USGS
science for a changing world

OVERVIEW AND TECHNICAL AND PRACTICAL ASPECTS FOR USE OF GEOSTATISTICS IN HAZARDOUS-, TOXIC-, AND RADIOACTIVE-WASTE-SITE INVESTIGATIONS

U.S. GEOLOGICAL SURVEY

Water-Resources Investigations Report 98–4145

Prepared in cooperation with the
U.S. ARMY CORPS OF ENGINEERS

Overview and Technical and Practical Aspects for Use of Geostatistics in Hazardous-, Toxic-, and Radioactive-Waste-Site Investigations

By C.R. Bossong, M.R. Karlinger, B.M. Troutman, and A.V. Vecchia

---

U.S. GEOLOGICAL SURVEY

Water-Resources Investigations Report 98–4145

Prepared in cooperation with the
U.S. ARMY CORPS OF ENGINEERS

Denver, Colorado
1999

U.S. DEPARTMENT OF THE INTERIOR
BRUCE BABBITT, Secretary

U.S. GEOLOGICAL SURVEY
Charles G. Groat, Director

For additional information write to:

District Chief
U.S. Geological Survey
Box 25046, Mail Stop 415
Denver Federal Center
Denver, CO 80225–0046

Copies of this report can be purchased from:

U.S. Geological Survey
Information Services
Box 25286
Federal Center
Denver, CO 80225

CONTENTS

FIGURES

TABLES

CONVERSION FACTORS, VERTICAL DATUM, AND ABBREVIATIONS

Multiply	By	To obtain
	Length	
kilometer	0.6214	mile
meter	1.094	yard

NOTATION

a—Angle for directional variogram.

c—Generic constant used for cutoff value in probability distribution or indicator transformation.

e—Kriging error.

$\tilde{e}$—Reduced kriging error.

f_j—Explanatory variables used in drift equations.

g—Nugget of variogram.

h—Lag or distance between two data points.

n—Number of data points.

m—Number of locations in a given block.

r—Range of variogram.

s—Sill of variogram.

w—Weight.

$\underline{x} = (u, v)$—Location based on coordinates u and v.

$z(\underline{x})$—Measurement of Z at location $\underline{x}$.

$\hat{z}(\underline{x})$—Kriging estimate using measured data.

A—Area of triangle.

B—Area designation in block kriging.

C—Population covariance function.

$\hat{C}$—Sample covariance function.

$C(\underline{x}_1, \underline{x}_2)$—Covariance of data values at locations $\underline{x}_1$ and $\underline{x}_2$.

$D_{i,j}$—Difference in values between data points i and j.

E—Expectation.

$I(.)$—Indicator function.

K—Number of variogram bins.

$N(.)$—Number of squared differences in variogram bin.

P—Probability.

S_n^2—Sample variance of n measurements.

T—Transformation.

V—Voronoi polygon.

Var—Population variance.

$W(\underline{x})$—Co-kriging random variable at location $\underline{x}$.

$Y(\underline{x})$—Transformed variable at location $\underline{x}$.

$\hat{Y}(\underline{x})$—Predictor or estimate of Y at location $\underline{x}$, obtained from kriging.

Z—Regionalized random variable.

$Z(\underline{x})$—Potential value of Z at location $\underline{x}$.

$\hat{Z}(\underline{x})$—Predictor or estimate of Z at location $\underline{x}$, obtained from kriging.

$Z^*(\underline{x})$—The residuals of $Z(\underline{x})$.

$\tilde{Z}(\underline{x})$—Arbitrary predictor of Z at location $\underline{x}$.

$\bar{Z}_n$—Sample mean of Z from n observations.

β—Regression coefficient used in polynomial representation for drift.

$\hat{\gamma}$—Sample variogram.

γ—Theoretical variogram.

$\gamma(h)$—Theoretical variogram for lag h.

λ—Optimization coefficient.

η—Parameter used in spline analysis.

$\rho(h)$—Correlation function as function of h.

$\sigma(\underline{x})$—Spatial population standard deviation at location $\underline{x}$.

$\sigma^2(\underline{x})$—Spatial population variance at location $\underline{x}$.

$\sigma_K(\underline{x})$—Kriging standard deviation at location $\underline{x}$.

$\sigma_K^2(\underline{x})$—Kriging variance at location $\underline{x}$.

$\mu(\underline{x})$—Spatial population mean of Z at location $\underline{x}$.

Overview and Technical and Practical Aspects for Use of Geostatistics in Hazardous-, Toxic-, and Radioactive-Waste-Site Investigations

By C.R. Bossong, M.R. Karlinger, B.M. Troutman, *and* A.V. Vecchia

Abstract

Technical and practical aspects of applying geostatistics are developed for individuals involved in investigations at hazardous-, toxic-, and radioactive-waste sites. Important geostatistical concepts, such as variograms and ordinary, universal, and indicator kriging, are described in general terms for introductory purposes and in more detail for practical applications.

Variogram modeling using measured ground-water elevation data is described in detail to illustrate principles of stationarity, anisotropy, transformations, and cross validation. Several examples of kriging applications are described using ground-water-level elevations, bedrock elevations, and ground-water-quality data.

A review of contemporary literature and selected public domain software associated with geostatistics also is provided, as is a discussion of alternative methods for spatial modeling, including inverse distance weighting, triangulation, splines, trend-surface analysis, and simulation.

1.0 INTRODUCTION

This report addresses the use of geostatistics at hazardous-, toxic-, and radioactive-waste (HTRW) site investigations. The report was prepared in cooperation with the U.S. Army Corps of Engineers (USACE) for use as a guidance document within the USACE. The USACE has distributed the report as an Engineer Technical Letter within their agency (USACE, 1997). One very fundamental aspect of perhaps all HTRW-site investigations that deal with environmental contamination is the need to characterize the extent and spatial distribution of contamination. Such a characterization usually includes describing and evaluating the spatial trends and variability of the contamination, using a variety of statistical or analytical tools. A principal difficulty in characterizing the contamination is the fact that measurements might be few or might be sparsely scattered over large regions. Another difficulty that arises naturally is how to interpolate between measured data in order to make predictions (or estimates) at points where measurements of contaminant concentration are not available. Such interpolation is referred to as point, or punctual, estimation in this report. Additionally, an investigator might need to determine a single representative value for an area that has several measured or estimated values, or both; this determination is referred to in this report as block estimation. Geostatistics is a set of statistical procedures designed to address these difficulties and needs. Geostatistics can be applied to many other problems, besides contamination, that occur at HTRW sites. Even though this report addresses only two-dimensional applications, geostatistics can be used in three dimensions as well. Indeed, there are many cases in which the third dimension, usually stratification, is desirable to address.

Kriging is the principal geostatistical technique described in this report. For introductory purposes, kriging can be defined as a technique for determining the optimal weighting of measurements at measured or sampled locations for obtaining predictions, or

estimates, at unmeasured or unsampled locations; additional definition of kriging is provided throughout this report. Kriging is well suited for making point and block estimates; however, much of the advantage of using geostatistical techniques, such as kriging, is not just in the point and block estimates but in the information provided concerning the uncertainty associated with the estimates. The uncertainty information is usually quantified by a kriging variance that is associated with a kriging estimate. The uncertainty also is sometimes referred to as the kriging standard deviation, which is simply the square root of the kriging variance.

Original geostatistical work involved making estimates for the areal extent and concentrations of economic mineral deposits in relation to mining. Today (1998), geostatistical techniques continue to have a function in mining. However, a well-developed method that is capable of interpolating a given set of measured values at discrete locations into estimates for new locations or developing an individual estimate for an area including many locations, or both, has attracted users from many disciplines, and there is a trend toward incorporating geostatistics as standard curriculum for most geoscience educational programs. The use of geostatistical techniques as part of HTRW-site investigations is becoming common because of the almost routine need for data interpolation as part of these investigations.

Once investigators have established that the data are adequate as to quality and quantity, geostatistics can be a powerful analytical tool that results in quantitative characterization of areas of special interest within the study area or the entire study area. These characterizations could be used to determine spatial variation; for example, where concentrations of contaminants in soils are relatively high or low, are less than or greater than a specified concentration, or even have a high or low probability of exceeding a certain concentration.

1.1 Purpose and Scope

The purpose of this report is to address the use of geostatistics in HTRW-site investigations by presenting an overview of geostatistical methods and discussing their technical and practical aspects. The report also includes a brief literature and software review, a presentation of kriging applications, a discussion of the review of kriging applications, and a discussion of more advanced geostatistical techniques, such as conditional simulation.

The scope of this report is limited to discussions and examples of two-dimensional point and block estimations using a geostatistical method known as kriging. The technical aspects of geostatistics are presented through discussion of the assumptions about, and the mechanics of, several types of kriging, including ordinary kriging, which is applicable when the mean for the variable of interest is constant over the region of interest, and universal kriging, which is applicable when the mean for the variable of interest changes gradually over the region. A specialized form of kriging known as indicator kriging and the use of information concerning uncertainty associated with kriging estimates also are discussed. The fundamental concepts of geostatistical kriging theory are discussed, but various references are provided for more detailed information.

1.2 Organization

This report is divided into eight sections described here.

- Section 1.0 presents introductory material and includes an overview of the use of geostatistics at HTRW-sites.
- Section 2.0 presents an overview of some important technical aspects of geostatistics with a minimum of theory and equations.
- Section 3.0 discusses the assumptions and theory behind kriging, including equations and concepts that are useful for obtaining a better understanding of the technical aspects, or mathematics, of kriging interpolation. Many of the concepts developed in section 3.0 are discussed in general terms in section 2.0, so those readers desiring only an overview of kriging concepts may wish to read only section 2.0 and bypass section 3.0.
- Section 4.0 reviews texts that contain much more detailed information regarding kriging theory than material included in section 3.0. Section 4.0 also provides a brief generic discussion of kriging software.

- Section 5.0 discusses detailed step-by-step variogram construction and demonstrates some pitfalls and solutions to this crucial process. Section 5.0 also discusses techniques that investigators may use to evaluate their variograms.
- Section 6.0 discusses practical aspects of geostatistics by presenting several examples of kriging applications using data from the HTRW field. The examples illustrate a few of the many different ways kriging can be used in HTRW-site investigations and are not presented with the same level of detail used in section 5.0.
- Section 7.0 provides additional detail on some crucial aspects of kriging applications and includes considerations that may be helpful to determine if kriging is feasible for the intended use.
- Section 8.0 briefly discusses other methods for spatial modeling and also includes discussion of advanced stochastic methods, such as simulation.

1.3 Overview of the Use of Geostatistics in Hazardous-, Toxic-, and Radioactive-Waste-Site Investigations

Investigations of HTRW sites involve complex administrative, scientific, and engineering functions and are truly interdisciplinary. For instance, administrative functions that are associated with fiscal, managerial, or regulatory input can guide or constrain scientific or engineering work. Similarly, scientific or engineering findings may define the scope of the administrative effort.

Scientists and engineers involved in HTRW-site investigations have found that they have an implicit need for many disciplines to fulfill the objectives of each particular investigation. Frequently, an HTRW-site investigation will benefit from specialized information available from earth-science disciplines such as geology, hydrogeology, and geochemistry, among others. Some HTRW-site investigations are large enough to use several individuals from each of these disciplines, as well as many others, for the duration of multi-year investigations. Most disciplines associated with HTRW-site investigations will benefit from knowledge or input from specialized or interdisciplinary branches, or both; the geologist, for example, will occasionally benefit from knowledge of geophysics. Interdisciplinary input also can be very helpful, especially in geostatistics, where earth-science disciplines rely on assistance from statisticians.

1.4 Acknowledgments

Several individuals have provided valuable technical review of this report. Dave Becker, USACE, HTRW Center for Expertise, not only provided a consistent and thorough technical review, but also coordinated additional technical reviews. Additional technical review from the USACE were provided by Terry Walker and Tom Georgian, both from the HTRW Center for Expertise; Brad Call, Earl Edris, Dave Kachek, Doug Mullendore, and Kerry Walker, all from USACE field offices; and Tommiann McDaniel from USACE headquarters. Additional technical reviews from outside the USACE were provided by Evan Englund, U.S. Environmental Protection Agency National Exposure Research Laboratory; Ed Gilroy, U.S. Geological Survey (USGS); Mohan Srivastava, Froidevaux, Srivastiva, and Schofield; and Wayne Woldt, University of Nebraska.

2.0 OVERVIEW OF SOME TECHNICAL ASPECTS OF GEOSTATISTICS

This section provides an overview of some of the procedures and concepts discussed in detail in this report. Some of the technical ideas and terminology are introduced in very general terms to familiarize the reader with geostatistics.

2.1 General Considerations in Spatial Prediction

The principal consideration in this report is spatial prediction or modeling values of a spatial process; in particular, to make best use of measurements of a variable (such as pollutant concentration) at sampled locations so as to make inferences (or predictions) about that variable at unsampled locations or for the region as a whole.

A spatial process can have a large-scale or a regional component and a small-scale or a local component; both these components need to be accounted for when modeling a spatial process. The large-scale component is referred to as the mean field and is most often modeled by a spatial trend that may or may not be constant over the region. The small-scale component is a random fluctuation that is mathematically combined with the trend to make up the sample at a point. On the average, the random fluctuation is assumed to be zero, but can be either positive or negative in individual samples. The separation of the trend from the random fluctuation is problem- and scale-dependent and needs to be determined carefully. There can be several solutions to the problem of separating the trend and the random fluctuation that may be useful for various geostatistical purposes when using a single set of data.

The small-scale fluctuation of the variable of interest (for example, water levels or contaminant concentrations at a sample point), although random, can indicate some association with the random fluctuations at nearby points. This association is referred to as spatial correlation. Positive spatial correlation between measurements indicates that the random fluctuation at both points tends to have the same sign, whereas negative correlation indicates that the random components tend to have the opposite sign. The large-scale trend and the positive spatial correlation of the small-scale fluctuations contribute to measurements at locations that are close together being more closely related than are measurements at locations that are farther apart.

The most obvious procedure to determine spatial prediction at unsampled locations is simply to take an average of the measured sample values and to assume that this average value gives a reasonable prediction at all locations in the region of interest. This procedure may work adequately in some cases, but there also are pitfalls. Using a single value for an entire region implicitly assumes spatial homogeneity. This assumption ignores any spatial trends that might exist in the data and also ignores spatial continuity. If the variable of interest does have a tendency to be spatially correlated, then a weighted average rather than a simple average could be used to make a spatial prediction by giving measurements at sampled locations that are nearer to the unsampled location more weight. This motivation is the basis for the geostatistical techniques discussed in this report. The technique known as kriging is a technique for determining the optimal weighting of measurements at sampled locations for obtaining predictions at unsampled locations. These optimal weights depend on spatial trends and correlations that may be present.

There are a number of ways to perform spatial prediction. The geostatistical technique of kriging belongs to a class of techniques known as stochastic techniques. In these techniques, the measurements, actual and potential, are considered to constitute a single realization of a random (or stochastic) process. One advantage of assuming the existence of such a random process is that measures of uncertainty, such as the variance used in kriging, can be defined. These measures of uncertainty permit the objective assessment of a spatial-prediction technique on the basis of how small such measures are. Once a measure of uncertainty has been selected, the weights to be used in spatial prediction may be determined to minimize the measure of uncertainty. In short, the use of stochastic techniques provides a way of objectively quantifying errors and determining weights. In practice, spatial predictions obtained using kriging are almost always accompanied by a measure of the associated error. Such an error evaluation is an integral part of a kriging analysis and is one of the principal advantages of using kriging (or stochastic techniques in general).

Nonstochastic techniques are generally applied strictly empirically; no assumptions concerning the existence of an underlying random process are made, and no theoretical framework is used to evaluate statistically the performance or optimality of the nonstochastic techniques. When applied in such a manner, whether such techniques would be expected to yield results that are satisfactory cannot be evaluated in advance. Two techniques that are commonly applied nonstochastically are simple averaging and trend analysis, which is a least-squares method for fitting a smooth surface to the data. Even though these two techniques are usually applied nonstochastically, performance can still be assessed if a stochastic setting is assumed. Generally, simple averaging would perform well if there were no trend and no spatial correlation, and trend analysis would perform well if there was a trend that can be modeled, but no spatial correlation. Lack of correlation in the measurements is one assumption that is made in ordinary statistical regression analysis, and trend analysis, if it is placed in a stochastic setting, is actually a special type of regression. The stochastic technique of kriging explicitly incorporates the spatial correlations that are ignored in trend analysis. In section 8.0, a few other common

techniques that are usually applied nonstochastically are discussed briefly. Most of these techniques are designed to incorporate spatial continuity, but the way it is incorporated may be subjective. Use of kriging provides an objective means of incorporating partial correlation and makes the background assumptions explicit.

2.2 Important Geostatistical Concepts

This section presents some of the key concepts in geostatistics that are discussed in detail in section 3.0. The concepts are presented in about the same order as they are discussed in section 3.0.

2.2.1 Variograms

A central concept in geostatistics is the use of spatial correlation to improve spatial predictions or interpolations. The variogram is the principal tool used to characterize the degree of spatial correlation present in the data and is fundamental to kriging. The correlation between measurements at two points is usually assumed to depend on the separation between the two points. This dependence can be examined by squaring the difference between the measured values at each pair of locations and then categorizing the squared differences according to the separating distance between the paired locations. For small separations, or lags, the squared differences are usually small and increase as the lag increases. A plot of the squared differences per sample pair as a function of lag is referred to as the sample variogram.

The general behavior of the points in the sample variogram is affected by the spatial correlation between sample sites and can provide investigators with qualitative information about the spatial process, but to use this information rigorously as a basis for interpolation, a function, that has specific properties, needs to be fit to the sample variogram points. The fitting passes a smooth curve through the scattered points. The curve, which can be represented by a mathematical expression or function, is called a model. There are several models introduced in section 3.0 that have characteristic features that are commonly used in geostatistics. The variogram model is used to determine kriging weights for use in interpolation.

2.2.2 Directional Variogram and Anisotropy

Spatial correlation often depends not only on the distance between points, but also on the direction along which the points plot. For example, measurements at pairs of points that are 100 meters apart and are oriented north-south may have a different correlation than measurements at points that are the same distance apart, but that are oriented east-west. Correlations dependent on direction indicate anisotropy, and when anisotropy is present, a directional variogram needs to be used for the geostatistical analysis.

2.2.3 Kriging and Kriging Variance

Kriging yields optimal spatial estimates at points where no measurements exist using measurements at points where there are data. As discussed in section 2.1, placing an analysis in a stochastic framework enables precision in defining optimality. In kriging, a restriction that the predicted value at any point is a linear combination of the measured values is imposed first; that is, the kriging estimate is a linear predictor. Given this restriction, the values of the coefficients in this linear function are chosen to ensure the predictor to be optimal.

The first optimality criterion imposed is that the estimate be unbiased, or that on average, the difference between the predicted value and the actual value is zero. The second optimality criterion is that the variance of the predictions be minimized. The variance in the predictions is a statistical error measure defined to be the average squared difference between predicted and actual values. Because the kriging estimate minimizes this variance, the estimate, or prediction, is known as the best (minimum variance) unbiased linear predictor. This minimization is performed algebraically and results in equations known as the kriging equations, which are explicit representations of the optimal coefficients (weights) in terms of the variogram. These equations are presented in section 3.0.

An expression for the kriging variance also is discussed in section 3.0. This variance depends on the geometry of the data sites, with the variance at locations near measured points tending to be small. A variance then can be associated with any spatial prediction, which gives an indication of the uncertainty about that predicted value. The fact that kriging provides this measure of uncertainty is one of its principal advantages over many other techniques.

2.2.4 Trends and Universal Kriging

In kriging, special attention must be given to the question of whether there are spatial trends in the data. A trend is usually any detectable tendency for the measurements to change as a function of the coordinate variables, but also can be a function of other explanatory variables. For example, aside from random fluctuations, measurements of groundwater elevations may have a tendency to consistently increase in a certain direction. A kriging analysis in which there is no spatial trend is known as ordinary kriging; when a trend does exist, universal kriging should be considered. In universal kriging, the trends present are accounted for. For example, the trend might be represented as a linear function of coordinate variables. The form of the trend model then is incorporated into the universal kriging equations to obtain the optimal weights and account for the trend.

2.2.5 Block Kriging

The kriging discussed in sections 2.2.3 and 2.3.4 is usually known as point, or punctual, kriging. In point kriging, the goal is to predict the value of a variable at discrete locations. By contrast, in block kriging, the goal is to predict the average value of a variable for a specified region. As in point kriging, the optimal predictor is a linear combination of the measured values, and the degree of uncertainty is indicated by a block kriging variance. Block kriging variances tend to be smaller than point kriging variances because averages tend to be less variable than individual point values.

2.2.6 Prediction Intervals and Normality

A standard kriging analysis gives two values for any location: the optimal kriging estimate and the kriging variance. The variance provides a measure of uncertainty about the prediction. In some studies, the nature of the uncertainty needs to be specified beyond just giving the variance. One way to further specify the uncertainty is to obtain a prediction interval, which is an interval where there is a certain probability, generally 95 percent, that the actual value is within the interval. Finding such an interval often hinges on the probability distribution of the variables being sampled. An ideal situation is when the variable of interest, such as contaminant concentration, can be assumed to have a normal distribution. In this situation, and given the set of measured values, a potential value at an unsampled location has a normal distribution, with the average given by the kriging estimate and the variance given by the kriging variance. Thus, on the basis of classical statistics, the straightforward use of this normal distribution can be used to obtain a 95-percent prediction interval for a concentration at an unsampled location.

2.2.7 Transformations

A prediction interval generally is much more informative than the kriging estimate and kriging variance, so a common question is whether a normality assumption can be made for the data. When a normality assumption cannot be made, a transformation can be identified that will make the data normal, or almost normal. For example, for data that have values greater than 0, a logarithmic transformation is often tried; that is, a geostatistical analysis is performed on logarithmically transformed values rather than on the original data. Prediction intervals obtained using the transformed values can be readily converted to corresponding intervals on untransformed variables. However, there are subtleties that need to be considered in back-transforming the kriging estimate and the kriging variance; these subtleties are discussed in more detail in section 3.0.

2.2.8 Indicator Kriging

In indicator kriging, analysis is performed using indicator variables of the measured data rather than the measured data. An indicator variable is a special kind of transform of the measured data and can have one of two possible values: 0 or 1. To obtain the indicator variables to be analyzed, a threshold value is specified (c), which, for example, may represent a contaminant concentration level of particular importance. At each measurement location, the indicator variable then is assigned a value of 1 if the measured value is less than or equal to c and is assigned a value of 0 if the measured valued is greater than c. This kind of transform allows censored data, or data reported as less than some reporting limit, to be included in the analysis if the reporting limit is less than or equal to the threshold, or cutoff, value of c. After the indicator transform has been determined, the kriging analysis then is done using these indicator variables in the

usual manner; first, a variogram is obtained, and then the kriging equations yield the optimal linear predictor and the kriging variance for the indicators.

Although the indicator kriging analysis uses only 0's and 1's, the interpolated estimates are not restricted to these two values. In most analyses, the estimates are between 0 and 1, which is interpreted to be the probability that the actual value is less than or equal to the threshold c. Use of this analysis for a number of different threshold values can provide information about the probability distribution of contaminant values at a location, which may be used to obtain prediction intervals. Such prediction intervals may even be more valuable than having only the optimal predictor and variance provided by the usual kriging analysis, particularly if the behavior of extremes may be of interest. An advantage of using indicator kriging to obtain prediction intervals is that there is no need to assume a normal distribution for the data.

3.0 TECHNICAL ASPECTS OF GEOSTATISTICS

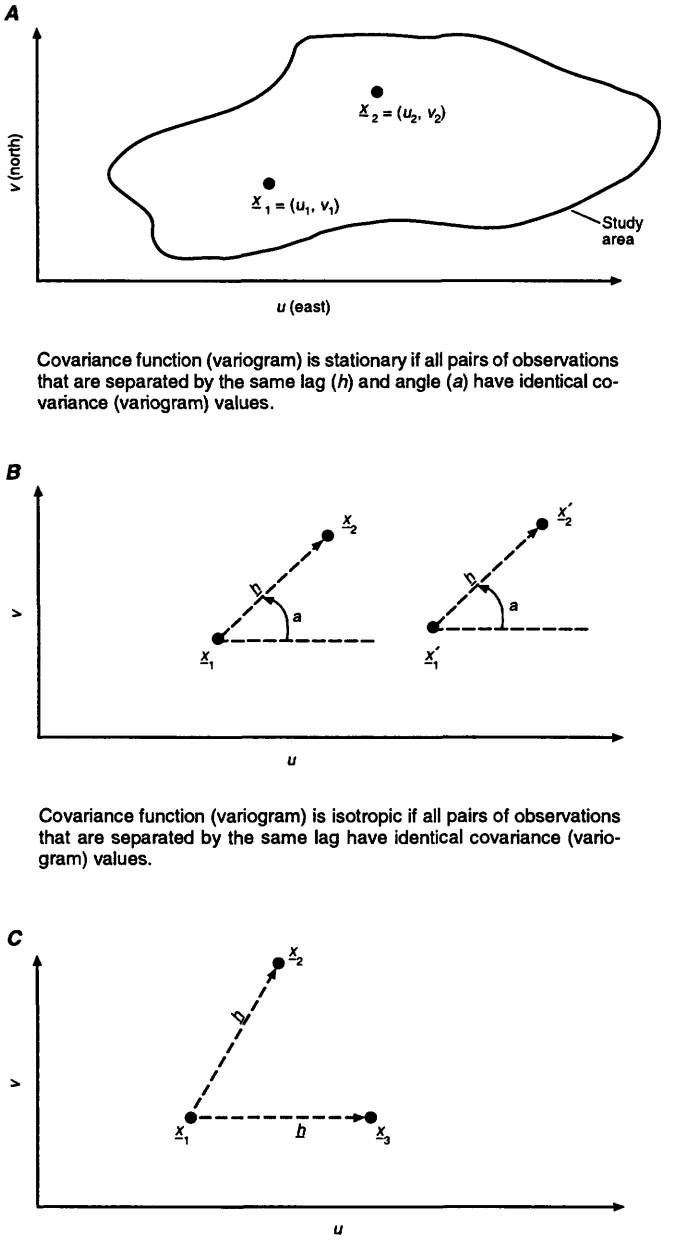
This section provides the technical aspects or the necessary theoretical background for understanding kriging applications. Emphasis is placed on presentation of the basic ideas; long formulae or derivations are minimized. Statistical terms that are commonly used in geostatistical applications are in bold text and are briefly defined as they are introduced; notation used in this report also is listed in the "Notation" section. More thorough discussions of these fundamental concepts are indicated by references cited in section 4.0. A knowledge of engineering statistics at the level of Devore (1987) and Ross (1987) would help in understanding some parts of this section. Readers who have limited statistical experience may wish to briefly scan this section and refer back to it after reading the remaining sections.

In section 3.1, regionalized random variables are discussed. Regionalized random variables constitute the random process that is sampled to obtain the observed data that are available for analysis. Basic ideas related to probability distributions, averages, variances, and correlation are introduced. In section 3.2, the variogram, which is the fundamental tool used in geostatistics to analyze spatial correlation, is introduced. In section 3.3, the use of kriging to obtain the best weights for spatial prediction is discussed, and the computation of the average-, or mean-, squared prediction error for these predictions also is discussed. In section 3.4, co-kriging, which is prediction of one variable based on measurements of that variable and other variables, is discussed. Finally, in section 3.5, the application of kriging to determine not just optimal spatial predictions, but also probabilities associated with various events, such as extreme events that may be of importance in risk-based analyses, is discussed.

3.1 Regionalized Random Variables

Suppose the extent of ground-water contamination by a particular pollutant in a given study area is being determined. To simplify the presentation, all data are assumed to be distributed over a two-dimensional region. In three-dimensional ground-water flow systems, one could study the depth-averaged concentration of a pollutant or the concentration of the pollutant in a particular horizontal stratum of the flow system. Let a vector $\underline{x} = (u,v)$ denote an arbitrary spatial location in the study area. Unless otherwise stated, u is assumed to be the east-west coordinate and v the north-south coordinate (fig. 1*A*). Use $z(\underline{x})$ to denote a measurement at location $\underline{x}$, such as the concentration of a pollutant. The ultimate goal is to determine $z(\underline{x})$ for all locations in the study area. However, without explicit knowledge of the ground-water flow and transport field, this goal cannot be achieved. Therefore, suppose that the goal is to estimate the values of $z(\underline{x})$ with a given error tolerance. For some studies, a small estimation error for some parts of the study area (for instance, near a domestic water supply) may need to be obtained, while allowing larger estimation errors in other parts of the study area. The theory of regionalized random variables is designed to accomplish these goals.

In the regionalized random-variable theory, the true measurement, $z(\underline{x})$, is assumed to be the value of a **random variable**, $Z(\underline{x})$. A random variable $Z(\underline{x})$ is associated with a true measurement $z(\underline{x})$ to characterize the degree of uncertainty in the quantity of interest at point $\underline{x}$. If there is no measurement obtained at $\underline{x}$, then the values acquired by $Z(\underline{x})$ represent potential measurements at $\underline{x}$; that is, $Z(\underline{x})$ represents possible values that might be expected if a measurement were



Hypothetical study area with two sampling locations. The coordinate axes are in the east-west (u) and north-south (v) directions.

Covariance function (variogram) is stationary if all pairs of observations that are separated by the same lag (h) and angle (a) have identical covariance (variogram) values.

Covariance function (variogram) is isotropic if all pairs of observations that are separated by the same lag have identical covariance (variogram) values.

Figure 1. Covariance function properties—*A*, hypothetical study area; *B*, stationary covariance functions; and *C*, isotropic covariance function.

obtained at $\underline{x}$. Because there is uncertainty associated with $Z(\underline{x})$, the random variable needs to be characterized by a **probability distribution**, defined by $P[Z(\underline{x}) \leq c]$, where P denotes probability and c is any constant. This distribution is a function of c and, to be completely defined, needs to be known for all values of c. The distribution is used to make certain evaluations. For example, suppose there is no measurement of the concentration of a certain contaminant at $\underline{x}$, but the distribution is known and a threshold value of $c = 8$ milligrams per liter is of interest. If $P[Z(\underline{x}) \leq 8] = 0.60$, and if a measurement were made at $\underline{x}$, there is a 60-percent chance of obtaining a value less than or equal to 8 milligrams per liter. The distribution also may be used to calculate other probabilities, such as the probability of obtaining a value in some specified interval.

An important concept in all geostatistical applications is the **support** of the regionalized random variable. The support of $Z(\underline{x})$ is the in-situ geometric unit represented by an individual sample. For example, in a soil-contamination study, $Z(\underline{x})$ might represent the concentration of a contaminant in a vertical soil core 0.1 meter in diameter and 1 meter in length and centered at location $\underline{x}$. Thus, although $Z(\underline{x})$ is defined at a particular point, it represents a volume of soil. Changing the support of $Z(\underline{x})$ usually changes its probability distribution. Therefore, all the measurements in a geostatistical analysis need to have the same support. The technique called point, or punctual, kriging, described in section 3.3, is designed to predict values of $Z(\underline{x})$ with the same support as the sample data.

A concept closely related to support is that of **estimation block**, which is a geometric unit larger than the support of a single measurement, for which a single representative value is desired. For example, in the example soil-contamination study, an estimate of the average concentration of a contaminant in a truckload of soil excavated from a block 6 meters long, 6 meters wide, and 0.3 meter thick may be necessary. Using a method called block kriging, also described in section 3.3, the block average can be predicted based on individual measurements.

Although the distribution of $Z(\underline{x})$ completely characterizes $Z(\underline{x})$ at any particular location, the distribution indicates nothing about the relations among the values of $Z(\underline{x})$ at different locations, which is very important because geostatistics is based on using a measured value of a regionalized variable at one location to gain information about values of the variable at another location. The distribution of $Z(\underline{x})$ at a single location can be readily generalized to two or more locations. For two locations, if $\underline{x}_1$ and $\underline{x}_2$ are two distinct locations, then the

joint probability distribution is defined to be the probability $P[Z(\underline{x}_1) \le c_1, Z(\underline{x}_2) \le c_2]$ for any constants c_1 and c_2. This latter probability means the probability that both $Z(\underline{x}_1) \le c_1$ and $Z(\underline{x}_2) \le c_2$. If the variables $Z(\underline{x}_1)$ and $Z(\underline{x}_2)$ are statistically independent of one another, then the joint probability distribution can be obtained as the product of the individual probability distributions,

$$P[Z(\underline{x}_1) \le c_1, Z(\underline{x}_2) \le c_2] = P[Z(\underline{x}_1) \le c_1]\ P[Z(\underline{x}_2) \le c_2]. \quad (3\text{–}1)$$

However, in most applications, $Z(\underline{x}_1)$ and $Z(\underline{x}_2)$ are not statistically independent, and their joint distribution cannot be obtained from the individual distributions. When this joint description is applied to more than two locations, specification of the full spatial distribution of Z would need the joint distribution of $Z(\underline{x}_1), \ldots, Z(\underline{x}_n)$ for any set of n spatial locations and for any n; however, except in very special cases, working with the full set of distribution functions of $Z(\underline{x})$ is not feasible and is not done.

To simplify the problem even further, various parameters of the distributions are used rather than using the entire distributions. The parameter most commonly used to characterize a distribution is the **mean**; because the mean in geostatistical applications depends on the spatial variable $\underline{x}$, the mean may be called the **spatial mean**, or the **drift**. In statistics, the mean is referred to as the **expectation** (E) of the random variable $Z(\underline{x})$, and the symbol μ is used in this report to denote this expectation. Thus,

$$\mu(\underline{x}) = E[Z(\underline{x})] \quad (3\text{–}2)$$

is used to denote the mean, or expected value, of the bracketed term, in this case $Z(\underline{x})$. Thinking of the expectation as an average can be helpful. In fact, if the distribution of $Z(\underline{x})$ assigned equal probability to a finite number of values, then the expectation of $Z(\underline{x})$ would indeed be the simple average of these numbers. However, in geostatistics, $Z(\underline{x})$ is usually assumed to take on any value in a continuous range of possible values rather than being limited to a discrete set of values. Therefore, calculus needs to be used to define the expectation. The following example illustrates the difference between averages and expectations.

3.1.1 Example 3.1.1

An experiment consists of injecting a conservative tracer at a particular well in a steady-state groundwater flow system and measuring the concentration, $Z_1(\underline{x})$, of the tracer in a neighboring well 24 hours later. The tracer then is allowed to flush from the system, and the experiment is repeated a second time to obtain another concentration measurement, $Z_2(\underline{x})$, at the same location. If this process is repeated n times, n concentration measurements $Z_1(\underline{x})$, $Z_2(\underline{x})$, ..., $Z_n(\underline{x})$ would be obtained, all at locations $\underline{x}$. The average concentration at location $\underline{x}$ is

$$\bar{Z}_n(\underline{x}) = \frac{1}{n}[Z_1(\underline{x}) + Z_2(\underline{x}) + \ldots + Z_n(\underline{x})], \quad (3\text{–}3)$$

which would change depending on n and on the actual values obtained for $Z_1(\underline{x})$, $Z_2(\underline{x})$, ..., $Z_n(\underline{x})$. However, in the limit as n increases, $Z_n(\underline{x})$ becomes closer and closer to the true mean, or expected, concentration $\mu(\underline{x})$:

$$\bar{Z}_n(\underline{x}) \to \mu(\underline{x}) \text{ as } n \text{ increases.} \quad (3\text{–}4)$$

This theoretical limit is a constant value, or **population parameter**, as opposed to $\bar{Z}_n(\underline{x})$, which is a random variable, or a property of the particular **sample** that is obtained.

In example 3.1.1, no assumptions were needed concerning whether the mean changed with spatial location because all sampling was done at one sampling location, $\underline{x}$. In most HTRW-site applications, the mean probably changes, depending on the sampling location. In addition, usually only one measurement is available at any particular location. Therefore, some assumptions regarding the structure of $\mu(\underline{x})$ must be made. For example, to assume that $\mu(\underline{x}) = \mu$ is constant for all $\underline{x}$ sometimes is appropriate, in which case, $Z(\underline{x})$ has a **stationary mean**. For example, data that have no underlying trend, such as hydraulic conductivity in a homogeneous aquifer, might be assumed to have a constant mean. If the mean is constant, estimating it with the sample average of n measurements obtained at different spatial locations $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n$ is reasonable; therefore,

$$\bar{Z}_n = \frac{1}{n}[Z(\underline{x}_1) + Z(\underline{x}_2) + \ldots + Z(\underline{x}_n)]. \quad (3\text{–}5)$$

However, in contrast to example 3.1.1, $\bar{Z}_n$ may not get closer to μ as n increases, as defined in equation 3–4. Because of the possible spatial correlation in the data, the size of the sampling region needs to be large, compared to the correlation length for $\bar{Z}_n$, to accurately estimate μ.

In addition to the mean of $Z(\underline{x})$, its variability or dispersion also is of interest, and this variability is most commonly measured by the **spatial variance**, defined to be the mean of squared deviations of $Z(\underline{x})$ from $\mu(\underline{x})$ and denoted by $\sigma^2(\underline{x})$,

$$\sigma^2(\underline{x}) = E\{[Z(\underline{x}) - \mu(\underline{x})]^2\}. \quad (3\text{–}6)$$

The **spatial standard deviation** $\sigma(\underline{x})$ is the square root of the variance. The following example illustrates the difference between the population variance, which has been defined in equation 3–6, and a sample variance.

3.1.2 Example 3.1.2

If the scenario presented in example 3.1.1 is used again, the sample variance $S_n^2(\underline{x})$ of the n measurements could be computed as follows:

$$S_n^2(\underline{x}) = \frac{1}{n-1}\sum_{i=1}^{n}[Z_i(\underline{x}) - \bar{Z}_n(\underline{x})]^2. \quad (3\text{–}7)$$

This equation gives a measure of dispersion of the $Z_i(\underline{x})$ values from their sample mean. The sample variance depends on n and on the particular values measured for $Z_1(\underline{x}), Z_2(\underline{x}), \ldots, Z_n(\underline{x})$. However, in the limit as n increases, $S_n^2(\underline{x})$ gets closer and closer to a constant value, which is denoted by $\sigma^2(\underline{x})$. Thus, $\sigma^2(\underline{x})$ is a population parameter, and $S_n^2(\underline{x})$ is a random variable.

The mean and the variance both can be calculated from the probability distribution of $Z(\underline{x})$. Again, in geostatistics, the relations among regionalized variables at different locations are of interest. From the joint distribution of $Z(\underline{x}_1)$ and $Z(\underline{x}_2)$, the **spatial covariance function,**

$$C(\underline{x}_1, \underline{x}_2) = E\{[Z(\underline{x}_1) - \mu(\underline{x}_1)][Z(\underline{x}_2) - \mu(\underline{x}_2)]\} \quad (3\text{–}8)$$

may be obtained. This function is key in geostatistical analyses. It is a measure of the association between values obtained at point $\underline{x}_1$ and the values obtained at point $\underline{x}_2$. If values at these two spatial locations tend to be greater than average or less than average at the same time, then the covariance is positive. However, if the values vary in the opposite direction (that is, one value tends to be larger than average when the other value is less than average, or vice versa), the covariance is negative.

Because $C(\underline{x}_1, \underline{x}_2)$ is an unknown population parameter, it too needs to be estimated using a statistic computed from sample data. To make this estimate possible, the covariance function is sometimes assumed to depend only on the distance between points, which is the **lag** h, and not on the relative location or orientation of the points:

$$C(\underline{x}_1, \underline{x}_2) = C(h), \qquad h = \sqrt{(u_1 - u_2)^2 + (v_1 - v_2)^2}. \quad (3\text{–}9)$$

Under this assumption, $C(h)$ can be estimated by pooling all pairs of measurements that are approximately h units apart and computing a **sample covariance function**

$$\hat{C}(h) = \text{average}\{[Z(\underline{x}_i) - \bar{Z}][Z(\underline{x}_j) - \bar{Z}]: h - \Delta h < h_{ij} < h + \Delta h\}, \quad (3\text{–}10)$$

where h_{ij} is the distance between $\underline{x}_i$ and $\underline{x}_j$ and the average is from all pairs of points so that h_{ij} is between $h - \Delta h$ and $h + \Delta h$. The quantity Δh is called the **lag tolerance.** There are more effective ways to estimate $C(h)$ besides using equation 3–10; for example, see Isaaks and Srivastava (1989). However, because the emphasis in this report is on the variogram (to be defined below) rather than on the covariance function, this method of estimating the covariance function does not need to be used.

A covariance function is called **stationary** if it does not depend on the origin of the coordinate system; that is,

$$C(\underline{x}_1 + \underline{b}, \underline{x}_2 + \underline{b}) = C(\underline{x}_1, \underline{x}_2), \qquad (3\text{–}11)$$

for any given vector, $\underline{h}$ (fig. 1*B*). The covariance function (eq. 3–9) is stationary because changing the origin does not change the distance between the points. Substituting $\underline{x}_1 = \underline{x}_2 = \underline{x}$ in equation 3–9 yields

$$C(\underline{x}, \underline{x}) = C(0) \ , \qquad (3\text{–}12)$$

which, when combined with the definitions in equations 3–6 and 3–8, becomes

$$\sigma^2(\underline{x}) = C(0) \text{ for all } \underline{x}. \qquad (3\text{–}13)$$

Therefore, when $Z(\underline{x})$ has a stationary covariance function, the variance of $Z(\underline{x})$ is constant for all $\underline{x}$. The covariance function then can be standardized by dividing it by the variance. The resulting dimensionless function of h is called the **spatial correlation function**,

$$\rho(h) = \frac{C(h)}{C(0)} \ . \qquad (3\text{–}14)$$

The correlation function is a scale-independent measure of linear association between values of Z at different locations. The spatial correlation is always between −1 and +1, with zero indicating no linear association.

In addition to being stationary, the covariance function in equation 3–9 has another important property. It also is **isotropic**, or **omnidirectional**, because the function does not depend on the direction between the two locations. In many HTRW applications, the correlation between values of Z at two locations is affected by direction as well as lag. For example, contaminant concentrations in a ground-water flow system might be more highly correlated along a transect in the direction of flow than along a transect perpendicular to the flow. Therefore, the covariance function depends on the lag, h, and on the angle, a, between locations,

$$\begin{aligned} C(\underline{x}_1, \underline{x}_2) &= C(h, a), \\ h &= \sqrt{(u_1 - u_2)^2 + (v_1 - v_2)^2}, \\ a &= \text{atan}\left(\frac{v_2 - v_1}{u_2 - u_1}\right). \end{aligned} \qquad (3\text{–}15)$$

In this example, a is the angle measured counterclockwise from the east (fig. 1*B* and *C*). In many geostatistical publications or computer software, the angle may be defined as clockwise from the north, so the appropriate angle in any application needs to be carefully defined. A covariance function that satisfies equation 3–15 is called **anisotropic**, or **directional**.

To summarize, the basic model framework that is used throughout this report is the following: the value of a measurement $z(\underline{x})$ (concentration, porosity, hydraulic head, and so on) at location $\underline{x}$ in a two-dimensional region is the value of a regionalized random variable, $Z(\underline{x})$, with mean $\mu(\underline{x})$ and stationary covariance function $C(h,a)$. Other assumptions may be added in the applications sections of this report to analyze specific data sets, but this framework is the basic framework from which many of the results are derived. In some situations, the covariance-stationarity assumption may be relaxed; for instance, when using the linear variogram described in the next section.

3.2 Variograms

Regionalized random variables differ from classical (ordinary least-squares) regression models in that the **residuals**, defined as the deviations of the regionalized random variable from its mean and denoted by

$$Z^*(\underline{x}) = Z(\underline{x}) - \mu(\underline{x}), \qquad (3\text{–}16)$$

are related to one another, whereas the residuals in a regression model are generally assumed to be independent. Thus, in the regionalized random-variable model, measured values of the residuals at measurement locations contain valuable information when predicting the value of $Z(\underline{x})$ at unsampled locations. The relation between the residuals can be understood by examining the variogram, which is a tool that is widely used in geostatistics for modeling the degree of spatial dependence in a regionalized random variable.

Although the variogram is closely related to the covariance function, there are some important differences between the variogram and the covariance function that are described in this section. The covariance function and the related correlation function are more commonly used in basic statistics courses than the variogram, so many readers may be more familiar with the former concepts. However, the variogram is more widely used in geostatistics and is used as the primary tool for analyzing spatial dependence in the remainder of this report.

As with the covariance function, it is necessary to distinguish between the theoretical variogram, which is based on population parameters, and the sample variogram, which is an estimator of the theoretical variogram obtained from measured data. The **theoretical variogram** of a regionalized random variable, $\gamma(\underline{x}_1, \underline{x}_2)$, is defined as one-half the variance of the difference between residuals at locations $\underline{x}_1$ and $\underline{x}_2$:

$$\gamma(\underline{x}_1, \underline{x}_2) = \frac{1}{2}\mathrm{Var}[Z^*(\underline{x}_1) - Z^*(\underline{x}_2)]. \quad (3\text{–}17)$$

Because the residuals have been mean-centered, as shown in (3–16), they have a mean of zero. Therefore, using the well-known formula for the variance of a random variable, X,

$$\mathrm{Var}(X) = E(X^2) - (EX)^2, \quad (3\text{–}18)$$

equation 3–17 is equal to

$$\gamma(\underline{x}_1, \underline{x}_2) = \frac{1}{2}E[Z^*(\underline{x}_1) - Z^*(\underline{x}_2)]^2. \quad (3\text{–}19)$$

The theoretical variogram is always nonnegative; a small value of γ indicates that the residuals at locations $\underline{x}_1$ and $\underline{x}_2$ tend to be similar and a large value of γ indicates that the residuals tend to be different. Although equation 3–19 is sometimes called a **semi-variogram** because of the multiplication by 1/2, it is referred to in this report as a variogram.

Knowing the theoretical variogram before taking measurements would be ideal, but the theoretical variogram is typically estimated using sample data. To facilitate variogram estimation, it is usually assumed that, as with the covariance function, γ depends only on the lag,

$$\gamma(\underline{x}_1, \underline{x}_2) = \gamma(h), \quad h = \sqrt{(u_1 - u_2)^2 + (v_1 - v_2)^2}, \quad (3\text{–}20)$$

or possibly, on the lag and angle between locations,

$$\gamma(\underline{x}_1, \underline{x}_2) = \gamma(h, a), \quad h = \sqrt{(u_1 - u_2)^2 + (v_1 - v_2)^2}, \quad a = \mathrm{atan}\left(\frac{v_2 - v_1}{u_2 - u_1}\right) \quad (3\text{–}21)$$

(fig. 1). Equation 3–20 is called an **isotropic variogram**, and equation 3–21 is a **directional variogram** at angle a.

For the isotropic variogram, the **sample**, or **empirical, variogram** is obtained by averaging the square of all computed differences between residuals separated by a given lag:

$$\hat{\gamma}(h) = \frac{1}{2}\mathrm{ave}\{[Z^*(\underline{x}_i) - Z^*(\underline{x}_j)]^2: \; h - \Delta h < h_{ij} < h + \Delta h\}, \quad (3\text{–}22)$$

where

h_{ij} is the distance between $\underline{x}_i$ and $\underline{x}_j$.

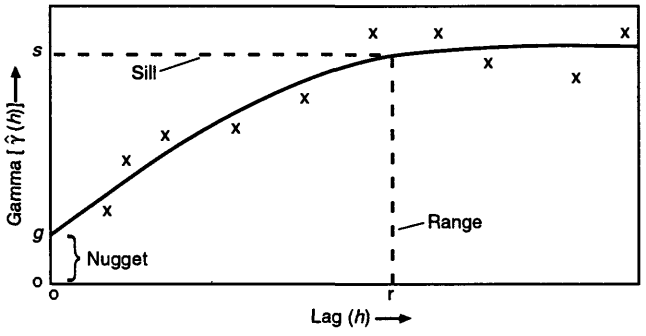
For a given h, as more and more points that are separated by distance $h \pm \Delta h$ are sampled and as Δh decreases, $\hat{\gamma}(h)$ approaches the theoretical variogram. More detail on variogram estimation is presented in section 5.0, including the directional case. In this present section, some general properties of isotropic variograms are described that are referred to numerous times in section 5.0.

A plot of the sample variogram versus h often has a considerable degree of scatter (fig. 2), which is especially evident if the sample size, n, is small. However, the points can usually be fitted by a smooth curve that represents a theoretical variogram selected from a suite of possible choices. Usually, the

theoretical variogram is monotonically increasing, signifying that the farther two measurements are apart, the more their residuals tend to differ, on average, from one another. Several properties common to many theoretical variograms are shown in figure 2. If the variogram either reaches or becomes asymptotic to a constant value as h increases, that value is called the **sill** (fig. 2). The distance (value of h) after which the variogram remains at or close to the sill is called the **range**. Measurements whose locations are farther apart than the range have the same, or even no, degree of association and are assumed to be uncorrelated. Often, a variogram has a discontinuity at the origin, signifying that even measurements obtained very close together are not identical. Such variation in the measurements at small scales is called the **nugget effect**. The size of the discontinuity is called the nugget. Although the nugget effect is sometimes confused with the measurement error, there is a subtle difference between these two concepts that is explained in section 3.3. A simple monotonic function is usually selected to approximate the variogram. Four such functions that are often used in practice are:

The **exponential variogram** (parameters: sill, $s > 0$; nugget, $0 < g < s$; range, $r > 0$),

$$\gamma(h) = \begin{cases} g + (s-g)\left[1 - \exp\left(-3\frac{h}{r}\right)\right], & h > 0 \\ 0, & h = 0 \end{cases} \quad (3\text{–}23)$$



NOTE: The x's denote hypothetical sample variogram points computed from observed data. The smooth curve represents a theoretical variogram fitted to the sample variogram points.

Figure 2. Variogram and features.

the **spherical variogram** (parameters: sill, $s > 0$; nugget, $0 < g < s$; range, $r > 0$),

$$\gamma(h) = \begin{cases} s, & h > r \\ g + (s-g)\left[1.5\frac{h}{r} - 0.5\left(\frac{h}{r}\right)^3\right], & 0 < h \le r \\ 0, & h = 0 \end{cases} \quad (3\text{–}24)$$

the **Gaussian variogram** (parameters: sill, $s > 0$; nugget, $0 < g < s$; range, $r > 0$),

$$\gamma(h) = \begin{cases} g + (s-g)\left[1 - \exp\left(-3\left(\frac{h}{r}\right)^2\right)\right], & h > 0 \\ 0, & h = 0 \end{cases} \quad (3\text{–}25)$$

the **linear variogram** (parameters: nugget, $g > 0$; slope, $b > 0$),

$$\gamma(h) = \begin{cases} g + bh, & h > 0 \\ 0, & h = 0 \end{cases}. \quad (3\text{–}26)$$

Although there are many other functions, these four describe the variogram models most commonly used (Journel and Huijbregts, 1978); the four models are shown in figure 3. The exponential, spherical, and Gaussian models are similar in that they all have a sill and a range. However, they have different shapes near zero lag ($h = 0$) that result in substantial differences in the prediction results using the three models as discussed in section 5.0. The linear model is quite different from the other three, in that it does not reach a sill, but increases linearly without bound. This fact has important implications on the prediction results using a linear variogram. Because the squared differences between residuals increase without bound as the lag increases, a regionalized random variable based on a linear variogram has ever increasing variability about its mean as the size of the sampling region increases. In applications involving the linear variogram, the variogram is usually truncated at a sill corresponding to the value of the variogram at maximum lag, h_{max}. This is illustrated in figure 3 where a linear variogram is shown with a sill and range.

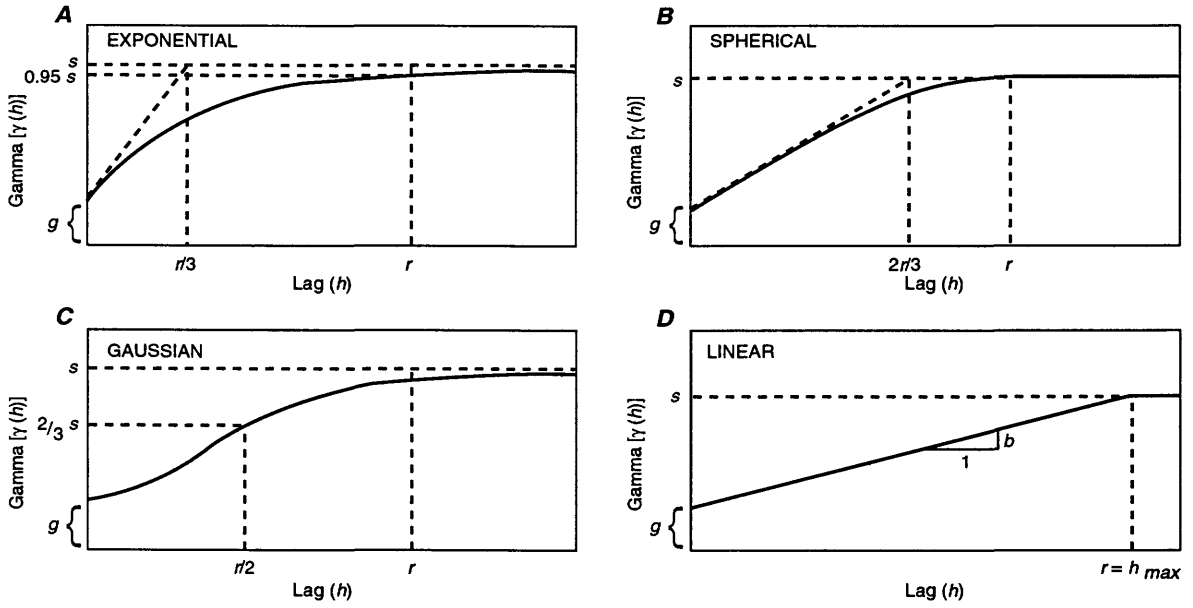


Figure 3. Theoretical variograms indicating *A*, exponential; *B*, spherical; *C*, Gaussian; and *D*, linear models.

Although the variogram is commonly used in a geostatistical analysis, an intuitive understanding of geostatistical techniques may be more easily obtained by using the covariance function, or equivalently, the spatial variance and the correlation function. When $Z(\underline{x})$ has a stationary, isotropic covariance function (eq. 3–9), there is a one-to-one correspondence between the variogram and the covariance function:

$$\gamma(h) = C(0) - C(h) \quad . \tag{3–27}$$

As long as $C(h)$ approaches zero as h increases (a minor technicality that can always be assumed in practice), then, the variogram reaches a sill and the sill equals $C(0)$ as indicated by equation 3–27. Therefore, using a regionalized random variable that is covariance stationary, the variogram and the spatial covariance function contain the same information. By factoring out $C(0) = s$ from equation 3–27 and using equation 3–14, the relation between the spatial correlation function and the variogram can be obtained,

$$\rho(h) = 1 - \frac{\gamma(h)}{s} \quad . \tag{3–28}$$

From equation 3–28, high values of $\gamma(h)$ (that is, close to s) signify low values of $\rho(h)$. In fact, $\rho(h) = 0$ whenever $\gamma(h) = s$, indicating that measurements whose locations are farther apart than the range are uncorrelated. As h decreases, a nugget in $\gamma(h)$ is reflected in a correlation that is less than 1,

$$\rho(h) \rightarrow 1 - \frac{g}{s} \text{ as } h \rightarrow 0 \, . \tag{3–29}$$

Therefore, the larger g is in relation to s, the less correlated nearby observations are. The case when $g = s$, called a **pure nugget variogram**, results in $\rho(h) = 0$ for all $h > 0$. In that case, neighboring measurements are uncorrelated no matter how closely they are spaced.

Occasionally, $\gamma(h)$ may not reach a finite sill, as in the linear variogram (eq. 3–26). In that case, it is not possible to define a correlation function as in (3–28). The corresponding regionalized random variable is said to be **intrinsically stationary** (Journel and Huijbregts, 1978), which is more general than covariance stationary. The theory behind intrinsically stationary variograms is not discussed in this report. As long as a pseudo-range, h_{max}, is defined, all of the computations described below can be used for the linear variogram model.

3.3 Kriging

Given a regionalized random variable $Z(\underline{x})$ that has a known theoretical variogram, how can the value of $Z(\underline{x})$ be predicted at an arbitrary location, based on measurements taken at other locations? To answer that question, suppose that Z is measured at n specified locations: $Z(\underline{x}_1)$, ..., $Z(\underline{x}_n)$. For example, Z could represent hydraulic conductivity and the locations might correspond to n preexisting wells in an aquifer. Let a new location be given by $\underline{x}_0 = (u_0, v_0)$ and denote the ith measurement location by $\underline{x}_i = (u_i, v_i)$. Suppose that, based on prior knowledge of the geology in the study area, there are no prevailing trends in hydraulic conductivity, so the mean of $Z(\underline{x})$ is assumed to be constant over the entire study area:

$$\mu(\underline{x}) = \mu \text{ (constant)}. \tag{3–30}$$

Suppose the value of $Z(\underline{x}_0)$ is to be predicted by using a **linear predictor**, $\hat{Z}(\underline{x}_0)$, which is defined as a weighted linear combination of the measured data,

$$\hat{Z}(\underline{x}_0) = \sum_{i=1}^{n} w_i Z(\underline{x}_i), \tag{3–31}$$

where

w_i is the weight assigned to $Z(\underline{x}_i)$.

To determine specific values for the weights, some criteria need to be specified for $\hat{Z}(\underline{x}_0)$ to be a good predictor of $Z(\underline{x}_0)$. The first criterion is that $\hat{Z}(\underline{x}_0)$ needs to be an **unbiased predictor** of $Z(\underline{x}_0)$, which is expressed as

$$E[\hat{Z}(\underline{x}_0) - Z(\underline{x}_0)] = 0. \tag{3–32}$$

An unbiased predictor neither consistently overpredicts nor underpredicts $Z(\underline{x}_0)$ because the statistical expectation of the prediction errors is zero. The second criterion for a good predictor is that it have small **prediction variance** as defined by

$$\begin{aligned} &\mathrm{Var}[\hat{Z}(\underline{x}_0) - Z(\underline{x}_0)] \\ &= \{E[\hat{Z}(\underline{x}_0) - Z(\underline{x}_0)]^2\}. \end{aligned} \tag{3–33}$$

The smaller the prediction variance, the closer $\hat{Z}(\underline{x}_0)$ is (on average) to the true value $Z(\underline{x}_0)$. The geostatistical technique of kriging computes the **best linear unbiased predictor** of $Z(\underline{x}_0)$, which is the linear unbiased predictor (eqs. 3–31 and 3–32) that has the smallest possible prediction variance (eq. 3–33).

The best linear unbiased predictor depends on the mean of $Z(\underline{x})$. For example, if $Z(\underline{x})$ has a constant mean (eq. 3–30) and a pure nugget variogram [$\gamma(h) = s$ for all $h > 0$], the best linear unbiased predictor of $Z(\underline{x}_0)$ is the average of the measured data,

$$\hat{Z}(\underline{x}_0) = \frac{1}{n} \sum_{i=1}^{n} Z(\underline{x}_i). \tag{3–34}$$

Because the variogram is the same for all $h > 0$ and there is no trend in the data, there is no reason to favor any of the measurements over any of the other measurements. Therefore, the weights are all the same. Ordinary kriging, which is discussed in section 3.3.1, deals with the constant-mean model (the assumption in eq. 3–30) in which the variogram is not a pure nugget variogram. The weights of the best linear unbiased predictor reflects the information in the variogram and results in an improved predictor over the sample mean. In section 3.3.2, universal kriging, which is the extension of ordinary kriging to a nonconstant mean, is discussed. Universal kriging is a very powerful tool that can be used to combine regression models and spatial prediction into one unifying theory. Other, more specialized types of kriging that are discussed in this section are block kriging (section 3.3.3), co-kriging (section 3.4), and indicator kriging (section 3.5.2).

There also is a prediction technique in geostatistics known as **simple kriging**, which uses the best linear unbiased prediction in the case when the mean of $Z(\underline{x})$ is fixed and known. Simple kriging is not discussed in this report because, in most applications, the mean is not known and has to be estimated.

3.3.1 Ordinary Kriging

Let $Z(\underline{x})$ be a regionalized random variable with a constant mean (eq. 3–30) and an isotropic variogram (eq. 3–20). Also, assume that the variogram reaches a sill so the variance of $Z(\underline{x})$ is $C(0) = s$, and the correlation function is given by equation 3–28. Although the prediction equations can be expressed in terms of a variogram, they are defined in this report in terms of the sill (variance) and the correlation function.

Consider linear unbiased predictors from equation 3–31 with the condition in equation 3–32 holding. The unbiased condition is equivalent to

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

for any μ, which holds if, and only if,

$$\sum_{i=1}^{n} w_1 = 1.$$

Therefore, all linear unbiased predictors need to have weights that sum to 1. There are many sets of weights that satisfy this condition, including the set in which all the weights equal $1/n$, as in the sample mean (eq. 3–34). However, the unique set of weights that minimize the prediction variance (eq. 3–33) can be shown to satisfy the following set of $n + 1$ **ordinary kriging equations** (Isaaks and Srivastava, 1989, chap. 12):

$$\sum_{j=1}^{n} w_j \rho_{ij} + \frac{\lambda}{s} = \rho_{i0}, i = 1, 2, ..., n, \qquad (3\text{–}35a)$$

$$\sum_{j=1}^{n} w_j = 1 \qquad (3\text{–}35b)$$

where

$\rho_{ij} = \rho(h_{ij})$ is the correlation between measurements,

i and j, h_{ij} is the distance between locations i and j, and

λ is a coefficient resulting from the constrained optimization.

Furthermore, the resulting **ordinary kriging variance** is

$$\begin{aligned} \sigma_K^2(\underline{x}_0) &= E\{[\hat{Z}(\underline{x}_0) - Z(\underline{x}_0)]^2\} \\ &= s\left(1 - \sum_{j=1}^{n} w_j \rho_{j0}\right) - \lambda. \end{aligned} \qquad (3\text{–}36)$$

The system of equations 3–35a and 3–35b can easily be solved for the w_j's and λ, after which the kriging variance can be obtained from equation 3–36. The ordinary kriging variance changes depending on the prediction location, $\underline{x}_0$, even though the variance of $Z(\underline{x}_0)$ itself (eq. 3–6) is constant for all $\underline{x}_0$.

3.3.1.1 Example 3.3.1.1

Let the mean of $Z(\underline{x})$ satisfy equation 3–30, and suppose that the residual $Z^*(\underline{x})$ (eq. 3–16) has an isotropic exponential variogram (eq. 3–23). Consider predicting $Z(\underline{x}_0)$ based on $n = 2$ measurements, $Z(\underline{x}_1)$ and $Z(\underline{x}_2)$, where the three locations ($\underline{x}_0$, $\underline{x}_1$, and $\underline{x}_2$) are distinct. Using equations 3–23 and 3–28, the correlation function is

$$\rho(h) = \begin{bmatrix} \left(1 - \frac{g}{s}\right)\exp\left(-3\frac{h}{r}\right), h > 0 \\ 1, h = 0 \end{bmatrix}. \qquad (3\text{–}37)$$

Suppose that

$$\frac{g}{s} = p, 0 \le p \le 1, \qquad (3\text{–}38)$$

where

p is a fixed proportion.

The quantity p is sometimes referred to as a **relative nugget.**

The ordinary kriging equations 3–35a and 3–35b are given by

$$w_1 + w_2\rho_{12} + \frac{\lambda}{s} = \rho_{10} \qquad (3\text{–}39a)$$

$$w_1\rho_{12} + w_2 + \frac{\lambda}{s} = \rho_{20} \qquad (3\text{–}39b)$$

$$w_1 + w_2 = 1. \qquad (3\text{–}39c)$$

These three equations have three unknowns w_1, w_2, and λ; the solution is

$$w_1 = \frac{1}{2} + \frac{1}{2}\frac{\rho_{10} - \rho_{20}}{1 - \rho_{12}} \qquad (3\text{–}40a)$$

$$w_2 = \frac{1}{2} - \frac{1}{2}\frac{\rho_{10} - \rho_{20}}{1 - \rho_{12}} \qquad (3\text{–}40b)$$

and

$$\lambda = \frac{s}{2}(\rho_{10} + \rho_{20} - \rho_{12} - 1). \qquad (3\text{–}41)$$

The resulting kriging variance is

$$\sigma_K^2(\underline{x}_0) = s\left[\frac{3}{2} - w_1\rho_{10} - w_2\rho_{20} - \frac{1}{2}(\rho_{10} + \rho_{20} - \rho_{12})\right]. \qquad (3\text{–}42)$$

Although there are only three sample locations in this example (two actual and one potential), the example indicates several properties of best linear unbiased prediction that generally hold. For example,

Effect of sill: The kriging weights depend on s only through the relative nugget, p. However, the kriging variance is directly proportional to s. The sill is called a scaling parameter because scaling each measurement by a constant, c, has the effect of scaling s by c^2. When the relative nugget is allowed to vary so that s and g can change independently, the effect of s is somewhat more complicated.

Effect of nugget: Increasing p has the effect of drawing each of the weights closer to 1/2. As p approaches 1, both weights equal 1/2. The larger g is compared to s, the more small-scale variability there is in the data, and the less important the correlation between neighboring locations becomes. The increased small-scale variability also causes an increase in the kriging variance.

Effect of correlations: If $Z(\underline{x}_0)$ is more highly correlated with $Z(\underline{x}_1)$ than with $Z(\underline{x}_2)$, then w_1 is larger than w_2, indicating that the measurement at the first location has more predictive information than the measurement at the second location. Also, correlation in the data always decreases the kriging variance compared to the variance using uncorrelated data.

Effect of data clumping: If $Z(\underline{x}_1)$ and $Z(\underline{x}_2)$ are highly correlated, as indicated by ρ_{12} being close to 1, then the two measurements contain much of the same information. Two situations then can occur: $\rho_{10} = \rho_{20}$, where the weights are both equal, or $\rho_{10} > \rho_{20}$ [$\rho_{10} < \rho_{20}$], where w_1 is much larger [or smaller] than w_2. In either case, the kriging variance increases to reflect the same information in the two measurements. The automatic adjustment of the kriging weights and kriging variance to account for data clumping is an important property of the kriging predictor.

3.3.1.2 Example 3.3.1.2 (Nugget Effect Versus Measurement Error)

In example 3.3.1.1, all three locations $\underline{x}_0$, $\underline{x}_1$, and $\underline{x}_2$, were assumed to be distinct. When a prediction location coincides with a measurement location, an important distinction needs to be made between a true nugget effect and a measurement error. Suppose that, in example 3.3.1.1, $\underline{x}_0$ and $\underline{x}_1$ are the same. If there is only small-scale variability and no measurement error, then repeated measurements at the same location would be identical, that is, $\rho_{10} = 1$. In this situation, the kriging equations result in $w_1 = 1$, in $w_2 = 0$, in $\lambda = 0$, and in a kriging variance of zero. That is, $Z(\underline{x}_1)$ is a perfect predictor of $Z(\underline{x}_0)$. This property, called **exact interpolation**, is a property of kriging when the data are assumed to contain no measurement errors. However, suppose that the nugget is interpreted as a measurement error rather than a small-scale variability. Repeated measurements at the same location would not be perfectly correlated, but rather, $\rho_{10} = 1 - g/s$. Substituting this correlation into the kriging equations and solving the equations results in a predictor that does not exactly interpolate the data, but smooths the measured data to account for the measurement error. In this report, prediction locations are assumed not to coincide with measurement locations, in which case no distinction needs to be made between the nugget and the measurement error.

3.3.2 Universal Kriging

Universal kriging is an extension of ordinary kriging and can be important in HTRW-site investigations because environmental data often contain drift. Universal kriging addresses the nonconstant mean $\mu(\underline{x})$. Generally, the mean is assumed to have a functional dependence on spatial location of the form

$$\mu(u, v) = \sum_{j=1}^{p} \beta_j f_j(u, v) \qquad (3\text{–}43)$$

where the $f_j(u,v)$'s are known deterministic functions of $\underline{x} = (u,v)$ (that is, these functions serve as independent variables) and the β_j's are regression coefficients to be estimated from the data. Suppose $Z(\underline{x})$ is hydraulic head in an aquifer. If the flow is in a steady state, the mean of $Z(\underline{x})$ could be assumed, in a given study, to have a unidirectional ground-water gradient that is expressed by

$$\mu(u, v) = \beta_1 + \beta_2 u\,. \qquad (3\text{–}44)$$

In this example, there are two independent variables,

$$f_1(u, v) = 1$$

and (3–45)

$$f_2(u, v) = u,$$

and two regression coefficients (β_1 and β_2). The mean can include other independent variables besides the simple algebraic functions of u and v. For example, if the aquifer is not of uniform thickness, an independent variable that involves the aquifer thickness at location (u,v) could be included.

The form of the mean in equation 3–43 also is generally used in standard linear-regression analysis. In regression, ordinary least squares is used to solve for the coefficients; when this is done, the residuals are assumed to be independent and identically distributed. Universal kriging is an extension of ordinary least-squares regression that allows for spatially correlated residuals. Assuming that $Z(\underline{x})$ is a regionalized random variable with a mean as in equation 3–43 and a residual correlation function as in equation 3–28, the best linear unbiased predictor can be obtained from the following $n + p$ equations, called the **universal kriging equations** (Journel and Huijbregts, 1978):

$$\sum_{j=1}^{n} w_j \rho_{ij} + \frac{1}{s} \sum_{k=1}^{p} \lambda_k f_k(\underline{x}_i) = \rho_{i0}\,,$$

$$i = 1, 2, ..., n \qquad (3\text{–}46a)$$

$$\sum_{j=1}^{n} w_j f_k(\underline{x}_j) = f_k(\underline{x}_0),$$

$$k = 1, 2, ..., p \qquad (3\text{–}46b)$$

where, in contrast to the ordinary kriging equations (eqs. 3–35a and 3–35b), there are now p coefficients $\lambda_1, ..., \lambda_p$ resulting from the unbiased condition on the predictor. The first term in the mean (eq. 3–43) is usually a constant, or an intercept, for which $f_1(\underline{x}) = 1$. Therefore, the universal kriging model includes ordinary kriging as a special case. The **universal kriging variance** is given by

$$\sigma_K^2(\underline{x}_0) = s\left(1 - \sum_{i=1}^{n} w_i \rho_{i0}\right) - \sum_{k=1}^{p} \lambda_k f_k(\underline{x}_0). \qquad (3\text{–}47)$$

Equations 3–46a and 3–46b and equation 3–47 can be easily solved to obtain universal kriging predictors and kriging variances for any location. The estimated trend surface does not need to be computed to obtain the universal kriging predictor. If a particular application needs an estimate of the trend surface, then generalized least-squares regression can be used to estimate the coefficients (β_j's) in the regression equation.

3.3.3 Block Kriging

In the previous sections, the problem of predicting the value of a regionalized random variable at a specified location in the region for which

the variable is defined has been discussed. Implicit in this discussion is the assumption that the support of the variable being predicted is defined in the same way as the variables that make up the measurements. However, there may be applications where estimating the average value of Z for an estimation block of much larger area than is represented by an individual sample is necessary. For example, an estimate of the average concentration of a contaminant in an entire aquifer that is based on point measurements at various locations might be needed. In other applications, an estimate of the average concentration of soil contaminant, in daily excavation volumes that are much larger than the volume of an individual sample, may be needed. Let Z_B be the average value of $Z(\underline{x})$ for a particular block B,

$$Z_B = \frac{1}{m}\sum_{i=1}^{m} Z(\underline{x}_{0i}) \qquad (3\text{–}48)$$

where $\underline{x}_{0i}$, $i = 1,...,m$, denotes m prediction locations in block B. The object is to predict this average rather than the regionalized random variable at a single location. In many applications, the locations $\underline{x}_{0i}$ might correspond to nodes of a regular grid or finite-element nodes in a ground-water model. The results of the block kriging are dependent on m and on the prediction locations. Selecting a large number of locations in block B, where each location has approximately the same representative area, probably is the best approach to block kriging (Isaaks and Srivastava, 1989, chap. 13).

The objective of block kriging is to obtain the best linear unbiased predictor of Z_B and an estimate of the block kriging variance based on the measurements. The model for $Z(\underline{x})$ can be the constant-mean model (eq. 3–30) assumed for ordinary kriging or the more general linear-regression model (eq. 3–43) assumed for universal kriging. For either, the predicted value of Z_B coincides with the average of the predicted values of the individual measurements in the block; that is,

$$\hat{Z}_B = \frac{1}{m}\sum_{i=1}^{m} \hat{Z}(\underline{x}_{0i}). \qquad (3\text{–}49)$$

In this model, the individual predicted values are obtained from either the ordinary or the universal kriging equations. However, computation of the block kriging variance is not as simple as computation of the point kriging variance because the individual kriging estimates are not independent of one another. There are simple modifications to the kriging equations, discussed in sections 3.3.1 and 3.3.2, that can be used to directly compute the kriging estimate of Z_B and its kriging variance (Isaaks and Srivastava, 1989, chap. 13). The equations are not presented in this report. The computer packages described in the next section can be used to compute block kriging estimates. In general, kriged values of block averages are less variable than kriged values at single locations. Consequently, the blocked kriging variance tends to be smaller than the kriging variance at a single location.

3.4 Co-Kriging

Kriging as discussed so far provides a way of predicting values of a regionalized variable $Z(\underline{x})$ at a location $\underline{x}_0$ based on measurements of the same variable at locations $\underline{x}_1, \underline{x}_2, ..., \underline{x}_n$. In some situations, however, measurements could be available not only of $Z(\underline{x})$, but also of one or more other variables that can be used to improve predictions of $Z(\underline{x}_0)$. The variable $Z(\underline{x})$ is called the primary variable because it is the one to be predicted, and the other variables are called secondary variables. **Co-kriging** is a technique that uses the information contained in secondary variables to predict a primary variable. For example, suppose that $Z(\underline{x})$ is a regionalized variable representing the hexavalent chromium concentration, a relatively difficult determination, and suppose that the hexavalent chromium concentration needs to be predicted at a location $\underline{x}_0$ based on measurements of hexavalent chromium at other locations. However, there also are measurements of a second, relatively easily determined contaminant, such as lead, that, for the purposes of this example, tend to be correlated with hexavalent chromium concentration, and these data are to be used as well. Denote the second variable, lead, by a regionalized variable $W(\underline{x})$, and assume that measurements have been made on W at m locations $\underline{x}'_1, \underline{x}'_2, ..., \underline{x}'_m$. The co-kriging predictor of $Z(\underline{x}_0)$ then is

$$\hat{Z}_C(\underline{x}_0) = \sum_{i=1}^{n} w_i Z(\underline{x}_i) + \sum_{j=1}^{m} w'_j W(\underline{x}'_j). \quad (3\text{–}50)$$

This extension of the kriging predictor in equation 3–31 is straightforward. Analogous to kriging, co-kriging produces the weights w_i and w'_j so that the resulting predictor is the best linear unbiased predictor. Also, as with kriging, co-kriging uses modeling of the variogram for Z, but co-kriging presents an additional necessity of modeling the variogram for W and the **cross variogram** for Z and W. The optimal weights then are expressed in terms of all these variogram properties. More than one secondary variable may be included in the co-kriging predictor, and theory has been developed for co-kriging in the presence of drift (universal co-kriging) and block co-kriging. Details are not included in this report, but Isaaks and Srivastava (1989) and Deutsch and Journel (1992) have more discussion and citations of other references.

One situation for which co-kriging might be useful is when the primary variable is undersampled, so any additional information, such as that given by secondary variables, would be helpful. However, although co-kriging can be a useful tool, joint modeling of several variables tends to be demanding as to data and computational requirements. Thus, undersampling of the primary variable may present problems for co-kriging and for one-variable kriging. Also, unless the primary variable of interest is highly correlated with the secondary variable(s), the weights assigned to the secondary variable(s) are often small, and the effort needed to include the additional variable(s) may not be worthwhile. For these reasons, co-kriging is not used extensively in practice.

Although co-kriging is similar to universal kriging in that both techniques use extra variables to predict $Z(\underline{x})$, there is an important distinction between the two techniques. In universal kriging, the independent variables in equation 3–43 need to be known with certainty at the prediction location $\underline{x}_0$. For example, aquifer thickness might be an independent variable in predicting aquifer head if the thickness can easily be determined at any location. However, aquifer thickness may need to be considered a secondary variable in a co-kriging procedure if the thickness is only known at a few selected locations in the aquifer.

3.5 Using Kriging to Assess Risk

The kriging predictor of $Z(\underline{x}_0)$ has certain desirable properties based on how close it is to the actual value of $Z(\underline{x}_0)$; it is unbiased and has the smallest variance among all linear predictors. However, when possible, the relation between the predicted and observed values could be specified further, and ideally, probability statements could be made. For example, if $Z(\underline{x}_0)$ is the concentration of a contaminant, a 95-percent certainty that the true concentration is within 0.05 microgram per liter of the predicted concentration might be desired. In other situations, the probability that the actual concentration exceeds a given target concentration might need to be estimated. Knowledge of the entire distribution function of $Z(\underline{x})$, as opposed to knowledge of only the mean and variogram of $Z(\underline{x})$, can be used for risk-qualified inferences in situations when extremes might be of more interest than averages.

A discussion of the concept of a **conditional probability-distribution function** of the regionalized variable $Z(\underline{x})$ is appropriate at this point. The concept also is applicable in section 8.0 when conditional simulation is discussed. The conditional probability-distribution function is defined much like the probability-distribution function in section 3.1, except the probability that $Z(\underline{x}) \leq c$ is computed conditional on, or given, information at other spatial locations. Geostatistics is used to make predictions at a location $\underline{x}_0$ using information at measurement locations $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n$; therefore, in conditional distributions, the focus is on $P[Z(\underline{x}_0) \leq c \mid Z(\underline{x}_1), Z(\underline{x}_2), \ldots, Z(\underline{x}_n)]$. The vertical bar denotes the conditioning and is read "given." The conditional probability distribution needs to be determined to make probability statements about the regionalized variable at location $\underline{x}_0$. Also, **conditional mean** and **conditional variance** can be defined in the same way that mean and variance for distribution functions were defined in section 3.1.

Section 3.5.1 contains methods for using kriging output to obtain prediction intervals or quantiles when the regionalized random variable is either normally distributed or can be transformed to a near-normal distribution. Section 3.5.2 discusses indicator kriging, which is a nonparametric technique for obtaining quantiles when data cannot be adequately transformed to a normal distribution.

3.5.1 Normal Distributions and Transformations

For prediction at a location $\underline{x}_0$, a kriging analysis produces the predictor $\hat{Z}(\underline{x}_0)$ and the associated kriging variance $\sigma_K^2(\underline{x}_0)$. If more informative probability assessments are to be made, the ideal situation is when $Z(\underline{x})$ is assumed to be a Gaussian, or normal, process, which means that $[Z(\underline{x}_1),..., Z(\underline{x}_n)]$ has a joint normal probability distribution for any set of n locations and any value of n. Then the conditional probability distribution of $Z(\underline{x}_0)$, given the n measurements, is a normal distribution that has a conditional mean equal to the kriging predictor $\hat{Z}(\underline{x}_0)$ and conditional variance equal to the kriging variance $\sigma_K^2(\underline{x}_0)$. This normal distribution can be used to obtain a **prediction interval** for $Z(\underline{x}_0)$ (conditional on the measured data). For example, from a table of the normal distribution, a value of 1.96 corresponding to a 0.95 (two-sided) probability can be obtained. Then the assertion that there is a 95-percent chance that $Z(\underline{x}_0)$ is in the 95-percent prediction interval $[\hat{Z}(\underline{x}_0) - 19.6\sigma_K(\underline{x}_0), \hat{Z}(\underline{x}_0) + 1.96\sigma_K(\underline{x}_0)]$ can be made. Knowing this interval is much more useful than simply knowing the kriging predictor and variance.

To illustrate quantile estimation, suppose that contaminant concentrations are being studied, and a concentration that has only a 1-percent chance of being exceeded at location $\underline{x}_0$ needs to be determined. The appropriate (one-sided) value from a normal table is 2.33, so the desired estimate is $\hat{Z}(\underline{x}_0) + 2.33\sigma_K(\underline{x}_0)$.

Even if $Z(\underline{x})$ is not Gaussian, a **transformation**, $Y(\underline{x}) = T[Z(\underline{x})]$, can often be found so that $Y(\underline{x})$ is approximately Gaussian. When a transformation is made, the kriging analysis is performed using the transformed data $Y(\underline{x})$, and the inverse transformation may be applied to obtain prediction intervals for the original data. For example, the most common transformation is the (natural) **logarithmic transformation**, in which $Y(\underline{x}) = ln[Z(\underline{x})]$. A 95-percent prediction interval for $Z(\underline{x})$ then is $\{\exp[\hat{Y}(\underline{x}_0) - 1.96\sigma_K(\underline{x}_0)], \exp[\hat{Y}(\underline{x}_0) + 1.96\sigma_K(\underline{x}_0)]\}$. As long as the transformation is a one-to-one function, such as a logarithmic transform, prediction intervals for the original data can be obtained by simply back-transforming prediction intervals for the transformed data.

Although prediction intervals and probabilities can be easily obtained using simple back-transformation, obtaining a predictor of the untransformed data that is unbiased and optimal in some sense is more difficult. For example, using a logarithmic transformation, a kriging analysis using the transformed data yields a predictor $\hat{Y}(\underline{x}_0)$, which is the best linear unbiased predictor of $Y(\underline{x}_0)$. However, the back-transformed value $\hat{Z}(\underline{x}_0) = \exp[\hat{Y}(\underline{x}_0)]$ does not possess the same optimality properties as a predictor of $Z(\underline{x}_0)$. The technique known as log-normal kriging, and more generally as trans-normal kriging, has been developed to obtain predictors when transformations are made (Journal and Huijbregts, 1978), but because of the complexity involved, the technique is not usually used by practitioners. For example, if a predicted value corresponding to $Z(\underline{x}_0)$ needs to be obtained for contour plotting, the kriging predictions $\hat{Y}(\underline{x}_0)$ may be back-transformed and plotted, as long as the investigator realizes that such values do not have the usual kriging optimality properties.

3.5.2 Indicator Kriging

There may be situations when a transformation that makes $Z(\underline{x})$ approximately normal cannot be easily determined. In such situations, **indicator kriging** can be used to infer the probability distribution of $Z(\underline{x})$. Because no distributional assumptions are made, this technique is known as a **nonparametric** statistical technique. An example of indicator kriging is included in section 6.0, and an article by Journel (1988) is a good reference for additional information about indicator kriging.

To perform indicator kriging, a special transformation, known as an indicator transformation, is applied to $Z(\underline{x})$:

$$I(\underline{x}, c) = \begin{cases} 1, & Z(\underline{x}) \le c \\ 0, & Z(\underline{x}) > c \end{cases} . \qquad (3\text{–}51)$$

If, as in the usual kriging scenario, the data set at hand consists of measurements of the regionalized variable $Z(\underline{x})$ at n locations, c needs to be fixed first, and then the indicator transformation is applied by replacing values that are less than or equal to c with 1 and values that are greater than c with 0. The variogram and kriging analysis then is performed using these 0's and 1's rather than the raw data.

Kriging predictors using the indicator data are equal to their measured values of 0 or 1 at the measurement locations $\underline{x}_i$, $i = 1,...,n$. However, at locations different from the measurement locations, predictions may be between 0 and 1. In interpreting these predictions, the power of indicator kriging becomes apparent. A predicted value at $\underline{x}_0$ is an estimate of the conditional probability distribution $P[Z(\underline{x}_0) \leq c \mid Z(\underline{x}_1), Z(\underline{x}_2),..., Z(\underline{x}_n)]$. This analysis may be performed for a range of values of c; therefore, the entire distribution function can be estimated. This estimate of the distribution function can be used to obtain prediction intervals or estimates of quantiles. For example, to estimate the value that has a 1-percent chance of being exceeded at location $\underline{x}_0$, the value of c for which the kriged indicator prediction is 0.99 at that location is determined.

One advantage of indicator kriging is that the indicator variogram is robust with respect to extreme outliers in the data because no matter how large (or small) $Z(\underline{x})$ is, the indicator variable is either 0 or 1. Indicator variables also may be used in block kriging. For example, a spatial average of $I(\underline{x},c)$ over a block B equals the fraction of block B for which $Z(\underline{x})$ is less than c. Another advantage of indicator kriging is that it can be used when some data are censored.

Despite the relative ease of implementation, there are several drawbacks to indicator kriging, and this technique might be used only when other techniques, such as normality transformations, produce unacceptable results. For example, the kriged values of $I(\underline{x},c)$ may be less than 0 or larger than 1. Also, the kriged prediction for $I(\underline{x},c_1)$ may be larger than the kriged prediction for $I(\underline{x},c_2)$ even if $c_1 < c_2$, which is not compatible with a valid probability distribution. There are several more advanced techniques for solving these problems (Isaaks and Srivastava, 1989, chap. 18); however, those techniques are beyond the scope of this report.

4.0 GEOSTATISTICAL RESOURCES AND TOOLS

Since the mid-1970's, there have been texts and articles published that are either totally dedicated to geostatistical methods or discuss geostatistics in detail. As well as being separately published, numerous computer programs and software packages on geostatistics and kriging are included in these texts and articles. Although only a few of these resources are briefly described in this report, their references can provide lists of other geostatistical topics or software not specifically covered in the resources.

4.1 Texts on Geostatistics

The geostatistical texts presented in this section can be classified into two broad categories: instructional texts or reference texts. For one who is delving into geostatistics for the first time, Clark's (1979) book can be a starting point. Simple explanations of the basic kriging techniques are applied to an example data set. A more detailed treatment of the kriging techniques is described by Isaaks and Srivastava (1989). This text book presents discussions of many of the background statistical tools and concepts needed in geostatistical applications, including histograms and distributions (univariate and bivariate), sampling, correlation, and spatial continuity. The text also discusses how to treat the subtleties of kriging using three example data sets. As well as being instructional, the book also can be used as a reference.

Texts by Cressie (1991) and Journel and Huijbregts (1978) describe the tools of geostatistics, but also include a comprehensive theoretical background on the techniques. Cressie's (1991) text is a treatment of spatial processes in general and reviews a wide range of statistical techniques in the analysis and stochastic modeling of spatial data. There is a four-chapter section on geostatistics, with a complete discussion of variogram estimation, kriging (including universal kriging), intrinsic random functions, and comparisons of kriging to other spatial prediction techniques. The text is written from a statistician's point of view and is, in places, written at a fairly high level mathematically. Nevertheless, it contains numerous examples and illustrations using real-world data. Journel and Huijbregts (1978) maintained a mining-geological perspective. Two other texts

written by statisticians that present general treatments of spatial processes, but that lack detailed discussions of kriging, are by Cliff and Ord (1981) and by Ripley (1981).

David's (1977) text was the first extensive discussion of the practice of geostatistics and kriging in mining applications, and the discussion is presented from a practitioner's viewpoint. The text references many specific mining applications and results for geostatistics. A broad statistics text by Davis (1973) with a bent toward geological applications, serves as a reference for standard statistical procedures needed in geological applications of geostatistics. A book by Bras and Rodriguez-Iturbe (1985) that discusses a range of techniques for stochastic modeling in hydrology includes a chapter on applications of kriging. There is a fairly complete mathematical development of kriging with details of an application to predict mean areal precipitation. In an article prepared for the U.S. Environmental Protection Agency, Journel (1993) discussed geostatistics as it relates to environmental science. Finally, Olea (1991) presented a useful glossary of geostatistical terms.

4.2 Journals

The journal *Mathematical Geology* by the International Association for Mathematical Geologists reports new developments in the theory and application of kriging. Although many of the articles present new applications of kriging tools, many articles also are dedicated to the derivation of statistical properties of the variogram, to kriging estimation, and to cross-validation results. Journals such as *Water Resources Research*, published by the American Geophysical Union, and *Groundwater,* published by the Association of Ground Water Scientists and Engineers, contain articles describing special applications of kriging techniques in the environmental arena. *Water Resources Research* includes many theoretical articles. Other journals that may contain information addressing geostatistics are the *Journal of Environmental Engineering*, published by the American Society of Civil Engineers; *Stochastic Hydrology and Hydraulics*, published by Springer International; and the *North American Council on Geostatistics*, published by the Colorado School of Mines.

4.3 Software

The geostatistics software described in this section is limited to a few readily available public-domain packages that are executable at least on the DOS and sometimes on the UNIX platforms. There are several commercial packages that are being marketed, but these packages are not reviewed in this report. Each of the software packages described in this report are listed in table 1, which may serve as a reference guide to other software packages.

One of the earliest interactive kriging software packages was developed by Grundy and Miesch (1987). Overall, this general statistics package, known as STATPAC, contains a series of programs that can handle two-dimensional kriging, including universal kriging. The package has capabilities for univariate statistics, transformations, variogram analysis, and cross validation (table 1). The graphics in the package are limited to simple line-printer plots of the sample variogram points and data maps. The menu-driven package includes a tutorial using all of the kriging routines. The package is distributed with not all, but most source codes, and, therefore, can be modified by the user if desired. All two-dimensional kriging routines can be executed from the command line, which provides users with the opportunity for batch processing.

The geostatistical environmental-assessment software, known as GEO-EAS (Englund and Sparks, 1991), also is an interactive, menu-driven kriging software package for performing two-dimensional kriging. It has no direct provisions for universal kriging (table 1). GEO-EAS does have an advantage over STATPAC through its enhanced graphics capabilities, which are useful in the interactive fitting of theoretical variograms to sample variogram points. In addition, in the computation of the sample variogram points, GEO-EAS allows for variable bin sizes, the use of which are further discussed in section 5.0.

STATPAC and GEO-EAS were originally developed for the personal computer. Since then, versions of GEO-EAS have been developed for some types of work stations. The kriging routines in STATPAC have not been adapted to workstations.

A third software package, the geostatistical software library known as GSLIB (Deutsch and Journel, 1992), is a suite of programs developed over the years at Stanford University, Stanford, California.

Table 1. Geostatistical software characteristics

[Note: STATPAC, statistical software package developed by Grundy and Miesch (1987); GEO-EAS geostatistical environmental-assessment software developed by Englund and Sparks (1991); GSLIB geostatistical software library developed by Deutsch and Journel (1992); GMS, ground-water modeling system developed for the U.S. Department of Defense]

Characteristic	STATPAC	GEO-EAS	GSLIB	GMS2.0
Operating system	DOS	DOS/UNIX	Independent (requires FORTRAN compiler)	WINDOWS 95 UNIX
Menu driven	Yes	Yes	No	Yes
Batch processing	Yes	No	Yes	Yes
User modifications	Yes, source code provided	No	Yes, source code provided	No
Data-set constraints	Yes, modifications possible via source code	Yes	Yes, modifications possible via source code	Yes
ASCII output	Yes	Yes	Yes	Yes
Univariate statistics	Yes	Yes	Yes	Yes
Additional exploratory capabilities	Yes	Yes	Yes	Yes
Graphical support for analysis	Yes	Yes	Yes	Yes
Transformations	Yes	Yes	Yes	Yes
Back-transformations	No	No	Yes	Yes
Variogram construction	Yes	Yes	Yes	Yes
Variogram analysis	Yes	Yes	Yes	Yes
Variogram graphics	Yes	Yes	Yes	Yes
Cross-validation operations	Yes	Yes	Yes	Yes
Ordinary kriging	Yes	Yes	Yes	Yes
Universal kriging	Yes	No	Yes	Yes
Block kriging	Yes	Yes	Yes	Yes
Indicator kriging	Yes	Yes	Yes	No
Conditional simulation	Perhaps with batch processing	No	Yes	No
Three-dimensional kriging	Perhaps with batch processing	No	Yes	Yes
Mapping	Yes	Yes	Yes	Yes
Contouring	Yes	Yes	Yes	Yes
Gray-scale maps	Yes	Yes	Yes	Yes
Line printer	Yes	No	Yes	Yes
High-resolution screen	No	Yes	Yes, via postscript	Yes
High-resolution printer	No	Yes	Yes	Yes
Postscript	No	No	Yes	Yes

It is presented as a collection of routines that are machine independent (table 1) and are intended to be used as a modular concept. The package is distributed as a suite of FORTRAN source codes that need to be compiled. To use GSLIB effectively, a relatively high level of familiarity with geostatistics is required. Like the other two software packages, GSLIB handles variogram analysis and kriging techniques (table 1). Two of its primary advantages over the other two packages are its simulation techniques and its ability to analyze three-dimensional data sets. Such techniques are useful especially in estimating potential extreme outcomes in a geostatistical analysis.

The U.S. Department of Defense Groundwater Modeling System (GMS) is a fourth software package that has kriging capabilities. GMS is a windows-based integrated modeling environment for site characterization, ground-water flow and transport modeling, and visualization of results. The GSLIB software has been implemented in GMS to facilitate two- and three-dimensional kriging and interactive variogram modeling. GMS also provides comprehensive visualization techniques and other interpolation techniques that can be used as alternatives to kriging. The GMS system was developed for the U.S. Department of Defense by the Brigham Young University Engineering Computer Graphics Laboratory. GMS may be obtained from the U.S. Army Groundwater Modeling Technical Support Center, Waterways Experiment Station, Vicksburg, MS 39180.

In geostatistical software and literature, there can be differences in jargon or notation. These differences may cause some initial confusion if users or readers do not familiarize themselves with the jargon or notation. For example, some authors may use the term "semi-variogram" rather than "variogram"; others may express random variables as other than Z (which has been used in this report); and different software often has different references for directional angles when discussing anisotropy.

5.0 PRACTICAL ASPECTS OF VARIOGRAM CONSTRUCTION AND INTERPRETATION

Section 3.0 presented the mathematical foundation for geostatistics and the kriging technique. One theme that pervades the technique is the importance of the theoretical variogram. The theoretical variogram, or what is often referred to simply as the variogram, is a mathematical function or model that is fitted to sample variogram points obtained from data. Permissible models, which include those models discussed in section 3.0, belong to a family of smooth curves having particular mathematical properties and are each specified by a set of parameters. Section 5.0 describes a sequence of stages for estimating and investigating sample variogram points and a calibration procedure for specifying the parameters of the variogram model eventually to be fitted to the sample points. Although the calibration procedure is largely an objective way for evaluating theoretical variograms, the process of obtaining sample variogram points and finalizing a theoretical variogram remains an art as much as a science. An understanding of the material presented in section 3.0 and professional judgment achieved through experience in geostatistical studies are important in effectively using the guidelines presented in section 5.0.

An accurate estimate of a variogram from a kriging perspective is needed because the correlation matrix used to obtain the kriging weights is developed from the variogram values. Even more directly, the variogram affects the computation of the kriging variance (eqs. 3–36 and 3–47) through the product of the kriging weights and the correlation values. An accurate variogram also can be used outside the strict context of kriging. For example, in augmenting a spatial network with new data-collection sites, the range parameter of the variogram could be used as the minimum distance of separation between the new sites and between the new and the existing sites to maximize overall additional regional information. In another nonkriging-specific application, the variogram is used in dispersion variance computations in which the variance of areal or block values is estimated from the variance of point-data values (Isaaks and Srivastava, 1989, p. 480).

The stages of variogram construction are described based on an example data set of groundwater elevations measured near Saratoga, Wyoming (Lenfest, 1986). The data set is summarized in table 2 and the relative locations of the data are shown in figure 4.

The sequence of steps in computing sample variogram points depends on the stationarity properties of the regional variable represented by the data. If the mean of the regional variable is the same for all locations, then the mean is said to be spatially stationary; if the mean changes with location, then

Table 2. Univariate statistics for example data sets

[Note: Base unit for Saratoga, water level A and B and bedrock A and B is meters; base unit for water quality A is log concentration, concentration in micrograms per liter]

Example identifier	Number of measurements	Transformation	Minimum (base units)	Maximum (base units)	Mean (base units)	Median (base units)	Standard deviation (base units)	Skewness (dimensionless)
Saratoga	44	Drift	614	687	646	641	17.3	0.45
Water level A	83	Drift	7.80	20.0	12.9	11.7	3.09	1.03
Water level B	74	Drift	7.80	20.0	13.1	11.8	3.23	0.87
Bedrock A	107	None	6.91	24.5	13.5	13.1	3.28	0.89
Bedrock B	90	None	7.75	21.1	13.3	13.2	2.62	0.26
Water quality A	66	Natural log	2.08	8.01	5.19	5.59	1.75	–0.42

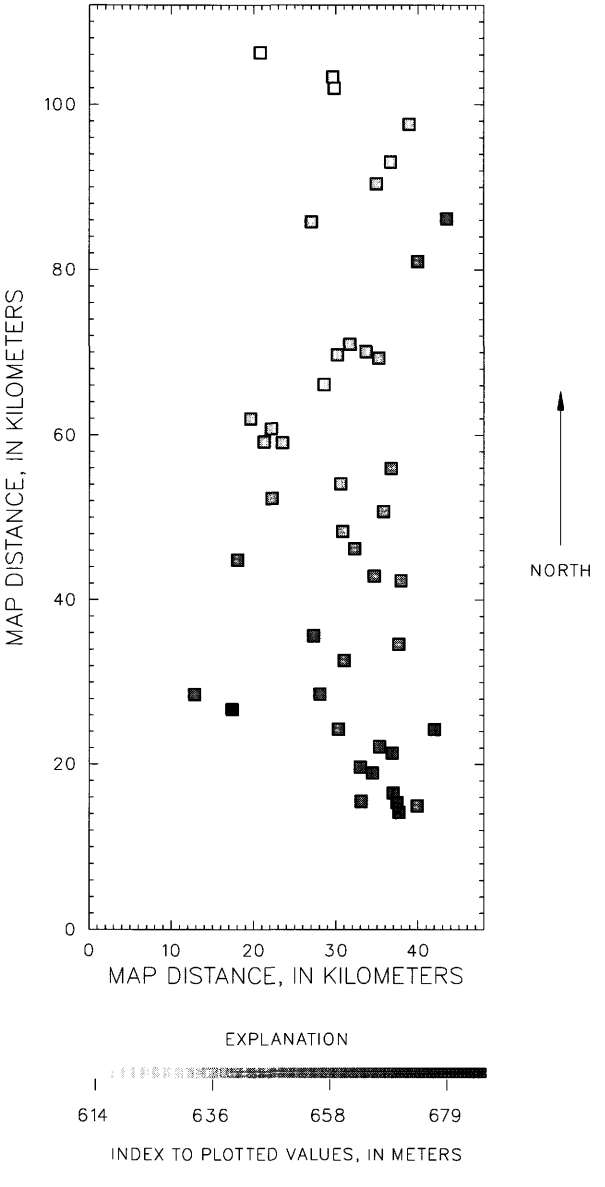


Figure 4. Measured water levels from Saratoga data.

it is spatially nonstationary. If the data have a stationary spatial mean, the discussions in sections 5.2 and 5.6, which address nonstationarity and additional trend considerations, can be skipped. If the spatial mean is not stationary, as in this example data set, then sections 5.2 and 5.6 become important, and the sequence of stages for obtaining a variogram becomes an iterative procedure. All variogram and kriging computations for the Saratoga ground-water-level example were performed using the interactive kriging software STATPAC described by Grundy and Miesch (1987).

5.1 General Computation of an Empirical Variogram

As described in section 3.2, the variogram, $\gamma(h)$, characterizes the spatial continuity of a regional variable for pairs of locations as a function of distance or lag, h, between the locations. This variogram is sometimes called the theoretical variogram because it is assigned a continuous functional form that expresses the spatial correlation for any lag in the region of analysis. The function is estimated by fitting one of the equations in section 3.2 to empirical or sample variogram points, $\hat{\gamma}(h)$, using data whose locations contribute only a finite number of lags. Although $\hat{\gamma}(h)$ characterizes the spatial correlation of the data, it is computed from residuals of the data from the spatial mean. Therefore, without prior knowledge of nonstationarity in the underlying spatial process, the first step in computing the sample variogram is to identify existing nonstationarity indicated for the spatial mean.

The approximation to equation 3–19 begins by computing squared differences, D_{ij}^2, from the data values $z(\underline{x}_1)$, $z(\underline{x}_2)$,...$z(\underline{x}_n)$ collected at locations $\underline{x}_1, \underline{x}_2 \cdots \underline{x}_n$

$$D_{ij}^2 = [z(\underline{x}_i) - z(\underline{x}_j)]^2 . \qquad (5\text{–}1)$$

If the spatial mean is stationary, then the squared differences of the data are equal to the squared differences of the residuals, and sample variogram computations can be continued using the data themselves. If the spatial mean is strongly nonstationary, the plot of equation 5–1 versus the distance between associated points may indicate a trend or drift that needs to be removed before further variogram computations can be made. Drift needs to be considered in HTRW studies such as determining contaminant concentrations areally dispersed from localized sources or determining ground-water elevations that follow a local or regional gradient. In such studies, sample variogram computations need to be made using residuals obtained by subtracting the estimated drift value at each location from the value of the datum at the location.

The differencing of the data in equation 5–1 is done without considering the relative direction between the locations; that is, D_{ij}^2 is isotropically computed. A plot of D_{ij}^2 versus h_{ij} for all i,j $(i > j)$, where $h_{ij} = |\underline{x}_i - \underline{x}_j|$, produces a cloud of points whose properties govern the behavior of $\hat{\gamma}$. The central tendency of the cloud generally increases with h. A substantial increase in the central tendency that persists for large h can indicate a nonstationary spatial mean. The cloud computed for the Saratoga data, with ground-water levels (z) in meters and lag (h) in kilometers, is shown in figure 5 and does show increasing D^2 (meters squared) with increasing h, indicating potential nonstationarity.

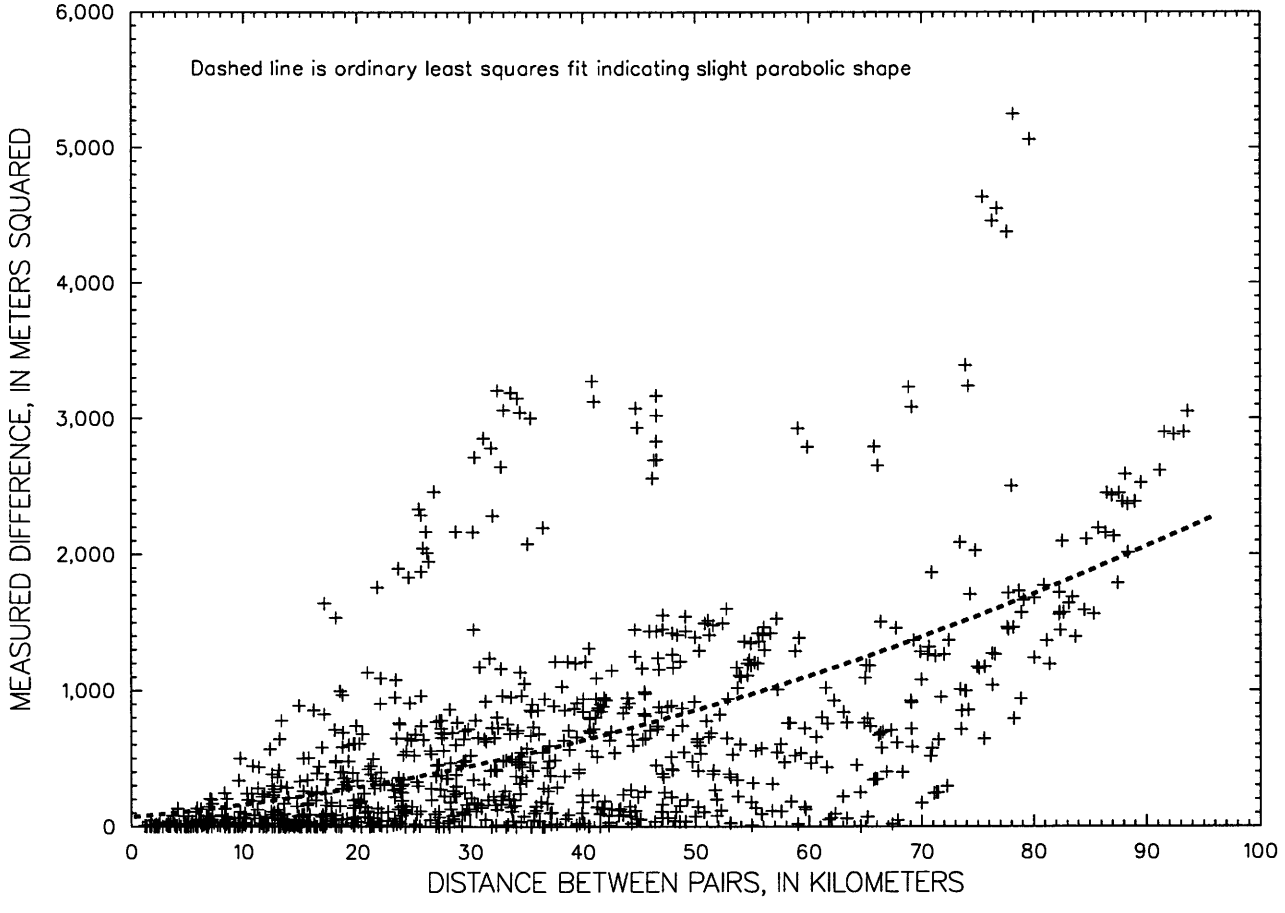


Figure 5. Squared differences of values for all possible pairs of points for Saratoga data.

Generally, there is a large amount of scatter in these plots, as seen in figure 5, and this scatter can conceal the central tendency of D^2 with h. One way to estimate the central tendency and to minimize the effect of aberrant data is to collect the D^2 into K bins or lag intervals of width $(\Delta h)_k$, $k = 1,...K$ and assign to $\hat{\gamma}$ the average of the values of D^2 in each bin. This process is similar to the way data are placed in bins for obtaining histograms. The expression for the kth average bin value is

$$\hat{\gamma}(h_k) = \frac{1}{2N(h_k)} \sum D_{ij}^2 I_k(h_{ij}) \qquad (5\text{–}2)$$

where

$N(h_k)$ is the number of squared differences that fall into bin k, and

h_k is the lag distance associated with bin k.

$I_k(.)$ is an indicator function that has a value of 1 if h_{ij} falls into bin k and 0 otherwise [$I_k(.)$ only allows values of D_{ij}^2 in the calculation that have an h_{ij} that falls into the bin].

The lag value h_k can be the midpoint of the bin or it can be the average of the actual lag values for the points that fall in the bin.

To establish bins, equal bin widths are specified and the distance between the two most separated data points, h_{max}, is subdivided according to these equal increments, or a K is chosen that defines the bin width. For the Saratoga data, a bin width of about 8 kilometers established $K = 12$ bins for $\hat{\gamma}$. The $\hat{\gamma}$ points computed from the binned D_{ij}^2 values in figure 5 are shown in figure 6. The lag plotting positions are the average h values in the bin. The symbol x indicates that $N(h)$ is less than 30 pairs for the particular bin, and this differentiation is discussed in section 5.3. Although the sample variogram is still preliminary, its general behavior at this

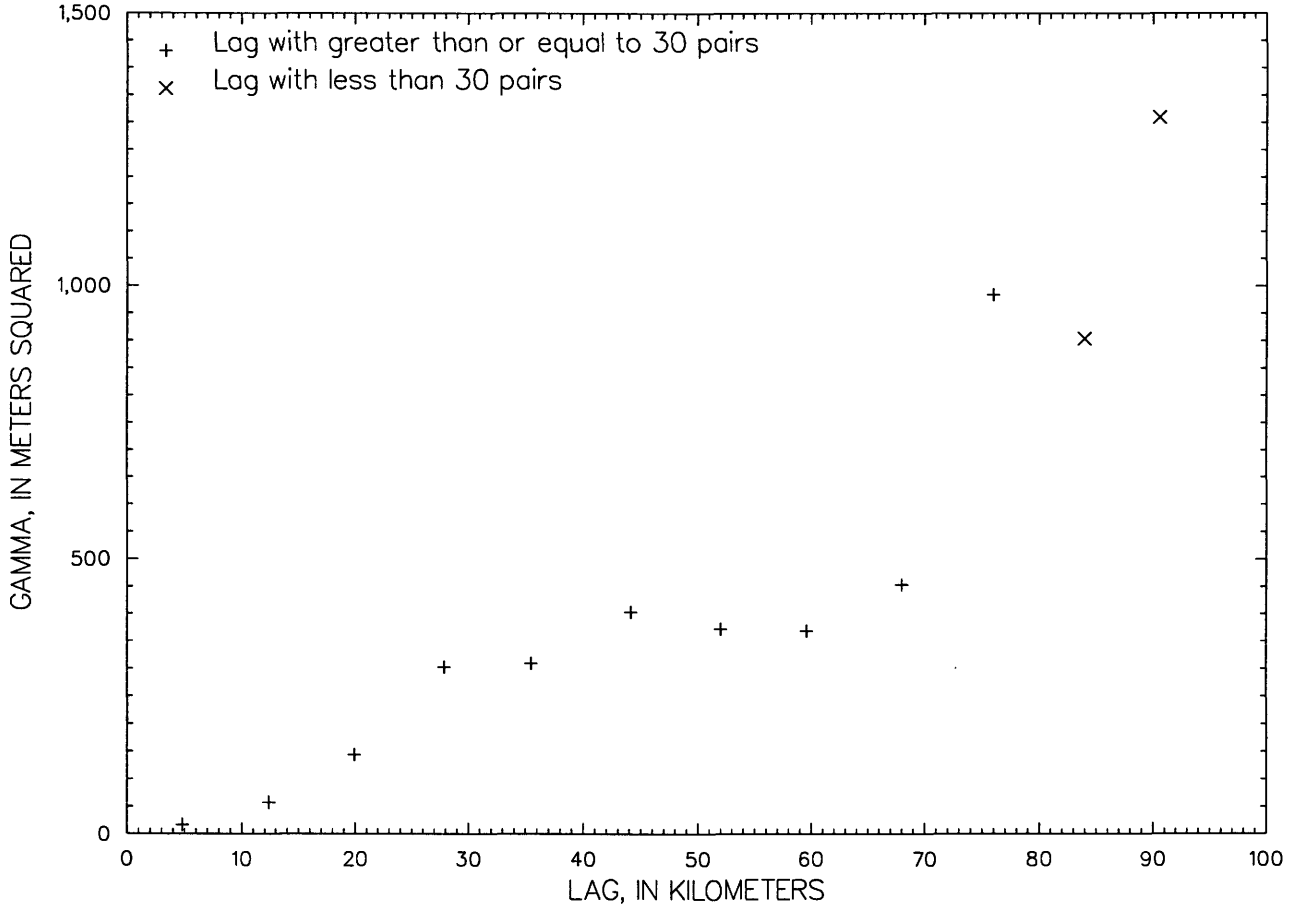


Figure 6. Initial sample variogram points for Saratoga data.

stage is adequate to indicate if nonstationarity needs to be addressed before sample variogram refinement is done.

5.2 Nonstationarity

An indication of substantial nonstationarity or drift in the spatial mean would be a parabolic shape through all lags in a plot of $\hat{\gamma}$. This shape occurs because differences between data contain differences in the drift component that increase as h increases. If equation 3–16 is inserted into equation 3–17, squaring the differences in μ greatly amplifies the increase with h. In this case, drift, generally a low-order (less than three) polynomial drift in (u,v), is fitted to the data and subsequently subtracted from the data to obtain residuals. Trend surfaces are not necessarily limited to polynomial forms. For example, a numerical model of ground-water flow may be used to obtain residuals of ground-water head data.

In theory, the polynomial trend indicates a slowly varying drift in the spatial mean and, as such, one regional trend surface should be fitted to all the data. However, often the drift and residuals are obtained locally; that is, using moving neighborhoods of locations. Therefore, estimates of these values at any point are made using a decreased number (usually between 8 and 16) of surrounding locations, which is done because, ultimately, the kriging estimates are made using only the data values in the given neighborhood. Manipulating the kriging matrices takes less time when a small number of data are used to make estimates, and these efficiencies can be substantial for dealing with large data sets. Little accuracy is lost because the nearest neighbors have the most effect in the kriging weighting scheme.

A parabolic shape for $\hat{\gamma}$ in the Saratoga data is shown in figure 6 for the sample variogram points plotted for lags up to about 32 kilometers (the first four points) and for lags greater than about 56 kilometers. A parabolic shape in the sample variogram points was not surprising because analysis of the data indicates a north-south gradient in the ground-water levels. The simplest polynomial trend, linear in u and v, was fitted to all the data using ordinary least-squares estimation. Residuals obtained by subtracting this regional trend surface from the data were used to reestimate $\hat{\gamma}$ in equation 5–2, and the sample variogram for the residuals is shown in figure 7.

5.3 Variogram Refinement

In section 5.1, an initial $\hat{\gamma}$ was specified by points computed from equation 5–2. In general, the larger $N(h_k)$ is for any bin k, the more reliable are the points defining $\hat{\gamma}(h_k)$. Also, the larger K is, the greater the number of sample variogram points shaping $\hat{\gamma}$. However, $N(h_k)$ and K are competing elements of $\hat{\gamma}$. Journel and Huijbregts (1978) suggested that each bin k could have $N(h_k)$ equal to at least 30 pairs. The American Society for Testing Materials (1996) suggested 20 pairs for each lag interval. For small data sets, the number of intervals may have to be small to guarantee either number of suggested pairs in all bins.

The minimum number of data, n, needed to satisfy the $N(h_k)$ requirements for all bins of a sample variogram is difficult to determine. Simple combinatorial analysis can establish a sample size needed for a given total number of distinct pairs obtained from the sample, but the analysis does not address the spatial considerations needed for proper lagging. For example, for data collected on a uniform grid and equal-sized bins, fixing an n to just satisfy the minimum $N(h_k)$ for the small lags would yield insufficient data pairs to meet the minimum $N(h_k)$ for the larger lags. Fixing an n to ensure the minimum $N(h_k)$ for the large lags would generally have $N(h_k)$ much greater than the minimum for the small lags. Therefore, the question of how much data are needed to adequately compute a variogram also needs to address the relative locations of the data-collection sites.

The first 10 of the 12 bins for $\hat{\gamma}$ for the Saratoga data contained more than 30 data pairs. Therefore, the bin width was decreased to have more points define the early part of $\hat{\gamma}$. These bin-width adjustments were made to refine $\hat{\gamma}$ whether it was computed from the data or from the residuals. A plot of $\hat{\gamma}$ for the residuals for the Saratoga ground-water elevations with the bin width narrowed to about 6.5 kilometers is shown in figure 8.

Spatial data usually are not collected on a uniform grid, but occur in a pattern that reflects problem areas, accessibility, and general spatial coverage. In the Saratoga data set, nonuniform data spacing resulted in the number of data pairs in each bin being highly variable among the bins, although there were still greater than 30 data pairs. This variability yields different reliabilities for the points

defining $\hat{\gamma}$. To establish a balance for $N(h_k)$ among the bins, variable bin sizes can be used so that each bin contains approximately the same (large) number of points. A bin having few points can be coalesced with an adjacent bin to form a wider bin having a large number of points. Conversely, a bin having an excessive number of points can be subdivided into adjacent, narrower bins. The coalescing and subdividing procedure is largely trial and error until the distribution of the pairs of points is satisfactory.

The values of $\hat{\gamma}$ at the small lag values are the most critical in defining the appropriate γ. Therefore, the trade-off between the number of bins and the number of data pairs within each bin can be varied for different regions of the sample variogram. At small lags, the numbers of data pairs per bin can be closer to the minimum $N(h_k)$ so more bins can be defined. At larger lags, a smaller number of wider bins may be adequate. Knowing that the variogram is a smooth function, the analyst visually decides when the sample variogram is sufficiently defined at all lags to adequately approximate a theoretical variogram.

5.4 Transformations and Anisotropy

5.4.1 Transformations

A transformation is applied to a data set generally for one of two interrelated purposes. First, a transformation can decrease the variability of highly fluctuating data. This variability especially occurs with contaminant concentrations where order-of-magnitude changes in data at proximate sites are not uncommon. The effects of such data would be erratic sample variogram points indicated by a large amplitude, ill-defined sawtooth pattern of the lines connecting the points.

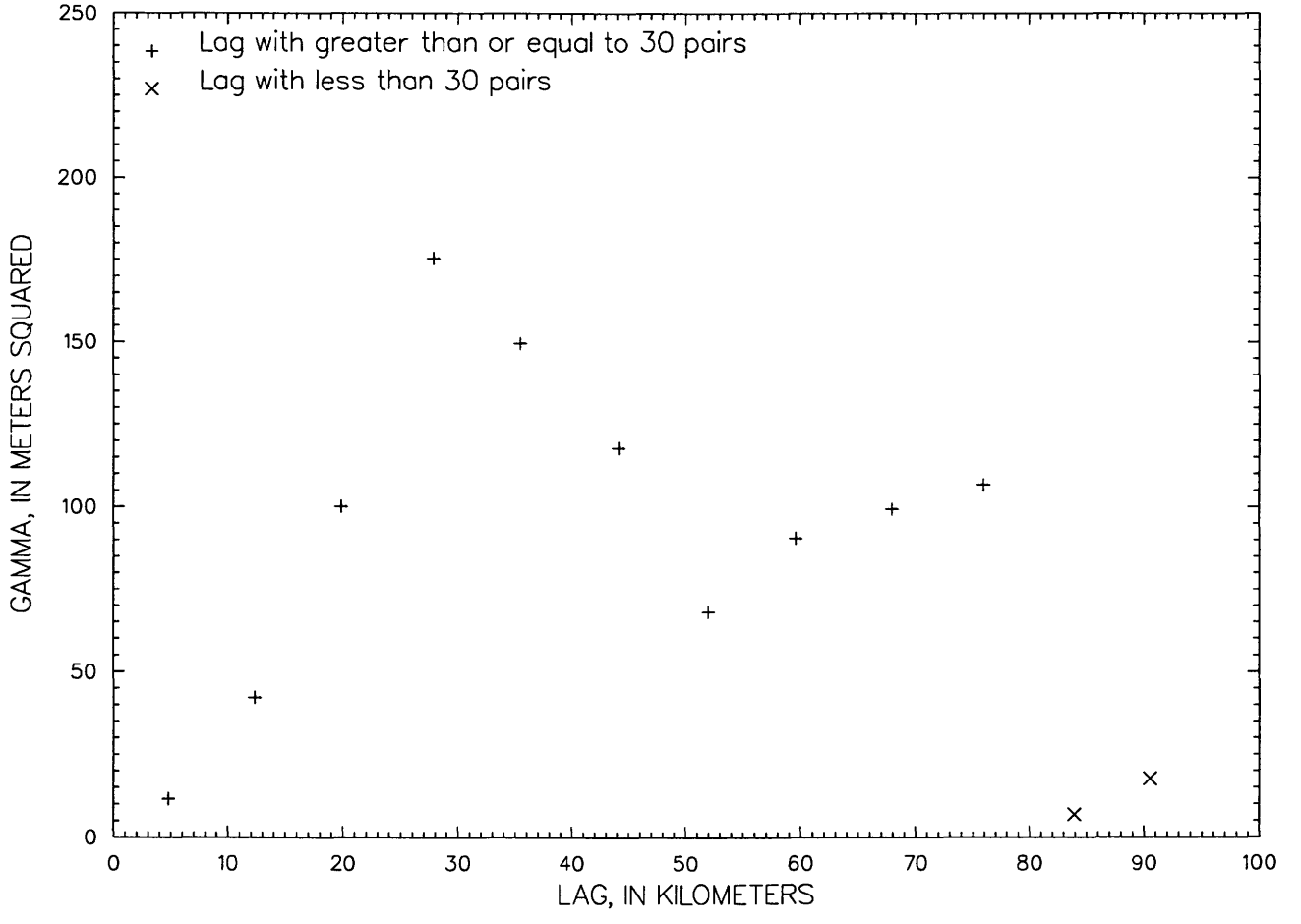


Figure 7. Sample variogram points for ordinary least-squares trend residuals for Saratoga data.

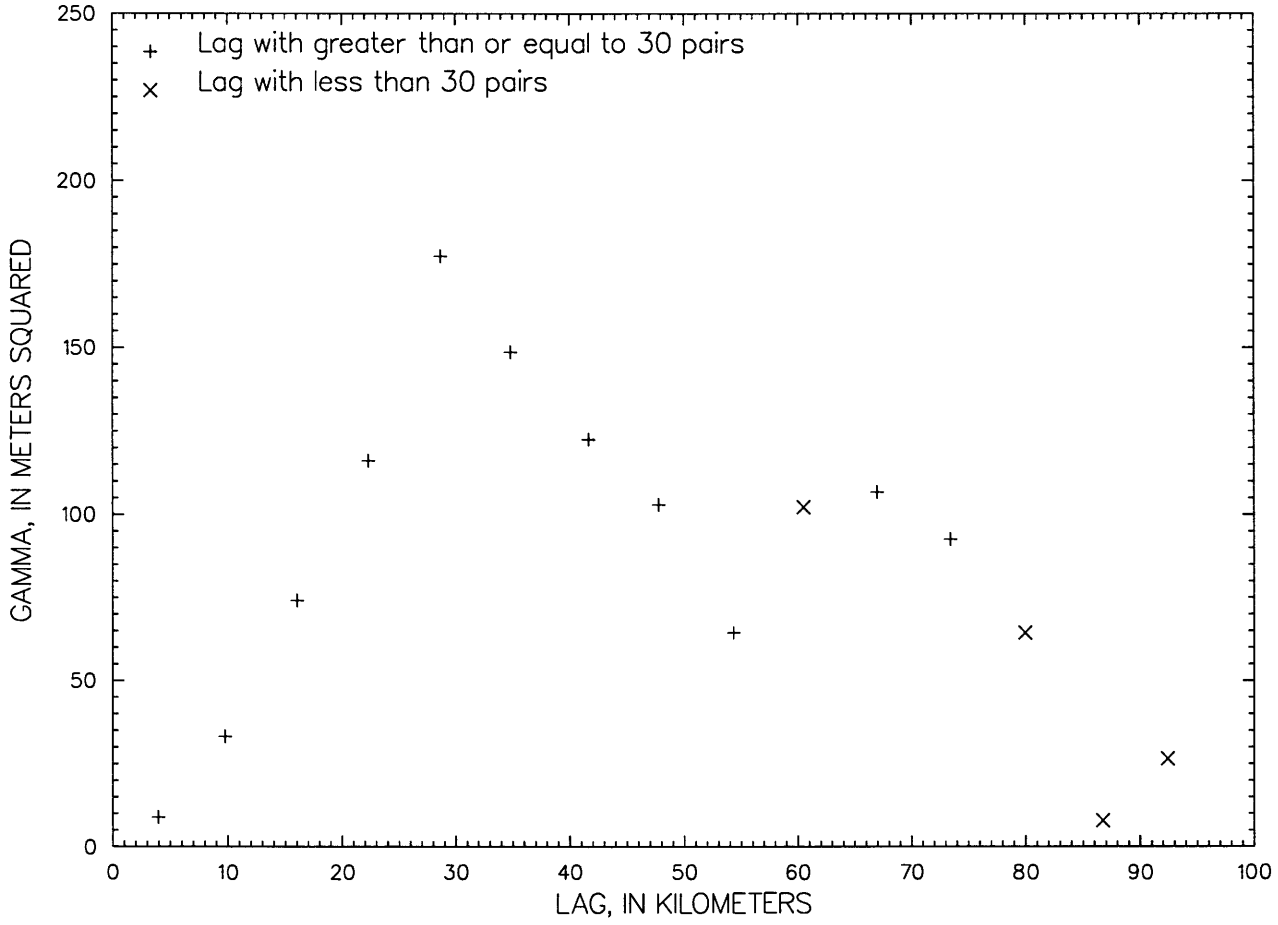


Figure 8. Sample variogram points for ordinary least-squares trend residuals for Saratoga data binned to 6.5 kilometers.

Second, a proper transformation of data, whose probability distribution is highly skewed, often produces a set of values that can be approximately normally distributed by mitigating the effect of problematic extreme data. A data set having a normal distribution is important in kriging when confidence levels of the estimates are desired. The use of confidence levels in a kriging analysis is discussed in section 6.0.

Among the more common transformations is the natural logarithmic (log) transform. For example, in this transformation, the $\hat{\gamma}$ is the sample variogram of logarithms, and subsequent kriged estimates are logarithms. Another transformation that is often used, especially in spatial analyses of contaminant levels, is the indicator transformation described in section 3.5.2. Although a transformation might result in a better distribution of sample variogram points, there are subtleties in interpreting the kriging results of the transformed data or in back-transforming kriging results into the untransformed (original) units, as discussed in section 3.5.1. If a satisfactory variogram of the original data cannot be achieved and a transformation is indicated, the computation of a sample variogram needs to begin again with equation 5.2. Even though no transformation was needed for the Saratoga data, an example using a logarithmic transformation and an example using the indicator transformation are presented in section 6.0.

5.4.2 Directional Variograms and Anisotropy

Anisotropy in the data can be investigated by computing sample variograms for specific directions. Locations included in a given direction from an original location are contained in a sector of a circle of

radius h_{max} centered on the original location. The sector is specified by two angular inputs. The first is a bearing defining the specific direction of interest [measured counterclockwise from east (= 0 degrees)] and the second is a window angle defining an arc in both directions from the bearing. Thus, in the terminology used here, the total angle defining a direction is equal to twice the window angle. Differences in sample variograms computed using these angle windows specified for different directions can be an indication of anisotropy.

Anisotropy is generally either geometric or zonal. Geometric anisotropy is indicated by directional theoretical variograms that have a common sill value but different ranges. The treatment of geometric anisotropy is dependent on the software used. The lags of the directional variograms can be scaled by the ratio of their ranges to the range of a standard or common variogram. In some cases, the lags of all directional variograms are scaled by their respective ranges, and a common variogram that has a range parameter of 1 is used. Ground-water contaminant plumes often have geometric anisotropy in which the prevailing plume direction has a greater range than the range of the transect of the plume.

Zonal anisotropy is indicated by directional variograms that have the same range but different sills. Pure zonal anisotropy is usually not seen in practice; generally, it is found combined with geometric anisotropy. Such mixed anisotropy may be present when evaluating the variograms of three-dimensional HTRW-sampling results. Variability of such data (as indicated by the sill of the variogram) may be substantially higher and the range substantially shorter in the vertical direction than in the horizontal direction. To model this mixture of anisotropic variograms, the overall variogram is set to a weighted sum of individual models of the directional variograms and scaled by their ranges. This process is called nesting, in which the choice of weights uses a trial-and-error approach with a constraint that the sum of the weights equals the sill of the overall variogram. Isaaks and Srivastava (1989, p. 377–390) contains further information on both types of anisotropy.

For a given number of data locations, directional sample variograms will have fewer points for any lag when compared to the points for the same lag in the omnidirectional variogram. Hence, point values in directional variograms are less reliable, which could be a critical constraining factor for small data sets or for a data pattern that does not conform to a direction of anisotropy. To determine the adequacy of the data for determining anisotropy, the computations of anisotropic sample variograms can be initially limited to two orthogonal directions with window angles of 45 degrees.

Directional sample variograms also can be used to further delineate nonstationarity of the spatial mean. If the omnidirectional sample variogram indicates a drift in the data, the directional variograms may determine the dimension of the drift. That is, although the directional sample variograms may not establish the degree of the polynomial in the drift equation, they can indicate the relative strengths of the drift in the u and v directions.

The computed sample variograms for north-south and east-west directions and window angles of 45 degrees for the Saratoga data are shown in figure 9. The north-south variogram is specified by a direction angle of 90 degrees and a window angle of 45 degrees. The north-south variogram shows the preferential north-south data alignment by mimicking the omnidirectional (direction angle = 0 degrees and window angle = 90 degrees) sample variogram in figure 6. The east-west variogram is specified by a direction angle of 0 degrees and a window angle of 45 degrees. The lack of pairs of locations for the east-west variogram precludes a good analysis for this direction, but the overlap of the few sufficiently defined variogram points with the north-south variogram indicates a consistency of drift in the two directions. Because of this consistency, an isotropic variogram is assumed for the Saratoga residuals. An example of kriging using anisotropic variograms is described in section 6.0.

5.5 Fitting a Theoretical Variogram to the Sample Variogram Points

The importance of adequately defining the bin values of a sample variogram is substantiated by the need to accurately generalize the data-based behavior of the sample variogram by a theoretical variogram, γ. The parameters controlling the specific behavior of theoretical variograms are the nugget value, the range, the sill, or for case of a linear variogram, a slope. Of these parameters, the nugget and the sill can be related to properties and statistics of the data.

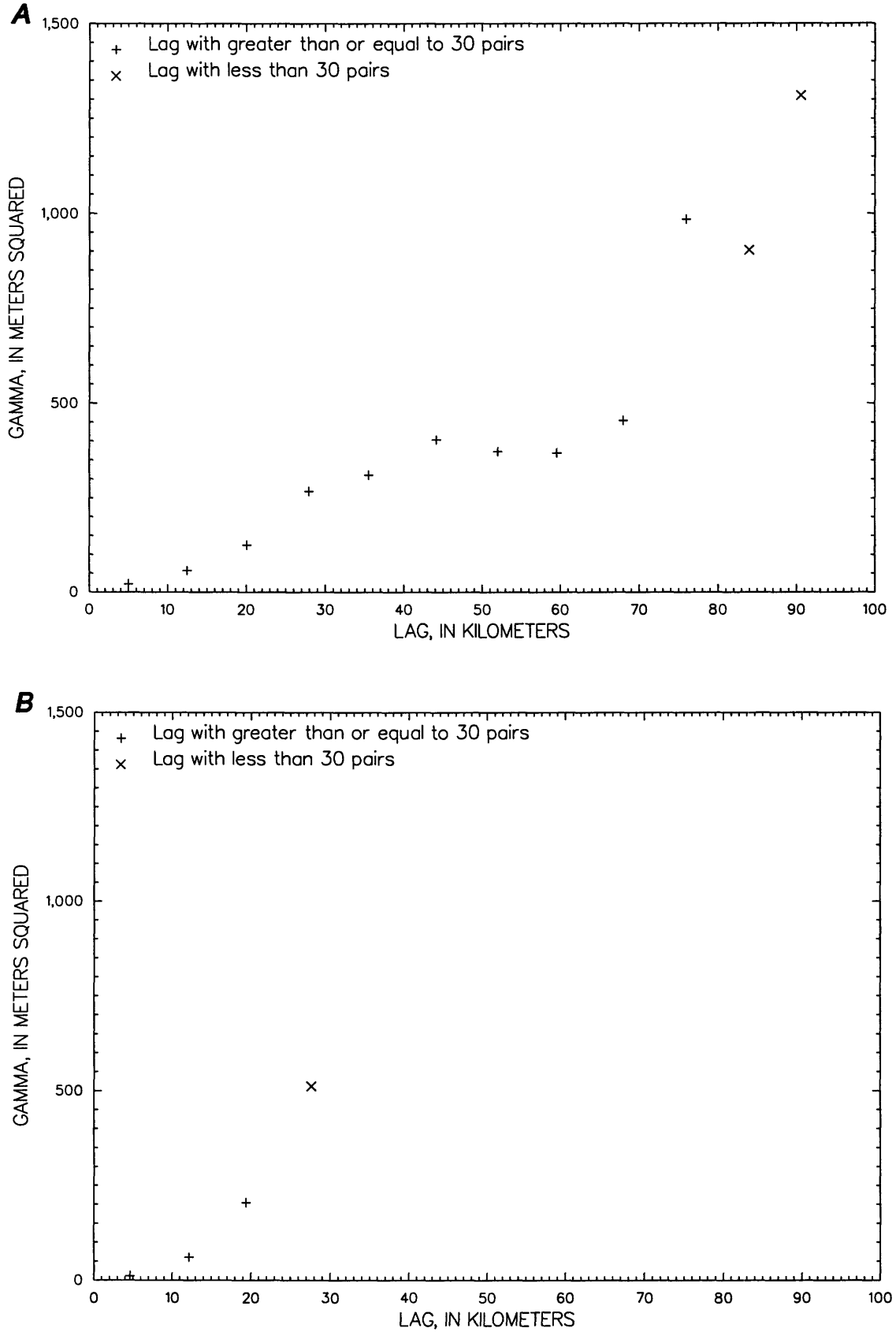


Figure 9. Initial directional sample variogram points for raw Saratoga data—*A*, north-south and *B*, east-west.

The nugget is essentially the extrapolation of the sample variogram to a lag of zero. It indicates the uncertainty of the variogram at lags that are smaller than the minimum separation between any two data locations. The nugget can include measurement error variance, and an estimate of this variance approximates a minimum value of the extrapolation.

The sill determines the maximum value of a variogram and approximates the variance of the data. However, the points defining $\hat{\gamma}$ take precedence over the sample variance in determining the sill. Some variograms are unbounded, and others may only reach a sill value asymptotically. A defined sill allows conversion of the variogram to a covariance function using equation 3–27, which is generally done because computations in the kriging algorithms are more efficiently performed using a covariance function.

Fitting a function to the sample variogram values can range from a visual fit to a sophisticated statistical fit. A statistical fit is an objective method as long as the choice of bins and the weighting of the sample variogram points remain fixed. However, because the inputs vary, inherent subjectivity persists as in a visual fit. A final calibration of the variogram parameters is based on the kriging algorithm; thus, either of the initial fitting methods at this stage would suffice.

Because the initial part of the variogram has the most effect on subsequent kriging output, a good estimate of the nugget value becomes a most important first step. The range and the sill, in that order, complete the ranking of the effect of variogram parameters on the output of a geostatistical analysis. Whatever the fitting method used, the theoretical variogram needs to be supported by the sample variogram values. For variograms that have a range parameter, this support could extend to the range. Journel and Huijbregts (1978) suggested that this support would be through one-half the dimension of the field or essentially through one-half the maximum lag distance of the sample data.

Most geostatistical studies can be successfully completed using one of the following four singular theoretical variogram forms: exponential, spherical, Gaussian, and linear functions (fig. 3); however, positive linear combinations of these forms also are acceptable as theoretical variograms (see section 5.4.2). Geometric relations for obtaining parameters for the four variogram forms are described in the following sections and are shown in figure 3.

5.5.1 Exponential Variogram

The exponential variogram (eq. 3–23) is specified by the nugget, g; sill, s; and a practical range, r. The range is qualified as practical because the sill is reached only asymptotically. From the nugget, the sample variogram points indicate a convex behavior that persists through all lags, although to a much lesser degree at larger lag values. A nugget and a sill are first specified based on the $\hat{\gamma}$ points. The practical range is chosen so the value of the resulting exponential function at the practical range lag is 95 percent of the sill. The specified exponential function meshes with the sample variogram points at least through the practical range lag. An initial estimate of the practical range can be made if the intersection of the sill with a line tangent to the variogram at the nugget is at a lag value equal to one-third of the assumed practical range as shown in figure 3. Examples of the exponential variogram are included in spatial studies of sulfate and total alkalinity in ground-water systems (Myers and others, 1980).

5.5.2 Spherical Variogram

The spherical variogram parameters (eq. 3–24) are a nugget, g; a range, r; and a sill, s. At small lag values, the sample variogram points indicate linear behavior from the nugget that then becomes convex and reaches a sill at some finite lag (fig. 3). A sill is estimated, and a line drawn through the points of the initial linear part of the variogram intersects the sill at a lag value approximately equal to two-thirds of the range. With these estimates of the parameters, a spherical variogram is defined that should be supported by the sample variogram points. If the spherical variogram does not plot near the sample variogram points, adjustments need to be made to the parameter estimates and the subsequent fit evaluated. Although the spherical variogram model is one of the most often used models for real-valued spatial studies, it also seems to be a predominant model for indicator values at various cutoff levels as, for example, in a study of lead contamination (Journel, 1993).

5.5.3 Gaussian Variogram

The Gaussian variogram parameters (eq. 3–25) are a nugget value, g, and a sill, s; and this variogram also has a practical range, r. The Gaussian variogram is horizontal from the nugget, becomes a concave upward function at small lags, inflects to concave downward, and asymptotically approaches a sill (fig. 3). After a nugget and sill are specified based on the $\hat{\gamma}$ points, the variogram value at a lag of one-half the estimated practical range is two-thirds of the sill. Again, this fitted variogram needs to be supported by the $\hat{\gamma}$ points to a reasonable degree. As is described in the example using the Saratoga data, the Gaussian variogram often is used where the analyzed variable is spatially very continuous, such as a ground-water potentiometric surface.

5.5.4 Linear Variogram

Parameters for a linear variogram (eq. 3–26) are a nugget value, g, and a slope, b. Sample points that indicate a linear variogram increase linearly from the nugget and fail to reach a sill even for large lags (fig. 3). Using the nugget as the intercept, the slope is computed for the line passing through the $\hat{\gamma}$ points. A pseudosill can be defined as the value of the line at the greatest lag, h_{max}, between any two locations. This lag becomes the defacto range, r, for a linear variogram. Examples of the use of the linear variogram are in hydrogeochemical studies of specific conductance and in studies of trace elements, such as barium and boron (Myers and others,1980).

5.6 Additional Trend Considerations

If a drift in the data is indicated as in section 5.2, the theoretical variogram of residuals that has been fitted thus far is used to update the drift equation. Although ordinary least squares often suffices for computing a polynomial drift equation, drift determination is a function of γ when the data are spatially correlated. But γ cannot be estimated until a drift equation is obtained to yield the residuals. Therefore, obtaining a sample variogram and a subsequent theoretical variogram from drift residuals of a specified drift form is an iterative process (David, 1977, p. 273–274) using the following steps:

1. An initial variogram is specified and drift coefficients are computed to obtain residuals. For this step, a pure nugget (that is, a constant) variogram can be used to compute the initial estimates of the drift coefficients. These initial coefficients yield an ordinary least-squares estimate of the drift and a first-iteration sample variogram of residuals.
2. A theoretical variogram is fitted to the sample variogram of the residuals and is used to obtain updated drift coefficients.
3. The residuals from the drift that were obtained in step 2 are used to compute an updated sample variogram.
4. The sample variogram computed at the end of step 3 is compared to the sample variogram from step 2. If the two sample variograms compare favorably, then the theoretical variogram from step 2 is accepted as the variogram of residuals for subsequent kriging computations. If the sample variogram from step 3 differs markedly from the sample variogram from step 2, steps 2 through 4 are repeated using the sample variogram from the most recent step 3.

Generally, the plot of the points of $\hat{\gamma}$ from a set of residuals initially increases with h, reaches a maximum, and then decreases as shown in figure 7. This typical haystack-type behavior, discussed by David (1977, p. 272–273), is attributed to a bias resulting from the estimation error in the drift and its coefficients. This behavior in the variogram of the residuals generally would more readily occur with a high degree of drift polynomial and need not prohibit acceptable variogram determination because the initial points of the sample variogram of residuals are still indicative of the theoretical variogram. For example, the lag associated with the maximum of $\hat{\gamma}$ of the residuals can be a good first approximation for the range of the theoretical variogram.

5.7 Outlier Detection

Outliers in a data set can have a substantial adverse effect on $\hat{\gamma}$. However, divergent data can be screened for evaluation using a Hawkins statistic (Hawkins, 1980), which is described in the context of kriging by Krige and Magri (1982). A neighborhood

containing 4 to 10 data points that are approximately normally distributed around each suspected outlier needs to be defined. Despite potential outliers in the data set, a best guess initial theoretical variogram also is needed.

The Hawkins statistic is obtained by comparing a suspect datum to the mean value of the 4 to 10 surrounding data, the smaller number being sufficient if the variability is low. The spacing between these surrounding points is accounted for by the properties of the chosen variogram. A statistic of 3.84 or higher would indicate an outlier on the basis of a 95-percent confidence interval. A large number of surrounding points has the direct effect of increasing the magnitude of the statistic. Anomalous points are removed from the data set, and the procedure described for obtaining the sample variogram is repeated for the small data set. There were no outlier problems in the Saratoga data.

There is debate among geostatisticians regarding the merit of automated outlier-detection procedures. A procedure such as that described is presented as an investigative tool with the understanding that attendant justification and a Hawkins-type statistic need to be used to ultimately decide if a data value is discarded as a true outlier or retained as a valid measurement. In some situations, highly problematic data are removed for computation of the sample variogram points, but are reinstated for kriging.

5.8 Cross Validation for Model Verification

The parameters of the theoretical variogram obtained from the initial fitting and refinement of the sample variogram are calibrated using a kriging cross-validation technique. In this technique, the fitted theoretical variogram is used in a kriging analysis in which data are individually suppressed and estimates are made at the location using subsets of the remaining data. As described in section 5.2, these subsets are the data points in a moving neighborhood surrounding the location under consideration. The calibration estimate made at each data location needs a matrix inversion, which could be very time consuming if all the data locations were used to construct the matrices rather than just the data within a neighborhood of a limited search radius.

After kriged values at all data locations have been estimated in the above manner, the data are used with their kriged values and the kriging standard deviations to obtain cross-validation statistics. A successful calibration is based on criteria for these statistics, which are described in section 5.8.1. If the criteria cannot be reasonably met by adjusting the parameters in the given theoretical variogram function, then the calibration needs to be reinitialized using a different theoretical variogram model. In some data sets that have a nonstationary spatial mean, the drift polynomial and the variogram may have to be changed to achieve a satisfactory calibration.

5.8.1 Calibration Statistics

The kriging cross-validation error, e_i, which corresponds to measurement $z(\underline{x}_i)$, is defined as

$$e_i = z(\underline{x}_i) - \hat{z}(\underline{x}_i) \tag{5–3}$$

where

$\hat{z}(\underline{x}_i)$ is the kriged estimate of $z(\underline{x}_i)$.

The kriged estimate is obtained by ordinary kriging if the spatial mean is constant or by universal kriging if the spatial mean is not stationary. A reasonable criterion for selecting a theoretical variogram is to minimize the squared errors, Σe_i^2, with respect to the variogram parameters. However, unlike ordinary least-squares regression, which also minimizes the sum of squared errors, simply minimizing the squared errors is not sufficient for kriging because the resulting model can yield inconsistent estimates of the kriging variances, $\sigma_K^2(\underline{x}_i)$ at location $\underline{x}_i$. This simple minimization would give unrealistic measures of the accuracy of the kriging estimates. To guard against such bias, an expression for the square of a reduced kriging error is defined:

$$\tilde{e}_i^2 = \frac{e_i^2}{\sigma_K^2(\underline{x}_i)} \tag{5–4}$$

where the kriging variances are computed using either equation 3–36 or equation 3–47. If the kriging variance is a consistent estimate of the true mean-squared error of estimate, then the reduced kriging errors have an average of about 1. Therefore, the standard cross-validation technique for evaluating a theoretical variogram is:

$$\min \left(\frac{1}{n} \sum_{i=1}^{n} e_i^2 \right)^{0.5} \tag{5–5}$$

$$\text{subject to} \left(\frac{1}{n} \sum_{i=1}^{n} \tilde{e}_i^2 \right)^{0.5} \approx 1 .$$

The expression to be minimized is called the kriging root-mean-squared error and the constraint is called the reduced root-mean-squared error. The reduced root-mean-squared error needs to be well within the interval having endpoints

$$1 + \left(2\sqrt{\frac{2}{n}} \right)$$

and

$$1 - \left(2\sqrt{\frac{2}{n}} \right)$$

(Delhomme, 1978). An additional check on the goodness of the cross-validation results is the unbiasedness condition where

$$\frac{1}{n} \Sigma e_i \approx 0 .$$

As indicated in section 3.0, if probabilistic statements concerning an actual value of Z at an unmeasured location are to be made compared to the kriged estimate and the kriging variance at the location, the distribution of the cross-validation kriging errors needs to be analyzed. In particular, the reduced errors, $\tilde{e}_i$, $i = 1,2,...,n$, need to be approximately normally distributed with mean 0 and variance 1. A histogram or normal probability plot of the reduced kriging errors can be used to assess the validity of assuming a standard normal distribution for the reduced kriging errors. Additionally, if the distribution of reduced kriging errors can be assumed to be standard normal, outliers not detected using the technique discussed in section 5.6 may be detected by comparing the absolute values of the reduced kriging errors to quantiles of the standard normal distribution.

Using the Saratoga data, a spherical variogram was fitted to the refined sample variogram of the residuals. The estimated nugget was about 1.49 meters squared, the sill was 133.8 meters squared, and the range was about 48 kilometers. Because of the difficulty in determining an exact extrapolated value for the nugget, the value of 1.49 meters squared was selected based on an estimated measurement error related to obtaining water levels at the well depths in the Saratoga Valley.

After two iterations using drift residuals, as described in section 5.6, a final variogram was chosen that had a nugget of 1.49 meters squared, a sill of 148.6 meters squared, and a range of 44.8 kilometers (fig. 10). These parameters defined the theoretical variogram used to obtain the cross-validation errors through universal kriging with an assumed linear drift. The best combination of statistics that could be obtained after several attempts at refining the model were a root-mean-squared error of 3.45 meters and a reduced root-mean-squared error of 0.5794. The reduced root-mean-squared error is too small, indicating that the kriging variances produced by the model are relatively large compared to the actual squared errors. This fact, coupled with the rather large root-mean-squared error, warranted additional variogram refinements. In section 5.8.2, a Gaussian variogram was fitted to the data; the Guassian variogram produces much better cross-validation results than the results from the spherical variogram.

5.8.2 Variogram-Parameter Adjustments

If any of the cross-validation statistics vary unacceptably from their suggested values, minor adjustments to the variogram parameters can be made to attempt to improve the statistics. Modifications made to the parameters should not have to be so severe that the variogram function drastically deviates from the sample variogram points. If the support of the sample variogram points is compromised to achieve acceptable cross-validation results with the given drift-variogram model, a different drift-variogram combination needs to be investigated.

A reduced root-mean-squared error that is unacceptable may be improved by adjusting the range parameter or the nugget value of the variogram. Modifying the range parameter needs to be considered first, and any shifts in the nugget value need to be minimal and made only as a final recourse. The calibration errors are relatively insensitive to minor adjustments of the sill.

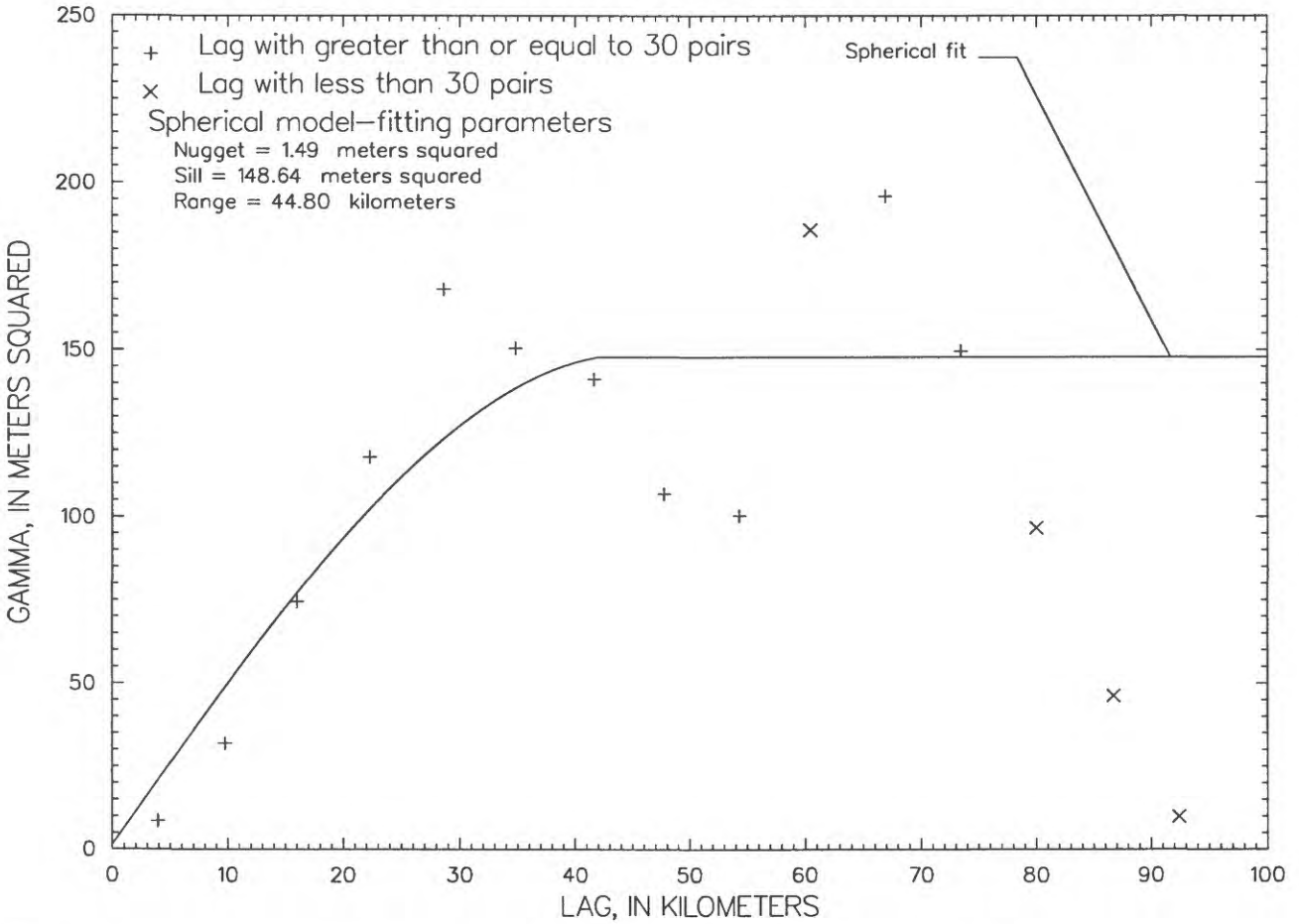


Figure 10. Sample variogram points and theoretical spherical fit for iterated Saratoga residuals.

If the reduced root-mean-squared error is too small, as in the Saratoga example, extending the range (equivalent to decreasing the slope for a linear variogram) decreases the kriging variance and, thus, increases the reduced root-mean-squared error. If a shift in the nugget is needed, a decrease in the nugget decreases the kriging variance. If the reduced root-mean-squared error is too large, then a contraction of the range or a positive shift in the nugget can be made, based on the priority and the extent of the changes. Generally, changes in these parameters also have an effect on the mean-squared error. The larger the nugget is as a percentage of the sill, the larger the mean-squared error is. In general, improvements in one statistic are usually made at the expense of the other statistics. The optimization of the statistics as a set is, in effect, a trial-and-error procedure that is operationally convergent.

The reduced kriging errors may not approximate a standard normal distribution. If so, a transformation of the data may be needed to achieve a more normal distribution, and the variogram estimation procedure would be repeated.

Because no convergence could be reached for parameter values of a spherical variogram for the Saratoga data, a Gaussian theoretical variogram was fitted to the sample variogram of residuals in figure 8. This choice was made because the initial sample variogram points seemed to have a slight upward concavity, but eventually reached a sill. This behavior can be attributed to correlation rather than to further drift. After an iterated cross validation using the Gaussian parameters, a Gaussian variogram that had a nugget of 1.49 meters squared, a sill of 185.81 meters squared, and a range of 27.52 kilometers (fig. 11)

yielded a root-mean-squared error of 2.33 meters and a reduced root-mean-squared error of 1.083 meters. The mean cross-validation error was 0.0195 meter. These values represented an improvement over the spherical variogram and were deemed acceptable for the Gaussian variogram.

A probability plot of the reduced kriging errors using the final Gaussian variogram is shown in figure 12. The plot is reasonably linear between two standard deviations and, thus, approximates a standard-normal-distribution function. A plot in figure 13 of the measured data versus their kriged estimates indicates that the linear drift/Gaussian variogram model selected for the Saratoga data would produce accurate estimates of ground-water elevations for interpolation or contour gridding in the region.

6.0 PRACTICAL ASPECTS OF GEOSTATISTICS IN HAZARDOUS-, TOXIC-, AND RADIOACTIVE-WASTE-SITE INVESTIGATIONS

In this section, several example applications are described. The applications have been developed using hydrologic, geologic, and contaminant data from established and well-studied hazardous-waste sites. The real nature of the data enables discussion of some problems that can occur during HTRW-site investigations that originate, not only from natural field conditions, but also from typical problems that are associated with the types of data involved. In addition, the real nature of the example data provides an opportunity for comparison between kriging estimates and the real data; these comparisons are brief

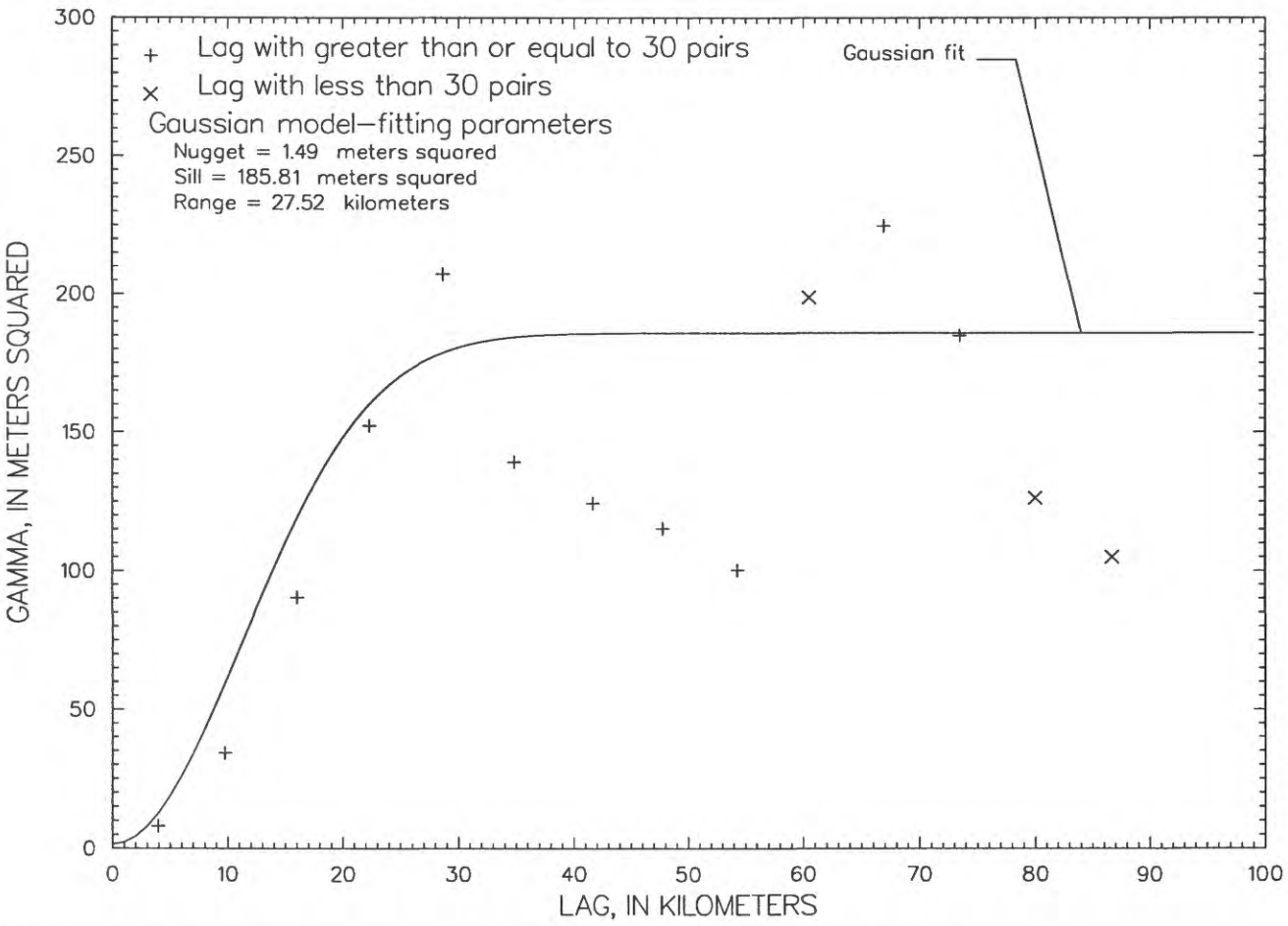


Figure 11. Sample variogram points and theoretical Gaussian fit for iterated Saratoga residuals.

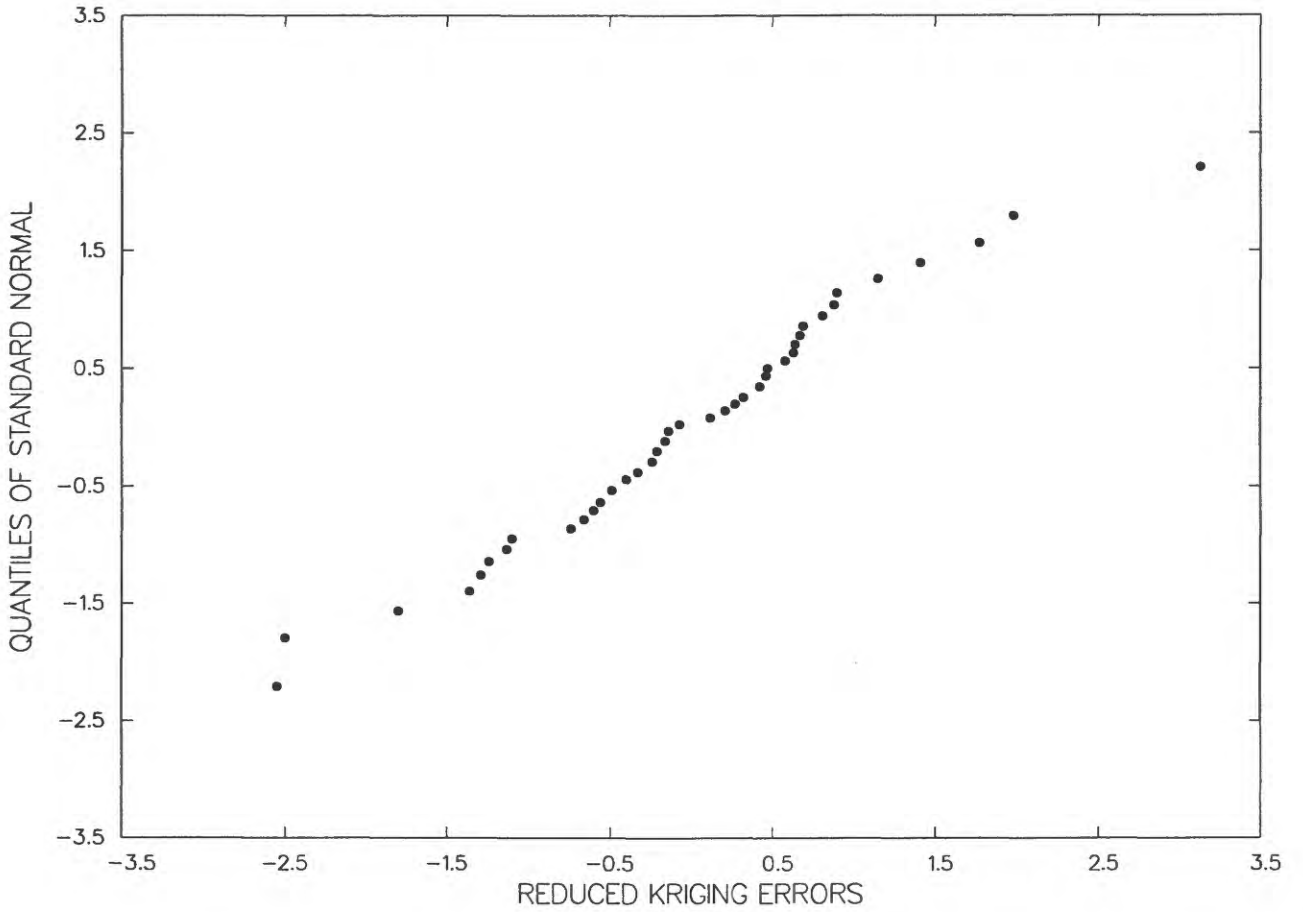


Figure 12. Cross-validation probability plot for Saratoga data.

and general. This report does not provide comprehensive analyses of data that are available in other more elaborate studies.

The principal intent of the examples is to provide systematic descriptions for a few of the large number of possible applications that may be used during HTRW-site investigations. The examples are not intended to provide guidance for comprehensive analysis of the included data. However, this report presents some fundamental problems that can occur in geostatistical applications and, in some examples, indicates some possible alternatives.

With each example, a purpose is established and a general environmental setting is described. Most aspects of variogram construction and calibration are briefly described and are shown in figures and listed in tables. A comprehensive treatment of variogram construction has been presented in section 5.0.

The GEO-EAS software has been used whenever the example data did not require universal kriging; for those examples, STATPAC was used. As indicated in section 4.0, both of these software packages run on the DOS platform (table 1), which is probably most convenient to readers. The results of kriging estimates are portrayed by gray-scale maps rather than by contours because of the objective nature of the gray-scale format. North is at the top of all maps presented although this orientation may represent some deviation from the real data.

6.1 Ground-Water-Level Examples

The principal purpose of the ground-water-level examples is to familiarize the reader with a kriging exercise using ground-water levels and to indicate simply how kriging standard deviations may be useful in evaluating monitoring networks. The data are from

a water-table setting in unconsolidated sediments where the local relief for the land surface is about 30 meters. The data involved in this example are considered virtually free of actual measurement error.

The location of measured water levels is shown in figure 14*A*, and the basic univariate statistics for this data set are listed in table 2 (water level A); modifications to the measured data, in the form of removal and addition of measured values, are shown in figures 14*B* and *C*. The techniques described in section 5.0 were used to guide the following steps for variogram construction:

1. A raw variogram analysis and basic hydrologic knowledge of water-level behavior indicated that universal kriging would be needed for this analysis.
2. To obtain a stable variogram of residuals, an iterative, generalized least-squares operation was initially used to remove prominent linear drift of the form $a + bu + cv$, observed in the measured water levels, where a, b, and c are constants determined in the iterative process.
3. After drift was removed, residuals were determined to be stationary and universal kriging with a linear drift was appropriate.
4. A Gaussian model was used to fit the stabilized variogram of residuals (fig. 15*A*), which has a nugget of 0.09 meter squared, a sill of 2.69 meters squared, and a range of 1,219 meters (table 3).

Cross validation was performed, and the results are shown in figures 15*B* and *C* and listed in table 3. The cross-validation statistics conform to the criteria discussed in section 5.0.

Linear drift is commonly observed in ground-water-elevation data where there are no major anthropogenic activities, such as large ground-water withdrawals.

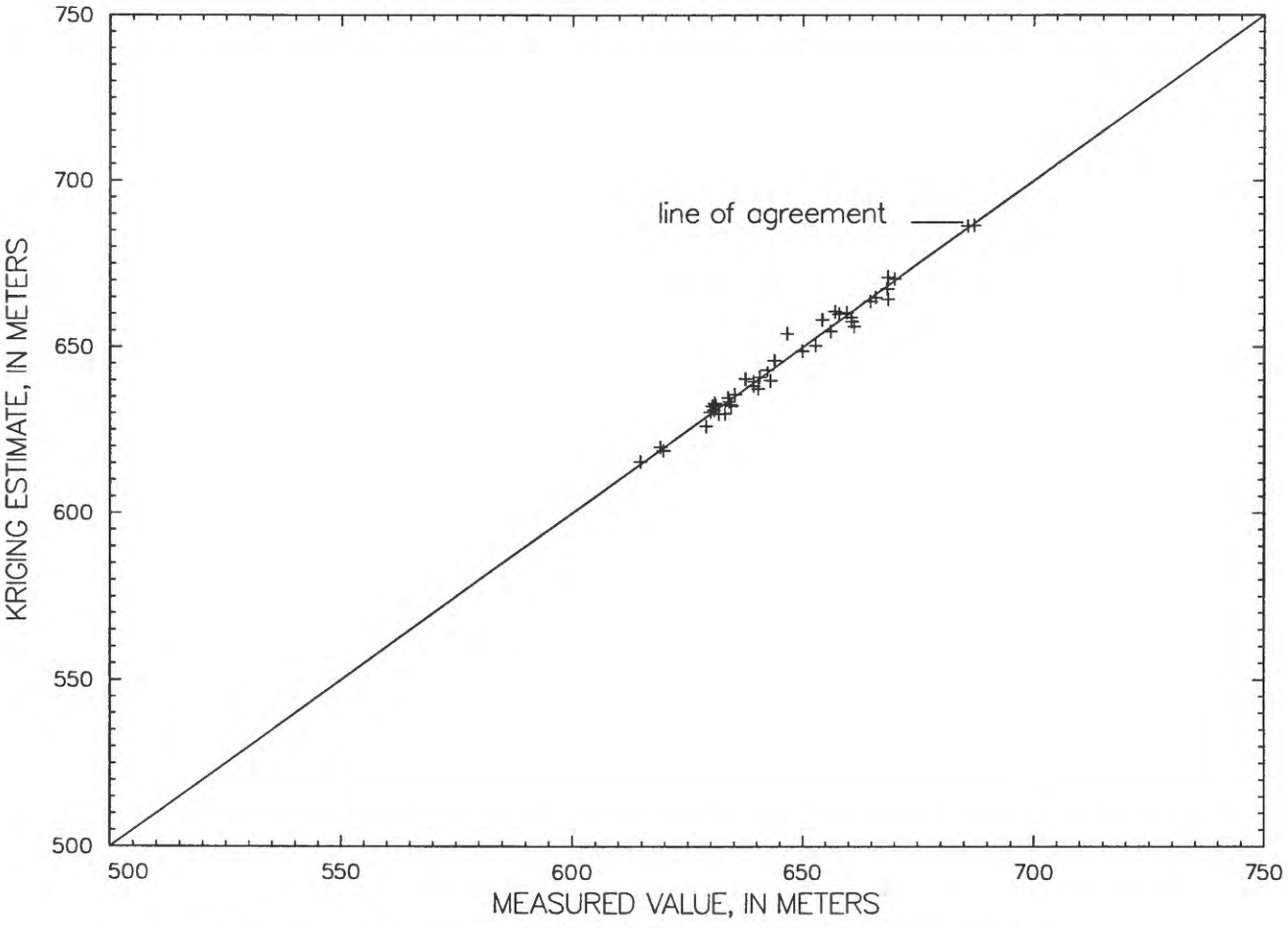


Figure 13. Scatterplot of measured versus kriging estimates from cross validation of Saratoga data.

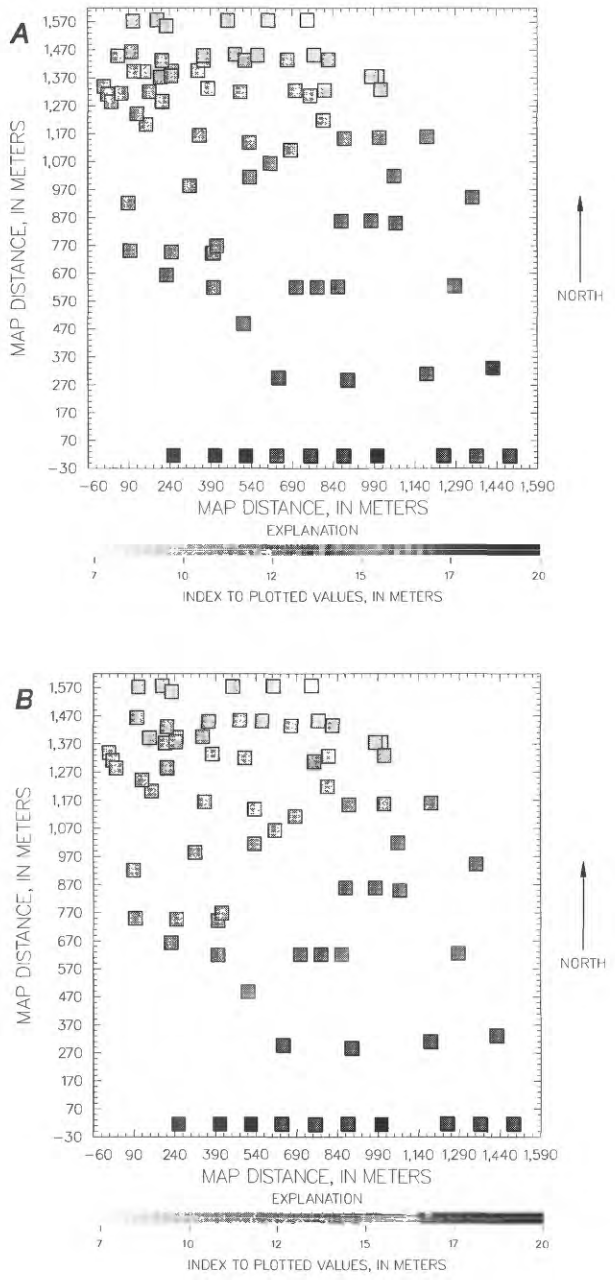


Figure 14. Location of measured data for ground-water-level examples—*A*, original data; *B*, original data without dropped sites; and *C*, original data with added sites (added sites indicated with +).

Under these circumstances, there is usually a fairly uniform and general ground-water movement along a flow path. This uniform and general nature introduces a nonstationary element to the data that, in geostatistics, is referred to as drift. As indicated in section 5.0, the presence of drift is indicated by a parabolic variogram shape. In this example, the initial variogram in the raw variogram analysis had a characteristic parabolic shape, and a linear drift was identified. Once the drift was identified and characterized, universal kriging procedures were used.

A Gaussian model is usually appropriate for variograms of highly continuous variables, such as ground-water-elevation data, and this model is particularly appropriate in this example. The variogram (fig. 15*A*) at small lags beyond the nugget has an upward concavity that cannot be fit with a linear, spherical, or exponential model. The observed shape was interpreted as a function of continuous small-scale variability. The Gaussian model fits the bowl shape of the small lag data and other data well to a lag of about 610 meters, but it is not flexible enough to closely fit the points much beyond 610 meters, indicating that kriging estimates should be computed using neighborhoods with a search radius less than 610 meters. In section 5.0, the initial part of the variogram was described as having the most effect on subsequent kriging estimates.

The established variogram then was used with the measured data to produce universal kriging estimates for all points in a 26-by-26 grid that had a grid size of about 61-by-61 meters. A gray-scale map of the kriging water levels is shown in figure 16*A* and basic univariate kriging estimate statistics are listed in table 4 (water level A). The kriging results are a good representation of the results from other more elaborate studies.

The kriging standard deviations for the kriging estimates are shown in figure 16*B*. The magnitude of kriging standard deviations can provide investigators with a direct indication of where the uncertainty associated with kriging estimates is relatively high or low. The areas of the greatest uncertainty for the kriged water levels are in the upper right and the lower left corners of figure 16*B*, where standard deviations are as high as about 1.4 and 0.8 meters. These areas are where the density of the measured data is relatively low. Throughout much of the remainder (about 70 percent) of figure 16*B*, the kriging standard deviation is almost constant at about 0.35 meter.

To use the kriging standard-deviation values more quantitatively, some assurance is needed that the measured data and the reduced kriging errors are approximately normally distributed and that the assumption of stationary residuals after drift removal is correct. If assumptions are valid, then the basic statistical principles involving confidence intervals can be applied. In this example, the kriging standard deviation of 0.35 meter throughout most of the map indicates that there is a 95-percent chance that the true value at a location where there is a kriging estimate is within 0.76 meter (twice the kriging standard deviation) of the kriging estimate.

As an example of evaluating network density and the accuracy of kriging estimates, two new maps were developed. To compile the first map, a decrease in network density was effected by removing nine measured locations from the northwest part of the area (fig. 14*B*) where sampling density was high and kriging standard deviations were low. Kriging estimates were produced for the same grid and the basic univariate kriging estimate statistics are listed in table 4 (water level B). The map shown in figure 16*C* indicates that the ratio of the original kriging standard deviations and the kriging standard deviations with the nine measured locations removed is always very close to 1.00, which indicates that there is very little difference between the two sets of kriging standard deviations and that water levels are oversampled in the area where the nine measured locations were removed.

To produce the third map (fig. 14*C*), nine locations were added in the southwest corner where the sampling density was relatively low and the kriging standard deviation was relatively high. In section 3.3, equation 3–47 indicates that the universal kriging variance depends on the variogram, the type of trend, and the measurement locations; in this respect, the kriging standard deviation does not depend on the values at the measurement locations. Consequently, values of zero were used for the nine new measurement locations and only the resultant map of kriging standard deviations (fig. 16*D*) is of interest. The map shows that the kriging standard deviations in the lower left corner, which formerly had values of about 0.8, approximately have been decreased by a factor of about 0.25, which indicates that the kriging estimates, based on the geometry of the network, are more reliable.

6.2 Bedrock-Elevation Examples

The following examples are for bedrock elevations. The principal purposes of the examples are to familiarize the reader with a kriging exercise using bedrock elevations and to describe block kriging. The data are from an area where bedrock consists of a series of intercalated terrestrial deposits that have been weathered somewhat and then covered with alluvium. The opportunity for measurement error in these data is inevitable because the determination of just where bedrock begins is complicated and subjective.

The set of measured locations, set A, is shown in figure 17*A*, and the basic univariate statistics are listed in table 2 (bedrock A); modifications to the measured data, such as removal of sites, are shown in figure 17*B*. The techniques described in section 5.0 were used to guide the following steps for variogram construction:

1. The raw variogram indicated a stationary spatial mean. The data were assumed to be suitable for ordinary kriging.
2. An isotropic Gaussian model was used to fit the variogram, which had a nugget of 0.65 meter squared, a sill of 12.54 meters squared, and a range of 914 meters (table 3, bedrock A).
3. Cross validation was performed, and the results (table 3, bedrock A) were not acceptable.

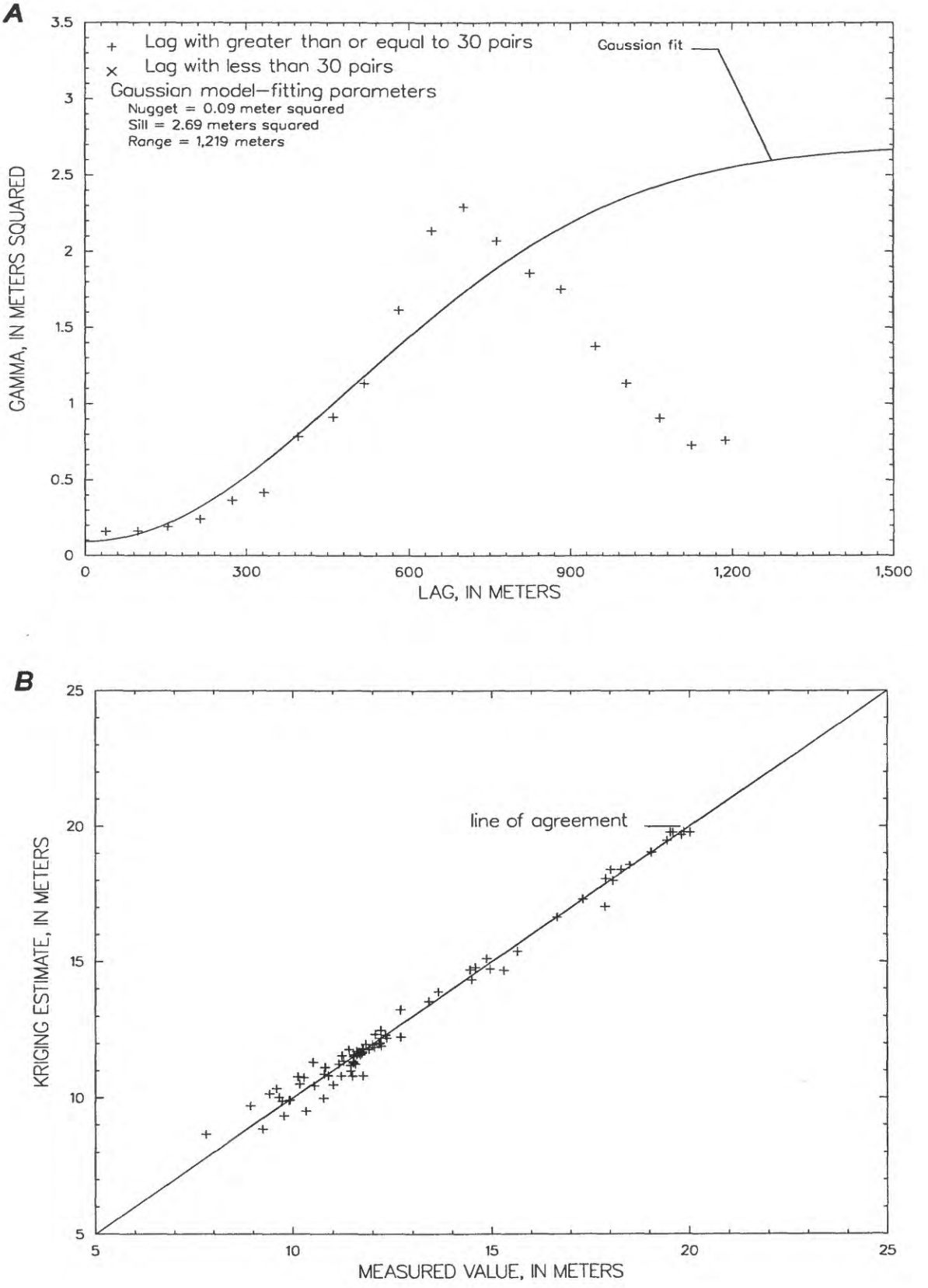


Figure 15. Variogram and variogram cross-validation plots for residuals in water-level examples—*A*, theoretical variogram; *B*, cross-validation scatterplot; and *C*, cross-validation probability plot.

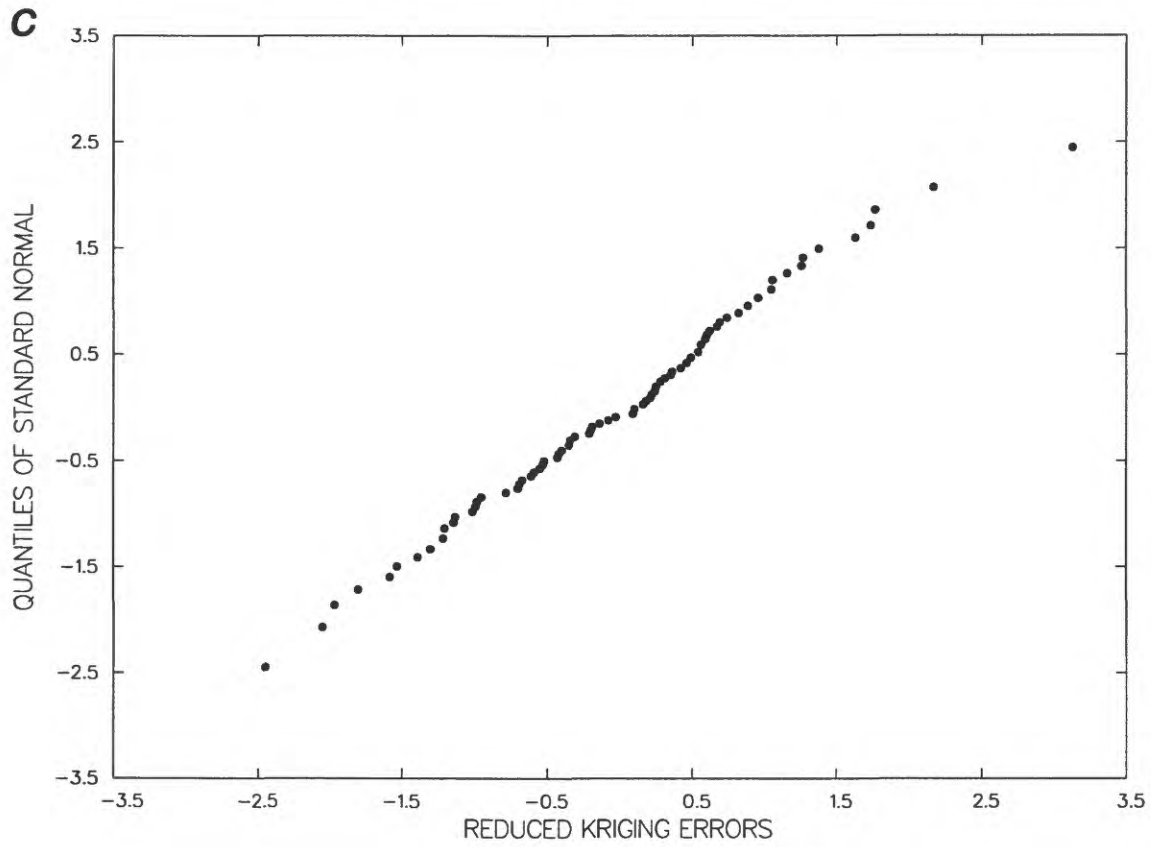


Figure 15. Variogram and variogram cross-validation plots for residuals in water-level examples—*A*, theoretical variogram; *B*, cross-validation scatterplot; and *C*, cross-validation probability plot—Continued.

Table 3. Variogram characteristics and cross-validation statistics

[Note: NA, not applicable; base unit for water levels and bedrock is meters; base unit for water quality A is log concentration, concentration in micrograms per liter; base unit for water quality B is indicator units]

Example identifier	Variogram characteristic						Cross-validation statistics		
	Transformation	Direction/ tolerance	Model	Nugget (base units squared)	Sill (base units squared)	Range (meters)	Average kriging error (base units)	Kriging root-mean-squared error (base units)	Reduced root-mean-squared error (dimensionless)
Water levels	Drift	0/NA	Gaussian	0.09	2.69	1,219	−0.0006	0.37	1.083
Bedrock A	None	0/NA	Gaussian	0.65	12.54	914	0.045	2.53	2.146
Bedrock B	None	0/NA	Gaussian	0.74	8.36	732	−0.010	1.34	1.192
Water quality A	Natural log	150/45	Exponential	1.00	3.20	1,295	0.105	1.54	0.938
Water quality A	Natural log	240/45	Exponential	1.00	3.20	228	0.105	1.54	0.938
Water quality B	Indicator	150/45	Spherical	0.05	0.25	610	NA	NA	NA
Water quality B	Indicator	240/45	Spherical	0.05	0.25	213	NA	NA	NA

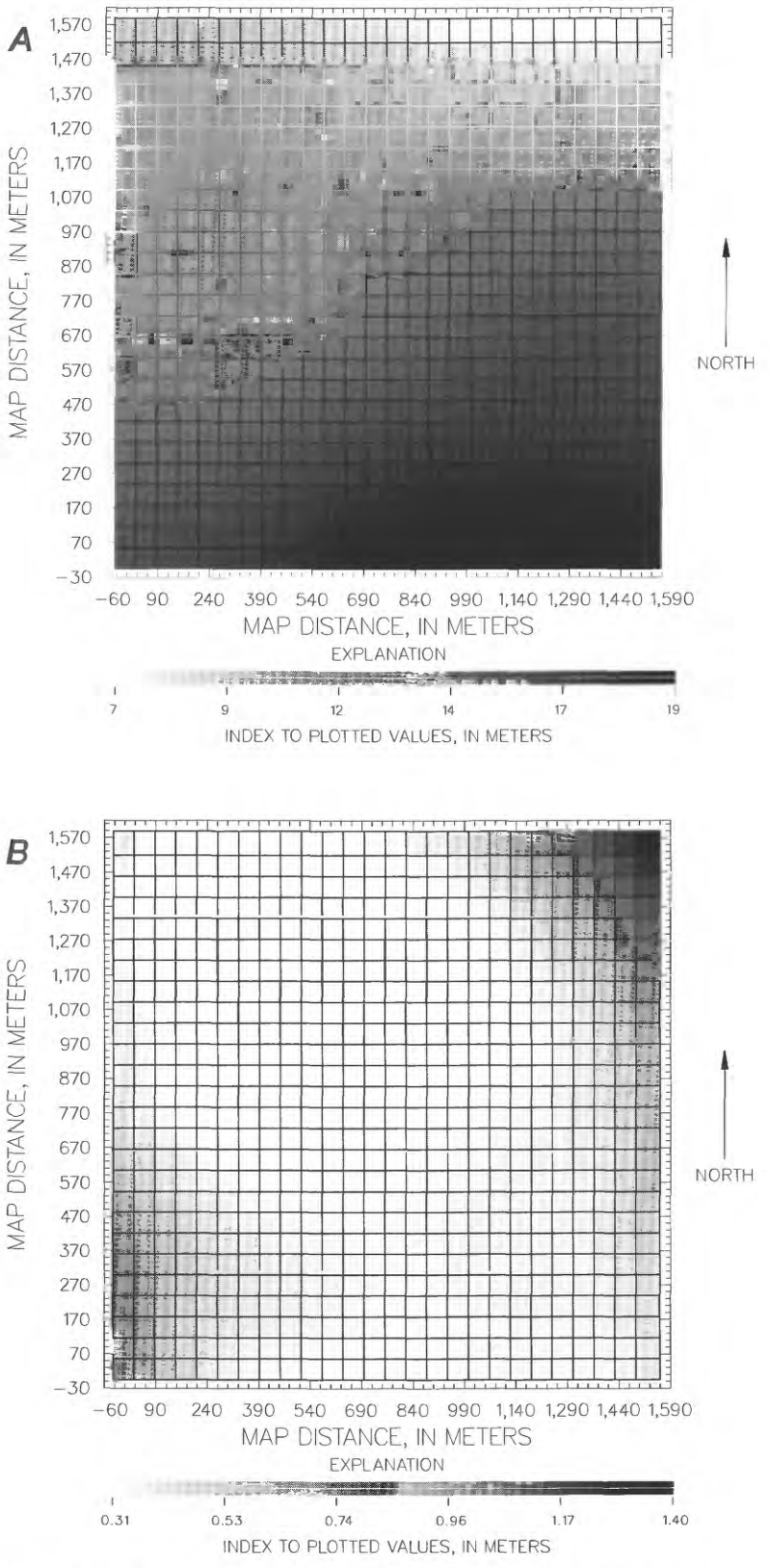
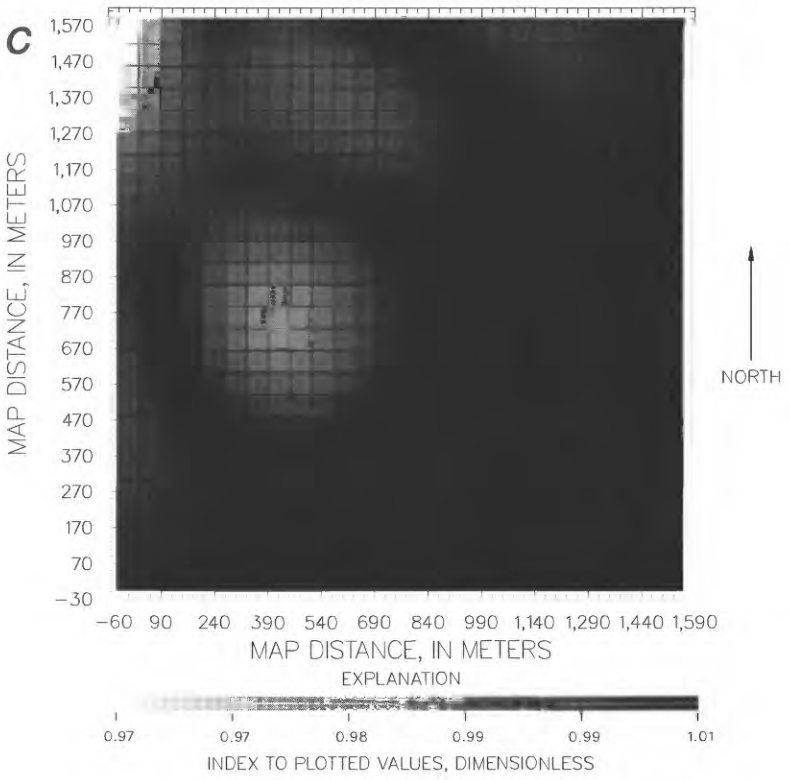


Figure 16. Kriging results for ground-water-level examples—*A*, kriging estimates for original data; *B*, kriging standard deviations for original data; *C*, ratio (original data to original with dropped sites) of kriging standard deviations; and *D*, kriging standard deviations for original data with added sites.



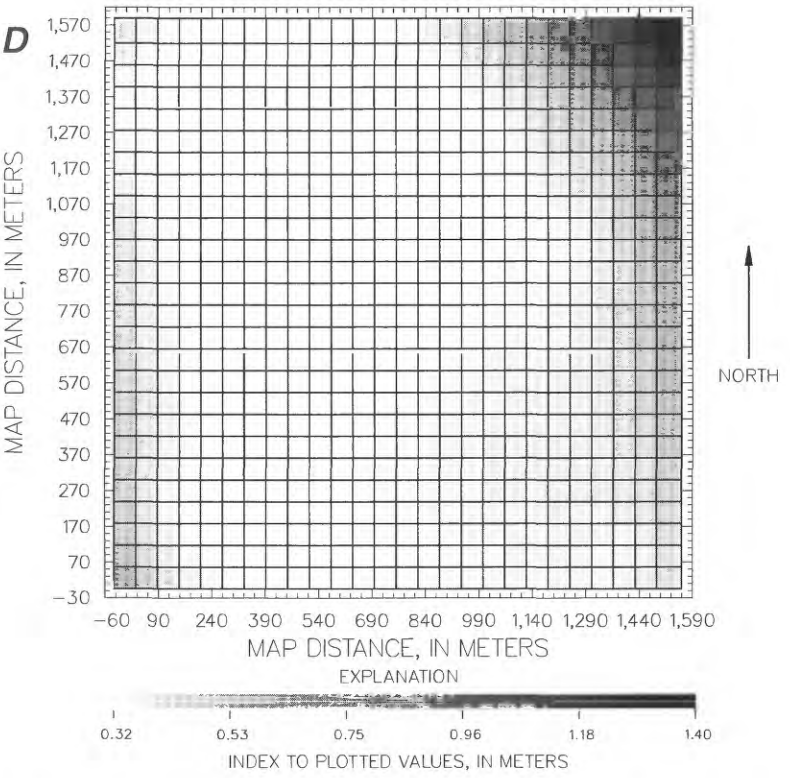


Figure 16. Kriging results for ground-water-level examples—*A*, kriging estimates for original data; *B*, kriging standard deviations for original data; *C*, ratio (original data to original with dropped sites) of kriging standard deviations; and *D*, kriging standard deviations for original data with added sites—Continued.

Table 4. Univariate statistics for gridded kriging estimates in example applications

[Note: Base unit for water level A and B and bedrock B and C is meters; base unit for water quality A is log concentration, concentration in micrograms per liter]

Example identifier	Transformation	Minimum (base units)	Maximum (base units)	Mean (base units)	Median (base units)	Standard deviation (base units)	Skewness (dimension-less)
Water level A	Drift	7.42	19.8	14.0	13.6	3.09	0.11
Water level B	Drift	7.49	19.8	14.0	13.5	3.09	0.11
Bedrock B	None	7.96	19.8	12.6	12.1	2.35	0.82
Bedrock C	None	8.14	19.8	12.6	12.1	2.33	0.82
Water quality A	Natural log	2.92	7.07	5.17	5.03	0.72	−0.06

The cross-validation exercise produced a reduced root-mean-squared error of 2.146 (table 3, bedrock A), which indicates that the kriging variance is underestimated. Further attempts to fit the Gaussian model to the sample variogram points produced better cross-validation statistics; however, the Gaussian curve began to depart substantially from the sample variogram points at the low lag sample points. As a result, the distribution of the residuals was examined, and the eastern, and especially northeastern, parts of the area were determined to contain problematic data that rendered the distribution nonhomogeneous. The nonhomogeneous nature was related to an incised channel present on the bedrock surface. Therefore, the measured data were restricted to exclude the outlying measurements. Before the restriction, two alternative techniques for dealing with the outlying measurements were considered and deemed beyond the scope of this effort. However, a brief discussion of the alternatives is appropriate.

The first alternative was to fit a contrived and nongradual surface to the measured data to remove the outlier effect. A splined surface might be capable of producing the desired result. The decision whether or not to pursue such an alternative becomes somewhat philosophical. In a relatively simple example, as in this bedrock example, such an alternative may be entirely appropriate; however, this alternative may actually involve two unique and homogeneous domains. Therefore, the second alternative, distributing the kriging process so that each homogeneous domain is addressed independently, becomes more attractive. In more complicated applications where a large number of domains are present, a distributed approach may be necessary to avoid an undue amount of compromise.

The restriction of measured data, set B, is shown in figure 17*B*, and the basic univariate statistics are listed in table 2 (bedrock B). The restriction exercise resulted in removing 18 measured locations and in the truncation of the northeastern part of the area so that the area became polygonal rather than rectangular. The techniques described in section 5.0 were used to guide the following steps for variogram construction:

1. A Gaussian model was used to fit the variogram, which had a nugget of 0.65 meter squared, a sill of 8.36 meters squared, and a range of 732 meters. The variogram indicated a stationary spatial mean.
2. Initial cross validation was performed, and the nugget was changed from 0.65 meter squared to 0.74 meter squared to improve cross-validation statistics. The final variogram is shown in figure 18*A*, and the characteristics are listed in table 3.
3. Final cross validation was performed, and the results, shown in figures 18*B* and *C* and listed in table 3 (bedrock B), were acceptable.

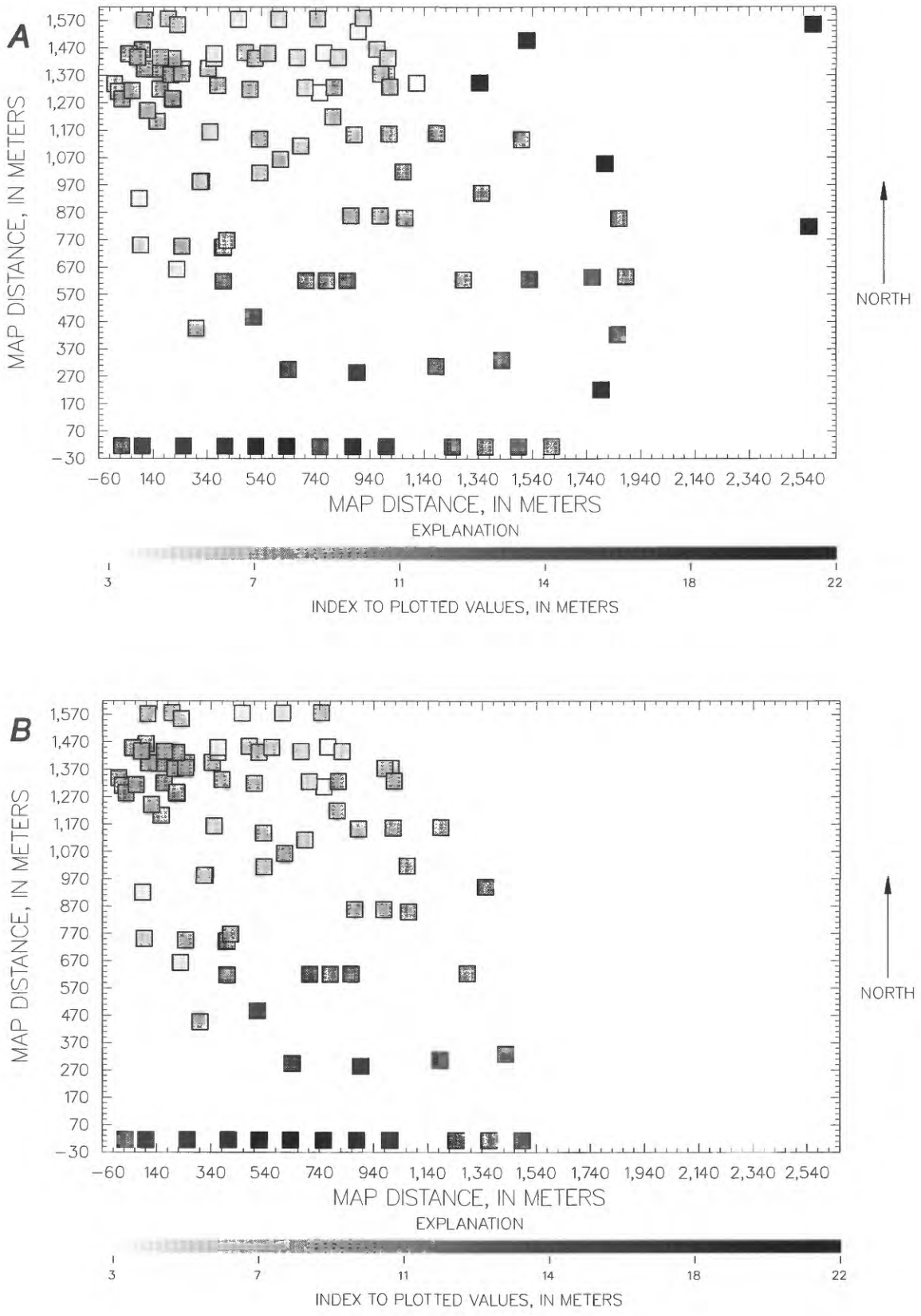


Figure 17. Location of measured data for bedrock-elevation examples—*A*, original data and *B*, restricted data.

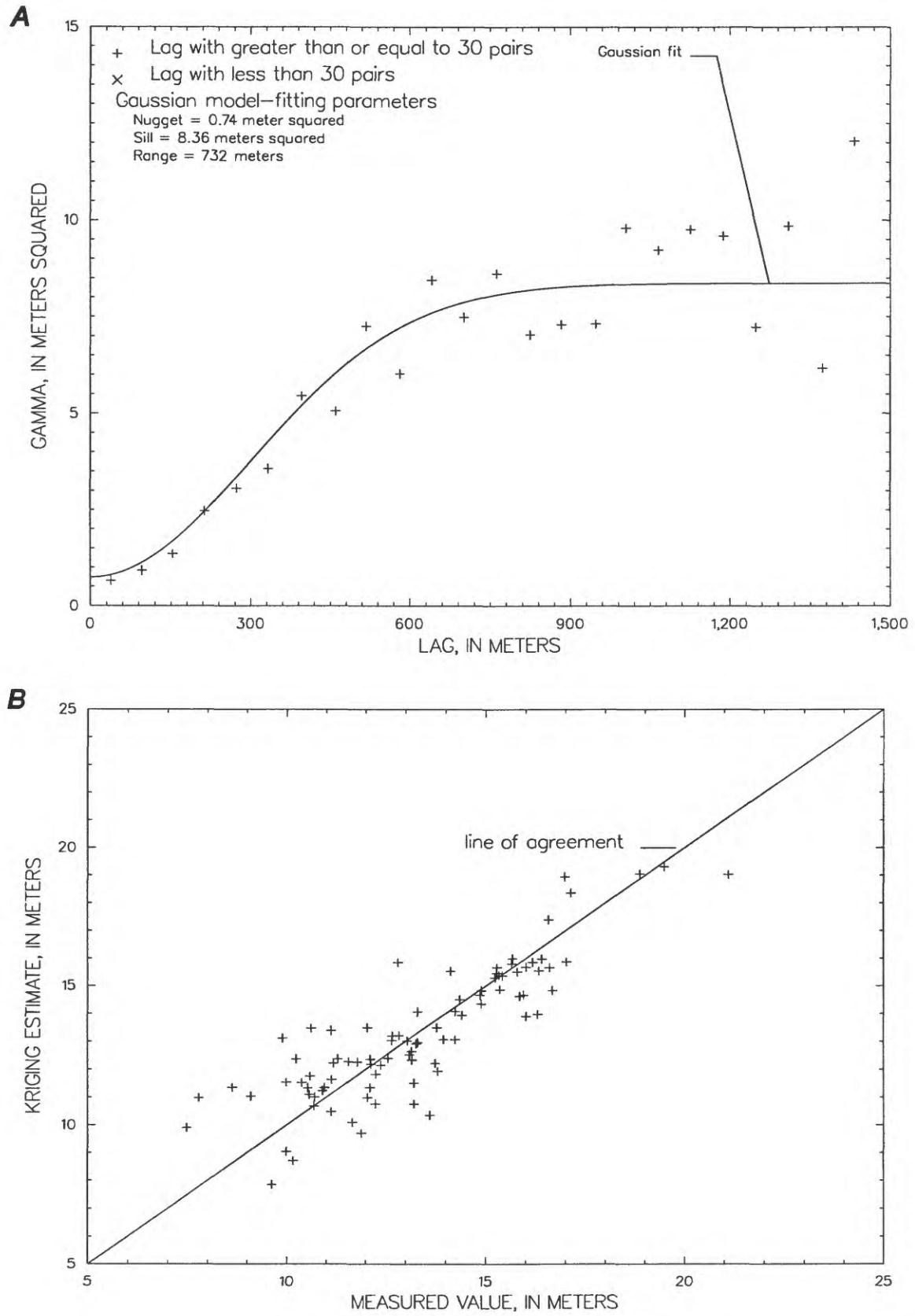


Figure 18. Variogram and variogram cross-validation plots for bedrock-elevation examples—*A*, theoretical variogram; *B*, cross-validation scatterplot; and *C*, cross-validation probability plot.

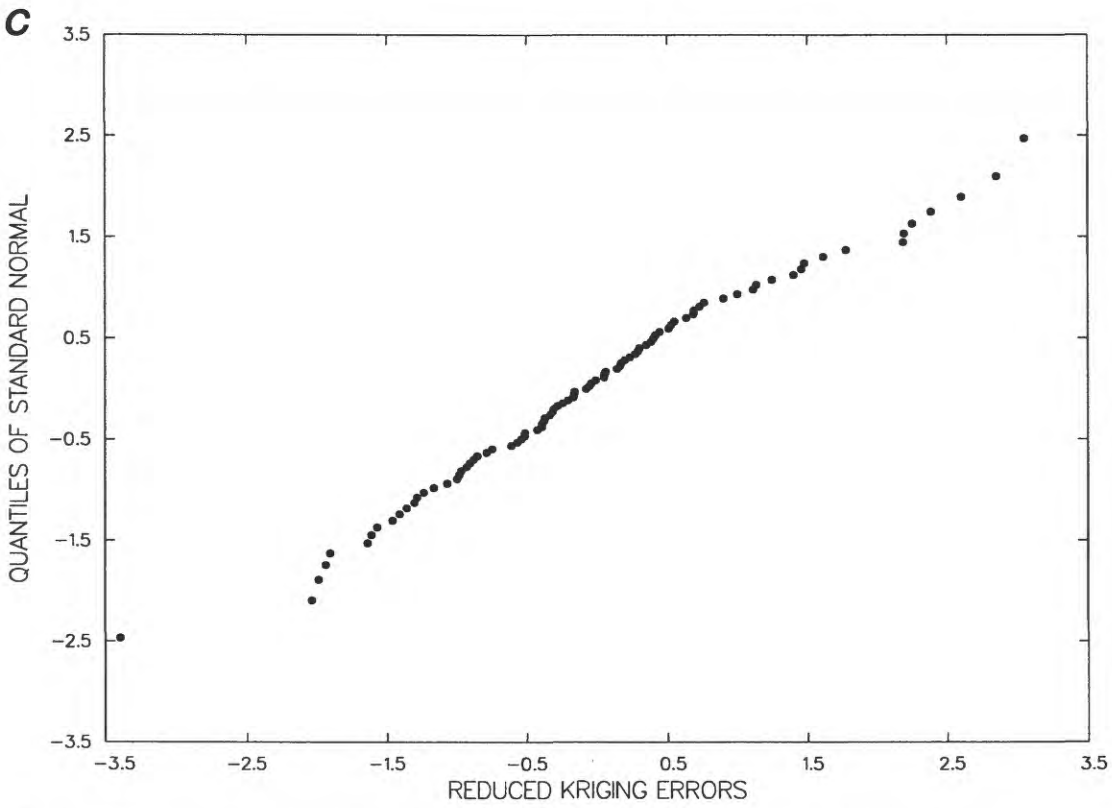


Figure 18. Variogram and variogram cross-validation plots for bedrock-elevation examples—*A*, theoretical variogram; *B*, cross-validation scatterplot; and *C*, cross-validation probability plot—Continued.

The large difference between the sill defined for the initial data set and the sill for the restricted data set [12.54 meters squared and 8.36 meters squared (table 3)] supports the hypothesis that the original data set is actually two different domains. The final variogram then was used, along with the measured data, to produce ordinary kriging estimates for all points in a 52-by-52 grid that had a spacing of about 30-by-30 meters, which was truncated along the northeastern border because of the restriction operation. For the kriging procedure, a search radius of about 914 meters, with a maximum of 16 and a minimum of 8 surrounding locations, was specified. It is not uncommon to specify a search radius that is greater than the variogram range; this practice helps ensure that, in this case, between 8 and 16 points would be obtained to develop the kriging estimate. Gray-scale maps of the kriging estimates and kriging standard deviations are shown in figures 19*A* and *B*, respectively, and the univariate kriging estimate statistics are listed in table 4 (bedrock B). The kriging results indicate channel-like features in the bedrock surface and a prominent bedrock high at the south border of the area; the results are a good representation of the results from other more elaborate studies.

For an example of block kriging, an investigative goal of establishing block values of bedrock elevation for a finite-difference ground-water-model grid having about 120-by-120-meter cells was assumed. The same variogram and search criteria were used to estimate block values for a 13-by-13 grid that had about 120-by-120-meter spacing; a 4-by-4 block was specified. Each kriging value shown in figure 19*C* is an estimate of the average value of bedrock elevation throughout the about 120-by-120-meter block. The standard deviation for the block estimates is less than the standard deviation for the point estimates (table 4). Gray-scale maps of the kriging estimates and of the kriging standard deviations are shown in figures 19*C* and *D*, and the univariate kriging estimate statistics are listed in table 4 (bedrock C).

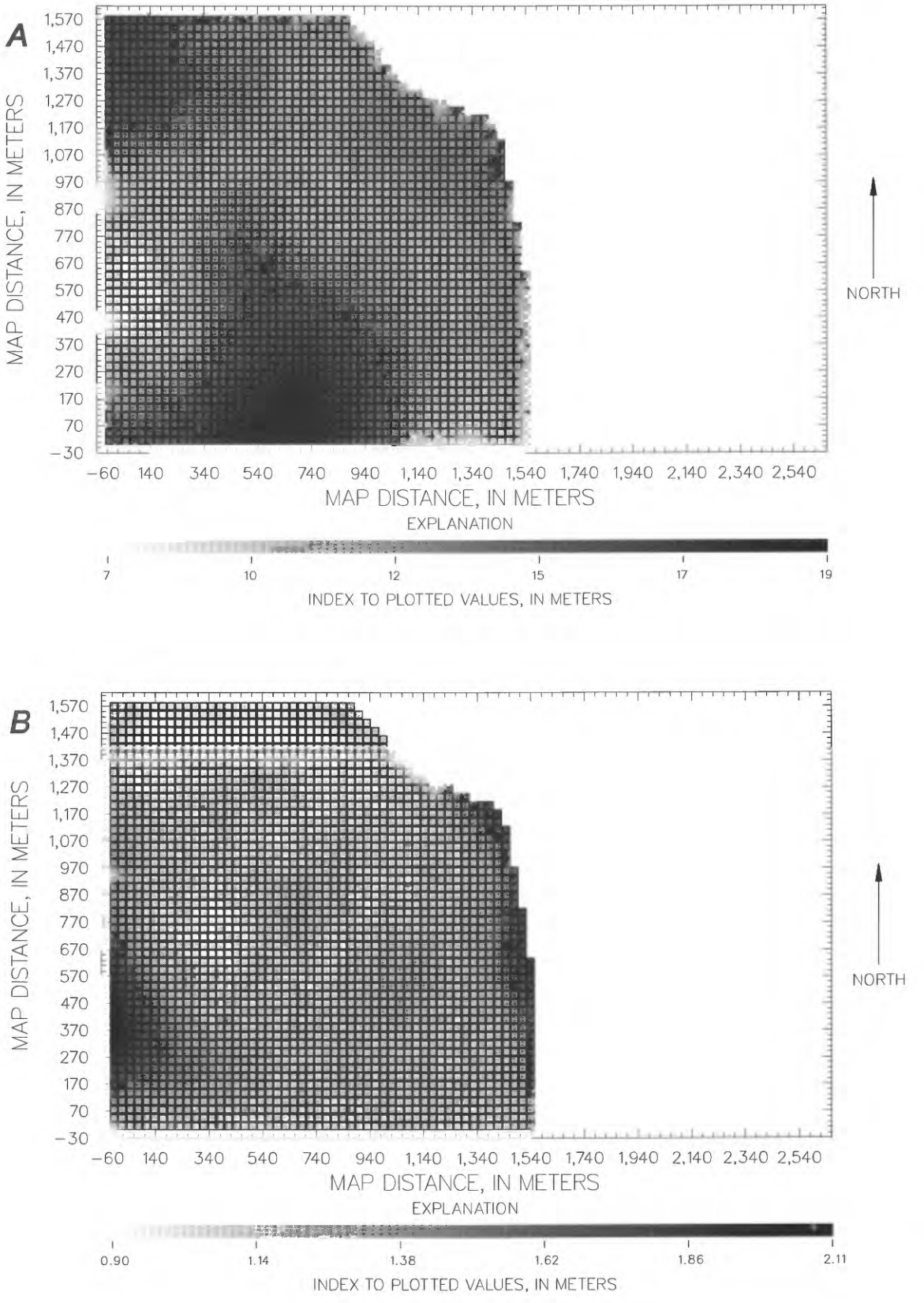
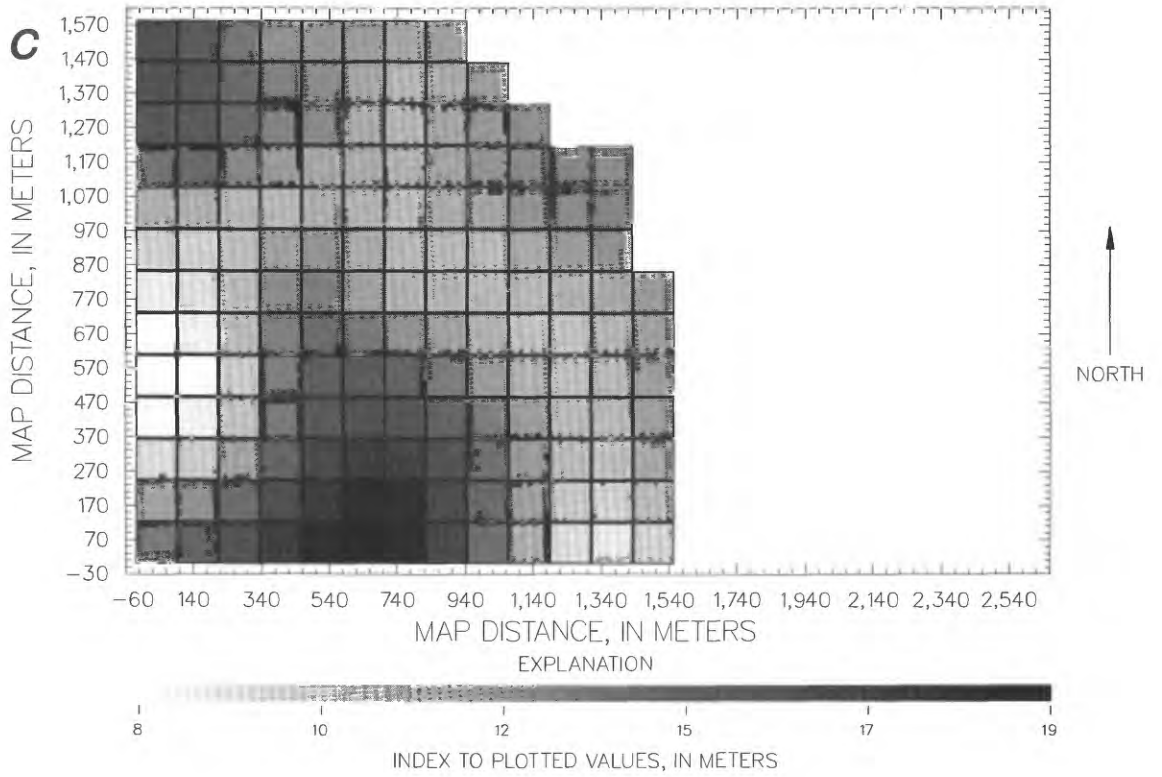


Figure 19. Kriging results for bedrock-elevation examples—*A*, kriging estimates; *B*, kriging standard deviations; *C*, block kriging results; and *D*, block kriging standard deviations.



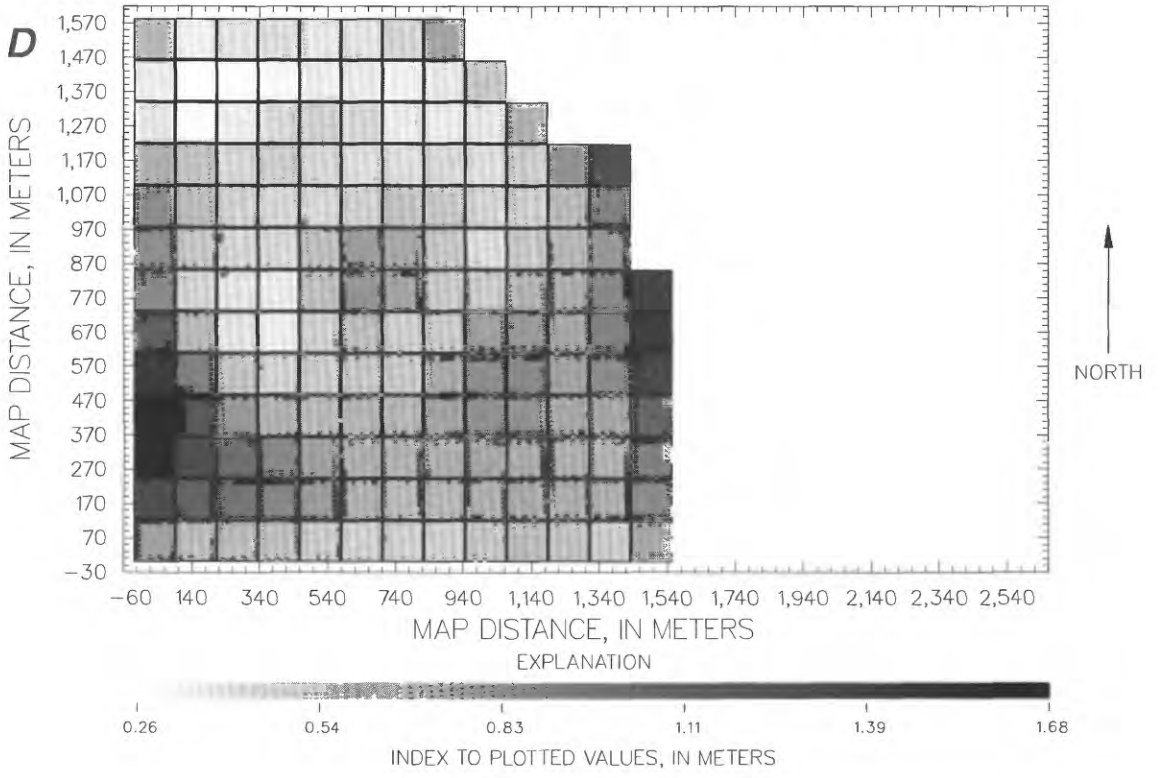


Figure 19. Kriging results for bedrock-elevation examples—*A*, kriging estimates; *B*, kriging standard deviations; *C*, block kriging results; and *D*, block kriging standard deviations—Continued.

6.3 Ground-Water-Quality Examples

The following examples are for ground-water-quality information consisting of concentrations determined for a contaminant. The principal purposes of the examples are to familiarize the reader with a kriging exercise using ground-water-quality information and to illustrate indicator kriging. The examples also are to familiarize the reader with data that are strongly anisotropic and need transformation. The data are from a water-table aquifer developed in alluvial sediments where the depth to water was less than about 23 meters. Several analytical laboratories were involved in measuring the concentration of the contaminant in the water-quality examples. Each of the analytical laboratories had to follow rather comprehensive guidelines that specified tests of instrument performance before sample determinations were made, as well as measurement of extraction efficiencies. Because of these performance guidelines, the opportunity for errors due to instrument error was considered to be either known or relatively low. In addition to using performance guidelines, field quality-assurance samples also were collected. These samples can be used to evaluate other possible errors, such as cross contamination and representativeness of the sample. Duplicate samples for the contaminant in the water-quality examples indicate as much as about 15-percent variability in reported results. This variability is not entirely unusual and is most likely related to the integrity of the analytical method or the method for aggregating the sample media during sample collection.

The set of measured locations is shown in figure 20 and the basic univariate statistics are listed in table 2 (water quality A). An initial review of the data indicated three important features:

1. The data seemed to have strong anisotropy at about 150 counterclockwise degrees to the east-west base line.
2. The data required a natural logarithmic (log) transformation so the distribution was approximated by a normal distribution.

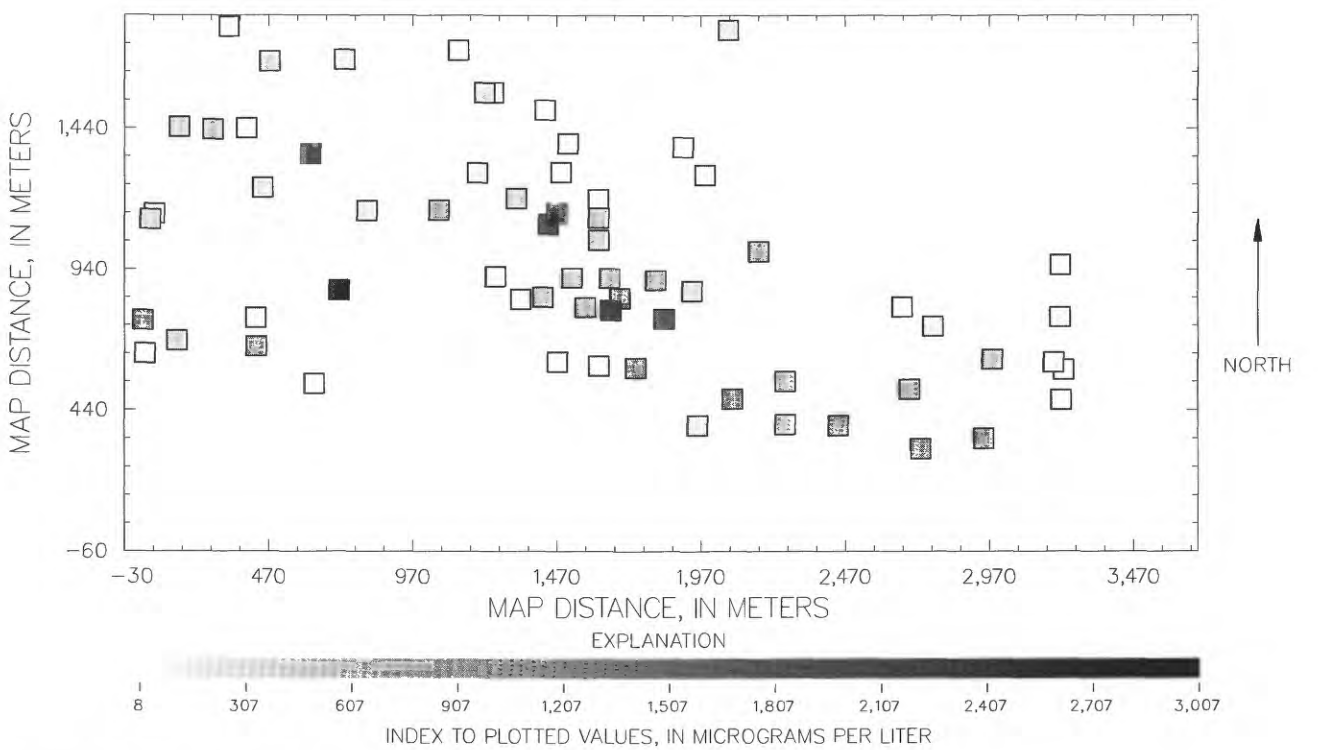


Figure 20. Location of measured data for ground-water-quality examples.

3. No trends were indicated during preliminary exploration, and ordinary kriging was tentatively selected as the appropriate technique.

Natural log transformations are routinely needed for concentration data that vary over several orders of magnitude, which is common in areas of contaminant plumes. The data were transformed to log space and fit acceptable criteria for normality. After transformation to log space, the techniques described in section 5.0 were used to guide the following steps for variogram construction:

1. An exponential model was used to fit a directional variogram at an angle of 150 counterclockwise degrees to the east-west base line. The variogram had a nugget of 1.00 log concentration squared, a sill of 3.20 log concentration squared, and a range of 1,295 meters [fig. 21*A* and table 3 (water quality A)].
2. An exponential model also was fit to a directional variogram at an angle of 240 counterclockwise degrees to the east-west base line. The variogram had a nugget of 1.00 log concentration squared, a sill of 3.20 log concentration squared, and a range of 228 meters [fig. 21*B* and table 3 (water quality A)].
3. Cross validation was performed using the geometric anisotropy of the two variograms and the results [figs. 21*C* and *D*, and table 3 (water quality A)] were acceptable.

The residuals are symmetrically distributed (fig. 21*D*). However, the scatterplot (fig. 21*C*) indicates that small concentrations were overestimated and that large concentrations were underestimated. This discrepancy in the estimates does not indicate an error in the model, but rather, indicates a consequence of data that have a large nugget compared to the sill; in this example, the nugget is approximately 30 percent of the sill. The large nugget decreases the predictive capacity of the model and increases the smoothing introduced by kriging.

The established variogram then was used, along with the measured locations, to produce ordinary kriging estimates for all points in a 40-by-20 grid using a grid spacing of about 91-by-91 meters. For the kriging procedure, a search radius of about 1,524 meters with maximum of 16 and a minimum of 8 locations was specified. Gray-scale maps of kriging estimates, in back-transformed and log-space concentrations, as well as the kriging standard deviations in log space, are shown in figures 22*A*, *B*, and *C*.

The back-transformation procedure was a simple exponentiation of the log-space kriging estimates. Such a back-transformation does not use bias-correction factors to deal with moment bias; consequently, the back-transformed values need to be interpreted as median values rather than average, or mean, values. A simple back-transformation, however, is convenient and was performed, principally, to enhance visual interpretation of the kriging estimates. Univariate statistics for the log-space kriging estimates are listed in table 4 (water quality A). The kriging results do have noticeable smoothing; however, they also indicate a plume emanating from a location just northwest of the center of the area and indicate movement and some dispersion to the southeast; the estimates are a very good representation of the results from other more elaborate studies.

Additionally, to indicate the effect of the log transform on probabilities in converting, or back-transforming, kriging estimates, the kriging estimates and the kriging standard deviations, in log space, were used to estimate the one-sided 95th percentile at each kriging-estimate location according to the formula:

$$\hat{C}_{0.95} = \exp[\hat{Z}(\underline{x}_o) + 1.645\sigma_K(\underline{x}_o)] \qquad (6\text{–}1)$$

where

$\hat{Z}(\underline{x}_o)$ is the kriging estimate at location, $\underline{x}_o$, in log space; and

$\sigma_K(\underline{x}_o)$ is the corresponding kriging standard deviation in log space.

The resulting map is shown in figure 22*D* and can be used to indicate areas where the true concentration has only a 5-percent chance of exceeding the value indicated.

To perform indicator kriging, the indicator transformation, as described in section 3.0, was applied. An indicator cutoff equal to the median value of 270 micrograms per liter for the untransformed measured data was selected. The model for indicator kriging estimates the probability that the concentration would be less than the indicator cutoff.

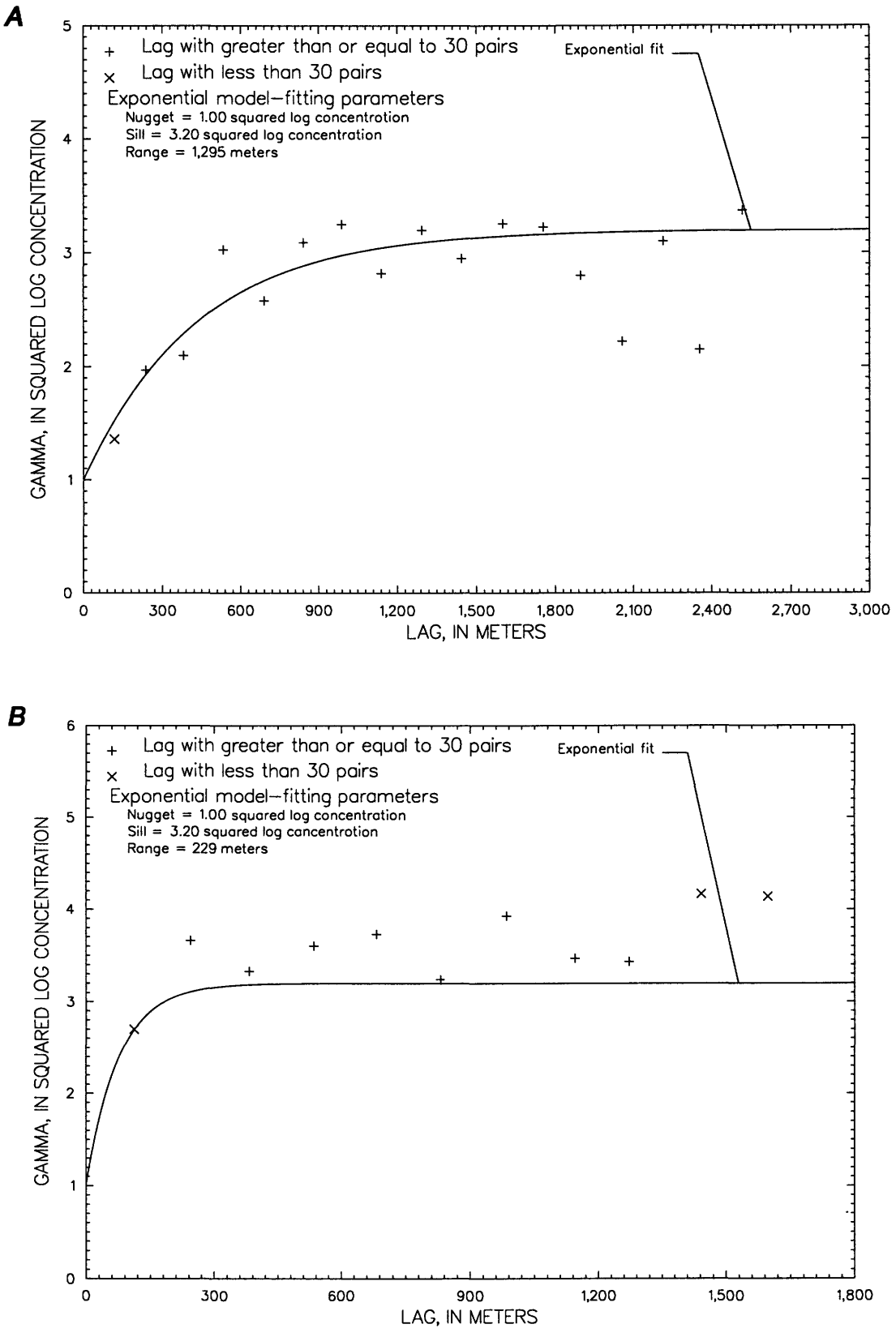
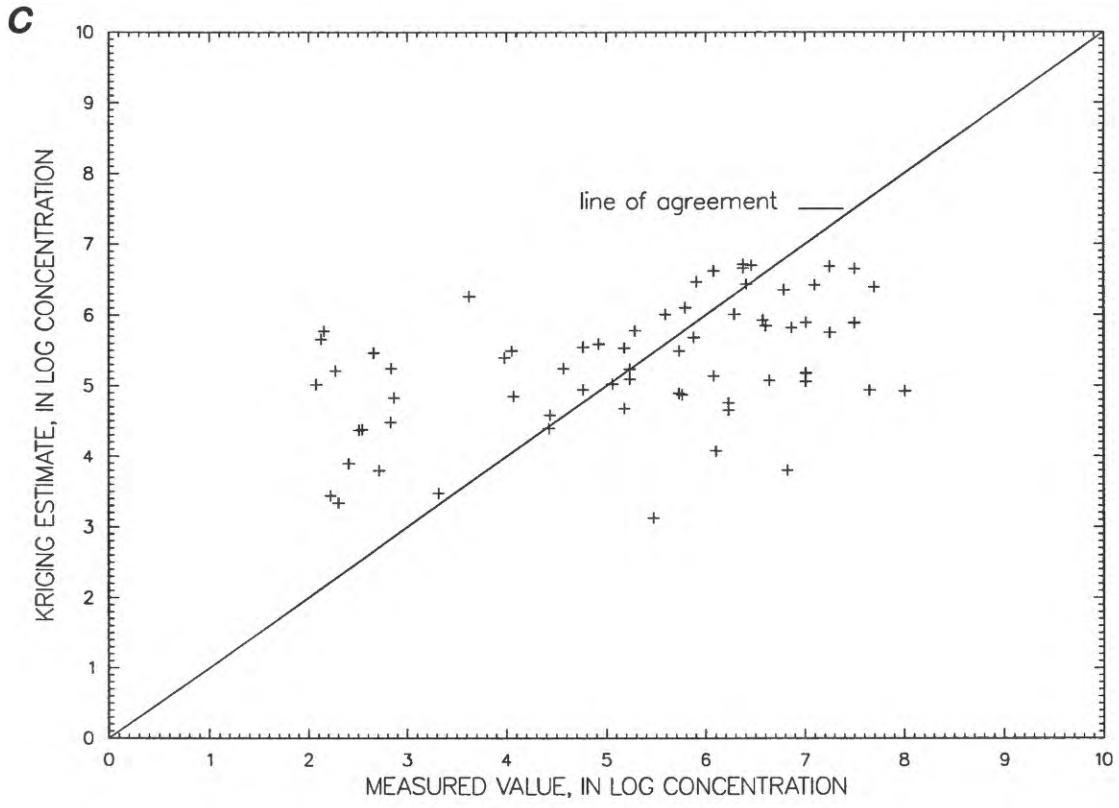


Figure 21. Directional variograms and variogram cross-validation plots for ground-water-quality examples—*A*, theoretical major-direction variogram (southeast); *B*, theoretical minor-direction variogram (northeast); *C*, cross-validation scatterplot; and *D*, cross-validation probability plot.



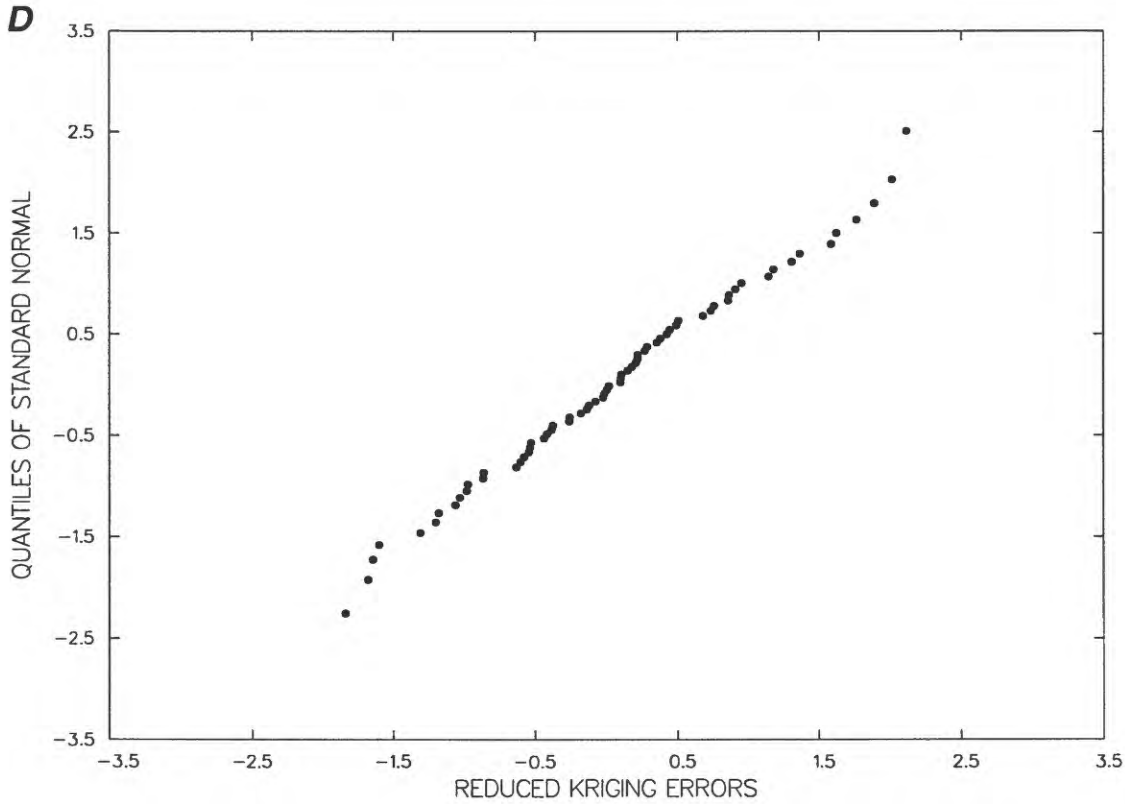


Figure 21. Directional variograms and variogram cross-validation plots for ground-water-quality examples—*A*, theoretical major-direction variogram (southeast); *B*, theoretical minor-direction variogram (northeast); *C*, cross-validation scatterplot; and *D*, cross-validation probability plot—Continued.

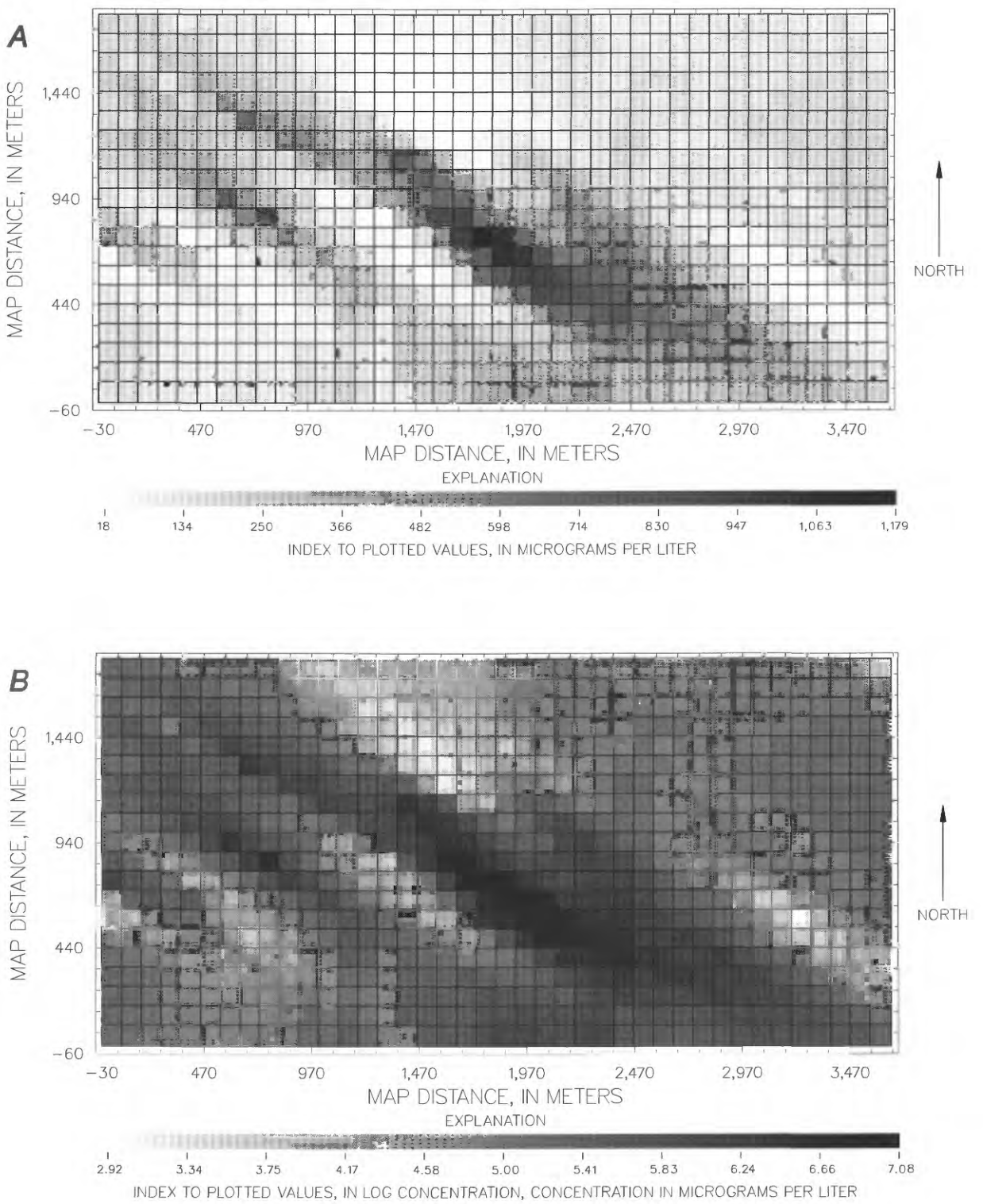
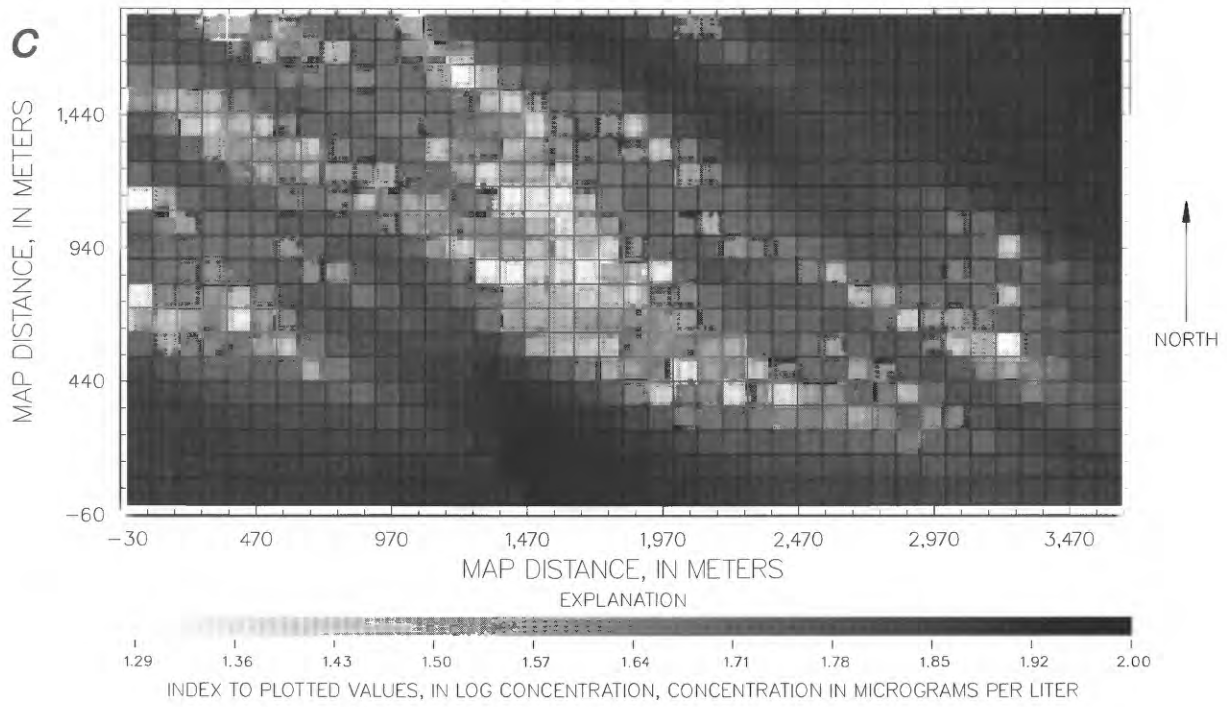


Figure 22. Kriging results for ground-water-quality examples—*A*, kriging estimates back-transformed; *B*, kriging estimates in log space; *C*, kriging standard deviations in log space; and *D*, 95-percent confidence level for kriging estimates back-transformed.



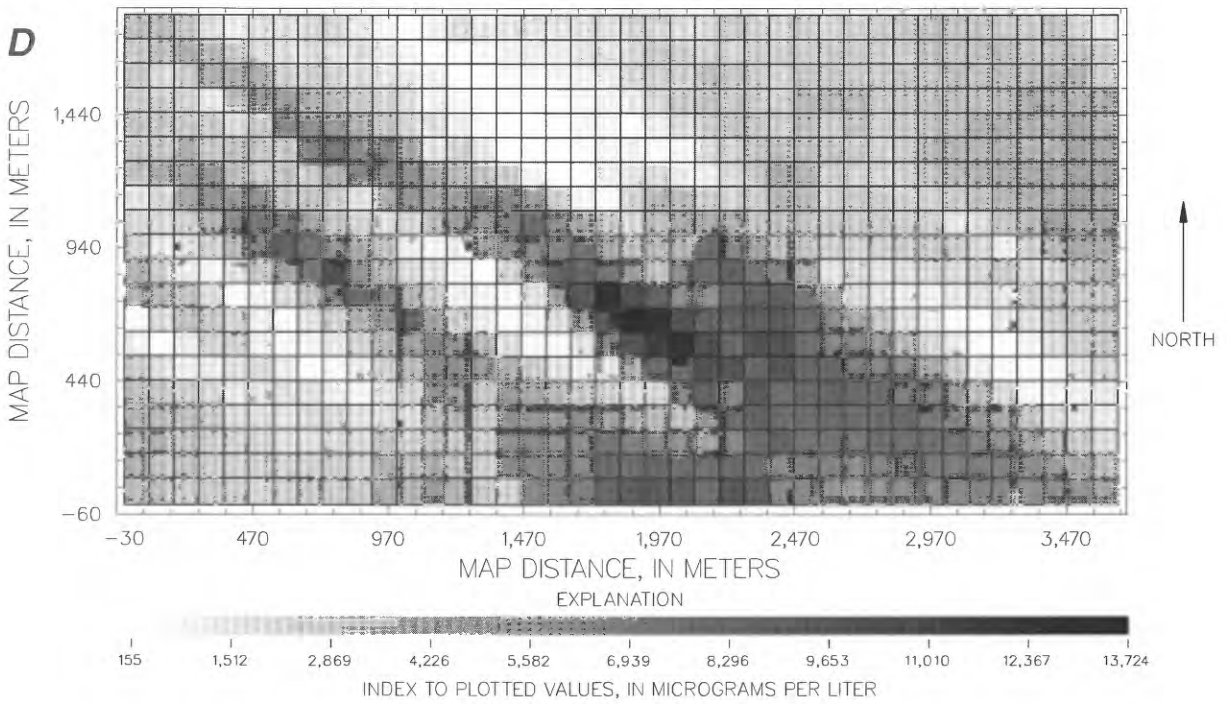


Figure 22. Kriging results for ground-water-quality examples—*A*, kriging estimates back-transformed; *B*, kriging estimates in log space; *C*, kriging standard deviations in log space; and *D*, 95-percent confidence level for kriging estimates back-transformed—Continued.

The techniques described in section 5.0 were used to guide the following steps in variogram construction:

1. No trends were indicated during preliminary exploration, and ordinary kriging was tentatively selected as the appropriate technique.
2. A spherical model was used to fit an anisotropic variogram at an angle of 150 counterclockwise degrees to the east-west base line. The variogram had a nugget of 0.05 indicator units squared, a sill of 0.25 indicator units squared, and a range of 610 meters [fig. 23*A* and table 3 (water quality B)].
3. A spherical model also was fit to an anisotropic variogram at an angle of 240 counterclockwise degrees to the east-west base line. The variogram had a nugget of 0.05 indicator units squared, a sill of 0.25 indicator units squared, and a range of 213 meters [fig. 23*B* and table 3 (water quality B].

The established variogram, and the indicator transform of the measured data were used to produce ordinary kriging estimates for the same grid and search criteria as the first ground-water-quality example. A gray-scale map of the kriging estimates is shown in figure 24. The kriging indicator map provides a gridded estimate for the probability of contaminant values being less than the indicator cutoff, which is a concentration of 270 micrograms per liter in this example.

The cutoff value selected for the preceding indicator kriging example is probably higher than many investigators involved in HTRW-site investigations would like to use. The number of measurements [66 in table 2 (water quality A)] used in this example is probably a high number of measurements for typical HTRW-site investigations; yet, even with this high number of measurements, it was not possible to construct a variogram for indicator values much lower than the median. An alternative to this problem would be to assume the log-transformed kriging model developed in the first water-quality example is correct and to rely on the kriging estimates from that model to determine areas greater than or lesser than some indicator value. The same estimates also could be used to compute the probability that the concentration was less than some arbitrarily selected value.

7.0 REVIEW OF KRIGING APPLICATIONS

This section presents a brief discussion of three principal topics—applicability of kriging techniques, important elements that need to be addressed in kriging applications, and errors in measured data. Much of the information presented in this section has been gathered from other sections of this report and is presented collectively here. The items identified as important to kriging applications may be helpful in assessing kriging applications that are under review.

7.1 Applicability of Kriging

In the preceding sections of this report, the theory of kriging techniques has been summarized, and examples have been given to indicate the use of kriging techniques in HTRW-site investigations. The examples presented were selected so that kriging would provide satisfactory results or be applicable. Additionally, the examples were designed so that, for the purposes of demonstration, some sort of adjustment of the data was needed; that is, drift was removed or transformations were made.

Investigators are very likely to have data for which, in a strict sense, kriging may be applicable, but results may be unsatisfactory. Much of the fundamental information that might be used to establish how satisfactory the application of kriging techniques may be has been presented in the preceding sections. In particular, section 5.0 includes a detailed discussion on variogram construction, which is the preliminary step in any kriging application, and systematically describes many decisions in variogram construction that need attention. If a variogram that has structure, or some identifiable dependence on lag, cannot be obtained from the data or be obtained from other means such as institutional knowledge, the results of a kriging application may not be satisfactory. Some additional discussion that is designed to aid in evaluating the amount of data that may be required for kriging applications is presented in this section. This discussion assumes that the measured data are correct; a separate and brief discussion of measurement errors also is presented in this section.

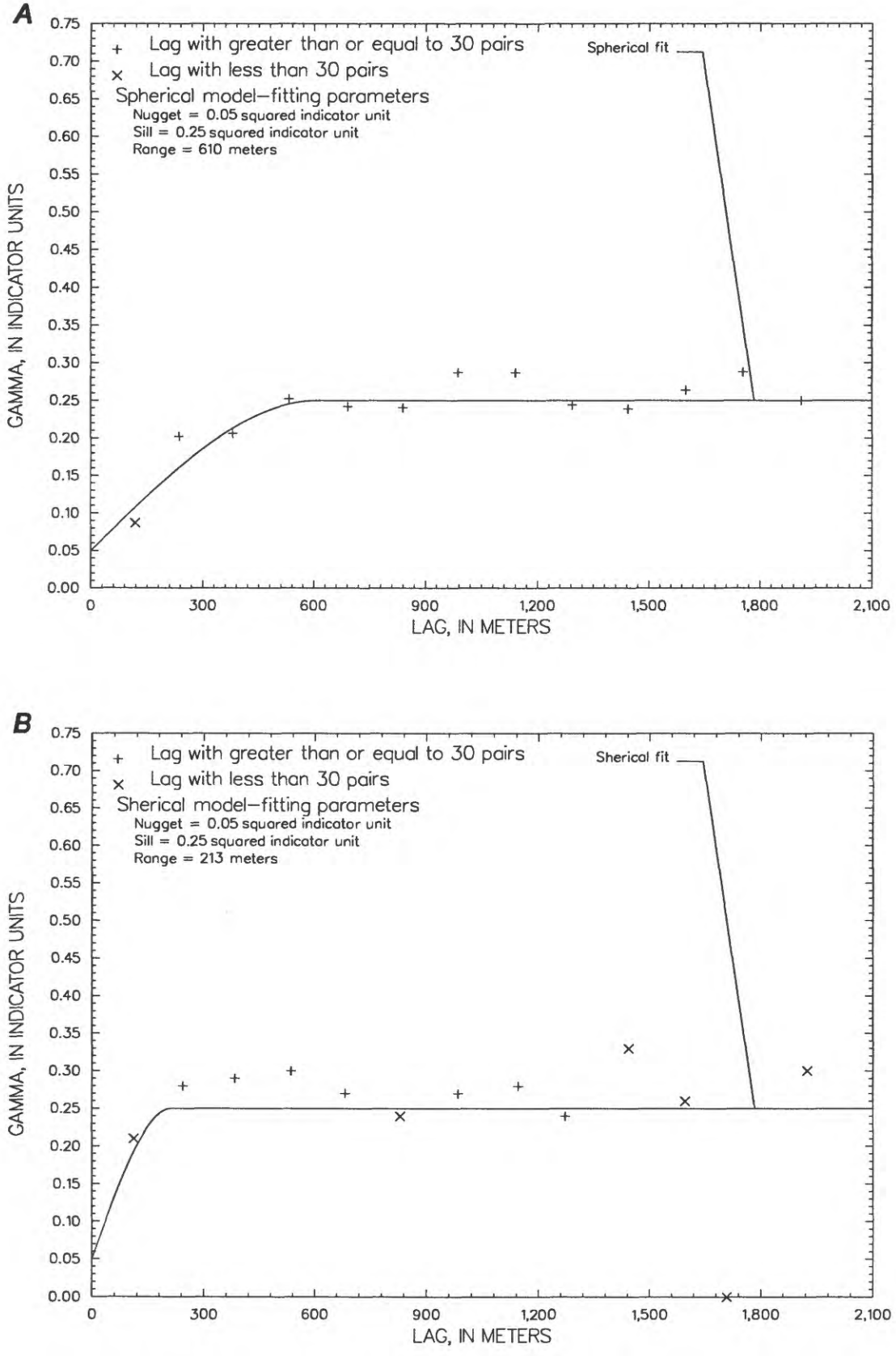


Figure 23. Directional variogram plots for indicator kriging ground-water-quality example—*A*, theoretical major-direction variogram and *B*, theoretical minor-direction variogram.

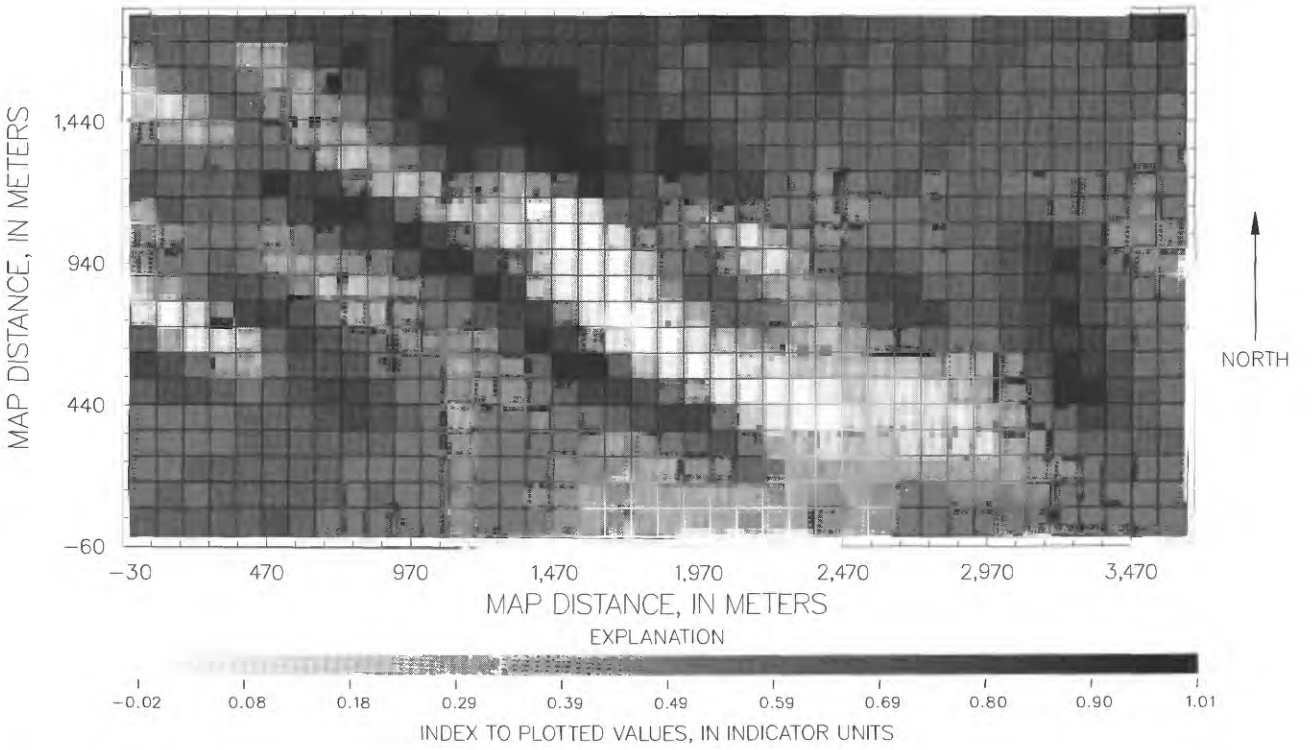


Figure 24. Indicator kriging results for ground-water-quality example.

Initially, many investigators have a tendency to focus on the amount of measured data that is available as an initial consideration; however, the applicability of kriging techniques cannot be based simply on the amount of measured data. Unless the investigator is presented with a reliable variogram, the amount and spatial distribution of measured data can be a constraint. If, for instance, there are fewer than 25 measured values at optimal locations from the field, there may not be enough data to confidently estimate Gaussian variogram parameters; however, a small amount of measured data may be suitable for other variogram models.

How much data are needed to apply kriging techniques is not easy to determine, but information in this report, especially in section 5.0, and the literature cited can provide some guidance. Section 5.3 points out that a good minimum for the number of pairs of locations in each variogram lag is 30, and the American Society for Testing and Materials (1996) has suggested that 20 may work well also. Most investigators would probably feel comfortable defining a Gaussian form, which has more inflection and, consequently, is more difficult to fit compared to the other standard variogram models, with 8 to 10 optimally located sample variogram points (enough points to define the nugget, two areas of curvature, and the sill). In this ideal case, about 25 measured values would be needed to fulfill the conservative minimum of 30 pairs per lag. In this case, the relatively few measured data points need to be systematically located so that the optimally located variogram points can be computed. If the measured data were not located systematically, as is usually the case, then more measured data would be needed.

Once sample variogram points meeting the required number of pairs can be defined, the resultant variogram still must have structure. The variogram, for instance, may simply exhibit noise about a horizontal line, a case that has no structure. If measured data are clustered and the number of lags has been minimized to meet the required number of pairs of locations, the variogram may seem horizontal because it is dominated by small-scale effects in the clustered data. The investigator then has latitude to adjust the lags and attempt to balance the lag spacing and required

number of pairs per lag interval, as described in section 5.3. However, the variogram also could seem horizontal because the actual sill is reached within a very small lag. If that lag is smaller than the minimum spacing of measured data, obtaining structure in the variogram would not be possible. In such cases, the measured data need to be considered independent, and kriging techniques, at the lag of the measured data, would be ineffective, or at least, offer little advantage over other interpolation techniques.

7.2 Important Elements of Kriging Applications

Many important elements of kriging applications have been discussed in this report. These discussions have been presented as a systematic and sequential method designed to provide guidance in kriging applications. Occasionally, an investigator is presented with the results of a previous kriging application and needs to evaluate the application before deciding whether or not to use the results. This section presents a brief review of some important elements of kriging applications that can be used in that evaluation. For a more detailed discussion of important elements of geostatistical applications, the reader is referred to the American Society for Testing and Materials (1994) for content of geostatistical investigations.

The presence of or lack of stationarity in the spatial mean needs to be demonstrated definitively. If the spatial mean is nonstationary, then drift is indicated and appropriate measures to address the nonstationarity, which are similar to the measures presented in section 5.2, need to be part of the application. In ideal situations, nonstationarity occurs as a gradual change. HTRW-site investigations may present cases, especially when using with groundwater-quality data in and around plumes, that have abrupt step-like changes at plume boundaries and do not appear as regional drift. In these cases, the investigator needs to be aware that, without knowledge of the plume boundaries, points from within the plume will be grouped with points from outside the plume in computing the sample variogram. The effect of this problem is minimized as long as the investigator can define lags that allow data points within the plume to be grouped together.

The construction of the variogram needs to be described and included as part of the kriging application documentation. The description needs to address the number of pairs of locations in each variogram lag and to demonstrate that the variogram has structure. A plot of the variogram is helpful to demonstrate the presence or absence of structure. The variogram construction discussion also needs to establish the presence of or lack of isotropy. If anisotropy is present, its nature needs to be established, and it needs to be addressed by variogram adjustments similar to the adjustments presented in section 5.4.2.

The variogram cross-validation statistics described in section 5.8 are useful and, if available, they can aid in the evaluation of a kriging application; authoritative and definitive kriging applications should include cross validation. Often, the most useful variograms have cross-validation statistics that conform to the guidelines discussed in section 5.8. Section 5.8.1 indicates that the cross-validation exercise needs to balance minimizing the kriging cross-validation errors with efforts to guard against bias. Also, as discussed in section 5.8.1, if probabilistic statements are part of the kriging application, there needs to be some investigation of the normality of the reduced kriging error, such as the cross-validation probability plots included with the examples in section 6.0.

Maps of the kriging estimates and standard deviations need to be presented or discussed. The maps of kriging estimates need to conform to any qualitative information about the information portrayed on the maps that is available to the investigator. The maps of kriged standard deviations can be used to delineate areas of large uncertainty in the kriging estimates.

Finally, the variogram and kriging algorithms are intended for interpolation rather than extrapolation tools. Once the application extends to areas beyond the geographic extremes of the measured data, or perhaps those extremes plus the range, there needs to be some qualification of the area of extrapolation. For instance, in universal kriging, the practitioner would need to have some assurance that the conditions of drift defined in the study area continue into the area of extrapolation.

7.3 Errors in Measured Data

Data associated with HTRW-site investigations have the same possibilities for errors that most investigations do. The errors may involve, among others, bias, inaccuracy, or lack of representativeness. The classical nature of these errors is described in a publication by the U.S. Army Corps of Engineers (1995), which describes HTRW data-quality design.

The presence of contamination may complicate the function of errors in HTRW-site investigations. Because these investigations often concern contamination, there can be large ranges of values for data involving contaminant concentrations, and these large ranges have a tendency to increase the incidence of data that may seem to be statistical outliers. Even more complicating is the presence of high concentrations of organic materials that may create challenging analytical problems in laboratory determinations that also may result in reported values that seem to be statistical outliers. In either case, the kriging practitioner is likely to find that the apparent outliers have a strong effect on the results of the kriging application.

When HTRW-site investigations find data that seem to be outliers, the data need to be very carefully evaluated before removal is seriously contemplated. Automated outlier detection tools, as suggested in section 5.7, may be best used to identify points that may be outliers and warrant further investigation. Often data that appear to be outliers may be the most important and meaningful data of all measurements. For example, in the first case described in the preceding paragraph, apparent outliers often are representative values. In the second case, the reported value may be an erroneous determination that has been affected by the extremely contaminated nature of the sample matrix. The investigator needs to either possess or have access to qualitative or institutional knowledge of the study area that aids in outlier interpretation.

8.0 OTHER SPATIAL PREDICTION TECHNIQUES

In this section, some alternative approaches to spatial prediction are discussed. At the beginning of section 3.0, the distinction between stochastic and nonstochastic techniques for spatial prediction was discussed. Kriging is a stochastic technique because of the structure that is imposed in terms of an underlying random process (the regionalized variables) with joint probability distributions that obey certain assumptions. Kriging yields the predictor that is statistically optimal in that it is the best linear unbiased predictor, given certain assumptions that are detailed in section 3.0. There are other stochastic techniques that are less well known, such as Markov-random-field prediction and Bayesian nonparametric smoothing (Cressie, 1991), but these techniques are not discussed here.

Several techniques that are often applied in a nonstochastic setting are discussed. Techniques applied in a nonstochastic setting are generally applied strictly empirically and are not evaluated by rigorous statistical criteria, such as mean-squared prediction error—although, as discussed in section 3.0, such criteria may be applied in certain of the techniques, such as simple average and trend analysis. As indicated in this report, there are some compelling advantages for assuming some kind of stochastic setting. However, the simplicity of not postulating and justifying the structure and assumptions inherent in stochastic analyses might be considered one advantage of nonstochastic techniques, and a nonstochastic analysis may be perfectly adequate for certain problems. In addition to statistical optimality and simplicity, there are other considerations in selecting a spatial-prediction technique—including properties such as ease of computation, sensitivity to data errors, and whether the predictors are exact interpolators; that is, the interpolators match the measurements exactly at the measurement locations $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n$. The last property is one that needs to be given careful consideration. Kriging, as usually applied, is an exact interpolator. Questions may be raised, however, about whether exact interpolation is a desirable property if the measurements are contaminated with a considerable measurement error. One advantage of stochastic techniques is that, generally, the existence of measurement error may be incorporated objectively; in fact, some kriging software packages (including STATPAC) have incorporated this feature, resulting in a surface that is not an exact interpolator. Several of the nonstochastic techniques discussed in this section depend on a parameter that controls the deviation from exact interpolation. The ability to adjust such a parameter when using these techniques lends a degree of flexibility, but selecting the best value may not be straightforward and may involve considerable subjectivity.

In most of the following techniques, the predictor of the process at location $\underline{x}_0$ is a linear combination of the measurements at locations $\underline{x}_i$, $i = 1, 2, ..., n$. Using $\tilde{Z}(\underline{x}_0)$ to denote an arbitrary predictor [this notation distinguishes the predictors to be discussed in this section from the kriging predictor, which is denoted by $\hat{Z}(\underline{x}_0)$], the definition of $\tilde{Z}(\underline{x}_0)$ is

$$\tilde{Z}(\underline{x}_0) = \sum_{i=1}^{n} w_i Z(\underline{x}_i) . \qquad (8\text{–}1)$$

Although this form is the same form that is taken by the kriging predictor, the difference is in the way the coefficients w_i are computed.

8.1 Global Measure of Central Tendency (Simple Averaging)

The predictor for the process at any location $\underline{x}_0$ is the simple average of the measurements; that is, the weights w_i are all equal and are given by (Cressie, 1991)

$$w_i = \frac{1}{n} . \qquad (8\text{–}2)$$

This predictor represents the smoothest possible predictor surface. In using this predictor, a certain degree of spatial homogeneity is assumed. No attempt is made to incorporate any detectable patterns (or trends) in the mean or variance of the data as a function of location, and the fact that measurements made at points that are close together may be related is disregarded. Such a predictor has the advantage of being very simple to compute; it needs no estimation of a variogram or other model parameters. The disadvantage is that representing the spatial field by a single value ignores much of the relevant and interesting structure that may be very helpful in improving predictions. As discussed in section 3.3, if applied in a stochastic setting, this predictor would be optimal (best linear unbiased) if there is no drift and if residuals are uncorrelated and have a common variance.

8.2 Simple Moving Average

Let h_{i0} be the distance of $\underline{x}_0$ from $\underline{x}_i$, let $h_{[i0]}$ be the ordered (from smallest to largest) distances, and fix $1 \le k \le n$. Then the weights w_i are (Cressie, 1991)

$$w_i = \begin{cases} \frac{1}{k}, & h_{i0} \le h_{[k0]} \\ 0, & h_{i0} > h_{[k0]} \end{cases} . \qquad (8\text{–}3)$$

Thus, this predictor is the average of the measurements at the k nearest locations from $\underline{x}_0$.

If k is equal to n, this predictor is identical to the simple average, with weights as given in equation 8–2. A choice of k smaller than n assumes that the predictor needs to incorporate more of the local fluctuation measured in the data, or, equivalently, that measurements at locations near $\underline{x}_0$ need be more informative than measurements at other locations in predicting $z(\underline{x}_0)$; the smaller k is, the more variable the predictor. If $k = 1$, the predictor is an exact interpolator and is constant on the Voronoi polygons (see section 8.4) induced by the measurement locations.

There are several variations of this predictor. In one such variation, a distance r may be fixed (rather than fixing k) and averages over locations that are within distance r of $\underline{x}_0$ may be obtained. Additionally, a moving median may be used rather than a moving average. Sorting and testing distances can slow computations compared to obtaining the simple average, and use of medians rather than means results in a predictor that is more resistant to outliers.

8.3 Inverse-Distance Squared Weighted Average

The weights w_i are (Journel and Huijbregts, 1978)

$$w_i = \frac{\frac{1}{h_{i0}^2}}{\sum_{j=1}^{n} \frac{1}{h_{j0}^2}} , \qquad (8\text{–}4)$$

where again h_{i0} is the distance of $\underline{x}_0$ from $\underline{x}_i$.

In the simple moving average, weights are the same, provided the measurement locations are sufficiently close to the prediction locations and are zero otherwise. For the inverse-distance squared method, weights are forced to decrease smoothly as distance from the prediction location increases. This predictor again has the advantage of being easy to compute. Another feature of this predictor is that it is an exact interpolator. In addition, the exponent 2 of h_{i0} may be changed to any positive number, providing some flexibility in determining the rate of decrease of weights as a function of distance from $\underline{x}_0$. Isaaks and Srivastava (1989, p. 257–259) presented an example describing the effects on weights of changing the exponent.

8.4 Triangulation

To compute this predictor, the region R is partitioned into what are referred to as Voronoi polygons $V_1, V_2, ..., V_n$, with V_i being the set of locations closer to measurement location $\underline{x}_i$ than to any other measurement location. If any two polygons, V_i and V_j, share a common boundary, $\underline{x}_i$ and $\underline{x}_j$ are joined with a straight line. The collection of all such lines defines what is known as the Delauney triangulation. One such triangle contains the prediction location $\underline{x}_0$; the vertices of this triangle, which are measurement locations, are labeled $\underline{x}_j$, $\underline{x}_k$, and $\underline{x}_l$. The spatial prediction at $\underline{x}_0$ is the planar interpolant through the coordinates $[\underline{x}_j, z(\underline{x}_j)]$, $[\underline{x}_k, z(\underline{x}_k)]$, and $[\underline{x}_l, z(\underline{x}_l)]$. By joining $\underline{x}_0$ and $\underline{x}_j$, $\underline{x}_k$, and $\underline{x}_l$, three subtriangles are formed. The weights w_i are (Cressie, 1991)

$$w_i = \begin{cases} \dfrac{A_i}{A_j + A_k + A_l}, & i = j, k, \text{ or } l \\ 0, & \text{otherwise} \end{cases} \qquad (8\text{–}5)$$

where

A_i is the area of the subtriangle opposite vertex $\underline{x}_i$.

These definitions are shown in figure 25. In this figure, the dashed lines depict the Voronoi polygons associated with points $\underline{x}_1, \underline{x}_2, ..., \underline{x}_6$, and the solid lines define the Delauney triangulation. Vertices of the triangle containing the prediction point $\underline{x}_0$ are $\underline{x}_1$, $\underline{x}_5$, and $\underline{x}_6$, and dotted lines represent the subtriangles defining the associated area A_1, A_5, A_6. For this example, j, k, and l in the general equation 8–5 are 1, 5, and 6, so the weights assigned to points $\underline{x}_1$, $\underline{x}_5$, and $\underline{x}_6$ are

$$w_1 = \frac{A_1}{A_1 + A_5 + A_6},$$

$$w_5 = \frac{A_5}{A_1 + A_5 + A_6}$$

and (8–6)

$$w_6 = \frac{A_6}{A_1 + A_5 + A_6},$$

The weight assigned to a point is proportional to the area of the triangle opposite the point.

Computation of this predictor is slower than computation of the predictors in sections 8.1, 8.2, and 8.3. The predictor is an exact interpolator, and the surface produced is continuous but not differentiable at the edges of the triangulation.

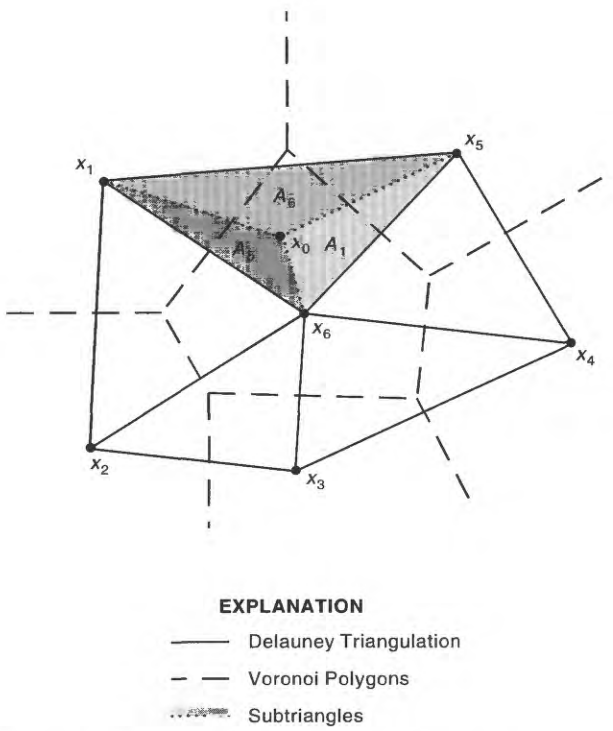


Figure 25. Diagram showing Voronoi polygons.

8.5 Splines

In spline modeling, the measurements are interpolated using combinations of certain so-called basis functions. These basis functions are usually piecewise polynomials of a certain degree that are determined by the user; let this degree be k. The coefficients of these polynomials are chosen so that the function values and the first k–1 derivatives agree at the locations where they join. The larger k is, the smoother the prediction surface is. Spline techniques are often applied in a nonstochastic framework; in such, they represent a way of fitting a surface that has certain smoothness properties to measurements at a set of locations with no explicit consideration of statistical optimality. There is, however, a considerable body of work in which this technique is applied in a stochastic setting. Splines may be used, for example, in nonparametric regression estimation problems (Wegman and Wright, 1983).

A typical approach to formulating a spline problem is to pose the problem as an optimization problem. In one special formulation, the first two derivatives of the prediction surface are assumed to exist, which imposes a certain degree of smoothness, and the spline function is assumed to minimize

$$\frac{1}{n}\sum_{i=1}^{n}[z(\underline{x}_i)-\tilde{z}(\underline{x}_i)]^2+\eta Q \ , \qquad (8\text{–}7)$$

where

Q is a term that depends on the first two derivatives of the predictor surface.

The parameter η is a nonnegative number that needs to be specified by the user; the value of this parameter indicates the trade-off between goodness of fit to the data, measured by the first term, and smoothness, as measured by Q. If η is 0, the spline is an exact interpolator and passes through all the data points. If $\eta > 0$, the spline is not an exact interpolator. (Splines that are not exact interpolators are referred to as smoothing splines.) There are a number of numerical procedures that may be used for fitting splines, but allowing the smoothing parameter η to be greater than 0 renders the computational problem more complex.

Under some conditions, a solution to the optimization problem (eq. 8–7) also may be obtained by a kriging algorithm, if the smoothing parameter η is equal to the variance of the measurement error and if a special form is chosen for the covariance function. In this situation, spline approximation is a special type of kriging. However, the variogram that needs to be used in the kriging equations to make the kriging predictor equal to the spline predictor is determined by the basis functions selected for the spline. Because the basis functions selected are subjective on the part of the user, the resulting equivalent variogram may not be representative of the true variogram of the data. Because kriging uses the data to indicate reasonable variogram choices, kriging has an important advantage over splines. Another advantage of using the kriging framework is the interpretation of the smoothing parameter in terms of measurement errors. Many times, an objective estimate of the magnitude of the measurement error can be obtained. The connections between kriging and splines are discussed further by Wegman and Wright (1983), Watson (1984), and Cressie (1991).

8.6 Trend-Surface Analysis

Trend-surface analysis is the process of fitting a function, such as that in equation 3–43 to the data, using least squares to determine the coefficients that yield the best fit. Computationally, trend-surface analysis is equivalent to universal kriging, with an assumption that the $Z^*(\underline{x}_i)$'s in equation 3–16 are uncorrelated. Thus, there is no need to estimate a variogram, and readily available regression packages may be used for estimating the coefficients. As in universal kriging, polynomial surfaces are the most commonly used. When trend surfaces are applied in a stochastic setting, the resulting predictor is optimal if deviations from the surface are uncorrelated and have a common variance.

8.7 Simulation

In this section, a regionalized random variable $Z(\underline{x})$, where $\underline{x}$ is a location in a two-dimensional study region R, is considered. Kriging is an interpolation algorithm that yields spatial predictions $\hat{Z}(\underline{x})$ that are optimal, as has been discussed in this report. The

mean-squared prediction error is smallest among all predictors that are linear in the measurements. This optimality property is local, in that the mean-squared error of predictions at unsampled locations, when considered one at a time, is minimized without specific regard to preservation of global spatial features. If, however, the actual realization $z(\underline{x})$ could be compared to the kriged prediction surface based on n measured values, the kriged surface would be much smoother than the actual surface, especially in areas of sparser sampling. Thus, the kriged surface is a good and realistic representation of reality in that the n measured values are honored, but the kriged surface is less realistic for global properties, such as overall variability.

The purpose of simulation is to produce one or more spatial surfaces (realizations) that are more realistic in preserving global properties than the surface produced by interpolation algorithms, such as kriging. These realizations are produced by using numbers that are drawn randomly (Monte Carlo) to impart variability to the simulated surface, making the simulated surface more representative of the overall appearance of the actual surface. Generally, simulation uses the idea that the true value of a random surface may be expressed as the sum of a predicted value (which is obtained by kriging) plus a random error, which varies spatially and depends on the random numbers drawn. A number of independent realizations are generated, and these realizations use equally probable representations of reality.

A simulation algorithm is said to be conditional if the resulting realizations agree with the measurements at measurement locations $\underline{x}_1, \underline{x}_2, \ldots, \underline{x}_n$. If the underlying process $Z(\underline{x})$ is assumed to be Gaussian (or if a transformation is found that makes the process Gaussian), the most common technique of conditional simulation is known as sequential Gaussian simulation (Deutsch and Journel, 1992, p. 141–143). Another, more complicated, Gaussian simulation technique that is particularly useful for three-dimensional simulations because of its computational efficiency is the turning-bands technique (Journel and Huijbregts, 1978; Deutsch and Journel, 1992).

In sequential Gaussian simulation, a set of grid points for which simulated values are desired is defined, and the points are addressed sequentially from location to location along a predetermined path. At each location, a specified set of neighboring conditioning data are retained, including the original data and simulated grid-location values at previous traversed grid locations along the path. Then, a random number is generated from a Gaussian distribution in which the conditional mean and the variance are determined using a kriging algorithm. The value of the random number determines the simulated process at this location. The conditional Gaussian distribution used in the simulation is identical to the conditional distribution discussed in section 3.5.1. An idea of the computational requirements can be obtained from the fact that a kriging algorithm needs to be applied for each simulation location. For multiple realizations, if the path connecting the grid points remains the same, the kriging equations need to be solved for only the first simulation. However, implementation of this procedure needs to account for the assumptions concerning the existence of drift; the details of such an implementation are beyond the scope of this report.

A sequential Gaussian simulation also may be applied in indicator kriging (see section 3.5.2). At each grid point along the path, a (Bernoulli) random variable that has only two possible values, 0 or 1, is generated, with the relative probability of these two values being determined by indicator kriging applied, as in the previous paragraph, to the original observed indicator data and the previously simulated indicator values.

To get an idea of how simulation results might be used in a risk-assessment setting, assume again that the underlying process is Gaussian and that 1,000 conditional realizations have been generated. If a single grid point $\underline{x}_0$ (which is not a measurement point) is used, then the simulation has produced 1,000 values at $\underline{x}_0$, which, when analyzed in histogram form, approximate the probability distribution of potential measurements at that location. If an interval that has exactly 25 (2.5 percent) of the values less than its lower end and 25 of the values larger than its upper end were established, the interval would almost correspond to the 95-percent prediction interval $\hat{Z}(\underline{x}_0) - 1.96\sigma_K(\underline{x}_0)$ to $\hat{Z}(\underline{x}_0) + 1.96\sigma_K(\underline{x}_0)$ discussed in section 3.5.1. For this single location, the simulation would not produce much more information than kriging alone would have produced. The real value of simulation is that realizations, not just

at a single location, but at all of the grid locations jointly, can be obtained. These realizations then can be used to calculate probabilities associated with any number of spatial locations together. For example, the probability that the largest (maximum) contaminant value over a certain subregion is greater than a particular concentration might be assessed. (If the word "largest" here were replaced with "average," then block kriging could be used to obtain the answer.)

A central point that needs to be emphasized is that simulation is especially useful when probabilities associated with complicated, usually nonlinear, functions of the regionalized variables for a region need to be analyzed. The maximum function mentioned in the preceding paragraph is one simple example. Another example is the problem of determining placement of ground-water monitoring wells to detect and monitor ground-water contamination emanating from a potential point source. Given an existing set of hydraulic-head data, kriging might be applied and flow paths determined from resulting hydraulic-head gradients. Intersection of the flow path from the point source with the regional boundary then might be used to determine monitoring-well placement. Conditional simulation would be useful to determine the uncertainty associated with well placement or to give an indication of how many monitoring wells might be appropriate. In this example, the variable of interest, well location, is a complicated function of hydraulic heads so this is a problem for which simulation is well suited. The reader may refer to Easley and others (1991) for a more detailed discussion of this application of kriging.

The complicated functions of interest in ground-water studies often involve physically based ground-water flow models. Conditional simulation may be used, for example, to generate a suite of hydraulic-conductivity realizations to be used as input to a model that produces as output a set of corresponding hydraulic-head realizations. Weber and others (1991) discussed how ground-water modeling might be used with conditional simulation to study the monitoring-well-placement problem discussed in the preceding paragraph.

9.0 SUMMARY

The geostatistical technique known as kriging can be used to determine optimal weighting of measurements at sampled locations for obtaining predictions, or kriging estimates, at unsampled locations. Kriging also provides information concerning the uncertainty associated with kriging estimates. The uncertainty information available from kriging, as well as the optimal weighting, distinguishes kriging from other techniques used for spatial modeling.

The theory of regionalized random variables is the basis for different forms of kriging. Ordinary kriging is used when the spatial mean is considered constant. Universal kriging is an extension of ordinary kriging that can be used to address a nonconstant spatial mean. Block kriging is used to obtain kriging estimates for a block of area that is larger than the area represented by an individual sample. Indicator kriging implements the kriging equations nonparametrically.

A fundamental step in kriging applications is development of a variogram. The variogram is usually developed from the results of measurements at many locations within the application area. The variogram describes spatial correlation within the application area and provides basic information required to determine optimal weights for measurements to be used in making kriging estimates. Information from the exercise of variogram development can be used to cross validate the variogram and the cross-validation statistics can, in turn, fine tune variogram development.

Example applications of kriging illustrate basic techniques and some constraints that apply to kriging. These applications also illustrate how different types of kriging, such as ordinary, universal, block, and indicator, can be used.

Other spatial modeling techniques include nonstochastic techniques such as simple averaging, inverse-distance squared weighted averaging, triangulation, splines, and trend-surface analysis. These nonstochastic techniques can be simpler to apply than kriging and may be appropriate to use for some problems, especially when it is not necessary to evaluate results with respect to statistical criteria. Another extension to kriging, simulation, is intended to preserve overall variability and to compensate for the tendency of kriging to smooth results.

10.0 REFERENCES

American Society for Testing and Materials,1994, Standard guide for content of geostatistical site investigations, D 5549–94: Philadelphia, Pa., American Society for Testing and Materials Committee D–18 on Soil and Rock, 5 p.

———1996, Standard guide for analysis of spatial variation in geostatistical site investigations D 5922–96: Philadelphia, Pa., American Society for Testing and Materials Committee D–18 on Soil and Rock.

Bras, R.L., and Rodriguez-Iturbe, Ignacio, 1985, Random functions and hydrology: Reading, Mass., Addison-Wesley Publishing Company, 559 p.

Clark, Isobel, 1979, Practical geostatistics: London, Applied Science Publishers, 129 p.

Cliff, A.D., and Ord, J.K., 1981, Spatial processes, models and applications: London, Pion Limited, 266 p.

Cressie, Noel, 1991, Statistics for spatial data: New York, Wiley, 900 p.

David, Michel, 1977, Geostatistical ore reserve estimation *in the collection* Developments in geomathematics 2: Amsterdam, Elsevier, 364 p.

Davis, J.C., 1973, Statistics and data analysis in geology: New York, Wiley, 646 p.

Delhomme, J.P., 1978, Kriging in the hydrosciences: Advances in Water Resources, v. 1, no. 5, p. 251–266.

Deutsch, C.V., and Journel, A.G., 1992, GSLIB, geostatistical software library and user's guide: New York, Oxford University Press, 340 p.

Devore, J.L., 1987, Probability and statistics for engineering and the sciences: Monterey, Calif., Cole Publishing Company, 672 p.

Easley, D.H., Borgman, L.E., and Weber, D., 1991, Monitoring well placement using conditional simulation of hydraulic head: Mathematical Geology, v. 23, no. 8, p. 1059–1080.

Englund, E.J., and Sparks, A.R., 1991, GEO-EAS (Geostatistical Environmental Assessment Software) users guide: Las Vegas, Nev., Environmental Monitoring Systems Laboratory Report EPA/600/8–91/008. [Available from National Technical Information Service, Springfield, VA 22161 as NTIS Report PB89–151 252.]

Grundy, W.D., and Miesch, A.T, 1987, Brief description of STATPAC and related statistical programs for the IBM personal computer—A, Documentation: U.S. Geological Survey Open-File Report 87–411–A, 34 p.

Hawkins, D.M., 1980, Identification of outliers: London, Chapman and Hall, 188 p.

Isaaks, E.H., and Srivastava, R.M., 1989, An introduction to applied geostatistics: New York, Oxford University Press, 561 p.

Journel, A.G., 1988, Non-parametric geostatistics for risk and additional sampling assessment, *in* Kieth, L., ed., Principles of environmental sampling: Washington, D.C., American Chemical Society, p. 45–72.

Journel, A.G., 1993, Geostatistics for the environmental sciences: Las Vegas, Nev., U.S. Environmental Protection Agency, 135 p.

Journel, A.G., and Huijbregts, C., 1978, Mining geostatistics: London, Academic Press, 600 p.

Krige, D.G., and Magri, E.J., 1982, Studies of the effects of outliers and data transformation on variogram estimates for a base metal and a gold ore body: Mathematical Geology, v. 14, no. 6, p. 557–564.

Lenfest, L.W., Jr., 1986, Ground-water levels and use of water for irrigation in the Saratoga Valley, south-central Wyoming, 1980–81: U.S. Geological Survey Water-Resources Investigations Report 84–4040, 23 p.

Myers, D.E., Begovich, C.L., Butz, T.R., and Kane, V.E., 1980, Application of kriging to the hydrogeochemical data from the National Uranium Resource Evaluation Program: Oak Ridge, Tenn., Union Carbide Corporation, Nuclear Division Oak Ridge Gaseous Diffusion Plant, Report GJBX–57–81, 124 p. [Available from U.S. Department of Energy, Grand Junction, Colorado.]

Olea, R.A., 1991, Geostatistical glossary and multilingual dictionary, *in the collection* Wiley series in probability and mathematical statistics: New York, Oxford University Press, 177 p.

Ripley, B.D., 1981, Spatial statistics: New York, Wiley, 252 p.

Ross, S.M., 1987, Introduction to probability and statistics for engineers and scientists: New York, Wiley, 492 p.

U.S. Army Corps of Engineers, 1995, Technical project planning—Guidance for hazardous, toxic, and radioactive waste data quality design: Washington, D.C., U.S. Army Corps of Engineers Engineer Manual 200–1–2.

_____1997, Practical aspects of applying geostatistics at hazardous, toxic, and radioactive waste sites: Washington, D.C., U.S. Army Corps of Engineers, Engineer Technical Letter ETL 1110–1–175.

Watson, G.S., 1984, Smoothing and interpolation by kriging and with splines: Mathematical Geology, v. 16, no. 6, p. 601–615.

Weber, D., Easley, D.H., and Englund, E.J., 1991, Probability of plume interception using conditional simulation of hydraulic head and inverse modeling: Mathematical Geology, v. 23, no. 2, p. 219–239.

Wegman, E.J., and Wright, I.W., 1983, Splines in statistics: American Statistical Association Journal, v. 78, no. 382, p. 351–365.