological setting to allow a qualitative comparison to Figure 1.4(b).

From a qualitative comparison of Figures I.4 (a) and (b), the connectivity and channeling of areas of higher hydraulic conductivity again shows similarity to the crescent shape of

extreme values in the true hydrogeological setting, but improved over the results of

inverse distance squared estimation method in Figure I.3(b). Notably, the radial patterns

of the inverse distance estimation technique in Figure I.3(b) are no longer artifacts in the

geostatistical estimation technique used in Figure I.4(b).

(a)   (b)
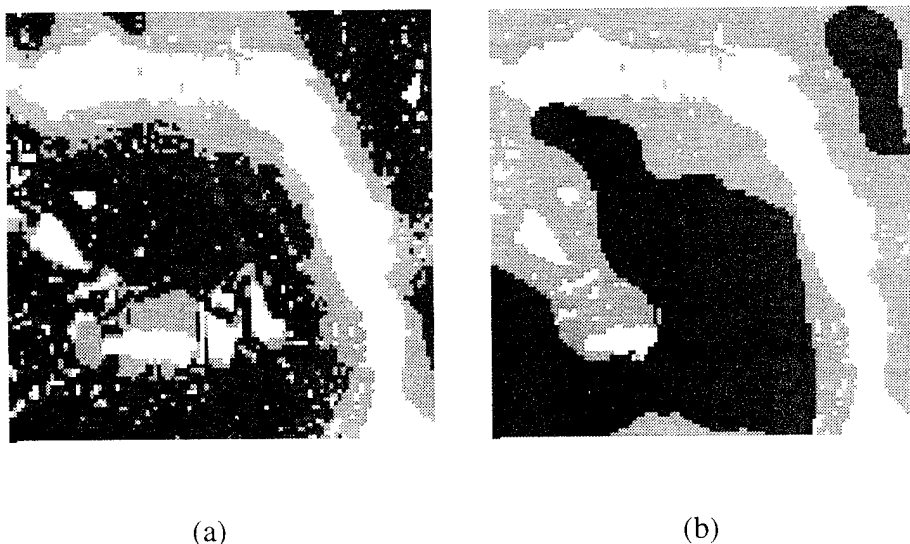
Figure I.4 Comparison of "truth" and ordinary kriging methodology results.

Continuing with the chain of events in Figure I.1, after using an appropriate estimation

technique to form a hydrogeological map (box 6), which is a numerical representation of

nature's hydrogeological setting, this map serves as input to a groundwater flow or

contaminant transport model (box 7). Mathematical models of solute transport in

groundwater are important tools in the assessment of the potential danger for

contamination and for the development of remediation techniques (Jussel et al., 1994).

Many modelers and hydrogeologists constantly search for ways to best conceptualize transfer functions of how groundwater flow and contaminant transport truly occur in nature. Development and improvement of mathematical models comprise much of the research. The output of mathematical models (box 8) indicates how groundwater is flowing or how a contaminant plume is developing.

Because models are merely simplified mathematical representations of nature's complex interacting processes, validation (box 9) of the output of mathematical models is necessary. Examples of validation include sampling of heads to validate modeled groundwater flow or sampling contaminant concentrations to validate modeled contaminant transport (box 3). Therefore, a sampling design and execution (box 10) collects samples in an attempt to validate model output with true groundwater flow or true contaminant transport. However, sampling attempts to measure reality, and sampling by its own nature introduces variability, error, and bias in environmental studies.

Numerous challenges exist in using the sampling - modeling - validating scheme. In an attempt to handle these challenges, numerical simulations are performed, creating a controlled setting where truth is known and measurement uncertainty does not exist. Hence, a comparison can be made to evaluate the worthiness of an investigated process. This has been done on occasion (Desbarats and Srivastava, 1991) and the application of geostatistical tools appears promising, but site specific. When constructing these hydrogeological maps, the need to draw as much information as possible from a costly

sampling effort is necessary. Therefore, because of time limitations, reduced expense, and ability for control, many studies guided towards a greater understanding of flow processes apply geostatistical techniques under the restrictions of hypothetical data sets (Desbarats and Srivastava, 1991).

In contrast to numerical simulations, Chapter II details a few recent experiments conducted at field sites which applied geostatistical techniques under conditions of various naturally occurring degrees of heterogeneities. In fact, the hydraulic conductivity field from the Macrodispersion Experiment (MADE) at Columbus Air Force Base MS is the focus of analysis in Chapter III. In this study although "truth" is not known, a high density data set is available using all of the state-of-the-art tools such as the borehole flowmeter. A methodology is applied to develop realistic estimates of K on a pre-existing grid used for groundwater and contaminant transport modeling on the MADE site by researchers at Tyndall Air Force Base (AFB). Therefore, this research demonstrates a practical estimation methodology using geostatistical techniques on field data collected at the MADE test site.

Scope and Limitations

To demonstrate the ability to quantify spatial variability of a naturally occurring hydraulic conductivity field in a practical problem, the scope of this analysis consists of hydraulic conductivity field characterization based on borehole flowmeter measurements at the

MADE site at Columbus AFB. This research is an analysis of that data using a prescribed approach of well-proven geostatistical tools. Although this research addresses data quality, the author did not design or execute the experiment, and sampling design and execution are not issues in the geostatistical analysis of the MADE data. The inclusion of aquifer characteristics and site geology in site characterization is for completeness; therefore, the geostatistical analysis of hydraulic conductivity does not include any qualitative information derived from studying the site geology. In addition, although there is an appreciation for chemistry, it is generally ignored in the physical groundwater flow and characterization of the hydraulic conductivity field.

II. Background

Overview

The purpose of Chapter II is to provide a methodology for the geostatistical estimation

process along with a theoretical basis for each step in the methodology. A discussion of

field studies reveals the application of a similar methodology to hydrogeological

experiments. Although other field studies are discussed, particular attention is given to

the MADE experiment at Columbus AFB. Specifically, the approach to calculating

hydraulic conductivity values, which are analyzed in Chapter III, is addressed.

Geostatistics

Geostatistics and its applications to groundwater problems are areas of active research.

Hydrogeologists rely heavily on site-specific data to describe geometry and properties of

formations of interest. A typical set of groundwater data, such as measured hydraulic

conductivity, consists of observations at a number of irregularly spaced points. Given

such a set, a hydrogeologist is often faced with estimating the hydraulic conductivity

value at an unmeasured point in the same aquifer. Because geostatistics deals with

making inferences with "incomplete" information, it relies on probability theory, and it is

the methodology to resort to when lacking sufficient data to make deterministic, or error-

free, inferences. In other words, there is an inability to predict with certainty the value of

head or hydraulic conductivity at a well, even if these quantities are measured at nearby wells. Geostatistics offers a systematic approach to making inferences about quantities that vary in space such as hydraulic conductivity (ASCE, 1990). Appendix A contains terminology commonly found in geostatistics.

Armstrong [1984] suggested a logical approach to geostatistics, and subsequent sections of this thesis make use of these fundamental steps. The steps are:

1. Data. Identify the distribution type of the random variable and its outliers using univariate statistics and looking for trends.

2. Experimental Variogram. Calculate the experimental variogram and choose an appropriate estimator.

3. Variogram Model. Seek out theoretical parameters from the empirical variogram. Important decisions are the type of theoretical model, the method of parametric estimation, and the range of empirical variogram used.

4. Application. The final step is application of material established from steps 1-3. For example, one application is kriging. (Woodbury and Sudicky, 1991)

(1) Univariate analysis of the data

Many geostatistical estimation tools work more favorably if the distribution of data is close to the Gaussian or normal distribution. The Gaussian distribution is one of many distributions for which a concise mathematical description exists. Also, it has properties

that favor its use in theoretical approaches to estimation. Therefore, it is important to know how close the distribution of the data values come to being Gaussian (Isaaks and Srivastava, 1989).

Before determining a univariate description of the data, the random variable of interest must be defined. It is nearly impossible to describe a regionalized variable (in this case hydraulic conductivity) with precision. Therefore,

$$K(\mathbf{x}) = m(\mathbf{x}) + \xi(\mathbf{x})$$

where $K(\mathbf{x})$ is a random variable representing $k(\mathbf{x})$, the hydraulic conductivity at $\mathbf{x}$; $m(\mathbf{x})$ is the mean value of $K(\mathbf{x})$; $\xi(\mathbf{x})$ is a fluctuation value at $\mathbf{x}$, or residual (ASCE, 1990).

One of the most common and useful presentations of a data set is a histogram. Also, summary statistics often capture the important features of most histograms. Taken together, statistics such as the mean, median, variance, standard deviation and coefficient of skewness provide valuable information on the random variable.

Theoretically, let $K(\mathbf{x})$ and $K(\mathbf{x} + \mathbf{h})$ be two components of a random field representing hydraulic conductivity at locations $\mathbf{x}$ and $\mathbf{x} + \mathbf{h}$, respectively, where $\mathbf{h}$ denotes the distance vector. $K(\mathbf{x})$ and $K(\mathbf{x} + \mathbf{h})$ can assume any possible value, each with a given probability. Also, the expected value of $K(\mathbf{x})$ and $K(\mathbf{x} + \mathbf{h})$ are $m(\mathbf{x})$ and $m(\mathbf{x} + \mathbf{h})$, respectively. Therefore, the expected value of the squared difference of $K(\mathbf{x})$ from its expected value

$m(\mathbf{x})$ is the variance of $K(\mathbf{x})$.

$$Var[K(\mathbf{x})] = E[K(\mathbf{x}) - m(\mathbf{x})]^2$$

The variance of $K(\mathbf{x} + \mathbf{h})$ is similarly defined. Hence, the variance is a measure of spread of a random variable about its mean. (ASCE, 1990)

The covariance between $K(\mathbf{x})$ and $K(\mathbf{x} + \mathbf{h})$ is a measure of the mutual variability between them, defined as

$$C(\mathbf{x}, \mathbf{x} + \mathbf{h}) = E\{[K(\mathbf{x}) - m(\mathbf{x})][K(\mathbf{x} + \mathbf{h}) - m(\mathbf{x} + \mathbf{h})]\}$$

The mean and the covariance are important measures of central tendency and spread of a random field. However, to estimate them based on observed values and to use them in an ordinary kriging estimation process, an assumption of stationarity of order two is required. The stationarity of order two assumption implies (1) the mean $m(\mathbf{x})$ is the same everywhere ; and (2) the covariance exists and is a unique function of the separation distance. Therefore, there is no trend, the differences in values of $K(\mathbf{x})$ are only due to the random component $\xi(\mathbf{x})$, and $C(\mathbf{x}, \mathbf{x} + \mathbf{h})$ is simply $C(|\mathbf{h}|)$. (ASCE, 1990)

For ordinary kriging it is important to determine if the data to be analyzed is stationary. If the mean tends to follow some spatial trend causing the random function to become nonstationary, there is probable cause for trend removal. Although estimation procedures such as universal kriging automatically calculate a trend, it is wise to check the

calculation of this automatic method to see if it makes sense. In many situations, it is often wiser to select a trend based on available information, subtract this trend from observed values, perform a geostatistical analysis of the residuals, and add the trend back at the conclusion of estimation. (Isaaks and Srivastava, 1989)

To illustrate the concepts of trend and stationarity, it is useful to consider the following example. The curves in Figure II.1 represent hydraulic conductivity versus distance along a certain path. If samples are taken every 5 feet, the results for curves A and B will show the average difference in hydraulic conductivity as position changes 5 feet in either direction is zero. This is equivalent to saying $m(x)$ is a constant for curves A and B. For curve C, $m(x)$ increases as position changes in the positive direction along the path. Curve A and curve B have no trend, but curve C exhibits a linear trend in the mean.

Curve A also exhibits second order stationarity. The variance in the difference between sample values 5 feet apart is the same throughout the curve. From inspection of curve B, the variance of the difference in values 5 feet apart is greater as position moves in a positive direction. Therefore, second order stationarity does not exist for curve B. Of course, second order stationarity does not exist for curve C because of its linear trend.
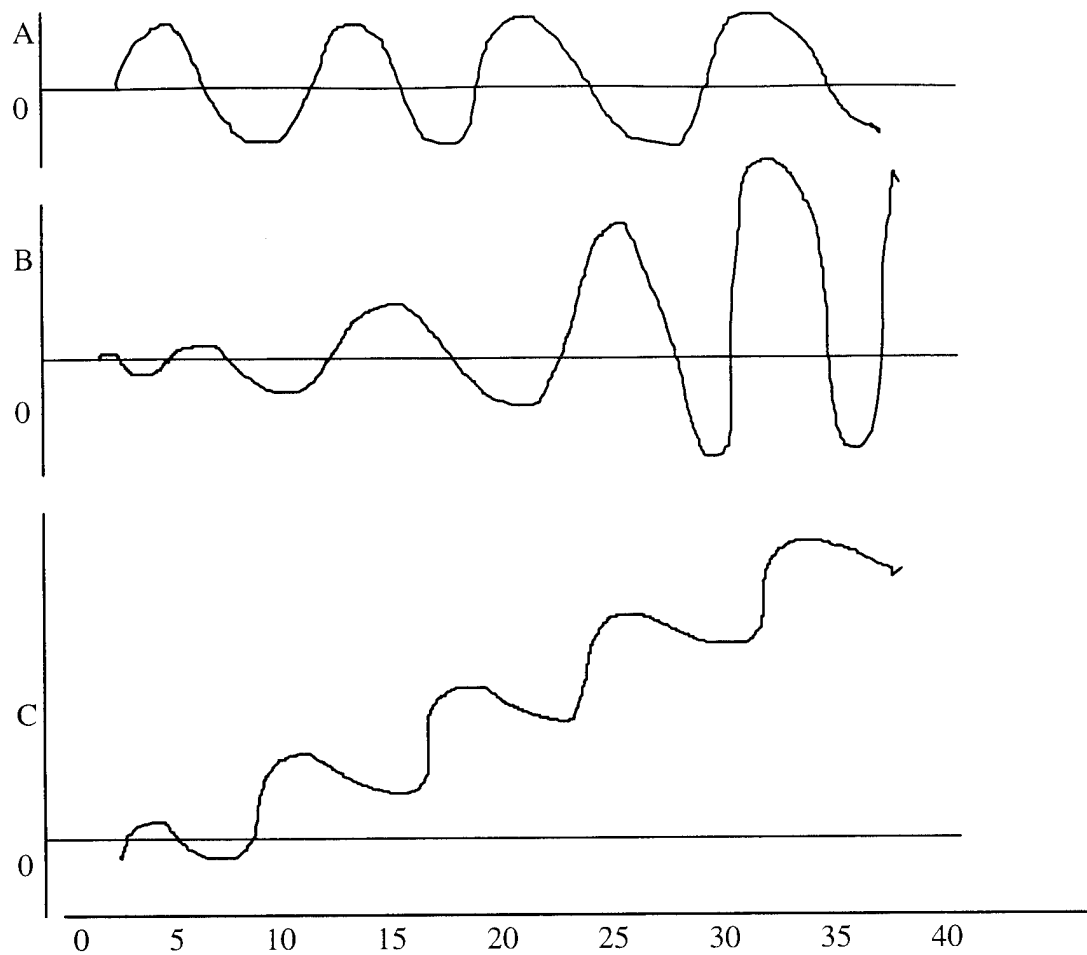
Figure II.1 Illustrative example of stationarity and trend.

However, there are methods for dealing with trend. In case C, if the trend is calculated by assuming a linear drift in m($\mathbf{x}$) and finding a linear equation to approximate the trend, then second order stationarity can be achieved and ordinary kriging performed. At each sample location, the fitted trend value at m($\mathbf{x}$) is subtracted from the observed value, resulting in residual values, $\xi(\mathbf{x})$, at each location. Experimental variograms are constructed from these residuals, variogram models are fitted to these experimental variograms, and an ordinary kriging technique can be applied to estimate residuals at

locations without values. The trend m(**x**) is then added to each known and estimated ξ(**x**), providing K(**x**) hydraulic conductivity estimates each point on curve C.

(2) Experimental variogram

Statistical models resort to weaker stationarity assumptions, such as intrinsic hypothesis postulating (1) the mean is the same everywhere ; and (2) for all distances, the variance of K(**x**) - K(**x** + **h**) is defined and is a unique function of **h**. A system satisfying stationarity of order two also satisfies the intrinsic hypothesis, but the converse is not true. The intrinsic hypothesis allows determination of a statistical structure without demanding prior estimation of the mean. In this case the semivariogram, or commonly called the variogram, defined as

$$\gamma = \frac{1}{2} \text{Var}[K(\mathbf{x}) - K(\mathbf{x} + \mathbf{h})]$$

represents correlation structure.

When the variance exists, there is a relationship between $\gamma(|\mathbf{h}|)$ and $C(|\mathbf{h}|)$,

$$\gamma(|\mathbf{h}|) = \sigma^2 - C(|\mathbf{h}|)$$

where the assumption of constant variance σ^2 exists over the domain. (ASCE, 1990)

To apply geostatistical estimation, analysts must perform structural analysis to determine experimental semivariograms from the data. For each distance class, identified by the average distance in that interval, the calculation of the experimental variogram is

$$\gamma * (\mathbf{h}) = \frac{1}{2N(\mathbf{h})} \sum_{(i,j)|\mathbf{h}_{i,j}=\mathbf{h}} \left[k(\mathbf{x}_i) - k(\mathbf{x}_j) \right]^2$$

where $k(\mathbf{x}_i)$ is a hydraulic conductivity measurement value at point $\mathbf{x}_i$; $k(\mathbf{x}_j)$ is a hydraulic conductivity measurement value at point $\mathbf{x}_j$; $\mathbf{h}$ separate (i,j) pairs; and $N(\mathbf{h})$ is the number of pairs of data points belonging to the distance interval represented by $\mathbf{h}$ (ASCE, 1990). Figure II.2 illustrates a typical empirical variogram which is defined only at discrete separation distances.

Occasionally, construction of an experimental variogram will indicate the variance at a distance zero is not zero. This phenomenon called the nugget effect will cause the variogram constructed on the data to have a classical variogram shape but intersects the 'γ' axis at some value greater than zero. It is also an indication that observed values do not always gradually change as position moves a short distance. If the nugget effect is equal to the sill, the variogram looks like a straight line, indicating no spatial correlation between samples, or independence.
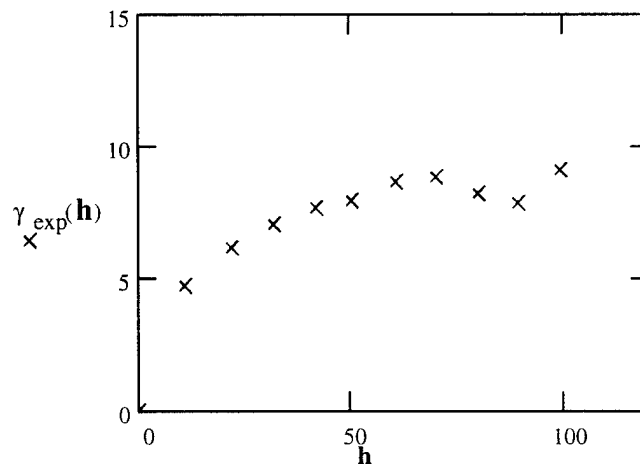

Figure II.2 Experimental variogram.

Some experimental variograms, known as directional variograms, also validate the common assumption of isotropy or confirm the existence of anisotropy. For example, if the variogram obtained by analyzing all points oriented north and south of each other is the same as the variogram obtained by analyzing all points oriented east and west of each other, then the spatial field is under a condition called isotropy. If this uniformity does not occur in every direction, it is inappropriate to assume isotropy.

(3) Variogram models

The next step is fitting a semivariogram model or combination of models to the estimated values of $\gamma^*(|\mathbf{h}|)$. When constructing an ordinary kriging system, the variogram values between sampled locations are not the only requirements. Variogram values between all sample locations and locations for estimates are necessary. Additionally, it is possible the separation vector between the sample locations and the future estimated locations involve a direction and distance unaccounted for with a sample variogram value. Therefore, to build the matrices later to be defined for ordinary kriging, the selected model needs to provide a variogram value for any possible separation vector. In addition to building the matrices for ordinary kriging, correlation scale and covariance parameters given by the continuous variogram model provides the necessary inputs to stochastic transport equations.

Figure II.3 illustrates commonly used functions in variogram models. Each model has a sill equal to the variance σ^2 for large separation distance h.
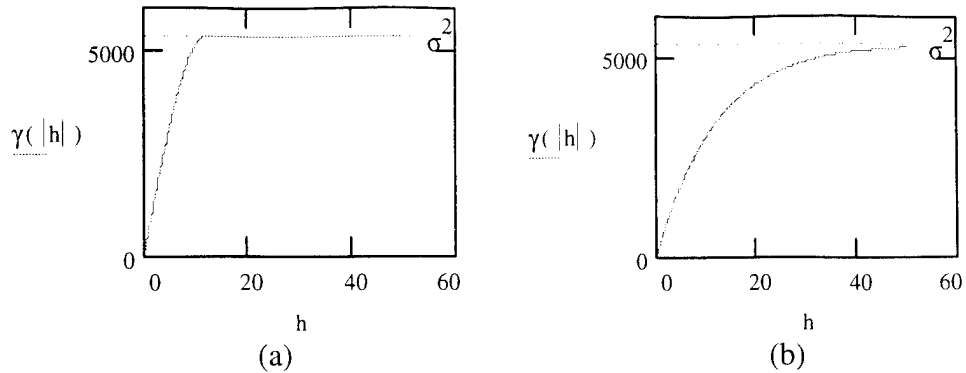
Figure II.3 (a) Spherical and (b) exponential variogram models.

The equation for the spherical model in Fig. II.3 (a) is

$$\gamma(|\mathbf{h}|) = \begin{cases} \sigma^2 \left[\dfrac{3}{2}\left(\dfrac{|\mathbf{h}|}{\ell}\right) - \dfrac{1}{2}\left(\dfrac{|\mathbf{h}|}{\ell}\right)^3 \right], |\mathbf{h}| \leq \ell \\ \sigma^2, |\mathbf{h}| > \ell \end{cases}$$

and the exponential semivariogram model in Fig. II.3 (b) is

$$\gamma(|\mathbf{h}|) = \sigma^2\left(1 - \exp\frac{|\mathbf{h}|}{\ell'}\right)$$

where ℓ is the range and ℓ' is integral or correlation scale. The spherical variogram shows $\gamma = \sigma^2$ when $|\mathbf{h}| > \ell$, meaning the variance reaches the sill at a distance equal to range ℓ. It also implies the lack of correlation among points whose separation distances are larger than ℓ. The exponential model approaches sill σ^2 asymptotically. The range for the exponential variogram is ℓ for which $\gamma(\ell) = 0.95\sigma^2$. The exponential model is important because it is often assumed by researchers in stochastic hydrology (Woodbury and Sudicky, 1991).

For the ordinary kriging system to have one and only one stable solution, analysts ensure the mathematical condition of positive definiteness for the covariance matrix to be defined in the ordinary kriging system. One way of satisfying the positive definiteness condition is to use a few functions known to be positive definite. Then, combinations of these positive-definite functions are available to form new functions for variogram fitting which are also positive definite. (Isaaks and Srivastava, 1989).

The choice of variogram model is a prerequisite for ordinary kriging. The variogram model chosen for ordinary kriging provides a powerful ability to customize the ordinary kriging estimation procedure. In practice the basis for selection of the pattern of spatial continuity for the random function model is the spatial continuity evident in the sample data set (Isaaks and Srivastava, 1989). Figure II.4 illustrates a possible fitting to the experimental variogram in Figure II.3.
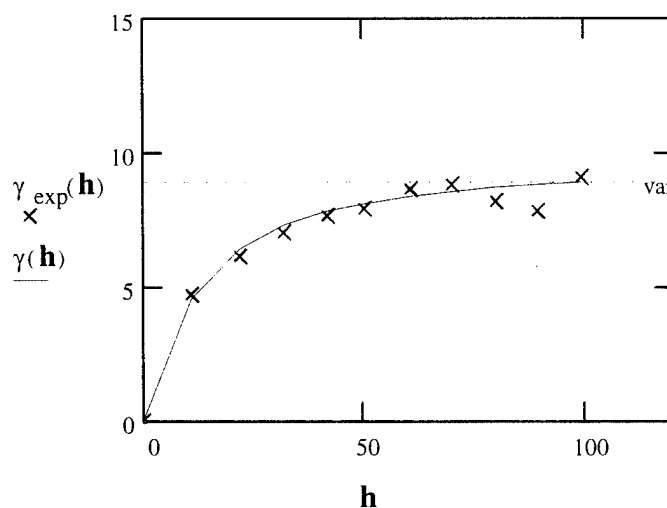
Figure II.4 Fitting a model variogram to an experimental variogram.

Although various methods to fit the rising portions of experimental variograms such as ordinary least squares are available, they are outside the scope of this thesis. Modelers also often use personal judgment in fitting a combination of variogram models to experimental variograms.

(4) Application - kriging

Ordinary kriging is a means of providing values at unknown locations based on known locations by obtaining the most information possible from the samples taken. The selection of this geostatistical estimation technique increases the ability to faithfully reproduce the natural variability and, in turn, provide a better input map to a groundwater model.

In ordinary kriging the mean value is assumed constant, but unknown. In fact, the unbiased condition dictates that the mean become a part of the solution. To accomplish the unbiasedness, $K*(\mathbf{x}_0)$ is a linear combination of the measured values as follows:

$$K*(\mathbf{x}_0) = \sum_{i=1}^{n} w_i k(x_i)$$

where $K*(\mathbf{x}_0)$ is the estimated value of hydraulic conductivity K at $\mathbf{x}_0$ based on known values of hydraulic conductivity at sampled locations, and w_i's are weights chosen to satisfy two statistical conditions. The first condition requires $K*(\mathbf{x}_0)$ to be an unbiased estimator:

$$E[K*(\mathbf{x}_0) - K(\mathbf{x}_0)] = 0$$

substituting for $K*(\mathbf{x}_0)$ yields

$$E\left\{\left[\sum_{i=1}^{n} w_i K(x_i)\right] - K(\mathbf{x}_0)\right\} = 0$$

Taking the expectation value of each term and equating it to the constant mean m gives

$$\sum_{i=1}^{n} w_i E[Z(x_i)] - E[Z(\mathbf{x}_0)] = \sum_{i=1}^{n} w_i m - m = 0$$

resulting in the unbiasedness condition satisfied for only those weights which sum to 1 as follows:

$$\sum_{i=1}^{n} w_i = 1$$

The second condition requires the estimator $K^*(\mathbf{x}_0)$ to have minimum variance shown to be

$$V[Z^*(\mathbf{x}_0) - Z(\mathbf{x}_0)] = 2\sum_{i=1}^{n} w_i \gamma_{i0} - \sum_{i=1}^{n}\sum_{j=1}^{n} w_i w_j \gamma_{ij}$$

where unbiasedness condition above minimizes $\gamma_{ij} = \gamma(|\mathbf{x}_i - \mathbf{x}_j|)$ by introducing a Lagrangian multiplier μ and writing

$$\frac{\partial \text{Var}[Z^*(\mathbf{x}_0) - Z(\mathbf{x}_0)]}{\partial \lambda_i} - 2\mu = 0 \quad \text{for } i = 1, 2, ..., n. \text{ (ASCE, 1990)}$$

These n equations plus the unbiasedness condition $\sum_{i=1}^{n} w_i = 1$ yield the ordinary kriging system:

$$
\begin{bmatrix}
\tilde{\gamma}_{11} & \tilde{\gamma}_{12} & \tilde{\gamma}_{13} & \rightarrow & \tilde{\gamma}_{1n} & 1 \\
\tilde{\gamma}_{21} & & & & \tilde{\gamma}_{2n} & 1 \\
\tilde{\gamma}_{31} & & & & \tilde{\gamma}_{3n} & 1 \\
\downarrow & & & & \downarrow & \downarrow \\
\tilde{\gamma}_{n1} & \tilde{\gamma}_{n2} & \tilde{\gamma}_{n3} & \rightarrow & \tilde{\gamma}_{nn} & 1 \\
1 & 1 & 1 & \rightarrow & 1 & 0
\end{bmatrix}
\begin{bmatrix} w_1 \\ w_2 \\ w_3 \\ \downarrow \\ w_n \\ \mu \end{bmatrix}
=
\begin{bmatrix} \gamma_{10} \\ \gamma_{20} \\ \gamma_{30} \\ \downarrow \\ \gamma_{n0} \\ 1 \end{bmatrix}
$$

$$\mathbf{C} \quad\cdot\quad \lambda \;=\; \mathbf{D}$$

Taken by itself, the **D** matrix provides a weighting scheme similar to inverse distance methods. Like an inverse distance weight, the variance between sample and the point being estimated increases as the sample gets farther away. However, unlike inverse distance weights, the calculated variances come from a much larger family of functions than $|h|^p$. (Isaaks and Srivastava, 1989)

Additionally, what really distinguishes ordinary kriging from other inverse distance methods is the role of the **C** matrix. The **C** matrix records distances between each sample and every other sample, providing the ordinary kriging system with information on clustering of the sample data. If two samples are very close to each other, the corresponding entry in the **C** matrix records a small value. If the two samples are far apart, this corresponding entry records a large value. (Isaaks and Srivastava, 1989)

Therefore, the ordinary kriging system takes into account two important aspects of estimation problems - distance and clustering:

$$\lambda \;=\; \mathbf{C}^{-1} \cdot \mathbf{D}$$

Clustering   Distance

Finally, given this four-step process of data, experimental variogram, variogram model, and application, the geostatistical analysis portion of Chapter III employs this methodology to sampled data from the MADE site at Columbus AFB. Because of the vast computational expense of calculating semivariogram values and kriging with thousands of pairs of data, the Geostatistical Software Library (GSLIB) [Deutsch and Journel, 1992] proves a tremendous aid in the characterization of the hydraulic conductivity random field of the MADE site in Chapter III.

Macrodispersion Investigations in Field Studies

There is a link between the advective and dispersive qualities of an aquifer and the hydraulic conductivity field. Some form of the advection-dispersion equation is typically used to quantify the transport of dissolved contaminants in groundwater. The advection portion of this equation describes the bulk transport of contaminants relative to the mean groundwater velocity - a function of the velocity vector. The dispersive portion describes the spreading of solutes in the porous medium caused by hydrodynamic dispersion and molecular diffusion. Hydrodynamic dispersion arises from the non-uniform flow field existing in individual pore channels and, on a larger scale, from the non-uniform flow caused by local hetreogeneity within the aquifer material. This latter form of spreading, called macrodispersion, generally dominates smaller-scale processes, including molecular diffusion, in porous media of high permeability (Hess et al., 1992). Because advection

and macrodispersion are both functions of the velocity vector, which is a function of the hydraulic conductivity tensor, knowledge of the spatial distribution of hydraulic conductivity is vital to adequately describing these phenomena.

Macrodispersion cannot readily be studied in the laboratory. A three-dimensional dispersivity tensor characterizes dispersion, but one-dimensional tracer experiments in aquifer cores typify measurements of dispersivity at the laboratory scale. Velocity variations within individual pore channels predominantly influence these measurements. Therefore, laboratory measurements may have little relevance to field scale dispersion because they do not incorporate the heterogeneity of the aquifer. (Dagan, 1986; Gelhar 1986; Hess et al., 1992).

To adequately characterize macrodispersion, major field studies at the Borden site (Sudicky, 1986), the Cape Cod site (Hess et al., 1992), and the Columbus AFB site (Boggs et al., 1992) were designed to study macrodispersion on the field scale under conditions of various naturally occurring degrees of heterogeneity. For example, the Borden aquifer consists of medium-grained, fine-grained and silty fine-grained sand with infrequent silt, silty-clay and coarse sand layers (Sudicky, 1986). Also, the aquifer at Otis Air Base on Cape Cod consists of medium-to-coarse sand and gravel, containing typically less than 1% silt and clay (Hess et al., 1992). Glaciofluvial sediments composed both of these aquifers; thus, researchers also consider them relatively homogeneous (Boggs et al., 1992). The MADE site at Columbus AFB, however, is very heterogeneous with

hydraulic conductivity values typically ranging over 2-4 orders of magnitude at each

sampling test well site (Boggs et al., 1992). In fact, the MADE hydraulic conductivity

field is the focus of analysis in Chapter III.

In the Borden experiment, field investigations were made in a sand quarry on Canadian

Forces Base Borden, Ontario. After extracting core samples, permeameter analyses

provided estimated measurements of K. The coring locations were chosen along two

transects, one along the direction of mean groundwater flow and one transverse to the

mean flow direction. Visual inspection of histograms confirmed the hydraulic

conductivity followed a lognormal probability density function and the natural

logarithmic transformation of K resembled a Gaussian distribution with a slight skew.

Instead of using the variogram, Sudicky [1986] used an equally valuable geostatistical

method to characterize spatial variability - the autocorrelation function. The correlation

scale of the fitted exponential models were used to estimate macrodispersivities in a

three-dimensional porous media using stochastic transport theories of Gelhar and Axness

[1983] and Dagan [1987]. (Sudicky, 1986)

In the experiment at Cape Cod, field investigations were made in an abandoned gravel pit

south of Otis Air Base on Cape Cod MA. Two field methods were used to estimate

measurements of K at a small scale: permeameter analyses of the cores and flowmeter

tests. The spatial distribution of flowmeter wells and the coring locations were chosen to

facilitate construction of variograms used to obtain horizontal correlation (integral) scales

for the K distribution. Summary statistics for K of the two methods indicated that flowmeter data have a significantly higher mean and variance values than do the permeameter data, implying a consistent difference between K values obtained from these two methods. Compaction of the samples during coring appeared to be the most significant source of bias in the permeameter results. A geostatistical analysis of the aquifer displayed an underlying correlation structure using the basic tool of analysis - the experimental variogram. An exponential variogram model was chosen to represent the discontinuous empirical variograms with emphasis placed on fitting the rising portion of the variogram. The correlation scale of the fitted exponential variogram models were also used to estimate macrodispersivities using the stochastic transport theories of Gelhar and Axness [1983] and Dagan [1987]. (Hess et al., 1992)

In the MADE experiment at Columbus Air Force Base, direct and indirect methods were used to estimate K. Among them were flowmeter tests, grain size analysis, slug tests, and permeameter analyses of the cores. Of all the methods evaluated, the borehole flowmeter provided the largest data set for analysis of the spatial distribution of the K field. Summary statistics for ln K of the methods indicated that flowmeter data have a significantly higher variance values than the grain size analysis and slug tests. The disparity between variance estimates could be attributed to several causes. For example, approximately 45% of the grained-size conductivity estimates were from core samples collected outside of the general region where flowmeter measurements were taken. Also, slug tests were of relatively short duration, causing the auger installation of the

piezometer to have a significant impact on the conductivity estimate. Compaction of the samples during coring again appeared to be the most significant source of bias in the permeameter results. Therefore, a geostatistical analysis of the aquifer assuming second order stationarity of the ln K field and another analysis using a detrended ln K field was based on borehole flowmeter measurements. An exponential variogram model was chosen to represent the discontinuous empirical variograms with emphasis placed on fitting the rising portion of the variogram. Table II.1 is a set of summary and geostatistical statistics resulting from the stationary and nonstationary analyses of the flowmeter data. Noting the smallest variance resulted from third-order polynomial detrending, in a similar fashion to the other two field experiments the correlation scale of the fitted exponential variogram models in the detrended case were used to estimate macrodispersivities using stochastic transport theories of Gelhar and Axness [1983] and Dagan [1987]. (Rehfeldt et al., 1992)

Table II.1 Summary statistics of 2187 borehole flowmeter measurements.

	μ	Minimum	Maximum	σ^2	λ_h'	λ_v'
ln K	-5.2	-10.1	0.4	4.5	12.8	1.6
3rd order residuals	-0.616	-5.67	3.95	2.7	5.3	0.7

The variance σ^2 and correlation scale λ' for the horizontal and vertical directions in each field experiment is summarized in Table II.2. Because the variance and correlation scales are relatively indistguishable between the Borden and Cape Cod when compared to the MADE experiment, Figure II.5 allows a comparison between the horizontal variogram

models associated with Cape Cod and the MADE site. Because the MADE site is very heterogeneous in comparison to the Cape Cod site, the sill, or variance, in the variogram is much higher than that of Cape Cod. However, in a heterogeneous aquifer, properties which affect hydraulic conductivity, such as grain size, are not completely disordered in space (Hufschmied, 1986). With each degree of variability, or heterogeneity, there may exist a spatial covariability corresponding to the spatial distribution of the heterogeneity. This covariability manifests itself into the rising portion of the variogram. Therefore, the correlation scale of a very heterogeneous site can be significantly larger than the correlation scale of a relatively homogeneous site. In contrast, if with each added degree of variability there existed no spatial covariability, the effect on the variogram would correspond to the addition of pure variance, or nugget effect.

Table II.2 Variance and correlation scales of ln K in three field experiments.

	$\lambda_h{}'$	$\lambda_v{}'$	σ^2
Borden	2.8	0.12	0.29
Cape Cod	2.6	0.19	0.24
Columbus AFB	12.7	1.6	4.5

Modeled Horizontal Variograms

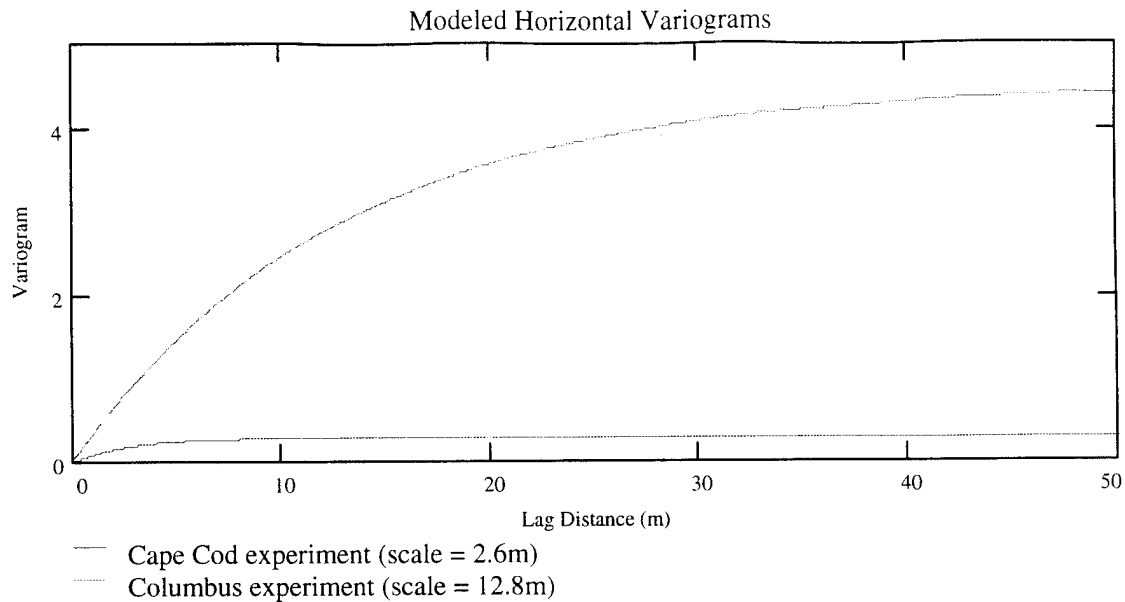Cape Cod experiment (scale = 2.6m)
Columbus experiment (scale = 12.8m)

Figure II.5 Horizontal variograms from Cape Cod and Columbus AFB experiments.

In each of these experiments, knowledge of the statistical properties of the spatial

distribution of K allowed calculation of the macrodispersivity tensor using stochastic

transport theories (Gelhar and Axness, 1983; Dagan, 1987). These stochastic transport

theories treat small-scale variation of hydraulic conductivity as a random process, and

solve stochastic forms of the transport equations. The dispersivity terms of these

stochastic models are functions of covariance of the hydraulic conductivity field.

Applications of these stochastic models require a large number of hydraulic conductivity

measurements in three dimensions to estimate the required covariance parameter.

(Rehfeldt et al., 1992)

In general, results from the Borden site, the Cape Cod site and the Columbus AFB site

showed good agreement between longitudinal dispersivities calculated from the statistical

properties of the K distribution using the stochastic theory of Gelhar and Axness [1983] and those dispersivities observed in tracer tests. However, in each experiment the stochastic theory underestimated transverse components of dispersivity. (Rehfeldt et al. 1992)

Because hydraulic conductivity field characterization was crucial in each of these studies to further research about macrodispersion and stochastic transport theory, all of these studies used geostatistical tools to characterize that particular site's hydraulic conductivity field. In fact, the analytical methodology in each of these experiments was relatively the same:

(1) Univariate analysis of the data

(2) Experimental variogram

(3) Variogram model

(4) Application.

Although the application focus of each field experiment was estimation of macrodispersivity, the focus of this thesis is the estimation of the K field for the Columbus AFB site using ordinary kriging. However, each step in the geostatistical analysis in Chapter III parallels the steps taken for analysis in each of these studies conducted by field experts.

<u>The Columbus AFB Site</u>

The purpose of the Macrodispersivity Experiment (MADE) at Columbus AFB was to predict and confirm macrodispersion and degradation of hydrocarbons at a heterogeneous site. Researchers performed injections of several aromatic hydrocarbons and a nonreactive tracer into the unconfined aquifer. By monitoring plume development of these solutes, and by measuring a number of physical characteristics, such as hydraulic conductivity, and chemical characteristics of the aquifer, the study provided data on those properties which significantly control the propagation of dissolved contaminants in groundwater systems (Stauffer and Manoranjan, 1994).

Figure II.6 is a location map of Columbus AFB with the experiment site annotated. Figure II.7 is a map showing the locations of 85 monitoring wells used to experimentally determine hydraulic conductivity by the borehole flowmeter method. The design of this well network used three closely spaced well cluster installations to define the spatial correlation of the hydraulic conductivity field. Other wells, spaced sufficiently far apart to avoid correlation, were installed to estimate the variance of the hydraulic conductivity field. (Rehfeldt et al., 1992)
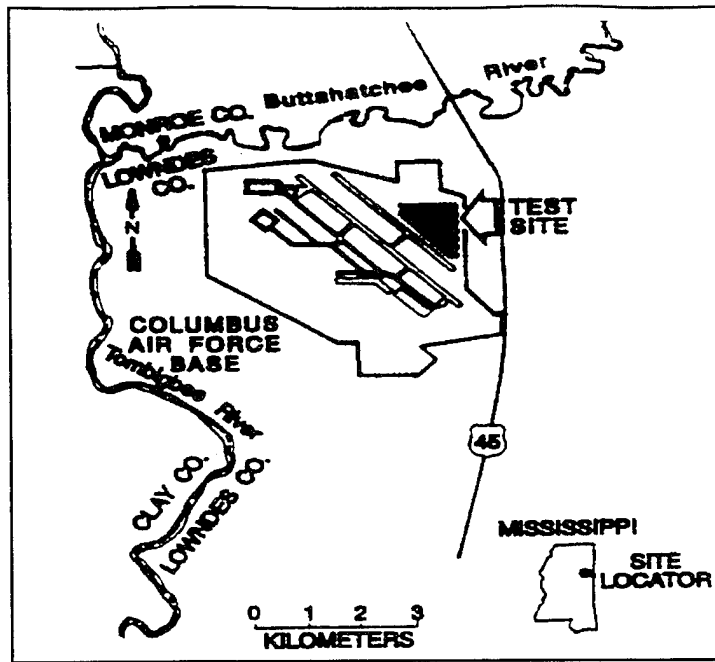
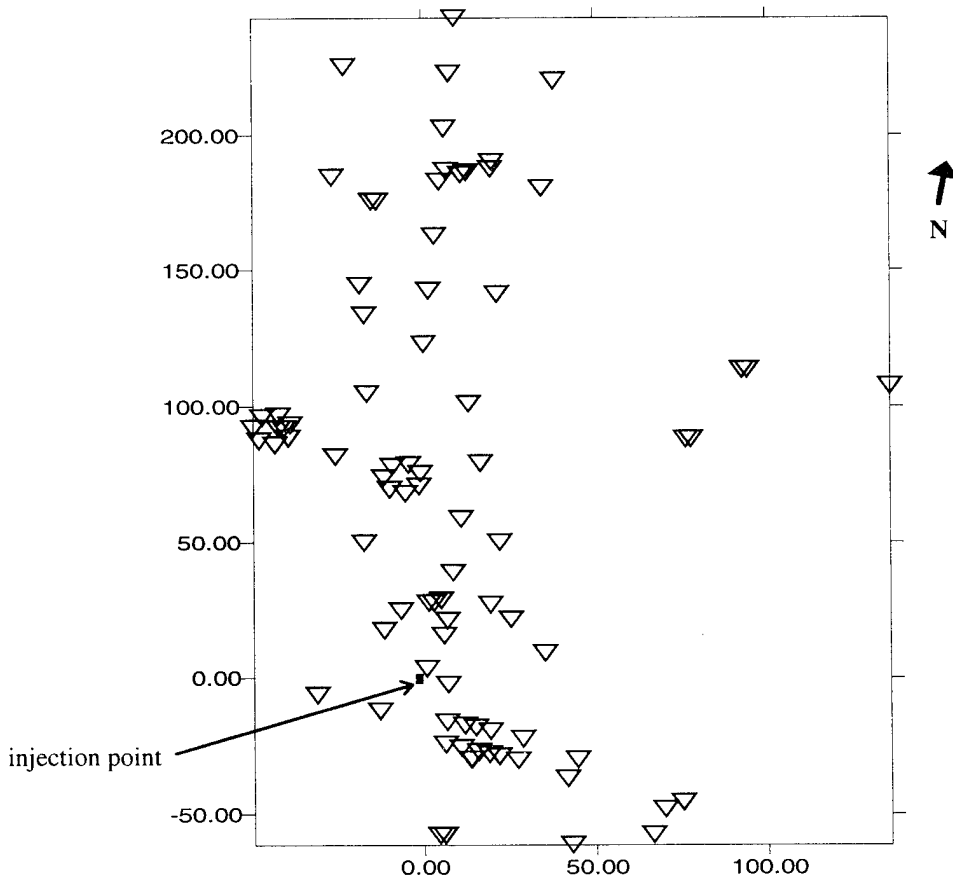Figure II.6   Location map of the MADE site. (Boggs et al., 1992)

Figure II.7 Monitoring wells tested by the borehole flowmeter.

Calculation of Hydraulic Conductivity from the Borehole Flowmeter Data

To appreciate the technique of indirectly measuring K values in the field using the

borehole flowmeter, one needs to recall Darcy's experiments. In Darcy's experiments,

water passed through a sand column with the volumetric flow rate Q and cross-sectional

area A. From these measurements a velocity Q/A, termed the specific discharge was

expressed as:

$$\frac{Q}{A} = q = -K\frac{h_2 - h_1}{L} \cong -K\frac{\partial h}{\partial L}$$

where q is the specific discharge and K is hydraulic conductivity (Darcy, 1856).

When Darcy performed his experiments, K was a constant of proportionality unique to

the uniform sand and the one-dimensional nature of his column. Porous material in an

aquifer may contain various types of soil with different properties, grain sizes and

orientation leading to heterogeneous, anisotropic hydraulic conductivity fields or tensors.

Figure II.8 denotes various combinations of uniformity and isotropy assumptions of two

points in space. As an example of how grain size and orientation affects conductivity,

clay particles in Figure II.9 have a flattened shape resulting in increased conductivity

parallel to platelet orientation. On a larger scale anisotropy is caused by the bedding of
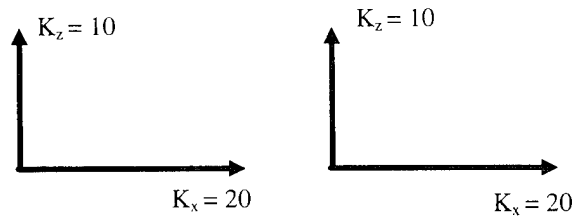
materials.

Uniformity: Homogeneous: K(x,y,z) = constant
               Heterogeneous: K(x,y,z) ≠ constant   (K is not constant in space)

Isotropy: Isotropic: $K_x = K_y = K_z$
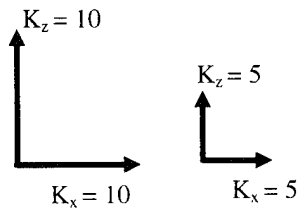   Anisotropic: $K_x \neq K_y \neq K_z$ (K is dependent on direction)

Homogeneous/Isotropic   Homogeneous/Anisotropic

$K_z = 10$ $K_z = 10$ $K_z = 10$ $K_z = 10$

$K_x = 10$ $K_x = 10$ $K_x = 20$ $K_x = 20$
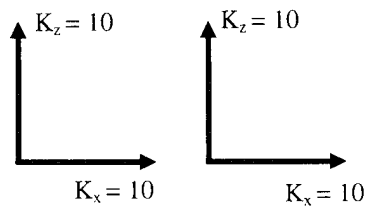
Heterogeneous/Isotropic   Heterogeneous/Anisotropic

$K_z = 10$ $K_z = 10$

               $K_z = 5$ $K_z = 5$

$K_x = 10$ $K_x = 5$ $K_x = 20$ $K_x = 10$

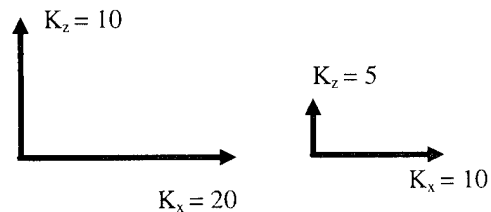Figure II.8 Directional qualities of heterogeneity and anisotropy. (Adapted from lecture
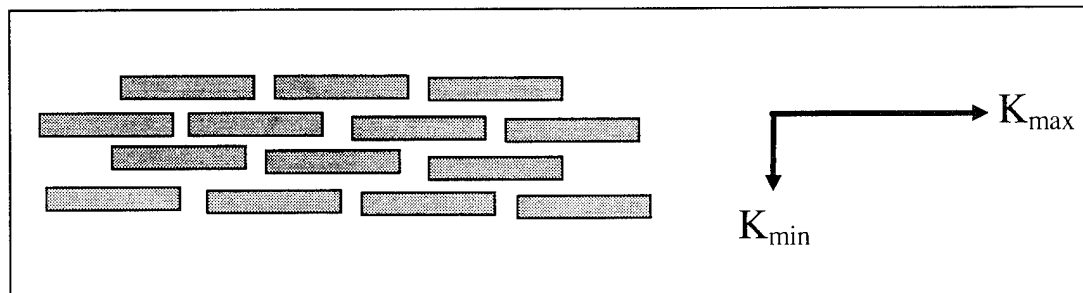by Maj. Edward Heyse, 1995)

K_{max}

K_{min}

Figure II.9 Effect of grain orientation on hydraulic conductivity. (Adapted from lecture
by Maj. Edward Heyse, 1995)

Currently, there are only a few tests available to measure the field hydraulic conductivity K. The methodology selected by Rehfeldt et al. [1992] for the MADE site was that of the borehole flowmeter. The hydraulic conductivity was obtained by solving the mass balance equation in terms of well discharge Q. The steady state mass balance equation characterizing two-dimensional horizontal flow under homogeneous, isotropic conditions is:

$$0 = K\left(\frac{\partial^2 h}{\partial x^2} + \frac{\partial^2 h}{\partial y^2}\right)$$

Rewriting the equation in radial coordinates:

$$0 = \frac{K}{2}\left(\frac{\partial^2 h}{\partial r^2} + \frac{1}{r}\frac{\partial h}{\partial r}\right)$$

Therefore, Figure II.10 indicates steady state flow to a well in a confined aquifer assuming horizontal, radial flow, homogeneous, isotropic conditions for one layer. Substituting the appropriate boundary conditions for a confined aquifer of thickness b, the steady-state solution in terms of well discharge, Q is:

$$Q = qA = K\frac{\partial h}{\partial r}(2\pi r\, b) \qquad \text{or} \qquad Q = \frac{2\pi Kb\left(h_0 - h_w\right)}{\ln\left(R / r_w\right)}$$

By substituting collected layer flowmeter data for Q, individual layer K values for layer i are:

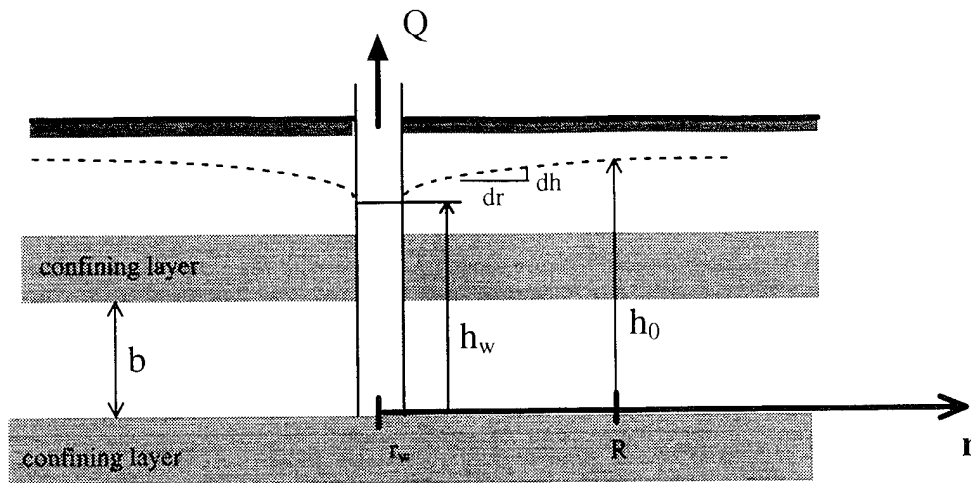$$K_i = \frac{Q_i \ln\left(R / r_w\right)}{b_i\, 2\pi\Delta h_i}$$

Figure II.10 Drawdown in a well with aquifer represented as a single layer. (Adapted from lecture by Maj. Edward Heyse, 1995)

Similar to Figure 10, Figure II.11 is a schematic presentation of flow from a multiple-layered aquifer. Similar to the volumetric flow into a single layer in the previous two-dimensional example, there is a corresponding volumetric flow rate from entering the well from each layer of the aquifer. A constant flow rate Q_{well} discharges from the well. As a result, there is a volumetric flow rate profile inside the well measured by the flowmeter at discrete locations along length of the screen. The well's volumetric flow rate measured by the borehole flowmeter is plotted vs. elevation above the bottom of the screen. Flow rate is zero at bottom of well and increases to Q_{well}, the constant flow rate, at the top of the well. The discharge into the well from each layer (Q_i) is the difference in flow rate at the top, Qz_i, and the bottom, Qz_{i-1}, of each layer. Assuming the horizontal flow discharge from each layer is proportional to the layer hydraulic conductivity, K_i, and

layer thickness, b_i, the normalized flow from each layer, (Q_i /b_i) for each layer is proportional to the hydraulic conductivity of that layer. (Rehfeldt et al., 1989)

Figure II.11 indicates prescribed elevations between hydrologically distinct materials; however, the depths of the actual geological interfaces were unknown to MADE experimenters. Therefore, the MADE experimenters used prescribed depths uniformly spaced along the well length for convenience. (Rehfeldt et al., 1989). Hence, the basis for the calculation of hydraulic conductivity is based on a questionable assumption of homogeneous layering of the aquifer. This is an exmple of a source of error in the sampling design and execution blocks of Figure I.1.
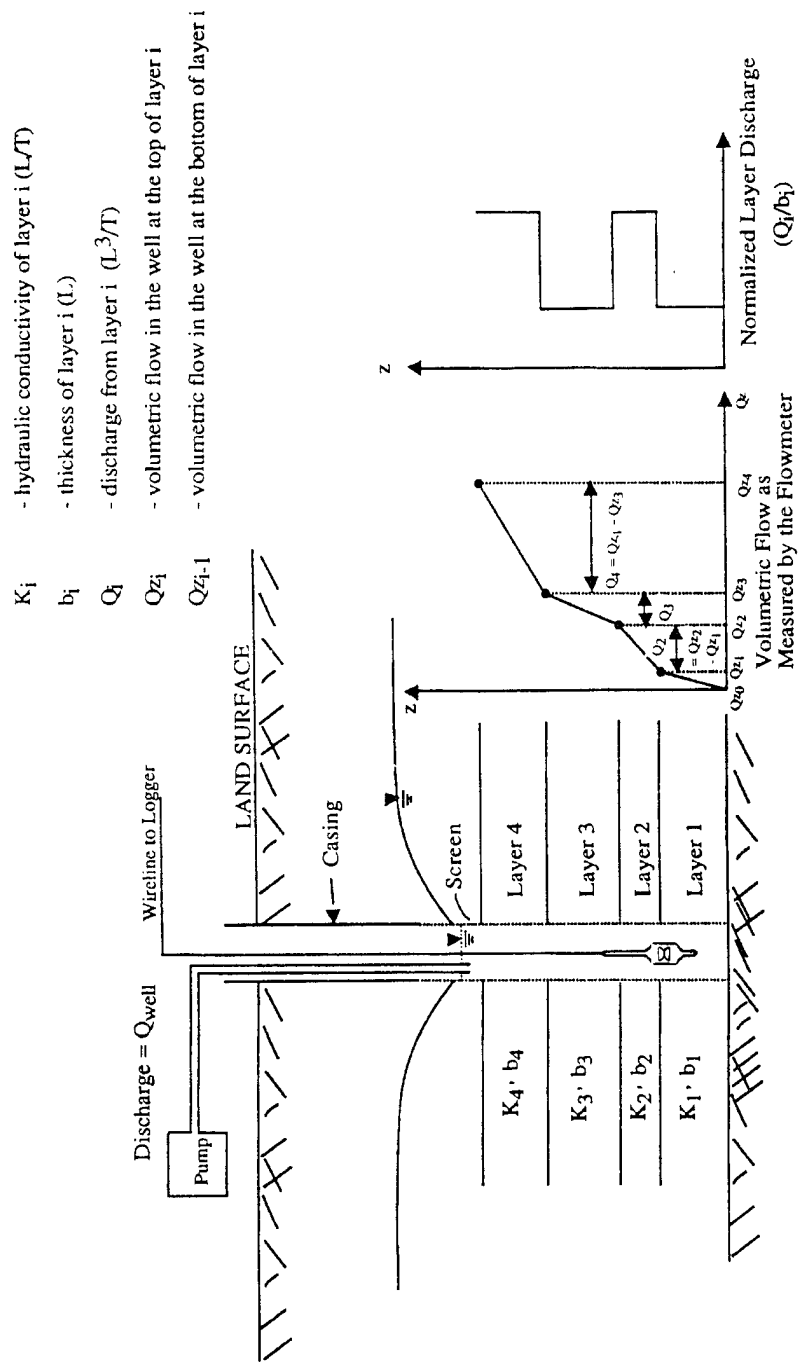
Figure II.11 Schematic of a pumping well in a layered aquifer with flowmeter

measurements. (Rehfeldt et al., 1989)

III. Spatial Characterization of the Hydraulic Conductivity Field

<u>Overview</u>

The purpose of Chapter III is to demonstrate the ability to quantify spatial variability of a naturally occurring hydraulic conductivity field at the MADE site. The analysis consists of hydraulic conductivity field characterization based on borehole flowmeter measurements at the MADE site at Columbus AFB. These analytical procedures parallel steps taken in the field studies discussed in Chapter II - in particular the steps taken by Rehfeldt et al. [1992] for the MADE site. Aquifer characteristics and data quality are addressed as well as a geostatistical analysis of the MADE data.

<u>Aquifer Characteristics</u>

Before an analysis of the hydraulic conductivity data is performed, it is important to examine the qualitative aspects of the hydrogeological setting from which the data was gained. Subsequent to analysis, this type of information often provides some qualitative validation of the methodology employed.

The shallow unconfined aquifer immediately underlying the MADE site at Columbus AFB MS consists of alluvial trace deposits averaging approximately 11m in thickness. The aquifer is composed of poorly- to well-sorted sandy gravel and gravelly sand with a

variable content of silt and clay. Sediments are generally unconsolidated and
cohesionless below the water table, occurring as irregular horizontal layers and lenses.
Marine sediments belonging to Eutaw Formation as well as marine sediments consisting
of clays, silts and fine-grained sands form an aquitard beneath the aquifer. (Boggs et al.,
1992).

The hydraulic head field exhibits complex temporal and spatial variability produced by
the heterogeneity of the aquifer and large seasonal changes of the water table. Hydraulic
conductivity measurements were taken at a water table elevation of approximately 62m,
but a 20-30% variation in the saturated thickness of the aquifer causes the water table to
periodically reach elevations of 65m. Figures III.1 and III.2 illustrate potentiometric
surfaces constructed using head measurements from deep and shallow piezometer
networks to indicate the spatial variability. The general direction of groundwater flow
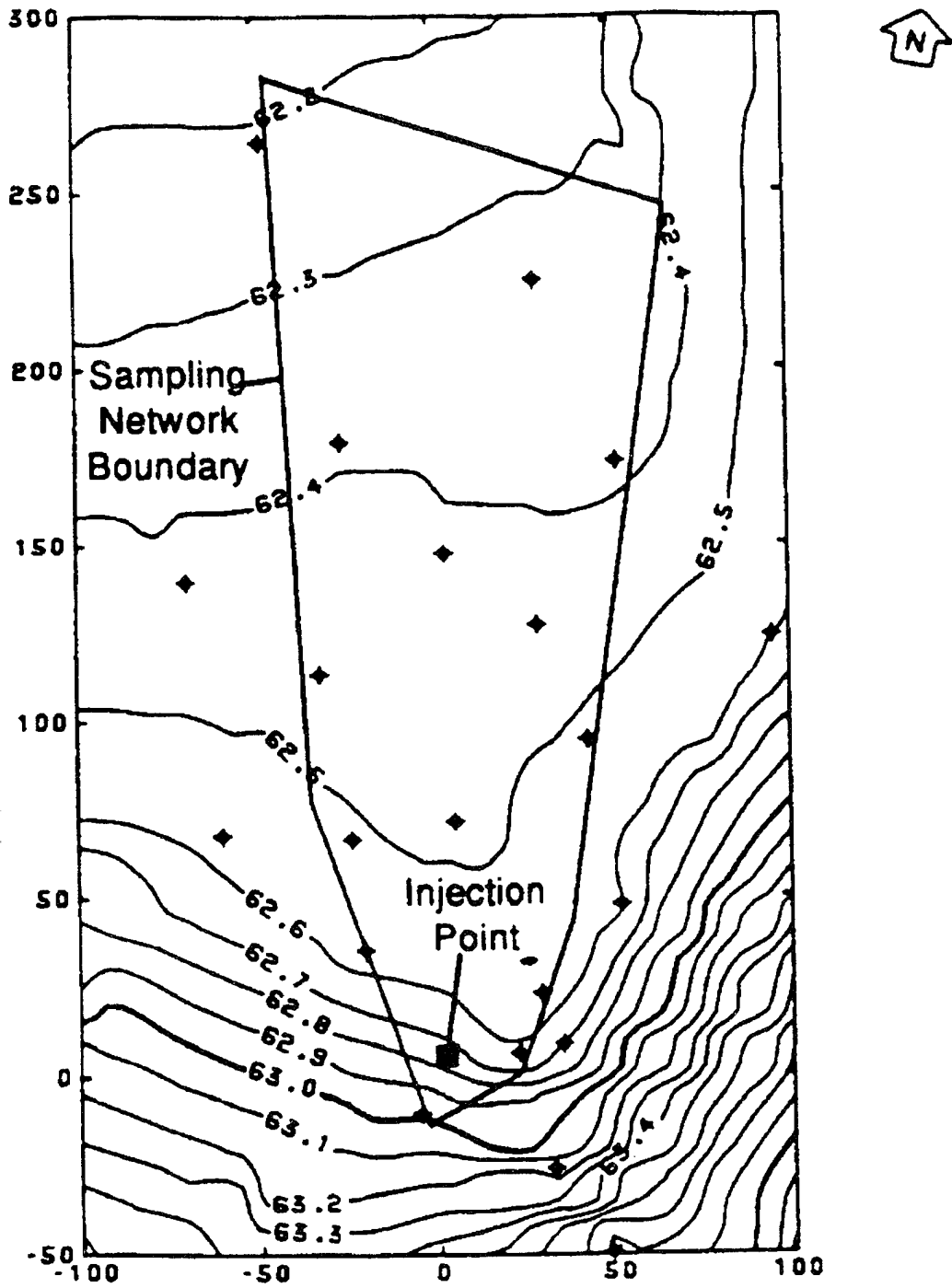inferred from both potentiometric surfaces is northward. (Boggs et al., 1992).

Figure III.1 Potentiometric surface maps derived from head measurements in shallow observation wells at mean well screen elevation of 61.1m. (Boggs et al., 1992)
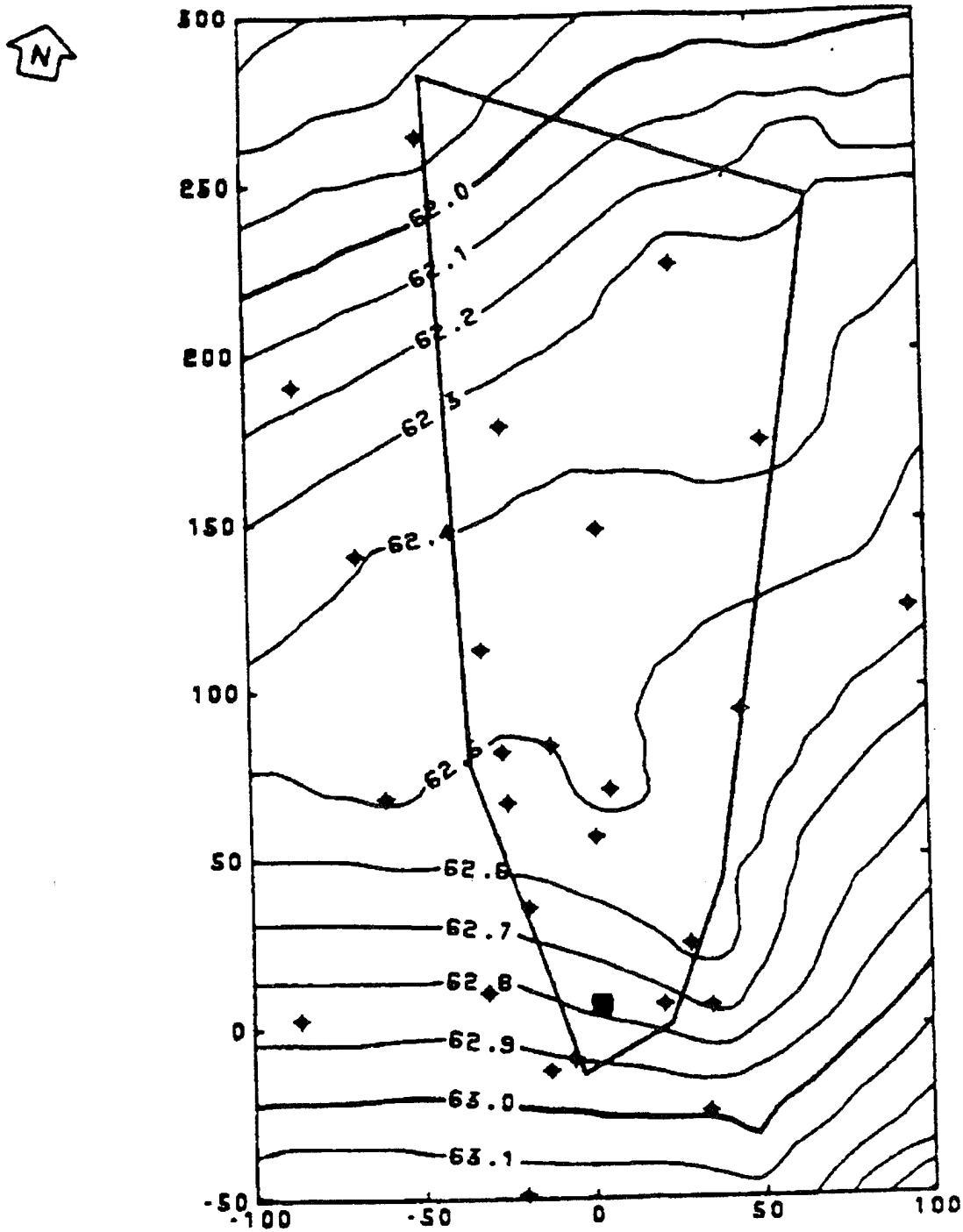
Figure III.2 Potentiometric surface maps derived from head measurements in deep

observation wells at mean well screen elevations at 56.3m. (Boggs et al., 1992)

Data Quality

Upon initial overview of the hydraulic conductivity data, there was a disproportionate number of zero measurements. In fact, of the 3647 samples available measurements taken from borehole flowmeter tests conducted in 85 fully penetrating wells, 308 measurements were zero or less than zero. Further study revealed the insertion of zeros into the data set indicated blank sections of pipe connecting screened sections (Rehfeldt et al., 1989). Therefore, 122 zero entries in 34 wells were deleted as a non-measurement due to blank screens, leaving 186 zero or less than zero entries without explanation in the analyzed data.

These 186 zero entries, or 5.1% of analyzed measurements, provided the largest frequency count of hydraulic conductivity measurements in a histogram. Given the assumption of lognormally distributed hydraulic conductivity values, 5.1% of all measurements would be lost to the logarithmic transformation.

It is important to distinguish the difference between no data and non-detectable measurements. Reporting zero in place of measurements less than the lowest detectable level throws away data that is useful to an analyst. Gilliom, Hirsch, and Gilroy (1984) showed through the use of Monte Carlo experiments that trends are more likely to be detected if data sets are not censored (Gilbert, 1987). The mean μ and variance σ^2 could be estimated in one of the following ways:

1. Ignore non-detectable measurements and compute $\bar{x}$ and s^2 using only "detected" values.

2. Replace non-detectable values with zeros and then compute $\bar{x}$ and s^2.

3. Replace non-detectable values by some value between zero and the lowest detection limit, such as one half of the lowest detection limit; then compute $\bar{x}$ and s^2.

The first two methods promote bias for both μ and σ^2. The third method has an unbiasedness condition for μ, but not σ^2. Kushner (1976) concluded biases in using the midpoint would be overshadowed by measurement error. (Gilbert, 1987).

Because of the inability to distinguish between no data and non-detectable measurements in all wells, the author uses method one in the analysis of this data, ignoring non-detectable values and causing a possible bias in estimate of μ in the positive direction and a smaller estimate for σ^2.

Trends and Detrending

The borehole flowmeter provided 3334 measurements for analysis of the spatial distribution of hydraulic conductivity at the Columbus site. It is useful to examine some general features of the field before proceeding with the structural analysis. It is evident from data inspection that the location of the injection point at coordinates (0, 0) is in a

region of relatively low hydraulic conductivity (i.e., on the order of 10^{-3}), while hydraulic conductivity in the far-field is 1 to 2 orders of magnitude larger. A southeast-northwest trend is evident from the hydraulic head fields indicated in Figure III.1. The closeness of the head contours near the injection point indicates lower hydraulic conductivity in the area, and the hydraulic gradient flattens in the region of larger hydraulic conductivity.

In addition, regions of higher hydraulic conductivity exist in the central portion of the aquifer and appear to correspond with a former river meander evident in aerial photographs. The river meander cuts across the sampling network and trends southwest to northeast. Because this heterogeneity is a large-scale feature, it can be accounted for with large-scale detrending by fitting a continuous polynomial function (deterministic) trend to the field. (Rehfeldt et al., 1992)

Because a trend does exist, it violates a condition of the stationarity of order two assumption where a constant mean exists across the site, making the hydraulic conductivity field a nonstationary random field. A nonstationary random field is a process composed of a stochastic component represented by small-scale, high frequency variations, and either a deterministic or random component representing large-scale, low-frequency variations. Estimation of covariance parameters for the ordinary kriging system in the case of a nonstationary field requires prior determination and trend removal before a variogram analysis. (Rehfeldt et al., 1992)

There are a number of proposed methods to remove trends from nonstationary random fields. The ordinary least squares method (OLS) is by far simplest to implement. Although the robustness of OLS is outside the scope of this thesis, the work of Russo and Jury [1987a,b] indicate the removal of a trend by OLS is valid, and Rehfeldt et al. [1992] used the OLS technique to detrend the Columbus AFB hydraulic conductivity field.

When fitting the first-, second-, and third-order trends to this random field, the method of weighted OLS was used to fit polynomial functions to the base 10 logarithm of the hydraulic conductivity measurements with no constraints applied to the fit. In an attempt to explain the measurement uncertainty, the weight associated with each conductivity measurement in the OLS technique is $W_i = K_i^{0.25}$ (Boggs et al., 1990). Consequently, the uncertainty associated with low hydraulic conductivity measurements is greater than the uncertainties associated with high hydraulic conductivity. Subsequent to its fit, the polynomial trend was numerically removed from each corresponding $\log_{10} K$ datum with residuals transformed into ln K residuals. The polynomial coefficients for the first-, second-, and third-order coefficients for the analysis of 3334 measurements are in Appendix B. (Rehfeldt et al., 1992)

Univariate Analysis

The general approach in this univariate analysis is the application of classical statistical techniques to the log hydraulic conductivity field and the residual log hydraulic conductivity field resultant of detrending.

It is common in stochastic groundwater analysis to assume hydraulic conductivity values are distributed lognormally (Freeze, 1975). Then, a logarithmic transformation is made so the data in this transformed space are normally distributed. Figure III.3 is a histogram of 3334 measurements of hydraulic conductivity, and Table III.1 contains the corresponding summary statistics including mean, variance, minimum, and maximum values. The histogram validates the assumption that hydraulic conductivity data are distributed lognormally and validates the employment of this assumption by Rehfeldt et al. [1992] in their analysis of a 2187 measurement subset of this data. The subsequent logarithmic transformation of the 3334 measurements yields the histogram in Figure III.4, illustrating the validity of normally distributed data in the transformed space. The summary statistics of the logarithmically transformed data are also in Table III.1 and they are relatively similar to those summary statistics calculated by Rehfeldt et al. [1992] in Table II.1.

Table III.1 Summary Statistics

	μ	Minimum	Maximum	σ2
K	0.0257	6.69 x 10^{-6}	1.446	0.00591
ln K	-5.6275	-11.915	0.3688	4.05
1st order residuals	-0.878	-7.766	4.317	3.721
2nd order residuals	-0.697	-6.714	4.748	3.056
3rd order residuals	-0.638	-6.250	3.693	2.801

Additionally, although slightly skewed, the removal of the third-order polynomial trend has produced ln K residuals that are also normally distributed in Figure III.5. When compared to parameter estimates of the original ln K data in Table III.1, the variances of the first-, second-, and third-order ln K residuals are somewhat smaller.
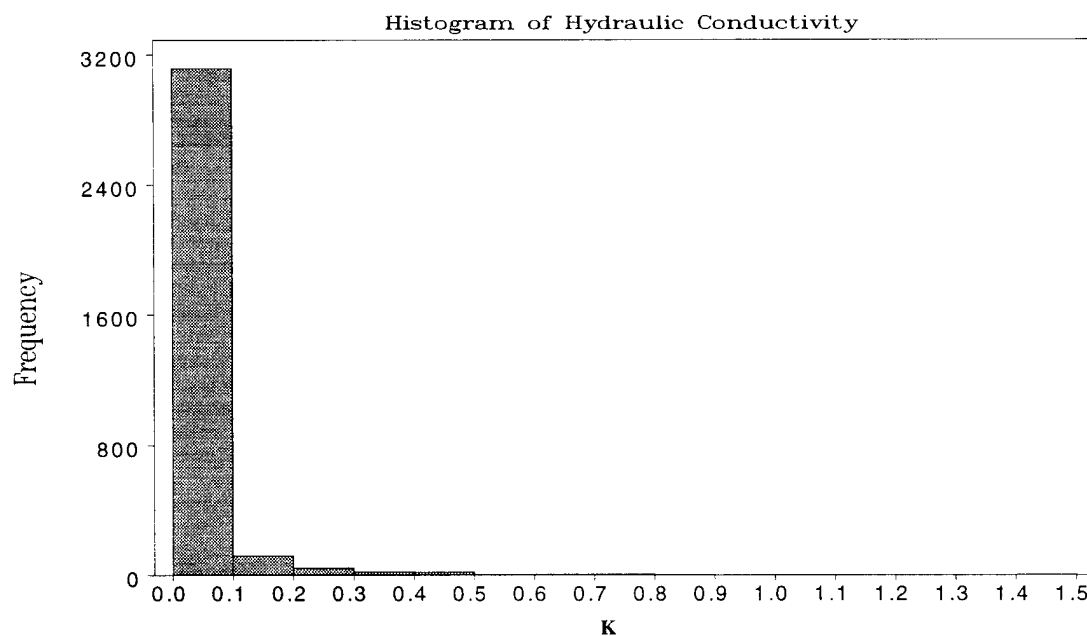
Figure III.3 Histogram of 3334 hydraulic conductivity measurements.

Rehfeldt et al. [1992] selected the third-order trend to characterize the ln K residual field because it was the lowest order trend that was physically compatible with the major

features of the collected plume data associated with the experiment. Because the characteristics of the third-order residuals, such as the variance, is very similar to that calculated by Rehfeldt et al. [1992] (see Table II.1), the following structural analysis, variogram modeling, and ordinary kriging estimation makes use of the third-order ln K residual field. Upon completion of the ordinary kriging application, the trend is numerically replaced.
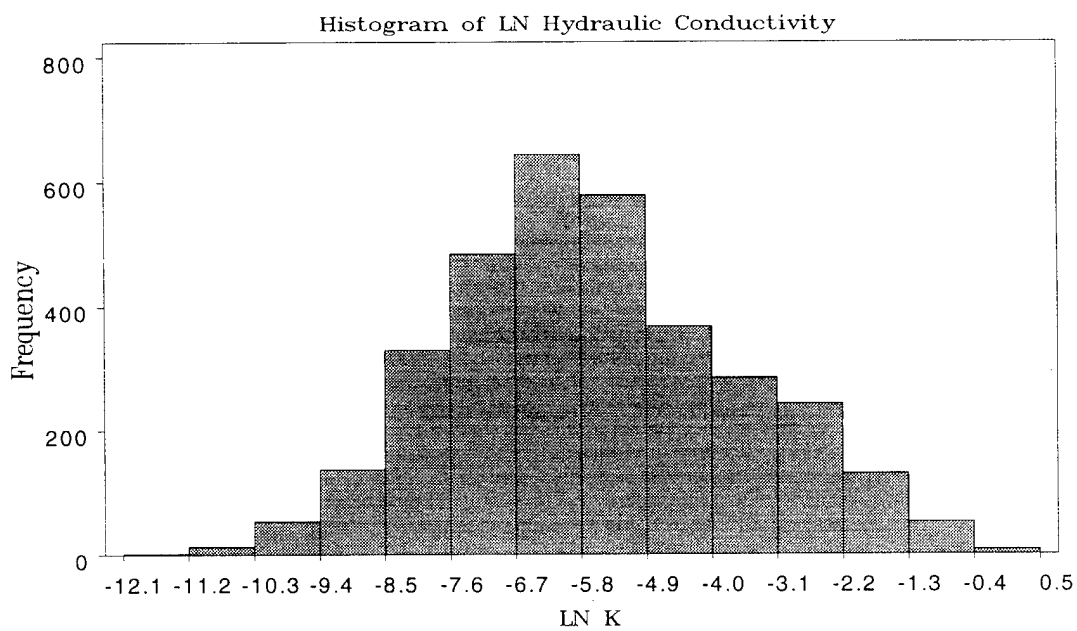
Histogram of LN Hydraulic Conductivity

Figure III.4 Histogram of 3334 ln K transformations.
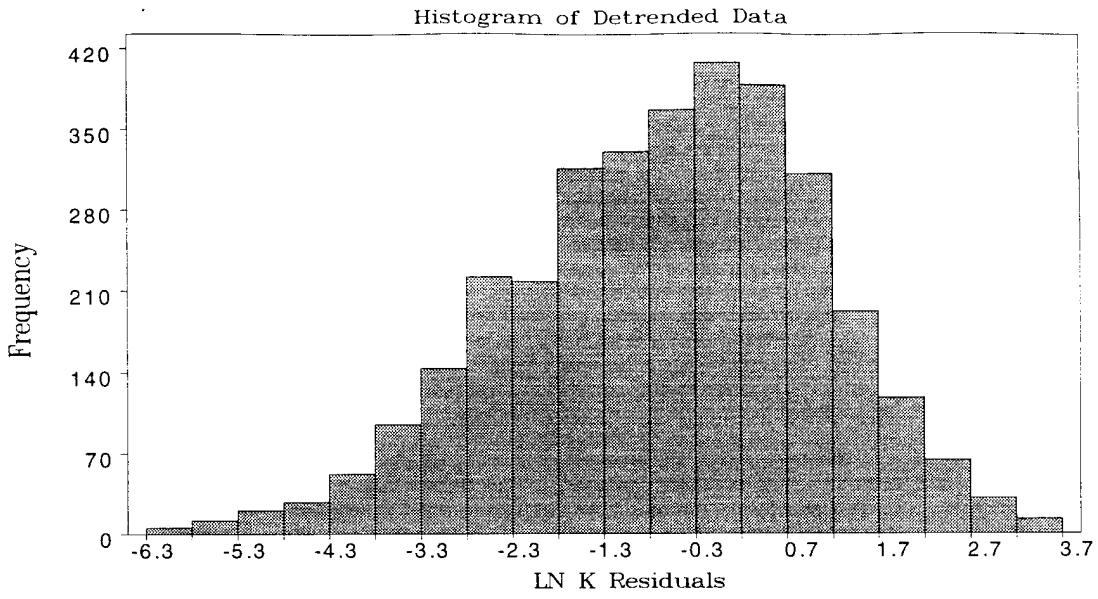
Figure III.5 Histogram of ln K residuals after polynomial detrending.

<u>Spatial Description</u>

The spatial description of the data has two major components - vertical and horizontal.

In the methodology of Rehfeldt et al. [1992], the empirical variogram in the vertical direction limited its consideration to data pairs located only directly above or below each other; i.e., no pairs involving more than one well were in the search region. Therefore, to calculate the empirical variogram in the vertical direction in Figure III.6, a band in the vertical direction is defined by an angle of $90° \pm 2.5°$ to a maximum horizontal distance of ± 0.1 meters, with the lag, or step, in the vertical direction defined as .16 m ± 0.08 m.
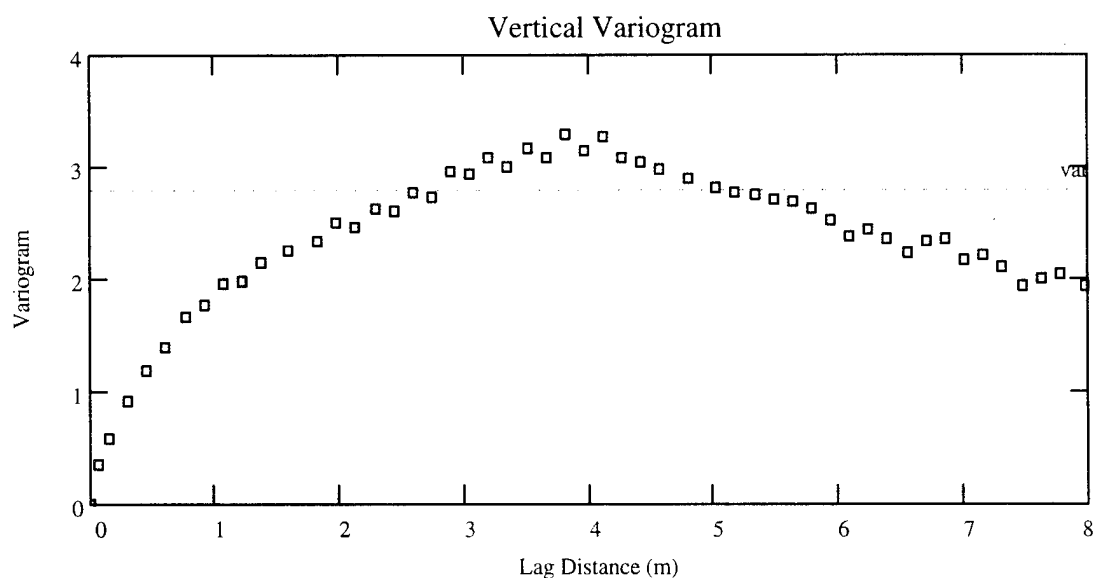
Figure III.6 Experimental vertical variogram.

Similarly, to calculate the empirical variogram in the horizontal direction in Figure III.7, a band in the vertical direction isdefined by an angle of $0° \pm 2.5°$ to a maximum vertical

distance of ± 0.16 m, with the lag, or step, in the horizontal direction defined as 5.0 m $\pm$

2.5 m. Figure III.7 is a plot of variograms calculated along different directions (0°, 45°,

90°, 135°) measured counterclockwise from the x-axis. A band of 22.5° on both sides of

each direction insured the inclusion of each data pair in one of the directions. Because of

the similarity of the directional empirical semivariograms, it is difficult to ascertain any

definite horizontal anisotropy.

To better visualize any definite horizontal anisotropy, the contour plot of the directional

variograms in Figure III.8 is an attempt to reveal more. The darker areas of the contour

plot indicate lower variance or higher correlation and lighter areas indicate higher

variance and lower correlation. Because no definite elliptical pattern of the shading

suggests horizontal anisotropy, the horizontal plane can be characterized with a single

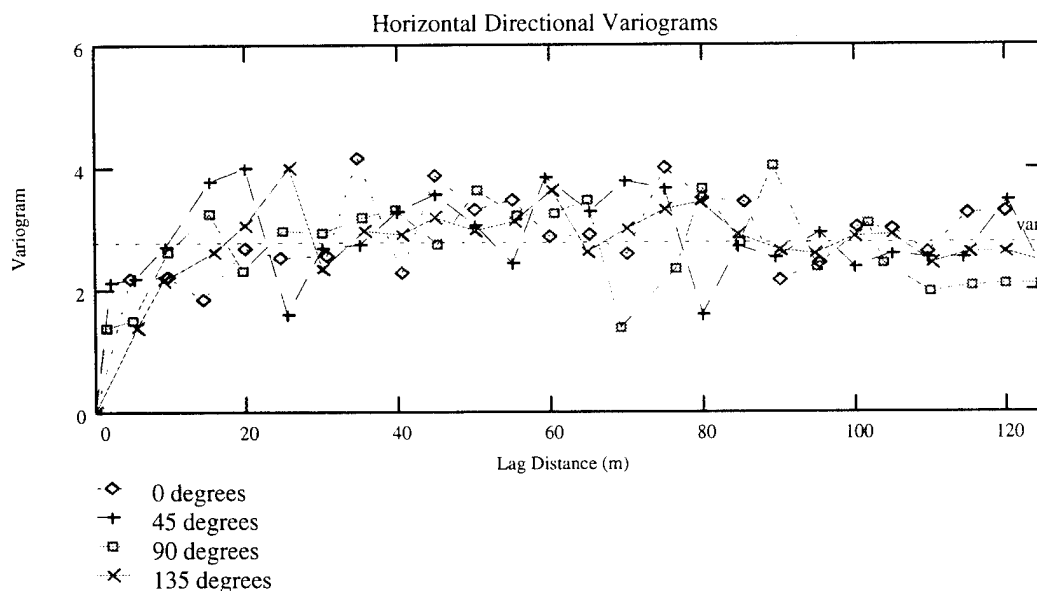omnidirectional variogram depicted in Figure III.9.

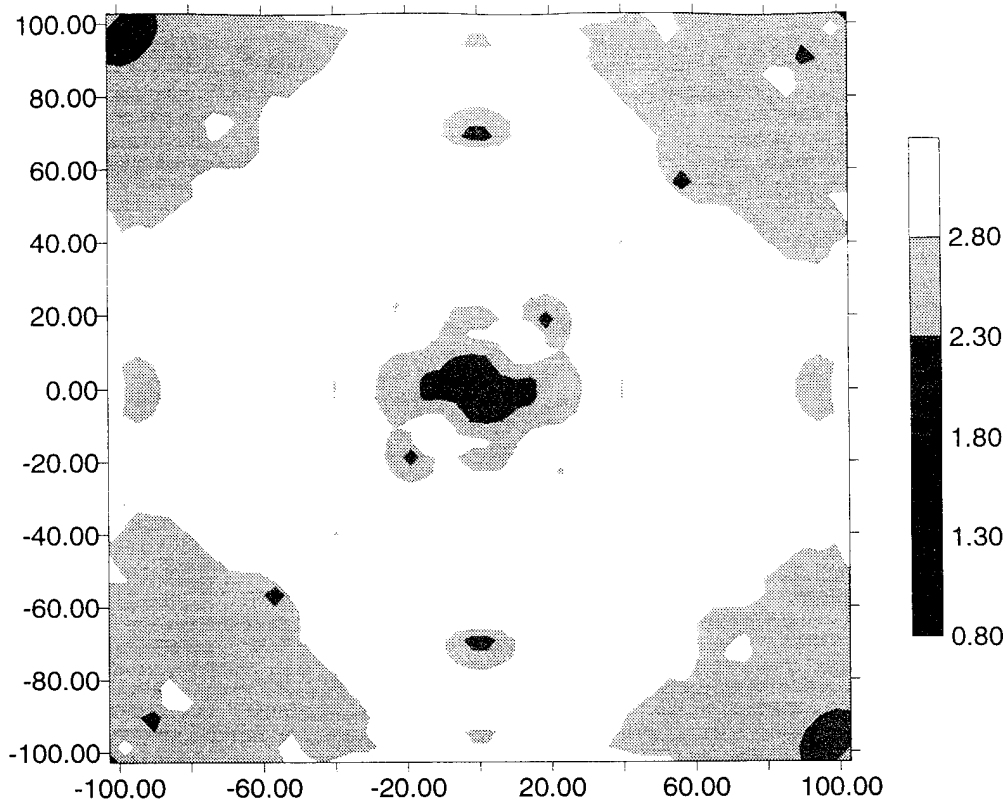Figure III.7 Experimental horizontal directional variograms.

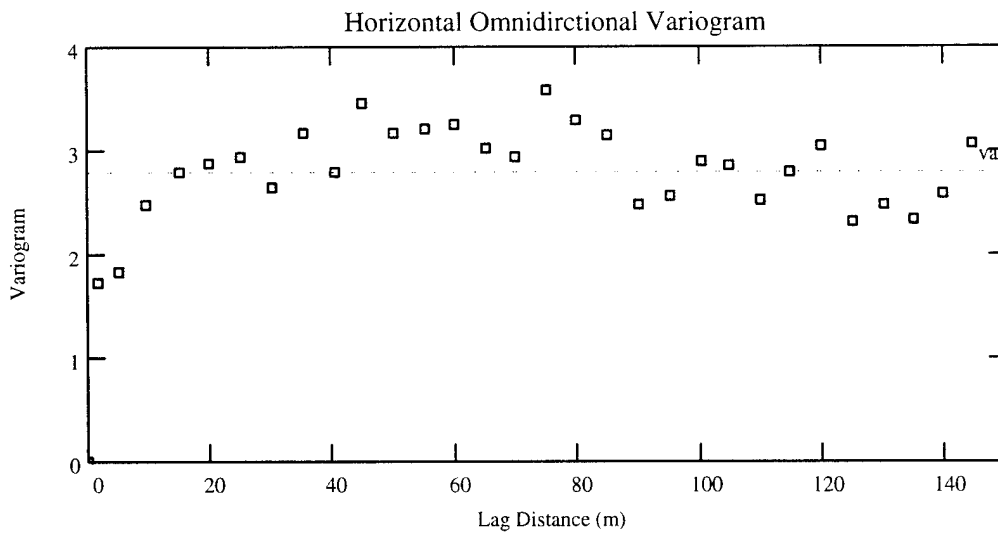Figure III.8 Contour plot of directional variograms.

Figure III.9 Experimental horizontal omnidirectional variogram.

Modeling the Sample Variogram

Because of the lack of existence of anisotropy in the horizontal plane, model variograms are fitted to the experimental vertical variogram as well as to the horizontal omnidirectional variogram in Figures III.10 and III.11, respectively. Woodbury and Sudicky [1991] deemed the exponential model important because it is often used by researchers in stochastic hydrology, and Rehfeldt et al. [1992] used the exponential model in fitting the vertical and horizontal variogram. To estimate the integral (correlation) scale in each direction, the exponential variogram model is judgmentally fitted to each sample variogram by adjusting the integral scale of the exponential variogram model while holding the sill equal to the sample variance. The resulting correlation scales in the vertical direction $\lambda_v{}'$ and horizontal direction $\lambda_h{}'$ are 0.85m and 3.5m, respectively.
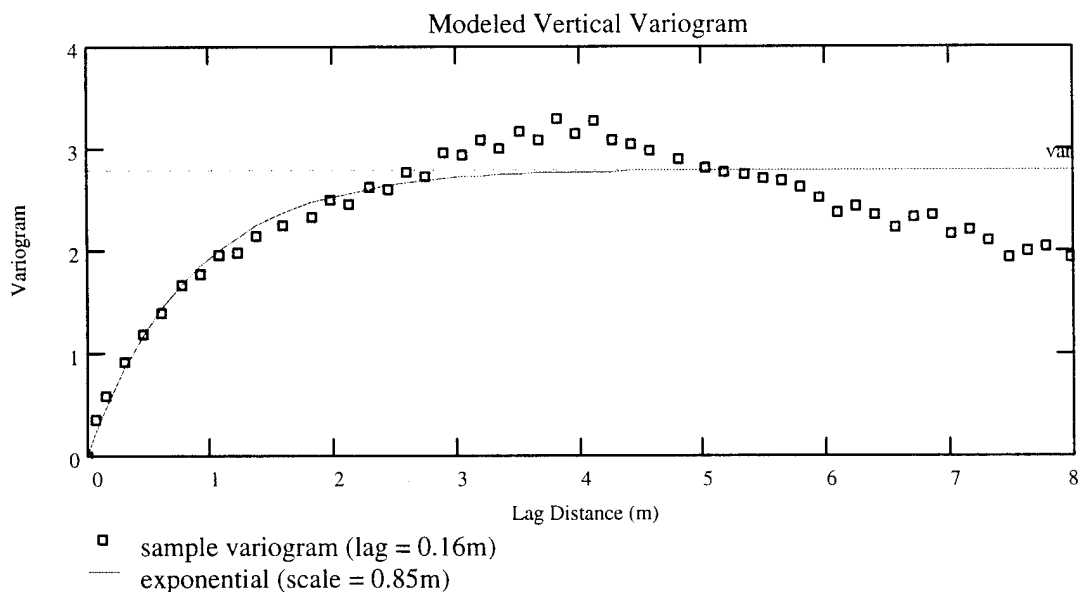
Modeled Vertical Variogram

□   sample variogram (lag = 0.16m)
—   exponential (scale = 0.85m)

Figure III.10 Modeled vertical exponential variogram.

Modeled Horizontal Variogram

□ sample variogram (lag = 5m)
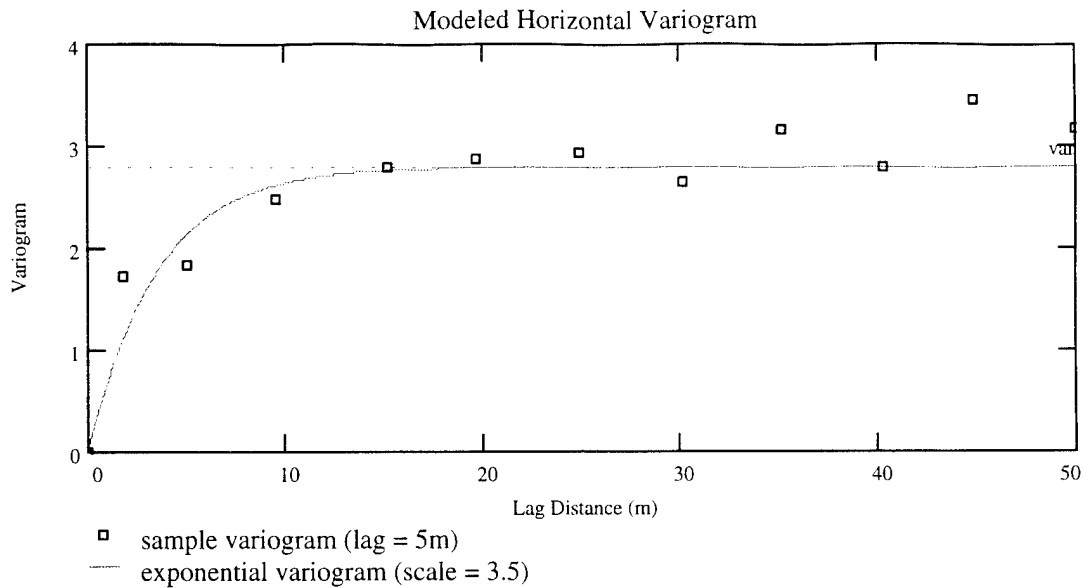— exponential variogram (scale = 3.5)

Figure III.11 Modeled horizontal exponential variogram.

Estimation of the Hydraulic Conductivity Field

Before performing ordinary kriging, the grid system must be defined. The output of the

ordinary kriging algorithm is a three-dimensional estimated hydrogeological map

describing the hydraulic conductivity field. Assuming this map serves as input into a

numerical model such as Modular Flow Model (MODFLOW), it is important to define

spatially location and spacing of the kriged estimates. If the discretization interval, or

grid spacing, is chosen to be much larger than the integral scale of the variogram model,

these heterogeneities described by variogram analysis will be greatly overestimated. In

particular the grid spacing must never exceed the integral scale (Neuman, 1982);

otherwise, spatial correlation will not be reflected in the grid.   Therefore, discretization

intervals for the MADE data must not exceed the integral scale for the modeled

variograms - 3.5m in the horizontal plane and 0.85m in the z-direction. Because

researchers at Tyndall AFB defined their experimental grid with spacings of 5.0m in the

x- and y- directions and 1.0m in the z-direction, for compatibility purposes, choices for

spacing in this grid system are 2.5m in the x- and y- directions and 0.5m in the z-

direction.

In addition to a defined grid system, a search pattern must also be defined for the kriging

algorithm of GSLIB. Most kriging algorithms consider a limited number of nearby

conditioning data. The MADE data are very closely spaced in a vertical direction and are

generally more sparsely in the horizontal direction. Because of some large distances

between sampling locations in the horizontal plane and the large anisotropy in the

horizontal and vertical directions, the search strategy needs to capture more information

in the horizontal region and less from the vertical directions. Therefore, only those data

falling within a search ellipsoid, depicted in Figure III.11, centered at the location being

estimated are considered (Deutsch and Journel, 1992). This anisotropic ellipsoid, used in

ordinary kriging of the MADE data, is oriented parallel with the horizontal plane with

equal major axis radii in x- and y- directions of 100m and a minor axis radius in the z-

direction of 5m.

Based on the well locations in Figure II.3, the pre-defined coordinates for the 8 corners of

the three-dimensional grid system are (-52.5, -27.5, 54.0), (52.5, -27.5, 54.0),

(-52.5, -27.5, 65.0), (52.5, -27.5, 65.0), (-52.5, 277.5, 54.0), (52.5, 277.5, 54.0),

(-52.5, 277.5, 65.0), and (52.5, 277.5, 65.0). With spacing defined at 2.5m in the x- and y-directions and 0.5m in the z-direction, a total of 121,647 estimates of hydraulic conductivity result from the kriging technique based on 3334 sample values and exponential variogram models with integral scale 3.5m in the x-y plane and 0.85 in the z-direction. These 121,647 estimates of hydraulic conductivity form a three-dimensional hydrogeological map for groundwater flow or contaminant transport modeling.
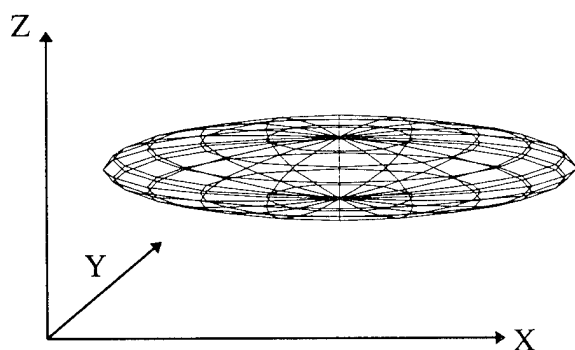

Figure III.12 Ellipsoidal search pattern used in GSLIB kriging algorithm.

Subsequent to ordinary kriging, the polynomial trend was reintroduced to estimates, providing an estimated K field. With estimates for K ranging from 4.92×10^{-10} cm/s to 0.55 cm/s, Appendix C contains 23 horizontal maps, representing each 0.5m discretized interval in the z- direction of the gridded system. Darker shading indicates regions of lower hydraulic conductivity while lighter shading indicates regions of higher hydraulic conductivity. Regions of high hydraulic conductivity cross the site in each case. In general the pattern of hydraulic conductivity is somewhat inconsistent with what is expected when considering the potentiometric surface maps derived from head

measurements in Figure III.1 and Figure III.2. However, the high conductivity transverse region appears consistent with a former river meander cutting across the site trending southwest to northeast. Hydraulic conductivity data derived from the flowmeter suggest the presence of the channel with regions of $K > 10^{-2}$ cm/s.

IV. Conclusions and Recommendations

Summary of Findings

There are several estimation methodologies available to characterize a hydraulic conductivity field. Because every point in space cannot physically be sampled, the estimation methodology selected must use the most information provided by the available samples. The most effective methodologies are geostatistical techniques because they use Euclidean distance information as well as statistical information to more adequately characterize spatial variability.

The focus of this thesis was to demonstrate a practical methodology for characterizing hydraulic conductivity in heterogeneous aquifers. This demonstration was accomplished by applying a defined four-step geostatistical process to the hydraulic conductivity data set from the MADE site at Columbus AFB.

During the MADE experiment, the borehole flowmeter method provided measurements for flow which are later transformed into hydraulic conductivity by direct calculation. Based on the 3334 measurements analyzed, an estimate for σ^2_{lnk} of 4.05 was obtained from the natural logarithm of the hydraulic conductivity measurements assuming second-order stationarity of the hydraulic conductivity field. Because the hydraulic conductivity field was clearly nonstationary, the method of weighted least squares was used to fit

three-dimensional polynomials of orders 1, 2, and 3 to log conductivity measurements. Each polynomial trend was numerically removed from the measured data, and an estimate for σ^2_{lnk} was obtained from the residuals of the natural logarithm of the hydraulic conductivity measurements. In agreement with Rehfeldt et al. [1992], because of its best fit and smallest variance, the third-order polynomial trend was judged the best representation of the conductivity trend with the lowest estimate for $\sigma^2_{lnk\ residuals}$ of 2.8.

A structural analysis revealed significant anisotropy comparing the horizontal to the vertical direction. Closer inspection of directional variograms in the horizontal plane indicated no significant anisotropy. Therefore, exponential model estimates for $\sigma^2_{lnk\ residuals}$, $\lambda_h{}'$, and $\lambda_v{}'$, of 2.8, 3.5, and 0.85, respectively, were calculated from fitting the empirical variograms of third order detrended data. These estimates were applied to a three-dimensional ordinary kriging algorithm of GSLIB with a defined grid system and search pattern resulting in 121,647 estimates.

Ideally, these 121,647 hydraulic conductivity estimates provide a realistic map as inputs to a groundwater flow or contaminant transport model; however, it must be noted that ordinary kriging, although a powerful tool, is not a "black box" technique to apply arbitrarily. For example, the grid system employed in this particular estimation process was defined to be compatible with an experimental grid system defined by researchers at Tyndall AFB. The grid definition considered that water table fluctuations allowed the saturated thickness of the aquifer to periodically extend to an elevation of 65m.

However, at the time of sampling for hydraulic conductivity data, the water table constrained data collection to an elevation approximately 62m or less. Consequently, if the distance between a small number of sample values (e.g. elevation 62m) and an estimate location (e.g. elevation 65 m) is beyond the correlation scale of the vertical variogram model (0.85 m), a large-scale smoothing effect becomes evident and the estimates reflect the mean of the sampling distribution. In addition, if estimate locations on the grid did not coincide exactly with sample locations, actual sample values may not appear on the grid. In particular, the maximum sampled K value was 1.446 cm/s where the maximum ordinary kriging estimate for K is 0.55 cm/s resulting from the same smoothing effect. Therefore, extreme values are most likely to fall prey to the smoothing effect.

From inspection of the estimated K field maps it is obvious the meander channel had a significant impact in the estimation process. There are two possible reasons for this phenomena:

(1) Polynomial detrending and reinsertion "overwhelmed" the estimation process, and/or

(2) Many estimates were a simple average of the sample points captured by the search ellipsoid. This occured because many samples were located outside of the correlation scale and the range of the modeled exponential variograms from points being estimated.

For the ordinary kriging algorithm to provide an estimate at every grid location for this pre-defined grid, the definitions of the search ellipsoid radii in Figure III.10 were forced large enough to accommodate the grid dimensions. By forcing the vertical (minor) radius to be 5m, there were many instances during the search algorithm that considered a over 80% of the aquifer's thickness when constructing an estimate at a grid location. Horizontal layering of an aquifer often creates pronounced anisotropy with the horizontal correlation much larger than the vertical one (Dagan, 1984). With sampling locations in the vertical direction tenths of meters apart, the search ellipsoid captured mostly vertical component data in spite of the larger major horizontal radii of 100m. Therefore, a significantly smaller vertical radius should be employed in order to consider a larger proportion of sample values in the horizontal direction. If the ordinary kriging estimation grid were more compatible with the sampling program (i.e. confined to elevations less than 62m), a search ellipsoid with smaller major radii than 100m and significantly minor radius than 5m could be employed. Also, a grid system with spacing defined at 4.0m in the x- and y- directions and 0.75m in the z-direction would satisfy correlation scale constraints and increase computational speed.

Recommendations

The methodology demonstrated in this thesis lays a general base for future research. Therefore, primary emphasis for future research should be focused in the following areas:

a. Evaluate various methods of detrending data. This would provide residuals with smaller variability for structural analysis and kriging.

b. Evaluate results of deterministic models such as MODFLOW with the ordinary kriging estimates obtained in Chapter III and a weighted inverse distance technique. Researchers at Tyndall AFB employed a weighted inverse distance methodology on the MADE hydraulic conductivity data on the pre-defined grid.

c. Evaluate other geostatistical estimating techniques such as indicator kriging, cokriging, and simulated annealing on the MADE hydraulic conductivity data set, and evaluate results of deterministic models using these techniques.

d. Evaluate the use of the covariance parameter and the uncertainty associated with the covariance parameter estimates for characterizing macrodispersion using stochastic transport theories.

e. Evaluate various approaches to fitting positive-definite models to variograms.

Appendix A - Definitions

Anisotropy: the directional quality used to describe the measurement of an aquifer characteristic from some representative sample as not equal in all directions. (Domenico and Schwartz, 1990)

Aquifer: a geologic unit that contains sufficient saturated porous medium to transmit significant quantities of water.

Geostatistics: a collection of techniques for the solution of estimation problems involving spatial variables.

Hydraulic Conductivity: the ability of a porous medium to transmit water.

Heterogeneous: a description used to define hydrogeologic properties that vary spatially.

Homogeneous: a description used to define hydrogeologic properties that are identical spatially.

Isotropy: the directional quality used to describe the measurement of an aquifer characteristic from some representative sample as the same in all directions. (Domenico and Schwartz, 1990)

Random Variable: a variable whose values are randomly generated according to some

probabilistic mechanism (Isaaks and Srivastava, 1989)

Range: The distance at which the variogram reaches its plateau (Isaaks and Srivastava,

1989).

Sill: The plateau the variogram reaches at its range (Isaaks and Srivastava, 1989)

Variogram: Graphical representation of how the variance of a random variable changes as

a function of distance and direction (Isaaks and Srivastava, 1989).

Appendix B - Trend Removal

The detrending of the log hydraulic conductivity data was performed by multiple linear regression. The following polynomial equation represented the log K trend:

$$\log_{10} K \; \text{trend}(x,y,z) = a_0 + a_1 x + a_2 y + a_3 z + a_4 x^2 + a_5 xy$$
$$+ a_6 xz + a_7 y^2 + a_8 yz + a_9 z^2 + a_{10} x^3 + a_{11} x^2 y + a_{12} x^2 z$$
$$+ a_{13} xy^2 + a_{14} xyz + a_{15} xz^2 + a_{16} y^3 + a_{17} y^2 z + a_{18} yz^2 + a_{19} z^3$$

where x, y, and z are normalized spatial coordinates and a_i are coefficients. Normalization of the X, Y, and Z spatial coordinates corresponding to hydraulic conductivity measurements was of the following form:

$$x = \frac{X - \min(X)}{\max(X) - \min(X)} \qquad y = \frac{Y - \min(Y)}{\max(Y) - \min(Y)} \qquad z = \frac{Z - \min(Z)}{\max(Z) - \min(Z)}$$

Coefficients for trends of order 1, 2, and 3 are as follows:

Trend order 1

$$a_0 = -3.09590 \qquad a_1 = 0.08575$$

$$a_2 = 0.34176 \qquad a_3 = 1.59131$$

Trend order 2

$a_0 = -4.08101$ $a_1 = 4.20151$ $a_2 = 4.11785$

$a_3 = 0.59034$ $a_4 = -5.06850$ $a_5 = 1.96772$

$a_6 = -1.33913$ $a_7 = -5.20048$

$a_8 = 1.42269$ $a_9 = 0.41870$

Trend order 3

$a_0 = -3.27318$ $a_1 = 4.55202$ $a_2 = 4.66162$

$a_3 = -6.36161$ $a_4 = -7.65759$ $a_5 = -9.12083$

$a_6 = 10.1255$ $a_7 = -5.04245$ $a_8 = 2.97014$

$a_9 = 12.2957$ $a_{10} = 4.96045$ $a_{11} = -1.30882$

$a_{12} = -6.19003$ $a_{13} = 8.94255$ $a_{14} = 9.08261$

$a_{15} = -11.0601$ $a_{16} = 1.29871$ $a_{17} = -9.69249$

$a_{18} = 5.93121$ $a_{19} = -7.67433$

Appendix C - Horizontal Planar Views of the Kriged Hydraulic Conductivity Field

Elevation Z = 54.0m   Elevation Z = 54.5m   Legend K (cm/s)

5×10^{-3}

5×10^{-4}

5×10^{-5}

N
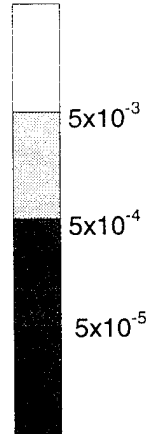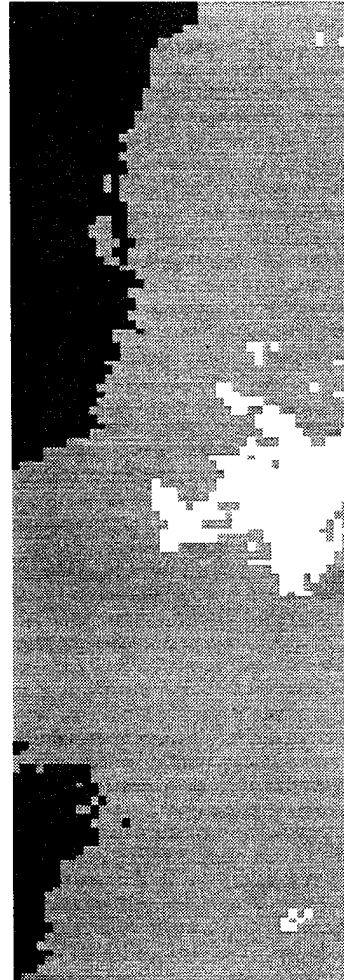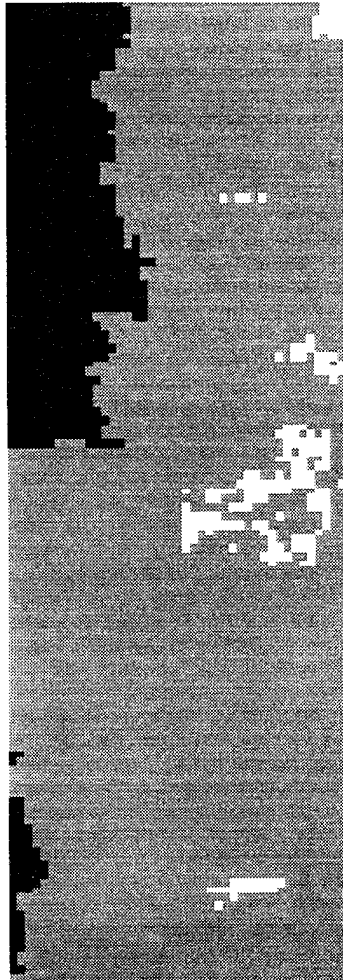
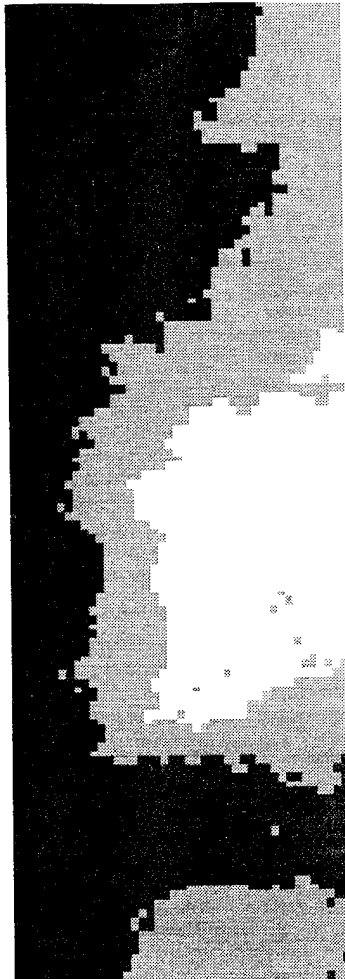Elevation Z = 55.0m   Elevation Z = 55.5m
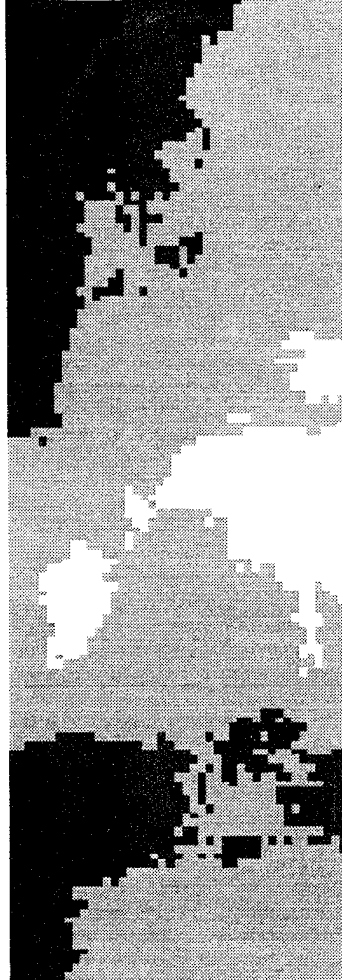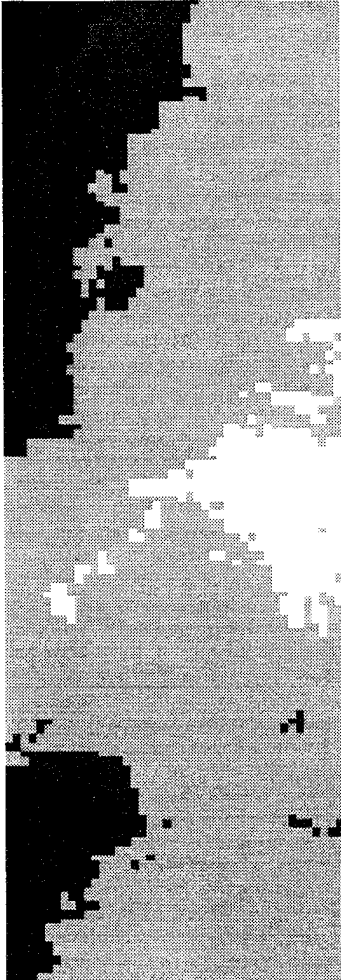
Elevation Z = 56.0m   Elevation Z = 56.5m
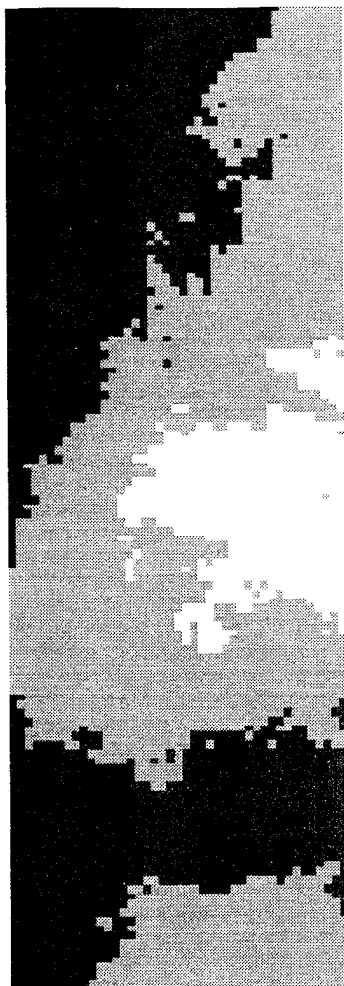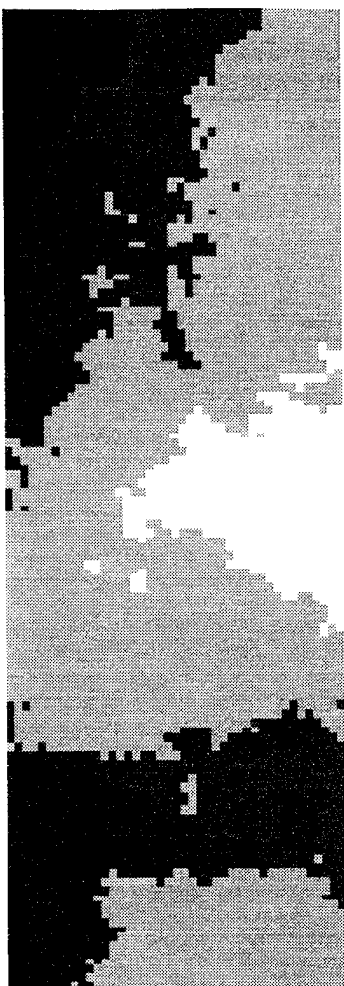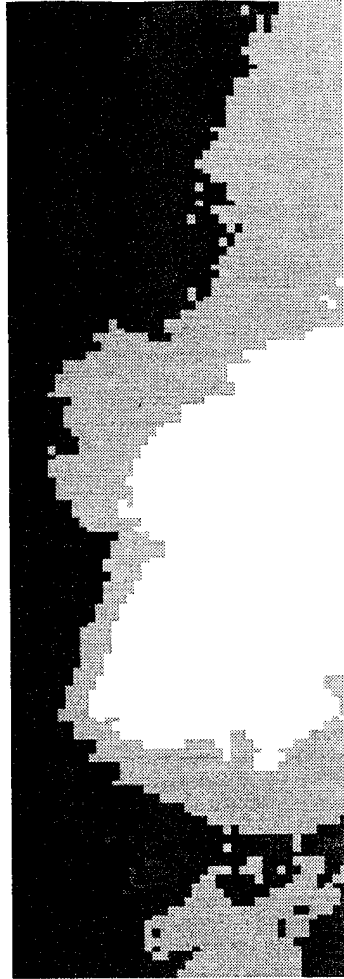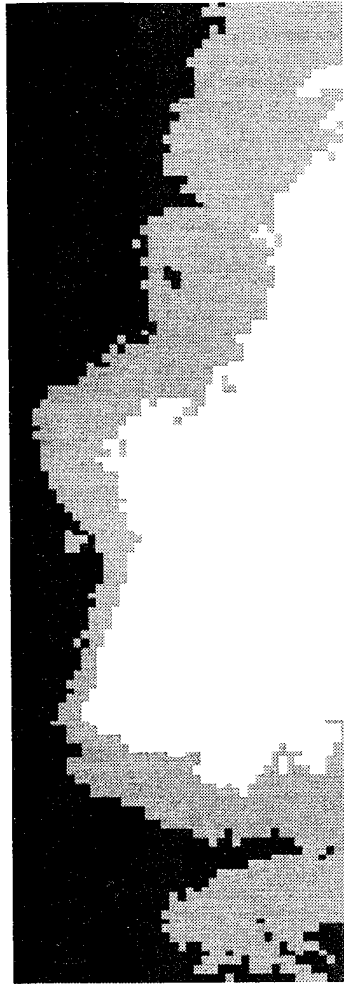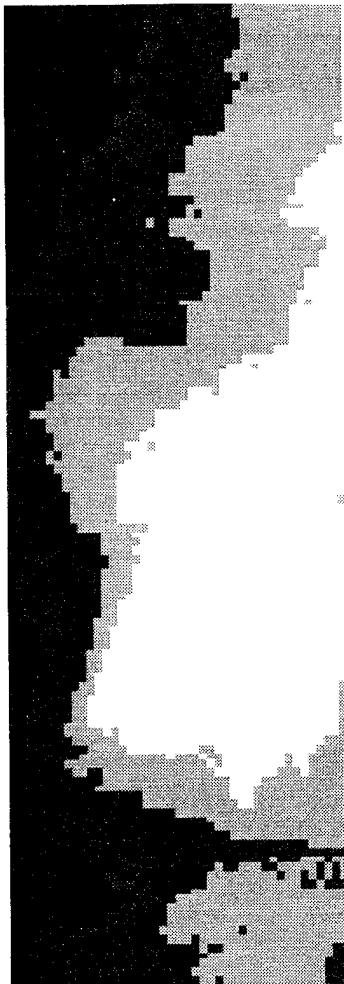
Elevation Z = 57.0m   Elevation Z = 57.5m

Elevation Z = 58.0m   Elevation Z = 58.5m
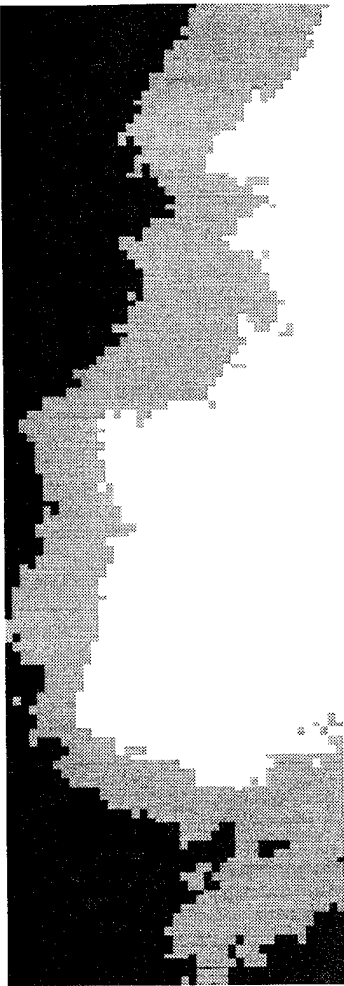
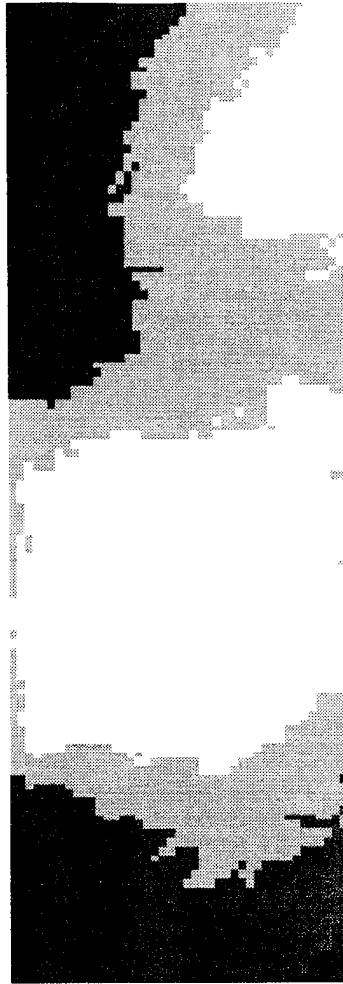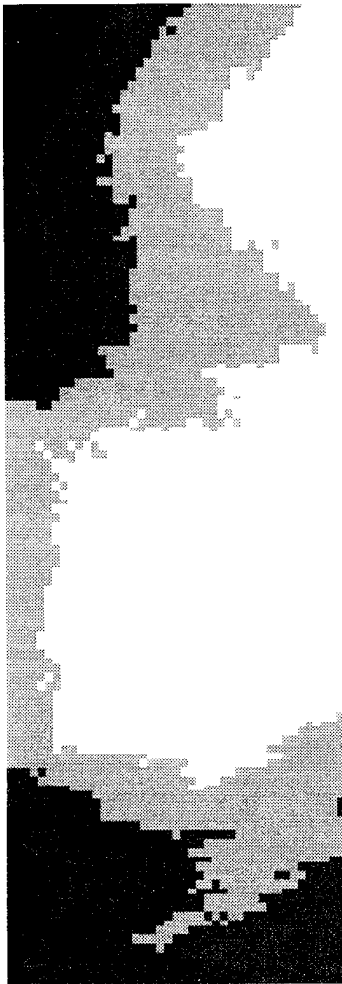Elevation Z = 59.0m   Elevation Z = 59.5m

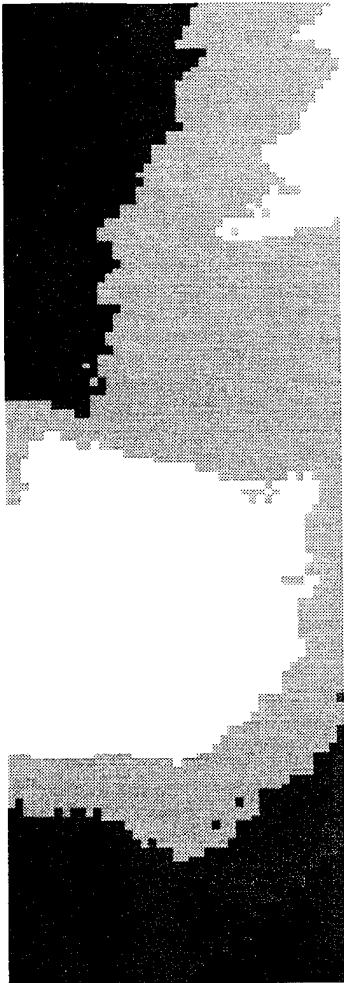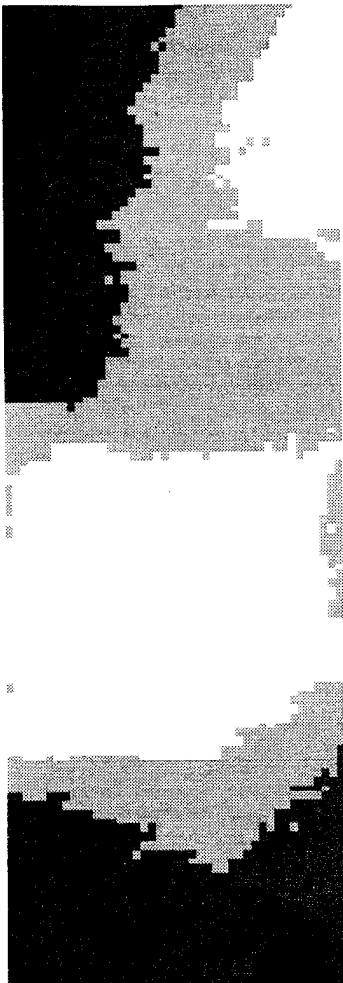Elevation Z = 60.0m   Elevation Z = 60.5m

Elevation Z = 61.0m   Elevation Z = 61.5m
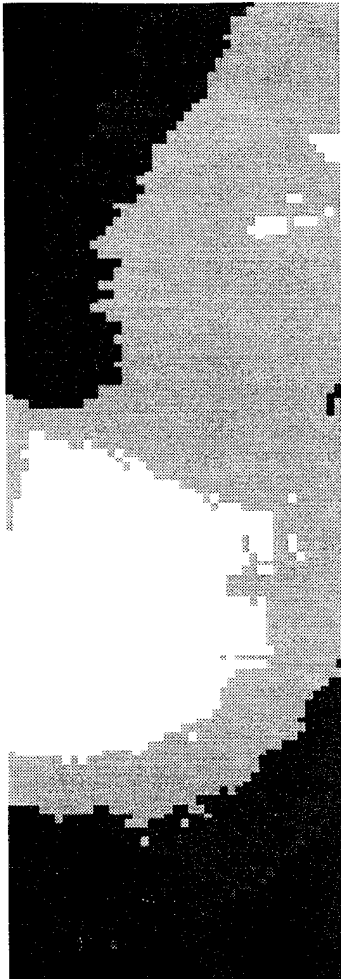
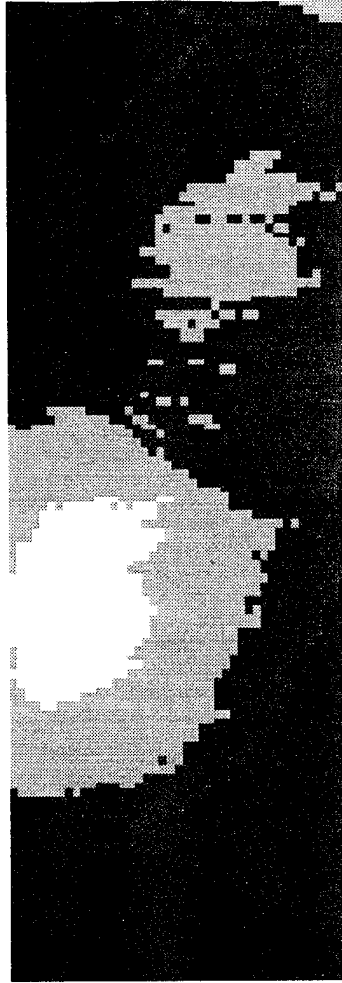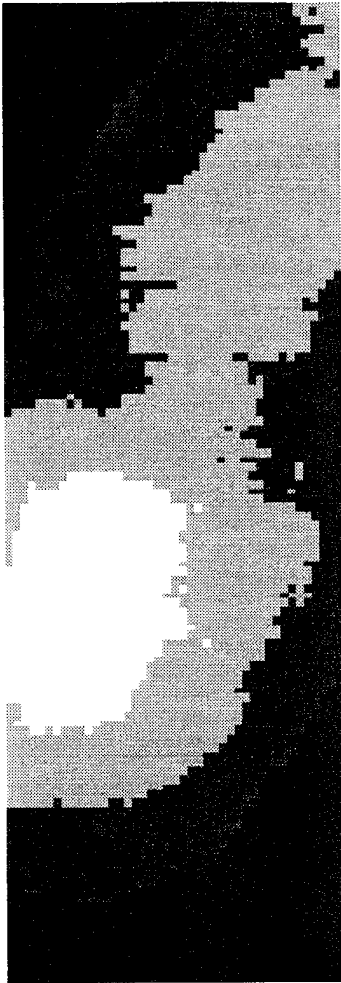Elevation Z = 62.0m   Elevation Z = 62.5m

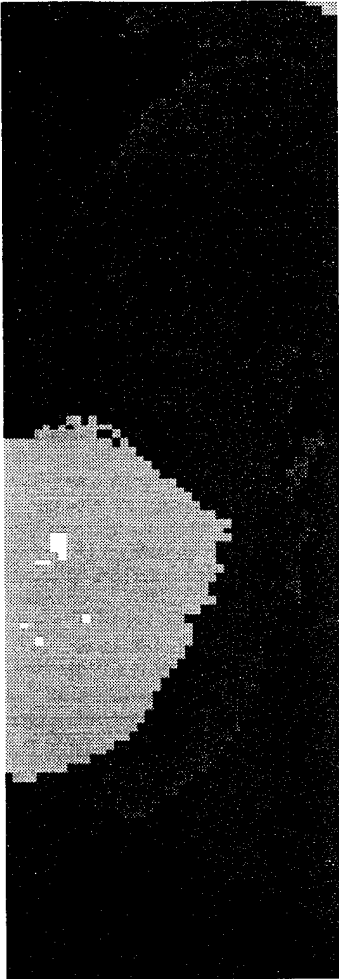Elevation Z = 63.0m   Elevation Z = 63.5m

Elevation Z = 64.0m   Elevation Z = 64.5m

Elevation Z = 65.0m

Bibliography

ASCE Task Committee on Geostatistcal Techniques in Geohydrology of the Ground Water Hydrology Committee of the ASCE Hydraulics Division, Review of Geostatistics in Geohydrology, I: Basic Concepts, Journal of Hydraulic Engineering, Vol. 116, No. 5, May 1990.

Boggs, J. Mark, Steven C. Young, Lisa M. Beard, Lynn W. Gelhar, Kenneth R. Rehfeldt, and E. Eric Adams, Field Study of Dispersion in a Heterogeneous Aquifer, 1, Overview and Site Description, Water Resources Research, Vol. 28, No. 12, pp. 3281-3291, December 1992.

Bredehoeft, J.D., Hazardous Waste Remediation: A 21st Century Problem, Groundwater Monitoring Review, pp. 95 - 100, 1994.

Dagan, Gedeon, Solute transport in hetrogeneous porous formations, Journel of Fluid Mechanics, No. 145, pp. 151 - 177, 1984.

Dagan, Gedeon, Theory of solute transport by groundwater, Annual Review Fluid Mechanics, Vol. 19, pp 183 - 215, 1987.

Darcy, H.P.G., Les fontaines publiques de la Ville de Dijon: Victor Dalmont, Paris, 1856.

Desbarats, A.J. and R.M. Srivastava, Geostatistical Characterization of Groundwater Flow Parameters in a Simulated Aquifer, Water Resources Research, Vol. 27, No. 5, pp. 687-698, May 1991.

Deutsch, Clayton V. and Andre' G. Journel, GSLIB Geostatistical Software Library and User's Guide, Oxford University Press, New York, 1992.

Freeze, R.A., A stochastic-conceptual analysis of one-dimensional groundwater flow in nonuniform homogeneous media, Water Resources Research, Vol. 11, No. 5, pp 725-742, 1975.

Gelhar, Lynn W. and Carl L. Axness, Three-Dimensional Stochastic Analysis of Macrodispersion in Aquifers, Water Resources Research, Vol. 19, No. 1, pp. 161-180, February 1983.

Gelhar, Lynn W., Stochastic Subsurface Hydrology From Theory to Applications, Water Resources Research, Vol. 22, No. 9, pp. 135S-145S, August 1986.

Gilbert, Richard O., Statistical Methods for Environmental Pollution Monitoring, Van Nostrand Reinhold, New York, 1987.

Gilliom, R.J., R.M. Hirsch, and E.J. Gilroy, Effect of Censoring Trace-Level Water-Quality Data on Trend Detection Capability, Environmental Science and Technology, Vol. 18, pp. 530-535, 1984.

Hess, Kathryn M., Steven H. Wolf, and Michael A. Celia, Large-Scale Natural Gradient Tracer Test in Sand and Gravel, Cape Cod, Massachusetts, 3, Hydraulic Conductivity Variability and Calculated Macrodispersivities, Water Resources Research, Vol. 28, No. 8, pp. 2011-2027, August 1992.

Hufscmied, P., Estimation of three-dimensional statistically anisotropic hydraulic conductivity fields by means of single well pumping tests combined with flowmeter measurements, Hydrogologie, No.2, pp. 163-174, 1986.

Isaaks, Edward H. and R. Mohan Srivastava, An Introduction to Applied Geostatistics, Oxford University Press, New York, 1989.

Jussel, Peter, Fritz Stauffer, and Themistocles Dracos, Transport modeling in heterogeneous aquifers, 1, Statistical description and numerical generation of gravel deposits, Water Resources Research, Vol. 30, No. 6, pp 1803-1817, 1994.

Kushner, E.J., On Determining the Statistical Parameters for Pollution Concentration from a Truncated Data Set, Atmospheric Environment, Vol 10, pp. 975 - 979, 1976.

Jensen, Lawrence J. "safe Drinking Water Act," in Environmental Law Handbook, Twelfth Edition, Rockville MD: Government Institutes, Inc., 1993.

Journel, A.G. and F.G. Alabert, Focusing on Spatial Connectivity of Extreme-Valued Attributes: Stochastic Indicator Models of Reservoir Heterogeneities, Society of Petroleum Engineers, SPE 18324, 63d Annual Technical Conference and Exhibition of the Society of Petroleum Engineers, Houston TX, October 1988.

Neuman, Shlomo P., Statistical Characterization of Aquifer Heterogeneities: An Overview, Recent Trends in Hydrogeology, Special Paper 189, Geological Society of America, Boulder CO,1982.

Olsen, Roger L. and Michael C. Kavanaugh, Can Groundwater Restoration Be Achieved?, Water Environment Technology, Vol. 5, pp. 42-47, March 1993.

Rehfeldt, K.R., P. Hufschmied, L.W. Gelhar, and M.E. Schaefer, Measuring Hydraulic Conductivity With the Borehole Flowmeter, Electrical Power Research Institute Topical Report, EN-6511, September 1989.

Rehfeldt, Kenneth R., J. Mark Boggs, and Lynn W. Gelhar, Field Study of Dispersion in a Heterogeneous Aquifer, 3, Geostatistical Analysis of Hydraulic Conductivity, Water Resources Research, Vol. 28, No. 12, pp. 3309-3324, December 1992.

Ritzi, R.W., Jr. and D.F. Dominic, Evaluating Uncertainties in Modeling Flow And Transport in Heterogeneous Buried-Valley Aquifers, Proceedings of the Ground Water Modeling Conference, 1993.

Russo, D. and W. Jury, A Theoretical Study of the Estimation of the Correlation Scale in Spatially Variable Fields, 1, Stationary Fields, Water Resources Research, Vol. 23, No. 7, pp. 1257-1268, 1987 a.

Russo, D. and W. Jury, A Theoretical Study of the Estimation of the Correlation Scale in Spatially Variable Fields, 1, Nonstationary Fields, Water Resources Research, Vol. 23, No. 7, pp. 1257-1268, 1987 b.

Stauffer, T.B., and V.S. Manoranjan, The Use of Grain-Size Analysis with Limited Field Data to Study Hydraulic Conductivity Variability, Aviation, Space, and Environmental Medicine, May 1994.

Woodbury, Allan D., E.A. Sudicky, The Geostatistical Characteristics of the Borden Aquifer, Water Resources Research, Vol. 27, No. 4, pp. 533-546, April 1991.

Vita

Captain Craig S. Biondo was born on 5 December 1966 in Thibodaux, Louisiana. He graduated from Louisiana State University with a Bachelor of Science in Electrical Engineering in 1989. Capt Biondo worked as an Electrical Design Engineer and Officer-In-Charge, Utilities for the 23d Civil Engineering Squadron at England Air Force Base from 1989 to 1992. He then worked as Chief, Simplified Acquisition for Base Engineer Requirements (SABER) and Design Team Chief for the 394th Civil Engineering Squadron at Lackland Air Force Base from 1992 to 1994. Captain Biondo entered the Graduate Engineering and Environmental Management Program, School of Engineering, Air Force Institute of Technology in May 1994. Capt Biondo's immediate family includes wife Renee' Bourgeois Biondo and children Ryan Paul and Stefan John.

Permanent Address: 613 Magnolia St, Thibodaux, Louisiana 70301

REPORT DOCUMENTATION PAGE

Form Approved
OMB No. 0704-0188

Public reporting burden for this collection of information is estimated to average 1 hour per response, including the time for reviewing instructions, searching existing data sources, gathering and maintaining the data needed, and completing and reviewing the collection of information. Send comments regarding this burden estimate or any other aspect of this collection of information, including suggestions for reducing this burden. to Washington Headquarters Services, Directorate for Information Operations and Reports, 1215 Jefferson Davis Highway, Suite 1204, Arlington, VA 22202-4302, and to the Office of Management and Budget, Paperwork Reduction Project (0704-0188), Washington, DC 20503.

1. AGENCY USE ONLY *(Leave blank)*	2. REPORT DATE	3. REPORT TYPE AND DATES COVERED
	December 1995	Master's Thesis

4. TITLE AND SUBTITLE

GEOSTATISTICAL ANALYSIS OF HYDRAULIC CONDUCTIVITY IN HETEROGENEOUS AQUIFERS

5. FUNDING NUMBERS

6. AUTHOR(S)

Craig S. Biondo, Capt, USAF

7. PERFORMING ORGANIZATION NAME(S) AND ADDRESS(ES)

Air Force Institute of Technology,
WPAFB OH 45433-6583

8. PERFORMING ORGANIZATION REPORT NUMBER

AFIT/ENP/GEE/95D-01

9. SPONSORING / MONITORING AGENCY NAME(S) AND ADDRESS(ES)

10. SPONSORING / MONITORING AGENCY REPORT NUMBER

11. SUPPLEMENTARY NOTES

12a. DISTRIBUTION / AVAILABILITY STATEMENT

Approved for public release; distribution unlimited

12b. DISTRIBUTION CODE

13. ABSTRACT *(Maximum 200 words)*

Observations of the spatial variability of hydraulic conductivity of the heterogeneous alluvial aquifer at Columbus Air Force Base in Mississippi are analyzed using parametric geostatistical approaches. Field studies have revealed that the heterogeneity of the aquifer hydraulic conductivity field controls the movement and dispersion of groundwater solutes (Rehfeldt, et al., 1992). Therefore, a means of quantifying spatial variability is essential for the application of flow and solute transport models to practical problems. Application of these models requires a large number of hydraulic conductivity measurements. Geostatistical analysis and kriging estimation procedures assist in providing these large numbers of values when sampling designs have provided sparse data.

The purpose of this research is to demonstrate a practical methodology for characterizing hydraulic conductivity variability in heterogeneous aquifers. Using kriging estimation procedures, provide estimations of a hydraulic conductivity field for deterministic groundwater flow.

14. SUBJECT TERMS	15. NUMBER OF PAGES
Geostatistics, Groundwater, Aquifer, Heterogeneity, Kriging	99
	16. PRICE CODE

17. SECURITY CLASSIFICATION OF REPORT	18. SECURITY CLASSIFICATION OF THIS PAGE	19. SECURITY CLASSIFICATION OF ABSTRACT	20. LIMITATION OF ABSTRACT
Unclassified	Unclassified	Unclassified	UL