SYSTEMATIC SAMPLING OF SPATIAL FUNCTIONS

1984

R. A. OLEA

SYSTEMATIC SAMPLING OF SPATIAL FUNCTIONS

R. A. OLEA

1984

KANSAS GEOLOGICAL SURVEY
UNIVERSITY OF KANSAS, LAWRENCE, KANSAS
SERIES ON SPATIAL ANALYSIS NO. 7

TABLE OF CONTENTS

LIST OF FIGURES

PAGE

LIST OF TABLES

ABSTRACT

A procedure is presented to minimize the number of samples required to estimate a spatial function at a specified level of accuracy. The technique is based on universal kriging, which is an estimation method derived from the theory of regionalized variables. Statistical characteristics of spatial functions, such as spatial continuity, statistical dependence among closely spaced sample elements, and the practical impossibility of making error-free estimates, are explicitly considered in the method. Universal kriging is used to compute the average standard error and maximum standard error of the estimates over the sampling domain; these are used as indices of sampling efficiency. The procedure optimally selects the factors controlling the magnitude of the indices, such as the density and spatial pattern of the sample elements, and the number of nearest sample elements to be used in estimation. As a demonstration, the network of observation wells used to monitor the water table in the Equus Beds of Kansas is enhanced. This example illustrates the ease and convenience of the procedure, which can be used equally well to design sampling programs for other spatial variables, as the procedure is not limited by the physical nature of the function.

CHAPTER 1

INTRODUCTION

Over the years, engineers and scientists in different disciplines have faced the problem of characterizing n-dimensional functions which are only partly known at discrete points. Great emphasis has been given to the search for different methods to analyze such data. Some methods have a theoretical basis, others are justified purely on empirical grounds, and some have no apparent justification (Henley, 1981, p. 10). Although the selection of an analytical procedure has a great influence on the results obtained in a study, the answers also depend on the characteristics of the sample being analyzed, a fact that has received relatively little attention.

The nature of the sampling problem varies widely, ranging from difficulties caused by having too many sample elements to instances where nothing can be done because of the lack of sufficient information. Despite the massive use of computers, advances in the technology of data collection are shifting the analytical bottleneck from data gathering to data processing. Today it is not uncommon to have huge amounts of sample data obtained at high cost in a given area, and to be faced with the prospect of expending enormous additional amounts of money and time in order to see any final results. In these instances, judicious selection of a subset of the data may be of significant help in reducing costs and speeding up the analysis. A final, exhaustive processing of the data can be postponed until preliminary results demonstrate that more detail is required. The preliminary study can be used to pinpoint subareas of interest and to schedule priorities for data processing. The success of a preliminary study depends greatly on the selection of an appropriate subset of the sample.

Rather than too much data, the problem facing a researcher is more often a partial or total lack of information. In such circumstances a careful sampling design is critical in order to maximize the information that can be collected within a usually limited budget. The designer of the sampling scheme would like to know, in a short time and at a minimum cost, the implications of alternative sample arrangements, and the effects of varying the number of sample elements and of adjusting the sampling uncertainty within specified limits.

Although regionalized variable theory has been extensively described in the geomathematical and statistical literature, almost all discussions focus on either the problem of estimation or on simulation. Only a few studies have marginally addressed the question of the efficient arrangement of sample elements (Alldredge and Alldredge, 1978; Ripley, 1981, p. 214-241). This work addresses directly the systematic analysis of spatial functions from the viewpoint of sampling. The primary objectives of the research which led to this report were:

1. To analyze the mathematical nature of spatial functions and to select an estimation procedure which would best describe them.

2. To identify parameters characterizing the quality of a sample for the purpose of forecasting unknown values at locations not considered in the sample.

3. To determine the factors that control values of the parameters that characterize sampling quality.

4. To present a methodology for optimizing the quality of a sample by properly selecting the controlling parameters.

5. To apply the optimal sampling methodology to a real case.

6. To generate recommendations leading to the best use of existing sampling systems and the optimal implementation of future data-gathering networks.

CHAPTER 2

REVIEW OF THEORY AND PRINCIPLES

Certain basic statistical concepts are necessary in order to understand the methodology used to determine optimal sampling schemes for spatial functions. These concepts are presented following an original line, but the theories, equations, and concepts are derived from the existing literature. References have been selected on the bases of scientific accuracy and clarity, and should provide a greater understanding of the general topics of sampling theory and spatial functions.

2.1 SPATIAL FUNCTIONS

2.1.1 Nature

Spatially arranged measurements occur in a surprisingly wide variety of scientific and engineering disciplines. Meteorologists study changes in temperature and barometric pressure recorded at weather stations. Foresters and agricultural scientists investigate the lateral variation in soils and in harvest yields. Geologists and petroleum engineers estimate variations in formation thickness and reservoir porosity. Mathematically, a **spatial function** is an association of numbers to a domain of geographical coordinates. Spatial functions such as the annual amount of rainfall are time dependent; others, such as the thickness of a geological formation, are invariant at the human scale of time.

Typical spatial functions are continuous and uniquely defined over sizable domains. Some, such as geothermal gradients, are not easily measured and are expensive and time-consuming to characterize accurately. Commonly, such functions are only partially known through a set of scattered observations. In statistical jargon, the selected observations are called a **sample** (James and James, 1976, p. 339), and each individual measurement is called a **sample element** (Williams, 1978, p. 27).

Even if the observations have been carefully taken to avoid measurement error, the spatial function is known with certainty only at the sampled locations. The exact value of the spatial function at unsampled locations is subject to uncertainty, as no method has yet been devised which will yield error-free estimates. Figure 2.1 is a diagrammatic cross-section based on two sample elements at locations A and B, where the spatial function is known. Any surface is a potential description of the real spatial function at unsampled locations, such as C. The four alternatives presented in Figure 2.1 are simply a small subset of all possible cases. For some arbitrary location such as C, a table can be prepared containing all the estimated values at that location. The minimum and maximum values in the table will provide an interval which encloses all likely answers to the problem. Based on supplementary information, there could be a concensus that solution (a) in Figure 2.1 is less likely than solution (d). The insight provided by the table of possible values for the spatial function at location C can be enhanced by incorporating the likelihood that a specific solution is the correct answer.

2.1.2 Modeling

A tabulation of events and their associated probability of occurrence corresponds to the statistical concept of a **probability density function** (Hogg and Craig, 1978, Section 1.6). Figure 2.2 represents a hypothetical probability density function for all likely values of the spatial function at some arbitrary unsampled location C in Figure 2.1. In statistical terms, a quantity which may take any of the values of a certain set with a specified relative frequency is called a **random variable** or a **variate** (Kendall and Buckland, 1971, p. 162). There will be as many random variables as unsampled locations. The set of all random variables associated with a given spatial function constitutes what is known as a **random function** (Matheron, 1971, p. 50). Even though there might be only one value of the spatial function at location C, the estimated value of the spatial function at any location not considered in the sample is more thoroughly described by a collection of likely values than by a single number.

Listing a random function describing a spatial function is a cumbersome task. It would be desirable to summarize the random function in an informative, short description rather than listing all its possible outcomes. For this purpose we may invoke the use of mathematical expectations. The average of all possible outcomes of a variate weighted by their probability of occurrence is the **mean** and represents a central value of the population. The weighted average of the squares of the differences between the outcomes and the mean is the **variance** (Hogg and Craig, 1978, p. 46). The square root of the variance is the **standard deviation.** As the

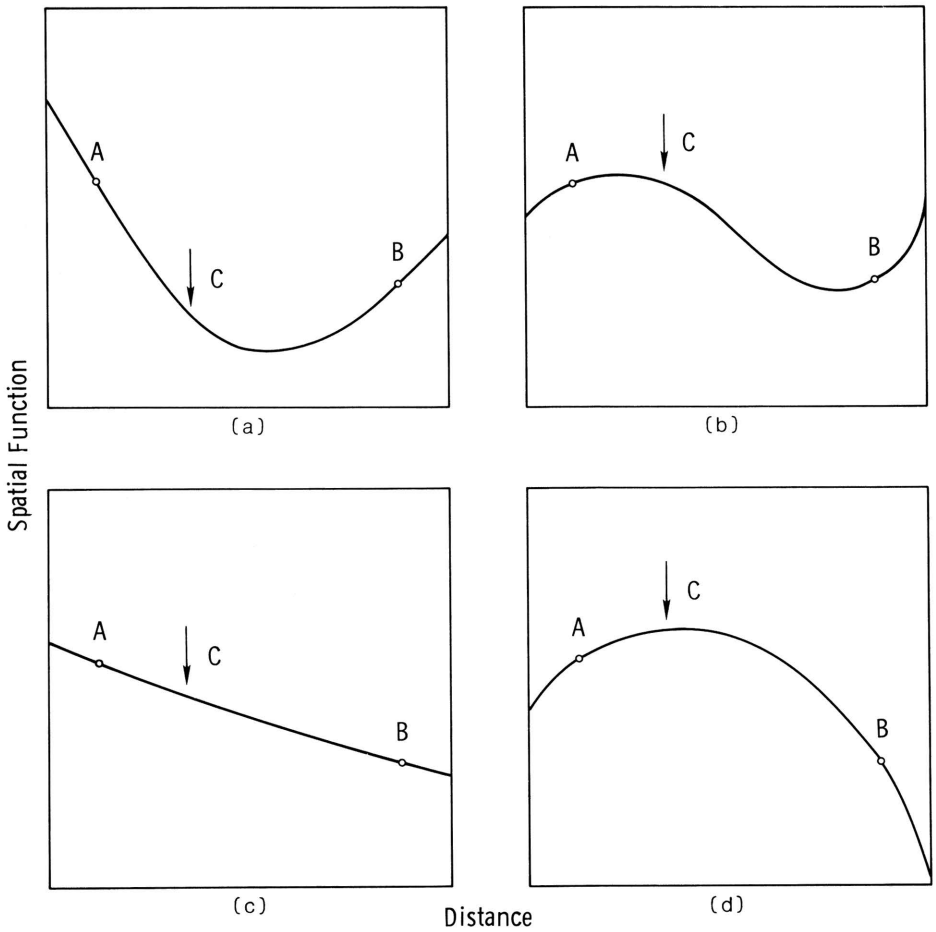


FIGURE 2.1. Cross-sectional view of a spatial function. The spatial function is known at locations A and B but is not known at other locations, such as C.

variance becomes larger when the differences increase, the variance and the standard deviation are a measure of the dispersion of the outcomes relative to the mean value. The standard deviation of Figure 2.2 also can be interpreted as a measure of the uncertainty as to the true value of the spatial function in C. A small standard deviation indicates the outcomes are clustered tightly around the central value over a relatively narrow range of possibilities. Conversely, a large standard deviation indicates that the actual value could be any of a larger range of possibilities.

2.2 THE ESTIMATION OF SPATIAL FUNCTIONS

2.2.1 Introduction

The methodology reviewed in this section was originally developed in France by Prof. G. Matheron (Matheron, 1965 and 1971) for the purpose of estimating spatial functions that arise in mine evaluation. However, Matheron's estimation methodology is based on a statistical theory which makes the procedure independent from the nature of the spatial function. Spatial functions satisfying the basic

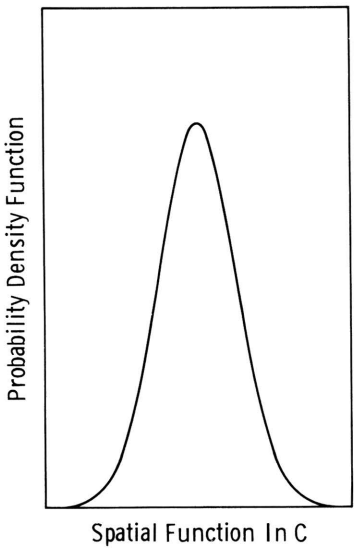


FIGURE 2.2. Probability density function of a spatial function at a location not considered in the sampling process, such as location C in Figure 2.1.

assumptions of the theory are called **regionalized variables.** These regionalized variables are assumed to be random functions, as described in the previous section, and thus are well suited for accomplishing the objectives of this study.

The term **geostatistics** has come to mean the specialized body of statistical techniques developed by Matheron and his associates to treat regionalized variables. Typical spatial functions amenable to estimation using geostatistics include the ore content of a mineralized body; the porosity of sedimentary rocks; the amount of precipitation per square mile; and the elevation of the tops of subsurface formations. Fluctuations in space- and time-dependent data are erratic and often unpredictable from one location to another, but there is an underlying trend in the fluctuations which precludes regarding the data as resulting from a completely random process. Typically, closely spaced outcomes are statistically autocorrelated. Sample elements may have a size, shape, and orientation, as, for instance, do drill cores used to assay ore grades. The geometric and spatial characteristics of a sample element constitute what is called the **support** of the sample.

An estimator which minimizes the variance of the estimate is usually considered the best for purposes of statistical inference. If, in addition, the expected value of the estimator is equal to the actual value of the parameter, the estimator is said to be **unbiased** (Williams, 1978, p. 59). Classical statistics provides numerous suggestions for reducing estimation variance, and for assessing its magnitude (Cochran, 1964; Deming, 1966; Williams, 1978). Unfortunately, classical statistical theory is seldom applicable to spatial functions; the basic assumption that elements in a sample are stochastically independent is almost always violated. The geostatistical method for estimating spatial functions is known as **universal kriging.** Basically, the method is the solution to an operations research problem in which the objective function is the estimation variance and the constraint is the condition of unbiasedness, considering that sample elements are autocorrelated.

2.2.2 Stationarity constraints

In order to make inferences, it would be desirable to define the probability density function that characterizes the regionalized variable. This is not possible in most real situations, but fortunately is unnecessary in any case. The minimum theoretical requirements do not go beyond knowledge of the first two moments of the function. As in conventional statistical inference, the spatial function can either be described by a mathematical model or given by a relative frequency analysis based on experimentation. The former approach is not practical because of the complexity of spatial functions; the latter is seriously limited by the maximum size of samples that can be collected. In fact, at any location it is only possible to collect a sample of size one, because most spatial functions are uniquely defined. To resolve this apparent impasse, geostatistics invokes a stationarity constraint similar in concept to the use of ergodicity in time series. The constraint is called the **intrinsic hypothesis** and is introduced in order to use outcomes of the regionalized variable for moment estimation. Under the hypothesis, outcomes from different locations can be considered as elements from the same population, thus allowing an unlimited increase in the sample size.

Let $Z(\vec{x})$ and $Z(\vec{x}+\vec{h})$ be two of the random variables comprising the regionalized variable. The arrow ($\rightarrow$) over the geographical coordinate indicates a vectorial property, implying orientation and magnitude in n-dimensional space. The difference between two random variables $[Z(\vec{x}) - Z(\vec{x}+\vec{h})]$ is yet another random variable. A regionalized variable is said to satisfy the intrinsic hypothesis if the difference $[Z(\vec{x}) - Z(\vec{x}+\vec{h})]$ is second-order stationary. In other words, a regionalized variable satisfies the intrinsic hypothesis if, for any displacement $\vec{h}$, the first two moments of the difference $[Z(\vec{x}) - Z(\vec{x}+\vec{h})]$ are independent of the location $\vec{x}$ and are a function only of $\vec{h}$:

$$E\,[Z(\vec{x}) - Z(\vec{x}+\vec{h})] = M(\vec{h}) \qquad (2.1)$$

$$E\,[\{Z(\vec{x}) - Z(\vec{x}+\vec{h}) - M(\vec{h})\}^2] = 2\gamma(\vec{h}) \qquad (2.2)$$

In the specialized language used in geostatistics, $M(\vec{h})$ and $\gamma(\vec{h})$ are referred to as the **drift** and the **semivariance**, or **intrinsic function.** The units of $M(\vec{h})$ are the same as the units of $Z(\vec{x})$, and the units of $\gamma(\vec{h})$ are the square of the units of $Z(\vec{x})$. Provided the intrinsic hypothesis is met, both moments can be estimated.

There are a few important circumstances under which geostatistics can be applied even if the intrinsic hypothesis does not hold. Of special interest is the situation when the regionalized variable is not first-order stationary because the drift has a systematic trend, as shown in Figure 2.3. Removing the drift from the regionalized variable results in a difference called the **residual.** Regionalized variable theory is still applicable if the intrinsic hypothesis holds for the residuals.

The assumption that the regionalized variable itself must be second-order stationary is stronger than the intrinsic hypothesis.

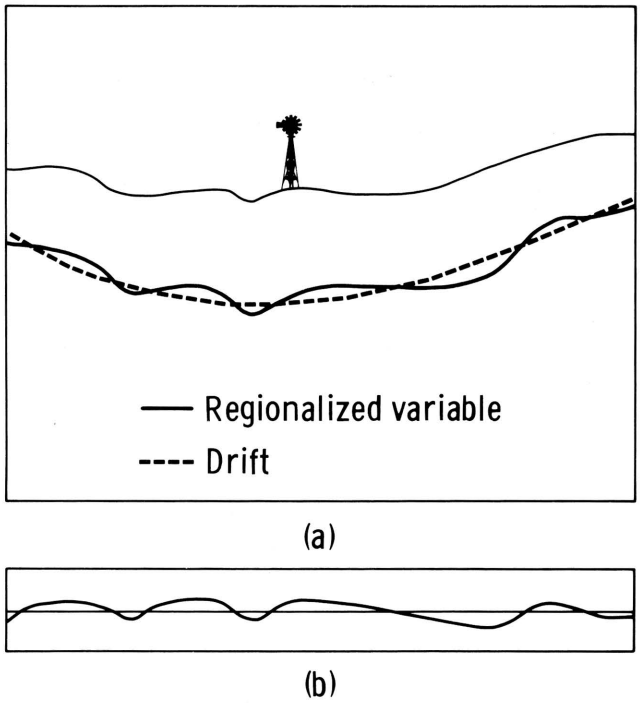


FIGURE 2.3. Basic elements in the theory of regionalized variables. (a) Regionalized variable and drift. (b) Residual after removal of drift.

Every regionalized variable which is second-order stationary satisfies the intrinsic hypothesis, but the converse is not true. For second-order stationary regionalized variables, the following relationship is true:

$$\sigma^2 = \gamma(\vec{h}) + \mathrm{cov}(\vec{h}) \qquad (2.3)$$

Here, σ^2 is the population variance, $\mathrm{cov}(\vec{h})$ is the autocovariance as used in classical statistics for a lag $\vec{h}$, and $\gamma(\vec{h})$ is the semivariance.

2.2.3 Structural analysis

The semivariance $\gamma(\vec{h})$ can be plotted against values of the lag $\vec{h}$ to yield the **semivariogram**, a graph analogous to the correlogram used in time series analysis. **Structural analysis** is the term applied to the study of semivariograms for the purpose of extracting information about the nature of the spatial variation in a regionalized variable. The objective of structural analysis is two-fold: to make genetic interpretations about the regionalized variable, and to provide parameters which are required in its estimation.

To obtain a semivariogram, it is necessary to sample the regionalized variable at regular intervals. Let

$$z(\vec{x}_1),\ z(\vec{x}_2),\ \ldots,\ z(\vec{x}_i),\ \ldots,\ z(\vec{x}_n)$$

be n values of either residuals or outcomes of the regionalized variable. Provided that the regionalized variable is first-order stationary and the intrinsic hypothesis holds, the following is an unbiased estimator of the semivariance (Olea, 1977, p. 20):

$$\gamma^*(h) = \frac{1}{2(k-p)} \sum_{j=k'}^{k'+k-p-1} [z(\vec{x}_j+\vec{h}) - z(\vec{x}_j)]^2 \qquad (2.4)$$

Here, $\vec{h}$ is p times the sampling interval $\vec{a}$; $k + k' \leq n$; and $p = 0, 1, 2, \ldots, k-1$. The study may be done along one traverse or, more desirably, along a series of traverses. The estimation of a semivariogram for observations of a regionalized variable is fairly straightforward, but this is not true for the semivariogram of the residuals. Before the residuals can be obtained, it is necessary to know the semivariogram. The problem is solved recursively by assuming a semivariogram, computing the drift and residuals, and comparing the resulting semivariogram to that assumed (Olea, 1975, p. 90-93; David, 1977, p. 272-274).

Satisfactory results for moderately tractable regionalized variables can be obtained by assuming drifts of the type

$$M^*(\vec{x}) = \sum_{i=0}^{n} a_i f^i(\vec{x}) \qquad (2.5)$$

where the a_i are n unknown coefficients to be determined and the $f^i(\vec{x})$ are functions of $\vec{x}$, typically monomials of the spatial coordinates up to degree 2. The smooth and slowly varying surfaces represented by $M^*(\vec{x})$ accord with the mathematical notion of the drift being a highly continuous function incorporating only the low frequency component of the regionalized variable and excluding any local fluctuations. The terms local and regional are relative and depend upon the scale of the regionalized variable. A feature that at the scale of a county could be a dominant element in the drift might appear as an anomaly when considering an entire state, and be completely negligible at a continental scale. Therefore, there is no single, unique drift for a given regionalized variable; as in curve fitting, the user must decide what should be fitted and what should be regarded as anomalous.

Certain major characteristics of the regionalized variable of interest in this study can be deduced from the semivariogram. These include:

1. Continuity. The shape of the semivariogram, and in particular its slope near the origin, is related to the regularity and smoothness of the regionalized variable. A parabolic semivariogram which is tangent to the x-axis at the origin means the variable is extraordinarily regular relative to the sampling interval. In contrast, highly erratic sequences will produce a semivariogram which is almost vertical at the origin.

2. Zone of influence. Figure 2.4 is a **transitive** type of semivariogram, expressing moderate continuity of the regionalized variable within a local neighborhood, and random behavior over larger distances. The semivariance steadily increases to a maximum value and then remains constant at a value called the **sill.** In instances where the regionalized variable is second-order stationary, the asymptote is equal to the sample variance in Equation 2.3. The arrow at the break in slope indicates the **range,** a distance which divides elements in a sample into two categories. Sample elements taken at distances which are smaller than the range are autocorrelated and are best suited to be used in common for estimation purposes. Sample elements spaced farther apart than the range are statistically independent and behave as independent random variables.

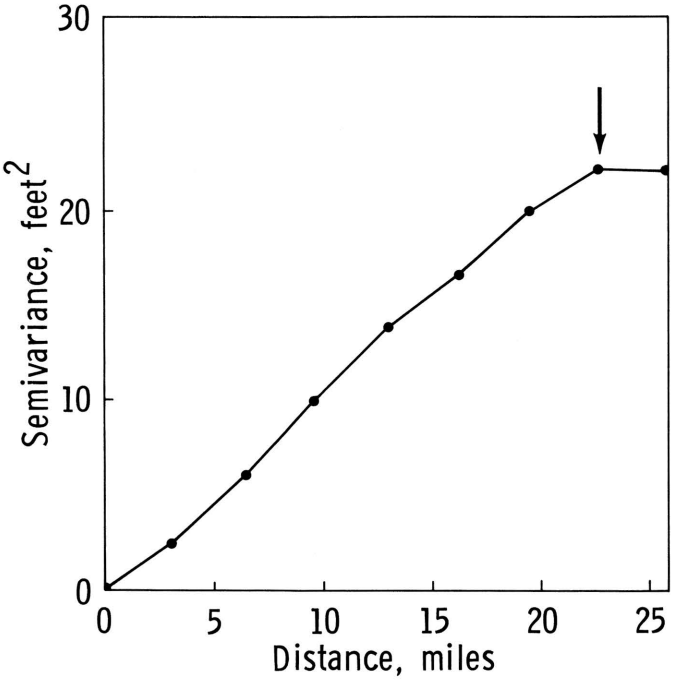


FIGURE 2.4. Transitive type of semivariogram with a range of 22.5 miles. The arrow indicates the left edge of the sill.

3. Anisotropy. Anisotropy is revealed by differing behavior of semivariograms which are computed along lines of different orientation. Differences between the semivariograms appear mainly in the slopes at the origin, in the ranges, and in the sills, if any. The semivariance of a regionalized variable can be anisotropic and at the same time the semivariance of the residuals can be isotropic, if the cause of the anisotropy is an underlying drift. This situation can be diagnosed if along the dip of the surface the semivariogram is highly regular with no sill and along the strike the semivariogram is transitive in form. If the regionalized variable is isotropic, the semivariance depends only on the magnitude h of the vector $\vec{h}$ and not on the direction h.

Although any function that fits the observed semivariance could be used as a model, sound geostatistical practice recommends the use of only those functions which are positive-definite (Journel and Huijbregts, 1978, p. 161-168; Armstrong and Jabin, 1981). Among the positive-definite models, the simplest one is linear, with a constant term equal to zero:

$$\gamma(\vec{h}) = \omega\vec{h} \qquad (2.6)$$

The linear model is convenient for representing a transitive semivariogram provided the argument h never becomes larger than the range. That is, the estimators are restricted to sample subsets inside the zone of influence.

When the model must incorporate the presence of a sill, the most commonly used function is the spherical model

$$\gamma(h) = \begin{cases} C\left[\frac{3}{2}\frac{h}{L} - \frac{1}{2}\left(\frac{h}{L}\right)^3\right], & h \leq L \\ C, & h > L \end{cases} \qquad (2.7)$$

2.2.4 Linear programming estimation problem

Universal kriging is a linear estimator of the regionalized variable and has the form:

$$Z^*(\vec{x}_0) = \sum_{j=1}^{k} \lambda_j Z(\vec{x}_j) \qquad (2.8)$$

Finding the set of weights which minimizes the estimation variance under the constraint that the estimator must be unbiased is a linear programming problem. After some algebraic manipulation (Olea, 1975, p. 67-73), the problem can be reduced to solving the system of equations

$$A\,X = B \qquad (2.9)$$

where A, B, and X are the matrices given in Equations 2.10, 2.11, and 2.12.

The Lagrange multipliers μ_i are additional slack variables which are introduced in order to solve the problem. $\gamma(\vec{x}_j,\vec{x}_p)$ is the semi-variance between two sample elements located a distance $(\vec{x}_j - \vec{x}_p)$ apart. We will restrict our attention to estimated locations having the same support as the sample elements.

$$A = \begin{bmatrix}
\gamma(\vec{x}_1,\vec{x}_1) & \gamma(\vec{x}_1,\vec{x}_2) & \dots & \gamma(\vec{x}_1,\vec{x}_k) & 1 & f^1(\vec{x}_1) & f^2(\vec{x}_1) & \dots & f^n(\vec{x}_1) \\
\gamma(\vec{x}_2,\vec{x}_1) & \gamma(\vec{x}_2,\vec{x}_2) & \dots & \gamma(\vec{x}_2,\vec{x}_k) & 1 & f^1(\vec{x}_2) & f^2(\vec{x}_2) & \dots & f^n(\vec{x}_2) \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
\gamma(\vec{x}_j,\vec{x}_1) & \gamma(\vec{x}_j,\vec{x}_2) & \dots & \gamma(\vec{x}_j,\vec{x}_k) & 1 & f^1(\vec{x}_j) & f^2(\vec{x}_j) & \dots & f^n(\vec{x}_j) \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
\gamma(\vec{x}_k,\vec{x}_1) & \gamma(\vec{x}_k,\vec{x}_2) & \dots & \gamma(\vec{x}_k,\vec{x}_k) & 1 & f^1(\vec{x}_k) & f^2(\vec{x}_k) & \dots & f^n(\vec{x}_k) \\
1 & 1 & \dots & 1 & 0 & 0 & 0 & \dots & 0 \\
f^1(\vec{x}_1) & f^1(\vec{x}_2) & \dots & f^1(\vec{x}_k) & 0 & 0 & 0 & \dots & 0 \\
f^2(\vec{x}_1) & f^2(\vec{x}_2) & \dots & f^2(\vec{x}_k) & 0 & 0 & 0 & \dots & 0 \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
f^i(\vec{x}_1) & f^i(\vec{x}_2) & \dots & f^i(\vec{x}_k) & 0 & 0 & 0 & \dots & 0 \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
f^n(\vec{x}_1) & f^n(\vec{x}_2) & \dots & f^n(\vec{x}_k) & 0 & 0 & 0 & \dots & 0
\end{bmatrix} \qquad (2.10)$$

$$X = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \cdots \\ \lambda_j \\ \cdots \\ \lambda_k \\ \mu_0 \\ \mu_1 \\ \mu_2 \\ \cdots \\ \mu_i \\ \cdots \\ \mu_n \end{bmatrix} \qquad (2.11) \qquad B = \begin{bmatrix} \gamma(\vec{x}_1,\vec{x}_0) \\ \gamma(\vec{x}_2,\vec{x}_0) \\ \cdots \\ \gamma(\vec{x}_j,\vec{x}_0) \\ \cdots \\ \gamma(\vec{x}_k,\vec{x}_0) \\ 1 \\ f^1(\vec{x}_0) \\ f^2(\vec{x}_0) \\ \cdots \\ f^i(\vec{x}_0) \\ \cdots \\ f^n(\vec{x}_0) \end{bmatrix} \qquad (2.12)$$

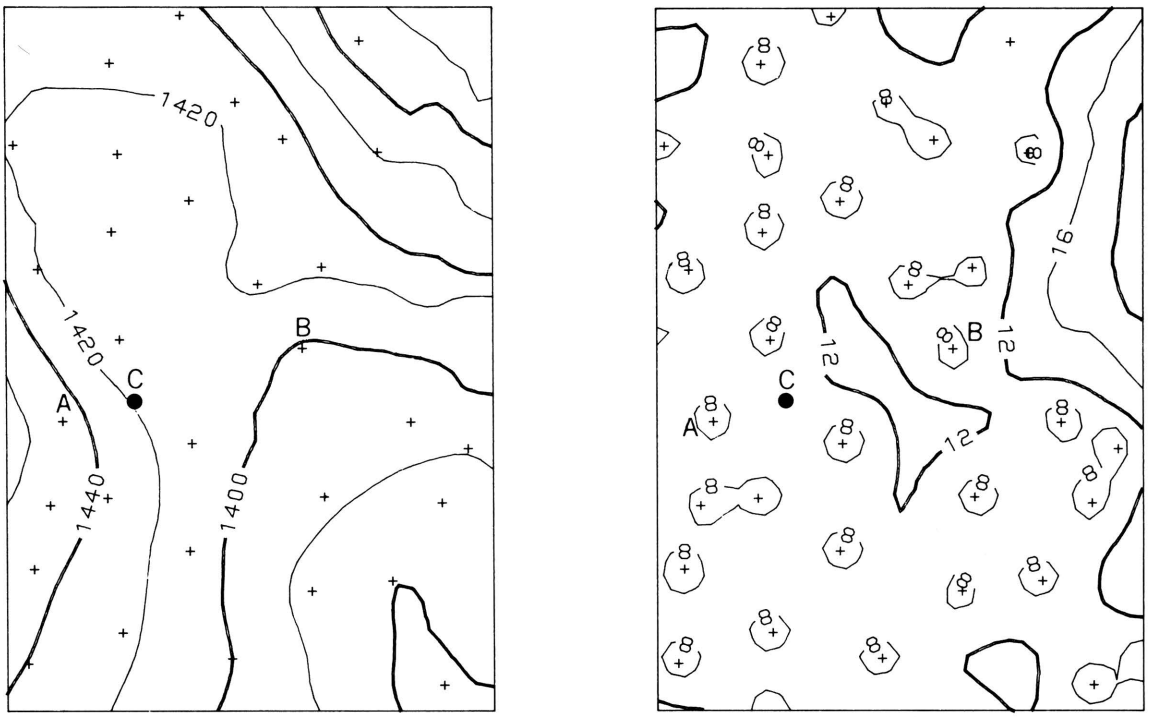


FIGURE 2.5. Graphic representation of the results of universal kriging of a two-dimensional spatial function. Crosses represent locations of sample elements. A-C-B shows location of cross-section in Figure 2.1. (a) Most likely representation of the spatial function. (b) Standard error.

The final product of universal kriging estimation at location $\vec{x}_0$ consists of two numbers. The first of these is the estimated value of the function at $\vec{x}_0$, given by inserting the values determined for the λ_j's into Equation 2.8. The second is the estimation variance, given by

$$\sigma_E^2(\vec{x}_0) = X^T B \qquad (2.13)$$

in which X^T is the transpose of matrix X.

The standard deviation of an unbiased estimator is called the **standard error** (James and James, 1976, p. 139). Since the standard error is given in the same units as the regionalized variable, it is more commonly used as a measure of uncertainty in the estimate than is the estimation variance. Universal kriging is a well established estimation method, and is essentially the only procedure that provides a standard error of the estimate for a spatial function. Appendix A presents a numerical example of standard error computation.

If universal kriging is repeatedly performed for regularly spaced discrete locations within the area of interest, the estimates can be passed to a graphical display package to create spatial representations of the regionalized variable. Possible forms of representation include block diagrams, cross-sections, or contour maps of either the variable itself or the standard error of the variable. Figure 2.5(a) is a graphical solution of the original problem in Figure 2.1, a continuous representation of a spatial function which is known only at discrete points. Locations A, B, and C from Figure 2.1 are also shown on Figure 2.5 to illustrate how universal kriging may be used to quantitatively define the probability density function mentioned in Part 2.1.2. Studies of several different regionalized variables have shown that the probability density functions for the discrepancies between actual values and those estimated by universal kriging closely follow a normal distribution, with slightly more pronounced mode and tails and a steeper slope between (Journel and Huijbregts, 1978, p. 48-51). For all practical purposes, a probability density function such as the one in Figure 2.2 can be regarded as a normal distribution with mean $Z^*(\vec{x}_0)$ and standard deviation $\sigma_E(\vec{x}_0)$.

By assuming normality, standard statistical tables can be used to compute confidence

intervals which bound the actual value of the regionalized variable within certain probability limits. For instance, the probability that the true value of a regionalized variable will be more than the estimate plus twice the standard error, or smaller than the estimate minus twice the standard error, is 4.5%. That is, from Figure 2.5, with a probability of 95.5%, the actual value of the regionalized variable at location C lies within the range 1419 +20.

For the remainder of this study, universal kriging will be accepted as a suitable estimator for spatial functions. The standard error will be used as an index to quantitatively compare the point accuracy of estimations made by universal kriging.

CHAPTER 3

THEORETICAL EXTENSIONS AND DEVELOPMENTS

Universal kriging provides a convenient way to estimate a spatial function and to simultaneously assess its reliability by determining the standard error of the estimates. Because of the definition of the estimator, universal kriging will result in an estimate having minimum standard error. This minimum is conditional upon the sample from which the estimation is made, as different samples will produce different standard errors of the same estimator and spatial function. The dependence of the standard error of kriging on sampling is discussed only in passing in a small number of publications. The specific problem addressed here is discussed only marginally in works devoted primarily to related but different problems (Quenouille, 1949; Matérn, 1960; Dalenius and others, 1961; Hannan, 1962; Deming, 1966; Payandeh, 1970; David, 1978; Delhomme, 1979; Ripley, 1981), or in situations where the mathematical assumptions are not valid for spatial functions (Madow and Madow, 1944; Madow, 1949; Barret, 1964; Cochran, 1964; King, 1969; Byth and Ripley, 1980; Diggle and Matérn, 1980). The potential of geostatistics in the optimal design of spatial sampling patterns has already been explored in the area of mining (Newton, 1973; Rendu, 1976), geohydrology (Olea, 1980) and soil science (McBratney and others, 1981; McBratney and Webster, 1981), but none of these studies are general in their analyses or assumptions.

This chapter presents original contributions to the optimal design of sampling patterns for spatial functions. The work is an extension of the geostatistical techniques of the previous chapter, used to find global indices measuring the performance of a spatial sample, and to disclose the nature of factors influencing the indices.

3.1 SAMPLING EFFICIENCY

3.1.1 Sampling efficiency indices

As shown in Equation 2.13, the magnitude of the point standard error is controlled by the individual values of each of the elements in matrices X and B. From Equation 2.9, matrix X is a function of matrices A and B. Consequently, the standard error is a function of matrices A and B.

From Equation 2.10, matrix A involves:

1. The semivariance of the spatial function.
2. The relative distance between every possible pair of sample elements.
3. Each of the terms of the analytical expression selected to represent the drift.
4. The location of sample elements.
5. The size of the sample subset used in the estimation.

In addition, from Equation 2.12, matrix B contributes:

6. The distance from the estimation location $\vec{x}_0$ to each of the sample locations $\vec{x}_j$.
7. The coordinates of the estimation location.

Note that the standard error does not depend on the sampled values themselves.

As explained in Part 2.2.3, the selection of models to represent the semivariance and the drift is somewhat arbitrary, but the freedom of choice is more apparent than real because only those models producing a good fit to reality are acceptable. In principle, the drift and semivariance are properties inherent in the spatial function. However, these properties may be represented by alternative models which describe them equally well.

For the purpose of this analysis, the relative arrangement of sample elements will be broken into two components, pattern and density. **Pattern** refers to the geometrical configuration of sample elements in space. An example is a square pattern, where the sample elements are in the center of a regular lattice of squares. Unfortunately, pattern is a nominal property possessing discrete categories which have no implicit ordering. **Density** is the degree of compactness of a sample, measured by the number of elements per space unit, such as the number of points per square mile. In some instances, pattern and density completely determine the

relative distance between sample elements, as in a square pattern with a density of one element per square foot (10.8 elements/m^2). In other instances, the nominal nature of pattern precludes a deterministic description. There may be an infinite number of alternative locations which fit the characteristics of the same pattern equally well, as, for example, in a random pattern. Numerous different configurations can be considered to be randomly arranged even though none may have a single point in the same location. In those circumstances where there is more than one possible configuration per pattern, a representative case has been selected for analysis.

Matrices A and B include terms containing both the coordinates of samples and the location of the estimation. However, the product is not a function of the arbitrarily chosen origin of the coordinate system, because the product of A^{-1} and B is invariant to changes in the origin (Journel and Huijbregts, 1978, p. 335).

There is an infinity of possible relative arrangements between the sample elements and the estimation location. Because the performance of a sample pattern cannot be judged meaningfully by the standard error at specific points, a better index of the efficiency of a sample is the average standard error over the sampling space. To express the spread around the average, a maximum and minimum standard error may also be of interest. The minimum standard error is trivial, as it is always zero at those points coinciding with sample locations. Therefore, only the maximum standard error will be considered in addition to the average standard error.

3.1.2 Sampling efficiency factors

The performance of samples of a spatial function will be judged by two indices, average standard error and the maximum standard error, which in turn depend on:

1. Unmanageable factors
   a. The semivariance
   b. The drift
2. Manageable factors
   a. The size of the sample subset considered by the estimate
   b. The sample pattern
   c. The sample density

The factors which influence the sampling efficiency indices are not alike; two have nothing to do with sampling and the third is only partially related to sampling. The semivariance and drift are inherent in the spatial function. The designer of the sampling program is limited to selecting models that provide better fits to the true semivariance and the drift. Specification of the size of the universal kriging matrix is primarily a problem in computational efficiency, with some sampling implications. This leaves pattern and density as the only two factors which offer wide flexibility in the design of the sampling scheme.

Interactions among factors is also an important concern. Independent variables are preferred because they can be analyzed one at a time. The semivariance depends upon the drift (Olea, 1982b, Appendix A), but all other factors are independent of one another. Since the drift relates to the spatial function itself and not to sampling, the drift is independent of the sample subset size, density, and pattern. Figure 3.1 illustrates the independence between the three factors related to the sampling process. In part (a), the number of sample elements considered in an estimation has been varied without modifying density or pattern. In part (b), only the density has been changed, and in part (c), the density as well as the subset size remains the same but the pattern has been altered.

The remainder of this chapter presents a sensitivity analysis of the effects that changes in sampling efficiency factors have on the sampling efficiency indices.

3.2 SEMIVARIANCE AND DRIFT

3.2.1 Modeling

The joint study of the semivariance of the residuals and the drift of a spatial function is the subject of **structural analysis,** a standard procedure in geostatistics.

The residuals are the difference between the spatial function and a smooth analytical function on the orthogonal geographic coordinates x and y. In this study, consideration will be limited to polynomials up to the second degree, having the form:

$$M(x,y) = a_0 + a_1x + a_2y + a_3x^2 + a_4xy + a_5y^2 \qquad (3.1)$$

A semivariogram will tend to be more irregular when computed from noisy and less regular observations. The practice is to fit idealized curves to the semivariograms computed from the data and use these ideal curves in the universal kriging system of equations rather than the computed semivariograms themselves (Henley, 1981, p. 19-20). The analysis will be restricted to linear models of the form

$$\gamma(h) = \omega h \qquad (3.2)$$

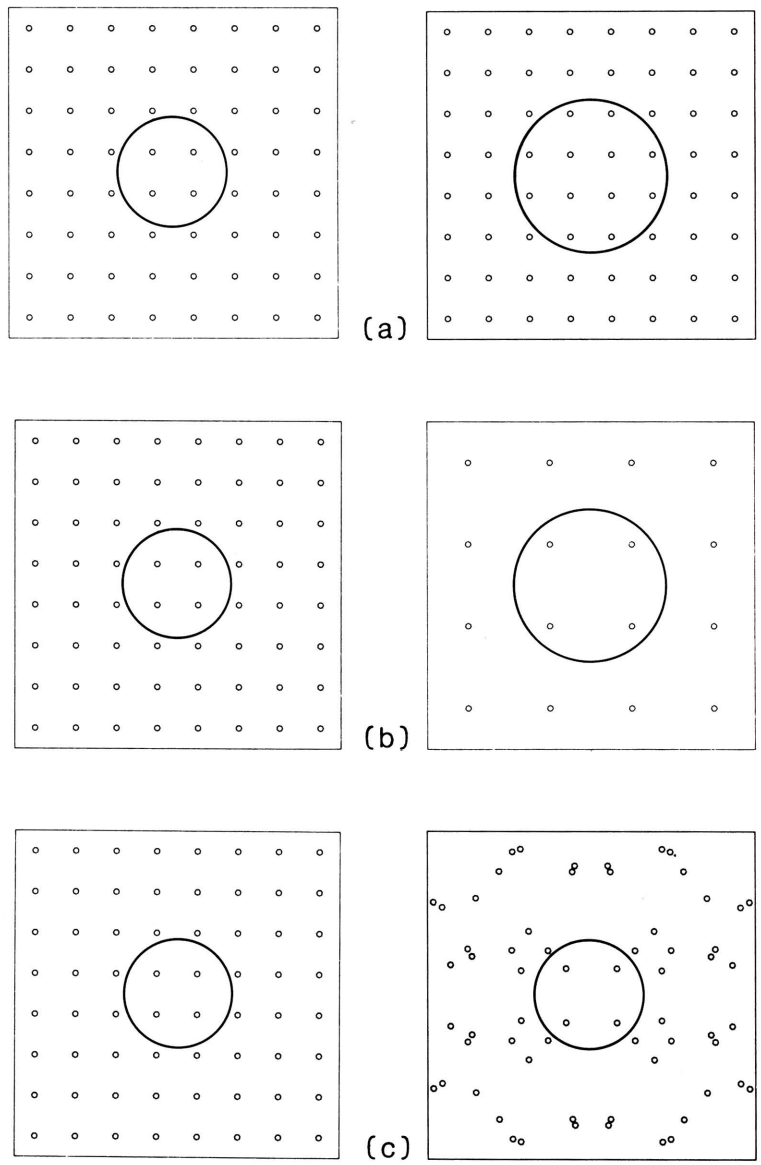


FIGURE 3.1. Independent sampling efficiency factors. Large circle encloses the actual number of sample elements considered in the estimation. Either (a) actual number of sample elements considered in the estimation, (b) density, or (c) pattern can be changed while holding the other two factors constant.

3.2.2 Linear semivariance

This part presents general characteristics of the standard error of regionalized variables whose semivariance of the residuals is linear.

Lemma 3.1

Let the semivariance of the residuals be isotropic and linear. Let λ_j and μ_i be the weights and slack variables in the universal kriging system of equations. Then λ_j and (μ_i/ω) are independent of the slope ω of the semivariogram.

Proof:

As the semivariance is linear and isotropic

$$\gamma(\vec{h}) = \omega\left|\vec{x}_p - \vec{x}_j\right| = \omega h_{pj} \qquad (3.3)$$

where the bars represent the modulus of the vector.

The universal kriging system of equations in Equation 2.9 for this particular case becomes

$$A'X' = B' \qquad (3.4)$$

where neither A' nor B' have terms depending on ω. Then, X' is independent of the value of ω. The matrices are shown on the following page as Equations 3.5 through 3.7.

Theorem 3.1

Let a regionalized variable have a semivariance of the residuals which is isotropic and linear. Then, the estimation variance is a linear function of the slope ω of the semivariance.

Proof:

From Equations 2.13 and 3.3

$$\sigma_E^2(\vec{x}_0) = \sum_{i=1}^{k} \lambda_j \omega h_{0j} + \mu_0 + \sum_{i=1}^{n} \mu_i f^i(\vec{x}_0) \qquad (3.8)$$

The equation is not changed by multiplying and dividing each μ_i term by ω.

$$\sigma_E^2(\vec{x}_0) = \sum_{i=1}^{k} \lambda_j \omega h_{0j} + \omega\{\mu_0/\omega\} + \omega \sum_{i=1}^{n} \{\mu_i/\omega\} f^i(\vec{x}_0) \qquad (3.9)$$

Factoring by ω

$$\sigma_E^2(\vec{x}_0) = \omega\left[\sum_{i=1}^{k} \lambda_j h_{0j} + \{\mu_0/\omega\} + \sum_{i=1}^{n} \{\mu_i/\omega\} f^i(\vec{x}_0)\right] \qquad (3.10)$$

From Lemma 3.1, neither λ_j, $\{\mu_0/\omega\}$, nor $\{\mu_i/\omega\}$ are functions of ω. As h_{0j} and $f^i(\vec{x}_0)$ are likewise not functions of ω, then the semivariance depends directly on the term ω.

Q.E.D.

We may consider two spatial functions whose factors are the same except that one has a linear semivariance with a slope k times larger than the other. Theorem 3.1 indicates that the spatial function with the steepest semivariance will have standard errors which are $\sqrt{k}$ larger than the other.

Another important corollary is that, although the standard error depends on the linear semivariance through the terms h_{pj}, the effect of the slope can be analyzed independently from all other factors. This means that a general study can be conducted which is valid for all n-dimensional spatial functions that can be represented by a linear model with unit slope for the semivariance of the residuals. The resulting standard error is constant for all such functions except for a factor equal to the square root of ω.

3.3 PATTERN

3.3.1 Pattern selection

Pattern is a nominal property that refers to different configurations of objects, which in this case are the sample elements. A pattern involves distances among the elements in the set. In this study patterns are arranged in two dimensions. Some patterns have unique characterizations; for example, there is one and only one way to arrange elements in a square pattern. All square patterns are the result of scaling, rotation, or translation of one basic type. In contrast, a random pattern has only a statistical definition, as there are an infinite number of possible configurations that fit the description of randomness.

Utility computer programs were prepared to generate sample element patterns covering the entire range of possible patterns (Olea, 1982a, Appendix B). From these, 14 patterns grouped into 7 categories were selected.

1. Regular (Fig. 3.2)
   a. Hexagonal
   b. Square
   c. Triangular
2. Orthogonal regular traverses (Fig. 3.3)
   a. Intersection every 2 points
   b. Intersection every 8 points
3. Stratified (Fig. 3.4)
   a. Hexagonal
   b. Square
4. Random (Fig. 3.5)
5. Bisymmetrical (Fig. 3.6)
   a. Random
   b. Regular clusters
6. Clustered (Fig. 3.7)
   a. One cluster
   b. Five clusters
7. Regular clusters (Fig. 3.8)
   a. Sixteen points per cluster
   b. Four points per cluster

Categories 1 and 2 are uniquely defined. The remaining patterns are representative examples from among an infinite set of possibilities.

$$
A' = \begin{bmatrix}
0 & h_{12} & \cdots & h_{1k} & 1 & f^1(\vec{x}_1) & f^2(\vec{x}_1) & \cdots & f^n(\vec{x}_1) \\
h_{21} & 0 & \cdots & h_{2k} & 1 & f^2(\vec{x}_1) & f^2(\vec{x}_2) & \cdots & f^n(\vec{x}_2) \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
h_{j1} & h_{j2} & \cdots & h_{jk} & 1 & f^1(\vec{x}_j) & f^2(\vec{x}_j) & \cdots & f^n(\vec{x}_j) \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
h_{k1} & h_{k2} & \cdots & 0 & 1 & f^1(\vec{x}_k) & f^2(\vec{x}_k) & \cdots & f^n(\vec{x}_p) \\
1 & 1 & \cdots & 1 & 0 & 0 & 0 \cdots & 0 & \\
f^1(\vec{x}_1) & f^1(\vec{x}_2) & \cdots & f^1(\vec{x}_k) & 0 & 0 & 0 \cdots & 0 & \\
f^2(\vec{x}_1) & f^2(\vec{x}_2) & \cdots & f^2(\vec{x}_k) & 0 & 0 & 0 \cdots & 0 & \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
f^i(\vec{x}_1) & f^i(\vec{x}_2) & \cdots & f^i(\vec{x}_k) & 0 & 0 & 0 \cdots & 0 & \\
\cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots & \cdots \\
f^n(\vec{x}_1) & f^n(\vec{x}_2) & \cdots & f^n(\vec{x}_k) & 0 & 0 & 0 \cdots & 0 &
\end{bmatrix} \qquad (3.5)
$$

$$
X' = \begin{bmatrix}
\lambda_1 \\ \lambda_2 \\ \cdots \\ \lambda_j \\ \cdots \\ \lambda_k \\ \mu_0/\omega \\ \mu_1/\omega \\ \mu_2/\omega \\ \cdots \\ \mu_i/\omega \\ \cdots \\ \mu_n/\omega
\end{bmatrix} \qquad (3.6) \qquad
B' = \begin{bmatrix}
h_{01} \\ h_{02} \\ \cdots \\ h_j \\ \cdots \\ h_k \\ 1 \\ f^1(\vec{x}_0) \\ f^2(\vec{x}_0) \\ \cdots \\ f^i(\vec{x}_0) \\ \cdots \\ f^n(\vec{x}_0)
\end{bmatrix} \qquad (3.7)
$$

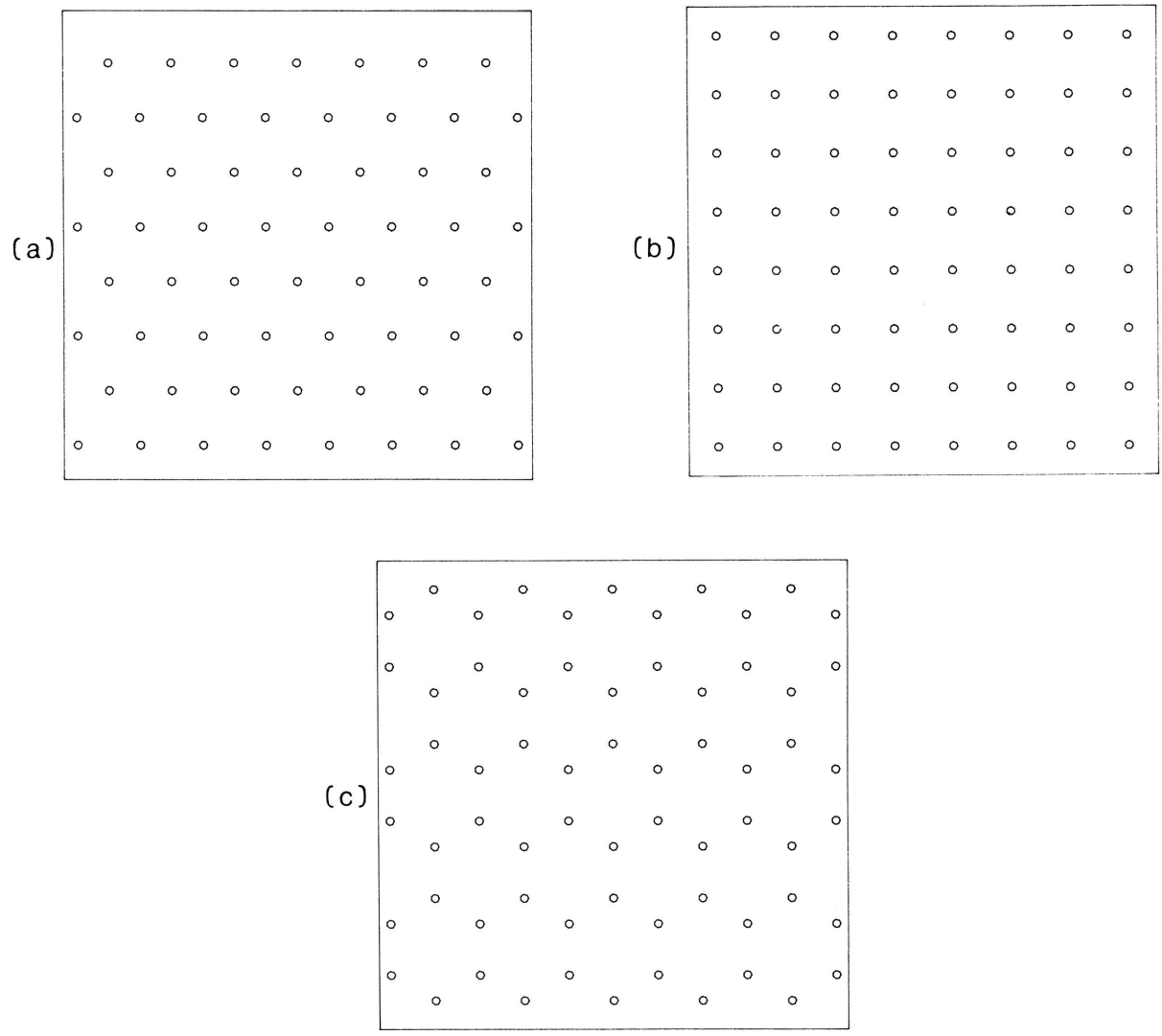


FIGURE 3.2. Regular patterns. (a) Hexagonal. (b) Square. (c) Triangular.

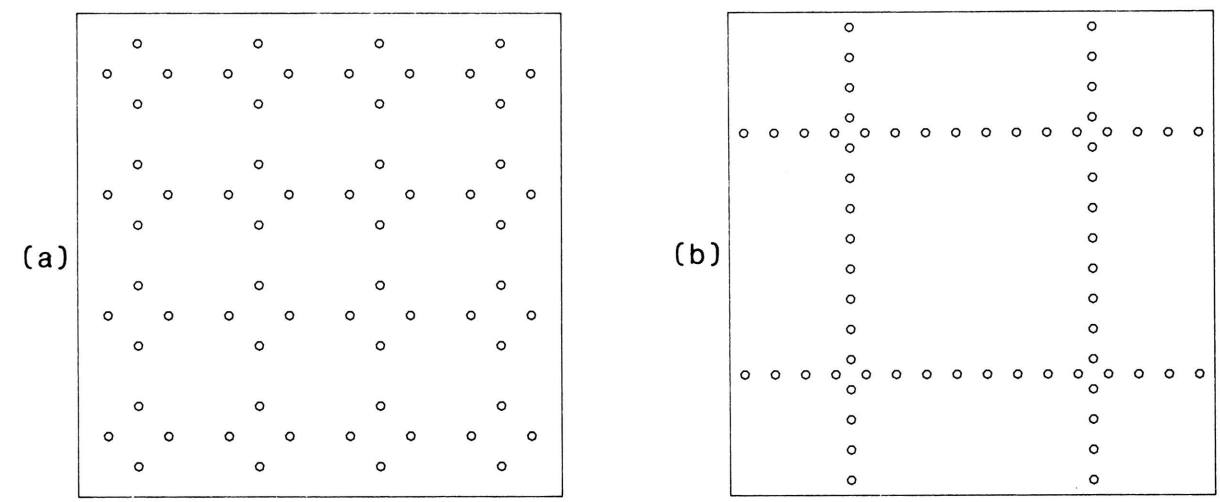


FIGURE 3.3. Orthogonal regular traverses. Intersection every (a) two points, (b) eight points.

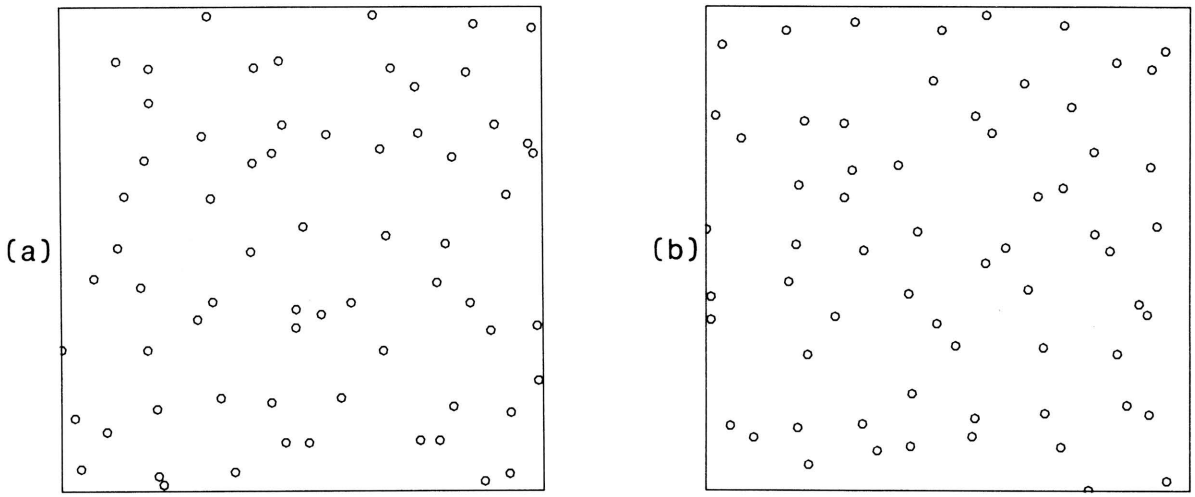


FIGURE 3.4. Stratified patterns. (a) Hexagonal. (b) Square.

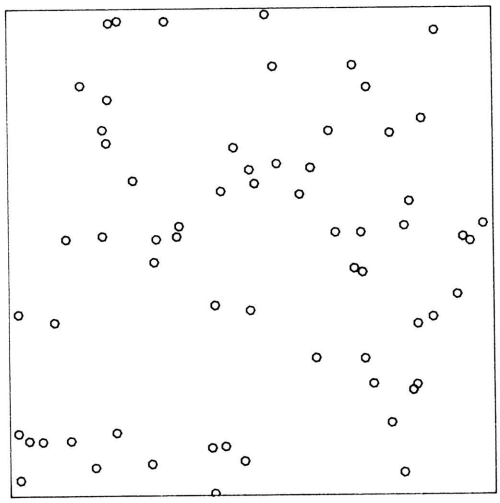


FIGURE 3.5. Random pattern.

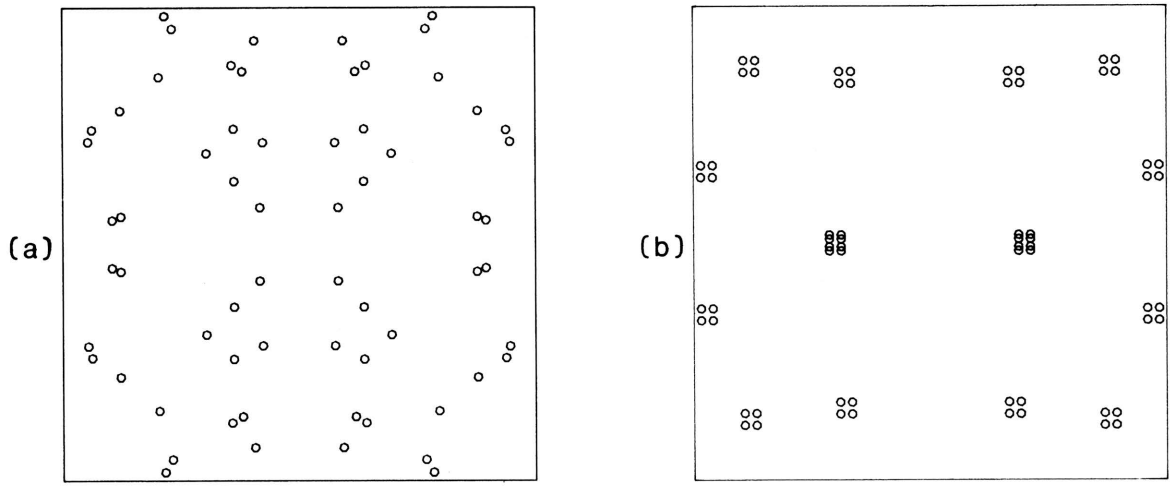


FIGURE 3.6. Bisymmetrical patterns. (a) Random. (b) Regular clusters.

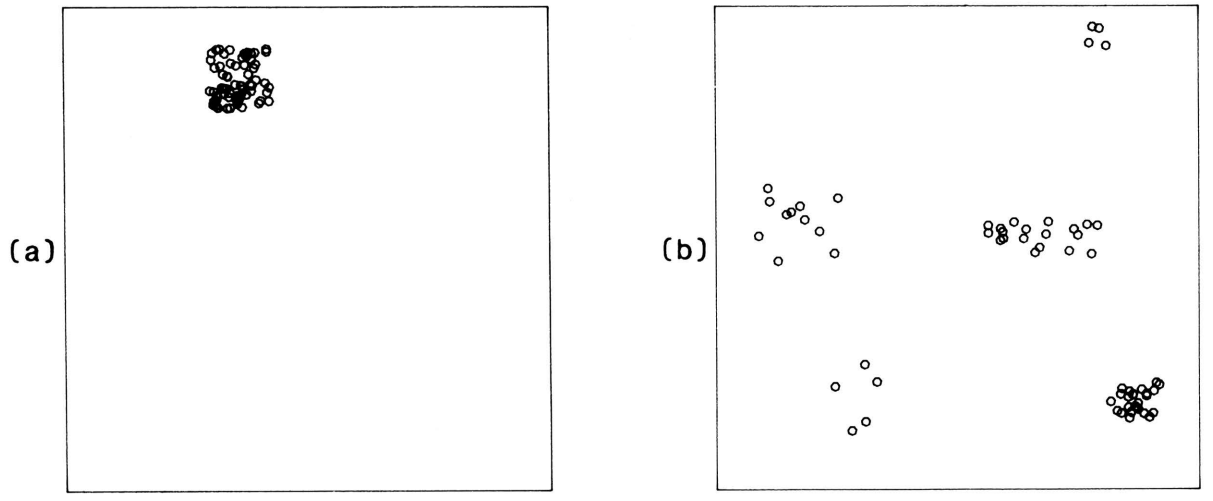


FIGURE 3.7. Clustered patterns. (a) One cluster. (b) Five clusters.

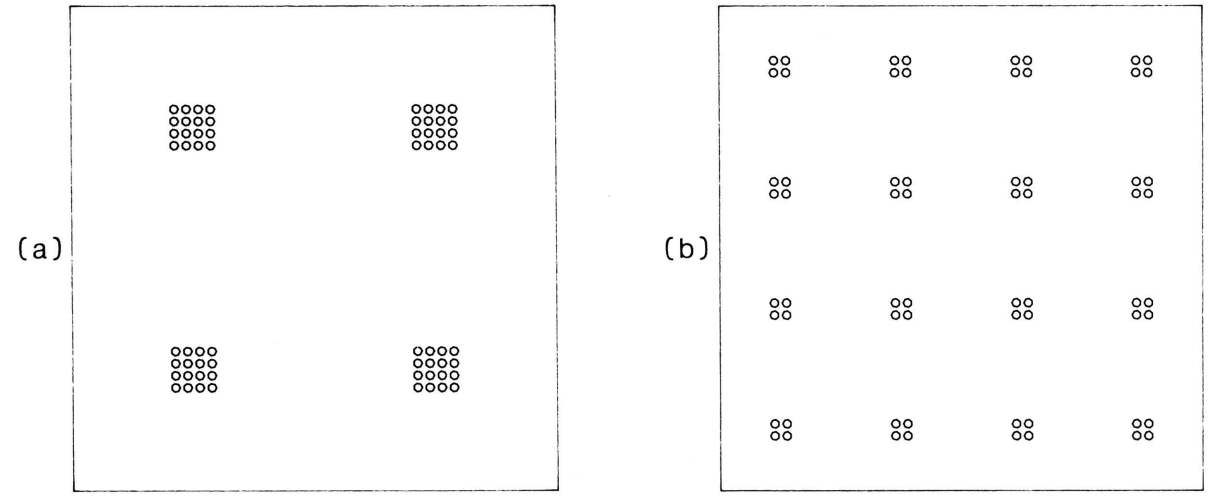


FIGURE 3.8. Regular clusters. (a) Sixteen points per cluster. (b) Four points per cluster.

Historically, the same regular pattern has been called by different names by various authors. The convention used here is based upon the idea that each point in a pattern can be considered to be a centroid of a Voronoi division. That is, each point can be surrounded by a mosaic of regular polygons (Rhynsburger, 1973), as in Figure 3.9. A hexagonal pattern, for instance, is the arrangement of the centroids at a set of regular hexagons (Abler and others, 1971, p. 105). Alternatively, patterns can be classified by joining each point to its nearest neighbors, as in Figure 3.10. The resulting polygons are the same only for the square pattern, but Voronoi hexagons become equilateral triangles and vice-versa. For this reason, what is called the hexagonal pattern in this study is also referred to by other authors as a face-centered hexagon (Cole and King, 1968, p. 178) or as a triangular network (Matérn, 1960, p. 73-74). This study will use a terminology based on Voronoi polygons, as these are more appropriate when density of pattern must be considered. It can be proven that the three patterns in Figure 3.9 are the only regular polygons which can be formed in a two-dimensional space (Matérn, 1960, p. 74).

Stratified sampling is a selection procedure in which the sample space is divided into mutually exclusive partitions and an element is then randomly taken from every partition (Ripley, 1981, p. 19-22). Figure 3.11 illustrates stratified sampling of a two-dimensional space partitioned into squares and into hexagons. The randomly selected points are the same as in Figure 3.4.

Clustered patterns offer the richest variety for sampling schemes; there are infinite

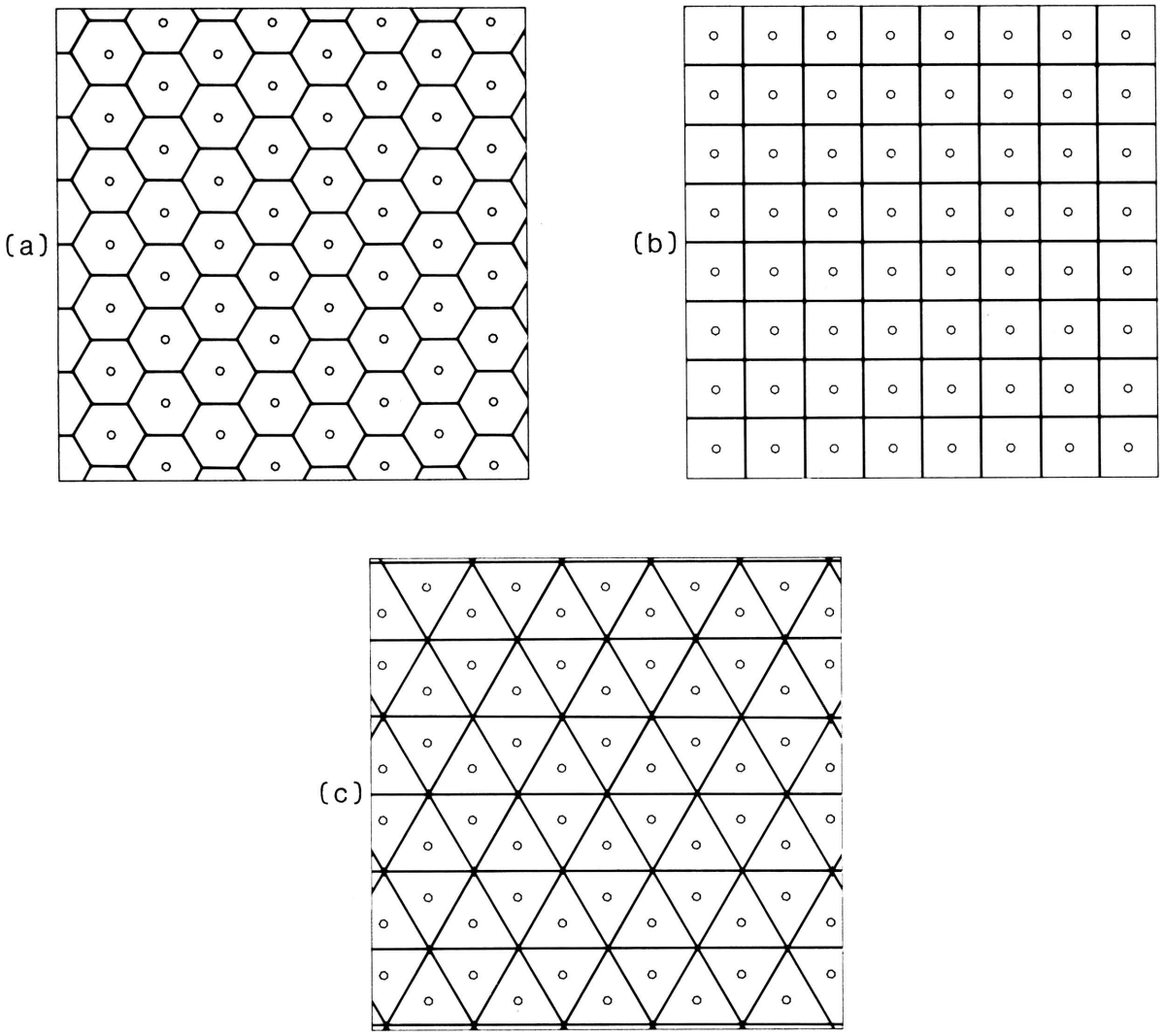


FIGURE 3.9. Voronoi polygons and regular patterns. Names given to patterns result from partitioning the plane into Voronoi polygons. (a) Hexagonal. (b) Square. (c) Triangular.

ways to vary the number of clusters, the relative positions of clusters, and the density of points inside the clusters. However, only a small number of these possibilities were investigated because, as will be shown later, clustered patterns do not provide acceptable solutions to the spatial sampling problem.

Geographers have been exceptionally active in the testing, modeling, and characterization of patterns. Clark and Evans (1954) introduced the distance from a point to its nearest neighbor as a measure of randomness. Let us consider a sample of n points. Each point will have a nearest neighbor which is some distance r_i away. The average distance to the nearest neighbor, $\bar{r}$, is defined as

$$\bar{r} = \frac{1}{n} \sum_{i=1}^{n} r_i \qquad (3.11)$$

Clark and Evans (1954) proved that the expected average distance to the nearest neighbor in a random pattern is

$$\bar{r}_E = \frac{1}{2\sqrt{\rho}} \qquad (3.12)$$

where ρ is the density of the pattern expressed as the number of sample elements per unit of area.

The distance index R is the observed average distance to the nearest neighbor divided by the value corresponding to a random pattern:

$$R = \bar{r} / \bar{r}_E \qquad (3.13)$$

One parameter is not sufficient to characterize a pattern, as configurations which

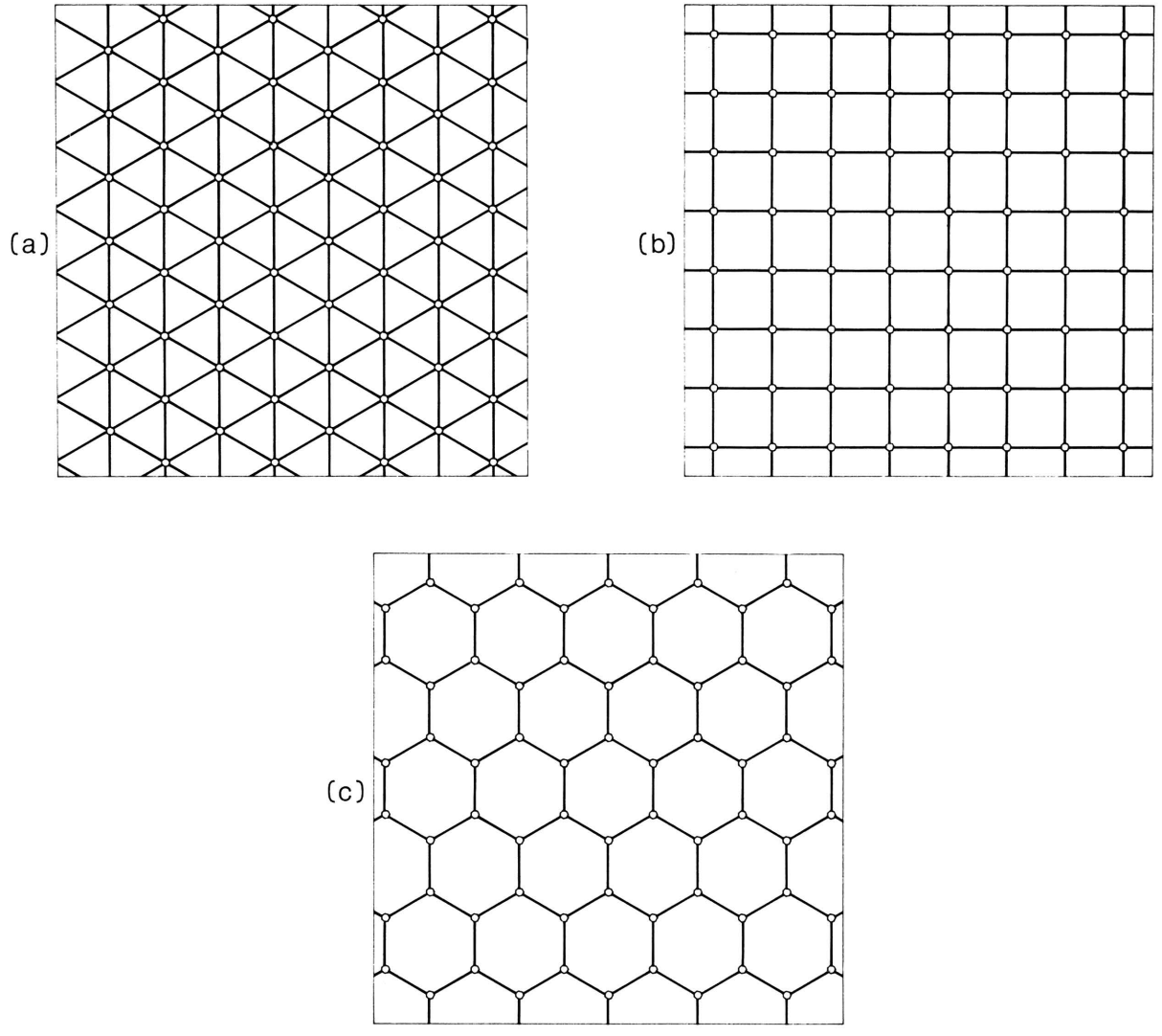


FIGURE 3.10. Nearest neighbors and regular patterns. (a) Voronoi hexagons in Figure 3.9 become equilateral triangles. (c) Squares remain as squares. (c) Voronoi triangles become hexagons.

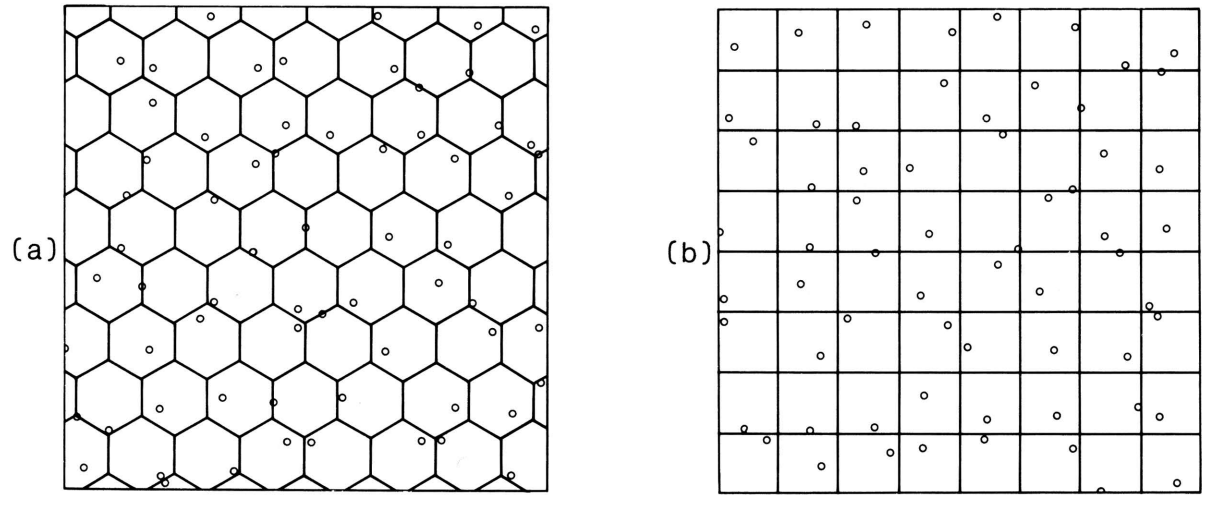


FIGURE 3.11. Sampling mechanism for stratified patterns. (a) One point is selected randomly from inside each hexagon. (b) One point is selected randomly from inside each square.

have completely different appearances may have the same distance index. Also, it seems intuitively reasonable that a two-dimensional pattern would require at least two parameters for its characterization. A second parameter was therefore created to express the entropy of a pattern. The maximum number of symmetry axes through any point in the plane was selected as a simple descriptor of order in the system. The distance index and maximum number of symmetry axes were calculated for all 14 patterns in Figures 3.2 to 3.8 using a computer program (Olea, 1982a, Appendix B). A graphical display of the results is given in Figure 3.12. Two patterns, one stratified and the other a regular cluster, are not included because they overlap in the diagram. The spread and spacing of the 14 patterns indicate that those selected for analysis cover the spectrum of possibilities.

3.3.2 Sensitivity to pattern

A computer program (Olea, 1972, Appendix H) was used to calculate the average and maximum standard error for each of the 14 patterns, assuming that the patterns extend to infinity with the same relative distance characteristics. The units of standard error are the same as the units for the corresponding spatial function. Since the units of the spatial function vary with the nature of the phenomenon, so do the units of standard error. The absence of units for the standard error in the remainder of this chapter is intended to emphasize the generality of the methodology, rather than to imply that the standard error is a dimensionless variable. In order to avoid border effects, the analysis considers only the central part of each window shown in Figures 3.2 to 3.8. The actual size of the window is 8 by 8 miles (12.9 km) and the point density is held constant at one point per square mile (0.39 points/km^2). The number of sample elements considered in the universal kriging system of equations was set to 32. The semivariance was assumed to be linear with unit slope. Table 3.1 is a condensation of the results. The average and maximum standard error for the particular instance of a drift of degree 1 is shown graphically in Figure 3.13.

From the analysis we can see that spatial sampling which is systematically ordered is superior. In general, the greater the distance index and the maximum number of symmetry axes through a point, the lower will be the standard error. A hexagonal sampling pattern produces both the lowest average and maximum standard errors. This performance is almost matched by the square pattern. The worst sampling schemes are those in which any form of clustering is present. For the same sample size, the standard error associated with clustered patterns may be several orders of magnitude greater than the standard error obtained using regular patterns.

If all other factors remain the same, a regionalized variable having a complex drift will have a larger standard error than a regionalized variable with a simple drift. However, the nature of the drift does not affect the standard errors of the eight most efficient patterns.

3.4 NUMBER OF NEAREST NEIGHBORS

3.4.1 Sample subset size

The complete set of data used in an actual study may consist of several thousand observations, and distances between the most distant pairs of points may be hundreds of miles. Points at such a great distance are seldom closely related. In addition, there are considerations of computational efficiency which suggest that the entire sample set should not be used in each estimation. Universal kriging requires the solution of a system of equations of order $(n + m)$ by $(n + m)$ where n is the number of sample elements and m is the number of unknown coefficients in the analytical expression for the drift. In a large sample, the n terms become dominant and the order of the coefficient matrix changes with the square of the number of elements in the sample. This has a pronounced effect on computation time and the cost of estimation. Fortunately, the nature of the estimation method is such that only the closest sample elements effectively contribute to the results, because of what is called the **screen effect** (Olea, 1975, Appendix G). Additionally, when a drift is present, the models selected to represent the semivariance and drift may not be valid over neighborhoods which are hundreds of miles in diameter. For these reasons, it is common practice to use only a subset of points, consisting of those closest to the estimation location. In other words, estimates are based on the nearest neighbors of the location being evaluated.

In some circumstances, an insufficient number of sample elements may be available within the neighborhood for which the structural analysis models are valid. Then, the number of nearest neighbors that are closest to the estimation location becomes a sampling concern; if there are not enough sample elements, the estimates may be unreliable or there may be blank areas in the final contour map.

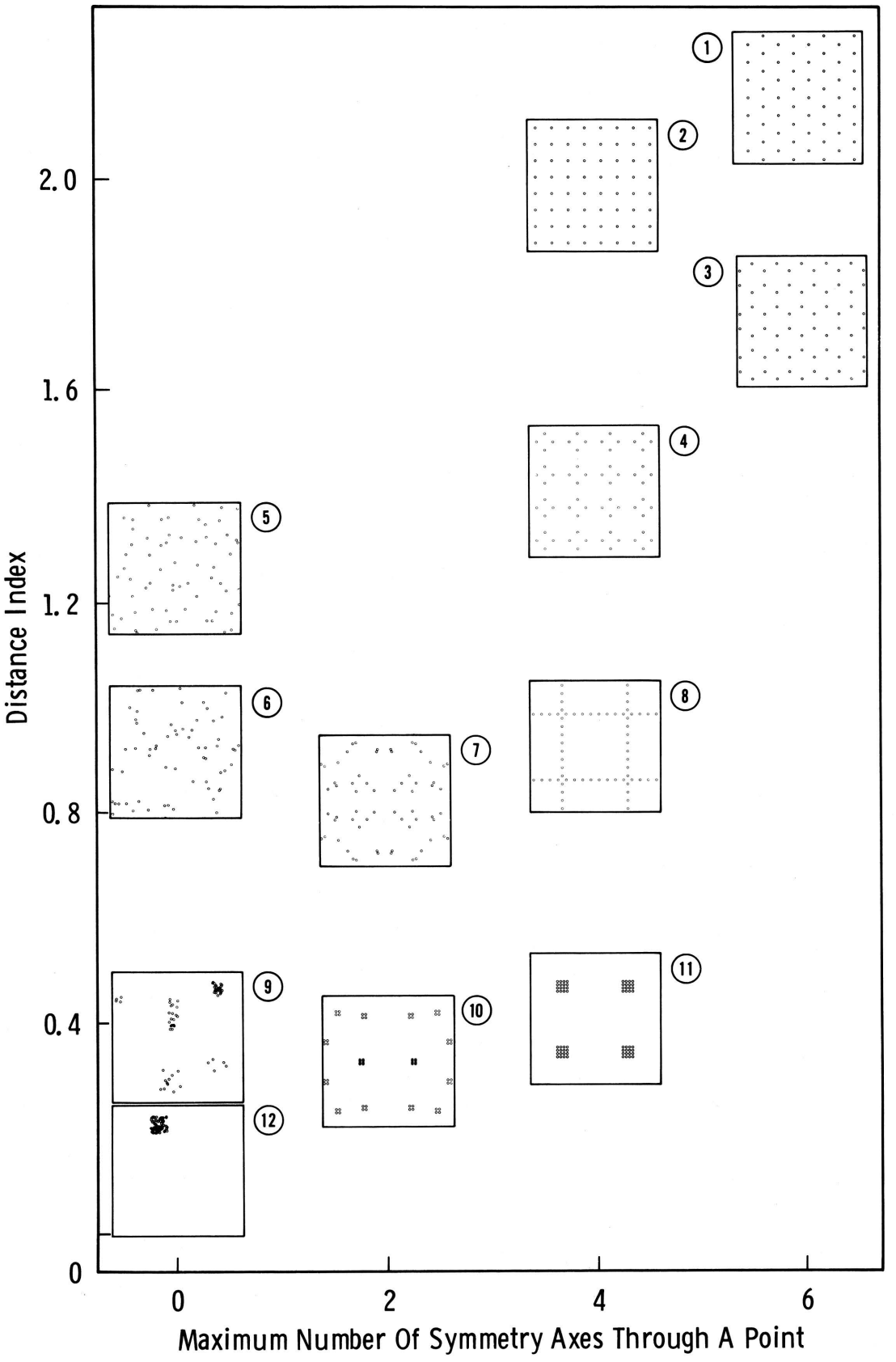


FIGURE 3.12. Two-dimensional point patterns at a fixed density. (1) Hexagonal. (2) Square. (3) Triangular. (4) and (8) Orthogonal traverses. (5) Stratified hexagonal. (6) Random. (7) and (10) Bisymmetric. (9) and (12) Clustered. (11) Regular cluster.

TABLE 3.1

SENSITIVITY OF THE SAMPLING EFFICIENCY INDICES TO CHANGES IN PATTERN. UNIT DENSITY, UNIT LINEAR SEMIVARIOGRAM SLOPE AND 32 NEAREST NEIGHBORS
(Number in parentheses corresponds to pattern identification number in Figure 3.12.)

Pattern	Distance Index	Symmetry Axes	Average Std. Error Drift 0	Drift 1	Drift 2	Maximum Std. Error Drift 0	Drift 1	Drift 2
Hexagonal (1)	2.15	6	0.63	0.63	0.63	0.72	0.72	0.72
Square (2)	2.00	4	0.64	0.64	0.64	0.74	0.74	0.74
Triangular (3)	1.75	6	0.66	0.66	0.66	0.80	0.80	0.80
Traverses every 2 points (4)	1.41	4	0.68	0.68	0.68	0.89	0.89	0.89
Hexagonal stratification (5)	1.26	-	0.69	0.69	0.69	0.86	0.86	0.86
Square stratification	1.28	-	0.69	0.69	0.69	0.91	0.91	0.91
Random (6)	0.91	-	0.71	0.71	0.71	1.05	1.05	1.05
Bisymmetrical random (7)	0.82	2	0.72	0.72	0.72	0.98	0.98	0.98
Traverses every 8 points (8)	0.92	4	0.81	0.81	0.84	1.23	1.23	1.45
Four points per regular cluster	0.40	4	0.83	0.83	0.84	0.99	0.99	1.00
Five clusters (9)	0.34	-	0.98	0.99	1.06	1.33	1.49	2.33
Bisymmetrical clusters (10)	0.33	2	1.03	1.03	1.11	1.22	1.22	1.38
Sixteen points per regular cluster (11)	0.40	4	1.13	1.17	1.53	1.51	1.85	5.13
One cluster (12)	0.13	-	2.19	5.01	61.50	2.94	8.27	148.00

3.4.2. Sensitivity to number of nearest neighbors

The minimum number of elements required to solve the universal kriging system of equations is equal to the number of unknown coefficients in the analytical expression of the drift. A sensitivity analysis was performed between the limits of the minimum number and a maximum of 32 elements, an interval long enough to study all patterns of practical interest. The 14 patterns selected in Section 3.3 were analyzed using the assumptions that the semivariance is linear with slope 1 and the density is one point per square mile (0.39 points/km^2). The sensitivity analysis shows that the screen effect depends upon drift and pattern. Table 3.2 presents critical values beyond which no improvement is made in sampling efficiency indices by adding extra sample elements. An entry of value 32+ indicates that the critical value is larger than the maximum tested value of 32. In general, there is a multiplicative effect on

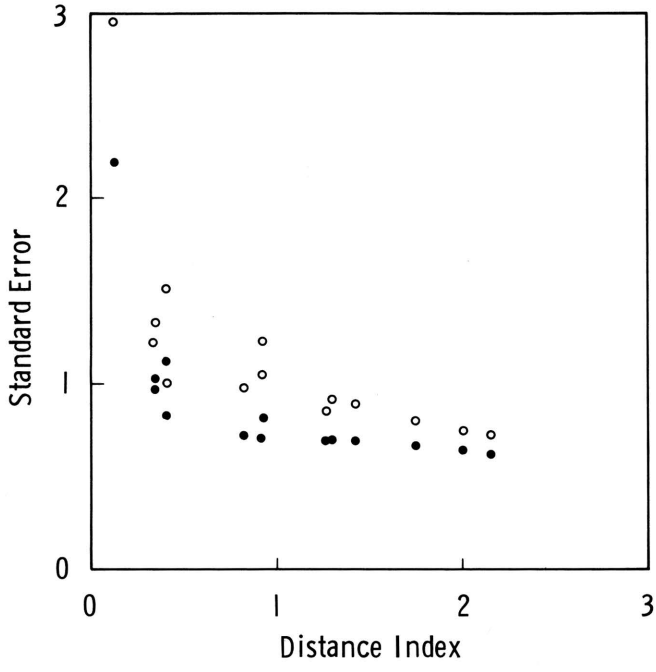


FIGURE 3.13. Standard error for the 14 basic patterns. Open circles indicate maximum standard error and solid circles average standard error. Results are valid for a linear semivariance with slope 1 and a first-degree polynomial drift, a density of 1 point per square mile, and solution of the universal kriging system of equations using 32 sample elements closest to the estimated location.

efficiency of a pattern; the smaller the limiting index in Table 3.1, the smaller will be the number of nearest neighbors required to achieve the full screen effect. To reach the limiting average standard error, while all other factors remain the same, a spatial function having a second-degree polynomial drift will require about twice as many nearest neighbors as a spatial function with a first-degree drift. If either first-degree or second-degree polynomial drifts are acceptable, the greater number of nearest neighbors required for the second-degree drift may suggest that the simpler drift is the more practical choice.

Ideally, the minimum number of sample elements used to solve the universal kriging system of equations is equal to the number of unknown coefficients in the drift model. If the drift model is a polynomial, the ideal minimum is 1 for a constant, 3 for a first-degree drift, and 6 for a second-degree drift. In practice, unstable systems of equations will result if the ideal minimum is used. Rather, a larger number of nearest neighbors is required, as listed for each of the 14 patterns in Table 3.3. The increase is particularly noticeable for second-degree models, but does not occur when the drift is a constant.

Figures 3.14-3.28 show changes in standard error for five selected patterns: hexagonal, orthogonal regular traverses intersecting every two points, hexagonal stratification, random, and five clusters. Although values

TABLE 3.2

MINIMUM NUMBER OF SAMPLE ELEMENTS REQUIRED TO OBTAIN FULL SCREEN EFFECT AND ASSURE MINIMUM VALUES FOR THE SAMPLING EFFICIENCY INDICES

(Number in parentheses corresponds to pattern identification number in Figure 3.12.)

Pattern	Average Std. Error Drift 0	Average Std. Error Drift 1	Average Std. Error Drift 2	Maximum Std. Error Drift 0	Maximum Std. Error Drift 1	Maximum Std. Error Drift 2
Hexagonal (1)	3	5	10	3	5	10
Square (2)	4	5	10	10	10	14
Triangular (3)	8	8	11	11	12	12
Traverses every two points (4)	12	12	20	14	15	26
Hexagonal stratification (5)	6	6	16	7	8	25
Square stratification	8	8	18	10	11	30
Random (6)	12	12	32+	12	14	32+
Bisymmetrical random (7)	8	8	32+	20	28	32+
Traverses every 8 points (8)	28	32+	32+	28	32+	32+
Four points per regular cluster	12	12	32+	22	28	32+
Five clusters (9)	32+	32+	32+	32+	32+	32+
Bisymmetrical clusters (10)	20	20	32+	26	28	32+
Sixteen points per regular cluster (11)	32+	32+	32+	32+	32+	32+
One cluster (12)	32+	32+	32+	32+	32+	32+

TABLE 3.3

MINIMUM NUMBER OF NEAREST NEIGHBORS REQUIRED TO SOLVE THE UNIVERSAL KRIGING SYSTEM OF EQUATIONS

(Number in parentheses corresponds to pattern identification number in Figure 3.12)

Pattern	Drift 0	1	2
Hexagonal (1)	1	3	6
Square (2)	1	3	7
Triangular (3)	1	3	7
Profiles every two points (4)	1	3	9
Hexagonal stratification (5)	1	3	6
Square stratification	1	3	6
Random (6)	1	3	6
Bisymmetrical random (7)	1	3	9
Profiles every eight points (8)	1	5	9
Sixteen regular clusters	1	3	9
Five cluster (9)	1	3	6
Bisymmetrical clusters (10)	1	5	12
Four regular clusters (11)	1	5	9
One cluster (12)	1	3	7

on the abscissa are discrete, the graphs are shown as continuous lines to facilitate interpretation. Appendix B gives results in tabular form for all classes analyzed.

The hexagonal pattern manifests the screen effect to the greatest extent possible. Because of this, it reaches the point of no additional return for additional samples sooner than in any other configuration.

Many patterns produce unstable systems of equations in the presence of a drift when only the minimum number of nearest neighbors is considered. In Figure 3.19, for instance, no values can be estimated if only six, seven, or eight nearest neighbors are considered. In general, larger minimum numbers of nearest neighbors are required to solve the system of equations when the drift is complex and the distance index is low.

In summary, a sensitivity analysis on the number of nearest neighbors shows that significant savings in computer processing time required to solve the universal kriging system of equations can be achieved. No spatial pattern is more efficient than a hexagonal pattern, because if all other factors are the same, the hexagonal pattern requires the fewest number of nearest neighbors to calculate the best possible efficiency indices.

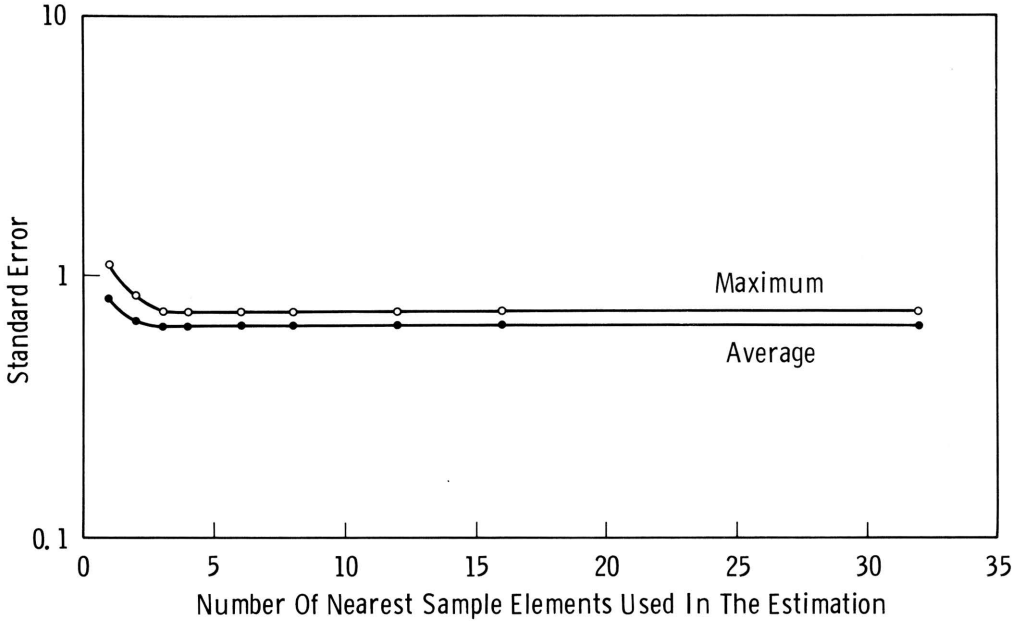


FIGURE 3.14. Sensitivity analysis of the number of nearest neighbors for a hexagonal pattern and no drift. Sampling density is one point per square mile and the semivariance is linear with slope 1.

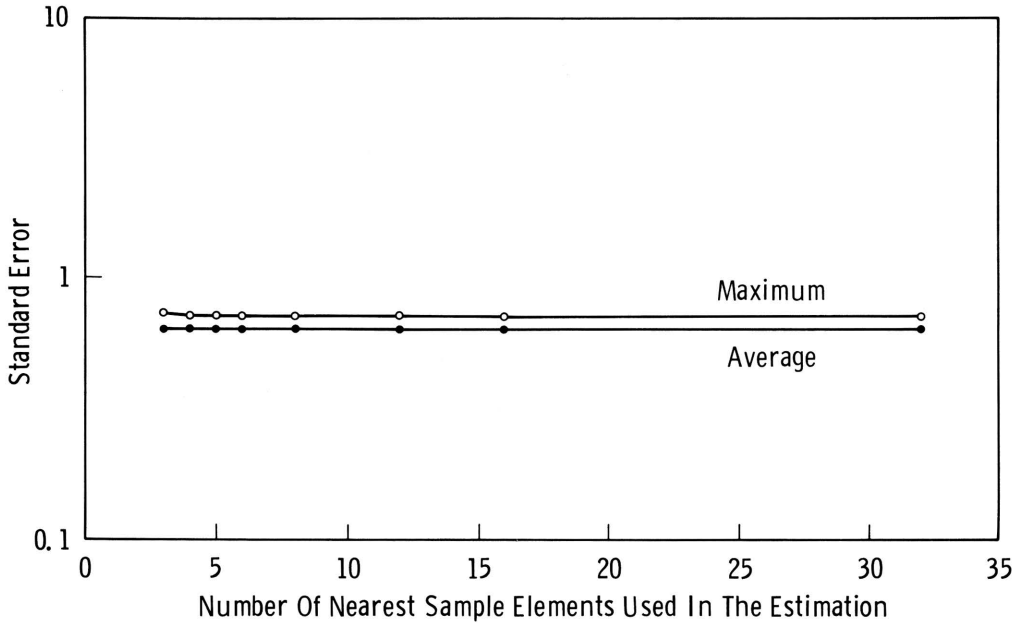


FIGURE 3.15. Sensitivity analysis of the number of nearest neighbors for a hexagonal pattern and a first-degree polynomial drift. Sampling density is one point per square mile and the semivariance is linear with slope 1.

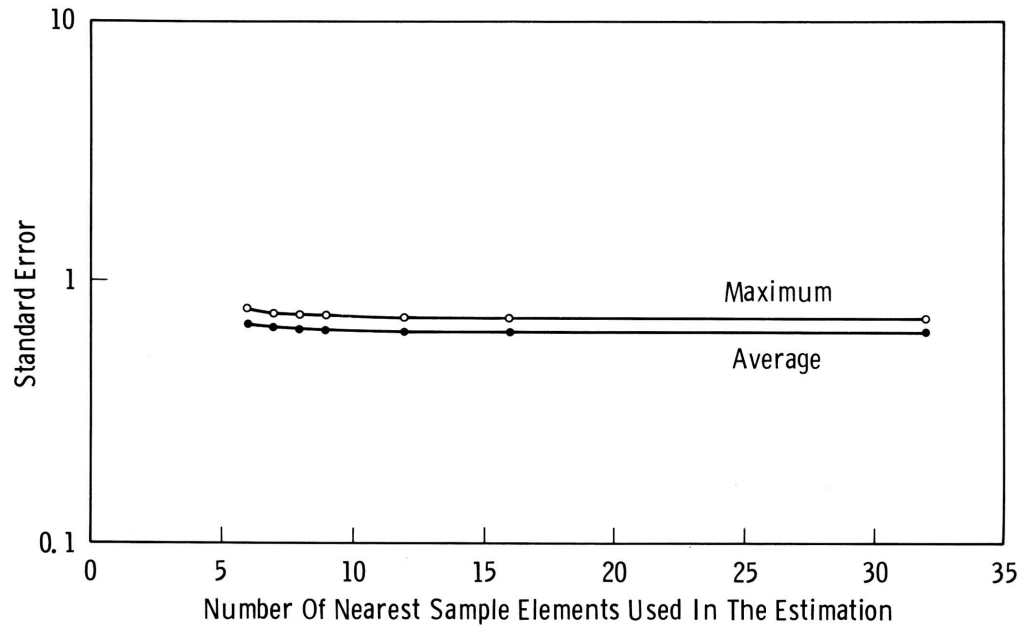


FIGURE 3.16. Sensitivity analysis of the number of nearest neighbors for a hexagonal pattern and a second-degree polynomial drift. The sampling density is one point per square mile and the semivariance is linear with slope 1.

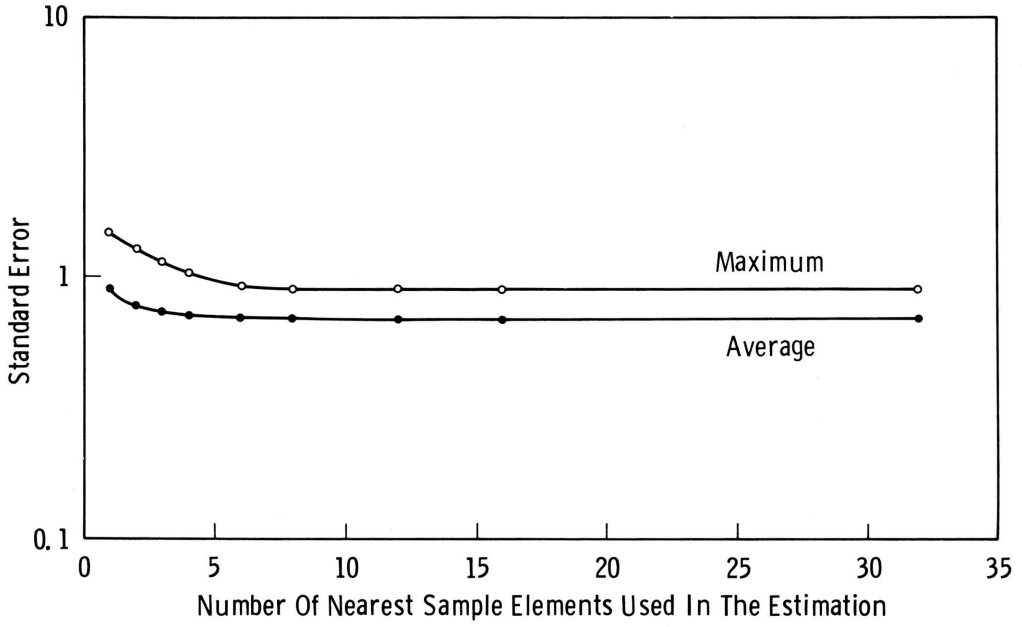


FIGURE 3.17. Sensitivity analysis of the number of nearest neighbors for orthogonal regular traverses intersecting every two points and no drift. The sampling density is one point per square mile and the semivariance is linear with slope 1.

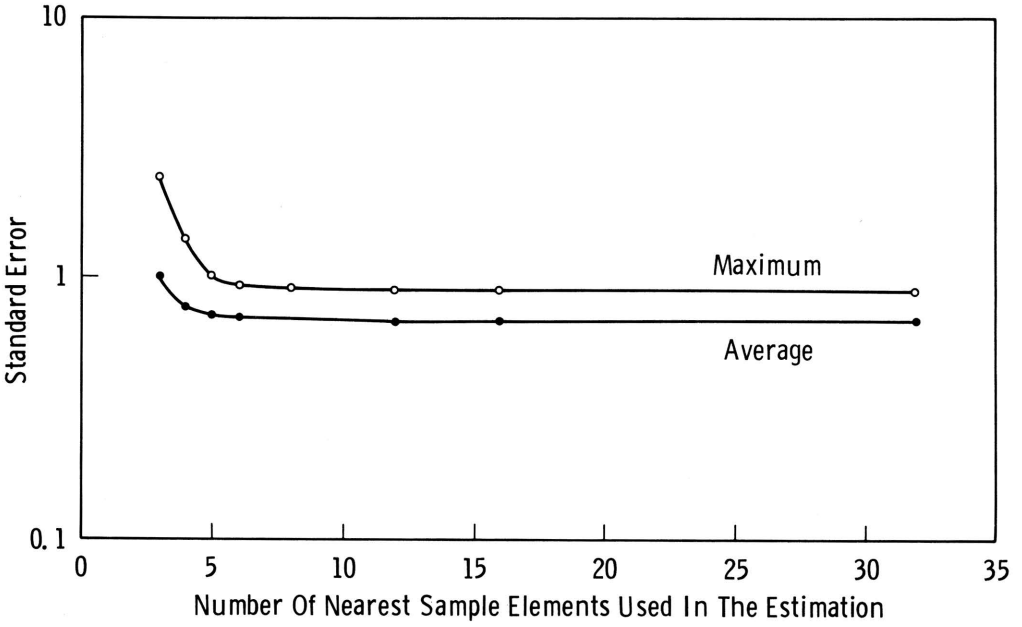


FIGURE 3.18. Sensitivity analysis of the number of nearest neighbors for orthogonal regular traverses intersecting every two points and a first-degree polynomial drift. The sampling density is one point per square mile and the semivariance is linear with slope 1.

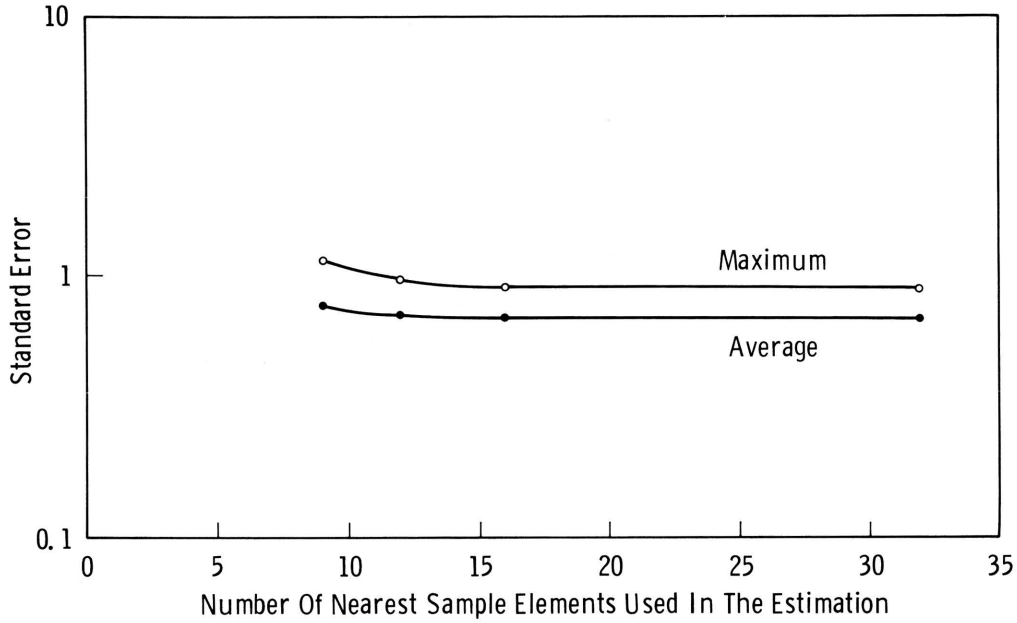


FIGURE 3.19. Sensitivity analysis of the number of nearest neighbors for orthogonal regular traverses intersecting every two points and a second-degree polynomial drift. The sampling density is one point per square mile and the semivariance is linear with slope 1.

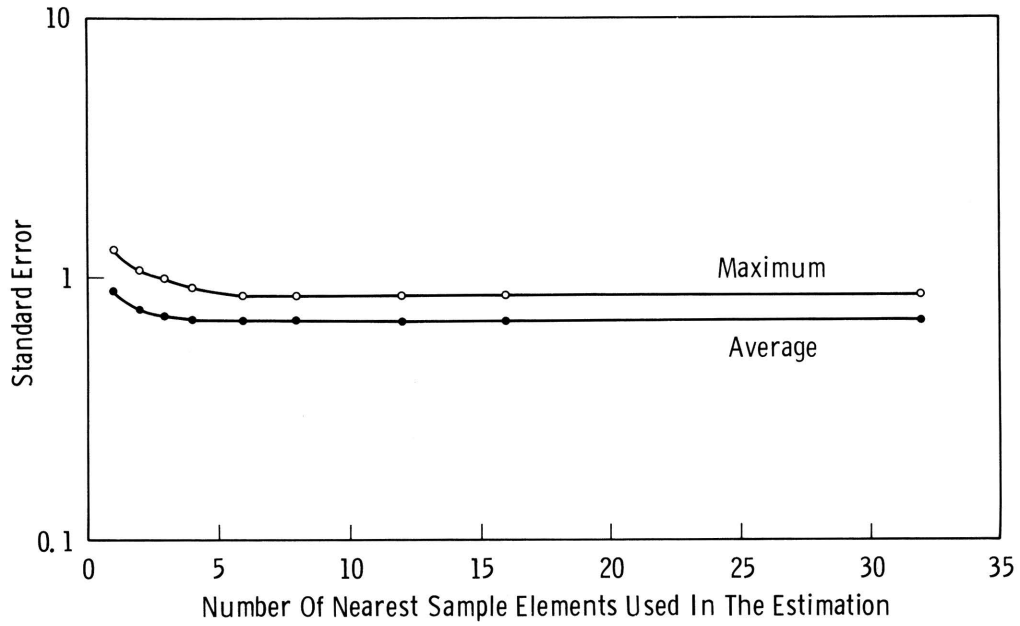


FIGURE 3.20. Sensitivity analysis of the number of nearest neighbors for a hexagonal stratified pattern and no drift. The sampling density is one point per square mile and the semivariance is linear with slope 1.

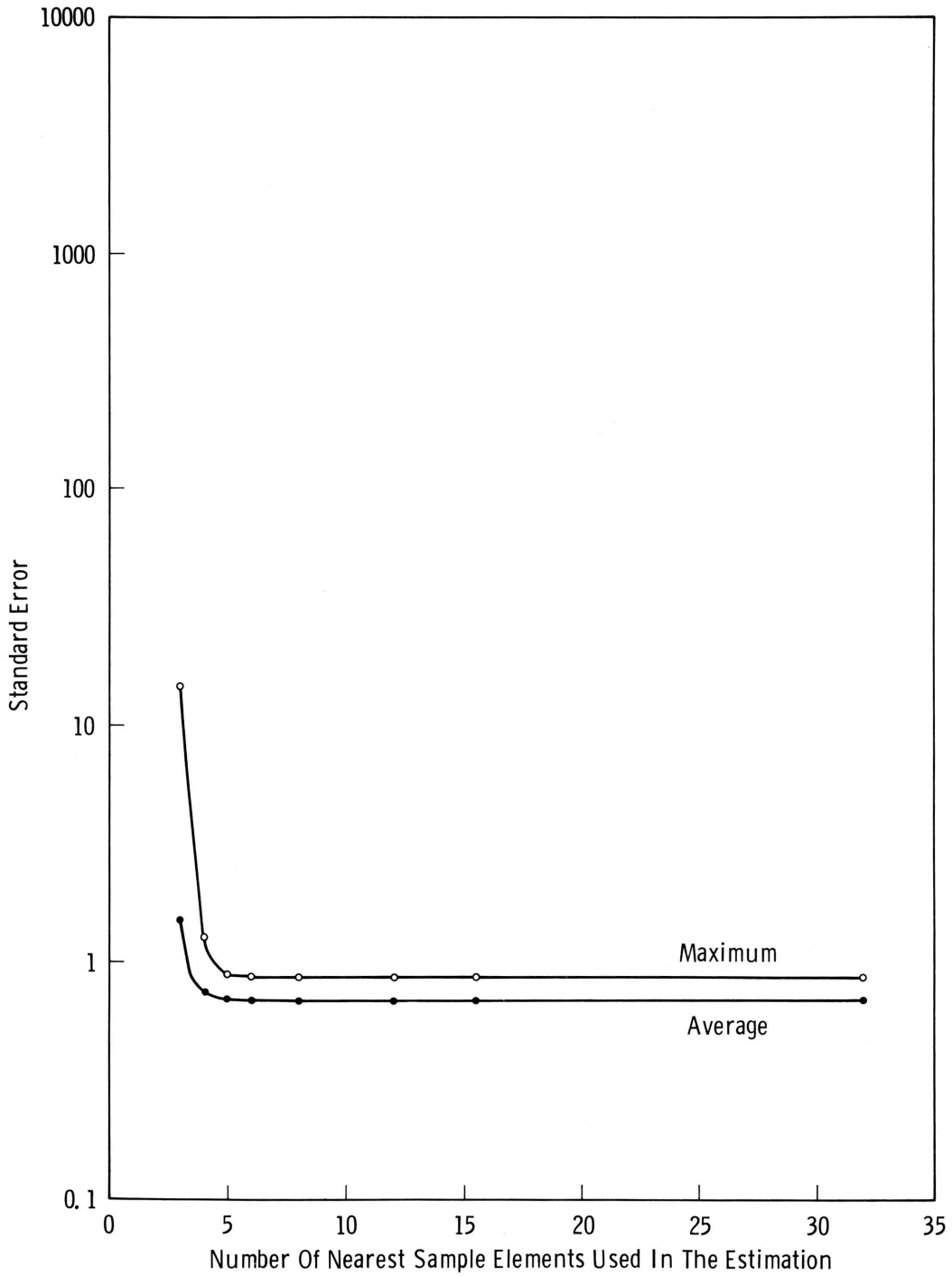


FIGURE 3.21. Sensitivity analysis of the number of nearest neighbors for a hexagonal stratified pattern and a first-degree polynomial drift. The sampling density is one point per square mile and the semivariance is linear with slope 1.

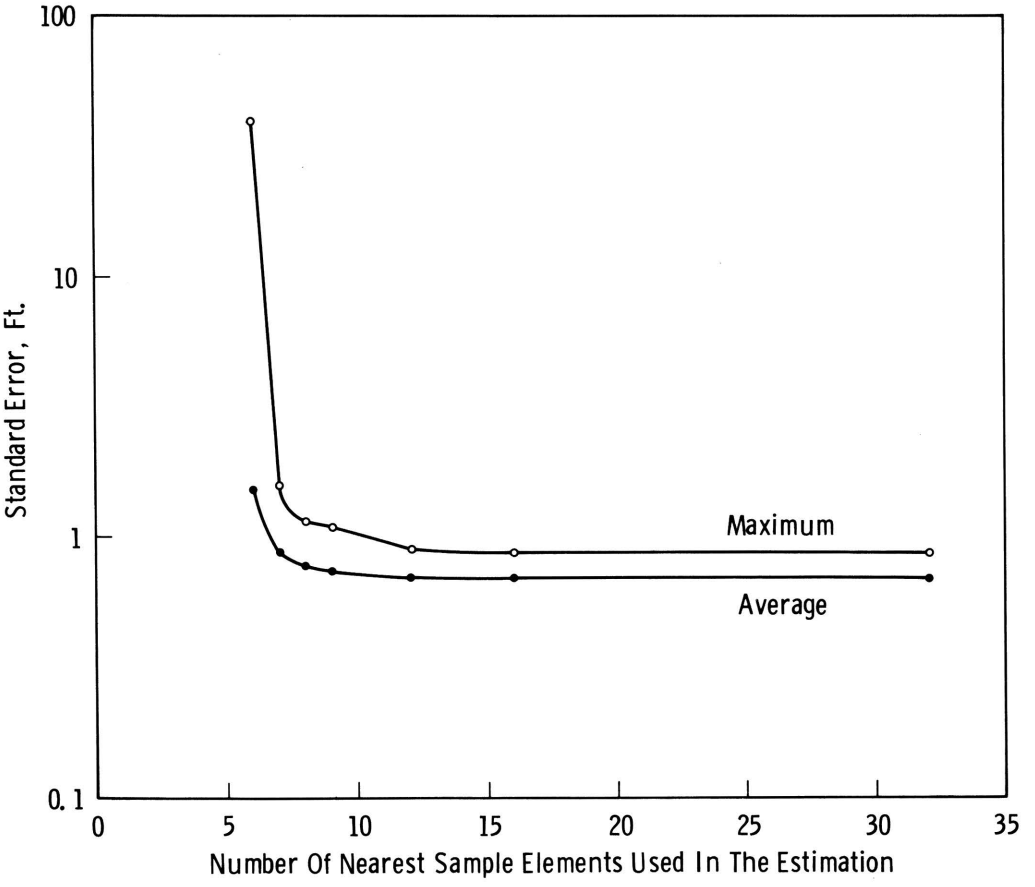


FIGURE 3.22. Sensitivity analysis of the number of nearest neighbors for a hexagonal stratified pattern and a second-degree polynomial drift. The sampling density is one point per square mile and the semivariance is linear with slope 1.

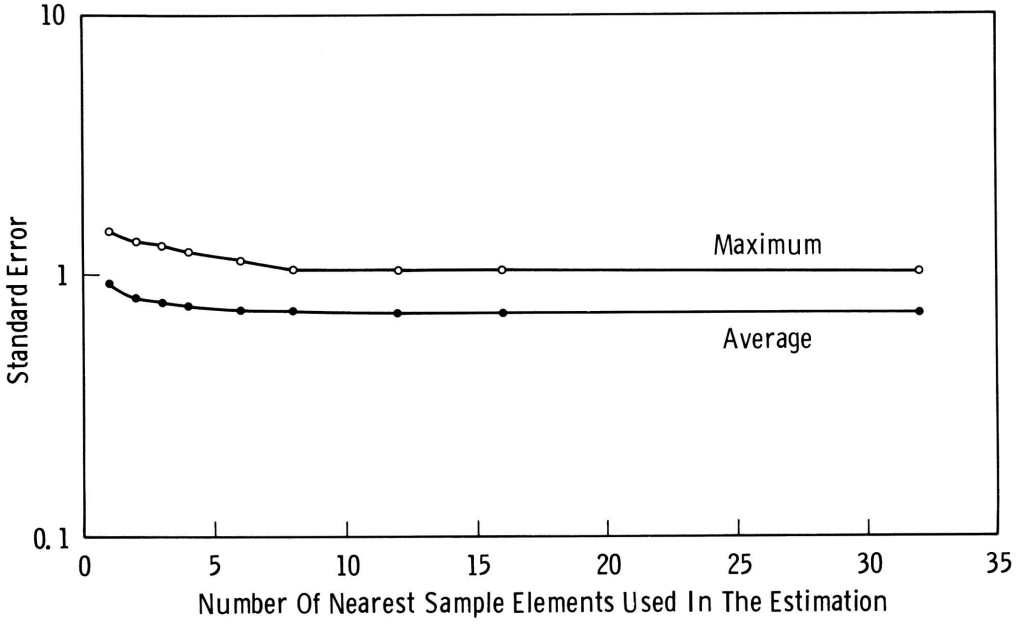


FIGURE 3.23. Sensitivity analysis of the number of nearest neighbors for a random pattern and no drift. The sampling density is one point per square mile and the semivariance is linear with slope 1.

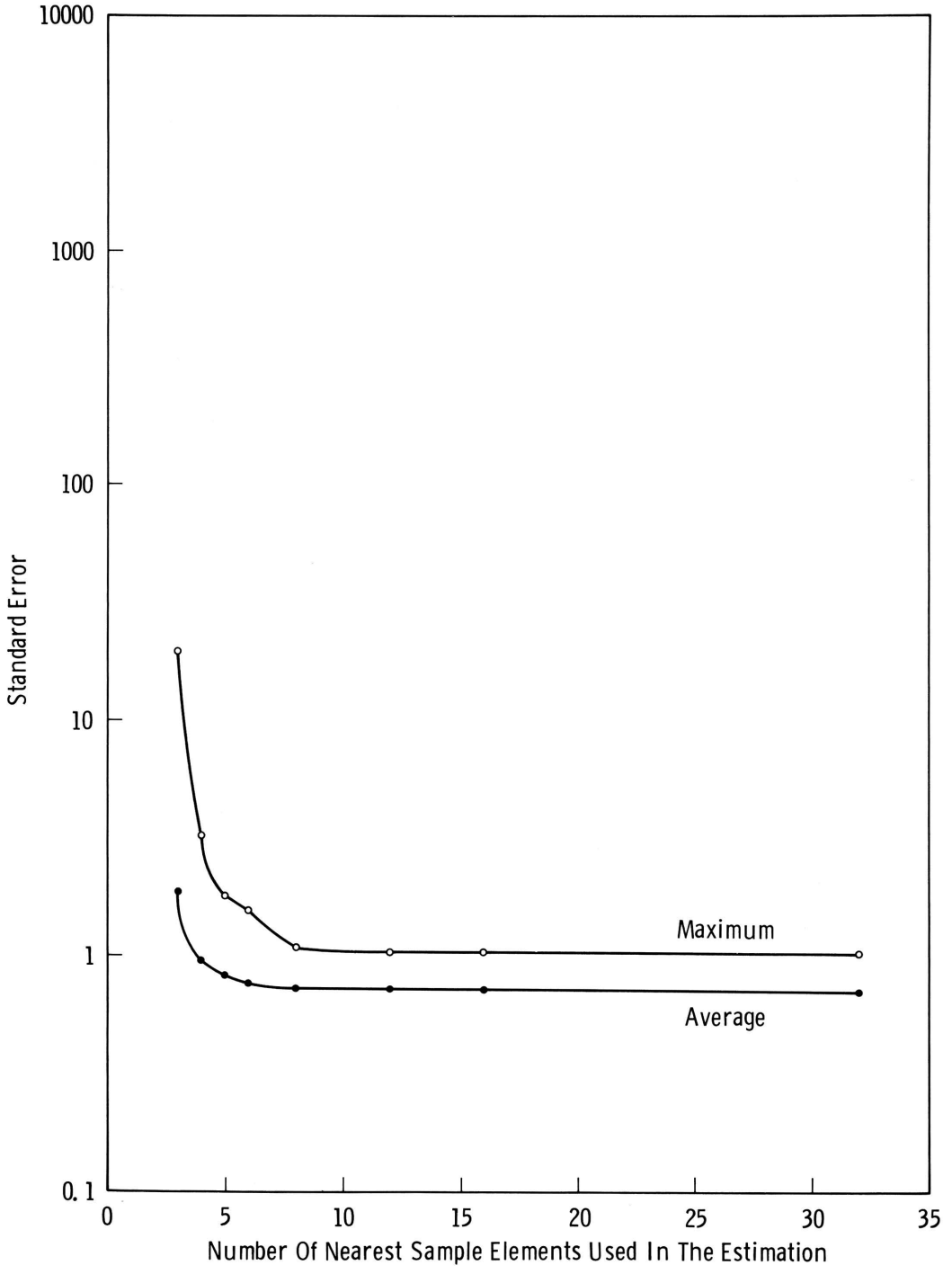


FIGURE 3.24. Sensitivity analysis of the number of nearest neighbors for a random pattern and a first-degree polynomial drift. The sampling density is one point per square mile and the semivariance is linear with slope 1.

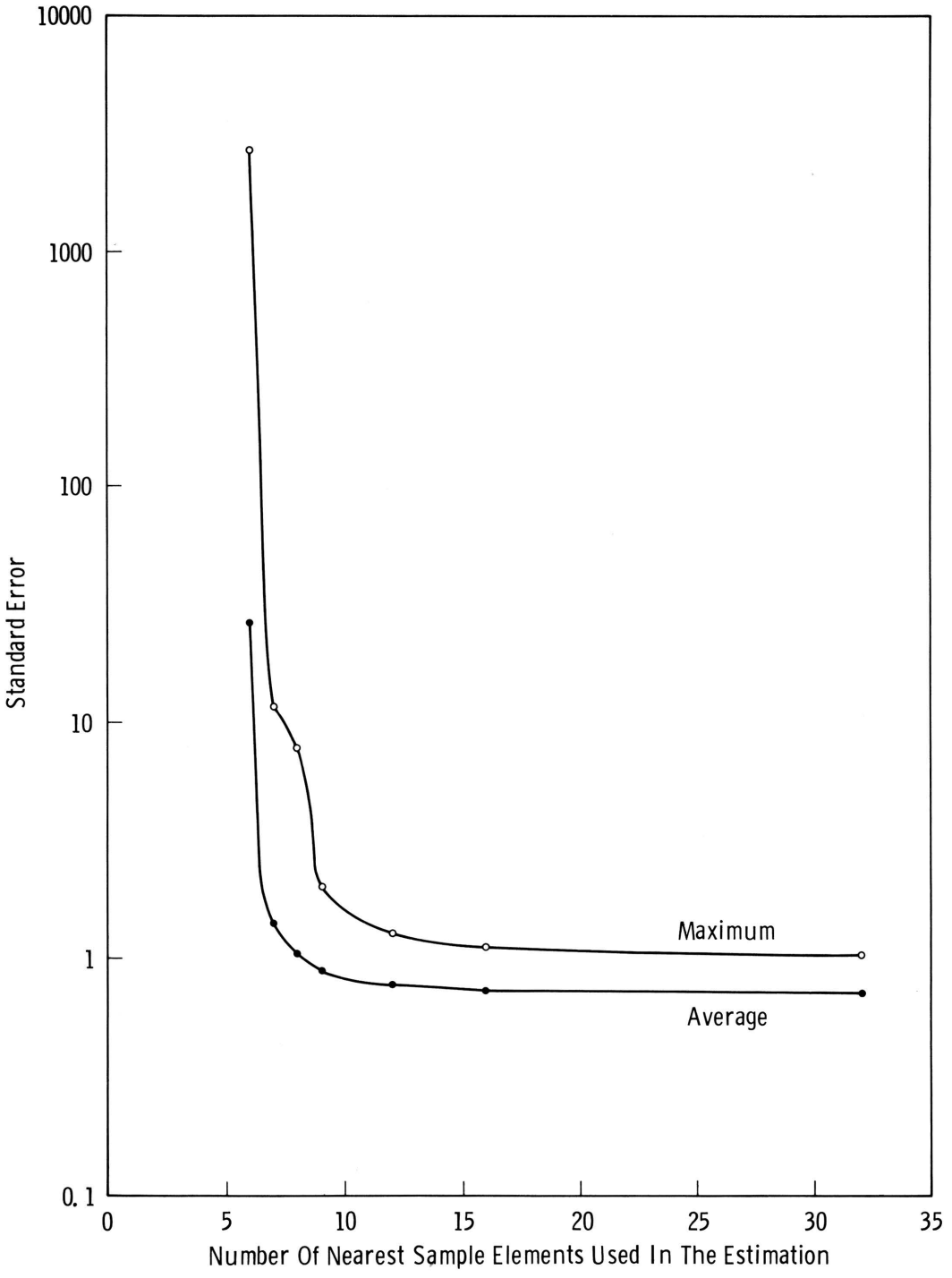


FIGURE 3.25. Sensitivity analysis of the number of nearest neighbors for a random pattern and a second-degree polynomial drift. The sampling density is one point per square mile and the semivariance is linear with slope 1.

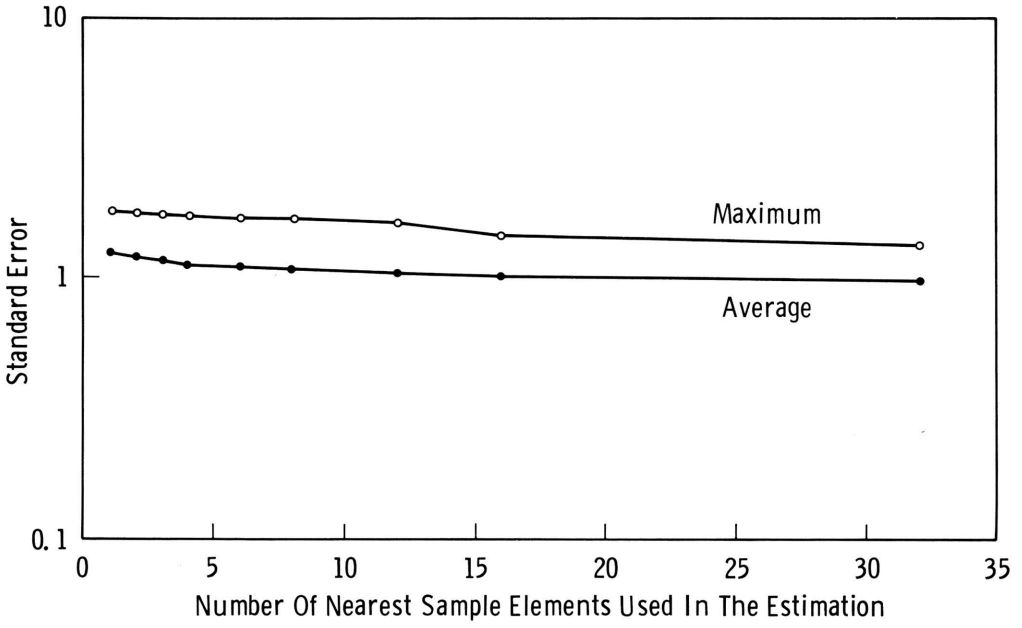


FIGURE 3.26. Sensitivity analysis of the number of nearest neighbors for a pattern of five clusters and no drift. The sampling density is one point per square mile and the semivariance is linear with slope 1.

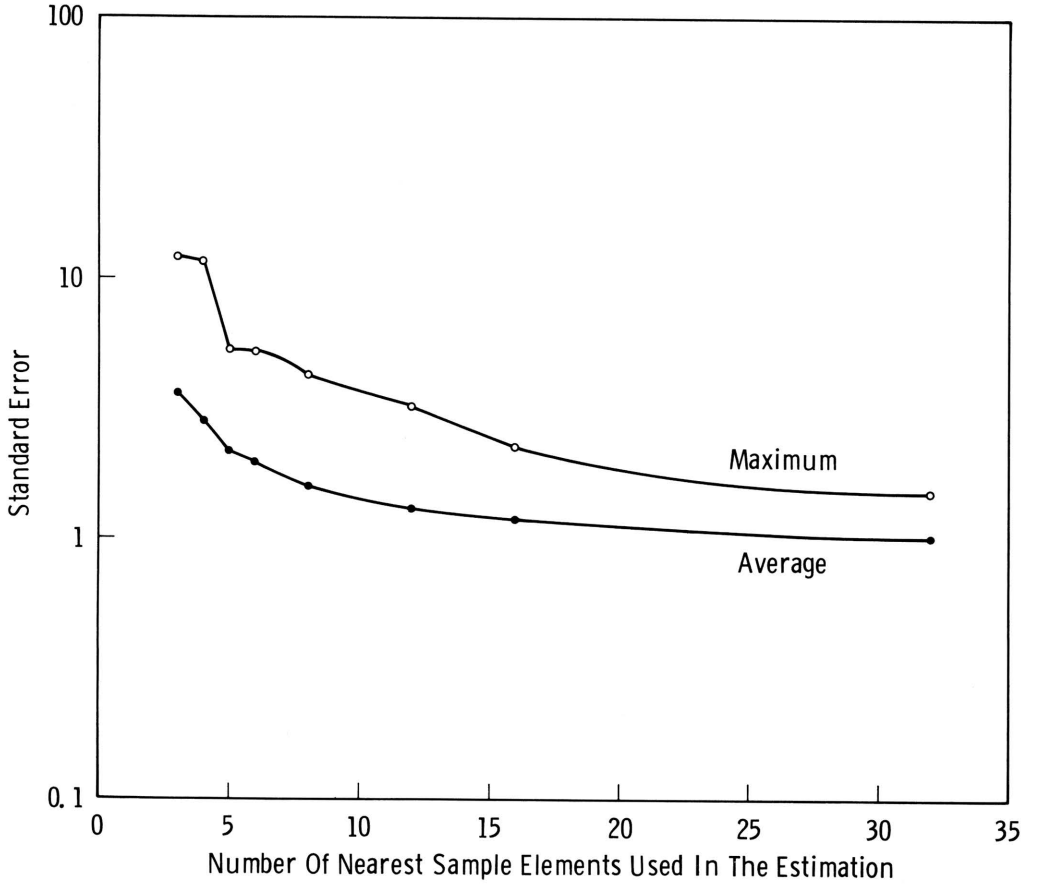


FIGURE 3.27. Sensitivity analysis of the number of nearest neighbors for a pattern of five clusters and a first-degree polynomial drift. The sampling density is one point per square mile and the semivariance is linear with slope 1.

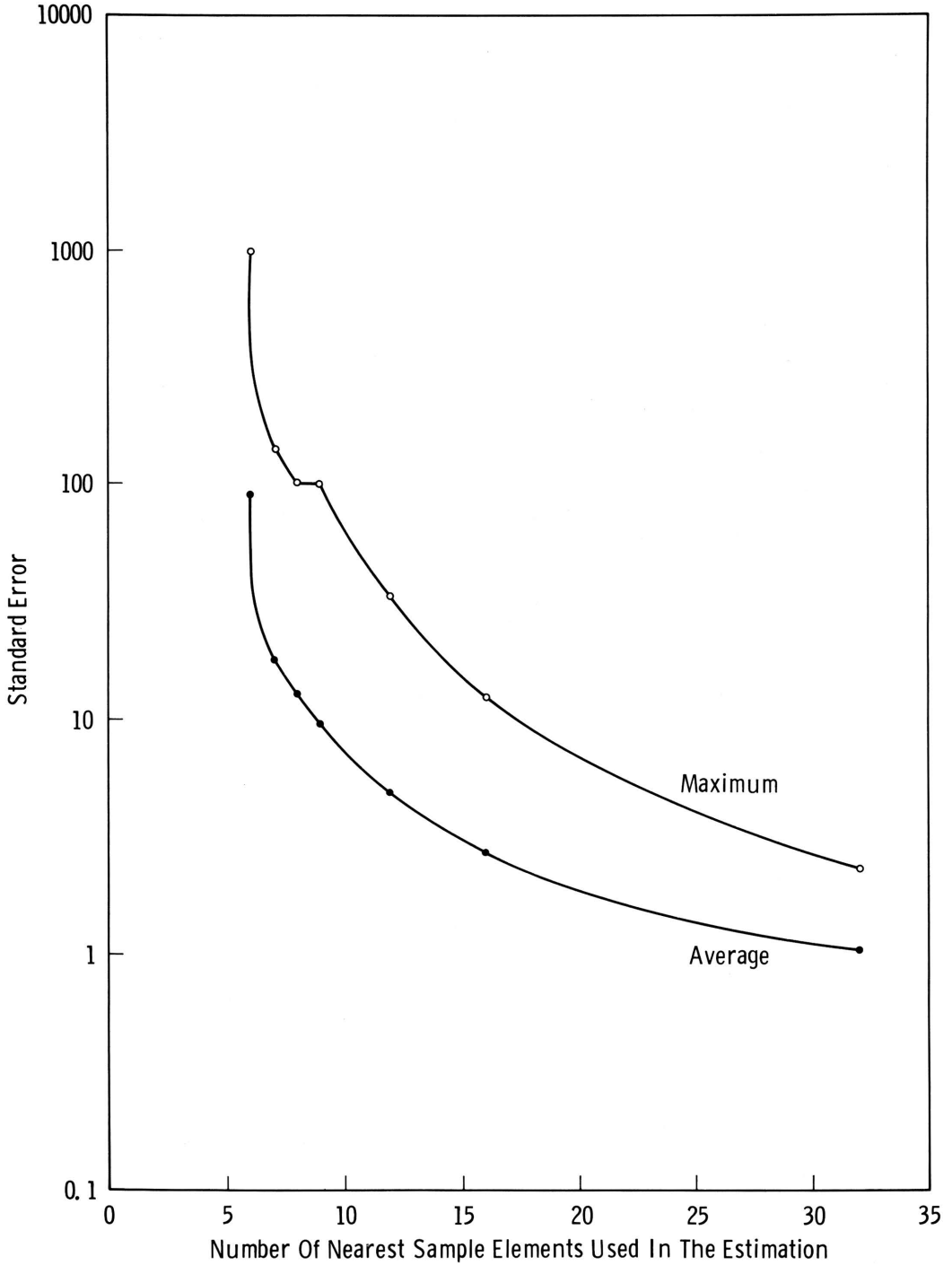


FIGURE 3.28. Sensitivity analysis of the number of nearest neighbors for a pattern of five clusters and a second-degree polynomial drift. The sampling density is one point per square mile and the semivariance is linear with slope 1.

3.5 DENSITY

3.5.1 Spacing

If a spatial pattern is held constant, any modification in density changes the relative distances between points by a constant factor. In order to perform a sensitivity analysis of density, the relative distances between elements of a sample can be changed without altering the pattern by scaling the n-dimensional space.

Lemma 3.2

Let the semivariance of the residuals be isotropic and linear and let the terms $f^i(\vec{x}_j)$ in the drift be distributive. Let λ_j and μ_i represent the weights and slack variables in the universal kriging system of equations. Then the vector X" is invariant under a scaling of the n-dimensional space by a factor α.

$$X'' = \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \cdots \\ \lambda_j \\ \cdots \\ \lambda_k \\ \mu_0/\omega\alpha \\ \mu_1 f^1(\alpha)/\omega\alpha \\ \mu_2 f^2(\alpha)/\omega\alpha \\ \cdots \\ \mu_i f^i(\alpha)/\omega\alpha \\ \cdots \\ \mu_n f^i(\alpha)/\omega\alpha \end{bmatrix} \qquad (3.14)$$

Proof:

The scaling will make every coordinate $\vec{x}_j$ go to $\alpha\vec{x}_j$. As each of the terms in the drift is distributive,

$$f^i(\alpha\vec{x}_j) = f^i(\alpha)f^i(\vec{x}_j) \qquad (3.15)$$

Because the semivariance is linear and isotropic,

$$\gamma(\alpha\vec{h}) = \omega\ \alpha\ |\vec{x}_p - \vec{x}_j| = \omega\ \alpha h_{pj} \qquad (3.16)$$

where the bars stand for the modulus of the vector. After scaling, the universal kriging system of equations in Equation 2.9 becomes

$$A'X'' = B' \qquad (3.17)$$

where A' and B' are the matrices in Equations 3.5 and 3.7, respectively. Notice that neither A' nor B' have any terms in α. Hence X" is independent of α.

Q.E.D.

Theorem 3.2

Let a regionalized variable have a semivariance of the residuals which is isotropic and linear. Then if the terms $f^i(\vec{x}_j)$ in the drift are distributive, the estimation variance changes linearly with scaling by a factor α of the n-dimensional space.

Proof:

From Equations 2.13, 3.15, and 3.16

$$\sigma_E^2(\alpha\vec{x}_0) = \sum_{i=1}^{k} \lambda_j\omega\ \alpha\ h_{0j} + \mu_0 + \sum_{i=1}^{n} \mu_i f^i(\alpha)f^i(\vec{x}_0) \qquad (3.18)$$

The equation can be algebraically transformed into

$$\sigma_E^2(\alpha\vec{x}_0) = \sum_{i=1}^{k} \lambda_j\omega\ \alpha\ h_{0j} + \omega\ \alpha\{\mu_0/\omega\ \alpha\} + \omega\ \alpha \sum_{i=1}^{n} \{\mu_i f^i(\alpha)/\omega\ \alpha\}f^i(\vec{x}_0) \qquad (3.19)$$

Factoring by α and ω:

$$\sigma_E^2(\alpha\vec{x}_0) = \alpha[\omega\{\sum_{i=1}^{k} \lambda_j h_{0j} + \{\mu_0/\omega\ \alpha\} + \sum_{i=1}^{n} \{\mu_i f^i(\alpha)/\omega\ \alpha\}f^i(\vec{x}_0)\}] \qquad (3.20)$$

From Equations 3.14 and 3.7,

$$\sigma_E^2(\alpha\vec{x}_0) = \alpha\ \omega\{X''\}^T B' \qquad (3.21)$$

By Lemma 3.2, X" does not depend on the scaling factor α, and since B' is also independent of α, the estimation variance is a linear function of the scaling factor α.

Q.E.D.

Notice that scaling does not alter the fact that the estimation variance varies linearly with the slope of the linear semivariance.

When α is equal to one in Equation 3.14, then the product $\omega\{X''\}^T B'$ is equal to the estimation variance $S_i^2(\vec{x}_0)$ before scaling. Then, Equation 3.21 can be rewritten as

$$S_f^2(\alpha\vec{x}_0) = \alpha\ S_i^2(\vec{x}_0) \qquad (3.22)$$

3.5.2 Sensitivity to density

Figure 3.29 illustrates Theorem 3.2 and Equation 3.22. Here, the spacing has been scaled by a factor of 4. The reduction in the spacing results in a scaling of relative distances among sample elements and isolines by the same factor. Isoline values do not remain the same, but are reduced by a factor of 2, which is the square root of 4.

Figure 3.29 also illustrates the isomorphism between distance scaling and insertion or deletion of sample elements in infinite patterns. From a scaling viewpoint, points A, B and C in the original pattern are brought to P,Q and R in Figure 3.29b. These new points are assumed to come from surrounding areas originally beyond the edges of Figure 3.29a. The additional points are new only within the area selected, but are not new in the sample. Alternatively, we can consider that the points A, B and C remain at exactly the same locations L, M and N in the second map, and that new sample elements have been added to the sample. Whatever the mechanism used to produce the change from Figure 3.29a to Figure 3.29b, the resulting patterns are identical and so are the standard errors. Remember that the universal kriging system of equations is invariant under a change in the origin and only depends on the relative distances between elements, and between the elements and the estimated location (Journel and Huijbregts, 1978, p. 335).

The insertion or deletion of sample elements implies a change in the density equal to:

$$\alpha = \left(\frac{\rho_f}{\rho_i}\right)^{-1/2} \qquad (3.23)$$

Thus, from Equations 3.22 and 3.23

$$S_f(\vec{\alpha x}_0) = \left(\frac{\rho_i}{\rho_f}\right)^{1/4} S_i(\vec{x}_0) \qquad (3.24)$$

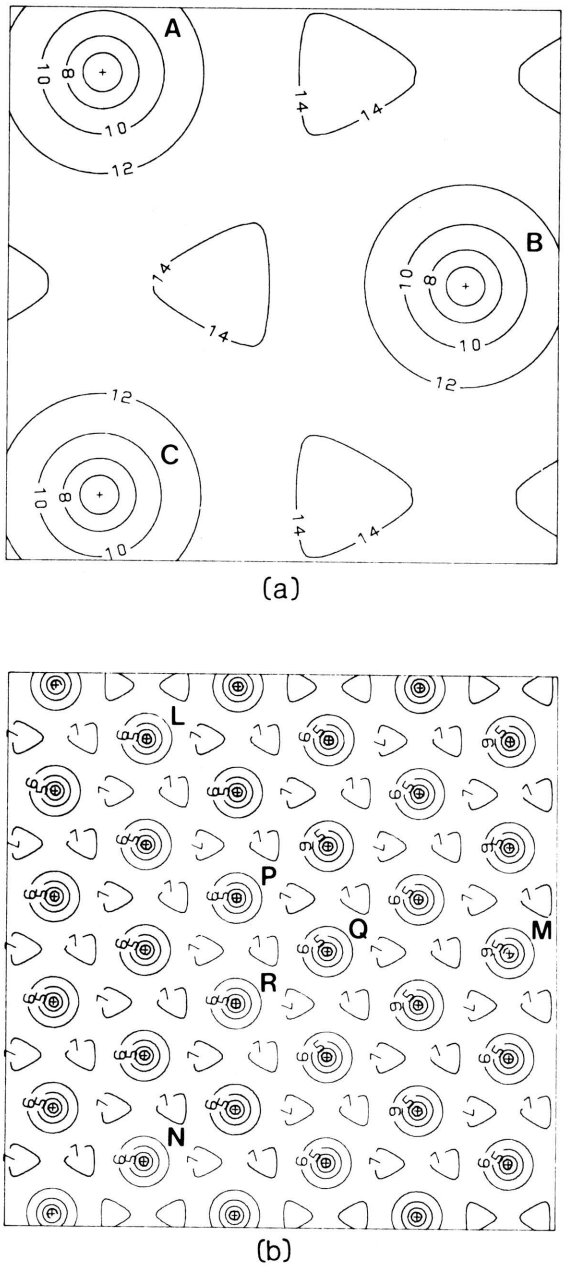


FIGURE 3.29. Spacing and standard error. Reduction in spacing of pattern (a) by a factor of 4 results in pattern (b), with standard error reduced by a factor of 2.

The standard error can be reduced by increasing the density, until a limiting value of zero is reached. In Figure 3.29, scaling by a factor of 1/4 increases the density relative to the original density by 16 times. Note that 1/16 raised to the 1/4 power is 1/2; the contour values are changed by exactly 1/2 in the example.

Because there is a one-to-one correspondence between points in the original pattern and points in the scaled pattern, the average standard error and the maximum standard error change in the same way as the individual points. From Theorem 3.1, Equation 3.24 can be generalized to

$$\rho = \omega^2 \left[\frac{I(1,1)}{I(\omega,\rho)} \right]^4 \qquad (3.25)$$

where ω is the slope of the linear semivariogram; $I(1,1)$ is the efficiency index in Table 3.1 or Appendix B for the given pattern at a density of one point per square mile (0.39 points/km^2) and a semivariogram slope of 1; and $I(\omega,\rho)$ is the desired level of the efficiency index.

In summary, if all other factors are the same for a spatial function sampled by two identical patterns which differ only in their density, the sampling efficiency indices of the pattern with the higher density ρ_u will differ from those of the pattern with the lower density ρ_l by a factor $(\rho_l/\rho_u)^{1/4}$.

3.5.3 Sampling frequency

It was shown in Section 3.3 that regular patterns are the most efficient configurations for sampling. A potential disadvantage of regular patterns is that a systematic bias in estimates could occur if the sampling frequency were equal to a dominant low frequency in the spatial function (Cochran, 1964, p. 218). Figures 3.30 to 3.32 show that such a possibility can be avoided easily by examining a previous structural analysis of the spatial function.

Three spatial functions were simulated and their corresponding semivariances estimated. Figure 3.30 shows a spatial function in profile and its semivariance. The spatial function has a base value of -1 except for two perfectly sinusoidal features with a wavelength of 360 feet (109.7 m). The semivariance for this idealized function intersects the sample variance at a distance of 140 feet (42.7 m), which is close to half the wavelength of the sinusoid. In Figure 3.31 the same spatial function is shown with the addition of a linear drift, which has the effect of tilting the profile. The semivariogram is

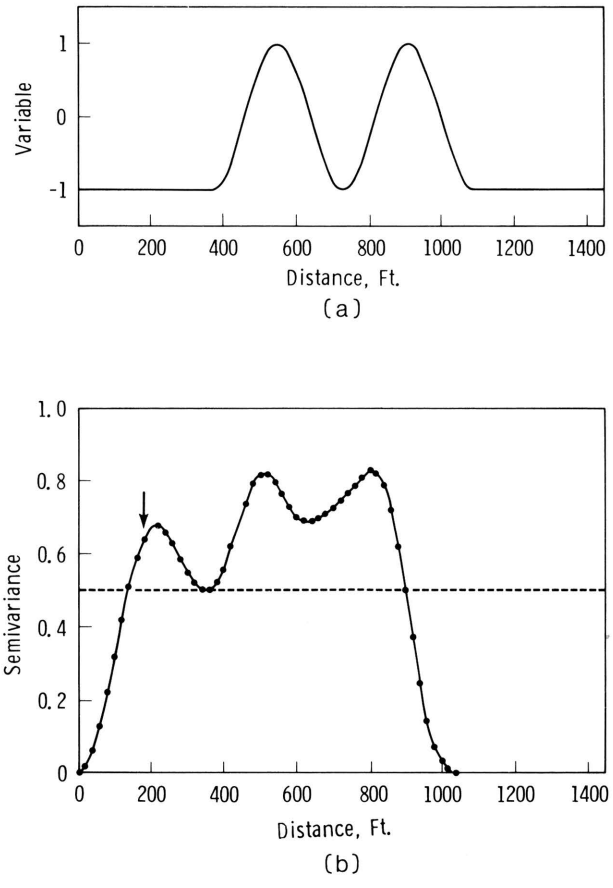


FIGURE 3.30. Semivariance of a spatial function which includes a sinusoidal feature with a wavelength of 360 feet and no drift. (a) Profile. (b) Semivariance. Dashed line represents the sample variance. Arrow indicates one-half the wavelength of features in the profile.

now tangent to the variance of the sample at 240 feet (73.2 m), which is again close to half the wavelength of the dominant frequency. Finally, a spatial function composed of sinusoidal elements with different wavelengths is considered in Figure 3.32. Features in the profile are wavelengths of 180 feet (54.9 m), 360 feet (109.7 m), and 720 feet (219.5 m). The semivariance reaches the sample variance at 260 feet (79.2 m), which is approximately half the average wavelength of the features in the profile.

The distance at which the semivariance is first equal to the population variance is thus a rough estimate of the average dominant low frequency of a profile. If the spacing of samples in a regular pattern is not allowed to reach the critical half wavelength

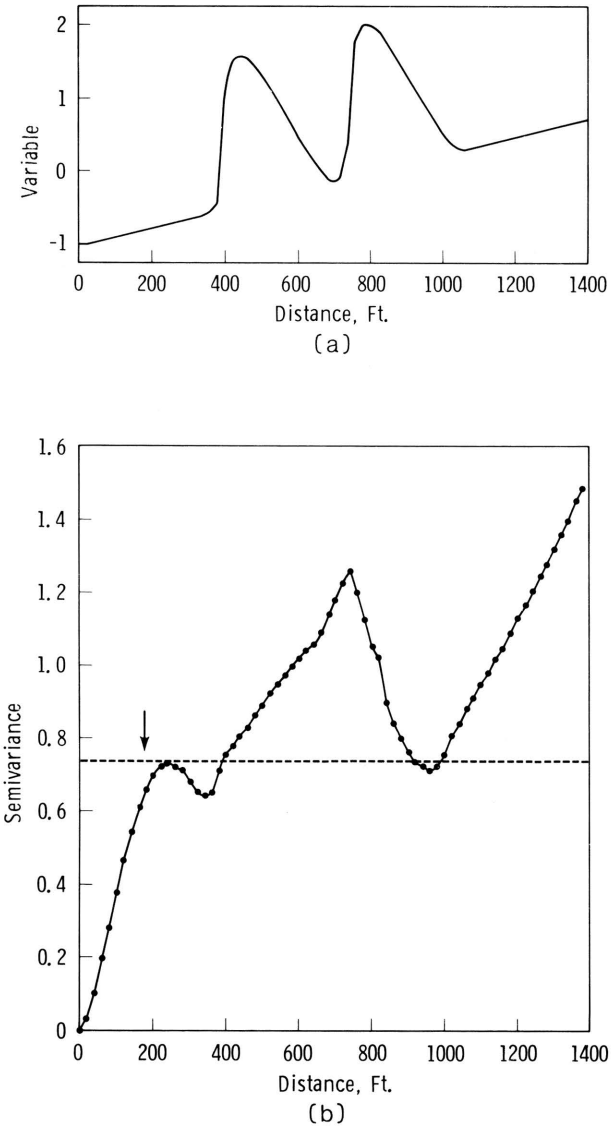


FIGURE 3.31. Semivariance of a spatial function which includes a sinusoidal feature and a linear drift. (a) Profile. (b) Semivariance. Dashed line represents the sample variance. Arrow indicates one-half the wavelength of the features in the profile.

given by the structural analysis, the possibility of a systematic sampling bias can be avoided.

3.6 OPTIMIZATION

3.6.1 Sensitivity analysis

Because pattern is a nominal variable and the number of nearest neighbors is an integer, the sampling efficiency indices are not continuously defined. Only the third factor, sample density, can be expressed as a continuous variable. A function of three variables is difficult to represent graphically; standard practice is to set one variable at selected values and to present the family of surfaces defined by the other two variables. Since pattern is discrete by its nature, it was selected as the parameter to be held constant. For ease of representation, and following the convention used in Figures 3.14 and 3.28, the number of nearest neighbors is represented as a continuous variable. Figure 3.33 is a diagrammatic representation of the average standard error for two patterns.

If maximum value of density is assumed to be one, then the side S-L-O-T is a graph like any two of the average standard error curves shown in Figures 3.14 through 3.28, except that the vertical axis now contains the square of the index. The two surfaces will intersect, as will all such surfaces, at infinite density where the average standard error becomes zero. As shown by Equation 3.24, for a specified number of nearest neighbors in any pattern, changes in density will result in straight lines, such as L-M and R-S, for the particular axes chosen in the block diagram. The envelopes through these families of lines resemble the surfaces R-S-T and L-M-N-O.

3.6.2 An approach to systematic sampling

Equation 3.25 provides a trivial solution to the sampling problem, selecting an infinite density of points which insures that efficiency indices will be equal to zero. When that alternative is not feasible, finding the best sampling procedure is an operations research problem: It is necessary to minimize the sampling requirements for a given sampling efficiency index level.

Let us assume that the desired average standard error is D in Figure 3.33. A horizontal plane through D intersects the average standard error surfaces along lines B-C and E-F. Point Q divides line B-C into two parts: a straight segment from B to Q that is parallel to A-D, and a curved part from Q to C. P is an arbitrary point between Q and C. Points on arc Q-P-C in the lowest surface are answers to the problem.

A solution such as P requires the minimum number of nearest neighbors for a given density. Conversely, given a density, there is no other pattern which requires fewer nearest neighbors to produce the specified average standard error. Should L-M-N-O be the surface corresponding to the hexagonal pattern, the optimum found would be an absolute optimum. Otherwise, the curve Q-P-C is one

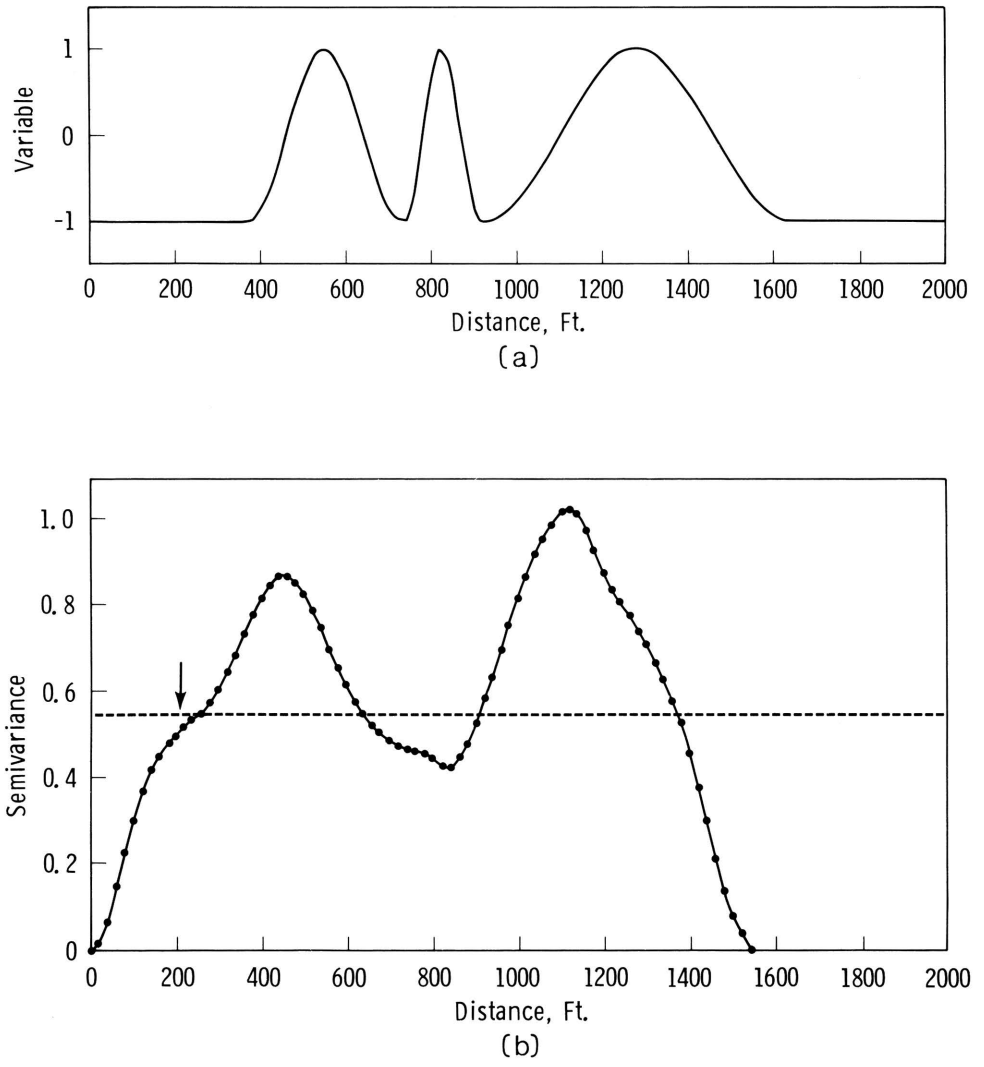


FIGURE 3.32. Semivariance of a spatial function which includes sinusoidal features of different wavelengths. (a) Profile. (b) Semivariance. Dashed line represents the sample variance. Arrow indicates one-half the average wavelength of the features in the profile.

relatively optimum solution among the subset of feasible patterns.

The graphic solution to the sampling problem can be organized as a systematic procedure which will yield an optimal solution if a solution is feasible. Due to the assumptions made in this analysis, the procedure is subject to the following constraints: (a) The residuals must satisfy the intrinsic hypothesis. (b) The physical size of the support of the sample elements and the estimated value must be the same. (c) The sampling space must be two dimensional. (d) The semivariogram of the residuals must be isotropic and linear, with no constant term.

The constraints are sufficiently weak that the method should be applicable to many other spatial variables as well. The second constraint is the strongest and excludes certain special functions which are of interest in ore reserve estimation. The third constraint is satisfied in all mapping problems. Even in studies of three-dimensional space, samples may be collected in a series of two-dimensional planes, which helps to satisfy the second constraint. The semivariance is a monotonically increasing function close to the origin. Because even complex functions can be approximated by linear piecewise interpolation procedures, the problem of satisfying the fourth constraint becomes one of determining how large a portion of the semivariogram can be approximated by a straight line. The shape of the semivariogram is immaterial if the nearest neighbors in the subset used in the estimations are all statistically correlated; that is, if their

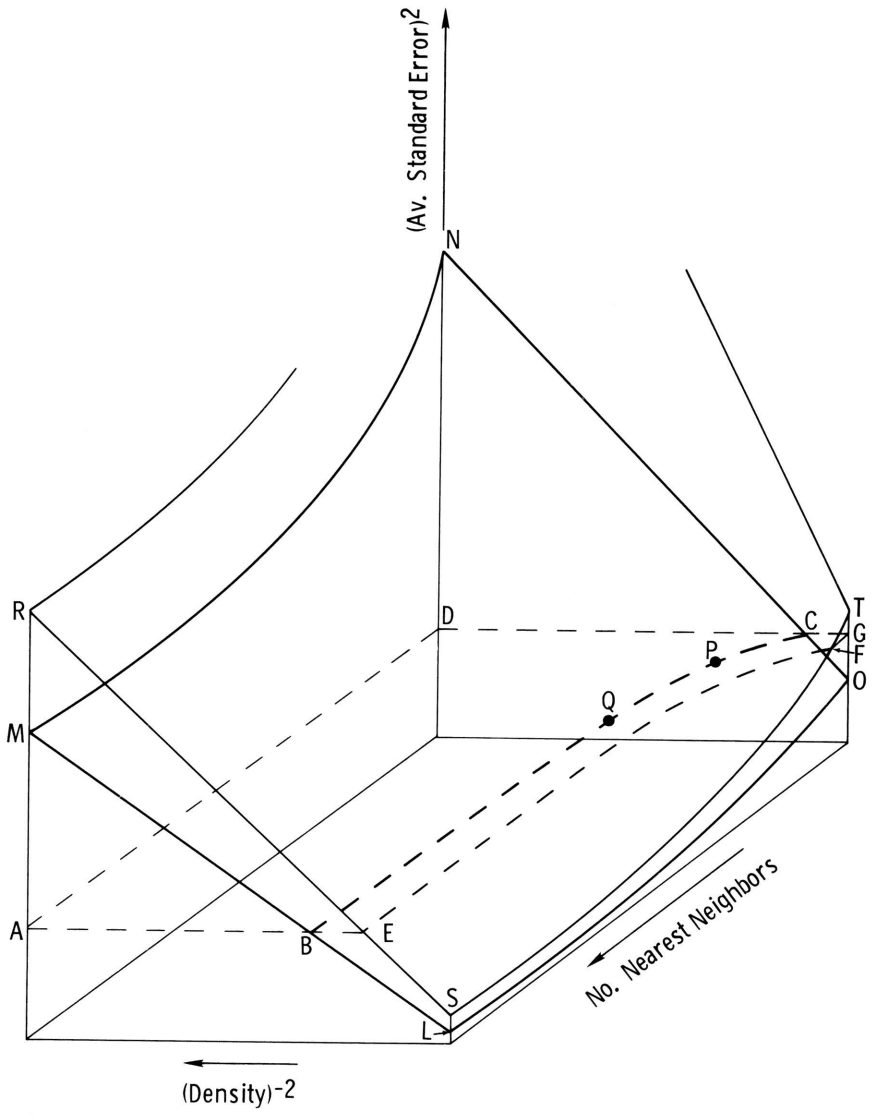


FIGURE 3.33. Average standard error surfaces.

relative distances are never larger than the range.

Algorithm 3.1

The following is a procedure for finding the optimal sampling method for a specified sampling efficiency index.

1. Perform a structural analysis.
2. Decide whether the average standard error or the maximum standard error is the index to be minimized.
3. Enter Table 3.1 for the specified index and appropriate drift. Choose the pattern with the lowest index in the table. In case of a tie, use the pattern with the minimum alternative index.
4. Specify a value for the sampling efficiency index.
5. Use Equation 3.25 to compute the required density using the minimum index of step 2.
6. Calculate the number of sample elements within the neighborhood for which the models in the structural analysis are valid.
7. Compare the number of sample elements inside the neighborhood with the minimum number of points necessary to achieve a solution of the universal kriging system of equations in Table 3.3. Should the number of points inside the neighborhood be insufficient, the solution is unfeasible. In case a solution is required, go back to

step 1 and redefine parameters. Otherwise, stop. If there are enough sample elements inside the neighborhood, proceed to the next step.

8. Enter Table 3.2 and determine the minimum number of sample elements required to obtain the full screen effect.

9. Compare the sample elements inside the neighborhood to the number of sample elements required to obtain the full screen effect. In case the minimum number of sample elements required to obtain the full screen effect is smaller than the number of points that can be placed inside the neighborhood, use a number of nearest neighbors equal to the minimum number of points required for full screen effect and stop. Otherwise use a number of nearest neighbors equal to the number of points inside the neighborhood.

10. Find in Appendix B the table for the selected index and appropriate drift. Take the pattern that minimizes the index for number of nearest neighbors computed in the previous step.

11. Insert the minimum index selected on the preceding step into Equation 3.25 to recompute the optimal density. Stop.

The procedure requires a prior knowledge of the spatial characteristics of the function as provided by a structural analysis. Note, however, that it is not necessary to actually implement the sampling procedure nor to calculate universal kriging estimates in order to determine the optimal pattern. The algorithm reduces to the use of a few simple formulae and tables.

CHAPTER 4

PRACTICAL SAMPLING DESIGN

This systematic approach to sampling has been developed primarily from theoretical considerations, based on the theory of regionalized variables and the use of synthetic patterns. The application of the approach to a practical problem will serve to demonstrate its general validity. In addition, the application will demonstrate the use of the methodology with real data.

4.1 A REAL SAMPLING SYSTEM

4.1.1 The Equus Beds

The spatial function selected to illustrate systematic sampling design is the water table elevation in an area that includes most of the Equus Beds (Fig. 4.1), a major aquifer in Kansas. The Equus Beds produce groundwater through more than 2,000 industrial, municipal, and irrigation wells in Harvey, Sedgwick, McPherson, and Reno counties in the south-central part of the State. The deposits consist of unconsolidated stream-laid material of the Pliocene Blanco and the Pleistocene Meade and Sanborn formations (Stramel, 1956; 1967). These poorly sorted sediments range in size from silt to gravel, with abundant clay lenses (Williams and Lohman, 1949).

The bedrock underlying the Equus Beds consists predominantly of shales of Permian and Cretaceous age which crop out to the north, west, and east (Petri and others, 1964; Albert and Stramel, 1966). Most of the area is drained by the Little Arkansas River; the southern part is drained by the Arkansas River. The Equus Beds range in thickness from 0 to 280 feet (85.3 m).

4.1.2 Data collection

Depth to water is measured in observation wells at least once a year in order to record systematic annual changes in water level. A small fraction of the observation wells measured in Kansas are used only for measuring depth to water, but most observations are made in producing irrigation wells. To minimize drawdown cone effects due to recent pumping in the well itself or in surrounding wells, measurements are taken during the winter when irrigation stops for several months.

Data used in this study were the latest available, consisting of measurements made from December 1980 to March 1981. These observations are called the "January 1981 data" because, except in special circumstances, annual measurements are scheduled to be taken during the first month of every year. The measurements used here are only a fraction of a state-wide survey. Details for individual measurements for the January 1981 data are listed separately (Olea, 1982a; 1982b).

The January 1981 data were used to estimate the form of the water table elevation in the Equus Beds by universal kriging. Figure 4.2 is the resulting contour map of the water table, Figure 4.3 is the corresponding standard error map, and Figure 4.4 is the relative

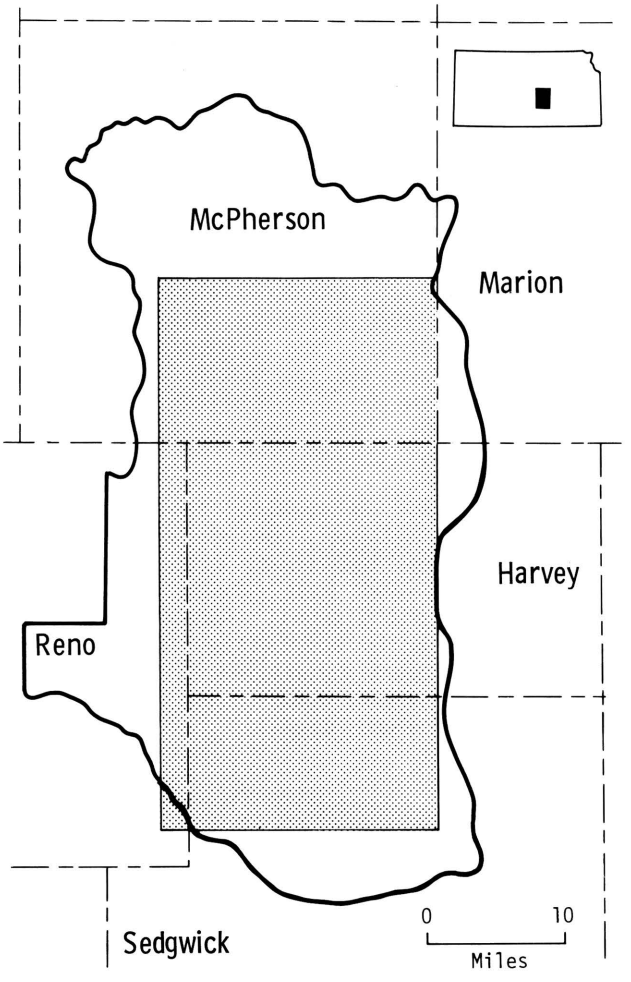


FIGURE 4.1. Location of the Equus Beds in south-central Kansas. Aquifer boundary is indicated by bold solid line; shading indicates the study area.

frequency of the standard error of the water table elevation in Figure 4.3.

4.1.3 Critical analysis

An analysis based upon the principles established in Chapter 3 reveals two major deficiencies in the present observation well network in the Equus Beds: First, the sampling pattern of observation wells is not optimal. The Equus Beds network in the study area consists of 244 observation wells in an area of 800 square miles (2071 km^2), for an average well density of 0.305 wells/square mile (0.12 wells/km^2). The average distance from a well to its nearest neighbor is 0.91 miles (1.46 km). From Equation 3.13, the distance index is 1.0, meaning that the actual well network is indistinguishable from a random arrangement of wells. Random patterns are not as efficient as other patterns, as indicated in Table 3.1 which ranks patterns in terms of their sampling efficiency indices. If other factors are the same, wells arranged in regular patterns, stratified patterns, or along closely spaced profiles should produce more satisfactory results than the present network.

Secondly, the well density is not homogeneous. About 80% of the observation wells are located in the southern half of the area, which causes the estimation accuracy to be uneven in the contour map. This concentration of wells is responsible for values of

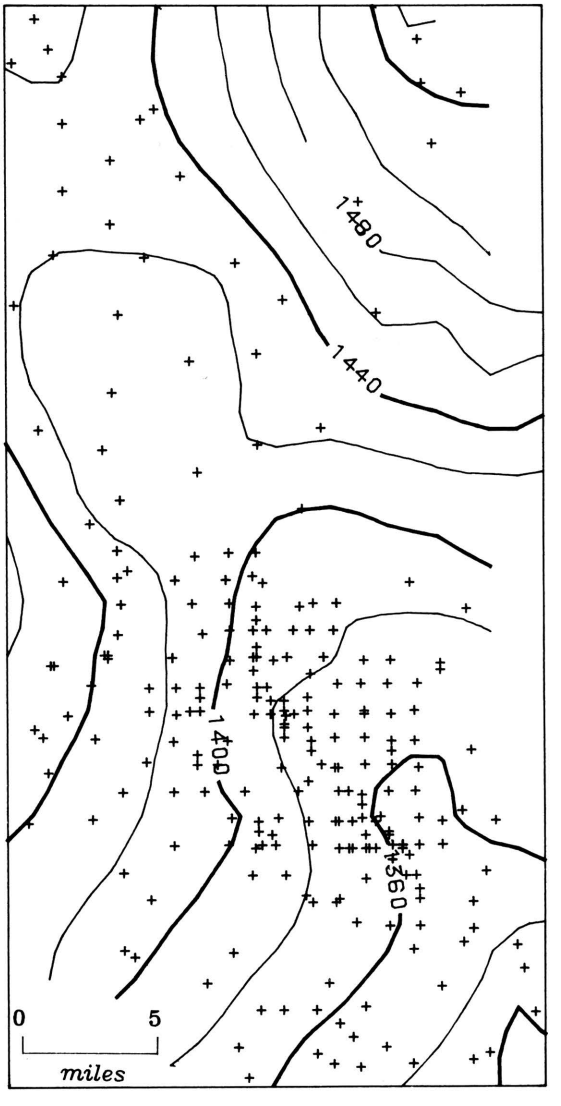


FIGURE 4.2. Water table elevation in the Equus Beds as perceived by the present network of 244 wells. Locations of observation wells are shown by crosses.

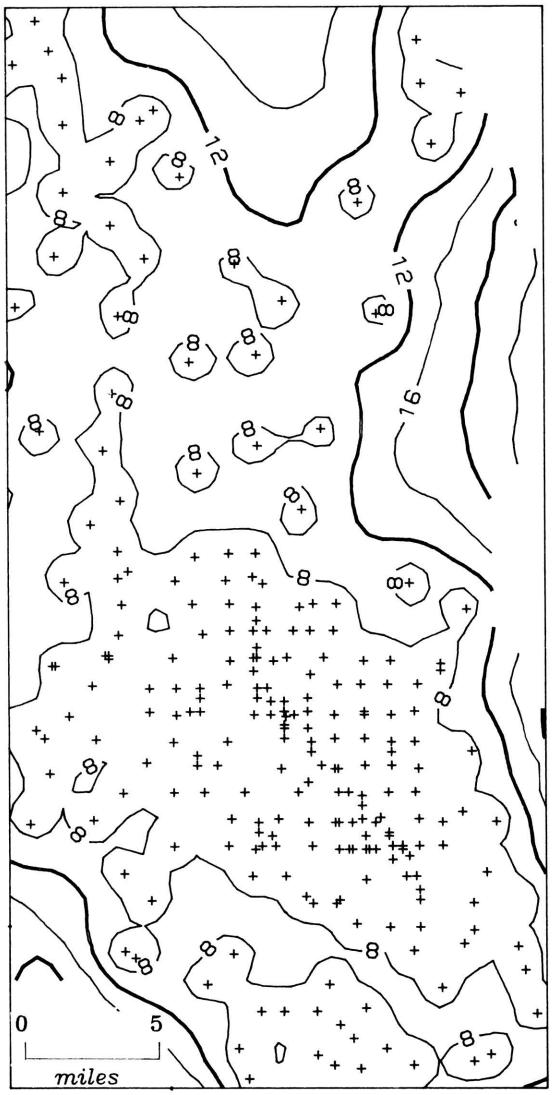


FIGURE 4.3. Standard error of the water table elevation in the Equus Beds as perceived by the present network of 244 wells. Locations of observation wells are shown by crosses.

the standard error that are mostly below 8 feet (2.4 m) in the southern part of the map shown in Figure 4.3, while standard errors higher than 8 feet (2.4 m) predominate in the northern part of the map. The uneven distribution of wells also is responsible for the bimodal distribution of standard error in Figure 4.4.

4.2 AN ALTERNATIVE NETWORK

4.2.1 Design specifications

An alternative observation well network for the Equus Beds which would eliminate the inadequacies of the present network has been designed to illustrate the potentials of the systematic approach of Algorithm 3.1. The water table elevation of the Equus Beds meets all the constraints implicit in the algorithm. The new network is designed to the following specifications:

a. No new observation wells have been permitted. The alternative network consists exclusively of existing observation wells.

b. The accuracy of the estimate of the water table has been kept uniform throughout the Equus Beds. Because no additional wells have been allowed, the alternative network cannot be more accurate than the present network in the portion with sparse control. The new average standard error is required to be within 5% of 12 feet (3.7 m). For purposes of this example, the value of 12 feet can be regarded

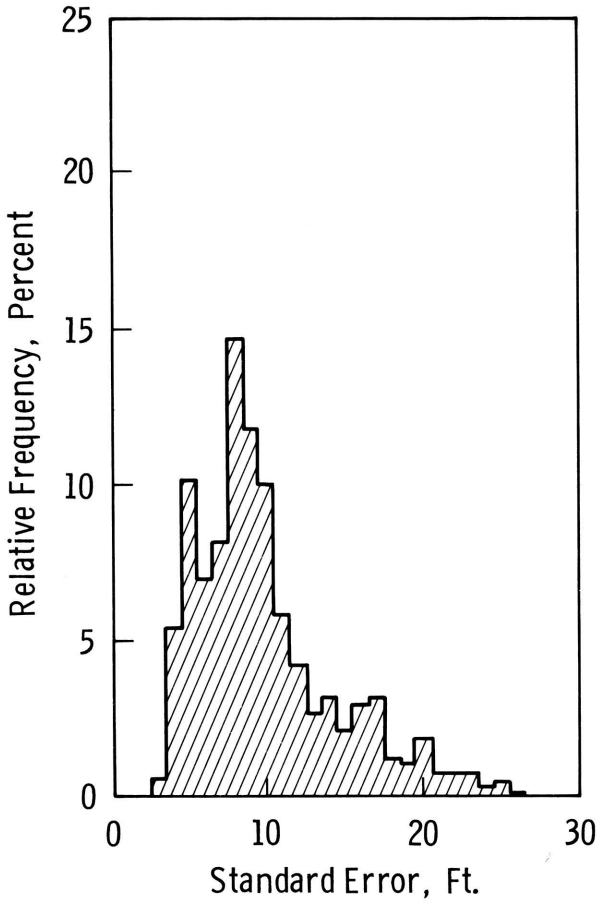


FIGURE 4.4. Relative frequency of the standard error of the water table elevation in the Equus Beds as perceived by the present network of 244 wells. The mean is 9.98 feet and the standard deviation is 4.77 feet.

as an arbitrary decision. However, there are theoretical and economic reasons which make this choice reasonable not only for the Equus Beds but also for the more extensive High Plains aquifer, of which the Equus Beds are part (Olea, 1982a, 1982b).

c. The number of observation wells has been reduced to a minimum.

d. The area estimated by the alternative network comprises at least 99% of the area estimated using the present network.

Provided the spatial function were isotropic, from Table 3.1, the absolute best sampling pattern for the new network would be a regular hexagonal arrangement of observation wells. Unfortunately, the random location of existing observation wells precludes a hexagonal pattern, or any other regular pattern. Therefore, the next most effective pattern was chosen, a stratified hexagonal arrangement which is partly irregular and sufficiently flexible to allow incorporation of existing wells.

The alternative network was designed to have an average standard error somewhat smaller than 12 feet (3.7 m) in most areas to compensate for higher-than-average standard errors which occur in boundary areas where no wells exist. Appendix C contains details of a design having a 10% penalty in the 12 feet (3.7 m) average standard error.

The alternative design requires random selection of one observation well from inside each of a set of regular hexagons that cover the study area. Every hexagon in the pattern contains 16 square miles. Notice that a previous knowledge of the structural characteristics of the spatial function was required, but no actual sampling or universal kriging calculation was necessary to improve the sampling design.

4.2.2 Design verification

Figure 4.5 shows the design for the alternative network of observation wells for the Equus Beds. The new network contains only 47 wells, and represents an 81% reduction from the original number of observations. Although a substantial number of observation wells have been deleted from the network, a comparison of the isolines in the water table elevation maps in Figures 4.2 and 4.5 shows that the loss of information is minimal. Corresponding isolines are almost identical on the two maps. The similarity between the

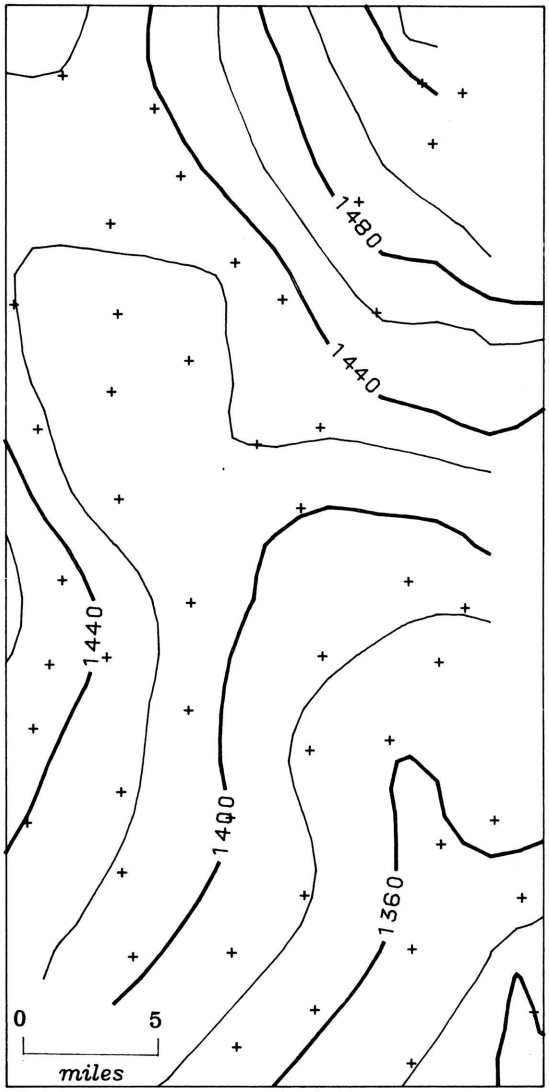


FIGURE 4.5. Water table elevation in the Equus Beds as perceived using an alternative network of 47 wells. Locations of remaining observation wells are shown by crosses.

maps is emphasized in Figure 4.6, which shows that the differences do not exceed 10 feet (3 m); and in Figure 4.7, which shows that 95% of the values in the original map grids are within 5 feet (1.5 m) of their corresponding values for the alternative network.

The actual average standard error increased only to 11.74 feet (3.5 m), which is within 2.2% of the desired level of 12 feet (0.37 m). The improvement in homogeneity of accuracy is shown by the more uniform pattern of isolines in Figure 4.8. The unimodal distribution of standard error in Figure 4.9 shows that 68.7% of the values are within a 5-foot (1.5 m) interval around 10 feet (3 m). There is a 35% reduction in the variance of

the standard error, from 22.8 feet^2 (1.37 m^2) to 14.8 feet^2 (2.11 m^2).

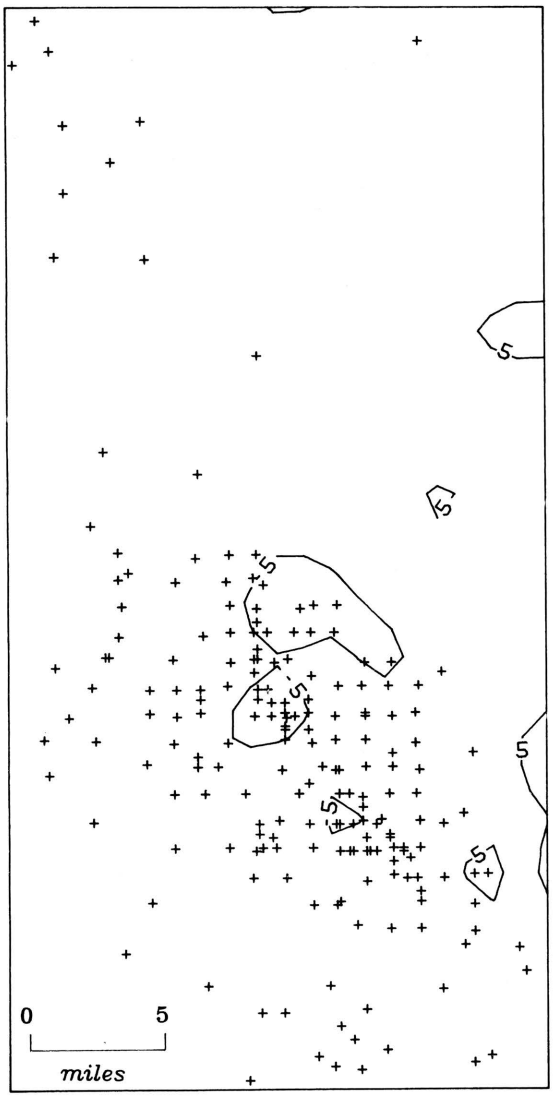


FIGURE 4.6. Map of the difference between water table elevations estimated using the alternative network of 47 wells and elevations estimated by the present network of 244 wells. Crosses show locations of 197 discarded wells.

Table 4.1 summarizes characteristics associated with the present and the alternative networks.

TABLE 4.1

COMPARISON BETWEEN PRESENT AND ALTERNATIVE NETWORKS

	Present	Alternative	Change, %
Pattern	Random	Hexagonal stratified	
Number of wells	244	47	-80.7
Average standard error, ft.	9.98	11.74	17.6
Variance in standard error, ft.^{2}	22.8	14.8	-34.9
Nodes estimated	819	823	0.05

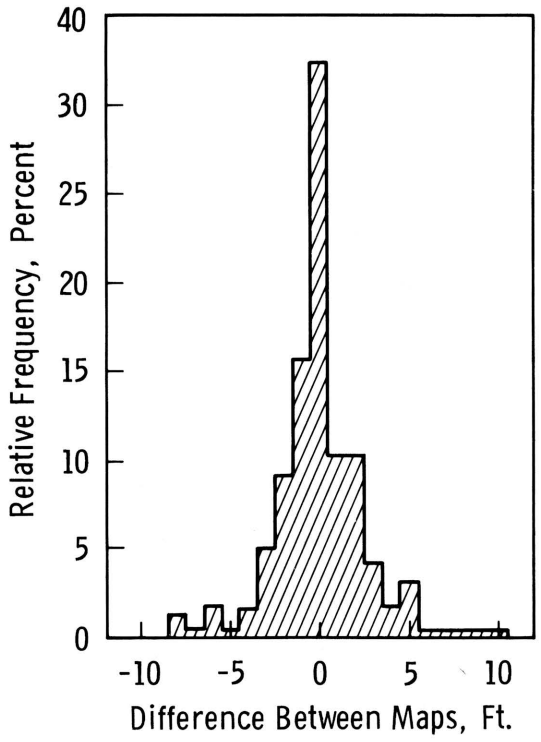


FIGURE 4.7. Relative frequency of difference between water table elevations estimated using alternative network of 47 wells and elevations estimated by present network of 244 wells. The mean is zero and the standard deviation is 2.6 feet.

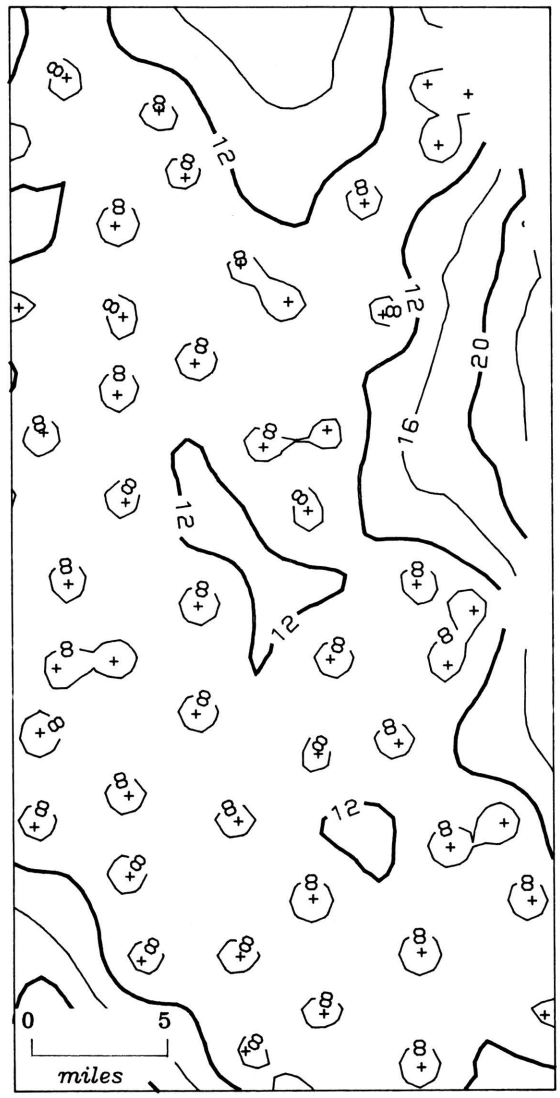


FIGURE 4.8. Standard error of the water table elevation in the Equus Beds as perceived using an alternative network of 47 wells. Locations of observation wells are shown by crosses.

CHAPTER 5

CONCLUSIONS AND RECOMMENDATIONS

This report demonstrates a methodology for systematically sampling spatial functions. Conclusions and recommendations are grouped below in terms of the original objectives of the study. These are concerned with aspects of minimizing the sampling requirements for estimation of a mappable spatial function. Included are considerations of the effects of sample pattern and density, and the nature of the spatial variability of the mapped property.

5.1 FINDINGS

5.1.1 Sampling and spatial functions

Spatial functions are partly stochastic and partly deterministic. Over extensive domains, spatial functions usually show a drift, which is a steady trend in mean value that can be modeled by analytical functions. Over short distances, the exact variation in a spatial function is unpredictable, but the function does possess spatial continuity. The statistical characteristics of spatial functions are summarized in the semivariance, and can be determined by a structural analysis. Closely spaced sample elements generally possess statistical dependence which monotonically deteriorates as distance between elements increases.

Incompletely surveyed spatial functions are characterized by an uncertainty in the actual value of the function at any location not contained in the sample. For an unsampled location it is realistic to assume,

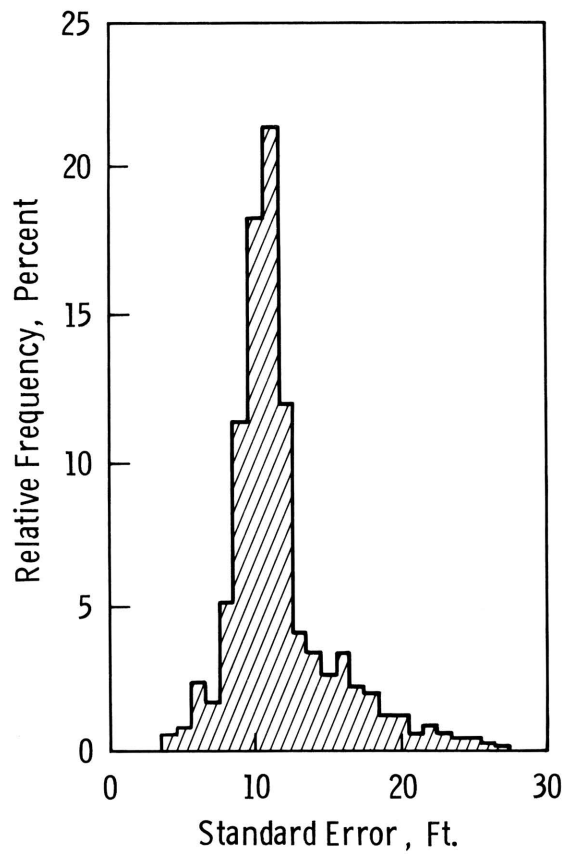


FIGURE 4.9. Relative frequency of the standard error of the water table elevation in the Equus Beds as perceived using the alternative network of 47 wells. The mean is 11.74 feet and the standard deviation is 3.85 feet.

rather than a single value, a set of values with associated probabilities of occurrence. The assumption leads to the selection of a collection of random variables as a convenient model for analyzing partially sampled spatial functions.

5.1.2 Estimation method

The theory of regionalized variables comprises a set of statistical principles well suited to describe all spatial functions and the sampling of such functions, without regard to the physical nature of the variable under study. Universal kriging is an estimation method based on regionalized variable theory. It can provide both unbiased estimates of spatial functions, and standard errors associated with the estimates.

5.1.3 Sampling efficiency factors

The standard error is a measure of accuracy in an estimate; the smaller the standard error, the narrower the expected discrepancy between the true value of the spatial function and the estimate. In this sense, the point standard error is ideally suited for locally assessing the reliability of a sample, and thus can be used to measure the performance of sampling procedures over spatial domains. At any given location, the standard error does not depend on the total number of sample elements nor on the individual values observed (when the samples are large), but rather on the spatial continuity of the function and the geometry of the sample set. The half-dozen sampling efficiency factors for point values of the standard error include:

1. Unmanageable factors
   a. The global variation of the function as given by the drift
   b. The semivariance of the residuals, which are the differences between the drift and the spatial function
2. Manageable factors
   a. The number of nearest sample elements considered by the universal kriging method
   b. The relative distances between sample elements, which can be collectively summarized by the notion of pattern
   c. The density of sample elements
   d. The location of the estimated point relative to the sampling elements

The semivariance is conditional on the drift model, both of which are inherent characteristics of the spatial function. The number of nearest elements, pattern, and density are independent variables associated with the sampling procedure and are under the control of the designer of the sampling scheme.

5.1.4 Sampling efficiency indices

By finding the average and maximum values of the standard error over the sampling domain, two indices of sampling efficiency can be generated that depend only upon the characteristics of the spatial function and on the sampling procedure, and not on the location of the estimated point. This reduces the number of sampling efficiency factors to five.

Often the semivariance of a spatial function can be approximated by a linear model. In such instances, changes in slope of the semivariogram will produce variations in standard error that are directly proportional to the square root of the ratio of the slopes of the semivariograms.

One important finding of this study is that more measurements do not necessarily achieve a more accurate portrayal of a regionalized variable. There are significant differences in the indices obtained using samples spaced according to a systematic sampling procedure and those obtained when sample elements occur haphazardly. A carefully designed sampling scheme using a modest number of samples can outperform an unplanned procedure that incorporates a much larger number of samples. Increasingly greater average and maximum standard errors will be produced by (1) regular, (2) stratified, (3) random, and (4) clustered sampling patterns.

Universal kriging estimates are dominated by those sample elements nearest to the location of the estimate. Judicious placement of the control points used in estimation can eliminate excessive sampling, and creation of blank areas where estimates cannot be made. This may significantly reduce computer time required to map an area.

Theoretically, efficiency indices can be reduced to a perfect value of zero by using an infinite density of observations. However, moderate increases in density of control points produce only minor improvements in the indices because the standard error varies with the fourth power of the inverse of the density.

5.1.5 Systematic sampling approach

Selection of the best sampling efficiency factors is an operations research problem. Pattern, number of nearest neighbors, and density are determined by minimizing the sampling requirements for a given index. The starting approximation for the minimization algorithm selects the pattern with the best index at a density of one sample element per square mile (0.39 elements/km^2) and 32 nearest neighbors. Successive approximations find the optimal density and number of nearest neighbors. The algorithm is valid for any spatial function when (1) the residuals satisfy the intrinsic hypothesis, (2) the sample elements have the same support among themselves and with relation to the estimated value, (3) the semivariance is isotropic and linear with no constant term, and (4) the space domain is two dimensional. The basic information required is simply the spatial characteristics of the function to be sampled, expressed as the semivariance and drift. The algorithm requires only the computation of simple formulae and the use of a few tables. It is not necessary to actually implement the sampling procedure, nor is it necessary to perform any universal kriging computations. Optimality is assured.

5.1.6 Case study

The systematic approach was used to redesign an existing groundwater observation network in the Equus Beds of Kansas, an area within the major High Plains aquifer. The study utilized information gathered during the period December 1980 to March 1981, consisting of depths to the water table in 244 observation wells. The application of the systematic sampling approach in Algorithm 3.1 to the Equus Beds observation well data demonstrates that the procedure can be performed easily and rapidly.

Despite the fact that the design of the alternative observation well network was limited to the use of existing wells, the procedure successfully enhanced the efficacy of the observation well network and compensated for the uneven density of wells. An even better network could have been designed if the study had not been constrained to the use of existing wells.

5.2 SUGGESTIONS

The present observation well network for the Equus Beds is a typical example of oversampling. The well density far exceeds the minimum requirements for proper characterization of the relevant features of the water table. Several considerations are necessary for the design of an optimal sampling scheme. Sampling efficiency index levels must be determined according to the economics of data collection, including a consideration of possible additional uses for the information to be gathered. The level of uncertainty acceptable for the study must be determined. An optimal sampling design should be devised following the procedures presented in this report. Ideally, development of such a design would take place before any measurements have been made, to ensure freedom in the selection of the sampling efficiency factors amenable to optimization: density, pattern, and number of nearest neighbors.

5.2.1 Design of unconstrained systems

The very best sample plan for a spatial function is continuous coverage. If this alternative is unfeasible or impractical, the best discontinuous sampling pattern is regular hexagonal, regardless of the nature of the spatial function. If all other factors remain constant, a regular hexagonal pattern has the following advantages: (1) The minimum average standard error, (2) The lowest maximum standard error, (3) The maximum screen effect, and (4) The minimum number of nearest neighbors needed to assure a stable solution to the universal kriging system of equations.

Practical considerations may necessitate the use of a regular square pattern which, in certain respects, is marginally inferior to the regular hexagonal pattern. It may, however, have advantages in terms of ease of implementation. The regular hexagonal pattern outperforms the regular square pattern by only 1% in average standard error, and by only 3% in maximum standard error. However, as shown in Table 3.2, the regular hexagonal pattern may require only a third of the number of nearest neighbors to obtain the full screen effect and reach the minimum value of the maximum standard error.

5.2.2 Design of constrained systems

Although regular patterns assure the best discontinuous sampling, historical, legal, logistic, or technical considerations may dictate the selection of less optimal alternatives. The designer should at least try to choose from the category of stratified patterns in order to avoid poor sampling efficiency indices or the necessity for costly additional sampling to compensate for the suboptimal location of sample elements. The placement of sample elements in clusters or at random should be avoided. A random pattern requires 4.5 times more elements than a regular hexagonal pattern to assure the same maximum standard error over the sampling

domain. The additional sample requirements for clustered patterns can be an order of magnitude larger.

Geologic samples, whether obtained by hand at low cost or at great expense through a drilling program, tend to be collected in an almost haphazard manner. Given a rudimentary knowledge of the spatial characteristics of the property to be described, and considerations derived from the theory of regionalized variables, a systematic sampling scheme can be devised. Collecting samples in an optimal, systematic pattern will result in better estimates of the variable and significantly reduce the cost of sample collection. Many geologic sampling schemes, such as observation well networks, slowly evolved over time without systematic planning. The inefficiencies of these observation networks may be so great that their continuance cannot be justified. Superior observation networks can easily be devised using techniques described in this report.

REFERENCES

Abler, R., **J.S. Adams**, and **P. Gould**, 1971, Spatial Organization: Prentice-Hall, Englewood Cliffs, New Jersey, 587 p.

Albert, C.D. and **G.J. Stramel**, 1966, Fluvial sediments in the Little Arkansas River Basin, Kansas: U.S. Geological Survey Water-Supply Paper 1978-B, 30 p.

Alldredge, J.R. and **N.G. Alldredge**, 1978, Geostatistics: A bibliography: International Statistical Review, v. 46, p. 77-88.

Armstrong, M. and **R. Jabin**, 1981, Variogram models must be positive-definite: Mathematical Geology, v. 13, no. 5, p. 455-459.

Barret, J.P., 1964, Correction for edge effect bias in point-sampling: Forest Science, v. 10, no. 1, p. 52-55.

Byth, K. and **B.D. Ripley**, 1980, On sampling spatial patterns by distance methods: Biometrics, v. 36, p. 279-284.

Clark, P.J. and **R.C. Evans**, 1954, Distance to nearest neighbors as a measure of spatial relationships in populations: Ecology, v. 35, no. 4, p. 445-453.

Cochran, W.G., 1964, Sampling Techniques: John Wiley & Sons, New York, 413 p.

Cole, J.P. and **C.A.M. King**, 1968, Quantitative Geography: John Wiley & Sons Ltd., London, 692 p.

Dalenius, T., **J. Hájek**, and **S. Zubrzycki**, 1961, On plane sampling and related geometrical problems: Proceedings of the Fourth Berkeley Symposium on Mathematical Statistics and Probability, v. 1, p. 125-150.

David, M., 1977, Geostatistical Ore Reserve Estimation: Elsevier Scientific Publishing Co., Amsterdam, 364 p.

David, M., 1978, Sampling and estimation problems for three dimensional spatial stationary and non-stationary stochastic processes as encountered in the mineral industry: Journal of Statistical Planning and Inference, v. 2, p. 211-244.

Delhomme, J.P., 1979, Spatial variability and uncertainty in groundwater flow parameters: A geostatistical approach: Water Resources Research, v. 15, no. 2, p. 269-280.

Deming, W.E., 1966, Some Theory of Sampling: Dover Publications, New York, 602 p.

Diggle, P.J. and **B. Matérn**, 1980, On sampling designs for the study of point-event nearest neighbor distributions in R^2: Scandinavian Journal of Statistics, v. 7, p. 80-84.

Hannan, E.J., 1962, Systematic sampling: Biometrika, University Press, Cambridge, England, v. 49, Parts 1 and 2, p. 281-283.

Henley, S., 1981, Nonparametric Geostatistics: Applied Science Publishers Ltd., London, 145 p.

Hogg, R.V. and **A.T. Craig**, 1978, Introduction to Mathematical Statistics, 4th Ed.: The Macmillan Co., New York, 415 p.

James, G. and **R.C. James**, 1976, Mathematics Dictionary: Van Nostrand Reinhold Co., New York, 509 p.

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

Kendall, M.G. and **W.R. Buckland**, 1971, A Dictionary of Statistical Terms: Oliver & Boyd, Edinburgh, 166 p.

King, L.J., 1969, Statistical Analysis in Geography: Prentice - Hall, Englewood Cliffs, New Jersey, 288 p.

Madow, W.G., 1949, On the theory of systematic sampling, II: The Annals of Mathematical Statistics, v. XX, p. 333-354.

Madow, W.G. and **L.H. Madow**, 1944, On the theory of systematic sampling, I: The Annals of Mathematical Statistics, v. XV, p. 1-24.

Matérn, B., 1960, Spatial variation: Meddelanden från Statens Skogsforskningsinstitut, Stockholm, v. 49, no. 5, p. 1-144.

Matheron, G., 1965, Les variables régionaliseés et leur estimation: Masson et Cie, Editeurs, Paris, 305 p.

Matheron, G., 1971, The theory of regionalized variables and its applications: Les Cahiers du Centre de Morphologie Mathématique, Ecole Nationale Supérieure des Mines, Paris, 211 p.

McBratney, A.B. and **R. Webster**, 1981, The design of optimal sampling schemes for local estimation and mapping of regionalized variables--II: Computers and Geosciences, v. 7, no. 4, p. 335-365.

McBratney, A.B., **R. Webster**, and **T.M. Burgess**, 1981, The design of optimal sampling schemes for local estimation and mapping of regionalized variables--I: Computers and Geosciences, v. 7, no. 4, p. 331-334.

Newton, M.J., 1973, The application of geostatistics to mine sampling patterns: International Symposium on Computer Applications in the Mineral Industry, 11th Session, Tucson, Arizona, D44-D58.

Olea, R.A., 1972, Application of regionalized variable theory to automatic contouring: Special Report, Project 131, American Petroleum Institute, Center for Research, University of Kansas, Lawrence, 191 p.

Olea, R.A., 1975, Optimum mapping techniques using regionalized variable theory: Kansas Geological Survey Series on Spatial Analysis No. 2, University of Kansas, Lawrence, 137 p.

Olea, R.A., 1977, Measuring spatial dependence with semivariograms: Kansas Geological Survey Series on Spatial Analysis No. 3, University of Kansas, Lawrence, 29 p.

Olea, R.A., 1980, Design of an observation well network, with an application to Kansas Groundwater Management District No. 2: Kansas Geological Survey Internal Distribution Publication, University of Kansas, Lawrence, 32 p.

Olea, R.A., 1982a, Systematic approach to sampling of spatial functions: Doctoral dissertation, Chemical and Petroleum Engineering Department, University of Kansas, Lawrence, 277 p.

Olea, R.A., 1982b, Optimization of the High Plains aquifer observation network, Kansas: Kansas Geological Survey Groundwater Series 7, University of Kansas, Lawrence, 73 p.

Payandeh, B., 1970, Relative efficiency of two-dimensional systematic sampling: Forest Science, v. 16, no. 3, p. 271-276.

Petri, L.R., **C.W. Lane**, and **L.W. Furness**, 1964, Water resources of the Wichita area, Kansas: U.S. Geological Survey Water-Supply Paper 1499-I, 69 p.

Quenouille, M.H., 1949, Problems in plane sampling: The Annals of Mathematical Statistics, v. XX, p. 355-375.

Rendu, J.M., 1976, The optimization of sample spacing in South African gold mines: Journal of the South African Institute of Mining and Metallurgy, v. 76, no. 9, p. 392-399.

Ripley, B.D., 1981, Spatial Analysis: John Wiley & Sons, New York, 252 p.

Rhynsburger, D., 1973, Analytic delineation of Thiessen polygons: Geographical Analysis, v. 5, p. 133-144.

Sampson, R.J., 1978 [1975], SURFACE II graphics system: Kansas Geological Survey Series on Spatial Analysis No. 1, University of Kansas, Lawrence, 240 p.

Stramel, G.J., 1956, Progress report on the ground-water hydrology of the Equus Beds area, Kansas: Kansas Geological Survey Bulletin 119, Part 1, University of Kansas, Lawrence, 59 p.

Stramel, G.J., 1967, Progress report on the ground-water hydrology of the Equus Beds area, Kansas--1966: Kansas Geological Survey Bulletin 187, Part 2, University of Kansas, Lawrence, 27 p.

Williams, C.C. and **S.W. Lohman**, 1949, Geology and ground-water resources of part of south-central Kansas: Kansas Geological Survey Bulletin 79, University of Kansas, Lawrence, 455 p.

Williams, W.H., 1978, A Sampler on Sampling: John Wiley & Sons, New York, 254 p.

NOMENCLATURE

The following table lists the symbols used in this publication, giving the name of the symbol and a brief explanation of its meaning. Also given are the dimensions of the quantity represented by the symbol.

Symbol	Explanation	Dimension
cov	Autocovariance	(same as Z)^2
E	Mathematical expectation	any
f	Monomial of degree j on the coordinates	L^j
$\vec{h}$	Distance in an n-dimensional space	L
I	Sampling efficiency index	same as Z
M	Drift	same as Z
R	Distance index	same as Z
S	Standard error	same as Z
$\vec{x}$	Location in an n-dimensional space	L
Z	Regionalized variable	any
z	Sample element of the regionalized variable	same as Z
γ	Semivariance	(same as Z)^2
λ	Universal kriging weight	
μ	Universal kriging slack variable	
ρ	Density in n-dimensional space	L^{-n}
σ^2	Variance	(same as Z)^2
ω	Semivariance slope at the origin	(same as Z)^2/L

APPENDIX A

VERIFICATION OF THE UNIVERSAL KRIGING PROGRAM

A substantial part of this project relies on results provided by an original version of a program for solving the universal kriging system of equations (Olea, 1972) and a modified version included in SURFACE II (Sampson, 1978). Care was taken to fully test the validity of the results under all possible circumstances of interest in this study. The following is an example of a verification which also illustrates the computation of standard error.

Problem:

Compute the standard error, as predicted by universal kriging, at location P in Figure A.1. The sample consists of only the four regularly spaced points shown in Figure A.1. The drift is a first degree polynomial of the geographical coordinates and the semivariance is linear with a slope of 1.

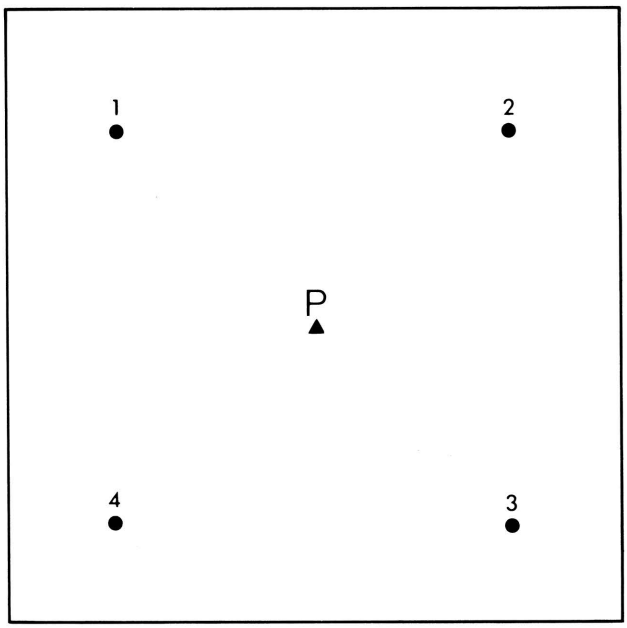


FIGURE A.1. Four sample elements in a square pattern. Elements identified as 1, 2, 3, and 4 are used to compute the standard error of the estimate at location P.

Solution:

As the semivariance is linear with slope 1, the semivariance is numerically equal to the distance between pairs of points. Let us assume that the reference axes are parallel to the square sides and that the origin is at the estimation location. From Equations 2.9-2.12:

$$
\begin{aligned}
a\lambda_2 + \sqrt{2a\lambda_3} + a\lambda_4 + \mu_0 - 0.5a\mu_1 + 0.5a\mu_2 &= 0.5\sqrt{2a} \\
a\lambda_1 + a\lambda_3 + \sqrt{2a\lambda_4} + \mu_0 + 0.5a\mu_1 + 0.5a\mu_2 &= 0.5\sqrt{2a} \\
\sqrt{2a\lambda_1} + a\lambda_2 + a\lambda_4 + \mu_0 + 0.5a\mu_1 - 0.5a\mu_2 &= 0.5\sqrt{2a} \\
a\lambda_1 + \sqrt{2a\lambda_2} + a\lambda_3 + \mu_0 - 0.5a\mu_1 - 0.5a\mu_2 &= 0.5\sqrt{2a} \\
\lambda_1 + \lambda_2 + \lambda_3 + \lambda_4 &= 1 \\
a(-0.5\lambda_1 + 0.5\lambda_2 + 0.5\lambda_3 - 0.5\lambda_4) &= 0 \\
a(0.5\lambda_1 + 0.5\lambda_2 - 0.5\lambda_3 - 0.5\lambda_4) &= 0
\end{aligned}
$$

The solution for the system of equations is

$\lambda_1 = 0.25$

$\lambda_2 = 0.25$

$\lambda_3 = 0.25$

$\lambda_4 = 0.25$

$\mu_0 = (0.25\sqrt{2} - 0.5)a$

$\mu_1 = 0$

$\mu_2 = 0$

From Equation 2.13,

$$\sigma_E^2(P) = (4 \times 0.25 \times 0.5\sqrt{2} + 0.25\sqrt{2} - 0.5)a$$

$$= 0.56066a$$

The standard error is $0.74877\sqrt{a}$. For a square having a side of 1, the standard error is identical, up to the fifth significant digit, to the value given by the computer program. This is the highest value that can be obtained for any location of P inside the square. The value 0.74877 corresponds exactly with the value shown in the second column of the second row of Table B.5.

APPENDIX B

EFFICIENCY INDICES AT UNIT DENSITY

The following tables contain values of average standard error and maximum standard error assuming the sample elements and the estimated value have the same support. The sampling density is one point per square mile (0.39 point/km^2) and the semivariogram is linear with no constant term and unit slope.

TABLE B.1

AVERAGE STANDARD ERROR
THE DRIFT IS A CONSTANT
UNIT DENSITY AND UNIT LINEAR SEMIVARIOGRAM SLOPE

Pattern	Number of nearest neighbors								
	1	2	3	4	6	8	12	16	32
Hexagonal	0.83	0.67	0.63	0.63	0.63	0.63	0.63	0.63	0.63
Square	0.86	0.68	0.65	0.64	0.64	0.64	0.64	0.64	0.64
Triangular	0.88	0.74	0.69	0.68	0.67	0.66	0.66	0.66	0.66
Orthogonal traverses every 2 points	0.92	0.78	0.73	0.71	0.70	0.69	0.69	0.68	0.68
Hexagonal stratification	0.90	0.76	0.72	0.70	0.69	0.69	0.69	0.69	0.69
Square stratification	0.90	0.78	0.72	0.71	0.70	0.69	0.69	0.69	0.69
Random	0.93	0.82	0.78	0.76	0.73	0.72	0.72	0.71	0.71
Bisymmetrical random	0.95	0.82	0.78	0.75	0.73	0.72	0.72	0.72	0.72
Orthogonal traverses every 8 points	1.08	0.99	0.95	0.93	0.88	0.86	0.84	0.83	0.81
Four points per regular cluster	1.11	1.03	0.99	0.97	0.87	0.86	0.83	0.83	0.83
Five clusters	1.27	1.21	1.17	1.14	1.12	1.08	1.05	1.03	0.98
Bisymmetrical clusters	1.37	1.30	1.28	1.27	1.20	1.19	1.06	1.04	1.03
Sixteen points per regular cluster	1.49	1.39	1.37	1.36	1.35	1.32	1.27	1.24	1.13
One cluster	2.26	2.25	2.24	2.23	2.22	2.22	2.20	2.20	2.19

TABLE B.2

AVERAGE STANDARD ERROR
FIRST DEGREE POLYNOMIAL DRIFT
UNIT DENSITY AND UNIT LINEAR SEMIVARIOGRAM SLOPE

Pattern	Number of nearest neighbors							
	3	4	5	6	8	12	16	32
Hexagonal	0.65	0.64	0.63	0.63	0.63	0.63	0.63	0.63
Square	0.66	0.65	0.64	0.64	0.64	0.64	0.64	0.64
Triangular	0.80	0.69	0.67	0.67	0.66	0.66	0.66	0.66
Orthogonal traverses every 2 points	1.02	0.77	0.72	0.70	0.69	0.69	0.68	0.68
Hexagonal stratification	1.47	0.74	0.70	0.69	0.69	0.69	0.69	0.69
Square stratification	1.81	0.75	0.71	0.70	0.69	0.69	0.69	0.69
Random	1.91	0.96	0.82	0.77	0.73	0.72	0.71	0.71
Bisymmetrical random	21.92	0.84	0.77	0.74	0.72	0.72	0.72	0.72
Orthogonal traverses every 8 points	--	--	1.24	1.14	1.02	0.88	0.83	0.81
Four points per regular cluster	1.99	1.56	1.09	1.01	0.92	0.83	0.83	0.83
Five clusters	3.69	2.71	2.20	1.98	1.60	1.33	1.18	0.99
Bisymmetrical clusters	--	--	2.50	2.24	1.93	1.18	1.07	1.03
Sixteen points per regular cluster	--	--	2.96	2.75	2.29	1.83	1.62	1.17
One cluster	148.00	33.87	19.74	13.92	9.17	7.15	6.31	5.01

TABLE B.3

AVERAGE STANDARD ERROR
SECOND DEGREE POLYNOMIAL DRIFT
UNIT DENSITY AND UNIT LINEAR SEMIVARIOGRAM SLOPE

Pattern	Number of nearest neighbors						
	6	7	8	9	12	16	32
Hexagonal	0.68	0.67	0.66	0.65	0.63	0.63	0.63
Square	--	0.67	0.67	0.66	0.64	0.64	0.64
Triangular	--	0.77	0.72	0.69	0.66	0.66	0.66
Orthogonal traverses every 2 points	--	--	--	0.76	0.71	0.69	0.68
Hexagonal stratification	1.55	0.86	0.77	0.74	0.70	0.69	0.69
Square stratification	3.27	0.90	0.78	0.74	0.71	0.70	0.69
Random	26.71	1.42	1.06	0.88	0.77	0.74	0.71
Bisymmetrical random	--	--	--	0.85	0.76	0.73	0.72
Orthogonal traverses every 8 points	--	--	--	1.69	1.30	1.16	0.84
Four points per regular cluster	--	--	--	1.11	1.02	0.95	0.84
Five clusters	88.99	17.76	12.80	9.68	4.90	2.71	1.07
Bisymmetrical clusters	--	--	--	--	9.92	1.62	1.11
Sixteen points per regular cluster	--	--	--	12.94	7.25	4.78	1.53
One cluster	--	2214.00	936.00	547.00	275.00	171.00	61.50

TABLE B.4

MAXIMUM STANDARD ERROR
THE DRIFT IS A CONSTANT
UNIT DENSITY AND UNIT LINEAR SEMIVARIOGRAM SLOPE

	Number of nearest neighbors								
Pattern	1	2	3	4	6	8	12	16	32
Hexagonal	1.11	0.84	0.72	0.72	0.72	0.72	0.72	0.72	0.72
Square	1.19	0.88	0.80	0.75	0.75	0.75	0.74	0.74	0.74
Triangular	1.32	1.14	0.98	0.88	0.81	0.81	0.80	0.80	0.80
Orthogonal traverses every 2 points	1.50	1.27	1.14	1.05	0.93	0.91	0.90	0.89	0.89
Hexagonal stratification	1.30	1.07	1.02	0.94	0.87	0.86	0.86	0.86	0.86
Square stratification	1.42	1.26	1.06	0.97	0.93	0.92	0.91	0.91	0.91
Random	1.48	1.34	1.34	1.23	1.16	1.06	1.06	1.05	1.05
Bisymmetrical random	1.60	1.45	1.33	1.32	1.05	1.00	0.99	0.99	0.98
Orthogonal traverses every 8 points	2.01	1.94	1.80	1.76	1.31	1.39	1.29	1.24	1.23
Four points per regular cluster	1.60	1.50	1.41	1.40	1.17	1.12	1.00	1.00	0.99
Five clusters	1.83	1.82	1.75	1.74	1.70	1.67	1.64	1.46	1.33
Bisymmetrical clusters	1.92	1.83	1.83	1.83	1.76	1.76	1.40	1.39	1.22
Sixteen points per regular cluster	2.19	2.06	2.06	2.05	2.05	1.90	1.90	1.74	1.51
One cluster	2.99	2.98	2.98	2.98	2.97	2.96	2.96	2.95	2.94

TABLE B.5

MAXIMUM STANDARD ERROR
FIRST DEGREE POLYNOMIAL DRIFT
UNIT DENSITY AND UNIT LINEAR SEMIVARIOGRAM SLOPE

Pattern	Number of nearest neighbors							
	3	4	5	6	8	12	16	32
Hexagonal	0.73	0.73	0.72	0.72	0.72	0.72	0.72	0.72
Square	0.84	0.75	0.75	0.75	0.75	0.74	0.74	0.74
Triangular	1.47	0.93	0.83	0.81	0.81	0.80	0.80	0.80
Orthogonal traverses every 2 points	2.45	1.40	1.02	0.94	0.91	0.90	0.89	0.89
Hexagonal stratification	15.39	1.30	0.89	0.87	0.86	0.86	0.86	0.86
Square stratification	23.01	1.75	0.99	0.93	0.92	0.91	0.91	0.91
Random	19.74	3.26	1.84	1.58	1.11	1.06	1.05	1.05
Bisymmetrical random	307.00	3.02	1.36	1.14	1.00	0.99	0.99	0.98
Orthogonal traverses every 8 points	--	--	2.58	2.21	1.77	1.45	1.25	1.23
Four points per regular cluster	3.94	2.38	2.11	1.87	1.54	1.00	1.00	0.99
Five clusters	12.35	11.76	5.33	5.23	4.24	3.26	2.28	1.49
Bisymmetrical clusters	--	--	6.04	5.12	4.22	2.26	1.98	1.22
Sixteen points per regular cluster	--	--	5.80	5.48	4.17	3.72	2.90	1.85
One cluster	793.00	118.00	98.19	46.45	18.47	12.55	10.13	8.27

TABLE B.6

MAXIMUM STANDARD ERROR
SECOND DEGREE POLYNOMIAL DRIFT
UNIT DENSITY AND UNIT LINEAR SEMIVARIOGRAM SLOPE

Pattern	Number of nearest neighbors						
	6	7	8	9	12	16	32
Hexagonal	0.78	0.76	0.76	0.74	0.72	0.72	0.72
Square	--	0.79	0.78	0.77	0.75	0.74	0.74
Triangular	--	1.07	0.99	0.87	0.80	0.80	0.80
Orthogonal traverses every 2 points	--	--	--	1.27	0.97	0.90	0.89
Hexagonal stratification	39.91	1.57	1.14	1.07	0.90	0.87	0.86
Square stratification	93.17	3.07	1.42	1.12	0.96	0.93	0.91
Random	2690.00	11.60	7.90	2.03	1.32	1.15	1.05
Bisymmetrical random	--	--	--	1.96	1.25	1.07	0.98
Orthogonal traverses every 8 points	--	--	--	8.84	4.48	3.27	1.45
Four points per regular cluster	--	--	--	1.67	1.43	1.22	1.00
Five clusters	986.00	142.00	101.00	101.00	32.87	12.30	2.33
Bisymmetrical clusters	--	--	--	--	107.00	10.03	1.38
Sixteen points per regular cluster	--	--	--	33.47	23.71	12.10	5.13
One cluster	--	10425.00	8763.00	1956.00	815.00	535.00	148.00

APPENDIX C

SAMPLING DESIGN EXAMPLE

Problem:

Find the best sampling procedure for the Equus Beds which will produce estimates of the water table elevation having an average standard error of 10.8 feet (3.3 m). The solution must be restricted to the existing observation wells. Ignore border effects and possible missing wells in the selected pattern.

Solution:

From Algorithm 3.1

1. From the structural analysis (Olea, 1982a, Appendix D; Olea, 1982b, Appendix A), the relevant results are:

   a. Slope, ω = 60 feet^2 per mile (3.5 m^2/km)

   b. The drift model is a first degree polynomial within a neighborhood which is 28 miles (45 km) in diameter.

2. From the statement of the problem, the sampling efficiency index is the average standard error.

3. From the statement of the problem, regular patterns are not feasible. From Table 3.1, the irregular patterns with lowest average standard error are the stratified patterns. A hexagonal stratification is preferred over a square stratification as the former offers a lower maximum standard error. The value I(1,1) is 0.69 feet (0.21 m).

4. From the statement of the problem, $I(\omega,\rho)$ is equal to 10.8 feet (3.3 m).

5. From Equation 3.25 and the steps above

$$\rho = 60^2\left[\frac{0.69}{10.8}\right]^4$$

$$= 0.06 \text{ point per square mile } (0.023 \text{ point/km}^2)$$

6. The number of points inside a 28 mile (45 km) diameter at a density of 0.06 point per square mile (0.023 point/km^2) is

$$N = \frac{\pi d^2}{4}\rho$$

$$= \frac{3.14159 \times 28^2 \times 0.06}{4}$$

$$= 36 \text{ points}$$

7. From Table 3.3, the minimum number of nearest neighbors required to solve the universal kriging system of equations is 3, which is an order of magnitude smaller than the number of points that can be contained inside the neighborhood for which structural analysis models are valid.

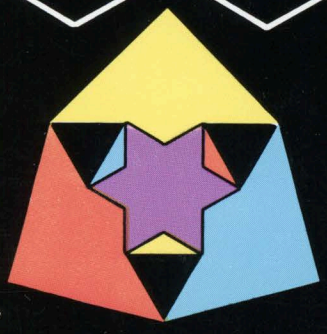
8. From Table 3.2, the minimum number of sample elements to assure a minimum average standard error is 6.

9. Since the 36 points that can be placed inside the structural analysis neighborhood is a larger number than the number of points required to assure the minimum average standard error, 6 nearest neighbors should be used in the solution of the universal kriging system of equations.

Hence, the best irregular sampling pattern for the water table elevation in the Equus Beds is a hexagonal stratified pattern. A density of 0.06 point per square mile (0.023 point/km^2) assures an average standard error of 10.8 feet (3.3 m), missing wells in the stratification and border effects not included. The number of sample elements to be used in the universal kriging system of equations should be 6.

SYSTEMATIC SAMPLING OF SPATIAL FUNCTIONS

R. A. OLEA



NUMBER SEVEN
SERIES ON SPATIAL ANALYSIS
KANSAS GEOLOGICAL SURVEY
1930 CONSTANT AVENUE
LAWRENCE, KANSAS USA 66044

KGS 1984