I. INTRODUCTION

The precision of statements that can be made about soil properties at any location depends largely on the amount of variation within the area sampled. As heterogeneity of soils increases, the precision of statements about their

I. INTRODUCTION

The precision of statements that can be made about soil properties at any location depends largely on the amount of variation within the area sampled. As heterogeneity of soils increases, the precision of statements about their
properties, behavior, and land use performance decreases. Soil scientists are restricted to limited observations of the earth's surface, necessitating the extrapolation of soil properties from locations where they are known to others where they are not known. The precision of such extrapolation is strongly influenced by the variability of soils both within sampling units and between locations. Similarly, the likelihood of successful transfer of land use experience from known to unknown sites is strongly influenced by the spatial and temporal variability of soil and other environmental parameters.

Recognition of the importance of spatial variability on land use has led to study of soil heterogeneity, ranging from a global scale (Food and Agricultural Organization, 1974) to changes in structural and chemical composition of soil minerals on a microscale (Sawhney, 1977). The nature of soil variability is itself scale-dependent because the soil-forming factors and processes interact over many different spatial and temporal scales (Burrough, 1983a). As a result, the nature of variability identified by spatial studies of soil properties depends largely on the scale of observation, the properties in question, and the methodology used to conduct the investigation (Wilding and Drees, 1983).

Soil classification and soil survey have traditionally been the most practical approaches to grouping similar and separating different soils on a regional scale. Variability of properties within soil mapping units and within smaller sampling units, such as fields, experimental plots, or pedons, is acknowledged and has often been described by classical statistical methods (Beckett and Webster, 1971; Wilding and Drees, 1978, 1983). Classical statistics assumes that the sampling unit mean is the expected value everywhere in the unit, with an estimation error expressed by the within-unit variance. This approach assumes that variability about the mean is random and contains no reference to the geographical distribution of differences within the sampling units. Several studies have shown that this "random" aspect of soil variability often contains a component that is spatially dependent (Campbell, 1978; Burgess and Webster, 1980a; Gajem et al., 1981; Vieira et al., 1981; Yost et al., 1982a; McBratney et al., 1982). This implies that within a given distance or range of spatial dependence differences in soil properties can be described as a function of their spatial separation. Traditional methods of soil classification and statistical analysis do not directly consider this aspect.

Recent developments in statistical theory enable spatial relationships among sample values to be quantified and used for interpolation of values at unsampled locations. These developments are based on the theory of regionalized variables (Matheron, 1963, 1965). This theory (under the name of "geostatistics") takes into account both the structured and random characteristics of spatially distributed variables to provide quantitative tools for their description and optimal, unbiased estimation. Geostatistical analysis
of spatial variability has been extensively applied to estimation of ore reserves in the mining industry (Journel and Huijbregts, 1978; David, 1977; Clark, 1979), water resources research (Delfiner and Delhomme, 1973; Delhomme, 1978, 1979), soil science (Burgess and Webster, 1980a,b; Vieira et al., 1981; Yost et al., 1982a,b), and archaeology (Zubrow and Harbaugh, 1979).

This article briefly reviews some of the traditional methods of describing soil variability, discusses geostatistical approaches to quantifying spatial dependence and their use for interpolation under different kinds of spatial variation, and, lastly, identifies some future applications of geostatistics to spatial studies in soil and agronomic research.

II. NATURE OF SOIL VARIABILITY

A. SYSTEMATIC AND RANDOM VARIATION

Spatial variation of soil properties contains systematic and random components. Wilding and Drees (1983) express systematic variability as gradual or distinct changes (or trends) in soil properties that can be understood in terms of soil-forming factors or processes at a given scale of observation. Sources of systematic variation may range from differences in topography, lithology, climate, biological activity, and age of soils in regional studies (Van Wambeke and Dudal, 1978) to differences in microfabric and physicochemical composition when soils are observed on a micro level (Blevins et al., 1970; Miller et al., 1971; Murphy and Banfield, 1978).

Associated with systematic variation are differences in observed soil properties which cannot be related to a known cause. In addition, there are often spatial, temporal, and measurement sources of variation which cannot be discerned by the nature or scale of the investigation (Ball and Williams, 1968). This unexplained heterogeneity is termed "random" or "chance" variation by Wilding and Drees (1983) or "noise" by Webster and Cuanalo (1975) and Burrough (1983a).

Clearly, as the soil body is studied in greater detail, heterogeneity that may have been previously defined as random variation may be found to contain a systematic component.

B. NESTED EFFECTS

Soil variability is the product of soil-forming factors operating and interacting over a continuum of spatial and temporal scales. Processes which
operate over large distances (e.g., climate) or long time periods (e.g., soil weathering) are modified by other processes which operate more locally (e.g., erosion and deposition of parent materials) or more frequently (e.g., weather). This nested nature of soil variation implies that the kind and causes of heterogeneity that are identified in variability studies depends largely on the scale or frequency of observation. Multistage sampling has often been necessary to effectively describe different orders of soil variation within field areas (Hammond et al., 1958; Nortcliff, 1978; Burrough, 1983a). Similarly, many soil properties (e.g., temperature, moisture content) have been measured over many different time intervals to determine their temporal fluctuation.

The change in spatial variability with increasing scale factor depends on the soil property in question and the soil factors determining spatial change (Wilding and Drees, 1983). Total variance will increase as sampling area increases (Beckett and Webster, 1971), but relative contributions of variance at different scales to the total variance follow no consistent pattern (Wilding and Drees, 1983). Much of the variability for some properties may occur over short distances within sampling units (McIntyre, 1967; Protz et al., 1968; Beckett and Webster, 1971), while for others a large proportion of the total variance is caused by long-range differences (Webster and Butler, 1976). The change in variability with increasing scale factor may be linear, curvilinear (upward or downward), or irregular, where different soil processes exert dominating effects over different spatial scales (Webster and Butler, 1976; Nortcliff, 1978; Burrough, 1983a).

In their study of changes in soil properties over a range of sampling distances, Webster and Butler (1976) found that most of the within-field variance of phosphorus occurred within distances of 5 m; of bulk density and water content, over 18 m; of soluble potassium, over 56–180 m; of pH, over 56 and 180 m; and of morphological properties, over 5 and 180 m. Schafer (1979) found that in natural soils 50–75% of the total variation in soil texture, color, root abundance, A horizon thickness, and depth to carbonates occurred at distances >500 m, while soils formed in recent mine spoils showed similar magnitudes of variation at <10 m. Differences in properties of natural soils were correlated to systematic variation of geomorphic factors, while soils from mine spoils were dependent on random variation in parent material composition as a function of mining and reclamation methods.

Babaloa (1978) found that large variation of hydraulic conductivity in a 0.3-ha plot relative to a 92-ha field was caused by local changes in particle size distribution and bulk density. This short-range variability of soil hydraulic properties is considered particularly important by Wilding and Drees (1983) and Bouma (1983) because many soil processes are dependent on rates and directions of water movement and chemical transport. Beckett
and Webster (1971) also acknowledged the importance of short-range variation by concluding that much of the variability within any soil sampling unit may be present within any square meter of it.

III. TRADITIONAL METHODS OF DESCRIBING SOIL VARIABILITY

Soil classification, its field application through soil survey operations, and statistical analysis have been the most commonly used methods for describing soil variation. These methods have been comprehensively reviewed by Beckett and Webster (1971) and Wilding and Drees (1978, 1983), so only their main concepts are discussed here.

A. SOIL CLASSIFICATION AND SOIL SURVEY

Soil classification and soil survey have been the most commonly used methods for partitioning field variation on a regional scale. By grouping soils that are similar and separating those that are different, this approach also forms the basis for establishing relationships between individual soils, predicting properties at unsampled locations, predicting soil behavior, and identifying potential uses (Buol et al., 1980).

One of the assumptions made in soil classification systems, and in soil survey practice, is that soil differences can be adequately characterized by relatively few diagnostic properties. The diagnostic properties that are used to define categories and taxa of soil classification systems are ideally those which have greatest independence of variation from each other but which have high covariance with many other nondiagnostic properties. This results in the variance within taxonomic units measured over all properties being minimized with respect to their total variance (Norris, 1971). This condition applies to all levels of hierarchical classification systems. Successively lower levels of classification are designed to partition more of the total variation, thus creating more homogeneous classes. Several statistical studies have shown that the highest categories of soil classification systems tend to account for much of the variance in specific properties that is explained by classification and that lower categories only contribute small additional amounts to improving the homogeneity of taxa (Beckett and Webster, 1971; Yost and Fox, 1981; Trangmar, 1984). Small variance reductions at lower levels often result from correlation of diagnostic criteria at lower levels with
those of the more general, higher levels (Yost and Fox, 1983; Trangmar, 1984).

Mapping the spatial distribution of soil taxonomic units involves systematically partitioning the landscape into soil mapping units that are reasonably homogeneous and that can be readily portrayed at the mapping scale used. Taxonomic impurity of soil mapping units is acknowledged, and the proportion of inclusions are commonly specified or estimated in soil survey reports. The proportion of inclusions allowed in simple soil mapping units ranges from 15% (Soil Survey Staff, 1951; Taylor and Pohlen, 1962) to 35% (Mapping Systems Working Group, 1981) before compound units are established.

Constraints of time and sampling effort usually restrict the soil surveyor to only a few field observations per mapping unit, with the result that heterogeneity may often exceed the desired limits. In their review of this topic, Beckett and Webster (1971) concluded that simple mapping units might actually average only 50% purity. Burrough et al. (1971) found that mapping unit purity varied with map scale and observation density. Purity ranged from 45–63% at a scale of 1:63,360 and 65–87% at a scale of 1:25,000. This apparently high degree of taxonomic variability within mapping units is often diminished in importance because impurities often differ only in minor definitive features and do not require different management (Bascomb and Jarvis, 1976).

If compound units are established, soil associations are used to portray groups of geographically associated soils (each of which is confined to a particular facet of the landscape) which occur in a predictable pattern (Dent and Young, 1981). Associations can be resolved into simple mapping units at a more detailed scale of investigation. If the soil pattern cannot be resolved because of its intricacy, it is mapped as a complex. Soil survey reports generally describe members of associations and complexes and indicate their relative proportions in the mapping unit. The areas of individual members within an association may be large enough to be managed separately from other members. Complexes generally have to be managed as complexes because of the small land areas covered by individual members (Cutler, 1977).

The importance of variation in soil properties depends on the kind and intensity of land use of the area in question. Clearly, soil properties differ in their effects on different kinds of land use, and some specific chemical or physical properties may have more dominance than others. As a result, spatial variability of specific properties within mapping units is also of considerable interest to the map user.

The variability of diagnostic properties of mutually exclusive taxonomic units is fixed and their distribution is truncated by limits between taxa. Not surprisingly, studies have shown that variability of such properties is smaller
within pedons than within series than within the corresponding mapping units (Beckett and Webster, 1971; Beckett and Burrough, 1971; Wilding and Drees, 1978, 1983).

The variability of most properties is usually less within mapping units than between units (Wilding et al., 1965), although where variable levels of management have been applied, within-unit variation may exceed that between units (McCormack and Wilding, 1969; Beckett and Webster, 1971). The Benchmark Soils Project found that variability of soil properties within the same soil family of Soil Taxonomy was sufficiently low to support the hypothesis that soils of the same family have similar responses to similar management practices (Silva, 1984).

Properties most affected by soil management (e.g., soluble phosphorus, exchangeable cations, sulphate-S, total sulfur) are commonly more variable than the morphological (e.g., color, A horizon thickness), physical (e.g., particle size, bulk density), and chemical (e.g., pH) properties used to define taxonomic units (Beckett and Webster, 1971; Adams and Wilde, 1976a,b; Wilding and Drees, 1978, 1983). In their summary paper, Wilding and Drees (1983) give mean coefficients of variation (CVs) for exchangeable calcium, magnesium, and potassium of 50–70% ranging up to 160%. They also concluded that the variability of physical properties such as Atterberg limits, particle size fractions, bulk density and water content (CVs of 10–53%) is often much less than hydraulic conductivity (CVs of 50–150%) measured over the same area.

As a result of such variation within sampling units, soil surveys cannot be expected to reliably predict variation of all properties, particularly those that are easily influenced by soil management.

B. Statistical Analysis

The application of statistics of soil variation has been summarized by Beckett and Webster (1971) and Wilding and Drees (1978, 1983). A more comprehensive treatment of the topic can be found in Webster (1977).

Classical statistics assumes that the expected value of a soil property $z$ at any location $x$ within a sampling area is

$$z(x) = \mu + \epsilon(x)$$

where $\mu$ is the population mean or expected value of $z$ and $\epsilon(x)$ represents a random, spatially uncorrelated dispersion of values about the mean. Deviations from the population mean are assumed to be normally distributed with a mean of zero and a variance of $\sigma^2$ (Sokal and Rohlf, 1969).
Many soil properties have skewed probability distributions and require transformation (e.g., natural log) to the normal distribution prior to statistical analysis (Cassel and Bauer, 1975; Wagenet and Jurinak, 1978). Other properties may be bimodally distributed (Smeck and Wilding, 1980), in which case each mode may be treated as a separate population for statistical analysis (Wilding and Drees, 1983).

Because mean values are used for estimation of properties at unsampled locations within sampling units, statistics of dispersion (e.g., coefficients of variation, standard deviation, standard error, confidence limits) are used to indicate precision of the mean as an estimator. These statistics have been used extensively to document the variation of soil properties within sampling areas such as soil mapping units (Wilding et al., 1965; McCormack and Wilding, 1969; Adams and Wilde, 1976a,b), fields (Cassel and Bauer, 1975; Biggar and Nielsen, 1976), experimental plots (Jacob and Klute, 1956; Nielsen et al., 1973), and pedons (Smeck and Wilding, 1980). Analysis of variance and subsequent statistical testing has been a common method for comparing variation among sampling units (Jacob and Klute, 1956; Cassel and Bauer, 1975; McBratney et al., 1982).

The influence of random sources on variance within sampling units has prompted much research into the sampling size required to estimate the sample mean at various levels of precision and confidence intervals (Ball and Williams, 1968; Beckett and Webster, 1971; Cassel and Bauer, 1975; Biggar and Nielsen, 1976; Adams and Wilde, 1976b). As within-unit variance increases, a proportionately larger number of samples is required to estimate the mean for a given level of confidence.

Classical statistical procedures assume that variation is randomly distributed within sampling units. Actually, soil properties are continuous variables whose values at any location can be expected to vary according to direction and distance of separation from neighboring samples (Burgess and Webster, 1980a). By so varying, soil properties exhibit spatial dependence within some localized region. Estimation using the classical model cannot be improved on if the initial classification of a region into discrete sampling or mapping units accounts for all the spatially dependent variance (McBratney et al., 1982). However, spatial dependence of soil properties will usually occur in most sampling units. The classical model is inadequate for interpolation of spatially dependent variables, because it assumes random variation and takes no account of spatial correlation and relative location of samples.

Several techniques which incorporate sample location to varying degrees have been used for interpolation of soil properties. These include proximal weighting (Van Kuilenburg et al., 1982), moving averages (Webster, 1978), weighted moving averages using inverse distance and inverse distance squared functions (Van Kuilenburg et al., 1982), trend surface analysis
These techniques are empirical, and although they may seem reasonable for many applications, they are theoretically unsatisfactory (Burgess and Webster, 1980a). Some provide good interpolation under optimal data configurations, but most give biased estimates that are not optimal; many do not provide estimates of the interpolation error and those that do, do not attempt to minimize that error (Burgess and Webster, 1980a).

**IV. REGIONALIZED VARIABLE THEORY AND GEOSTATISTICS**

Recent developments in statistical theory enable spatial dependence of soil properties to be directly considered in interpolation. These developments are based on the *theory of regionalized variables*, which takes into account both the random and structured characteristics of spatially distributed variables to provide quantitative tools for their description and optimal, unbiased estimation. These tools can augment the more commonly used methods in analysis of soil variability.

**A. DEVELOPMENT OF GEOSTATISTICS**

Interpolation based on spatial dependence of samples was first used by D. G. Krige (1951, 1960) for estimation of the gold content of ore bodies in the mining industry of South Africa. Classical statistical interpolation procedures were considered inappropriate in the mining industry because they were biased and nonoptimal in that they did not take local spatial dependence into account during estimation. Interpolation procedures which considered local changes in ore content and grade were developed to obtain a method which would enable optimal sample placement to minimize the high cost of sampling mineral deposits.

Krige’s practical methods were generalized and extended by Matheron (1963, 1965, 1969, 1970, 1971) into the theory of regionalized variables. This theory now forms the basis of procedures for analysis and estimation of spatially dependent variables. These procedures are known collectively as *geostatistics*. Blais and Carlier (1968) and Huijbregts and Matheron (1971) were among the first to apply *kriging* as an estimation procedure in mining engineering.

Geostatistical theory continued to develop in the 1970s to include analysis of variables having lognormal (Rendu, 1979; Journel, 1980) or complex
(Matheron, 1976; Journel and Huijbregts, 1978; Jackson and Marechal, 1979) probability distributions and estimation in the presence of trends (Olea, 1974, 1975; Delfiner, 1976; Journel and Huijbregts, 1978). While the use of geostatistics has centered on the mining industry, it is now being used extensively in water resources research (Delfiner and Delhomme, 1973; Delhomme, 1978, 1979), soil science (Campbell, 1978; Burgess and Webster, 1980a,b; Vieira et al., 1981; Yost et al., 1982a,b), and archaeology (Zubrow and Harbaugh, 1979).

B. Theory of Regionalized Variables

Geostatistics are based on the concepts of regionalized variables, random functions, and stationarity. A brief theoretical discussion of these concepts is necessary to appreciate the practical application of geostatistics to the analysis of soil variation. Comprehensive coverage of regionalized variable theory and its geostatistical applications are given by David (1977), Journel and Huijbregts (1978), Clark (1979), and Royle (1980).

1. Regionalized Variables and Random Functions

A random variable is a measurement of individuals that is expected to vary according to some probability distribution law (Henley, 1981). The random variable is characterized by the parameters of the distribution, such as the mean and variance of the normal distribution. A regionalized variable \( z(x) \) is a random variable that takes different values \( z \) according to its location \( x \) within some region (Journel and Huijbregts, 1978). As such, a regionalized variable \( z(x) \) can be considered as a particular realization of a random variable \( Z \) for a fixed location \( x \) within the region. If all values of \( z(x) \) are considered at all locations within the region, then the regionalized variable \( z(x) \) becomes a member of an infinite set of random variables \( Z(x) \) for all locations within the region. Such a set is called a random function because it associates a random variable \( Z \) with any location \( x \) (Huijbregts, 1975).

2. Stationarity

A random function \( Z(x) \) is said to be first-order stationary if its expected value is the same at all locations throughout the study region,

\[
E[Z(x)] = m
\]  
(2)
where \( m \) is the mean of classical statistics, and

\[
E[Z(x) - Z(x + h)] = 0
\]

(3)

where \( h \) is the vector of separation between sample locations.

Second-order stationarity applies if the spatial covariance \( C(h) \) of each \( Z(x) \) and \( Z(x + h) \) pair is the same (independent of position) throughout the study region and depends on \( h \):

\[
C(h) = E[Z(x) - m][Z(x + h) - m]
\]

(4)

As \( h \) gets larger, \( C(h) \) decreases and the spatial covariance decays (Fig. 1). Stationarity of \( C(h) \) implies stationarity of the sample variance \( s^2 \). The spatial covariance will approach the sample variance as the distance of separation tends to zero.

Second-order stationarity does not apply if a finite variance and covariance cannot be defined, as in the case of trend phenomena (David, 1977), and a weaker form of stationarity called the intrinsic hypothesis must be assumed (Journel and Huijbregts, 1978). Second-order stationarity implies the intrinsic hypothesis, but not the converse. The intrinsic hypothesis requires that for all vectors of \( h \), the variance of the increment \( Z(x) - Z(x + h) \) be finite and independent of position within the region, i.e.,

\[
\text{VAR}[Z(x) - Z(x + h)] = E[Z(x) - Z(x + h)]^2 = 2\gamma(h)
\]

(5)

Dividing by two yields the semi-variance statistic \( \gamma(h) \). The semi-variance \( \gamma \) depends on the vector of separation \( h \). Ideally, \( \gamma \) is zero at \( h = 0 \), but increases as \( h \) increases (Fig. 1).

---

**Fig. 1**. Relationship between the spatial covariance \( C(h) \) and the semi-variogram statistic \( \gamma(h) \). (From Journel and Huijbregts, 1978.)
V. ANALYSIS OF SPATIAL DEPENDENCE

The concepts of regionalized variables and stationarity provide the theoretical basis for analysis of spatial dependence using autocorrelation or semi-variograms.

A. AUTOCORRELATION

Analytical functions express the linear correlation between a spatial series and the same series at a further distance interval (Vieira et al., 1981). Their definition assumes second-order stationarity, in which case the autocorrelation is expressed as

\[ r(h) = \frac{C(h)}{s^2} \]  

where \( r(h) \) is the autocorrelation among samples at distance of separation, or lag, \( h \). A plot of the autocorrelation values \( r(h) \) versus the lag is called the autocorrelogram. The maximum value of \( r(h) \) is 1 at zero distance (\( h = 0 \)), and values decrease with increasing \( h \). The distance \( a \) at which \( r(h) \) no longer decreases defines the range over which samples of the variable are spatially dependent.

Values of the autocorrelation function are normalized to the range from \(-1\) to \(1\) inclusive, making for easy interpretation of data values. The mean, variance, and autocorrelation function completely characterize the random function \( Z(x) \), where \( Z(x) \) is normally or lognormally distributed (Gajem et al., 1981).

Autocorrelograms have been used to express spatial changes in field-measured soil properties and the degree of dependency among neighboring observations (Webster, 1973, 1978; Webster and Cuanalo, 1975; Vieira et al., 1981; Sisson and Wierenga, 1981). Such information aids identification of the maximum sampling distance for which observations remain spatially correlated and can be used in designing soil sampling schemes (Vieira et al., 1981) or defining minimum cell size for interpolation by moving average techniques (Webster, 1978). Webster and Cuanalo (1975) used autocorrelation analysis of soil chemical properties sampled along transects to locate soil boundaries. Russo and Bresler (1981) found that ranges of spatial dependence for soil moisture characteristics decreased with depth, indicating greater continuity of these properties in topsoils than in subsoils. Spatial analysis of soil properties using autocorrelograms has been restricted to data sampled at regular spacings along transects (Webster and Cuanalo, 1975; Gajem et al., 1981) or grids (Vieira et al., 1981).
Soil properties which do not show second-order stationarity do not have finite variances over the distance between sample locations, making it impossible to define the autocorrelation function (David, 1977). This nonstationarity can be removed by detrending, but it is often more convenient to assume the intrinsic hypothesis and use semi-variograms for quantifying spatial dependence (Vieira et al., 1981).

B. SEMI-VARIOGRAMS

1. Assumptions and Definitions

Structural analysis of spatial dependence using semi-variograms can be made using weaker assumptions of stationarity than are necessary for autocorrelation. Semi-variogram analysis has the added advantage of defining parameters needed for local estimation by kriging (Section VI).

The semi-variance statistic $\gamma(h)$ can be defined in terms of the variance $s^2$ and spatial covariance $C(h)$ of $Z(x)$ if second-order stationarity applies, i.e.,

$$\gamma(h) = s^2 - C(h)$$

Alternatively, the weaker intrinsic hypothesis can be assumed (Section IV,B).

The semi-variance $\gamma(h)$ describes the spatially dependent component of the random function $Z$. It is equal to half the expected squared distance between sample values separated by a given distance $h$, i.e.,

$$\gamma(h) = E[Z(x) - Z(x + h)]^2$$

Application of regionalized variable theory assumes that the semi-variance between any two locations in the study region depends only on the distance and direction of separation between the two locations and not on their geographic location. Based on this assumption, the average semi-variogram for each lag can be estimated for a given volume of three-dimensional space.

The semi-variance at a given lag $h$ is estimated as the average of the squared differences between all observations separated by the lag:

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

where there are $N(h)$ pairs of observations. The semi-variogram for a given direction is usually displayed as a plot of semi-variance $\gamma(h)$ versus distance $h$ (Fig. 2A).
Fig. 2. (A) Idealized semi-variogram with zero nugget variance and (B) observed semi-variograms for soil properties with nugget variance. (From Wilding and Drees, 1983.)

2. Parameters

The shape of the experimental semi-variogram may take many forms, depending on the data and sampling interval used. Ideally, the semi-variance increases with distance between sample locations, rising to a more or less constant value (the sill) at a given separation distance, called the range of spatial dependence,  \( a \) (Fig. 2A). The sill approximates the sample variance \( s^2 \) for stationary data. Samples separated by distances closer than the range are spatially related. Those separated by distances greater than the range are not spatially related because the semi-variance equals \( s^2 \), implying random
### Table I

Parameter Values of Some Isotropic Semi-variograms for Soils and Related Data

<table>
<thead>
<tr>
<th>Property</th>
<th>Sample spacing (m)</th>
<th>Range (m)</th>
<th>Nugget variance (% of sill)</th>
<th>Model&lt;sup&gt;a&lt;/sup&gt;</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>pH</td>
<td>0.5</td>
<td>4</td>
<td>4</td>
<td>S</td>
<td>Trangmar (1984)</td>
</tr>
<tr>
<td></td>
<td>20</td>
<td>320</td>
<td>23</td>
<td>S</td>
<td>McBratney and Webster (1981b)</td>
</tr>
<tr>
<td></td>
<td>1,000</td>
<td>14,000</td>
<td>23</td>
<td>M</td>
<td>Yost&lt;sup&gt;et al.&lt;/sup&gt; (1982a)</td>
</tr>
<tr>
<td>Exchangeable aluminum (meq/100 g)</td>
<td>0.5</td>
<td>4</td>
<td>26</td>
<td>S</td>
<td>Trangmar (1984)</td>
</tr>
<tr>
<td></td>
<td>500</td>
<td>4,200</td>
<td>63</td>
<td>S</td>
<td>Trangmar (1984)</td>
</tr>
<tr>
<td>Sodium content (meq/10 kg)</td>
<td>1.5</td>
<td>6</td>
<td>56</td>
<td>L</td>
<td>Burgess and Webster (1980a)</td>
</tr>
<tr>
<td>Phosphorus sorbed (ppm) at 0.2 mg P/liter</td>
<td>1,000</td>
<td>32,000</td>
<td>25</td>
<td>M</td>
<td>Yost&lt;sup&gt;et al.&lt;/sup&gt; (1982a)</td>
</tr>
<tr>
<td>Sand (%)</td>
<td>0.5</td>
<td>4</td>
<td>3</td>
<td>S</td>
<td>Trangmar (1984)</td>
</tr>
<tr>
<td></td>
<td>10</td>
<td>34</td>
<td>44</td>
<td>S</td>
<td>Vauclin&lt;sup&gt;et al.&lt;/sup&gt; (1983)</td>
</tr>
<tr>
<td>Bulk density (g/cc)</td>
<td>0.2</td>
<td>6</td>
<td>30</td>
<td>L</td>
<td>Gajem&lt;sup&gt;et al.&lt;/sup&gt; (1981)</td>
</tr>
<tr>
<td>Loam thickness (g/cc)</td>
<td>20</td>
<td>100</td>
<td>24</td>
<td>S</td>
<td>Burgess and Webster (1980a)</td>
</tr>
<tr>
<td>Rice grain yield (g/m&lt;sup&gt;2&lt;/sup&gt;)</td>
<td>0.5</td>
<td>18</td>
<td>55</td>
<td>S</td>
<td>Trangmar (1984)</td>
</tr>
<tr>
<td>Leaf phosphorus content (%) in sorghum</td>
<td>1.5</td>
<td>6</td>
<td>40</td>
<td>L</td>
<td>Trangmar (1982)</td>
</tr>
</tbody>
</table>

<sup>a</sup> Semi-variogram model: S, spherical; M, Mitscherlich; L, Linear.

Variation. The range also defines the maximum radius from which neighboring samples are drawn for interpolation by kriging (Section VI).

Semi-variogram ranges depend on the scale of observation and the spatial interaction of soil processes affecting each property at the sampling scale used. Reported ranges of spatial dependence of soil properties vary from 0.6 m for 15-bar water sampled at 0.2-m intervals (Gajem<sup>et al.</sup>, 1981) to 58 km for phosphorus sorbed at 0.2 mg P/liter sampled at 1–2-km intervals (Yost<sup>et al.</sup>, 1982a). Some ranges of semi-variograms for soil properties are given in Table I. An example of a well-structured semi-variogram is given in Fig. 3.

Semi-variances may also increase continuously without showing a definite range and sill, thus preventing definition of a spatial variance, indicating the presence of trend effects and nonstationarity (Webster and Burgess, 1980; Gajem<sup>et al.</sup>, 1981; Yost<sup>et al.</sup>, 1982b). Other semi-variograms show a
complete absence of spatial structure, implying that there is no easily quantifiable spatial relationship between sample values at the sampling scale used.

Ideally, the experimental semi-variogram should pass through the origin (Fig. 2A) when the distance of sample separation is zero. However, many soil properties have nonzero semi-variances as \( h \) tends to zero (Fig. 2B). This nonzero variance is called the "nugget variance" or "nugget effect" (Journel and Huijbregts, 1978). It represents unexplained or "random" variance, often caused by measurement error or microvariability of the property which cannot be detected at the scale of sampling. Some reported values of semi-variogram nugget variances are given in Table I.

The sum of the nugget variance \( C_0 \) and the spatial covariance \( C \) approximately equals the sill or sample variance \( s^2 \) for stationary data (Fig. 2B). The nugget variance can also be expressed as a percentage of the sill value (Table I) to enable comparison of the relative size of the nugget effect among properties (Yost et al., 1982a; Burrough, 1983a; Trangmar, 1984). Nugget variances of soil properties ranging from 0 (Vieira et al., 1981) up to 100% of the sill (Campbell, 1978; Luxmoore et al., 1981; Hajrasuliha et al., 1980) have been reported. A nugget variance of 0% of sill means that there is neither measurement error nor significant short-range variation present.

The experimental semi-variogram exhibits pure nugget effect (100% of sill) when \( \gamma(h) \) equals the sill at all values of \( h \). Pure nugget effect arises from very large point-to-point variation at short distances of separation and indicates a

![Figure 3. Example of a semi-variogram (for pH). (From Trangmar et al., 1984.)](image-url)
total absence of spatial correlation at the sampling scale used. Increasing the detail of sampling will often reveal structure in the apparently random effects of the nugget and pure nugget variances (Burrough, 1983a). According to Journel and Huijbregts (1978), a pure nugget effect at all scales of sampling amounting to a single discontinuity at the origin is exceptional. If this occurs, it implies that the mean is the best estimator at every point in the study region.

Part of the nugget variance may be caused by measurement and sampling error, so it also sets a lower limit to the precision of the sampling or measurement technique used (Burrough, 1983a). The size of the measurement error component is indicated if the nugget variance cannot be reduced by collecting additional samples at closer spacings. The true spatial component \( C \) of the sample variance is then also clearly defined (Fig. 2B). The magnitude of the nugget variance is important in kriging because it sets a lower limit to the size of the estimation variance and, therefore, to the precision of the interpolation.

Figure 4 presents a set of idealized semi-variograms that commonly occur for soil properties. If short-range effects predominate, the semi-variogram has a large nugget variance (curve 1), or if pure nugget effect occurs a straight line equal to the sill would be present. If a single, long-range process dominates, the semi-variogram is linear up to the sill, where it abruptly flattens out (curve 2). If several processes make important contributions to spatial dependence at different scales, the semi-variogram consists of several linear
portions, separated by marked slope changes at sampling intervals corresponding to the range of the soil process in question (curve 3). If several processes with similar contributions act over closely related scales, the resulting semi-variogram consists of a set of straight lines approximating a curve (curve 4). It is very difficult, if not impossible, to identify the relative contributions of each process for curves like type 4.

3. Estimation of Parameters

Parameters of experimental semi-variograms are commonly estimated using least squares regression, weighted for the number of sample pairs in each lag (Vieira et al., 1981; Yost et al., 1982a; Trangmar, 1984). This approach usually gives an adequate first approximation of semi-variogram model fitting against which the deviations of individual semi-variances from the overall structure can be assessed by critical review of the data. Minor errors in estimation of semi-variogram parameters make little difference to the reliability of interpolation because of the robustness of the kriging technique (David, 1977).

The equations most commonly used to estimate parameters of isotropic or unidirectional semi-variograms are the linear equation (Burgess and Webster, 1980a; Hajrasuliha et al., 1980; Vauclin et al., 1983) as in Fig. 4, curves 1 and 2, and a segmented quadratic form known as the spherical model (Burgess and Webster, 1980a; Vieira et al., 1981; Van Kuilenburg et al., 1982; Vauclin et al., 1983; Trangmar, 1984) as in Fig. 4, curve 4. A Mitscherlich model was also used by Yost et al. (1982a) for estimating semi-variogram parameters. Segmented models such as the double spherical model of McBratney et al. (1982) have been used to estimate semi-variograms in which breaks in slope mark different ranges of spatial dependence associated with different soil processes (Fig. 4, curve 3).

Other semi-variogram models that have been used in mining geostatistics (David, 1977), but which have not been used in soil science, include the De Wijsian (the linear model with the lag plotted on a log scale), the exponential (asymptotic convergence with the sill), and the “hole effect” model (for estimation of periodic semi-variances). The mathematical forms and detailed descriptions of the various models can be found in David (1977) and Journel and Huijbregts (1978).

It is important to choose the appropriate model for estimating the semi-variogram because each model yields quite different values for the nugget variance and range, both of which are critical parameters for kriging. The Mitscherlich and exponential forms have rarely been used because their infinite ranges imply very continuous processes (Journel and Huijbregts, 1978), which rarely occur in ore bodies or field soils. Yost et al. (1982a) found that an appropriate working range for the Mitscherlich form coincided with
the distance of separation at which the semi-variance equals 95% of the sill. When fitted to the same experimental semi-variogram, the spherical model generally gives longer ranges and smaller nugget variances than the linear form but yields shorter ranges and larger nugget variances than the Mitscherlich form. Over intermediate lags there is little difference between the spherical or Mitscherlich model in estimating the semi-variance.

4. Sampling

Choice of configuration and minimum spacing of samples for semi-variogram analysis has generally been based on the previous knowledge of variation within the study area, the objective of the study, and the costs of sampling and measurement. Sampling designs used for analysis of spatial dependence have included point samples collected along transects with regular (McBratney and Webster, 1981b; Gajem et al., 1981) or irregular spacings (Yost et al., 1982a), equilateral grids (Campbell, 1978; Burgess and Webster, 1980a; Hajrasuliha et al., 1980; Trangmar, 1982), equilateral grids with sampling at some shorter spacings in some “window areas” (Trangmar, 1984), and random sampling (Van Der Zaag et al., 1981; McBratney et al., 1982; Van Kuilenburg et al., 1982). Bulking of soil samples from within grid cells (Burgess and Webster, 1980a,b; McBratney and Webster, 1981a; Webster and Burgess, 1984) and areal measurements of crop parameters (Tabor et al., 1984; Trangmar, 1984) have also been made for semi-variogram analysis where spatial interpolation by block kriging is the study objective.

McBratney and Webster (1983b) suggest that for soil mapping purposes transect sampling can be used to obtain a working semi-variogram to initially identify spatial dependence parameters. This could then be used to design an optimal sampling scheme for kriging (Section VI,C), if necessary, and would only require a fairly small proportion of the total sampling effort needed for kriging. They also suggest that if mean estimation variances or standard errors of within-sampling unit variation are required, then regular grid sampling may be the best strategy, with the interval determined by the number of observations that can be afforded. In our experience, it seems desirable to collect a number of samples at distances smaller than the smallest grid spacing to reliably estimate the semi-variogram at short lags and to reduce the size of the nugget variance (Trangmar, 1984).

5. Interpretation of Semi-variograms

Analysis of spatial dependence using semi-variograms has contributed to our understanding of many aspects of soil variability, genesis, management, and interpretation. This section discusses some of these applications.
a. Isotropic and Anisotropic Variation. Soil properties are isotropic if they vary in a similar manner in all directions, in which case the semi-variogram depends only on the distance between samples, \( h \). One semi-variogram applies to all parts of the study region and defines a circular range of spatial dependence about each sample location.

Geometrical anisotropy occurs when variations for a given distance \( h \) in one direction are equivalent to variations for a distance \( kh \) in another direction. The anisotropy ratio \( k \) indicates the relative size of directional differences in variation. It characterizes an ellipsoidal zone of influence which is elongated in the direction of minimum variation. The direction of maximum variation is assumed to occur perpendicular to the direction of minimum variation (David, 1977). The anisotropy ratio would equal 1 and define a circular zone of influence if variation were the same in all directions, i.e., isotropic.

Differences in slopes of individual semi-variograms computed in different directions reveal the presence or absence of anisotropic spatial dependence (Webster and Burgess, 1980; Burgess and Webster, 1980a; McBratney and Webster, 1981a, 1983a; Trangmar, 1984; Tabor et al., 1984). If anisotropy occurs, the semi-variogram computed in the direction of maximum variation will have the steepest slope, while that in the direction of minimum variation will have the lowest slope.

Parameters of geometric anisotropic spatial dependence can be estimated by incorporating a directional component into the slope term of the semi-variogram. This involves fitting a single equation which defines a continuous envelope of estimated semi-variograms for all directions between those of maximum and minimum variation.

The anisotropic model used by Burgess and Webster (1980a), Webster and Burgess (1980), and Trangmar (1984) is

\[
\gamma(\theta, h) = C_0 + [A \cos^2(\theta - \psi) + B \sin^2(\theta - \psi)]h
\]  

(10)

where \( \gamma(\theta, h) \) is the semi-variance estimated in the direction \( \theta \) at distance of separation \( h \), \( C_0 \) the nugget variance, \( \psi \) the direction of maximum slope \( A \) (greatest variation), and \( B \) the slope of the semi-variogram at 90° to \( \psi \). The parameters \( A, B, \) and \( \psi \) are generally estimated by least squares fitting of Eq. (10) to the pooled directional semi-variograms, with each semi-variance value being weighted by the number of pairs in each lag \( h \). The anisotropy ratio is calculated as \( A/B \). Slopes estimated by Eq. (10) from pooled directional semi-variances compared closely with the slopes of the individual directional semi-variograms for the data of Burgess and Webster (1980a) and Trangmar (1984). Figure 5 shows Eq. (10) fitted to semi-variances pooled from four directions. The direction of maximum variation is northeast to southwest and that of minimum variation is southeast to northwest.
An alternative linear anisotropic model which gives similar results is that of McBratney and Webster (1981a, 1983a), in which the square root of the slope term of Eq. (10) is used. Equations for applying the spherical model to anisotropic data are given in David (1977) but have yet to be applied to soil properties.

The utility of anisotropic modeling lies in identification of changes in spatial dependence with direction which, in turn, reflect soil-forming processes. McBratney and Webster (1981a) found that the geometric anisotropy of peat thickness was related to the microtopography of the land surface prior to peat formation. The anisotropy was caused by directional differences in peat thickness across and up the slopes in the region. Trangmar (1984) found that the direction of maximum variation of particle size fractions occurred down the main axis of tuff fallout and deposition of alluvium. Anisotropy of pH and HCl-extractable phosphorus in the same area was caused by directional changes in the degree of soil weathering across geomorphic surfaces of different ages.

Anisotropy ratios of up to 5.4 have been reported for soil properties, but directional differences of this magnitude are probably unusual for soils because most of the ratios are in the 1.3–4.0 range (Table II). The relative degree of anisotropy between topsoils and subsoils in Table II does not show any clear pattern and probably depends on the particular properties and soil
Table II
Anisotropy Ratios of Some Semi-variograms for Soils and Related Data

<table>
<thead>
<tr>
<th>Property</th>
<th>Topsoil</th>
<th>Subsoil</th>
<th>Reference</th>
</tr>
</thead>
<tbody>
<tr>
<td>Peat thickness (cm)</td>
<td>1.88</td>
<td></td>
<td>McBratney and Webster (1981a)</td>
</tr>
<tr>
<td>Stone content (%)</td>
<td>5.42</td>
<td></td>
<td>Burgess and Webster (1980a)</td>
</tr>
<tr>
<td>Sand (%)</td>
<td>1.59</td>
<td>1.68</td>
<td>McBratney and Webster (1983a)</td>
</tr>
<tr>
<td></td>
<td>4.05</td>
<td>2.88</td>
<td>Trangmar (1984)</td>
</tr>
<tr>
<td>Silt (%)</td>
<td>1.71</td>
<td>1.79</td>
<td>McBratney and Webster (1983a)</td>
</tr>
<tr>
<td>Clay (%)</td>
<td>3.01</td>
<td>2.91</td>
<td>Trangmar (1984)</td>
</tr>
<tr>
<td>pH</td>
<td>4.37</td>
<td>2.80</td>
<td>Trangmar (1984)</td>
</tr>
<tr>
<td>HCl-extractable phosphorus (ppm)</td>
<td>2.40</td>
<td>1.54</td>
<td>Trangmar (1984)</td>
</tr>
<tr>
<td>Electrical resistivity (Ωm)</td>
<td>3.47</td>
<td>5.18</td>
<td>Trangmar (1984)</td>
</tr>
<tr>
<td>Cotton petiole nitrate (ppm)</td>
<td>1.29</td>
<td></td>
<td>Webster and Burgess (1980)</td>
</tr>
<tr>
<td>Cotton petiole nitrate (ppm)</td>
<td>2.7</td>
<td></td>
<td>Tabor et al. (1984)</td>
</tr>
</tbody>
</table>

processes being studied. McBratney and Webster (1983a) were able to reliably fit one common anisotropic model to semi-variograms of sand and silt fractions in both topsoils and subsoils. All four semi-variograms had similar anisotropy ratios and directions of maximum variation, thus giving one simple linear model for four different variables.

Soil properties which are highly correlated and whose auto-semi-variograms vary anisotropically often have anisotropic cross-semi-variograms (McBratney and Webster, 1983a). Similarly, properties whose auto-semi-variances are isotropic tend to have isotropic cross-semi-variances (Vauclin et al., 1983).

Zonal anisotropy is often expressed as nested semi-variogram structures in which the observed anisotropy cannot be reduced by a simple linear transformation of sample distance (Journel and Huijbregts, 1978). It may result in different sills or different forms of semi-variograms calculated for the same property in different directions (David, 1977). Zonal anisotropy is a common characteristic of properties showing geochemical or geophysical gradients caused by directional deposition of sediments or mineralization of ore bodies (Journel and Huijbregts, 1978). Zonal anisotropy can occur in
three dimensions and, although commonly observed in mineral deposits, it has not been described in the soils literature. The conceptual and mathematical models of zonal anisotropy are given in full by Journel and Huijbregts (1978).

b. Trends. Many regionalized variables do not vary randomly but show local trends or components of broader regional trends. Quasi-stationarity (Journel and Huijbregts, 1978) can be safely assumed for interpolation purposes where there is a regional trend but local stationarity because the trend is more or less constant within the estimation neighborhood. Regional trends are indicated by semi-variances that increase with distance of sample separation and either do not approach a sill (Gajem et al., 1981) or have a sill which considerably exceeds the general variance $s^2$ (Bresler et al., 1984). In this case, simple kriging is used locally and an appropriate radius for the kriging neighborhood is the distance at which the semi-variance intersects the general variance (David, 1977).

In the case of overall stationarity but locally occurring trends, the stationarity assumptions of Section IV.B.2 break down and universal kriging must be used for local estimation. The stationarity assumptions are violated because the expected value of the random function $Z$ is not always constant within the neighborhood and is no longer equivalent to the mean, but to a general quantity of drift, $m(x)$, which changes locally within the neighborhood. The significance of identifying locally changing drift lies in difficulties with kriging from nonstationary data (Section VII.E).

Local trends, or drift, are commonly identified by simply plotting values of the soil property as a function of distance or by examination of semi-variograms (David, 1977; Webster and Burgess, 1980). Bresler et al. (1984) also analyzed residuals from the regression of soil property values on location to identify the presence of trends. Ideally, changing drift produces gently parabolic semi-variograms of the raw data which are concave upward near the origin (David, 1977). In practice, however, Webster and Burgess (1980) considered that the presence of short-range variation in most soils and noisy data over short lags generally makes local trend identification difficult.

c. Periodic Phenomena. Periodicity of parent material deposition and repetition of land form sequences are often quoted sources of soil variation (Butler, 1959). Periodic variation is expressed in semi-variograms as a "hole effect" (Fig. 6), which is indicative of nonmonotonic growth of the semi-variance with distance (Journel and Huijbregts, 1978). The hole effect can appear on models with or without sills. Periodic behavior in ore bodies is said to indicate a continuous process of mineralization and is often characteristic of a succession of rich and poor zones (David, 1977). The continuity of the process is indicated by the smooth shape of the hole-effect semi-variogram.

The hole effect will usually be present only in certain directions because the
periodicity of geologic and soil processes does not generally operate isotropically. Trangmar (1984) noted periodic behavior of a semi-variogram for sand content calculated perpendicular to the axes of two rivers which were 25 km apart. The periodicity was caused by repetition of deposition patterns away from each river. Barnes (1981) obtained hole-effect electrical conductivity semi-variograms computed in directions perpendicular to irrigation lines. The electrical conductivity was low adjacent to the irrigation lines, where salts had been leached by the irrigation water, increased to a maximum equidistant between lines, where less water had been applied, and decreased again adjacent to the next line as water application and leaching increased.

The natural variability of soil properties is generally nested (Burrough, 1983a) so that evidence of periodic semi-variogram structures at a given scale of observation may often be confounded with structures operating over different scales, the consequence of which is a dampening of the hole effect.

d. Nested Variation. We have previously recognized that there may be many sources and scales of variability present in any spatial study of soil properties. In semi-variogram analysis, nested structures can be conveniently represented as the sum of a number of semi-variograms, each with its own range and nugget variance characterizing variability at a particular scale. The size of the nugget variance generally increases with sampling scale, largely due to the variance contributed by shorter-range processes.

A single semi-variogram calculated over several discrete, independent, nested processes may consist of several near-linear portions which coincide with the range of the process in question (Fig. 4, curve 3). McBratney et al. (1982) obtained this type of semi-variogram for copper and cobalt in
Scotland and found that agricultural effects arising from field-to-field variation and farm-to-farm differences were responsible for short-range variation, while a geologic component caused a longer-range effect.

A relatively new concept in dealing with nested variation of natural phenomena is that of fractals (Mandelbrot, 1977), which provide a means for assessing the relative balance between processes operating over different spatial scales. Analysis of fractal behavior is similar to that of scaling (Warrick et al., 1977; Simmons et al., 1979), in which variations identified at one scale are statistically equivalent to those seen at other scales and are related simply by a scaling factor.

An ideal fractal process operates singly, without interaction with other processes, over a discrete range of spatial scales. Burrough (1983a) applied fractal concepts to soils by using the slopes of the double log plot of semi-variograms to obtain a measure of the relative balance between short- and long-range sources of soil variation. An ideal fractal process would have a double log semi-variogram with a slope of 1, but values for most soil properties exceed 1.5, indicating that soils are nonideal fractals (Burrough, 1983a). Structural analysis of semi-variograms using a nested, nonideal fractal approach (Burrough, 1983b) may provide useful information on the scales and relative importance of discrete sources of variation when such variation is caused by superimposed, independently acting soil processes. Such information can, in turn, be incorporated in the design of sampling schemes for properties with nested variation.

e. Management Effects and Soil Genesis. Management practices may considerably alter the inherent spatial structure of soil properties, and subsequently crop growth. Directional application of water, as in furrow irrigation, may impose considerable anisotropy of soil moisture content (Gajem et al., 1981) and subsequent uptake of nutrients such as nitrogen (Tabor et al., 1984). Similarly, nonuniform application of fertilizers may alter the spatial structure of nutrient uptake by crops (Trangmar, 1982). The spatial variability of soils may be altered when land use changes occur, such as conversion of forested land to arable farmland. In this regard, Trangmar (1984) found that topsoil removal by mechanical land clearance combined with subsequent burning of forest trash resulted in considerable short-range variation of topsoil acidity, exchange characteristics, and subsequent crop growth.

Analysis of soil variability using semi-variograms has aided identification of soil mapping units (Campbell, 1978) and placement of mapping unit boundaries (McBratney and Webster, 1981b). Spatial dependence of soil-forming factors such as rainfall (Yost et al., 1982a), parent material composition (McBratney et al., 1982) and deposition (Trangmar, 1984), and soil genetic processes, such as age and degree of weathering (Yost et al., 1982a; Trangmar, 1984), have also been quantified using semi-variogram analysis.

Clearly, analysis of spatial dependence can aid identification of the
underlying spatial structure of many soil properties and contribute significantly to an understanding of the spatial effects of soil-forming factors and genetic processes.

VI. INTERPOLATION BY KRIGING

A. General

Kriging is a technique of making optimal, unbiased estimates of regionalized variables at unsampled locations using the structural properties of the semi-variogram and the initial set of data values. A useful feature of kriging is that an error term (estimation variance) is calculated for each estimated value, providing a measure of the reliability of the interpolation. The simplest forms of kriging involve estimation of point values (punctual kriging) or areas (block kriging) and assume that the sample data are normally distributed and stationary (Henley, 1981). Various other estimation procedures are available when sample data show departures from these assumptions (Fig. 7). Soil properties often exhibit lognormal or complex probability distributions, in which case lognormal or disjunctive kriging is more appropriate. Data showing weak forms of nonstationarity can be detrended or interpolated using universal kriging (Yost et al., 1982b; Hajrasuliha et al., 1980; Webster and Burgess, 1980). The concept of generalized covariances has been used in

![Diagram](image-url)

**Fig. 7.** Kriging methods used under different conditions of stationarity and probability distribution of the data. (From Henley, 1981.)
APPLICATION OF GEOSTATISTICS

71

the mining industry for interpolation in the presence of local trends (Matheron, 1973; Delfiner, 1976). Directional differences in variation can also be taken into account during interpolation by using the anisotropic semi-variogram model to obtain the weights in the kriging system (Burgess and Webster, 1980a; Tabor et al., 1984).

Detailed accounts of kriging theory and application are given in Journel and Huijbregts (1978) and David (1977). Less quantitative descriptions are provided by Clark (1979), Royle (1980), and Henley (1981). Only the main concepts of the various kriging procedures will be summarized here.

B. PUNCTUAL KRIGING

Simple point estimation is probably the most common kriging procedure used in soil science to date (Burgess and Webster, 1980a; Vieira et al., 1981; Van Der Zaag et al., 1981; Van Kuilenburg et al., 1982; Russo and Bresler, 1982; Trangmar, 1982; Trangmar et al., 1982).

1. Concepts of Punctual Kriging

Kriging is a means of local estimation in which each estimate is a weighted average of the observed values in its neighborhood. The interpolated value of regionalized variable \( z \) at location \( x_0 \) is

\[
\hat{z}(x_0) = \sum_{i=1}^{n} \lambda_i z(x_i)
\]

(11)

where \( n \) is the number of neighboring samples \( z(x_i) \) and \( \lambda_i \) are weights applied to each \( z(x_i) \). The weights are chosen so that the estimate \( \hat{z}(x_0) \) of the true value \( z(x_0) \) is unbiased, i.e.,

\[
E[\hat{z}(x_0) - z(x_0)] = 0
\]

(12)

and the estimation variance \( \sigma_k^2 \) is minimized, i.e.,

\[
\sigma_k^2 = \text{VAR}[\hat{z}(x_0) - z(x_0)] = \text{minimum}
\]

(13)

The weights placed on each neighboring sample sum to 1, and their unique combination for which \( \sigma_k^2 \) is minimized can be obtained when

\[
\sum_{j=1}^{n} \lambda_j \gamma(x_i, x_j) + \mu = \gamma(x_i, x_0) \quad \text{for all } i
\]

(14)

The values \( \gamma(x_i, x_j) \) and \( \gamma(x_i, x_0) \) are the semi-variances, or preferably the covariances (second-order stationarity), between observed locations \( x_i \) and \( x_j \).
and between the observed location $x_i$ and the interpolated location $x_0$, respectively. These values are obtained from the semi-variogram of $Z$. The quantity $\mu$ is the Lagrangian multiplier associated with the minimization of $\sigma_k^2$. Solution of the $n + 1$ equations of the kriging system [Eq. (12)] for each $\lambda_i$ and $\mu$ enables the kriged estimate $\hat{z}(x_0)$ to be determined by Eq. (11) and the estimation variance to be determined by solving for

$$\sigma_k^2 = \sum_{i=1}^{n} \lambda_i \gamma(x_i, x_0) + \mu$$

(15)

The set of $n + 1$ simultaneous equations of the kriging system is most efficiently solved using matrix methods, as outlined by Burgess and Webster (1980a).

The interpolated value at any unsampled location is the most precise possible from the available data and one that can be used with known confidence (McBratney et al., 1982). The estimation variance depends only on the semi-variogram and the configuration of the data locations in relation to the kriged points and not on the observed values themselves (Burgess and Webster, 1980a). Thus, it is a local error term giving more reliable estimation of interpolation precision at a given point than the global error terms of many other interpolation procedures (Giltrap, 1983a). Standard deviations or confidence limits can be calculated from the estimation variance, if necessary.

The condition of unbiasedness ensures that kriging gives exact interpolation in that estimated values are identical to observed values when a kriged location coincides with an observed location. In such cases, the weights on neighboring samples are zero and the estimation variance will equal the nugget variance of the semi-variogram. This feature of kriging is particularly desirable when the nugget variance is small and observations have been made with negligible error.

The appropriateness of the chosen semi-variogram model and the kriging assumptions of unbiasedness and minimum estimation variance can be tested by "jack-knifing" procedures (Efron and Gong, 1983). This involves deleting each sample in turn, then kriging it independently from all other points in the estimation neighborhood, and finally statistically testing the resulting mean prediction error for unbiasedness and the prediction mean square (variance of predicted minus observed values) for minimum estimation variance (Vieira et al., 1981; Vauclin et al., 1983; Tabor et al., 1984; Trangmar, 1984). Most other interpolation procedures do not have the optimality attributes of unbiasedness and minimum variance that kriging has for spatially dependent variables (Yost et al., 1982b).

Only sample locations which are spatially related to the kriged location (i.e., within the range of spatial dependence) are used in kriging. The nearest
few samples are the most heavily weighted, with the result that the semi-
variogram needs to be accurate only over the first few lags. Little is gained by
including distant points unless there are few samples close to the kriging
location, which may be the case for randomly sampled data. If so, then
estimation variances will be relatively large and additional sampling is
probably necessary. Sample points occurring in clusters carry less weight
than lone points, with the result that addition of just one additional sample in
sparsely sampled regions can markedly reduce estimation variances in such
regions. Sample locations lying between the kriged point and more distant
samples screen the distant ones so that the latter have less weight than they
otherwise would (Van Kuilenburg et al., 1982).

Distances closer than the range may be used to define the maximum radius
of the kriging neighborhood when the semi-variogram is well structured with
long range and there are sufficient samples at short lags (Vieira et al., 1981).
Reducing the kriging radius in such cases reduces computer costs and is an
efficient practice if reliable interpolation is still obtained from fewer neigh-
bors.

The number of neighbor samples required for reliable estimation depends
on the configuration of sampled to kriged locations and the degree of
anisotropy and will vary among data sets. Reported numbers range from 7
(Vauclin et al., 1983) to 25 (Burgess and Webster, 1980a) for grid-sampled
data and from 10 (Yost et al., 1982b) to 40 (Vieira et al., 1981) for irregularly
spaced samples. A conservative estimate would seem to be about 20 to 25
neighbors to achieve reliable interpolation.

Anisotropic spatial dependence can be readily taken into account in all
forms of kriging simply by using the anisotropic semi-variogram model for
estimation of the weights in Eq. (14) (Burgess and Webster, 1980a; Webster
and Burgess, 1980; McBratney and Webster, 1983a; Tabor et al., 1984;
Trangmar, 1984). Directional differences should be incorporated in the
kriging of anisotropic data because it results in more precise interpolation
than using the commonly assumed isotropic condition.

2. Application

Most of the early applications of kriging in soil science involved simple
point estimation for isoproperty mapping (Burgess and Webster, 1980a; Van
Der Zaag et al., 1981; Vieira et al., 1981) and use of estimation variances for
designing sampling schemes for future kriging operations (Burgess et al.,
1981; McBratney and Webster, 1981a; McBratney et al., 1981; Vieira et al.,
1981; Trangmar et al., 1982). Point estimation has also been used in co-
kriging (McBratney and Webster, 1983a; Vauclin et al., 1983) and universal
kriging (Webster and Burgess, 1980; Hajrasuliha et al., 1981; Yost et al., 1982b).

Punctual kriging for isoproperty mapping involves estimation of values for a fine grid of points through which isarithms (contours) or three-dimensional surfaces are drawn for spatial display as maps. Point interpolation has been carried out over many different spatial scales ranging from regional variation of soil chemical properties (Van Der Zaag et al., 1981; Yost et al., 1982b) to within-field variation (Burgess and Webster, 1980a) to microvariation of soil hydraulic properties (Vieira et al., 1981) and nutrient uptake by crops within experimental plots (Trangmar, 1982).

Van Kuilenburg et al. (1982) compared the estimation precision for punctual kriging of soil moisture content with the variance of mean values for soil mapping units, and with proximal and weighted average interpolation. They concluded that kriging was the most precise estimator but that its efficiency should be balanced with the multipurpose utility of a soil map. This result emphasizes the role of geostatistics in augmenting conventional soil survey methods in describing the spatial variability of specific properties.

Estimation variance maps of kriging from randomly sampled data have proved useful for identifying those areas where further sampling would provide the most additional information (Van Der Zaag et al., 1981). Trangmar (1984) was able to improve estimation precision by up to 40% by collecting additional samples in areas of high estimation variance and then recomputing the semi-variogram and re-kriging the data.

Despite the ease of computation and apparent utility of punctual kriging, it has some undesirable attributes. Punctual kriging is an exact interpolator and may produce local discontinuities where interpolated points coincide with sample locations. Isarithmic mapping often results in considerable local detail associated with these discontinuities, and broader regional soil patterns may be obscured (Burgess and Webster, 1980b; McBratney et al., 1982). The position of this local detail is itself a sampling effect because it depends on the particular places at which samples were collected. Moving the origin or orientation of the sampling pattern could result in a different map of kriged values and estimation variances (Burgess and Webster, 1980b). So, results obtained by punctual kriging depend strongly on the sampling methodology used.

The nugget or random variance is a component of the semi-variance at any lag and, although it does not influence the kriged value, it sets a minimum value to the estimation variance of each kriged location. As a result, punctual kriging may produce undesirably large estimation variances if the nugget variance is large.

These shortcomings of punctual kriging can be avoided by interpolation over areas, using block kriging (Section VI,C), which results in smoother maps and smaller estimation variances.
C. Block Kriging

Users of soils information are often interested in average estimates for discrete areas or blocks (e.g., management units) rather than point estimates obtained by punctual kriging. The "points" at which estimates are made by punctual kriging are actually volumes of soil of the same size and shape as the soil volumes (e.g., cores or pits) which were originally sampled. In many cases, one may wish to interpolate an average value for an area or block which is larger than the cross-sectional area of the soil volume actually sampled. Block kriging provides a method for achieving this and, at the same time, avoids some of the shortcomings of punctual kriging.

1. Concepts

In block kriging, a value for an area or block with its centre at \( x_0 \) is estimated rather than values at points as in punctual kriging. The kriged value of property \( Z \) for any block \( V \) is a weighted average of the observed values \( x_i \) in the neighborhood of the block, i.e.,

\[
\hat{z}(V) = \sum_{i=1}^{n} \lambda_i z(x_i)
\]

(16)

The only difference in Eq. (16) from the estimation equation (11) for punctual kriging is in the determination of the weighting coefficients. In the weighting procedure, the semi-variances between data points and the interpolated points of punctual kriging are replaced by the average semi-variances between the data points and all points in the block \([\gamma(x_i, V)]\). The optimum combination of the weights on sample locations is that for which the estimation variance is minimized.

The minimum estimation variance for block \( V \) is

\[
\sigma_V^2 = \sum_{i=1}^{n} \lambda_i \gamma(x_i, V) + \mu_V - \gamma(V, V)
\]

(17)

where \( \gamma(x_i, V) \) is the average semi-variance between the sample points \( x_i \) in the neighborhood and those in the block \( V \), \( \gamma(V, V) \) is the average semi-variance between all points within \( V \) (i.e., the within-block variance of classical statistics), and \( \mu_V \) is the Lagrangian parameter associated with the minimization. The within-block variance includes the nugget variance component plus variance among any samples occurring within each block (Table III). The estimation variance of block kriging is always less than that of punctual kriging because the within-block variance is removed from the error term [Eq. (17)]. Up to 20-fold improvements in average estimation precision have been achieved using block kriging compared to punctual kriging (Table
Table III

Variance Associated with Punctual and Block Kriging of Some Soil Properties

<table>
<thead>
<tr>
<th>Property</th>
<th>Stone content (%)&lt;sup&gt;2&lt;/sup&gt;</th>
<th>Loam thickness (cm)&lt;sup&gt;2&lt;/sup&gt;</th>
<th>Sodium content (meq/10 kg)&lt;sup&gt;2&lt;/sup&gt;</th>
<th>Total nitrogen&lt;sup&gt;a&lt;/sup&gt; (%)&lt;sup&gt;2&lt;/sup&gt;</th>
<th>Lime requirement (ton/ha)&lt;sup&gt;2&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Nugget variance</td>
<td>10.0</td>
<td>187.0</td>
<td>8.7</td>
<td>0.07</td>
<td>3.2</td>
</tr>
<tr>
<td>Within-block variance</td>
<td>15.1</td>
<td>369.2</td>
<td>10.5</td>
<td>0.09</td>
<td>3.9</td>
</tr>
<tr>
<td>Mean block kriging variance</td>
<td>0.6</td>
<td>32.5</td>
<td>0.8</td>
<td>0.04</td>
<td>0.9</td>
</tr>
<tr>
<td>Mean punctual kriging variance</td>
<td>12.8</td>
<td>320.0</td>
<td>10.7</td>
<td>0.10</td>
<td>4.0</td>
</tr>
<tr>
<td>Total sample variance s²</td>
<td>79.3</td>
<td>786.7</td>
<td>15.3</td>
<td>0.14</td>
<td>4.3</td>
</tr>
<tr>
<td>Reference&lt;sup&gt;b&lt;/sup&gt;</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
</tbody>
</table>

<sup>a</sup> Variances determined on log transformed values.
<sup>b</sup> References: 1, Burgess and Webster (1980b); 2, Trangmar (1984).

III). Table III also demonstrates that variance for the sample mean obtained in the classical manner will always overestimate the interpolation error achieved by block or punctual kriging for spatially dependent properties.

2. Applications

The most common use of block kriging has been for the production of isarithm maps of soil properties (Burgess and Webster, 1980b; McBratney et al., 1982; Trangmar, 1984).

Experience indicates that block kriging produces smoother maps than punctual kriging by interpolating average values for blocks, with the effect of smoothing local discontinuities. Such smoothing effects can be seen by comparing the plots in Fig. 8. These effects are particularly desirable when regional patterns of variation are of more interest than local detail.

Improvements in the estimation variance of block kriging over that of punctual kriging is most evident for properties with large nugget variances (Table III). If the block size exceeds the minimum sample spacing, much of this short-range variation is included in the within-block variance term, resulting in small estimation variances.
FIG. 8. Isarithm maps of cover loam thickness, Hole Farm, England, constructed from (A) punctual kriging and (B) block kriging of 400-m² cells. (From Burgess and Webster, 1980a,b.)
These desirable attributes of block kriging relative to punctual kriging make use of the former more prevalent in mining engineering (David, 1977), and this is likely to be the case in future spatial studies of soils.

The relatively smooth isolines and maps produced by block kriging may appear similar to those obtained by simpler procedures, such as spline interpolation. However, the similarity of such maps is misleading because the purpose of block kriging is to provide accurate local estimates with known error for variables over discrete blocks of land of predetermined area. Such attributes are rarely embodied in other interpolation techniques (Burgess and Webster, 1980b).

Block kriging has also been applied to interpolate spatial effects of crop response to variability imposed by soil management practices. Tabor et al. (1984) found that maps of block-kriged values for nitrate content of cotton petioles indicated a strong response to direction of planting rows and application of irrigation water. Trangmar (1984) used block kriging of growth and yield parameters of upland rice to demonstrate the inverse spatial relationship of soil aluminum saturation on crop growth within an experimental plot recently cleared of forest vegetation. These studies suggest that geostatistical analysis of environmental and crop parameters may be increasingly useful in evaluating the spatial effects of soil properties, pest and disease incidence, and other environmental effects on crop growth and in suggesting cost-effective strategies for amelioration of yield-reducing spatial effects.

The adaptation of volume-variance relationships for estimation of ore recovery in mining (David, 1977; Clark, 1979) to the agronomic situation offers potential for interpretation of critical levels of soil nutrients for determination of fertilizer or amendment needs in the presence of spatial heterogeneity of nutrients.

Block kriging of a very fine grid of cells forms the basis of interpolation procedures developed by Giltrap (1983a,b), which provide for prior stratification of the landscape into a number of land classes and can either restrict interpolation to cells within the same land class or allow interpolation across land classes using separately calculated autocorrelation functions. Using these procedures, Giltrap (1981) was able to rapidly and cheaply produce maps for many different soil properties at any scale smaller than that of the original interpolation grid. Such an approach has considerable potential to aid quantitative placement of mapping unit boundaries in soil surveys and areal estimation of inclusions within mapping units and to enable rapid generation of soil interpretive maps based on various combinations of field-measured values of soil properties stored in a data base system.

3. Sampling for Punctual and Block Kriging

Much of the punctual and block kriging of soil properties has been done on data originally sampled for other purposes, often in a nongeometric design
APPLICATION OF GEOSTATISTICS

(Yost et al., 1982b; McBratney et al., 1982; Van Der Zaag et al., 1981; Van Kuilenburg et al., 1982). Kriging from randomly sampled data usually results in uneven estimation variance maps arising from the irregular sampling density. Overall estimation variance is not minimized and such designs are, therefore, nonoptimal from a kriging standpoint.

The pooled value of estimation variances is minimized for any given sample if sampling is carried out on a grid basis, resulting in neighborhoods with the same number of samples. Estimation variances always increase along the margins of the study region, where the neighbors decrease in number irrespective of whether the data is grid or randomly sampled.

Burgess et al. (1981) and McBratney and Webster (1983b) found that, given a previously determined semi-variogram, an equilateral triangular grid is the optimal sampling for kriging of isotropic data in terms of minimizing the interpolation error. They also concluded that square grids were almost as precise and were often more convenient. Rectangular grids with their longer intervals aligned in the direction of least variation are optimal where there is simple anisotropy. The anisotropy ratio defines the ratio of sampling intervals in the direction of maximum and minimum variation. Using these approaches, McBratney and Webster (1983b) obtained 3- to 9-fold gains in sampling efficiency over that estimated by classical statistical theory for simple random sampling. McBratney et al. (1981) and McBratney and Webster (1981a) provided sampling theory and a computer program for the design of optimal sampling schemes for kriging and mapping of soil properties given a previously determined semi-variogram and knowledge of any anisotropy.

Given the semi-variogram and isotropic variation, Webster and Burgess (1984) showed that for block kriging, sampling of square grids with observations at the centers of a mosaic of cells within each block is the most efficient. They also concluded that bulking of samples from such a grid within blocks will always increase the estimation precision for additive properties compared to using individually measured samples at each cell center.

Establishment of such sampling grids depends on knowing the semi-variogram for the property of interest, which unfortunately is not generally the case until an initial sampling program has been carried out. McBratney and Webster (1983b) recommend that if the semi-variogram is not known and kriging over the whole region is the objective, then the best strategy is to sample on a regular grid, with the interval determined by the number of observations that can be afforded. One alternative approach is to sample along a series of transects to obtain the semi-variogram and then establish the sample pattern for kriging based on structural analysis of the semi-variogram. Another possible design in which the semi-variogram is unknown is that of Trangmar (1984), in which closely sampled transects were placed in four separate directions across a coarser-sampled square grid. The transects
enabled detection of anisotropy and establishment of the semi-variogram at short lags, while the coarser grid provided the basis for kriging interpolation.

D. Co-Kriging

The spatial distribution of any given property may often be closely related to that of other properties affected by the same regionalized phenomenon or spatial process. Such properties are said to be co-regionalized and are spatially dependent on one another. Co-kriging extends the principle of optimal estimation using regionalized variable theory from that of a single property to situations where there are two or more co-regionalized properties. Co-kriging is most efficiently used where one variable may not have been sampled sufficiently (due to experimental difficulties, high costs, etc.) to provide estimates of acceptable precision. Estimation precision can be improved by utilizing the spatial correlation between the undersampled (primary) variable and other, more frequently sampled covariables.

1. Concepts

The concepts of co-kriging discussed here assume only one covariable, but the equations are readily expanded to include additional covariables (Journel and Huijbregts, 1978).

The co-regionalization of two variables \( z_1 \) and \( z_2 \) is summarized by a cross-semi-variogram:

\[
\gamma_{12}(h) = \frac{1}{2N(h)} \sum_{i=1}^{n} [z_1(x_i) - z_1(x_i + h)][z_2(x_i) - z_2(x_i + h)]
\] (18)

where \( N(h) \) is the number of pairs of values separated by vector \( h \) (David, 1977). Variables \( z_1 \) and \( z_2 \) do not necessarily need to have the same number of samples, but the cross-semi-variogram is calculated using only the locations where both variables are measured. The spatial dependence of co-regionalized properties can also be determined using cross-correlograms (McBratney and Webster, 1983a). Unlike auto-semi-variances, cross-semi-variances can be negative if the relationship between \( z_1 \) and \( z_2 \) is negative. They may also be anisotropic, particularly if the auto-semi-variances are so (McBratney and Webster, 1983a).

The co-kriged value of the undersampled, or primary, variable, \( \hat{z}_2 \), is computed as a weighted average of the observed values of the covariable, \( z_1 \), and \( z_2 \) occurring in the estimation neighborhood of each kriged point. The
co-kriged value $\hat{z}_2$ is

$$\hat{z}_2(x_0) = \sum_{i=1}^{n_1} \lambda_{1i} z_1(x_{1i}) + \sum_{j=1}^{n_2} \lambda_{2j} z_2(x_{2j})$$

(19)

where $\lambda_{1i}$ and $\lambda_{2j}$ are the weights associated with $z_1$ and $z_2$, respectively, while $n_1$ and $n_2$ are the number of neighbors of $z_1$ and $z_2$ involved in estimating $\hat{z}_2$ at each location $x_0$, respectively.

The weights on observed values of $z_1$ and $z_2$ are chosen so that the estimate is unbiased with minimum variance, just as in auto-kriging. However, solution of the co-kriging system for the weights is obtained using the auto-semi-variances and the cross-semi-variances of each $z_1$ and $z_2$ with the kriged location $x_0$. The resulting system of equations is more complex and more costly to compute than in simple auto-kriging. Solution of the co-kriging system also yields the co-kriging estimation variance for each interpolated location. The equations of the co-kriging system are presented in full by Journel and Huijbregts (1978), McBratney and Webster (1983a), and Vauclin et al. (1983).

Co-kriging can be used for point estimation (punctual co-kriging) or block estimation (block co-kriging). In block co-kriging, average auto- and cross-semi-variances of samples within blocks are incorporated into the co-kriging system. The co-kriging system requires at least one sample point of both the primary variable and covariable properties within the estimation neighborhood. If the primary variable and covariable are present at all sampling sites in the neighborhood, then co-kriging yields the same estimate as auto-kriging of the primary variable alone. In such cases, co-kriging is unnecessary.

2. Application

Co-kriging has been applied only to point estimation of soil properties (McBratney and Webster, 1983a; Vauclin et al., 1983; Trangmar, 1984). McBratney and Webster (1983a) identified common anisotropic co-regionalization among particle size fractions and used it to co-krige topsoil silt content from more densely sampled subsoil silt and sand. Co-kriging using one and two covariables, respectively, reduced estimation variance in successively smaller increments relative to auto-kriging. Presumably, the increments in variance reduction gained by using more than one covariable must outweigh the increased complexity of the co-kriging system for practical uses in most situations.

In other soils-related applications of co-kriging, Vauclin et al. (1983) interpolated available water and 1/3 bar water content using sand content as
the covariable, while Trangmar (1984) co-kriged available phosphorus using more densely sampled observations of total phosphorus.

Maps of co-kriged values usually show the same broad pattern of variation as for auto-kriged values, but they tend to be more intricate because the covariables are sampled more densely. Co-kriging improves estimation precision only when there are few neighboring samples of the primary variable and the spatial correlation with the more frequently sampled covariable is taken into account. If several neighboring primary variables are present, they receive most of the weighting and the covariable merely adds another variance component to the estimation without resulting in much or any improvement to the overall estimation precision. The precision of co-kriging decreases near the boundaries of the study region as samples of the primary variable and covariable decrease in number.

Geometric configurations of the primary variable interspersed with a finer grid of covariable points provides the optimal sampling design for co-kriging (McBratney and Webster, 1983a; Vauclin et al., 1983). McBratney and Webster (1983a) demonstrate sampling configurations for different ratios of sampling intensity of the covariable relative to the primary variable and the subsequent spatial effects on estimation variance (Fig. 9). The position of maximum variance depends on the semi-variogram form and the strength of the cross-correlation, on the sampling interval, and on the sampling intensity ratio of covariable to primary variable. Once the semi-variogram is obtained, the maximum co-kriging variance corresponding to spacings of each variable can be determined by plotting co-kriging estimation variance values against sample spacings for a range of primary variable-to-covariable sampling ratios (Fig. 10). The optimal spacing and primary variable-to-covariable sampling ratio is that which acquires the desired precision at least cost. Anisotropy of the respective properties should be used to determine the relative dimensions of the sampling grids. Block co-kriging would yield smaller estimation variances and, therefore, narrower confidence limits than punctual co-kriging given the same sampling scheme (McBratney and Webster, 1983a).

There are numerous situations in agronomic research where properties that are cheap or easy to measure are likely to be co-regionalized with others that are of importance but less easily determined. These might include interpolation of available water capacity from textural components (Vauclin et al., 1983), use of pH for interpolation of exchangeable cations and aluminum, and interpolation of textural components from field estimates and laboratory determinations (Hodgson et al., 1976). Co-kriging of soil moisture content from limited ground observations and the large amounts of data generated by remote sensing systems provides another possible use of co-regionalized variables (Price, 1980). Similarly, co-regionalization of soil and
Fig. 9. Locations of maximum co-kriging estimation variance for sampling schemes on square grids with different intensities for the covariable. Symbol ○ shows where both primary and covariables are observed; • shows where only the covariable is observed. Maximum estimation variance can occur on the dashed lines and near the crosses. SR is the sampling ratio of covariable to primary variable. (From McBratney and Webster, 1983a.)

weather parameters with crop parameters could be used for spatial estimation of crop performance.

The cost in time and money of obtaining a larger number of covariable measurements must be considerably less than for the variable of primary interest to make co-kriging viable in a planned experiment. McBratney and Webster (1983a) estimated that measuring a primary variable would need to
Fig. 10. Graphs of punctual co-kriging variance against primary variable spacing for covariable-to-primary variable sampling ratios from 1 to 25. (From McBratney and Webster, 1983a.)

cost at least five times that of a covariable to make a design for co-kriging economically sound in their particular study. The success of co-kriging also relies on a well-structured cross-semi-variogram, which is not always present even if the auto-semi-variograms have strong structure.

Despite these qualifying conditions, co-kriging has the potential to make considerable time and cost savings in mapping properties for which there are cheap, co-regionalized surrogates. Its potential should be further tested.
E. Universal Kriging

Ordinary kriging is relatively well established, both theoretically and practically, and there are numerous applications in soil science and general agriculture. One of the troublesome assumptions required for ordinary kriging is that the data are stationary or specifically follow the “intrinsic hypothesis,” which requires that the expected value of the difference between any two samples depend on the distance between them but not on their location in the sampled region. Thus, the difference between any two samples the same distance apart should be generally similar throughout the region. This definition, however, does not indicate when a trend is sufficiently strong to require a method other than ordinary kriging. Universal kriging was developed to permit kriging in the presence of strong trends (Matheron, 1971; Dagbert and David, 1976; Delfiner, 1979).

1. Concepts

Universal kriging was designed to permit kriging in the presence of trends in the sample data, i.e.,

$$E(Z(x)) = m(x) = \sum_{i=1}^{n} a_i f_i(x)$$

where $E(Z(x))$ is the expected value of the sample data, $m(x)$ is the trend, $f_i$ are the terms in the polynomial, and $a_i$ are the coefficients of the polynomial that describes the trend. The estimation of the trend has been approached several ways. In early work (Olea, 1974; David, 1977), geostatisticians fitted response surface equations to the entire region to remove the trends and then used the residuals to estimate the semi-variograms. This approach can give rise to unusual concave-shaped semi-variograms, which have smaller semi-variances at the greater distances, rather than the usual case, where semi-variances become larger and approach the sample variance. Estimates of the nugget variance and range from such semi-variograms are much less reliable so that the inference of soil property zones of influence and genetic origin, for example, becomes difficult, if not impossible. There is a problem, however, in determining the correct form of the trend. Because of the typically highly correlated errors, statistical tests of model adequacy are not reliable. There also is the question of whether all trends can be described by polynomial equations.

A more recent approach to universal kriging has been to expand the matrices of the kriging equations and simultaneously estimate the weights
and remove the trend (be it linear, quadratic, cubic, etc.). With this approach the trend is removed from the neighborhood (the specific points used in the estimation) rather than from the entire region, as in the above case. This method also allows more flexibility because only the form of the trend is specified for each estimation neighborhood. This means that the order of the regression is determined, but the coefficients are independently determined for each neighborhood. Thus the estimation proceeds as a "moving neighborhood." This approach is therefore much more flexible and should remove more variation than the single-equation approach of trend analysis.

However, the questions remain, Was the trend correctly removed? Was it necessary to remove the trend in the first place? Further work has pointed out that trend removal is not always straightforward. Armstrong (1984) suggests that both of the above approaches suffer indeterminacy in both the drift and the semi-variogram, which precludes their accurate estimation unless additional constraints or data are determined. She suggests this indeterminacy is similar to that associated with the constant of integration. With both of the above approaches to universal kriging, one can increase the degree of the response surface equation from linear to quadratic and higher and it is not clear just where to stop. The problem is analogous to that of determining the appropriate degree for regression equations: how to determine when one is fitting the actual trend and when one is fitting local variation.

2. Application

It has been suggested that one can determine nonstationarity, and hence the need for universal kriging, from the appearance of the semi-variogram (Journel and Huijbregts, 1978; David, 1977). Specific evidence of nonstationarity is said to be displayed when the semi-variogram increases concave upward and does not level out to approach the population variance at large distance. The usual suggestion has been that if the semi-variance approaches but does not drastically exceed the overall variance at large distances, chances are that the data are stationary. Studies by Starks and Fang (1982), however, suggest that there may be more subtle distortion of the semi-variogram caused by nonstationarity. They suggest that nonstationarity can result in bias which is not manifested solely by concave-upward-shaped semi-variograms. In their simulation studies, both nugget variance and range were affected as well. They suggested the following to check for nonstationarity:

1. Visually inspect the data for large trends;
2. Compare kriged estimates with the actual data, i.e., "jack-knife" and compare the variance of actual kriged values with the estimation variances.
These authors point out that ordinary kriging is relatively robust to nonstationarity because (1) the ordinary kriged estimate is a good approximation for linear drift with small slope and (2) drift tends to increase the semi-variance, which causes a more conservative calculated estimation variance. Probably another factor is that nonstationarity seems to be expressed more at larger distances than at smaller ones. Ordinary kriging places particularly high emphasis on the closer points so that correct estimation of the semi-variance is most important at small distances, where nonstationarity probably is less frequent.

The need to apply universal kriging or unbiased kriging of the kth order in agriculture has been difficult to determine, and more experience is needed. In one of the first applications of universal kriging to soil science, Webster and Burgess (1980) concluded that universal kriging appeared to be neither universally applicable nor of particular benefit. Some of our preliminary studies (Yost et al., 1982b) with large areas and clear trends in the data suggested that universal kriging resulted in very little improvement over ordinary kriging when comparing the observed and estimated points. Our results suggested that ordinary kriging is quite robust to the presence of even strong trends in the data. This comparison was obtained by "jack-knifing," in which an observed data point was removed from consideration and then estimated from the surrounding data. The observed and predicted data were then compared. Such comparisons provide a good method of determining the accuracy and precision of alternative estimation procedures and are widely used in geostatistics.

Faith and Sheshinski (1979) also attempted to determine the effect of ignoring trends in the data. Their results also indicate that ordinary kriging is apparently quite robust to the presence of drift. In both of the above cases the nonstationarity occurred at large distances, not at small distances. David (1977) pointed out that there can also be nonstationarity at small distances. Clearly there are many questions remaining about how and when to apply universal kriging to soils and soil phenomena. In general, however, universal kriging remains a methodology of geostatisticians, and its use requires expert assistance.

Universal kriging is also known as unbiased kriging of the kth order (Journel and Huijbregts, 1978). This terminology derives from the concept of generalized covariances. The generalized covariances are presented in Matheron (1973) and provide yet another approach to handling nonstationarity. The trends are removed from the data by successive differencing operations similar to the removal of nonstationarity in time series by successive differencing (Box and Jenkins, 1976). The differencing operation consists of subtracting $Z(h_i)$ from $Z(h_{i+1})$ for all $i$. The procedure can be repeated as necessary to result in differences $\Delta Z$ which are stationary. The extent to
which the differencing also removes evidence of spatial dependence remains unknown. Matheron developed the procedure to ensure that the estimation variances are positive (Delfiner, 1976). Procedures and techniques have been described by Stark's et al. (1980) for use in coal mine contouring in Iowa. This method of handling trends in the data apparently serves well in mapping and contouring; however, the semivariogram is lost and with it the rich interpretations of the underlying soil-, geologic-, and soil genetic-related phenomena. Some of the earliest mapping and contouring procedures used by mining consultants (BLUEPACK and KRIGEPACK) used the generalized covariance approach. Evidence that the generalized covariance approach provides significantly more accurate contouring is scarce.

F. Kriging from Non-normally Distributed Data

Deviations from the assumptions of normality in regionalized variable theory has been a criticism of kriging (Henley, 1981). Transformation to normality prior to geostatistical analysis results in a nonlinear function of the original data, so that kriging estimates may not be made with minimum estimation variance and without bias. The serious of this violation increases as deviations from normality increase. Lognormal kriging of points or blocks has been developed for stationary, lognormally distributed data and disjunctive kriging for stationary variables of more complex probability distributions (Fig. 7). Kriging of nonstationary variables of complex distribution is still under theoretical development.

1. Lognormal Kriging

Kriging of lognormally distributed data has been widely used in the mining industry (Rendu, 1979; Parker et al., 1979; Journel, 1980) and is beginning to receive more attention in soil science (Van Der Zaag et al., 1981; Yost et al., 1982a,b; Trangmar, 1984). It simply involves computation of semi-variograms and kriging on natural log-transformed values of the original data using the same procedures as for simple linear kriging. The log-kriged values and estimation variances can be re-expressed in terms of the original data using procedures outlined by Journel and Huijbregts (1978; Eqs. VIII-14, VIII-15). These authors concluded that lognormal kriging methods yield smaller estimation variances than simple linear kriging of the untransformed data if the original data is clearly lognormally distributed. This approach has been found to be a practical one (Rendu, 1979; Yost et al., 1982b), yielding reasonable estimates despite the minor violation of the kriging assumptions.
One of the innovative aspects of ordinary kriging was that it placed the data before the statistical model, therefore requiring a new model. Ordinary kriging also did not depend on the statistical distribution of the data. Disjunctive kriging is a specialized form of kriging which differs from ordinary kriging by requiring consideration of the statistical distribution of the data. If the data are normally distributed, however, one can obtain additional probabilistic information, i.e., one can determine proportions and state confidence limits about the prediction which are more difficult with non-normally distributed data. Transforming the data so it is normally distributed permits drawing on the large body of statistical inference. However, to obtain the normally distributed data required for disjunctive kriging one has to determine the necessary transformation to convert the data into bivariately normally distributed data. This process is achieved by application of Hermite polynomials to the original data. A serious disadvantage of disjunctive kriging is the transformation step in which the form of the transformation must be determined so that the resultant transformed data are normally distributed. Typically a 12th-order polynomial is used for the transformation. The polynomial must be estimated and then incorporated into the standard kriging equations. This results in considerably increased computation time. Similar to ordinary kriging, disjunctive kriging assumes stationarity.

A recent study by Journel (1983) suggests another alternative to disjunctive kriging. He compared a multivariate Gaussian approach, which may be more straightforward to apply. As suggested above, by transforming the data into multivariate normal distribution, more powerful inferential statistics can be applied, leading to estimates of quantities, proportions, and grade-tonnage types of calculations. Currently these methods are too new and untested for general application to our needs. In our view, this area of research in geostatistical methods holds much promise for producing methods which will bring to fruition the application of geostatistics to agricultural problems.

VII. PERSPECTIVES: FUTURE USE OF GEOSTATISTICS IN SOIL RESEARCH

The main contribution of geostatistics to soil research lies in structural analysis of soil variation and its use for local estimation. The semi-variogram provides a quantitative tool for relating the inherent structure of variation in specific properties to spatial effects of soil-forming factors and processes,
including management effects. The development of procedures for quantifying anisotropy, trends, and periodic phenomena gives geostatistics a versatility for dealing with natural phenomena that few other interpolation methods have. Procedures for quantifying nested variation (Burrough, 1983a,b) and computer production of maps at a range of spatial scales from a finely kriged grid (Giltrap, 1983a) represent new approaches to dealing with scale-related effects of soil processes and scientists' perception of them.

The initial emphasis in geostatistical interpolation has been on punctual kriging, due to its ease of computation. The relative benefits accruing from block kriging, such as smoother maps, smaller estimation variances, and easier use for design of sampling schemes, are now generally acknowledged and are likely to result in increased use of areal interpolation in the future. In designing spatial studies for estimation and mapping of properties for which there are cheap surrogates, the co-regionalization of properties and use of co-kriging should be considered as a potential cost-saver in making field and laboratory measurements without loss in mapping precision.

Geostatistics could be used in soil survey operations for structural analysis of soil variation to aid understanding of soil genesis and for analysis of reconnaissance data for defining future sampling populations and configurations both within and among different terrain units. The cost-effectiveness of geostatistics-based sampling schemes in practical soil survey operations needs to be field tested in different types of terrain for comparison with traditional sampling techniques. Kriging can augment general-purpose information contained in conventional soil maps by interpolation of interpretive data and specific measured or derived properties, which may vary independently of mapping unit boundaries. The ease of data manipulation, speed of computation, and precision of computer-generated maps based on kriging of soil properties make geostatistical techniques particularly desirable in the face of user demand for quick and reliable soil survey results (Giltrap, 1983a).

The versatility and range of geostatistical software now available make spatial analysis of natural phenomena applicable to many areas of agronomic research. Block kriging appears to be particularly useful for estimating soil amendment requirements over areas the size of land management units. Adaptation of volume-variance relationships for estimation of ore recovery in mining (David, 1977; Clark, 1979) to the agronomic situation offers the potential for spatial interpretation of critical levels of soil constraints to crop production. Such an approach might be applied to using within-field variation of properties such as soil moisture content for improving the efficiency of irrigation water use, nutrient levels for fertilizer application, or soil chemical properties for amendment needs, such as liming.

Analysis of the spatial response of crop growth to the variability of soil properties, such as nutrient uptake in response to variation of soil nutrient
parameters (Trangmar, 1982), available moisture (Tabor et al., 1984), or root penetration, may further contribute to the agronomists' understanding of the role of spatial effects in soil-crop relations. Geostatistical analysis of the incidence of pest and disease attacks in crops might aid identification of spatial sources of such attacks. Identification of a spatially dependent component of "random" error may help further reduce the confounding effects of within-plot variability on treatment effects in agricultural experimentation. The use of spatial dependence in identifying optimal plot size and spacing of samples within plots has already been described by Vieira et al. (1981).

Geostatistical analysis of spatial variation in natural phenomena has a wide range of potential applications in soil and agronomic research. In applying geostatistics, it should be remembered that semi-variograms and kriging are tools constrained by their assumptions and, where these assumptions break down, other methods of spatial analysis may be more appropriate.

REFERENCES


APPLICATION OF GEOSTATISTICS

Agricultural Canada, Ottawa.


