WATER RESOURCES RESEARCH, VOL. 25, NO. 12, PAGES 2501–2510, DECEMBER 1989

Hydrostratigraphic Interpretation Using Indicator Geostatistics

NICHOLAS M. JOHNSON AND SHIRLEY J. DREISS

Earth Sciences, University of California, Santa Cruz

In this paper we describe an approach that uses indicator geostatistics to interpret qualitative borehole logs and compute experimental variograms for complex alluvial sediments. Borehole descriptions are first transformed into binary indicator values based on inferred relative permeability from the borehole descriptions. The resulting indicator data can then be used to compute variograms and construct three-dimensional variogram models. The ranges of computed indicator variograms for a groundwater contamination site in Santa Clara Valley, California, are very sensitive to the orientation of the search plane. These variograms are consistent with known stratigraphic features and describe the spatial structure of deposits from different depositional environments. Indicator kriging weighs all the available data on the basis of a three-dimensional, anisotropic variogram model and provides an estimate of uncertainty in the hydrostratigraphic correlation. Kriged indicator values represent probabilities that sediments at a specific location fall into one of two indicator categories. The location of the 0.5 indicator contour is approximately the boundary between high- and low-permeability sediments that might be constructed in a geologic cross section.

INTRODUCTION

A large number of groundwater contamination sites are located in alluvial sedimentary environments. The sediments at these sites are often geologically mapped in terms of facies assemblages and relative ages of deposition. Unless the sediments are exceptionally uniform, the mapped geologic units typically contain lenses or layers of contrasting grain size and permeability. These subunits can range in maximum thickness from less than a meter to tens of meters and may be geometrically anisotropic and discontinuous.

An accurate description of the location and geometry of subunits with contrasting permeabilities is important for defining flow field boundaries and preferential pathways of water and solute movement. In practice, these boundaries are usually inferred from borehole logs using geologic judgement. This approach tends to underutilize the detail and spatial correlation of available information and provides no means for representing uncertainties in the stratigraphic interpretation.

In this paper we illustrate the use of an indicator function approach for interpreting complex alluvial stratigraphy from qualitative borehole logs. At extensively studied sites the positions of contaminant plumes are usually determined from measurements at numerous boreholes. The logs of these boreholes provide a detailed, descriptive data base of the sediment lithology. Here we interpret borehole descriptions in terms of binary indicator values and use these values to compute experimental variograms. The resulting indicator variograms describe the spatial structure of aquifer-aquitard stratigraphy. Indicator kriging gives an estimate of the uncertainty in the hydrostratigraphic interpolations.

BACKGROUND AND RELATED RESEARCH

Numerous investigators have commented on the importance of incorporating information about aquifer heterogeneity and the spatial structure of geologic deposits into studies of fluid flow and contaminant transport [e.g., *Matheron and de Marsily*, 1980; *Smith and Schwartz*, 1980; *Gelhar and Axness*, 1983; *Fogg*, 1986*a*; *Guven et al.*, 1986]. Most investigations of aquifer heterogeneity describe the spatial variability of transmissivity, for which data are relatively abundant [e.g., *Delhomme*, 1979; *Clifton and Neuman*, 1982; *Aboufirassi and Marino*, 1984; *Hoeksema and Kitanidis*, 1984, 1985]. However, transmissivity is of limited use in studies of most contaminant plumes because the processes controlling transport are fundamentally three-dimensional [*Mackay et al.*, 1986].

The spatial variability of hydraulic conductivity has received considerable attention in studies of surficial soils in agricultural settings [e.g., *Vieira et al.*, 1981]. Recently, several investigations have examined the spatial variability of hydraulic conductivity within relatively uniform geologic deposits [e.g., *Smith*, 1981; *Sudicky*, 1986]. Because effective permeability and transmissivity are spatially averaged parameters, the scale of measurement is important when describing the spatial variability of these parameters [*Dagan*, 1986; *Gelhar*, 1986].

Approaches for interpreting stratigraphic correlation and interconnectedness have been considered for many years in the geologic literature. *Testerman* [1962] developed a method for estimating the correlation of high permeability zones in a petroleum reservoir by statistically comparing the permeabilities of two vertical zones in adjacent wells. Subsequent studies have used similar approaches to correlate strata using geologic parameters, such as electrical resistivity, fossil content, or lithology [e.g., *Kemp*, 1982; *Hawkins*, 1984]. *Schwarzacher* [1980, 1982] and *Royle and Hosgit* [1974] used geostatistical approaches to correlate bedding thickness in a sedimentary environment model and for an economic assessment of a gravel deposit. Recently, *Fogg* [1986*b*] simulated sand body interconnectedness in a hydrologic context using a geostatistical model of the spatial distribution of channel sand thickness.

Indicator geostatistics, as described by *Journel* [1983], have been applied to the evaluation of mineral deposits [e.g., *Davis*, 1984], earthquake intensities [*Carr et al.*, 1985], rock joint orientations [*Young*, 1987], and the distribution of pollutants [*Isaaks*, 1984]. These studies demonstrated that indicator variograms are relatively smooth and structurally

Paper number 89WR01380.
0043-1397/89/89WR-01380$05.00

informative. In a recent study, *Desbarats* [1987] used indicator geostatistics to represent relatively high- and low-permeability regions in a sandstone cross section with small-scale shale features. He then estimated an effective permeability for the sandstone by simulating flow in the resulting dual-permeability field. Our objective in this paper is to describe the boundaries of hydrologically significant stratigraphic features that are larger than those considered by Desbarats but smaller than alluvial sedimentary units traditionally mapped by geologists.

INDICATOR FUNCTION APPROACH

The basic premise in geostatistics is that the spatial structure of a variable can be described with a variogram, defined as the expected squared difference between pairs of data values [*Journel and Huijbregts*, 1978]. Assuming second-order stationarity, the variogram is considered to be independent of data locations and a function of only the direction and separation distance between points. An experimental semivariogram, $\gamma^*(h)$, can be computed from discrete observations $z(x_i)$ of the random variable $Z(x)$ as

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

where $N(h)$ is the number of data pairs separated approximately by the same distance h.

Indicator geostatistics use a transform of the data to estimate probabilities that a variable will be less than or equal to specified threshold values [*Journel*, 1983]. A binary indicator function for cutoff z_c is defined as

$$I(x; z_c) = 1 \quad \text{if } Z(x) \le z_c$$
$$I(x; z_c) = 0 \quad \text{if } Z(x) > z_c \quad (2)$$

The random function, $I(x; z_c)$, has a mean and variance given by

$$E[I(x; z_c)] = 1 \text{ Prob } [Z(x) \le z_c] + 0 \text{ Prob } [Z(x) > z_c]$$
$$= \text{Prob } [Z(x) \le z_c] = m(z_c) \quad (3)$$

$$\text{Var } [I(x, z_c)] = m(z_c)[1 - m(z_c)] \quad (4)$$

Experimental semivariograms of the indicator function

$$\gamma^*(h; z_c) = \frac{1}{2N(h)} \sum_{i=1}^{N(h)} [I(x_i + h; z_c) - I(x_i; z_c)]^2 \quad (5)$$

can reveal different spatial correlations of $I(x; z_c)$ for different threshold values z_c. Binary indicator kriging computes values within [0, 1], which are estimates of the probability that actual variable values are less than or equal to the cutoff z_c, given the data and variogram model used [*Journel*, 1983; *Journel and Alabert*, 1988].

In this study we wish to estimate the boundaries of relatively high- and low-permeability zones using descriptive information from borehole logs. Geotechnical investigators typically label intervals of borehole logs with group symbols of the Unified Soil Classification System (USCS) [cf. *Sowers*, 1979]. We infer a qualitative apparent permeability from these categories and assign indicator values of either 1 or 0 to intervals containing apparently high- or low-permeability materials (Table 1).

TABLE 1. Indicator Values for Alternative Interpretations of the United Soil Classification System

High Permeability ($I(x; z_c) = 1$)	Low Permeability ($I(x; z_c) = 0$)
Interpretation (a)	
GW SW GM SM	ML CL OL
GP SP GC SC	MH CH OH
Interpretation (b)	
GW SW	GM SM ML CL OL
GP SP	GC SC MH CH OH

G, gravel; S, sand; M, silt; C, clay; O, organics; W, well graded (i.e., poorly sorted); P, poorly graded; L/H, low/high plasticity.

The inference of relatively high and low permeability from the UCSC classification is subjective, as illustrated by the alternative interpretations shown in Table 1. For example, interpretation (a) distinguishes sediments composed of predominantly fine-grained materials from all other sediments. This criterion identifies zones of apparently low permeability and emphasizes the degree of hydraulic separation between high permeability layers. An alternative evaluation, interpretation (b), distinguishes clean sands and gravels from all other materials. This interpretation highlights the spatial structure of zones of high apparent permeability. The difference in the variograms for the two criteria depends on the amount of and spatial distribution of sediments with mixed compositions (GM, GC, SM, and SC) in the profiles.

A schematic example of how the structure of geologic deposits might be reflected by indicator variograms is illustrated in Figures 1 and 2. The block diagrams in Figure 1 show a hypothetical deposit of alluvial sediments and an interpretation using an indicator function representation. The deposit is an idealized layered system in which coarse-grained layers occur as lenses in a predominantly fine-grained matrix. The layers dip slightly in the x direction and are anisotropic because coarse-grained lenses are continuous and more elongated in the x direction than in the y direction.

Indicator semivariograms for the block are shown schematically in Figure 2. In this example, directional variograms in x and y represent spatial correlations along the strike and dip of sediment layers and the z direction variogram represents vertical spatial correlations. All three variograms have a sill equal to the indicator variance of the block, provided that the block is large relative to the correlation ranges. The anisotropic structure of the block is demonstrated by differences in the variogram ranges. Because the lenses are interconnected and more elongate in the dip direction, the x variogram approaches the sill at a relatively large distance. The variogram in the y direction approaches the sill at a shorter distance, reflecting the shorter elongation of the lenses along the strike. The z direction variogram crosses the sill at a still shorter distance.

Because of the periodicity in the illustration, all three variograms exhibit a dampened sinusoidal shape, or hole effect. The first maxima of the x and y variograms occur when h is about half the average lens length and width, respectively. Similarly, the oscillations in the x and y variograms have wavelengths, λ_x and λ_y, approximately equal to the length and width of the lenses. In the z variogram the first

maximum occurs when h approximately equals the average layer thickness. The wavelength of the oscillations, λ_z, is about twice the layer thickness if the average thickness of the coarse- and fine-grained layers are the same. In each variogram the amplitude of the hole effect diminishes with distance if the periodicity of the locations of the units is imperfect.

In actual layered sedimentary deposits, hole effects are most likely to occur in vertical variograms because repeated layers of similar thickness are more common than lenses with regular horizontal geometries. In any case, apparent nested structures, anisotropies, or trends in an empirical variogram should be consistent with the geologic setting.

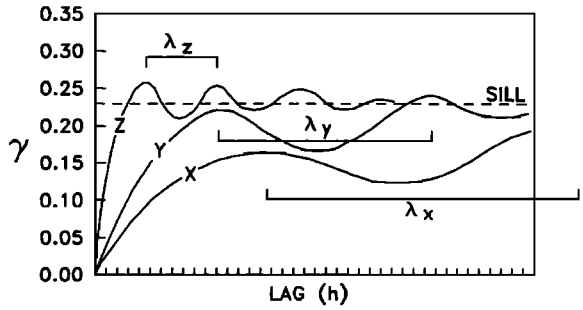

Fig. 2. Schematic indicator variograms for block in Figure 1b, with λ_x, λ_y, and λ_z corresponding to the dimensions illustrated.

CASE STUDY

As a demonstration, we consider a site in Santa Clara Valley, California, which contains a clearly defined contaminant plume and has been the subject of extensive field investigations (Figure 3). The site is in a narrow alluvial valley underlain by as much as 120 m of unconsolidated gravel, sand, silt, and clay. The surficial sediments have been mapped as two geologic units, younger (Q_y) and older (Q_o) alluvial fan deposits [*Helley and Brabb*, 1971]. These and underlying units contain discontinuous layers of sands and gravels interbedded with clays and silts.

During site studies to locate the contaminant plume, investigators drilled more than 150 boreholes within a 1.2 km^2 area. The borehole logs, developed from cuttings, cores, and geophysical data, reveal complex stratigraphy in the underlying sediments and three principal coarse-grained layers within 65 m of the ground surface (Figures 4 and 5). The layers dip approximately 0.1°–0.3° to the west-northwest in a direction similar to the orientation of the axis of the contaminant plume. Coarse-grained layers tend to be continuous in the dip direction and discontinuous along the strike.

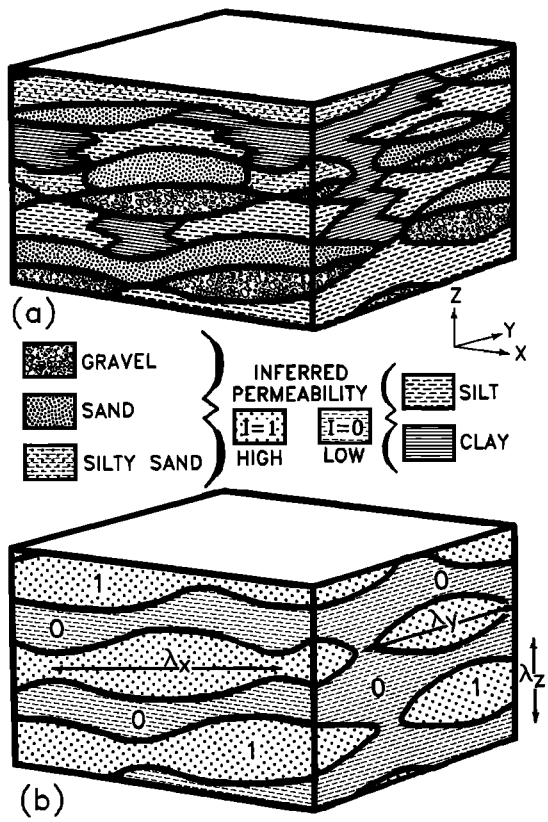

Fig. 1. Schematic block diagrams of (a) complexly layered alluvial sediments and (b) the structure of these sediments when interpreted with an indicator function. The length and width of a lens are given by λ_x and λ_y. The vertical distance between midpoints in low-permeability layers is λ_z.

We discretized logs of the boreholes shown in Figure 3 into 0.61 m (2 ft) intervals, to a maximum depth of 65 m. In doing so we obtained 7336 data locations, which allowed us to use an exceptionally large number of data pairs for most variogram calculations. We then assigned an indicator value to each location using the criterion in Table 1 (interpretation a). Defined as such, the data have a mean indicator value of 0.541 and a variance of 0.248 (Table 2). Because sediments of mixed grain size account for less than 10% of the borehole logs, use of the indicator function definition in Table 1 (interpretation b) rather than Table 1 (interpretation a) does not significantly alter the computed variograms. We use the criterion in Table 1 (interpretation a) throughout this example.

The three-dimensional structure of the deposits can be described with two variograms for near-horizontal, orthogonal directions and a third vertical variogram. We computed near-horizontal variograms with searches in dipping planes as illustrated in Figure 6. The orientations of the search planes can be described by conventional geologic strike and dip. Planes contain four search regions. Each region is defined by a search direction, a horizontal tolerance angle of ±22.5° centered about the search direction, and a vertical tolerance distance of ±0.33 m perpendicular to the plane. Separation distances or lags h between data pairs are multiples of selected unit lags l. Variograms were computed from (5) and adjusted by

$$\gamma_I'(h; z_c) = \gamma_I(h; z_c) \frac{\text{Var}\,[I(x_i; z_c),\ i = 1, \cdots, n]}{\text{Var}\,[I(x_i; z_c),\ i = 1, \cdots, N(h)]} \qquad (6)$$

where n is the total number of data. Equation 6 normalizes for differences between the overall indicator data mean and the mean of indicator data values at specific lags. This adjusts for a proportional effect as described by *Journel and Huijbregts* [1978]. Vertical variograms were found by pairing data down boreholes.

As a practical rule, experimental variograms are reliable only for lags equal to or less than half the areal extent of the data L [*Journel and Huijbregts*, 1978]. However, we chose to retain the variogram beyond $L/2$ when the number of data pairs was large and the behavior of this portion of the variogram was smooth and consistent with the part for $h < L/2$. In the plots of experimental variograms we distinguish

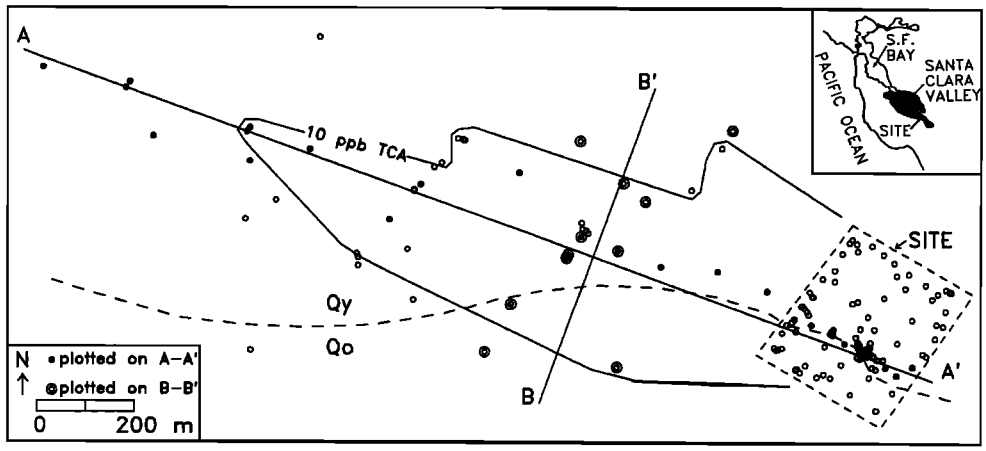

Fig. 3. Location of boreholes and contaminant plume, April and May 1982 (California Regional Water Quality Control Board, written communication, 1983).

between semivariances computed for $h < L/2$ and $h > L/2$ with closed and open symbols, respectively.

Experimental Variograms

We initially chose a search plane striking N20°E and dipping 0.3°WNW, based on the apparent orientation of the stratigraphic layers described in geologic reports. This dip direction approximately coincides with the orientation of contaminant plume migration and the transect A-A' (Figures 3 and 4).

Figure 7*a* illustrates variograms computed for the near-horizontal plane using a unit lag of 80 m. The variograms for the strike and the dip direction of the search exhibit an apparent anisotropy at lags greater than 200 m. The strike variogram crosses a sill equal to the data variance at a lag of approximately 250 m, while the variogram in the dip direction does not approach the sill until 1800 m or more, possibly beyond the areal extent of the data. Both the strike and the dip direction variograms exhibit a large nugget effect of approximately 0.15, about 60% of the variance.

The cluster of boreholes near the origin of the plume could bias the variogram. These closely spaced, on-site boreholes are those within the site boundary in Figure 3. Because many of the on-site boreholes were intentionally placed near a locally sinuous geologic contact to delineate its geometry, we suspected that data pairs that include points from these boreholes might not be representative of the spatial structure of deposits elsewhere in the study area. In particular, we expected that on-site data pairs would exhibit excessively high variability at short lags. On the other hand, nearly all long-lag data pairs would contain an on-site point, which might introduce artificial correlation. For these reasons we elected to delete the on-site data in computing near-horizontal variograms representative of the study site as a whole (Figure 8). In doing so we neglected information about the on-site sediments.

Figure 8 illustrates $\gamma'_I(h)$ computed with off-site data along the strike and dip using unit lags of 80 and 20 m. The variograms for the strike and dip again exhibit an apparent anisotropy at long lags (Figure 8*a*). The variogram along strike crosses a sill equal to the overall data variance, 0.248, at about $h = 400$ m, while the variogram in the dip direction approaches the sill at approximately 1400 m. Thus the large-scale hydrostratigraphic features appear to have greater spatial correlation along dip than along strike. The variograms along strike and dip are similar at lags less than 200 m, suggesting that the structure of the deposits is nearly isotropic at this smaller scale (Figure 8*b*). An apparent

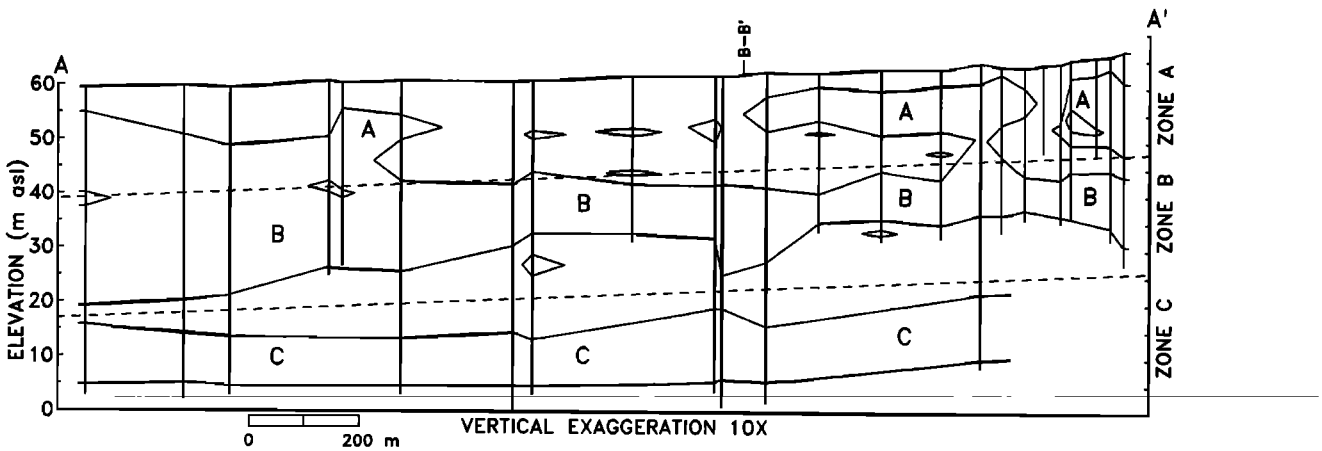

Fig. 4. Geologic cross section along A-A' (Figure 3). Letters A, B, and C designate high-permeability layers inferred from criteria in Table 1 (interpretation (a)).

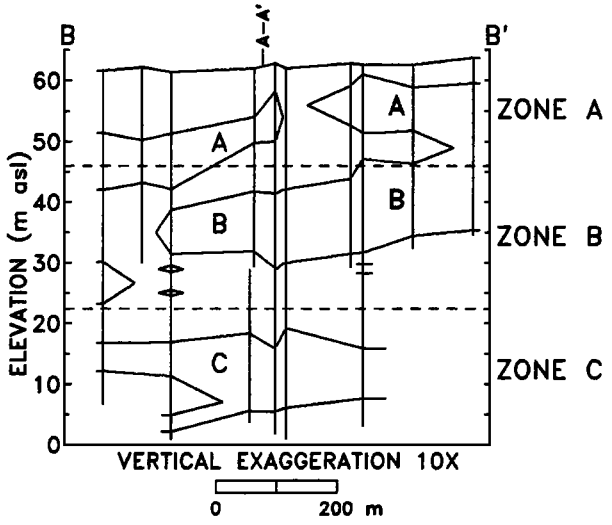

Fig. 5. Geologic cross section along B-B' (Figure 3). Letters A, B, and C designate high-permeability layers inferred from criteria in Table 1 (interpretation (a)).

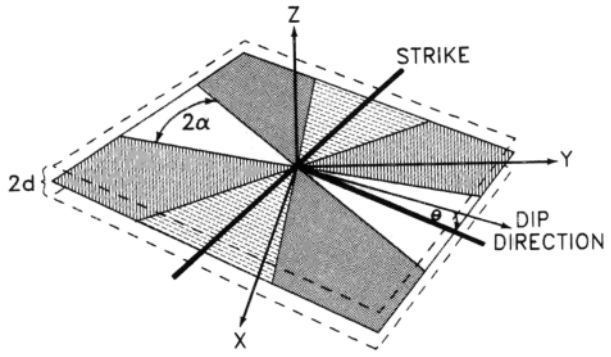

Fig. 6. Search plane for calculating experimental variograms. Shaded areas represent four possible search regions; θ, dip angle; α, angle tolerance; d, vertical tolerance.

nugget effect is present in both directions and is equivalent to about 25% of the overall data variance.

Figure 7*b* shows the vertical experimental variogram computed with a unit lag $l = 2.0$ m. This variogram has virtually no nugget effect. The hole effect reflects the vertical alternation of coarse- and fine-grained layers. The distance to the first maximum, 10 m, is approximately equal to the thickness of layers that can be recognized by inspection in the borehole logs.

As a means for comparing variograms and for later use in kriging, we fit the experimental variograms in Figures 7*b*, 8*a*, and 8*b* with a three-dimensional variogram model consisting of three nested, anisotropic spherical models:

$$\gamma(h_x, h_y, h_z) = \sum_{i=1}^{3} C_i Sph_{a_i}(h_i) \qquad (7)$$

where

$$Sph_{a_i}(h_i) = \frac{3}{2}\frac{h_i}{a_i} - \frac{1}{2}\left(\frac{h_i}{a_i}\right)^3 \qquad h_i < a_i$$
$$Sph_{a_i}(h_i) = C_i \qquad h_i \geq a_i \qquad (8)$$

$$h_i = \left[(h_x)^2 + \left(h_y \frac{a_{x_i}}{a_{y_i}}\right)^2 + \left(h_z \frac{a_{x_i}}{a_{z_i}}\right)^2\right]^{1/2} \qquad (9)$$

The C_i *and* a_i are the sills and ranges of the spherical models and h_x, h_y, and h_z are separation distances in the x, y, and z directions. In this case, x, y, and z are coincident with dip, strike, and near-vertical directions, respectively. We used a spherical model having the range of the vertical variogram in place of a nugget effect because of the high continuity displayed near the origin of the vertical variogram. Values of C_i and a_i, fit by inspection, are summarized in Table 3 (zone ABC).

Orientations of Maximum and Minimum Continuity

In computing and modeling the variograms in Figures 7–8 we assumed a stratigraphic dip and directions of maximum

TABLE 2. Summary of Indicator Data Sets and Search Plane Orientations Used to Compute Experimental Variograms

Zone*	Number of Points	Indicator Mean	Indicator Variance	Search Plane Orientation: Dip Direction	Strike	Dip Angle
			On and Off Site†			
ABC	7336	0.541	0.248	N70°W	N20°E	0.25°
			On Site†			
ABC	4008	0.540	0.248	N70°W	N20°E	0.20°
			Off Site†			
ABC	3328	0.543	0.248	N70°W	N20°E	0.20°
A	1355	0.374	0.234	N40°W	N50°E	0.20°
B	1307	0.730	0.197	N85°W	N5°E	0.25°
			On and Off Site†			
C	725	0.553	0.247	WEST	NORTH	0.15°

*See Figures 4 and 5 for zone locations.
†See Figure 3 for locations of on-site and off-site data.

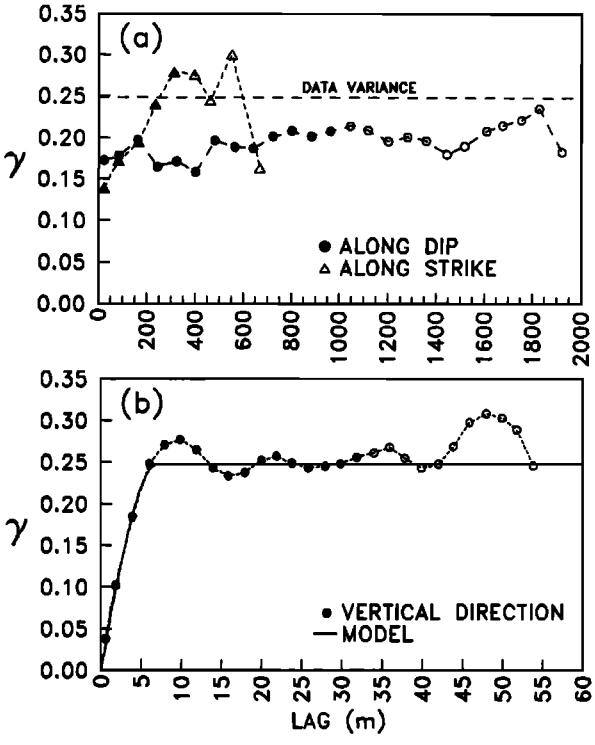

Fig. 7. Indicator variograms (*a*) along strike and dip and (*b*) in the vertical direction. Variograms were calculated using all available on-site and off-site data and unit lags of 80 m along strike and dip and 2 m in the vertical direction. Orientation of the search plane used in Figure 7*a* is given in Table 2. Parameters for the model in Figure 7*b* are given in Table 3.

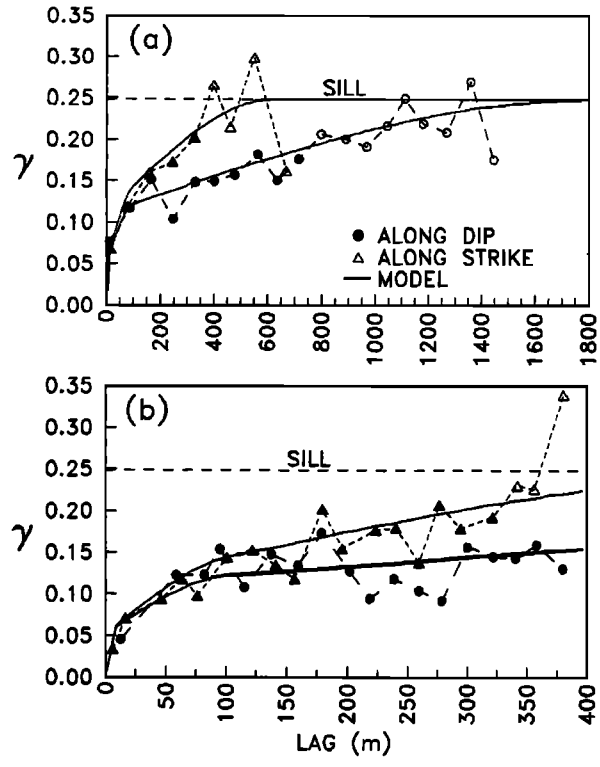

Fig. 8. Indicator variograms along strike and dip calculated using off-site data. A unit lag was used of (a) 80 m and (b) 20 m. See Table 2 for orientation of search plane and Table 3 for model parameters.

and minimum continuity using prior geologic information. We now examine the sensitivity of the variogram ranges to variations in the strike and dip of the search planes and describe the spatial structure of the sediment layers in more detail.

The variogram range should reach a maximum when the search direction is aligned with the direction of maximum continuity of the deposits and when the dip of the search plane is the same as the dip of the sediment layers. Figure 9 is a plot of the ranges of variograms computed along the strike and dip of various search planes dipping 0.2°. Ranges were determined by fitting each experimental variogram with nested sums of nugget and spherical models:

$$\gamma(h) = C_o + CSph_a(h) \tag{10}$$

where C_o is a nugget effect. We use a nugget effect instead of the two smaller spherical models in (7) because we are only interested in the range of the largest spherical model. The ranges of variograms computed along the dip direction of the search plane exhibit a broad maximum when the dip direction is between N30°W and west. Similarly, the ranges of variograms along strike are least for these orientations.

To examine the structure of the deposits in more detail, we explored the possibility that directions of minimum and maximum continuity may vary for the different stratigraphic layers. The data were subdivided into three depth zones A, B, and C, each centered around a coarse-grained layer (Figures 4 and 5). Both on-site and off-site data were used for zone C because of the sparsity of data at this depth. Figure 10 illustrates the sensitivity of the ranges of variogram models to various search plane orientations, each dipping 0.2°. The direction of maximum continuity is approximately due west for zones B and C and N40°W for zone A (Table 2). The directions of minimum continuity correspond generally to the strike of these same search planes. In addition, zone C exhibits greater spatial continuity than zones A or B.

Figure 11 shows the sensitivity of the variogram ranges for each zone to the dip angle of the search plane. The ranges are highly sensitive to small changes in dip. For all three

TABLE 3. Summary of Variogram Model Parameters

	Sph_{a_1}		Sph_{a_2}		Sph_{a_3}	
Zone	C_1^*	a_1, m	C_2^*	a_2, m	C_3^*	a_3, m
			Along Dip			
ABC	0.050 (20%)	7	0.060 (24%)	100	0.138 (56%)	1800
A	0.120 (51%)	7	···	···	0.114 (49%)	1800
B	0.080 (41%)	7	0.040 (20%)	300	0.077 (39%)	1800
C	0.050 (20%)	7	···	···	0.197 (80%)	3000
			Along Strike			
ABC	0.050	7	0.060	100	0.138	620
A	0.120	7	···	···	0.114	180
B	0.080	7	0.040	200	0.077	200
C	0.050	7	···	···	0.197	550
			Vertical			
ABC	0.050	7	0.060	7	0.138	7
A	0.120	7	···	···	0.114	7
B	0.080	7	0.040	7	0.077	7
C	0.050	7	···	···	0.197	7

Spaces without values indicate no intermediate range model.
*Percentage of total sill given in parentheses.

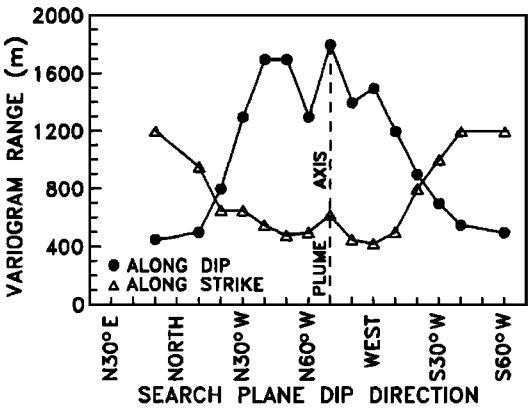

Fig. 9. Sensitivity of off-site variogram range along strike and dip to the orientation of the search plane.

zones the variogram range is greatest when the search plane dips between 0.15° and 0.25° (Table 2). Varying the dip by ±0.5° sharply reduces the variogram range. We selected a specific dip angle for each zone on the basis of minimum nugget effect and overall fit to the model. Using these criteria, zone C appears to be slightly more flat lying than zones A and B.

Figure 12 shows the variograms computed for zones A, B, and C using the search plane orientations found to maximize the variogram range in the dip direction and minimize the nugget effect. Tables 2 and 3 summarize the orientations, ranges, and sills of the fitted models. The ratio of ranges in the dip direction to ranges along strike is about 10:1 for zones A and B, 5:1 for zone C, and 3:1 for all zones combined. The relatively low range ratio for all zones combined results from the different hydrostratigraphic orientations of the individual zones.

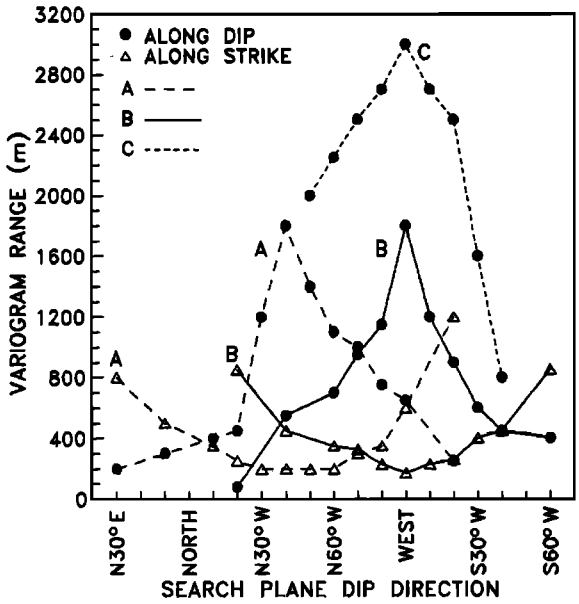

Fig. 10. Sensitivity of variogram range along strike and dip to the orientation of the search plane for individual zones A, B, and C. Curve along strike for zone C was not computed because of lack of data.

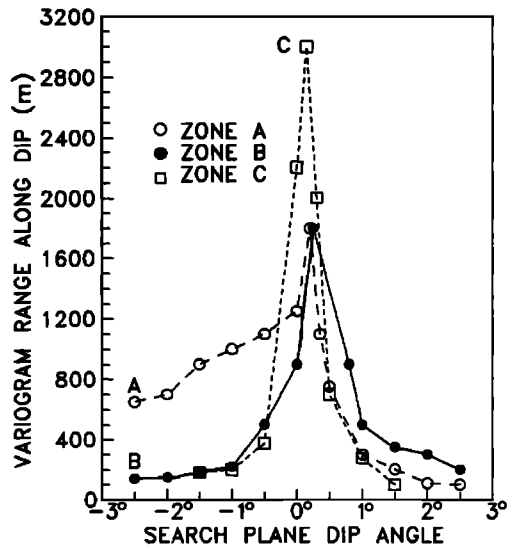

Fig. 11. Sensitivity of variogram range along dip to the search plane dip angle for zones, A, B, and C.

Discussion

The indicator variograms reflect observed stratigraphic features of the groundwater contamination site. The stratigraphic layering, dip, and anisotropy evident in cross-

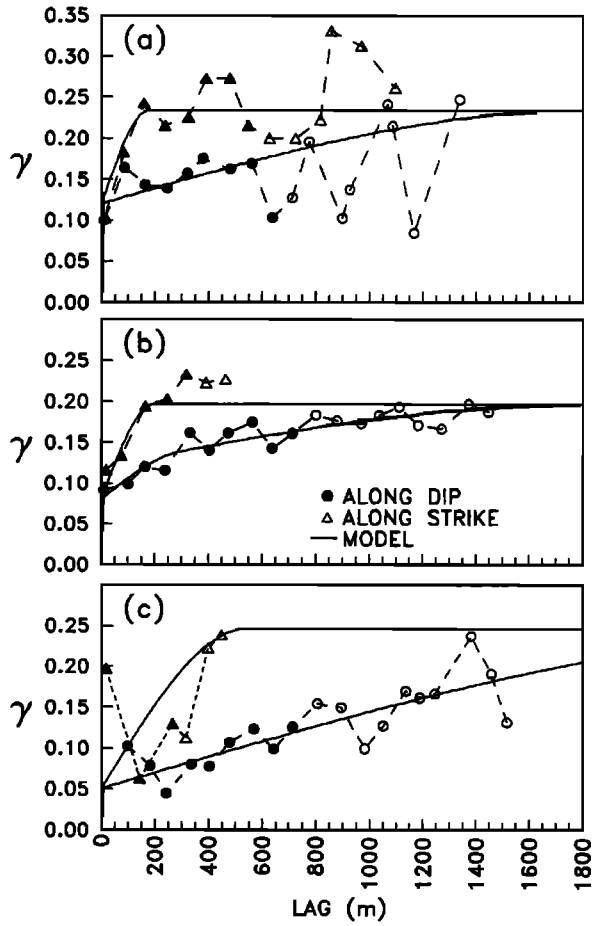

Fig. 12. Indicator variograms along strike and dip calculated for zones (*a*) A, (*b*) B, and (*c*) C. Unit lag = 80 m. See Table 2 for search plane orientations and Table 3 for model parameters.

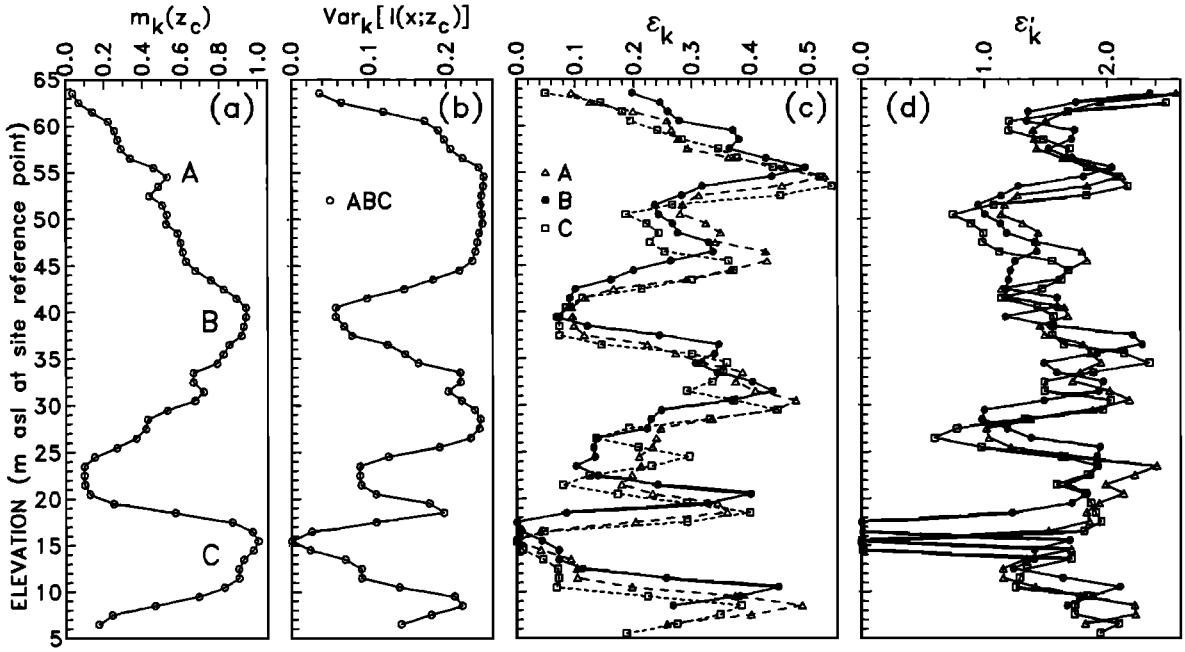

Fig. 13. (*a*) Mean, $m_k(z_c)$, and (*b*) variance, Var $_k[I(x; z_c)]$, of the indicator data in planar subsets oriented along the strike and dip of the variogram model for zone ABC (Table 2). (*c*) Mean absolute error ε_k, from cross validation of off-site planar data subsets using models for zones A, B, and C. (*d*) Relative error ε'_k, obtained by dividing ε_k by Var $_k[I(x; z_c)]$.

sections A-A′ and B-B′ (Figures 5 and 6) are expressed by the variograms' three-dimensional orientation and structure. The observed layering is reflected in the hole effect of the vertical variogram. Search plane dip angles associated with maximum continuity along dip (Figure 11) agree with observed dips. In addition, the anisotropy of the variogram ranges is consistent with the relative continuity of coarse-grained layers seen in cross section.

We can infer depositional environments from the variogram models that are reasonable given the geologic setting. The orientation of minimum and maximum continuity along the directions of stratigraphic strike and dip suggest elongated channel deposits. The 45° difference in the orientations of zones A and B may indicate a shift in streamflow direction between periods of deposition. The relatively high lateral continuity of zone C suggests that these sediments are sheet deposits.

Model Comparison and Kriging Demonstration

Although the differences among the variogram models for the individual depth zones are geologically informative, these differences may not be significant when kriging the available data. To evaluate whether the models are better able to compute known data within their respective zones, we performed a kriging cross validation following a procedure suggested by *Davis* [1987]. We divided the off-site data into 57 to 60 planar subsets, each with a vertical thickness of 1 m, oriented along the strike and dip of the variogram model under consideration. Figures 13*a* and 13*b* show the mean, $m_k(z_c)$, and variance, Var $_k[I(x; z_c)]$, respectively, of the indicator data in each plane k, using the variogram model for all zones combined (Table 2). The means approach 1.0 and 0.0 near the centers of high- and low-permeability layers, respectively. The variance is greatest around layer boundaries.

To compare the abilities of each model to estimate known data values we estimated indicator values, $I^*(x_i; z_c)$, by removing data values one at a time and reestimating the missing value using neighboring data and two-dimensional ordinary kriging with each variogram model. The kriging utilized data within a search ellipse with a principal semiaxis of 400 m in the dip direction and a minor semiaxis of 200 m along strike. For the purpose of comparing models we excluded data within a 100 m radius of the removed datum. Regardless of the model used, the mean difference between actual and estimated indicators for each zone is within ±0.02. This is sufficiently close to 0.0 to suggest that none of the models systematically overestimate or underestimate the indicator values. In addition, the errors do not exhibit a spatial trend, suggesting that the models are not spatially biased. The scattergram in Figure 14 of mean actual indicator values, $m_k(z_c)$, versus mean estimated indicator values, $m^*_k(z_c)$ for each planar subset k illustrates the good performance of all models at estimating removed values.

As a further comparison of the models for zones A, B, and C, we computed a mean absolute error ε_k by applying each of the models to the planar subsets. The ε_k are defined as

$$\varepsilon_k = \frac{1}{n} \sum_{i=1}^{n} [|I(x_i; z_c) - I^*(x_i; z_c)|] \qquad (11)$$

where n is the number of data in subset k. As with the indicator variance, ε_k for all these models are greatest near layer boundaries (Figure 13*c*). We obtained a relative error ε'_k, by dividing ε_k for each model by the indicator variance of data in planes oriented along the strike and dip of the respective models (Table 2). As shown in Figure 13*d*, ε'_k are

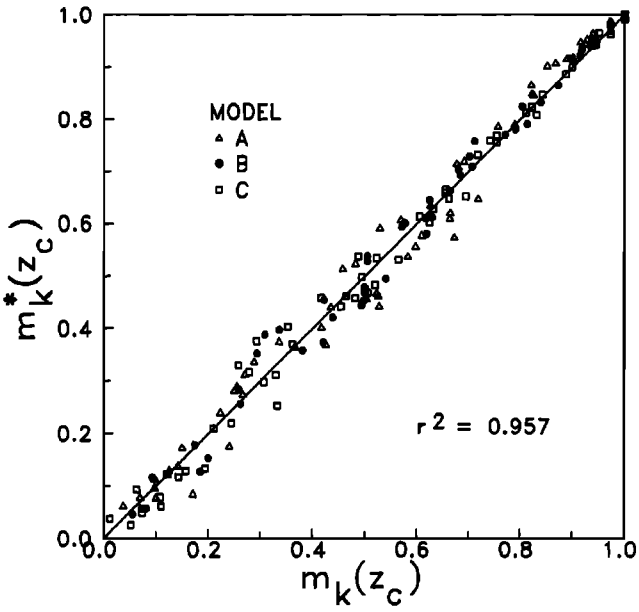

Fig. 14. Scattergram of mean actual indicator values, $m_k(z_c)$, versus mean estimated indicator values, $m_k^*(z_c)$, for each subset k. The r^2 is the goodness of fit about a 45° line.

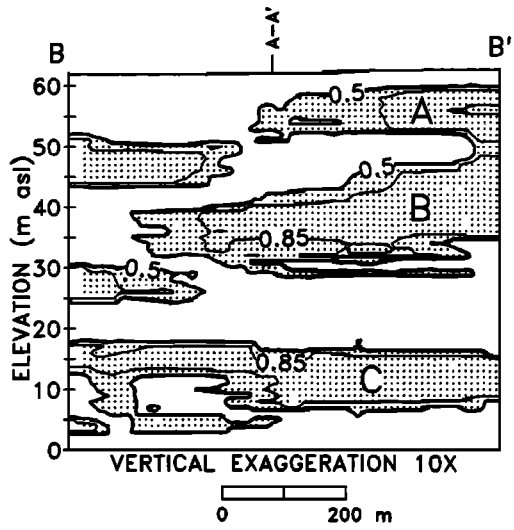

Fig. 16. Probability of high-permeability occurrence along cross section B-B' estimated by three-dimensional indicator kriging, using variogram models in Figure 12.

also similar for the three models and do not vary systematically through the depth profile, although they approach and equal zero in highly continuous portions of zone C.

The similarity of both ε_k and ε'_k for the three zonal variogram models shows that the models compute the removed indicator data values equally well, given available neighboring data (Figure 13). This reflects the similarity of the shape of the variogram models over the separation distances used in the kriging, $100 < h < 400$ m. Apparently, the errors are not sensitive to the differences in model orientation.

To complete this demonstration, we kriged the indicator data using a routine for three-dimensional ordinary kriging [*Kark*, 1985] and the indicator variogram models for zones A, B, and C. The resulting kriged values represent the probability that sediments at a particular location are of relatively high permeability. The cross sections of kriged values in Figures 15 and 16 are aligned along the A-A' and B-B' transects in Figure 3. We computed kriged values on a 10 × 10 × 1 m grid using a search ellipsoid defined by semiaxes of 400 and 150 m in the near-horizontal directions of maximum and minimum continuity of each variogram model. We used an ellipsoid semiaxis of 0.6 m vertically so that a single borehole contributed no more than two data values to each point estimate. This procedure used a maximum of four nearest data values from each octant of the ellipsoid, resulting in an average of about 10 data values for each point estimation.

Contours of the kriged values are shown in Figures 15 and 16. The similarity between the 0.50 contours and the geologic interpretations in Figures 4 and 5 is apparent. Also shown are contours representing an 85% likelihood of high-permeability occurrence. While it was difficult to represent all the nearest boreholes on Figures 4 and 5, the kriged sections reflect all the nearest data weighed anisotropically

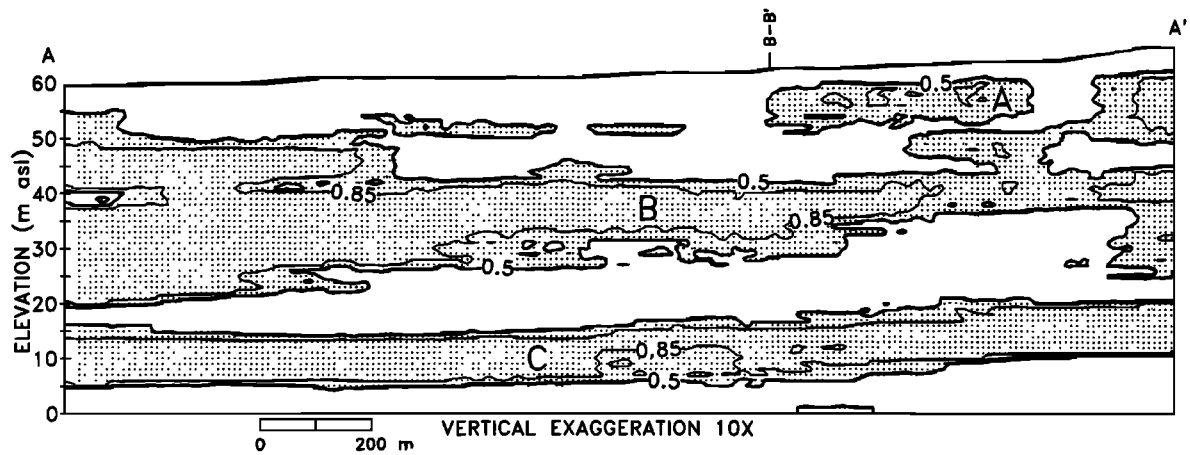

Fig. 15. Probability of high-permeability occurrence along cross section A-A' estimated by three-dimensional indicator kriging, using variogram models in Figure 12. Stippling represents locations of kriged point estimates in regions where $I(x_i; z_c) \geq 0.5$.

in three dimensions. Furthermore, estimates of uncertainty are provided by the kriged stratigraphic correlations.

SUMMARY

We have demonstrated the use of indicator geostatistics for interpreting hydrostratigraphy from qualitative borehole logs. While indicator variograms reflect observed stratigraphic features, they also reveal details that can be used to infer changes in depositional environments. Cross validation using the data available in our case study suggests that relatively subtle changes in the variogram with depth may not produce significantly different kriged results.

We used indicator kriging to estimate probabilities of inferred high-permeability occurrence. The 0.5 indicator contour appears to approximate hydrostratigraphic boundaries that would be inferred from geologic cross sections. The practical difficulty of projecting and weighing numerous boreholes onto a geologic cross section are handled in a consistent manner by kriging using a three-dimensionally anisotropic variogram model. Furthermore, this approach appears to provide a viable method for estimating uncertainties of hydrostratigraphic interpretation using qualitative data.

Acknowledgments. Financial support for this study was provided by the University of California Toxic Research and Teaching Program and the U.S. Geological Survey Water Resources Research Grant Program, grant 14-08-0001-G1311. We gratefully acknowledge the comments and suggestions of A. Journel and D. Freyberg of Stanford University, M. Reid of the U.S. Geological Survey, and G. Fogg of the University of California, Davis.

REFERENCES

Aboufirassi, M., and A. A. Marino, A geostatistically based approach to the identification of aquifer transmissivities in Yolo Basin, California, *Math. Geol.*, *16*(2), 125–137, 1984.

Carr, J. R., R. E. Bailey, and E. D. Deng, Use of indicator variograms for an enhanced spatial analysis, *Math. Geol.*, *17*(8), 797–811, 1985.

Clifton, P. M., and S. P. Neuman, Effects of kriging and inverse modeling on conditional simulation of the Avra Valley aquifer in southern Arizona, *Water Resour. Res.*, *18*(4), 1215—1234, 1982.

Dagan, G., Statistical theory of groundwater flow and transport: Pore to laboratory, laboratory to formation, and formation to regional scale, *Water Resour. Res.*, *22*(9), 120S–134S, 1986.

Davis, B. M., Indicator kriging as applied to an alluvial gold deposit, in *Geostatistics for Natural Resources Characterization*, edited by G. Verly et al., pp. 337–348, D. Reidel, Hingham, Mass., 1984.

Davis, B. M., Uses and abuses of cross-validation in geostatistics, *Math. Geol.*, *19*(3), 241–248, 1987.

Delhomme, J. P., Spatial variability and uncertainty in groundwater flow parameters: A geostatistical approach, *Water Resour. Res.*, *15*(2), 269–280, 1979.

Desbarats, A. J., Numerical estimation of effective permeability in sandshale formations, *Water Resour. Res.*, *23*(2), 273–286, 1987.

Fogg, G. E., Groundwater flow and sand body interconnectedness in a thick, multiple-aquifer system, *Water Resour. Res.*, *22*(5), 679–694, 1986*a*.

Fogg, G. E., Stochastic analysis of aquifer interconnectedness, with a test case in the Wilcox Group, east Texas, Ph.D. dissertation, 216 pp., Univ. of Tex. at Austin, 1986*b*.

Gelhar, L. W., Stochastic subsurface hydrology from theory to applications, *Water Resour. Res.*, *22*(9), 135S–145S, 1986.

Gelhar, L. W., and C. L. Axness, Three-dimensional stochastic analysis of macrodispersion in aquifers, *Water Resour. Res.*, *19*(1), 161–180, 1983.

Guven, O., F. J. Molz, and J. G. Melville, Comment on "An advection-diffusion concept for solute transport in heterogeneous unconsolidated geological deposits" by Gilham et al., *Water Resour. Res.*, *22*(1), 89–91, 1986.

Hawkins, D. M., A method for stratigraphic correlation of several boreholes, *Math. Geol.*, *16*(4), 393–406, 1984.

Helley, E. J., and E. E. Brabb, Geologic map of late Cenozoic deposits, Santa Clara County, California, *U.S. Geol. Surv. Open File Map*, 1971.

Hoeksema, R. J., and P. K. Kitanidis, An application of the geostatistical approach to the inverse problem in two-dimensional groundwater modeling, *Water Resour. Res.*, *20*(7), 1003–1020, 1984.

Hoeksema, R. J., and P. K. Kitanidis, Analysis of the spatial structure of properties of selected aquifers, *Water Resour. Res.*, *21*(4), 563–572, 1985.

Isaaks, E. H., Risk qualified mappings for hazardous waste sites: A case study in distribution free geostatistics, M.S. thesis, 85 pp., Stanford Univ., Stanford, Calif., 1984.

Journel, A. G., Nonparametric estimation of spatial distributions, *Math. Geol.*, *15*(3), 445–468, 1983.

Journel, A. J., and F. Alabert, Focusing on spatial connectivity of extreme-valued attributes: Stochastic indicator models of reservoir heterogeneities, paper presented at the 63rd Annual Technical Conference and Exhibition, Soc. of Pet. Eng., Richardson, Tex., 1988.

Journel, A. J., and C. Huijbregts, *Mining Geostatistics*, 600 pp., Academic, San Diego, Calif., 1978.

Kark, M. J., A software approach to linear geostatistics, M.S. thesis, 144 pp., Stanford Univ., Stanford, Calif., 1985.

Kemp, F., An algorithm for the stratigraphic correlation of well logs, *Math. Geol.*, *14*(3), 271–285, 1982.

Mackay, D. M., D. L. Freyberg, P. V. Roberts, and J. A. Cherry, A natural gradient experiment on solute transport in a sand aquifer, 1, Approach and overview of plume movement, *Water Resour. Res.*, *22*(13), 2017–2030, 1986.

Matheron, G., and G. de Marsily, Is transport in porous media always diffusive? A counterexample, *Water Resour. Res.*, *16*(5), 901–917, 1980.

Royle, A. G., and E. Hosgit, Local estimation of sand and gravel reserves by geostatistical methods, *Trans. Inst. Min. Metall.*, *83*, A53–A62, 1974.

Schwarzacher, W., Models for the study of stratigraphic correlation, *Math. Geol.*, *12*(3), 213–235, 1980.

Schwarzacher, W., Quantitative correlation of a cyclic limestone-shale formation, in *Quantitative Stratigraphic Correlation*, edited by J. M. Cubitt and R. A. Reyment, pp. 275–286, John Wiley, New York, 1982.

Smith, L., Spatial variability of flow parameters in a stratified sand, *Math. Geol.*, *13*(1), 1–21, 1981.

Smith, L., and F. W. Schwartz, Mass transport, 1, A stochastic analysis of macroscopic dispersion, *Water Resour. Res.*, *16*(2), 303–313, 1980.

Sowers, G. F., *Introductory Soil Mechanics and Foundations: Geotechnical Engineering*, 621 pp., Macmillan, New York, 1979.

Sudicky, E. A., A natural gradient experiment on solute transport in a sand aquifer: Spatial variability of hydraulic conductivity and its role in the dispersion process, *Water Resour. Res.*, *22*(13), 2069–2082, 1986.

Testerman, J. D., A statistical reservoir-zoning technique, *Trans. Soc. Pet. Eng. AIME*, *225*, 889–893, 1962.

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

Young, D. S., Indicator kriging for unit vectors: Rock joint orientations, *Math. Geol.*, *19*(6), 481–501, 1987.

S. J. Dreiss and N. M. Johnson, Department of Earth Sciences, University of California, Santa Cruz, CA 95064.

(Received November 30, 1988;
revised June 28, 1989;
accepted June 30, 1989.)