3 Geostatistical Methods Applied to Soil Science

A. W. WARRICK AND D. E. MYERS
University of Arizona
Tucson, Arizona

D. R. NIELSEN
University of California
Davis, California

3-1 INTRODUCTION

Variations in soil properties tend to be correlated over space—both vertically and horizontally. That is, two values taken close together tend to be more alike than two samples far apart. Most often, the classical approach in the field is to group soils together in like units or lay out small plots and assume variability within the plots or units is purely random. Conceptually, geostatistics offers an alternative approach in that spatial correlations are quantified. Estimates for a property at an unsampled location will be principally determined by measurements made close by, rather than by assuming a class (or plot) average.

Historically, the methodology for geostatistics began in mining engineering for assessment of ore bodies by D. G. Krige, for whom "kriging" is named. Matheron (1973) provided a sound theoretical basis by the formation of random functions. The approach is obviously not a panacea; nevertheless, some very difficult concepts are addressed. For example, spatial and inter-variable correlation is quantified, optimum interpolation schemes are designed, the scale of the sample is considered, and samples for different support volumes can be included. Additionally, new sampling locations can be defined in order to best improve estimates for a total population or location. When sampling locations are far apart, the approach reduces to classical random fields.

Dimensionally, applications of geostatistics could be for distances of a few molecules or kilometers. Methods can be used to analyze any number of soil properties (physical, chemical, biological) and can be extended to include plant response and crop yields. The development of the techniques was for application to very practical problems—e.g., optimizing the selection of blocks of ore to be processed on a sliding economic scale

 Methods of Soil Analysis, Part 1. Physical and Mineralogical Methods—Agronomy Monograph no. 9 (2nd Edition)

according to the market price of the end product. So far, applications to soil problems are somewhat embryonic, with versatility not yet fully exploited. Obvious choices include interpolation for preparing maps, for either transient or invariant properties. Also, sampling to attain a given precision or locating new sampling points is a logical choice. The theory for conditional simulation has not been applied to any extent and neither have the interrelationships of correlated properties.

A comprehensive treatment of the subject is by Journel and Huijbregts (1978). Other readily available texts are by Clark (1979), David (1977) and Rendu (1978). In addition, there has been an explosion of articles in soil science and hydrology journals within the past few years.

3–2 QUANTIFICATION OF SPATIAL INTERDEPENDENCE

Spatial variations with interdependence are commonly described with a correlogram or a variogram. In either case, we consider a set of values $Z(x_1)$, $Z(x_2)$, . . ., $Z(x_n)$ at x_1, x_2. . . ., x_n where each location defines a point in 1-, 2-, or 3-dimensional space. It is not a requirement that the value be for an exact point, but rather that each value is for a defined *support volume* which is centered at x. For the correlogram $\rho(h)$ (which we will define shortly), strong stationarity is required, that is

Strong stationarity (stationarity of order 2)

1. $E[Z(x)]$ exists and is equal to the same constant value for all x.
2. The covariance exists and is a unique function of separation distance h.

A weaker assumption is sufficient for the variogram function $\gamma(h)$ to be defined, namely

Weak stationarity (the intrinsic hypothesis)

1. $E[Z(x)]$ exists as above
2. For all vectors h the variance of $Z(x + h) - Z(x)$ is defined and is a unique function of h.

A system which satisfies the strong stationarity requirements also satisfies the intrinsic hypothesis, but the converse is not true.

3–2.1 Correlograms

The correlogram $\rho(h)$ of the regionalized variable Z is defined by

$$\rho(h) = \text{Cov}[Z(x),Z(x + h)]/\sigma^2 \qquad [1]$$

The covariance "Cov" is for any two values of Z at a distance h apart and σ^2 is the variance of Z. Thus, the correlogram is a series of correlations for a common variable where each couple is separated by distance h. In general, x and h are vector quantities and ρ will depend on the direction as well as the magnitude of h. The correlogram can have possible values from -1 to 1 just as can an ordinary correlation coefficient.

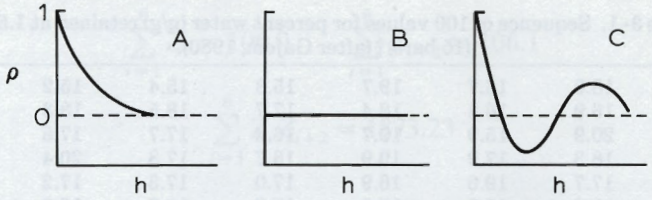


Fig. 3–1. Idealized correlograms: A is well-behaved and decreases monotonically; B is for independent or random values: and C is for a cyclical system.

Some "typical" correlograms are shown in Fig. 3–1. In Fig. 3–1A the maximum value of $\rho(h)$ is at $h = 0$ for which Eq. [1] trivially is 1. As h increases, the covariance decreases gradually until eventually, for large distances, ρ is zero and no correlation or spatial dependence exists. Thus, samples close together are alike; samples somewhat separated are less alike; and samples remote from each other are not correlated at all.

Figure 3–1B is for a variable which is not correlated over space. The value $\rho(0)$ is 1 as before, but for all other points $\rho(h)$ is zero, indicating the results are independent of separation distance. This is a purely random system. Fig. 3–1C shows results for a system showing a cyclical effect. As the distance h increases, the correlogram becomes alternatively positive and negative, and at large distances eventually approaches zero. Physical systems which may exhibit such a correlogram include sediment deposits for periodic flooding or compacted-noncompacted patterns due to wheel traffic in a row crop.

To estimate the correlogram function, let us assume $n(h)$ is the number of pairs of sample points a distance h apart. When dealing with a one-dimensional transect, perhaps all pairs of points will be in discrete classes; however, in general, a class size would be defined and $n(h)$ would refer to all pairs of points within that class. The sample covariance function is

$$C(h) = \left[\frac{1}{n(h) - 1}\right] \sum_{i=1}^{n(h)} [Z(x_i) - \overline{Z}][Z(x_i + h) - \overline{Z}] \qquad [2]$$

and the sample correlogram is

$$r(h) = C(h)/s^2 \qquad [3]$$

with $\overline{Z}$ and s^2 as estimates of the mean and variance. Ideally, the number of data pairs used would be very large for each value of h, but in practice this is normally not the case, especially for extremely small and extremely large h. Following is an example illustrating a sample correlogram.

3–2.1.1 EXAMPLE 1. A SAMPLE CORRELOGRAM FOR A 100-POINT TRANSECT OF 1.5 MPa (15-BAR) WATER VALUES.

Gajem (1980) and Gajem et al. (1981) reported values for the water retained in soil samples collected at the 50-cm depth of a Pima clay loam

Table 3-1. Sequence of 100 values for percent water (g/g) retained at 1.5 MPa (15 bars)† (after Gajem, 1980).

16.1	18.0	16.7	19.7	15.3	15.4	15.2	14.7
17.0	16.9	16.8	18.4	17.7	18.5	18.3	18.4
17.5	20.9	15.9	19.7	16.4	17.7	17.6	17.1
18.9	18.3	17.9	19.9	18.7	17.8	20.4	18.3
17.5	17.7	19.6	16.9	17.0	17.3	17.2	18.1
15.7	15.8	17.8	17.3	17.2	16.7	16.9	16.0
17.3	18.5	17.7	17.5	18.0	15.9	13.9	19.5
17.1	20.0	19.9	18.8	19.4	16.9	21.2	19.5
16.5	16.8	19.3	16.5	16.1	16.0	16.4	16.6
14.0	16.9	16.3	16.9	17.6	17.3	17.7	17.9
18.7	19.9	23.1	20.9	20.4	24.8	20.3	21.4
21.4	21.2	21.6	21.7	22.3	20.9	19.8	20.8
20.8	21.4	21.2	22.6				

† Samples were at 50-cm depth, 20 cm apart, on Pima clay loam. The sequence goes from left to right.

(fine silty, mixed thermic family of Typic Torrifluvents). In Table 3–1 are 100 values of water content at 15-bar suctions. The 100 soil samples were collected on a 20-cm spacing along a 2000-cm transect with a 7.5-cm diameter bucket auger. The depth increment was 40 to 60 cm; the experimental mean and standard deviations were 18.3 and 2.1.

For n points equally spaced at Δh, Eq. [2] may equivalently be written as

$$C(k\Delta h) = \frac{(n-k)\,\Sigma Z_i Z_{i+k} - \Sigma Z_i \Sigma Z_{i+k}}{(n-k)(n-k-1)} \qquad [4]$$

where $Z_i = Z(i\Delta h)$ and the summations are taken from $i = 1$ to $n - k$.

If we choose for example Row 5 of Table 3–1 and take $n = 8$ and $k = 1$, the corresponding sums in Eq. [4] are

$$\sum_{i=1}^{7} Z_i = 123.2 \qquad \sum_{i=1}^{7} Z_{i+1} = 123.8$$

$$\sum_{i=1}^{7} Z_i Z_{i+1} = 2178.19$$

Therefore the sample covariance for 20 cm ($k = 1$ and $\Delta h = 20$) is

$$C(20) = -4.83/42 = -0.115\,.$$

The corresponding estimate of the sample autocorrelation 1 for this row is

$$r(20) = C(20)/(s^*)^2 = -0.151$$

with $s^* = 0.873$ based on the eight values. Similarly for 40 cm ($k = 2$) and the same row:

$$\sum_{i=1}^{6} Z_i = 106.0 \qquad \sum_{i=1}^{6} Z_{i+2} = 106.1$$

$$\sum_{i=1}^{6} Z_i Z_{i+2} = 1873.23$$

resulting in

$$C(40) = -7.22/30 = -0.241$$

$$r(40) = -0.241/(0.873)^2 = -0.316$$

As demonstrated, the sample values are erratic for small series. There is no exact minimum number of points necessary for estimating $r(h)$, but generally 100 points or so will suffice.

Table 3–2 shows values of $r(h)$ for $h = 20$ to 500 cm. The value 1.0 at $h = 0$ is included for completeness. Values are also plotted as Fig. 3–2A. The $r(h)$ values begin at about 0.6 and gradually decrease towards zero at about 250 cm. The estimated values are small negative values for $h = 300$ to 480 cm, but can be interpreted as insignificant correlations. Fig. 3–2 compares well with the idealized Type A of Fig. 3–1, with the

Table 3–2. Autocorrelation values for moisture content values of Table 3–1.

Lag	h	$r(h)$	Lag	h	$r(h)$
0	cm			cm	
0	0	1.00	13	260	0.01
1	20	0.58	14	280	0.06
2	40	0.61	15	300	−0.07
3	60	0.58	16	320	0.00
4	80	0.49	17	340	−0.11
5	100	0.45	18	360	−0.07
6	120	0.38	19	380	−0.08
7	140	0.37	20	400	−0.12
8	160	0.28	21	420	−0.10
9	180	0.17	22	440	−0.07
10	200	0.11	23	460	−0.04
11	220	0.12	24	480	−0.04
12	240	0.01	25	500	0.06

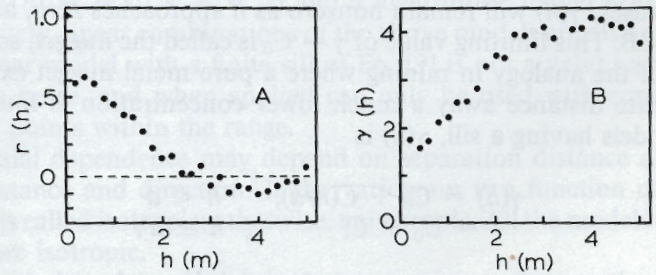


Fig. 3–2. (A) Sample correlogram and (B) variogram for the experimental values of Table 3–1.

exception that the values for small distance do not approach 1 but rather 0.6. This is not surprising, in view of the measurement error involved. On split samples from the same general area, the estimated standard deviation was determined to be approximately 0.009 (g/g) water, which would be about 45% of the estimate of the standard deviation. This will be addressed in more detail later with Example 2.

3–2.2 Variograms

The variogram $\gamma(h)$ is defined as

$$\gamma(h) = (1/2)\ \mathrm{Var}[Z(x) - Z(x + h)] \qquad [5]$$

with "Var" the variance of the argument. As for the correlogram, x and h are, in general, vectors. For 2 or 3 dimensions, both $\gamma(h)$ and $r(h)$ can be directionally dependent. Under the zero drift assumption $E[Z(x + h)] = E[Z(x)]$, and Eq. [5] is equivalent to

$$\gamma(h) = E[Z(x + h) - Z(x)]^2 . \qquad [6]$$

An estimate of γ is γ^* given by

$$\gamma^*(h) = \left[\frac{1}{2n(h)}\right] \sum_{i=1}^{n(h)} [Z(x_i + h) - Z(x_i)]^2 \qquad [7]$$

with $n(h)$ the number of pairs separated by a distance h.

If the strong stationarity conditions are met, then both $\rho(h)$ and $\gamma(h)$ exist and Eq. [1] and [6] may be used to show that

$$\gamma(h) = \sigma^2\ [1 - \rho(h)], \text{ (strong stationarity)} \qquad [8]$$

Examples of $\gamma(h)$ are given in Fig. 3–3. Fig. 3–3A shows a typical linear variogram starting at $\gamma(0) = 0$ and reaching a maximum or "sill" value $\gamma(h) = C$ for $h \geq a$. The value $h = a$ is called the *range* and is the maximum separation distance for which sample pairs remain correlated. In some cases, $\gamma(h)$ will remain nonzero as h approaches zero, as shown in Fig. 3–3B. This limiting value of $\gamma = C_0$ is called the *nugget*, so named because of the analogy in mining where a pure metal nugget exists and at any finite distance away a much lower concentration is found. For linear models having a sill, $\gamma(h)$ is

$$\begin{aligned} \gamma(h) &= C_0 + C(h/a), & h < a \\ &= C_0 + C, & h \geq a . \end{aligned} \qquad [9]$$

Another useful model is the spherical model illustrated by Fig. 3–3C and given by

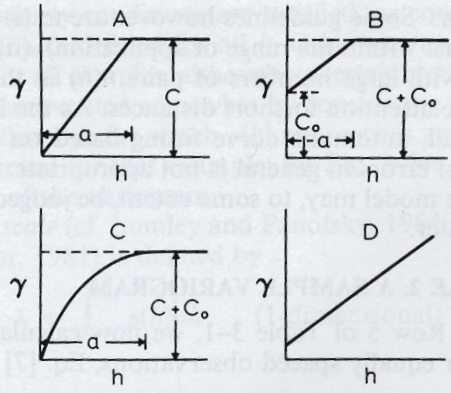


Fig. 3–3. Idealized variograms; A is a linear model with range a and sill C; B is a linear model with range a, sill C and nugget C_0; C is a spherical model with range a; and D is a linear model without a sill.

$$\gamma(h) = C_0 + C\,[(3/2)(h/a) - (1/2)(h/a)^3], \qquad 0 < h < a$$
$$= C_0 + C, \qquad h \geq a \qquad [10]$$

where, as before, C_0, $C_0 + C$, and a are the nugget, sill, and range, respectively.

The preceding examples of variograms assume that $\gamma(h)$ reaches a constant maximum value for h large. Such is not always the case. Figure 3–3D is for such a system and shows $\gamma(h)$ continuing to increase with h, at least on the scale over which the figure is drawn. For such a system, the intrinsic hypothesis is satisfied, but the strong stationarity conditions are not met. Consequently there can be no correlogram, as the variance is unbounded. However, the variogram function remains defined and is of value.

The choice of valid variogram models is restricted such that the negative of $\gamma(h)$ is a *positive-definite* function (Journel and Huijbregts, 1978; Armstrong and Jabin, 1981). It is best to use models known to behave properly. When a sill exists, these include a spherical model, an exponential model $\gamma(h) = 1 - \exp(-h/a)$, and a Guassian model $\gamma(h) = 1 - \exp(-h^2/a^2)$. When a sill does not exist, included are a power model $\gamma(h) = h^\alpha$ $(0 < \alpha < 2)$ and a logarithmic model $\gamma(h) = \log h$. Fortunately, linear combinations of the above models are also acceptable. The linear model with a finite sill of Eq. [9] is not a valid variogram in the true sense, and when applied can only be used with confidence in relating points within the range.

Spatial dependence may depend on separation distance only or on both distance and direction. If the variogram is a function of distance only, it is called isotropic; otherwise, anisotropic. All the models described above are isotropic.

The choice of the model used to approximate the sample variogram is often somewhat subjective. No fool-proof procedure exists for best-

fitting all situations. Some guidelines however are to (i) use a valid functional form (at least within the range of application), (ii) lend more credibility to points with large numbers of pairs $n(h)$ in the determination, and (iii) pay more attention to short distances. As the behavior at short distances is critical, automatic curve fitting based on minimization of sums of squares of errors in general is not appropriate. The adequacy of a given variogram model may, to some extent, be judged by *jack-knifing*, to be discussed later.

3–2.2.1 EXAMPLE 2. A SAMPLE VARIOGRAM

Returning to Row 5 of Table 3–1, we now calculate a sample variogram value. For equally spaced observations, Eq. [7] reduces to

$$\gamma^*(k\Delta h) = (1/2n)\sum_{i=1}^{n}(Z_{i+k} - Z_i)^2.$$

With $k = 1$ and $n = 7$, the result for Row 5 is

$$\gamma^*(20) = (1/14)[(17.7 - 17.5)^2 + (19.6 - 17.7)^2 + \ldots + (18.1 - 17.2)^2] = 0.85.$$

Similarly with $k = 2$ and $n = 6$, the result for $\gamma(40)$ is

$$\gamma^*(40) = (1/2)[(19.6 - 17.5)^2 + (16.9 - 17.7)^2 + \ldots + (18.1 - 17.3)^2] = 1.05.$$

Of course, for such a short series, the values are not reliable.

For the total 100-point transect of Table 3–1, sample variogram values were calculated for each lag and given as Fig. 3–2B. The value for small distances is about 1.7 and increases to a value of about 4 at $h =$ 250 cm. The values are shown up to about 500 cm or 1/4 of the total transect, roughly the limit of reliability. The results can be reasonably modeled by a linear or a spherical model (e.g. Fig. 3–3B, 3–3C) provided a nugget is included. Whether a sill exists is not clear; i.e. whether the variance would remain finite if the transect were much longer is not clear, but if it were, then Eq. [8] would relate $\rho(h)$ to $\gamma(h)$ and vice versa. We can compare for small h the approximate nugget value 1.7 to the sample $(2.1)^2[1 - (0.6)^2] = 2.8$. Similarly, the largest value shown on Fig. 3–2B is about 4 and can be compared to the sample variance $(2.1)^2$. If $\gamma(h)$ is extended beyond 500 cm with this data set, it decreases and then increases to about 6 at 1000 cm, but the results are not reliable for the longer distances.

3–2.3 Range of Influence and Integral Scale

A natural question to ask is "What is the range of influence beyond which values are independent of each other?" Such a limit exists provided

that strong stationarity conditions are satisfied, or equivalently that the covariance function is defined for all h. Otherwise a degree of interdependence exists for all samples, although uncertainty regarding large distances can be circumvented somewhat by considering *moving neighborhoods* only. The distance for which $\rho(h)$ approaches zero in Fig. 3–1A is the range of influence for that example; for Fig. 3–3A, 3–3B or 3–3C, it will be the range of the variogram.

The *integral scale* (cf. Lumley and Panofsky, 1964; Bakr et al., 1978; Russo and Bresler, 1981) is defined by

$$\lambda = \int_0^\infty \rho(h)dh \qquad \text{(1-dimensional)} \qquad [11]$$

or

$$\lambda = [2\int_0^\infty h\rho(h)dh]^{0.5} \qquad \text{(2-dimensional)}. \qquad [12]$$

The estimated value depends to some extent on the choice of sampling points. There is a tendency for samples over larger regions to result in larger integral scales. Dependence of integral scale to measuring area is a complex problem involving sample variation, scale of measurement and stationarity.

3–2.3.1 EXAMPLE 3. REPORTED VALUES FOR RANGE OF INFLUENCE AND INTEGRAL SCALE

Some reported values of range of influence and integral scales are reported in Table 3–3 for a number of different parameters and locations. The definition and method of determination of range or scale differ somewhat. Overall generalizations are difficult, but, for the most part, the larger the area sampled, the larger is the range. For example, on the same site Gajem et al. (1981) found ranges of 1.5, 21, and 260 m for pH values of 100-member transects spaced at 0.2, 2, and 20 m. Campbell (1978) found random values on 8 × 20 grids at 10-m spacings, and Yost et al. (1982) obtained values of 14.3 km on long transects in Hawaii. Interpretations and comparisons are extremely difficult, but the overall scale of the experiment is a determining factor on the ranges. Simple explanations such as row directions, changing soil types, and topography offer a deterministic answer for some of the results but will not explain many others. The question of whether homogeneity and proper stationarity exist is always valid. Fortunately for many applications, the exact range is a moot consideration, as observations at closer distances dominate the end results—e.g. kriging estimates.

3–3 PUNCTUAL KRIGING

A primary application of geostatistics is for estimating values at locations where measurements have not been made. The most common

Table 3-3. Reported values of integral scale or range.

Source	Parameter	Range or scale	Site
		m	
Al-Sanabani (1982)	Log of saturated EC	>90	Tucson fine loam, Typic Haplargids (Arizona) 10 ha, 101 random samples, 0–30 cm in depth
Burgess and Webster (1980a)	Sodium	61	Approx. 50 ha, Plas Gogerddan (Gr. Britain), 440 samples, 0–15 cm depth
	Depth cover loam	100	Approx. 18 ha, Hole Farm (Gr. Britain), 450 observations
Campbell (1978)	Sand content	30	Ladysmith series, mesic Pachic arguistolls (Kansas), 8 × 20 grid at 10-m spacing in B2 horizons
	Sand content	40	Pawnee series, mesic Aquic Argiudoll (Kansas) (as above)
	Soil pH	Random	Pawnee and Ladysmith
Clifton and Neuman (1982)	Log of transmissivity	9 600	Avra Valley (Arizona), about 15 × 50 km, 148 wells
Folorunso and Rolston (1984)	Flux of N_2 and N_2O at surface	<1	Yolo loam, Typic Xerorthents (California) 100- by 100-m area
Gajem et al. (1981)	Sand content	>5	Pima clay loam, Typic torrifluvents (Arizona), 20-m transect, 20-cm spaces, 50-cm depth
	Soil pH	1.5	Pima, as above, 4 transects
		21	Pima, as above but 4 transects, 2-m spacing
		260	Pima, as above, 1 transect, 20-m spacing, 100 points
	1.5 MPa	0.6	Pima, 20-cm spacing, 4 transects, each 100 points
		>32	As above, 2-m spacing
		150	As above, 20-m spacing
Hajrasuliha et al. (1980)	Saturated EC	<80	Clay loam to loam, Haft Tappeh Plantation (Iran). 0–1m depth, 150 ha, 232 points, (Site 1)
	Saturated EC	>1 200	As above, 455 ha, 710 points (Site 3)
Kachanoski et al. (1985)	Depth of A-horizon and mass of A-horizon	<2	Mix of Typic Haploborolls and Typic Argiborolls (Saskatchewan)
Liss (1983)	Water-soluble organic carbon	<8	Yolo loam, Typic Xerorthents (California) 100- by 100-m area
	Soil water content of 0–10 cm soil depth	<16	As above

(continued on next page)

scenario would assume that $Z(x_1)$, $Z(x_2)$, . . ., $Z(x_n)$ are known at locations x_1, x_2, . . ., x_n and that the variogram (or correlogram) is determined as in the previous section. The question that remains is to estimate the value Z^* at position x_0. The procedure leads to not only an optimal

Table 3-3. Continued.

Source	Parameter	Range or scale	Site
		m	
Russo and Bresler (1981a, 1981b)	Saturated conductivity	34	Surface, Harma Red Mediterranean, Rhodoxeralf (Israel). 30 random sites in 0.8 ha.
		14	90-cm depth, as above.
	Saturated water content	76	Surface, as above.
		28	90 cm, as above.
	Sorptivity	37	Surface
		39	90 cm, as above.
	Wetting front	16–30	Simulated for above site, 1 to 12.5 h
Sisson and Wierenga (1981)	Steady-state infiltration	0.13	Sandy clay loam, Typic torrifluvent (New Mexico). 6.4- by 6.4-m plot, transect of 125 contiguous 5-cm rings
van Kuilenberg et al. (1982)	Moisture supply capacity	600	Cover sand, 30 mapping units, 9 soil types including Haplaquods, Humaquepts, and Psammaquents (Netherlands). 2 by 2 km, 1191 borings
Vauclin et al. (1983)	Sand content	35	Sandy clay loam (Tunisia). 7 × 4 grid at 10-m spacing, 20–40 cm depth
	pF 2.5	25	Same
Vauclin et al. (1982)	Surface soil temperature	8–21	Yolo loam clay, Typic Xerorthents (California). 60 and 100 m transects, 1-m spacing
Vieira et al. (1981)	Steady-state infiltration	50	Yolo loam, Typic Xerorthents (California). 55- × 160-m area.
Wollum and Cassel (1984)	Log of most probable number of		Pocalla loamy sand, thermic Arenis Plinthic (N. Carolina),
	Rhizobium japonicum	1	0°, 3-m spacing
		>12	0°, 20-cm spacing
		Random	90°, 3-m spacing
		>12	90°, 20-cm spacing
Yost et al. (1982)	Soil pH	14 000–32 000	Various transects on Island of Hawaii at 1- to 2-km intervals, 0–15 cm depth.
	Phosphorus sorbed at 0.02 mg P/L	32 000	As above
	Phosphorus sorbed at 0.2 mg P/L	58 000	As above

solution for Z^*, but also an estimate of $\mathrm{Var}(Z - Z^*)$ which indicates the reliability of the result.

An estimate of Z^* is assumed to be a linear function (nonlinear estimates are rarely used) of known values:

$$Z^*(x_0) = \sum_{i=1}^{n} \lambda_i Z(x_i). \qquad [13]$$

The best linear estimate is found by choosing the weight factors λ_i such that the expected value and variance of $Z^*(x_0) - Z(x_0)$ are 0 and a minimum, respectively, that is,

$$\mathrm{E}[Z^*(x_0) - Z(x_0)] = 0 \qquad [14]$$

$$\mathrm{Var}[Z^*(x_0) - Z(x_0)] = \text{a minimum} . \qquad [15]$$

Of course, the "true value" $Z(x_0)$ is not known. If we assume the expected value of Z is not known, then the first condition (Eq. [14]) which guarantees $Z^*(x_0)$ to be an unbiased estimate results in

$$\sum_{i=1}^{n} \lambda_i = 1. \qquad [16]$$

This leaves $\mathrm{Var}[Z^*(x_0) - Z(x_0)]$ to be minimized subject to the constraint that the λ_i's sum to 1. This is done by introducing a Lagrangian multiplier -2μ and minimizing

$$\mathrm{Var}\left[Z^*(x_0) - Z(x_0)\right] - 2\mu\left(\sum_{i=1}^{n} \lambda_i - 1\right). \qquad [17]$$

By definition and by Eq. [13], it follows that

$$\mathrm{Var}\left[Z^*(x_0) - Z(x_0)\right] = -\sum_{i=1}^{n}\sum_{j=1}^{n} \lambda_i\lambda_j\gamma_{ij} + 2\sum_{j=1}^{n} \lambda_j\gamma_{0j} \qquad [18]$$

where γ_{ij} is defined by

$$\gamma_{ij} = \gamma(x_i - x_j) . \qquad [19]$$

Substitution of the right side of Eq. [18] for the "Var" term of Eq. [17] and taking the partial derivatives of the result with respect to each λ_i gives the set of linear equations (cf. Burgess and Webster, 1980a, esp. p. 310–321):

$$\sum_{j=1}^{n} \lambda_j\gamma_{ij} + \mu = \gamma_{i0}, \qquad i = 1, 2, \ldots, n. \qquad [20]$$

In matrix notation, the equivalence is

$$\mathrm{A}\begin{bmatrix} \lambda \\ \mu \end{bmatrix} = b \qquad [21]$$

where

$$A = \begin{bmatrix} \gamma_{12}\gamma_{21} & \cdots & \gamma_{n1} & 1 \\ \gamma_{21}\gamma_{22} & \cdots & \gamma_{n2} & 1 \\ \vdots & & \vdots & \vdots \\ \gamma_{n1}\gamma_{n2} & \cdots & \gamma_{nn} & 1 \\ 1 & \cdots & 1 & 0 \end{bmatrix} \quad [22]$$

and the column matrices given by (the transposes of)

$$\begin{bmatrix} \lambda \\ \mu \end{bmatrix}^T = [\lambda_1 \lambda_2 \ldots \lambda_n \mu] \quad [23]$$

$$b^T = [\gamma_{10}\ \gamma_{20} \ldots \gamma_{n0}\ 1] \quad [24]$$

Thus, the solutions for the λ's and μ are

$$\begin{bmatrix} \lambda \\ \mu \end{bmatrix} = A^{-1}b \quad [25]$$

Furthermore, by Eq. [18], the minimum estimation error is

$$\sigma_E^2 = b^T \begin{bmatrix} \lambda \\ \mu \end{bmatrix} \quad [26]$$

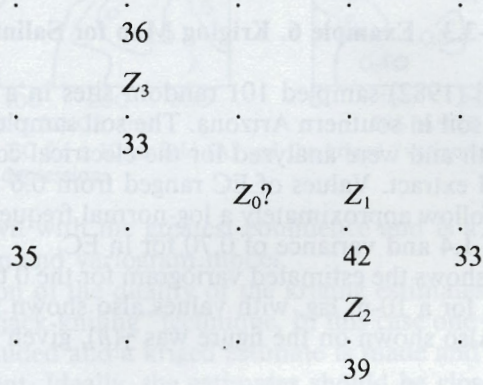
When the nugget is nonzero, Eq. [21] may be evaluated with γ_{ii} (for $i = 1$ to n) equal to the nugget or taken as 0. In either case, the weights λ_i are algebraically the same. The Lagrange multiplier when γ_{ii} is taken as 0 is $\mu + C_o$ where C_o is the nugget and μ the value when $\gamma_{ii} = C_o$.

An alternative to Eq. [21] is the "covariance" form where the γ_{ij} entries are replaced by the corresponding covariance (for i and j from 1 to n). The elements $a_{i,n+1}$ are changed from 1 to -1 for $i = 1$ to n. In place of Eq. [26], the kriging variance is $\sigma^2 - \mu - \Sigma \lambda_i C_{oj}$, where C_{oj} is the covariance corresponding to the distance between points "o" and "j".

The most tedious calculation in the procedure is the inversion of the matrix A. If the number of points used for the estimator is large, A and the necessary machine operations become unwieldy. Fortunately, the number of points necessary for the estimate may be relatively small (10 or less), as the inclusion of other weighting factors affects the results only negligibly. Another labor-saving feature, especially for data on regular grids, is that the A matrix is dependent upon the sampled locations only

3–3.2 Example 5. Punctual Kriging–Two-dimensional

Suppose Z_0 is to be estimated from its three nearest points, using the hypothetical values below, and with $\gamma(h) = 4h$. (The grid spacing is 1 unit.)



The closest three points are Z_1, Z_2, and Z_3. If relevant values for the variogram are

$$\gamma_{12} = \gamma_{10} = 4;\ \gamma_{13} = 4\sqrt{5};\ \gamma_{23} = 4\sqrt{8} = 8\sqrt{2}$$
$$\gamma_{20} = \gamma_{30} = 4\sqrt{2}$$

the system of Eq. [21] becomes

$$\begin{bmatrix} 0 & 4 & 8.9 & 1 \\ 4 & 0 & 11.3 & 1 \\ 8.9 & 11.3 & 0 & 1 \\ 1 & 1 & 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \lambda_3 \\ \mu \end{bmatrix} = \begin{bmatrix} 4 \\ 5.66 \\ 5.66 \\ 1 \end{bmatrix}$$

which has the solution

$$\lambda_1 = 0.398 \qquad \lambda_2 = 0.215$$
$$\lambda_3 = 0.387 \qquad \mu = -0.308.$$

Thus, the estimated value is

$$Z_0^* = \sum_{i=1}^{3} \lambda_i Z_i = 37.9$$

with a kriging variance of

$$\sigma_E^2 = \mu + \sum_{i=1}^{3} \lambda_i \gamma_{i0} = 4.69.$$

Note that the diagonal entries of A are all zero for this example, as the nugget is zero.

3–3.3 Example 6. Kriging Map for Salinity

Al-Sanabani (1982) sampled 101 random sites in a 10-ha field of a typic haplargid soil in southern Arizona. The soil samples were from the 0-to 30-cm depth and were analyzed for the electrical conductivity (EC) of the saturated extract. Values of EC ranged from 0.6 to 32 dS/m and were found to follow approximately a log-normal frequency distribution with a mean of 1.4 and variance of 0.70 for ln EC.

Figure 3–4 shows the estimated variogram for the 0 to 125 m. Values were calculated for a 10-m lag, with values also shown for 5-m lags out to $h = 20$ m. Also shown on the figure was $\gamma(h)$, given by the spherical model

$$\begin{aligned} \gamma(h) &= 0.3 + 0.6\,[1.5(r/160) - 0.5(r/160)^3], && r \leq 160 \\ &= 0.9 && r > 160\,. \end{aligned} \quad [27]$$

(This model was verified to give low error between measured and kriged estimates, as discussed later.) Directional variograms were also calculated, but little if any difference by direction was in evidence. The values are not shown beyond 125 m, as the field was only 300 by 350 m. The relationship shows considerable dependence out to at least 100 m.

A map was prepared using the variogram model and the 10 closest points on a 15- by 15-m grid, using the computer algorithm of Carr et al. (1983). Contours were drawn with the results shown in Fig. 3–5A. Generally, a low salt region exists through the center of the field from southwest to northeast with ln EC < 1. A high salt area is to the east with a sizeable area in excess of ln EC $= 2$ to the east. The kriging variance σ_E^2 is mapped as Fig. 3–5B. The kriging variance indicates which regions

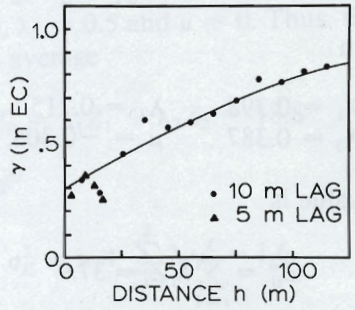


Fig. 3–4. Variogram for EC measurements. Solid line is spherical model with nugget, range, and sill as 0.3, 160, and 0.9, respectively. (Data from Al-Sanabi, 1982.)

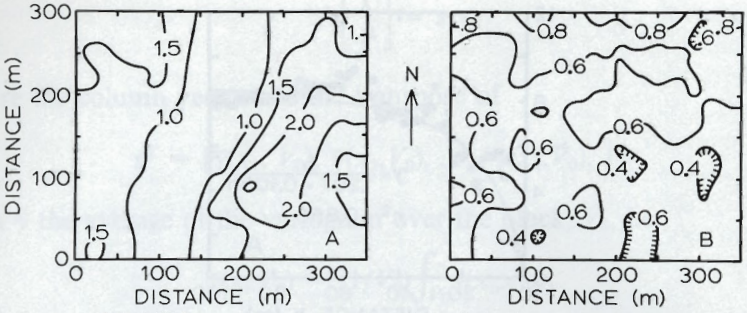


Fig. 3–5. Map of ln EC for a 10-ha field (A) and the kriged "variance" (B). Hatched line indicates interior depression.

tend to be known with the greatest confidence and is a function of the sampling pattern and variogram model.

An indication of the quality of the kriging estimates is by a "cross-validation" or "jack-knifing" technique. In this case one measured point at a time is excluded and a kriged estimate is made and compared with that measurement. Ideally, the estimates should be close to the experimental points and indicate no bias. Doing this for the 101 experimental points of the EC measurements and the variogram of Eq. [27] resulted in an average absolute difference of 0.010 with a variance of 0.57. As the difference is small, the model is judged to be adequate. The variance value of 0.60 is reasonably close to the average σ_E^2 of 0.49.

3–3.4 Example 7. Efficiency of Sampling—Infiltration

Vieira et al. (1981) measured limiting infiltration rates over a 160- by 55-m area of Yolo loam (fine-silty, mixed, nonacid, thermic Typic Xerothents). A total of 1280 measurements were made in eight columns of 160 measurements. Water was ponded in 46-cm diameter single rings for 36 hours, at which time steady rates were measured.

The 1280 values were approximately normally distributed with a mean value of 7.0 mm h^{-1} and variance of 7.8 mm^2h^{-2}. The sample values were highly dependent on position, as evidenced by the sample variogram of Fig. 3–6A. They addressed the question of what minimum number of samples would give results similar to the true 1280 (or more measured values). For a first trial, only 16 measured values are used for the entire field and the remaining 1264 positions was found by kriging. The 1264 kriged values resulted in a correlation coefficient (r) of about 0.16 with the actually measured points. The correlation coefficient is plotted by the first point on Fig. 3–6B. Repetition of the process was done assuming 32, then 64, 128, and 256 points were known resulted in increasing r values as shown in the figure. If an r^2 of 0.8 between measured and estimated values is acceptable, then 128 sampling points give similar results to the 1280 measured values. Figure 3–6C is a scatter diagram showing measured vs. kriged values based on 256 samples.

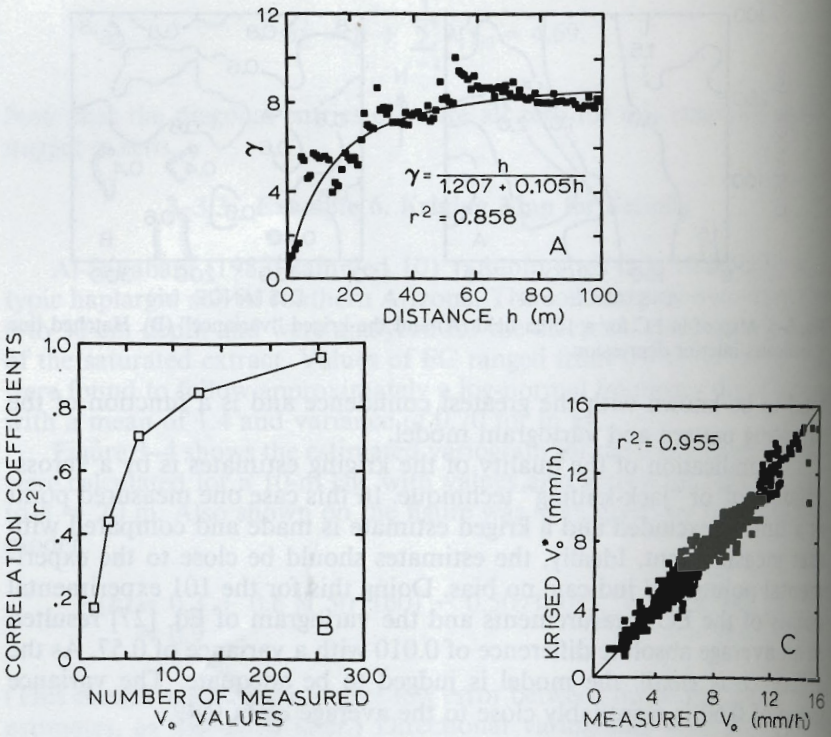


Fig. 3–6. (A) Variogram; (B) correlation for data points with kriged values; and (C) scatter diagram for 1024 data points with kriged values. (After Vieira et al., 1981.)

3–4 BLOCK KRIGING

Not only punctual (point) values are of interest, but also averages over finite regions. The average value $\overline{Z}(x_0)$ over the region V_0 is

$$\overline{Z}(x_0) = (1/V_0)\int_{V_0} Z(x)dx \qquad [28]$$

where V_0 is centered at x_0. The *block* V_0 can be a finite line, an area, or a volume, depending on whether Z is defined in one, two, or three dimensions. Block averages are smoother than point values and regional effects are displayed more clearly.

The kriging estimate $\overline{Z}^*(x_0)$ is

$$\overline{Z}^*(x_0) = \sum_{i=1}^{n} \lambda_i Z(x_i) \qquad [29]$$

(The support of the sample itself could be formally introduced into the discussion, but for simplicity we assume samples are taken at points.) Proceeding as before, Eq. [16] is still valid ($\Sigma\, \lambda_i = 1$) if the estimate is to be unbiased. Minimizing the variance, $\mathrm{Var}[\overline{Z}^*(x_0) - \overline{Z}(x_0)]$, leads to (cf. Burgess and Webster, 1980b, esp. p. 335)

$$A\begin{bmatrix} \lambda \\ \mu \end{bmatrix} = s \qquad [30]$$

where the column vector s is the transpose of

$$s^T = [\bar{\gamma}(x_1, V_0), \bar{\gamma}(x_2, V_0), \ldots, \bar{\gamma}(x_n, V_0), 1] \qquad [31]$$

with $\bar{\gamma}$ the average of the variogram over the block, V_0, i.e.

$$\bar{\gamma}(x_i, V_0) = (1/V_0)\int_{V_0} \gamma(x_i, x)dx \qquad [32]$$

The solution of the linear system of Eq. [32] is as for the punctual case. The resulting kriging variance, however, is

$$\bar{\sigma}^2 = \sum_{i=1}^{n} \lambda_i \bar{\gamma}(x_i, V_0) + \mu - (1/V_0)\int_{V_0} \bar{\gamma}(x, V_0)\, dx \qquad [33]$$

In general, the variance for block kriging is smaller than for point estimates by an amount corresponding to the last term. This last term of Eq. [33] is a "within block" term analogous to "within" effects of classical statistics. This is a special case of the extension variance of V by a support value v (cf. Journel and Huijbregts, 1978, p. 54):

$$\bar{\sigma}_E^2 = 2\bar{\gamma}(V,v) - \bar{\gamma}(V,V) - \bar{\gamma}(v,v) \qquad [34]$$

where

$$\bar{\gamma}(V,v) = [1/(vV)]\int\int_{v,V} \gamma(x_i, x_j)dvdV \qquad [35]$$

3–4.1. Example 8. Block Kriging for Salinity

The EC values of Example 6 were used with block kriging, resulting in the map of Fig. 3–7A. The kriging estimates are for the center of 50- by 50-m blocks based on the closest 10 points and were from the computer

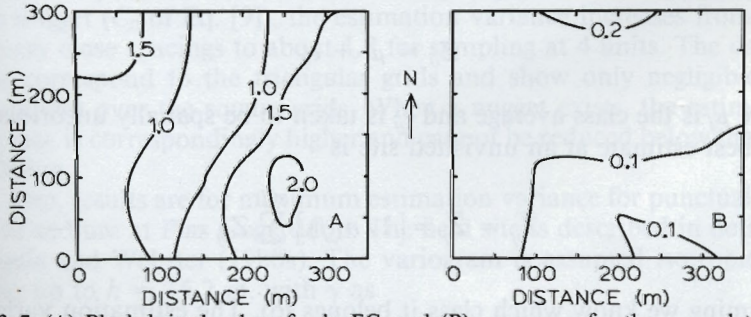


Fig. 3–7. (A) Block kriged values for ln EC, and (B) error map for the same data and variogram as for the punctual results in Fig. 3–5A.

code of Biafi (1982). The result is a smoothing of values compared to punctual kriging. The high value from southwest to northeast through the center is evident, but the boundaries are smoothed. The high value region (ln EC $>$ 2) is reduced to a single area to the lower right. The error map for the block kriging is given as Fig. 3–7B. The values are much smaller than the values for punctual kriging (about 0.1, compared to 0.4 to 0.8). This is evidence of the error reduction analogous to the "within block" component.

3–5 SAMPLING STRATEGIES FOR SPECIFIED ESTIMATION ERROR

Scientists almost without exception are confronted again and again with the question of how many locations to sample and how to best locate them. The best known relationship is that the sampling number n necessary to be within a specified value of the population mean is

$$n = z_\alpha^2 \sigma^2 / (x - \mu)^2 \qquad [36]$$

where σ^2 and μ are the population variance and mean values, $|x - \mu|$ the allowable deviation to be attained $(1 - \alpha)$ (100)% of the time, and z_α the two-tailed normalized deviate ($z_{0.05} = 1.96$; $z_{0.1} = 1.645$; $z_{0.5} = 0.842$). The assumptions are (i) independence of samples and (ii) n sufficiently large that the "central limit theorem" applies. For application, the best estimate of σ^2 and μ would be used. (For estimating confidence limits when s^2 and n are already specified, the appropriate Student t value is inserted for z_α, but it should not be used to estimate a sampling number required.) The problem of sampling strategies and estimation variance has recently been addressed by Burgess et al. (1981), McBratney et al. (1981), and McBratney and Webster (1983a).

The classical approach to reducing the sampling size required is to logically break the area (or appropriate population) into classes. The appropriate model is

$$Z_{ij} = \mu_j + \epsilon_{ij} \qquad [37]$$

where μ_j is the class average and ϵ_{ij} is taken to be spatially uncorrelated. The best estimate at an unvisited site is

$$\bar{Z}_{ij} = \bar{Z}_j = [1/n(j)] \sum_{i=1}^{n(j)} Z_{kj} \qquad [38]$$

(assuming we know which class it belongs to). The estimation variance σ_E^2 is

$$\sigma_E^2 = \text{Var}(\overline{Z}_{ij} - Z_{ij}) = \text{Var}(\overline{Z}_j) + \text{Var}(\epsilon_{ij})$$

or

$$\sigma_E^2 = \sigma^2 + \sigma^2/[n(j)] \qquad [39]$$

The σ^2 is the within-class variance and $\sigma^2/[n(j)]$ is the variance of the class mean. By increasing $n(j)$, better and better estimates of the class mean are found but the estimation variance at an unvisited site can never be less than the estimation variance σ^2.

The above approach implies that the initial subdivision accounts for all spatial variations and that ϵ_{ij} is purely random. In addition, it assumes that class lines are sharp. This may be the best approach if only very limited data are available, but leads to overly conservative estimates of precision. An alternate is to use Eq. [13] for punctual kriging or Eq. [29] for block kriging. Example 4 illustrates this.

3–5.1 Example 9. Sampling Error

Burgess et al. (1981) and McBratney et al. (1981) considered the design of optimal sampling schemes using regionalized variables. Square and triangular sampling grids are considered. The maximum distance between an interpolated point and the nearest observation will be 0.62 for unit triangles and $1/(2)^{1/2} = 0.71$ for unit squares at x_0. The maximum estimation variance will correspond to these points. The kriged variance at the maximum point can be calculated a priori by Eq. [26] and will depend on $\gamma(h)$ and the spacing. In general, the triangular (hexagonal) geometry is more efficient (although not by much, since x_0 for the square grid has four close neighbors rather than 3). The slight advantage of the triangular pattern is largely mitigated by the simplicity of sampling on a square grid.

Firstly, linear variograms are considered with unit slope and nugget of 0, 1, and 2. The estimation was by the 25 nearest neighbors (a computer code OSSFIM is given by McBratney and Webster, 1981). Resulting estimation variance as a function of sample spacing is in Fig. 3–8A. For zero nugget (C_0 of Eq. [9]), the estimation variance increases from zero for very close spacings to about 1.8 for sampling at 4 units. The dashed lines correspond to the triangular grids and show only negligible improvement over the square grids. When a nugget exists, the estimation variance is correspondingly higher and cannot be reduced below the nugget value.

Also, results are for maximum estimation variance for punctual kriging of sodium at Plas Gogerddon. The field site is described in detail by Burgess and Webster (1980a). The variogram is assumed isotropic and linear up to $h = 15.2$ m, with γ as

$$\gamma(h) = 8.7 + (1.69/15.1)h, \qquad h < 60.1\,. \qquad [40]$$

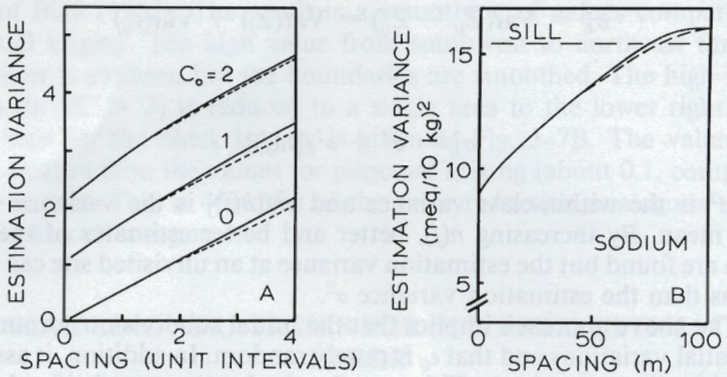


Fig. 3–8. Estimation of variance as a function of spacing. Effects of varying nuggets are shown in A. Sodium results at Plas Gogerddan are in B. Solid lines are for square patterns, dashed for triangular. (After Burgess et al., 1981.)

The resulting maximum estimation variance is given as Fig. 3–8B, with results again calculated by the nearest 25 neighbors. The value starts at 8.7 (the nugget) for very small spacings and increases with distance until it levels off at about 60 m which was the maximum distance over which dependence is assumed. The estimated variance overshoots the sill value for dependent samples. This is an anomaly of the calculations, in that 25 points were chosen, giving a long-range estimation of $[8.7 + (1.69/15.1)(60.1)](1 + 1/25) = 16.0$. Had more points been chosen, this would be reduced to the sill value of about 15.5. This reiterates that for spacings beyond the range of influence, kriging estimates reduce to the same results as for random sampling. Several more examples, including those for block kriging, are given in the above references.

3–6 FURTHER APPLICATIONS

3–6.1 Universal Kriging

Both the strong stationarity and intrinsic hypotheses imply an assumption of zero drift, that is,

$$E[Z(x + h) - Z(x)] = 0 \qquad [41]$$

In practice this assumption is scale-related and may be satisfied by partitioning the region into sub-regions. In other instances there may be a strong or pronounced drift, and Eq. [20] and [21] must be changed. Since this drift is generally unknown it is necessary to model or estimate it. The simplest model is given by

$$E[Z(x)] = \sum_{i=0}^{p} a_i f_i(x) \qquad [42]$$

where $f_0, f_1, \ldots, f_p$ are known, linearly independent functions, but $a_0, \ldots, a_p$ are unknown. While knowledge of these coefficients is necessary to estimate the variogram, it is not necessary for kriging. The estimator

$$Z^*(x_0) = \sum_{i=1}^{n} \lambda_i Z(x_i) \qquad [43]$$

is of the same form as Eq. [13] but Eq. [20] becomes

$$\sum_{i=1}^{n} \lambda_i \gamma(x_i, x_j) + \sum_{k=0}^{p} \mu_k f_k(x_j) = \gamma(x_0, x_j) \qquad [44]$$

and Eq. [42] becomes

$$\sum \lambda_i f_j(x_i) = f_j(x_0) \qquad j = 0, \ldots, p. \qquad [45]$$

The kriging variance is given by

$$\sigma^2_{UK} = \sum \lambda_i \gamma(x_i, x_0) + \sum_{k=0}^{p} \mu_k f_k(x_0). \qquad [46]$$

The μ_k's are Lagrange multipliers corresponding to the $p + 1$ constraints which are required to insure that the estimator is unbiased. The kriging variance is in general larger due to the uncertainty associated with modelling the drift.

The simplest choices for the drift functions are polynomials, [e.g., in one dimension $f_0(x) = 1, f_1(x) = x, f_2(x) = x^2, \ldots$].

The difficulty in utilizing universal kriging is the circular nature of the problem. Estimation/fitting of $\gamma(h)$ requires the drift function, but the coefficients in Eq. [42] can only be estimated optimally if the variogram is known. Some authors have used least squares to fit the drift. This will result in a bias in estimating the variogram. If a linear variogram is used this bias is known. Burgess and Webster (1980b) have given an example of universal kriging and show the computation of the bias.

Even in mining applications, universal kriging has received limited application because of these difficulties. There are several possible solutions. If the drift is at most second order, the samples are taken on a regular grid or transect, and a linear variogram is postulated, then least squares fitting can be used to model the drift and the bias in the variogram compensated. If the data are in two or three dimensions, the drift may be only in one direction and the variogram can be modeled using only data in other directions. In this case universal kriging may be used without difficulty. More recently a different formulation using generalized covariances has been developed. As yet there are few if any examples of applications to problems in soil science. There are several computer programs commercially available such as BLUEPACK (available through

Geomath, Inc., 4891 Independence Street, Suite 250, Wheatridge, CO 80023). The method is described in greater detail in Delfiner (1976).

3–6.2 Co-Regionalization

Kriging utilizes the spatial dependence of a particular soil characteristic. Some attributes such as clay and wilting percentage are dependent. This dependence can be used in estimation as well as the spatial dependence. When one or more variables are estimated by a linear combination using both the spatial and inter-variable dependence, the technique is known as *co-kriging* or *co-regionalization.*

Co-kriging is utilized for several kinds of problems. For example, Vauclin et al. (1983) reported available water content (AWC), pF 2.5 (pF = log of soil matric potential as cm of water), and sand values sampled on a regular grid at 10-m intervals. The AWC and pF 2.5 values were then co-kriged on a 5-m grid. It is also possible to replace missing or insufficient data for one variable by data on other variables by utilizing the intervariable dependence. This is known as the "undersampled problem." In both forms of co-kriging the objective is to reduce the kriging variance.

The most compact form of the co-kriging equations is given in matrix form. Let $Z_1(x),\ldots,Z_m(x)$ be the random functions representing the soil attributes where x is a location in 1-, 2-, or 3-space. If

$$\bar{Z}(x) = [Z_1(x), \ldots, Z_m(x)] \qquad [47]$$

and $x_1, \ldots, x_n$ are sample locations, the estimator for $\bar{Z}(x)$ is written as

$$Z^*(x) = \sum_{j=1}^{n} Z(x_j)\Gamma_j \qquad [48]$$

where $\Gamma_1, \ldots, \Gamma_n$ are $m \times m$ matrices with entries λ^i_{jk}. The λ^i_{jk} is the weight for location i given to Z_j in estimating Z_k.

The system of equations used to obtain the Γ_i's is given by

$$\sum \bar{\gamma}(x_i - x_j)\Gamma_i + \bar{\mu} = \bar{\gamma}(x_i - x_0)$$

$$\sum \Gamma_i = I \qquad [49]$$

Each $\bar{\gamma}$ is an $m \times m$ matrix whose entries are variograms (on the diagonal) and cross-variograms. For the undersampled problem the $\bar{\gamma}$'s must be modified by inserting zeros in appropriate places.

McBratney and Webster (1983b) and Vauclin et al. (1983) did not use the matrix form, since only two variables were considered. The details of the matrix form are found in Myers (1982, 1983).

To apply co-kriging it is necessary to model variograms for each variable separately as well as cross-variograms for all pairs. Examples are

found in McBratney and Burgess (1983a, 1983b) and Vauclin et al. (1983), with theoretical results given in Myers (1982). A computer program utilizing an iterative method for solving the large system of equations as well as the undersampled option is given in Carr et al. (1983).

3–6.3 Conditional Simulation

Kriging is formulated in terms of random functions, although only one realization of the random function is available. Often another realization is sought which exhibits the spatial variability and known values at the sample locations.

By the use of a random number generator one can generate many values of a random variable with a desired distribution. However, for one value of each of many dependent random variables a different technique is necessary. Generation of the set of values, one each for a set of spatially dependent random variables, is called simulation. Conditional simulation produces a simulation such that at the sample locations the estimate coincides with the sample value.

Unlike some laboratory or field experiments, a new value cannot be obtained by repeating the experiment. If the same locations are sampled then the same results (except for instrument or analysis error) should be obtained. Sampling new locations provides additional data on the same realization. For another, realization simulation is necessary. Delhomme (1979) has used conditional simulation to study the Bathonian aquifer in France. In particular, simulated values were obtained for the log of the transmissivity.

The spatial dispersion in soils of nutrients, water, or pollutants could be studied by producing many simulations. Conditioning or matching the simulation to the data at sample locations is accomplished by using kriging. With any location x where a simulated value is to be made, then

$$Z(x) = Z^*(x) + [Z(x) - Z^*(x)]$$

where Z^* is the kriging estimator. Moreover, the $Z(x) - Z^*(x)$ has a mean of zero; in fact, at a sample location $Z(x) - Z^*(x) = 0$. The procedure then is to simulate values of $Z(x) - Z^*(x)$ (except at sample locations) and add to Z*(x).

To produce simulations in 2- or 3-space, the *turning bands method* was developed by Matheron (1973) and described in more detail in Journel and Huijbrets (1978). The simulated value at a point in 3-space is the sum of 15 simulations on equally spaced lines. These lines are determined by the edges on an iscosahedron, which provides the optimal polygonal approximation to a sphere.

The simulations on lines are produced by a *moving average* as described in Box and Jenkins (1970). It is necessary to relate covariances in 3-space to a covariance in 1-space and then represent this covariance

as a convolution. This decomposition is known for the standard covariance models used in geostatistics. Lenton and Rodriquez-Iturbe (1977) and Smith and Freeze (1979b) have given examples of other methods for simulation of rainfall and groundwater flow.

3–6.4 Miscellaneous Comments and Notes

Variograms—As noted above, valid models must s
definite condition. The use of an invalid model ca
estimated variance. The collection of valid m
nesting, that is, by using positive linear combi
The shape of the basic models can then be mo
nesting and hence match the sample variogram m
also a method for constructing an anisotropic model fr
It should also be noted that replacing γ by $a + \gamma$ or
the kriged values, but the kriging variance is changed

Unique vs. Moving Neighborhoods—The kriging es
spatial correlation in such a way that sample locations close
for which the estimate is desired receive greater weights and
away lesser weights. The sum of the weights is 1.0, but the
not constrained to be non-negative. Thus, negative weights ca
negative estimated values. There is no theoretical contradiction,
physical phenomena negative values are often not realistic. One so
is to use only those locations that are close; that is, for each locatio
be estimated, a different neighborhood is used. Another advantage
using moving neighborhoods is that the coefficient matrix for the krigi
equations is much smaller. One disadvantage when contouring is that
the moving neighborhoods lead to discontinuities in the plot.

Computing—One of the advantages of kriging is its simplicity of application. Computing sample variograms requires finding average squared distances only, and the kriged values require solving only a linear system. In addition to the complexities of data handling, inverting large matrices (i.e., solving large systems) can create problems. In general, the coefficient matrix is not positive definite when variograms are used and pivotal methods do not work well. There are several ways to avoid this problem. Frequently the variogram is replaced by the corresponding covariance. This avoids all the zeros on the diagonal. Alternatively the matrix can be partitioned; the upper left block containing the variogram values is positive definite. Finally, iterative methods such as the projection method are useful. A version of this method is embedded in the co-kriging program of Carr et al. (1983). One advantage of the projection method is that the universality conditions are checked at each iteration. The subroutine used to solve linear systems is a major component of any kriging program.

Screening, Declustering and Smoothing—Because spatial correlation is incorporated in kriging, several sample locations in close proximity constitute related information. The weights assigned by kriging compen-

sate for this clustering. With this in mind, it is seen that cluster sampling is not efficient, although there may be other factors such as screening: locations screened by other locations will receive lesser weights. Finally it should be noted that the kriged random function Z^* is smoother than the original function Z; that is, the variance of Z^* is smaller than the variance of Z.

3–7 DISCUSSION

Several points regarding geostatistics should be reiterated or stated more explicitly:

1. The applications are relatively new, especially to soil science.
2. The applications are quite general in terms of processes, scale of measurement, and type of application.
3. The results very often give the same answer as more conventional statistics.
4. There are many "gray" areas, partially because of (1) but also because of general uncertainty in the assumptions.

Most applications to this date emphasize mapping and contouring. These include efforts towards general surveying as well as a number of physical and chemical parameters. Although these are of obvious benefit, they need not be the final product. In fact, the background of geostatistics is heavily rooted in practical economics of mining; thus possibilities for use in operational analysis in soil science seem feasible.

The parameters addressed have been strongly slanted towards chemical and physical properties. There are no inherent reasons to eliminate any spatially variable property—e.g., yields, plant nutrients, or microbes. Likewise applications have been mostly for a few centimeters to kilometers, but any meaningful scale can be used.

In many cases, results are trivial. In Example 2, after a lengthy analysis, the unknown value was estimated by the average of its two neighbors—hardly surprising in retrospect, but pointing out that the results should be reasonable and must be dependent on available information and data. In general, as spacings between samples get larger, the potential advantages of regionalized variables are less; results approach those based on independence of samples and agree fully with other statistical approaches.

There are many "gray" areas which underscore basic uncertainties regarding natural systems. The form of the variogram influences the analysis; yet it cannot be known with certainty, and its best approximation cannot be given with confidence. Whether and to what degree stationarity exists is usually not really known. In the case of universal kriging, the relationship between estimators of the drift and bias in the results becomes clumsily entangled. The complications, however, are mostly a consequence of a more detailed analysis of the system—the absolute status, which cannot be totally established.

The challenge then is to use geostatistics as a tool where favorable and advantageous. Recognition of best-suited situations hopefully will come into focus with time, experience, and further development. How to best complement existing knowledge and techniques is a key ingredient. Examples are how to best interface with existing soil survey and descriptions, or how to best address contemporary problems such as raising or maintaining productivity, protecting the environment, or utilizing less energy and water—all under difficult economic and/or social constraints.

3-8 REFERENCES

Agterberg, F. P. 1974. Geomathematics. Elsevier, New York.

Al-Sanabani, M. 1982. Spatial variability of salinity and sodium adsorption ratio in a typic haplargid soil. M.S. Thesis, The Univ. of Arizona, Tucson.

Armstrong, M., and R. Jabin. 1981. Variogram models must be positive definite. Math. Geol. 13:455–459.

Bakr, A. A., L. W. Gelhar, A. L. Gutjahr, and H. R. MacMillan. 1978. Stochastic analysis of spatial variability of subsurface flow. I. Comparison of one and three dimensional flows. Water Resour. Res. 14:263–272.

Biafi, E. Y. 1982. Program user's manual for basis geostatistics systems. Dept. of Mining and Geol. Engr., Univ. of Arizona, Tucson.

Box, G. P., and G. M. Jenkins. 1970. Time series analysis forecasting and control. Holden-Day, San Francisco.

Bresler, E., and G. Dagan. 1983. Unsaturated flow in spatially variable fields. 2. Application of water flow models to various fields. Water Resour. Res. 19:421–428.

Burgess, T. M., and R. Webster. 1980a. Optimal interpolation and isarithmic mapping of soil properties I. The variogram and punctual kriging. J. Soil Sci. 31:315–331.

Burgess, T. M., and R. Webster. 1980b. Optimal interpolation and isarithmic mapping of soil properties II. Block kriging. J. Soil Sci. 31:333–341.

Burgess, T. M., R. Webster, and A. B. McBratney. 1981. Optimal interpolation and isarithmic mapping of soil properties VI. Sampling strategy. J. Soil Sci. 31:643–659.

Campbell, J. B. 1978. Spatial variability of sand content and pH within continuous delineations of two mapping units. Soil Sci. Soc. Am. J. 42:460–464.

Carr, J., D. Myers, and C. Glass. 1983. Co-kriging: A computer program. Dept. of Mathematics, Univ. of Arizona, Tucson.

Clark, I. 1979. Practical geostatistics. Applied Sci. Pub., Ltd., London.

Clifton, P.M., and S.P. Neuman. 1982. Effects of kriging and inverse modeling on conditional simulation of the Avra Valley in southern Arizona. Water Resour. Res. 18:1215–1234.

Dagan, G., and E. Bresler. 1983. Unsaturated flow in spatially variable fields. 1. Derivation of models of infiltration and redistribution. Water Resour. Res. 19:413–420.

David, M. 1977. Geostatistical ore reserve estimation. Elsevier, New York.

Davis, J. C. 1973. Statistics and data analysis in geology. John Wiley and Sons, Inc., New York.

Delfiner, P. 1976. Linear estimation of non-stationary phenomena, *In* M. Guarascio, M. David, and Ch. Huijbregts (eds.) Advanced geostatistics in the mining industry. Reidel. NATO Symposium. pp. 49–68.

Delhomme, J. P. 1979. Spatial variability and uncertainty in groundwater flow parameters: a geostatistical approach. Water Resour. Res. 15:269–280.

DuBrule, O. 1983. Two methods with different objectives: splines and kriging. Math. Geol. 15:245–258.

Folorunso, O. A., and D. E. Rolston. 1984. Spatial variability of field measured denitrification gas fluxes. Soil Sci. Soc. Am. J. 48:1214–1219.

Gajem, Y. M. 1980. Spatial structure of physical properties of a typic Torrifluvent. M.S. Thesis, Univ. of Arizona, Tucson.

Gajem, Y. M., A. W. Warrick, and D. E. Myers. 1981. Spatial dependence of physical properties of a typic Torrifluvent soil. Soil Sci. Soc. Am. J. 46:709–715.

Hajrasuliha, S., N. Baniabbassi, J. Matthey, and D. R. Nielsen. 1980. Spatial variability of soil sampling for salinity studies in southwest Iran. Irrig. Sci. 1:197–208.

Hassan, H. M., A. W. Warrick, and A. Amoozegar-Fard. 1983. Sampling volume effects on determining salt in a soil profile. Soil Sci. Soc. Am. J. 47: 1265–1267.

Hawley, M. E., H. M. Richard, and J. J. Thomas. 1982. Volume-accuracy relationship in soil moisture sampling. J. Irrig. and Drain. Proc., ASCE 108:1–11.

Journel, A. G., and Ch. Huijbregts. 1978. Mining geostatistics. Academic Press, London.

Kachanoski, R. G., D. E. Rolston, and E. deJong. 1985. Spatial and spectral relationships of soil properties and microtopagraphy: I. Density and thickness of A horizon. Soil Sci. Soc. Am. J. 49:804–812.

Lenton, R. L. and I. Rodriquez-Iturbe. 1977. A multidimensional model for the synthesis of processes of areal rainfall averages. Water Resour. Res. 13:605–612.

Liss, H. F. 1983. Spatial and temporal variability of water soluble organic carbon for a cropped soil. Ph.D. Dissertation. Univ. of California, Davis.

Lumley, J. L., and A. Panofsky. 1964. The structure of atmospheric disturbance. John Wiley and Sons, Inc., New York.

Luxmoore, R. J., B. P. Spalding, and I. M. Munro. 1981. Areal variation and chemical modification of weathered shale infiltration characteristics. Soil Sci. Soc. Am. J. 45:687–691.

Matheron, G. 1973. The Intrinsic random functions and their applications. Adv. Appl. Prob. 5:239–465.

McBratney, A. B., and R. Webster. 1981. The design of optimal sampling schemes for local estimating and mapping of regionalized variables. II. Program and examples. Computers Geosci. 7:335–365.

McBratney, A. B., and R. Webster. 1983a. How many observations are needed for regional estimation of soil properties? Soil Sci. 135:177–183.

McBratney, A. B., and R. Webster. 1983b. Optimal interpolation and isorithmic mapping of soil properties. V. Co-regionalization and multiple sampling strategies. J. Soil Sci. 34:137–162.

McBratney, A. B., R. Webster, and T. M. Burgess. 1981. The design of optimal sampling schemes for local estimating and mapping of regionalized variables. I. Theory and method. Computers Geosci. 7:331–334.

Myers, D. E. 1982. Matrix formulation of co-kriging. Math. Geol. 14:248–257.

Myers, D. E. 1983. Estimation of linear combinations and co-kriging. Math. Geol. 15:633–637.

Myers, D. E. 1984. Co-kriging: new developments. p. 295–305. *In* G. Verly, M. David, A. Journel, and A. Marechal (ed.) Geostatistics for natural resource characterization. D. Reidel Publ. Co., Dordrecker, The Netherlands.

Rendu, J. M. 1978. An introduction to geostatistical methods of mineral evaluation. S. Afr. Inst. Mining and Metallurgy, Johannesburg.

Russo, D., and E. Bresler. 1981a. Effect of field variability in soil hydraulic properties on solutions of unsaturated water and salt flows. Soil Sci. Soc. Am. J. 45:675–681.

Russo, D., and E. Bresler. 1981b. Soil hydraulic properties as stochastic processes: I. An analysis of field spatial variability. Soil Sci. Soc. Am. J. 45:682–687.

Russo, D., and E. Bresler. 1982. Soil hydraulic properties as stochastic processes: II. Errors of estimates in a heterogeneous field. Soil Sci. Soc. Am. J. 46:20–26.

Sharma, M. L., G. A. Grander, and C. G. Hunt. 1980. Spatial variability of infiltration in a watershed. J. Hydrol. 45:101–122.

Sisson, J. B., and P. J. Wierenga. 1981. Spatial variability of steady-state infiltration rates as a stochastic process. Soil Sci. Soc. Am. J. 45:699–704.

Smith, L. 1981. Spatial variability flow parameters in a stratified sand. Math. Geol. 13:1–21.

Smith, L., and R. A. Freeze. 1979a. Stochastic analysis of steady state groundwater flow in a bounded domain, I. One-dimensional simulations. Water Resour. Res. 15:521–528.

Smith, L., and R. A. Freeze. 1979b. Stochastic analysis of steady state groundwater flow in a bounded domain, II. Two-dimensional simulations. Water Resour. Res. 15:1543–1559.

van Kuilenburg, J., J. J. DeGruijter, B. A. Marsma, and J. Bouma. 1982. Accuracy of spatial interpretation between point data on soil moisture supply capacity compared with estimations from mapping units. Geoderma 27:311–325.

Vauclin, M., S. R. Vieira, R. Bernard, and J. L. Hatfield. 1982. Spatial variability of surface temperature along two transects of a bare soil. Water Resour. Res. 18:1677–1686.

Vauclin, M., S. R. Vieira, G. Vachaud, and D. R. Nielsen. 1983. The use of co-kriging with limited field soil observations. Soil Sci. Soc. Am. J. 47:175–184.

Vieira, S. R., D. R. Nielsen, and J. W. Biggar. 1981. Spatial variability of field-measured infiltration rate. Soil Sci. Soc. Am. J. 45:1040–1048.

Webster, R. 1977. Quantitative and numerical methods in soil classification and survey. Oxford Univ. Press.

Webster, R., and T. M. Burgess. 1980. Optimal interpolation and isarithmic mapping of soil properties. III. Changing drift and universal kriging. J. Soil Sci. 31:505–524.

Wilding, L. P., and L. R. Drees. 1978. Spatial variability: a pedologist's viewpoint. pp. 1–12. *In* Diversity of Soils in the Tropics. Special Pub. 34. Am. Soc. of Agronomy, Soil Sci. Soc. Am., 677 S. Segoe Rd., Madison, WI 53711.

Wollum, A. G., II, and D. K. Cassel. 1984. Spatial variability of *Rhizobium japonicum* in two North Carolina soils. Soil Sci. Soc. Am. J. 48:1082–1086.

Yost, R. S., G. Uehara, and R. L. Fox. 1982. Geostatistical analysis of soil chemical properties of large land areas. I. Variograms. Soil Sci. Soc. Am. J. 46:1028–1032.