Mathematical Geology, Vol. 28, No. 4, 1996

Transition Probability-Based Indicator Geostatistics[1]

Steven F. Carle[2] and Graham E. Fogg[2,3]

Traditionally, spatial continuity models for indicator variables are developed by empirical curve-fitting to the sample indicator (cross-) variogram. However, geologic data may be too sparse to permit a purely empirical approach, particularly in application to the subsurface. Techniques for model synthesis that integrate hard data and conceptual models therefore are needed. Interpretability is crucial. Compared with the indicator (cross-) variogram or indicator (cross-) covariance, the transition probability is more interpretable. Information on proportion, mean length, and juxtapositioning directly relates to the transition probability; asymmetry can be considered. Furthermore, the transition probability elucidates order relation conditions and readily formulates the indicator (co)kriging equations.

KEY WORDS: cokriging, cross-covariance, cross-variogram, kriging.

INTRODUCTION

Indicator geostatistical methods are becoming increasingly popular in the earth sciences for estimation, mapping, and stochastic simulation because much geological data is categorical, for example, facies, soil classifications, mineralization phases, concentration ranges, etc. Furthermore, geological and geophysical data of continuous nature may not conform to Gaussian models, and may necessitate a nonparametric approach (Journel, 1983).

In applying an indicator geostatistical approach, spatial continuity modeling may be the most crucial and difficult step. Traditionally, sample indicator (cross-) variograms are fitted empirically with permissible and compatible mathematical functions such as the spherical or exponential models (Christakos, 1984; Armstrong, 1992; Deutsch and Journel, 1992, p. 23). In practice, data may not be sufficiently abundant to support direct computation of the model, especially in application to the subsurface, requiring some combination of empirical and subjective model fitting. Translation of subjective knowledge into

[1]Received 19 April 1995; accepted 26 July 1995.

[2]Hydrologic Science, University of California, Davis, California 95616; e-mail: carle@merri.ucdavis.edu.

[3]Department of Geology, University of California, Davis, California 95616.

0882-8121/96/0500-0453$09.50/1

spatial continuity modeling may be crucial to successful implementation of indicator geostatistics. According to Deutsch and Journel (1992, p. 58), ". . . it is subjective interpretation . . . that makes a good model; the data, by themselves, are rarely enough."

The subjective aspect of model development could be improved greatly by an ability to quantitatively translate concepts and observations into the spatial continuity model. Fundamental characteristics of indicator (categorical) variables that can be inferred subjectively are proportion, mean length (e.g., mean thickness in the vertical direction), and juxtapositioning patterns. "Juxtapositioning" herein refers to how one category locates in space preferentially or nonpreferentially with respect to another category, including nonrandom and direction-specific (asymmetric) patterns such as fining-upward cycles (see Allen, 1970). Anisotropy occurs when mean lengths differ with direction. Such conceptual information may be provided by geologic interpretations, for instance, of facies architecture or mineralization patterns.

Spatial continuity models and (co)kriging estimates for indicator variables also should obey basic laws of probability, otherwise referred to as "order relations" in geostatistics (Journel and Posa, 1990). Consideration of probability laws in spatial continuity model building helps avoid the *ad hoc* order relations corrections often required in indicator (co)kriging estimates (Deutsch and Journel, 1992, p. 77–81).

This paper shows how the transition probability, with respect to indicator (cross-) variograms and indicator (cross-) covariances, facilitates translation of subjective information and elucidates potential order relation problems. Furthermore, this paper shows that indicator (co)kriging equations can be formulated in terms of the transition probability. Thus, the transition probability can potentially replace the indicator (cross-) variogram and indicator (cross-) covariance throughout the implementation of indicator geostatistics.

DEFINITIONS

The indicator variable $I_k(\mathbf{x})$ can be defined, in general, over a region $\mathbf{D}$ by

$$I_k(\mathbf{x}) = \begin{cases} 1, & \text{if category } k \text{ occurs at location } \mathbf{x} \\ 0, & \text{otherwise} \end{cases} \tag{1}$$

where $\mathbf{x} \in \mathbf{D}$ and $k = 1, \ldots, K$. The categories may be defined as mutually exclusive units such as lithofacies or soil types (e.g., Goovaerts, 1994a) or, as in nonparametric estimation methods, by cutoff values z_k

$$I_k(\mathbf{x}) = \begin{cases} 1, & \text{if } Z(\mathbf{x}) \leq z_k \\ 0, & \text{otherwise} \end{cases} \tag{2}$$

where $Z(\mathbf{x})$ is a continuous variable (e.g., Journel and Alabert, 1989; Goovaerts, 1994b). The following discussion applies to the general situation [Eq. (1)], regardless of how the categories are defined.

Spatial Continuity Measures

Indicator (cross-) variograms or indicator (cross-) covariances traditionally are used in geostatistics to measure spatial continuity of indicator variables (see Deutsch and Journel, 1992, p. 39–60). As applied to indicator variables, the cross-variogram $\gamma_{mk}(\mathbf{h})$ is defined as

$$2\gamma_{mk}(\mathbf{h}) = E\{[I_m(\mathbf{x}) - I_m(\mathbf{x} + \mathbf{h})]\,[I_k(\mathbf{x}) - I_k(\mathbf{x} + \mathbf{h})]\} \tag{3}$$

and the cross-covariance $C_{mk}(\mathbf{h})$ is defined as

$$C_{mk}(\mathbf{h}) = E\{I_m(\mathbf{x})I_k(\mathbf{x} + \mathbf{h})\} - E\{I_m(\mathbf{x})\}E\{I_k(\mathbf{x} + \mathbf{h})\} \tag{4}$$

where $\mathbf{h}$ denotes a lag separation vector.

As applied to measuring spatial continuity, the transition probability $t_{mk}(\mathbf{h})$ denotes the conditional probability

$$t_{mk}(\mathbf{h}) = \Pr\{\text{category } k \text{ occurs at } \mathbf{x} + \mathbf{h} \,|\, \text{category } m \text{ occurs at } \mathbf{x}\}$$

Applying definitions (1) or (2) of an indicator to the definition of a conditional probability (Ross, 1988, p. 58), the definition of $t_{mk}(\mathbf{h})$ becomes

$$\begin{aligned} t_{mk}(\mathbf{h}) &= \Pr\{I_k(\mathbf{x} + \mathbf{h}) = 1 | I_m(\mathbf{x}) = 1\} \\ &= \Pr\{I_k(\mathbf{x} + \mathbf{h}) = 1 \text{ and } I_m(\mathbf{x}) = 1\}/\Pr\{I_m(\mathbf{x}) = 1\} \end{aligned} \tag{5}$$

Relations

The indicator (cross-) variogram, indicator (cross-) covariance, and transition probability are related to each other as different combinations of one-location marginal probabilities $p_k(\mathbf{x})$ defined as

$$p_k(\mathbf{x}) = \Pr\{I_k(\mathbf{x}) = 1\} = E\{I_k(\mathbf{x})\}$$

and two-location joint probabilities $p_{mk}(\mathbf{x}, \mathbf{h})$ defined as

$$p_{mk}(\mathbf{x}, \mathbf{h}) = \Pr\{I_m(\mathbf{x}) = 1 \text{ and } I_k(\mathbf{x} + \mathbf{h}) = 1\} = E\{I_m(\mathbf{x})I_k(\mathbf{x} + \mathbf{h})\}$$

Typically in practice, the assumption of a stationary model removes dependence on location $\mathbf{x}$ so that

$$E\{p_k(\mathbf{x})\} = p_k \qquad \forall \mathbf{x} \in \mathbf{D} \tag{6}$$

where p_k denotes a constant, and

$$E\{p_{mk}(\mathbf{x}, \mathbf{h})\} = p_{mk}(\mathbf{h}) \qquad \forall \mathbf{x} \in \mathbf{D} \tag{7}$$

where $p_{mk}(\mathbf{h})$ denotes a joint probability depending only on lag $\mathbf{h}$.

By applying (6) and (7), the indicator cross-variogram definition (3) becomes

$$\gamma_{mk}(\mathbf{h}) = p_{mk}(0) - [p_{mk}(\mathbf{h}) + p_{mk}(-\mathbf{h})]/2 \tag{8}$$

the indicator cross-covariance definition (4) becomes

$$C_{mk}(\mathbf{h}) = p_{mk}(\mathbf{h}) - p_m p_k \tag{9}$$

and the transition probability definition (5) becomes

$$t_{mk}(\mathbf{h}) = p_{mk}(\mathbf{h})/p_m \tag{10}$$

Substitution of (10) into (8) yields the relationship of the indicator cross-variogram with the transition probability

$$\gamma_{mk}(\mathbf{h}) = p_m\{t_{mk}(0) - [t_{mk}(\mathbf{h}) + t_{mk}(-\mathbf{h})]/2\} \tag{11}$$

and substitution of (10) into (9) yields the relationship of the indicator cross-covariance with the transition probability

$$C_{mk}(\mathbf{h}) = p_m[t_{mk}(\mathbf{h}) - p_k] \tag{12}$$

as shown by Ross (1988, p. 281).

INTERPRETABILITY

Interpretability of spatial continuity measures is important to a practitioner concerned with:

- evaluating spatial continuity data features in a geometric, probabilistic, or geologic context;
- incorporating geometric, probabilistic, or geologic concepts into spatial continuity model development;
- maintaining adherence to probability law (order relation conditions).

The spatial continuity measures $C_{mk}(\mathbf{h})$, $\gamma_{mk}(\mathbf{h})$, and $t_{mk}(\mathbf{h})$ applied to indicator variables can be estimated with similar ease (or difficulty) from the same dataset. If abundant measurements clearly constrain the spatial continuity model, then interpretability does not matter necessarily, and an empirical implementation of indicator geostatistics can proceed successfully. Otherwise, subjectivity must enter into model fitting, and interpretability becomes important for ensuring geologic plausibility.

Proportions, mean length, and spatial juxtapositioning patterns are general attributes that may be observed or inferred from geologic systems or models. Spatial juxtapositioning can be defined quantitatively in terms of asymmetry of spatial arrangement (e.g., fining or coarsening-upward cycles) and randomness of transitions from one category to another (see Miall, 1973). Rigorous procedures for infusing subjective information into spatial continuity modeling are needed to maintain consistency with conceptual models. These procedures, however, are not complete without means for avoiding order relation problems by maintaining adherence to probability law.

Proportions

Proportions, which typically can be inferred directly from the indicator data and conceptual models, can guide fitting of the sill of a spatial continuity model, whether expressed as $C_{mk}(\mathbf{h})$, $\gamma_{mk}(\mathbf{h})$, or $t_{mk}(\mathbf{h})$. Conversely, the sill of a spatial continuity model implies the assumed proportions. Thus, the relationship between proportions and the sill is integral to subjective development and conceptual understanding of the spatial continuity model.

Consider the problem of evaluating the proportions implied by models of either $C_{mk}(\mathbf{h})$, $\gamma_{mk}(\mathbf{h})$, or $t_{mk}(\mathbf{h})$. This problem arises during model building or interpretation, in order to verify that model sills are consistent with proportions known *a priori*. The sill of $t_{mk}(\mathbf{h})$ approaches p_k, the proportion of category k, whether $m = k$ or $m \neq k$. In contrast, the sills of $\gamma_{kk}(\mathbf{h})$ and $C_{kk}(0)$ approach $p_k(1 - p_k)$, which requires solving for "p" in a quadratic equation sill $= p(1 - p)$; but does p_k equal p or $1 - p$? The sill of $\gamma_{mk}(\mathbf{h})$ and $C_{mk}(0)$ for $m \neq k$ approaches $-p_m p_k$, an even more ambiguous situation.

For example, Figure 1 shows lithofacies interpretations for vertical (z-direction) boreholes drilled into an alluvial fan underlying the Lawrence Livermore National Laboratory, Livermore, California. In actuality, continuous data are available along each borehole; however, for this example, 80% of the data has been removed randomly to mimic a sparsely sampled geologic dataset. Horizontal location in Figure 1 is arbitrary. Proportions from the data shown in Figure 1 are debris flow (13%), floodplain (42%), overbank (27%), and channel (18%).

The ambiguity inherent to evaluating the proportions implied by a matrix of indicator cross-variogram (or cross-covariance) models is apparent especially in the most simple situation, a system with only two categories. Figures 2 and 3 show $t_{mk}(h_z)$ and $\gamma_{mk}(h_z)$ models fitted to measurements computed along the vertical (z) direction for a binary categorization as channel and not-channel facies. The $t_{mk}(h_z)$ model sills (dashed lines) clearly indicate an approximate 0.2 proportion of channel facies, and, correspondingly, a 0.8 proportion of not-channel facies. However, the $\gamma_{mk}(h_z)$ models and measurements for each transition mk in this binary system appear identical, except that the autovariograms are positive and the cross-variogram is negative. Ambiguity would arise in evaluating the proportions of channel facies implied by the $\gamma_{mk}(h_z)$ model sills

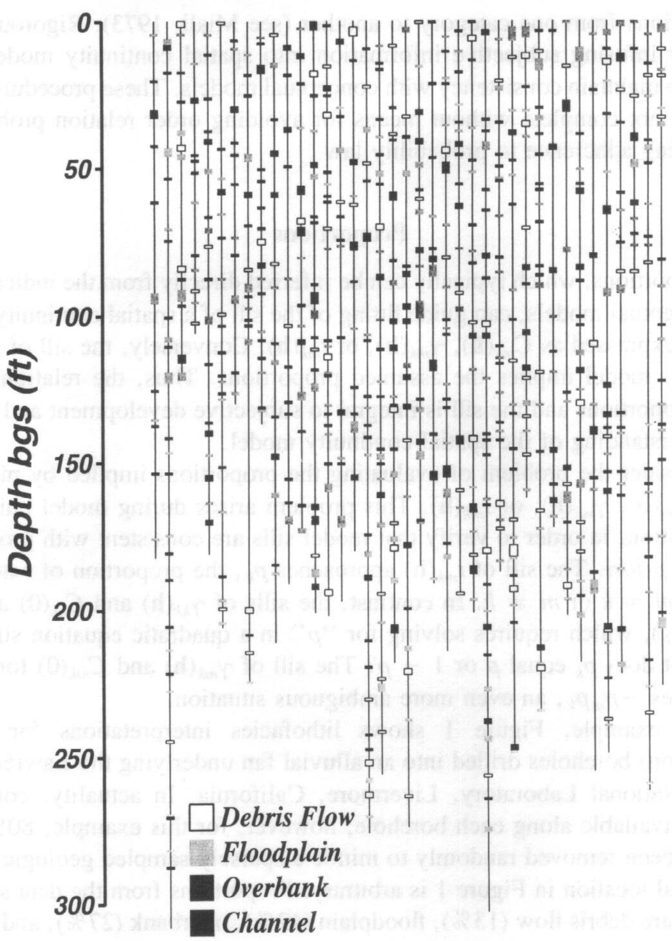

Figure 1. Vertical boreholes logged (where data are present) as four lithofacies: debris flow, floodplain, overbank, and channel. Note that horizontal location is arbitrary.

(dashed lines) from the indeterminate quadratic equation $0.15 = p(1 - p)$, which yields $p = 18\%$ or $p = 82\%$.

Difficulties intensify when evaluating the proportions implied by a matrix of indicator cross-variogram models for $K \geq 3$. Figures 4 and 5 show $t_{mk}(h_z)$ and $\gamma_{mk}(h_z)$ models and measurements for all four facies, debris flows, floodplain, overbank, and channel. The proportions p_k of each facies directly relate to the model sills (dashed lines) in each column of the $t_{mk}(h_z)$ matrix. Evaluating the proportions implied by $\gamma_{mk}(h_z)$ [or $C_{mk}(h_z)$] models requires solving systems of quadratic equations involving sill $= p_m(\delta_{mk} - p_k)$.

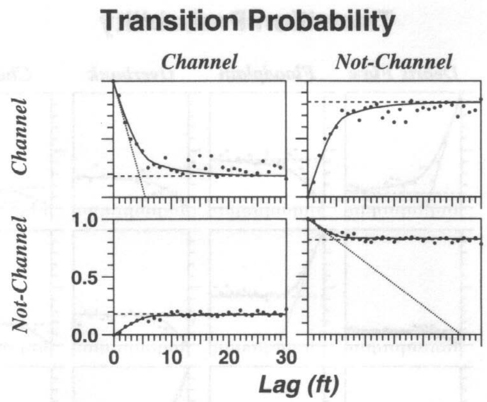

Figure 2. Vertical transition probability measurements (circles), proportions (dashed lines), and estimated slope lines (dotted lines) for data shown in Figure 1, categorized by channel and not-channel facies.

We are stressing the relationship between proportions and model sills not for the purpose of determining proportions from bivariate data, but rather to check for consistency of the spatial continuity model with proportions established by univariate data and conceptual information. Prior information on proportions (usually available) should guide subjective model fitting of the sill for either $C_{mk}(\mathbf{h})$, $\gamma_{mk}(\mathbf{h})$, and $t_{mk}(\mathbf{h})$. Of these measures, $t_{mk}(\mathbf{h})$ provides the most direct relationship between the model sill and proportions.

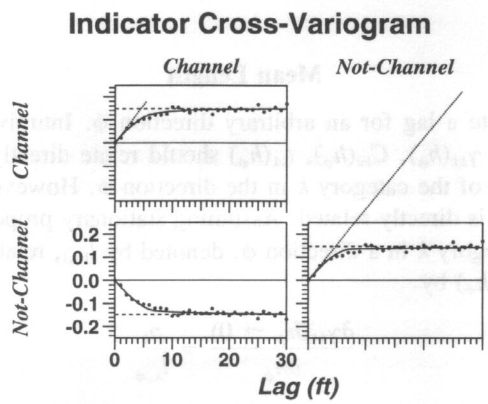

Figure 3. Vertical indicator cross-variogram measurements (circles), sills (dashed lines), and estimated slope lines (dotted lines) for data shown in Figure 1, categorized by channel and not-channel facies.

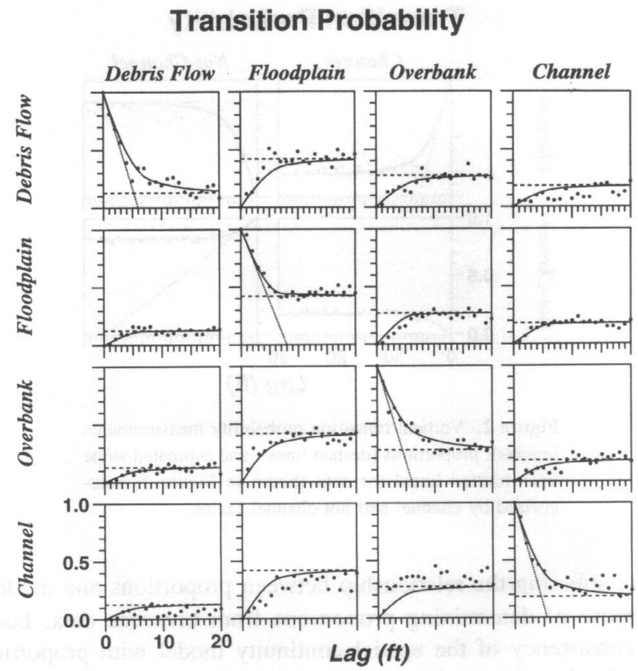

Figure 4. Corresponding to all four lithofacies shown in Figure 1: vertical transition probability measurements (circles), proportions (dashed lines), and estimated slope lines (dotted lines); autotransition probability models (solid curves, on diagonal) fitted by sum of proportions and spherical and exponential models and, corresponding to (16), random cross-transition probability models (solid curves, on off-diagonal).

Mean Length

Let h_ϕ denote a lag for an arbitrary direction ϕ. Intuitively it may seem that the range of $\gamma_{kk}(h_\phi)$, $C_{kk}(h_\phi)$, $t_{kk}(h_\phi)$ should relate directly to mean length (mean thickness) of the category k in the direction ϕ. However, it is the slope at the origin that is directly related. Assuming stationary proportions, the mean length of the category k in a direction ϕ, denoted by $\bar{l}_{k,\phi}$, relates to the slope at the origin of $\gamma_{kk}(h_\phi)$ by

$$\frac{\partial \gamma_{kk}(h_\phi \to 0)}{\partial h_\phi} = \frac{p_k}{\bar{l}_{k,\phi}} \tag{13}$$

to $C_{kk}(h_\phi)$ by

$$\frac{\partial C_{kk}(h_\phi \to 0)}{\partial h_\phi} = \frac{p_k}{\bar{l}_{k,\phi}}$$

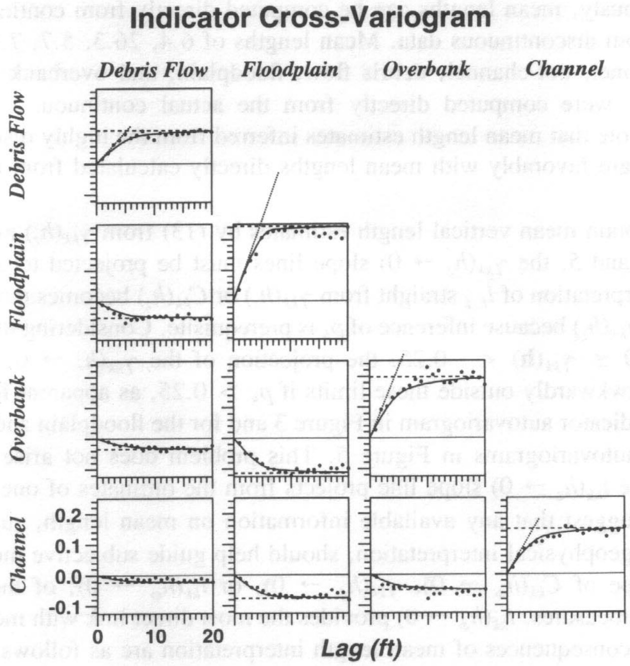

Figure 5. Corresponding to all four lithofacies shown in Figure 1: vertical indicator cross-variogram measurements (circles), sills (dashed lines), and estimated slope lines (dotted lines); autovariogram models (solid curves, on diagonal) fitted by sum of spherical and exponential models and, corresponding to (16), random cross-variogram models (solid curves, on off-diagonal).

(Dagan, 1989, p. 21–24), and to $t_{kk}(h_\phi)$ by

$$-\frac{\partial t_{kk}(h_\phi \to 0)}{\partial h_\phi} = \frac{1}{\bar{l}_{k,\phi}} \qquad (14)$$

(see derivation in Appendix A).

For example, the dotted lines in Figures 2 and 4 represent estimations of slope at the origin for $t_{kk}(h_z \to 0)$ in the "z" or vertical direction. We will refer to these dotted lines as the "slope lines." The abscissa value for the projection of a $t_{kk}(h_z \to 0)$ slope line to the intersection of the lag axis corresponds to an estimate of $\bar{l}_{k,z}$ by (14). For example, from Figure 2, mean vertical lengths of approximately 5.5 and 27.5 ft are inferred for channel and not-channel facies, respectively, by projecting the $t_{kk}(h_z \to 0)$ slope lines to the lag axis. From Figure 4, the mean vertical lengths of approximately 6, 8, and 6 ft are inferred for debris flow, floodplain, and overbank facies.

Obviously, mean lengths can be computed directly from continuous data, but not from discontinuous data. Mean lengths of 6.4, 26.3, 5.7, 7.9, and 6.1 ft for channel, not-channel, debris flow, floodplain, and overbank facies, respectively, were computed directly from the actual continuous dataset (not shown). Note that mean length estimates inferred from the highly discontinuous data compare favorably with mean lengths directly calculated from continuous data.

To obtain mean vertical length estimates by (13) from $\gamma_{kk}(h_z)$ as shown in Figures 3 and 5, the $\gamma_{kk}(h_z \rightarrow 0)$ slope lines must be projected to an ordinate of p_k. Interpretation of $\bar{l}_{k,z}$ straight from $\gamma_{kk}(h_z)$ or $C_{kk}(h_z)$ becomes more difficult than from $t_{kk}(h_z)$ because inference of p_k is prerequisite. Considering the practical limits of $0 \leq \gamma_{kk}(\mathbf{h}) < \sim 0.25$, the projection of the $\gamma_{kk}(h_z \rightarrow 0)$ slope line proceeds awkwardly outside these limits if $p_k > 0.25$, as apparent for the not-channel indicator autovariogram in Figure 3 and for the floodplain and overbank indicator autovariograms in Figure 5. This problem does not arise for $t_{kk}(h_z)$ because the $t_{kk}(h_z \rightarrow 0)$ slope line projects from the ordinates of one to zero.

We suggest that any available information on mean length, such as geological or geophysical interpretation, should help guide subjective model fitting of the slope of $C_{kk}(h_\phi \rightarrow 0)$, $\gamma_{kk}(h_\phi \rightarrow 0)$, or $t_{kk}(h_\phi \rightarrow 0)$; of these spatial continuity measures, $t_{kk}(h_\phi \rightarrow 0)$ provides the most direct link with mean length. Important consequences of mean length interpretation are as follows:

- if $\bar{l}_{k,\phi}$ is interpreted as greater than zero for a mutually exclusively defined category k, then the nugget for a $C_{kk}(h_\phi)$, $\gamma_{kk}(h_\phi)$, or $t_{kk}(h_\phi)$ model should be zero;
- if $\bar{l}_{k,\phi}$ cannot be computed directly (because of discontinuous data sampling), it can be interpreted from the $C_{kk}(h_\phi \rightarrow 0)$, $\gamma_{kk}(h_\phi \rightarrow 0)$, or $t_{kk}(h_\phi \rightarrow 0)$ slope line of the fitted model; and
- if data are too sparse to interpret $\bar{l}_{k,\phi}$, for example, in nonvertical directions for most subsurface investigations, the fitting of the slope for a $C_{kk}(h_\phi \rightarrow 0)$, $\gamma_{kk}(h_\phi \rightarrow 0)$, or $t_{kk}(h_\phi \rightarrow 0)$ model can be guided by reasonable estimates of mean length obtained from conceptual or other indirect information.

Thus, given information on proportions and mean lengths, the practitioner has most of the information needed to synthesize plausible models of either $C_{kk}(\mathbf{h})$, $\gamma_{kk}(\mathbf{h})$, or $t_{kk}(\mathbf{h})$. The slope line and proportion will dictate the sill and (effective) range if specific model structures such as spherical or exponential functions are assumed. In Figures 4 and 5, sums of spherical and exponential functions that agree with the proportions and inferred average lengths as noted were fitted to the sample $t_{kk}(h_z)$ and $\gamma_{kk}(h_z)$, as represented by the solid curves.

Asymmetry

Asymmetry implies that

$$p_{mk}(\mathbf{h}) \neq p_{mk}(-\mathbf{h}) \tag{15a}$$

or

$$p_{mk}(\mathbf{h}) \neq p_{km}(\mathbf{h}) \tag{15b}$$

By applying (15a) or (15b) to (9) and (10), it is apparent that indicator cross-covariances and transition probabilities can consider asymmetry. When $m \neq k$ it is possible that $C_{mk}(\mathbf{h}) \neq C_{mk}(-\mathbf{h})$ or $C_{mk}(\mathbf{h}) \neq C_{km}(\mathbf{h})$, and $t_{mk}(\mathbf{h}) \neq t_{mk}(-\mathbf{h})$ or $p_m t_{mk}(\mathbf{h}) \neq p_k t_{km}(\mathbf{h})$. However, for indicator cross-variograms the averaging of $p_{mk}(\mathbf{h})$ and $p_{mk}(-\mathbf{h})$ evident in (8) or (11) requires that $\gamma_{mk}(\mathbf{h}) = \gamma_{km}(\mathbf{h})$ or $\gamma_{mk}(\mathbf{h}) = \gamma_{mk}(-\mathbf{h})$. Thus, indicator cross-variograms cannot reveal asymmetry.

Figures 3 and 5 need only show the lower (or upper) triangle of the indicator cross-variogram matrix because of the symmetry assumption implicit to $\gamma_{mk}(\mathbf{h})$. This should not be taken necessarily as a practical advantage. Before employing $\gamma_{mk}(\mathbf{h})$ as a spatial continuity measure, the practitioner should consider the desirability or justifiability of an assumption of symmetry. Obviously, asymmetries do exist in geologic systems, for example, upward-fining/coarsening cycles of depositional units in fluvial environments (see Allen, 1970). If indicator geostatistical estimation and simulation methods are to be taken seriously as geologic modeling tools, asymmetry should be addressed.

Randomness

One of the main uses of transition probability measures by geologists has been to analyze how transitions from one geologic unit to another deviate, if at all, from randomness (Miall, 1973). Assuming a stationary model of $t_{mk}(\mathbf{h})$, $[1 - t_{mm}(\mathbf{h})]$ is the probability of transitioning to any category $k \neq m$ at $\mathbf{x} + \mathbf{h}$ given category m at $\mathbf{x}$. The relative proportions for the categories $k \neq m$ are $p_k/[1 - p_m]$. If $t_{mk}(\mathbf{h})$ for $k \neq m$ is assumed to depend randomly only on the relative proportions of categories $k \neq m$, then such a random transition probability model, denoted by $t_{mk}^{(r)}(\mathbf{h})$, is

$$t_{mk}^{(r)}(\mathbf{h}) = [1 - t_{mm}(\mathbf{h})]p_k/[1 - p_m] \qquad \text{for } k \neq m \tag{16}$$

In Figure 4, the solid curves in the off-diagonal plots represent $t_{mk}^{(r)}(h_z)$ as calculated by (16).

Note in Figure 4 that $t_{mk}^{(r)}(h_z)$ fits the sample $t_{mk}(h_z)$ fairly well, except for noticeably below random transitions for debris flow → channel, channel → debris flow, and channel → floodplain, and a well above random transition for channel → overbank. These nonrandom transitions can be interpreted in a stratigraphic context. The debris flow deposits tend not to be deposited or preserved (or recognized?) near channel deposits, perhaps because of erosion in channel areas and increased preservation potential in floodplain areas. The tendency for overbank deposits to overlie channel deposits is consistent with an application of Walther's Law (see Leeder, 1982, p. 122) to fluvial depositional models that show overbank deposits (such as levees and crevasse splays) occurring laterally proximal to channel deposits (see Allen, 1970). Conversely, stratigraphic interpretations of juxtapositioning patterns for depositional units could help guide fitting of the off-diagonal cross-transition probability models, especially to judge whether a model should fall below, at, or above randomness.

Applying (11) to (16) transforms $t_{mk}^{(r)}(h_z)$ to $\gamma_{mk}^{(r)}(h_z)$, a corresponding random indicator cross-variogram, also represented in Figure 5 by the solid curves. The $\gamma_{mk}^{(r)}(h_z)$ model fits the sample $\gamma_{mk}(h_z)$ well overall. Judging from $\gamma_{mk}(h_z)$ alone, one might interpret these data as representative of random vertical juxtapositioning relationships. However, the sample $\gamma_{mk}(h_z)$ ignores evidence for asymmetric structure in the data by averaging data pairs for opposing lag directions. Thus, the implicit assumption of symmetry in indicator cross-variograms could lead to misleading interpretations that spatial cross-relations generally are random when, in fact, considerably nonrandom, asymmetric juxtapositioning patterns may actually exist.

Order Relation Conditions

"Order relation" conditions in indicator geostatistics are imposed to assure that indicator (co)kriging estimates and indicator (cross-)variogram/covariance models conform with basic laws of probability theory.

Mutually Exclusive Categories. Let $i_k(\mathbf{x})$ denote an indicator data value prescribing $I_k(\mathbf{x}) = i_k(\mathbf{x})$ at a location $\mathbf{x}$. For indicator variables defined by mutually exclusive categories, such as geologic units, the indicator (co)kriging estimates $[i_k(\mathbf{x})]^*_{(\mathrm{co})K}$ constitute conditional probability estimates (Journel, 1983; Solow, 1986)

$$[i_k(\mathbf{x})]^*_{(\mathrm{co})K} = \Pr\{I_k(\mathbf{x}) = 1 \mid \text{surrounding data}\}$$

that should obey

$$0 \leq [i_k(\mathbf{x})]^*_{(\mathrm{co})K} \leq 1 \qquad \forall k$$

and

$$1 = \sum_{k=1}^{K} [i_k(\mathbf{x})]^*_{(\mathrm{co})K}$$

according to probability laws (Ross, 1988, p. 29, p. 85). Similarly, transition probability models should obey

$$0 \leq t_{mk} \leq 1 \quad \forall m, k \tag{17a}$$

and

$$1 = \sum_{k=1}^{K} t_{mk}(\mathbf{h}) \quad \forall m \tag{17b}$$

Corresponding order relation conditions for $\gamma_{mk}(\mathbf{h})$ and $C_{mk}(\mathbf{h})$ can be derived by applying (11) and (12) to (17a,b).

Cumulatively Defined Categories. For cumulatively defined indicator variables, such as categories defined by cutoffs of continuous variables as in (2), indicator (co)kriging estimates should obey

$$0 \leq [i_{k-1}(\mathbf{x})]^*_{(\mathrm{co})K} \leq [i_k(\mathbf{x})]^*_{(\mathrm{co})K} \leq [i_{k+1}(\mathbf{x})]^*_{(\mathrm{co})K} \leq 1 \quad \forall k \tag{18}$$

and joint probabilities should obey

$$0 \leq p_{m(k-1)}(\mathbf{h}) \leq p_{mk}(\mathbf{h}) \leq p_{m(k+1)}(\mathbf{h}) \leq p_m \quad \forall m, k \tag{19a}$$

and

$$0 \leq p_{(m-1)k}(\mathbf{h}) \leq p_{mk}(\mathbf{h}) \leq p_{(m+1)k}(\mathbf{h}) \leq p_k \quad \forall m, k \tag{19b}$$

Applying (10) to (19a, 19b), transition probabilities models should obey

$$0 \leq t_{m(k-1)}(\mathbf{h}) \leq t_{mk}(\mathbf{h}) \leq t_{m(k+1)}(\mathbf{h}) \leq 1 \quad \forall m, k \tag{20a}$$

and

$$0 \leq p_{(m-1)}t_{(m-1)k}(\mathbf{h}) \leq p_m t_{mk}(\mathbf{h}) \leq p_{(m+1)}t_{(m+1)k}(\mathbf{h}) \leq p_k \quad \forall m, k \tag{20b}$$

Journel and Posa (1990) warn that some of the order relation conditions for $\gamma_{mk}(\mathbf{h})$ and $C_{mk}(\mathbf{h})$ are "cumbersome and not practical to use" and "tedious and little informative." Instead of deriving the order relation conditions for $\gamma_{mk}(\mathbf{h})$ and $C_{mk}(\mathbf{h})$ pertaining to cumulatively defined categories, we suggest maintaining order relation conditions with $p_{mk}(\mathbf{h})$ by (19a, 19b) or with $t_{mk}(\mathbf{h})$ by (20a, 20b); $\gamma_{mk}(\mathbf{h})$ and $C_{mk}(\mathbf{h})$ functions that satisfy order relation conditions then can be obtained by applying (8) and (9) or (11) and (12).

REFORMULATION OF INDICATOR (CO)KRIGING SYSTEMS

Traditionally, the indicator (co)kriging equations are formulated in terms of $C_{mk}(\mathbf{h})$. Given the interpretive advantages of transition probabilities, one might desire to reformulate the indicator (co)kriging systems of equations in terms of $t_{mk}(\mathbf{h})$, and, thus, bypass the transformation step (12). Next we show this to be possible because $C_{mk}(\mathbf{h})$ and $t_{mk}(\mathbf{h})$ are linearly related by (12).

Indicator Kriging

Let $\mathbf{x}$ denote an estimation location in the region $\mathbf{D}$. The indicator kriging estimates assume that $\Pr\{I_k(\mathbf{x}) = 1 | \text{surrounding data}\}$ depends on a weighted sum of indicator data values $i_k(\mathbf{x}_\alpha)$ for category k only at $\alpha = 1, \ldots, N$ nearby locations. The simple indicator kriging estimate $[i_k(\mathbf{x})]^*_{SK}$ assumes

$$[i_k(\mathbf{x})]^*_{SK} = p_k + \sum_{\alpha=1}^{N} [i_k(\mathbf{x}_\alpha) - p_k]\lambda_{k,\alpha} \tag{21}$$

where each $\lambda_{k,\alpha}$ is a weighting coefficient pertaining to the category k and at the location $\mathbf{x}_\alpha$ (Solow, 1986; Deutsch and Journel, 1992, p. 73). The ordinary indicator kriging estimate assumes

$$[i_k(\mathbf{x})]^*_{OK} = \sum_{\alpha=1}^{N} i_k(\mathbf{x}_\alpha)\lambda_{k,\alpha} \tag{22}$$

subject to the "unbiasedness constraint"

$$\sum_{\alpha=1}^{N} \lambda_{k,\alpha} = 1 \tag{23}$$

(Deutsch and Journel, 1992, p. 74).

The kriging estimates for the weights $\lambda_{k,\alpha}$ are formulated to minimize the estimation error variance, subject to (23) for the ordinary kriging case (see Journel and Huijbregts, 1978, p. 304–306, 561–562). The resulting simple indicator kriging system of equations is

$$\begin{bmatrix} C_{kk}(\mathbf{x}_1 - \mathbf{x}_1) & \cdots & C_{kk}(\mathbf{x}_N - \mathbf{x}_1) \\ \vdots & \ddots & \vdots \\ C_{kk}(\mathbf{x}_1 - \mathbf{x}_N) & \cdots & C_{kk}(\mathbf{x}_N - \mathbf{x}_N) \end{bmatrix} \begin{bmatrix} \lambda_{k,1} \\ \vdots \\ \lambda_{k,N} \end{bmatrix} = \begin{bmatrix} C_{kk}(\mathbf{x} - \mathbf{x}_1) \\ \vdots \\ C_{kk}(\mathbf{x} - \mathbf{x}_N) \end{bmatrix} \tag{24}$$

and the ordinary indicator kriging system of equations is

$$\begin{bmatrix} C_{kk}(\mathbf{x}_1 - \mathbf{x}_1) & \cdots & C_{kk}(\mathbf{x}_N - \mathbf{x}_1) & 1 \\ \vdots & \ddots & \vdots & \vdots \\ C_{kk}(\mathbf{x}_1 - \mathbf{x}_N) & \cdots & C_{kk}(\mathbf{x}_N - \mathbf{x}_N) & 1 \\ 1 & \cdots & 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_{k,1} \\ \vdots \\ \lambda_{k,n} \\ \mu \end{bmatrix} = \begin{bmatrix} C_{kk}(\mathbf{x} - \mathbf{x}_1) \\ \vdots \\ C_{kk}(\mathbf{x} - \mathbf{x}_N) \\ 1 \end{bmatrix} \tag{25}$$

where $C_{kk}(\mathbf{h})$ is an indicator covariance model defined by (4) for $m = k$, and μ denotes a Lagrange parameter.

If the indicator covariance model assumes stationary proportions, then the simple indicator cokriging system of equations can be reformulated as

$$\begin{bmatrix} t_{kk}(\mathbf{x}_1 - \mathbf{x}_1) - p_k & \cdots & t_{kk}(\mathbf{x}_N - \mathbf{x}_1) - p_k \\ \vdots & \ddots & \vdots \\ t_{kk}(\mathbf{x}_1 - \mathbf{x}_N) - p_k & \cdots & t_{kk}(\mathbf{x}_N - \mathbf{x}_N) - p_k \end{bmatrix} \begin{bmatrix} \lambda_{k,1} \\ \vdots \\ \lambda_{k,N} \end{bmatrix} = \begin{bmatrix} t_{kk}(\mathbf{x} - \mathbf{x}_1) - p_k \\ \vdots \\ t_{kk}(\mathbf{x} - \mathbf{x}_N) - p_k \end{bmatrix} \tag{26}$$

by applying (12) for $m = k$ and dividing each row by p_k. Under the unbiasedness constraint (23), the relation

$$\sum_{\alpha=1}^{N} p_k \lambda_{k,\alpha} = p_k \tag{27}$$

holds, so that the ordinary indicator kriging system of equations can be reformulated as

$$\begin{bmatrix} t_{kk}(\mathbf{x}_1 - \mathbf{x}_1) & \cdots & t_{kk}(\mathbf{x}_N - \mathbf{x}_1) & 1 \\ \vdots & \ddots & \vdots & \vdots \\ t_{kk}(\mathbf{x}_1 - \mathbf{x}_N) & \cdots & t_{kk}(\mathbf{x}_N - \mathbf{x}_N) & 1 \\ 1 & \cdots & 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_{k,1} \\ \vdots \\ \lambda_{k,N} \\ \mu \end{bmatrix} = \begin{bmatrix} t_{kk}(\mathbf{x} - \mathbf{x}_1) \\ \vdots \\ t_{kk}(\mathbf{x} - \mathbf{x}_N) \\ 1 \end{bmatrix} \tag{28}$$

by applying (12) for $m = k$, dividing each row by p_k, and applying (27) to cancel terms. The systems of equations (26) and (28) are algebraically equivalent to (24) and (25), respectively, demonstrating that indicator kriging is formulated readily with transition probabilities.

Indicator Cokriging

Indicator cokriging estimates assume that $\Pr\{I_k(\mathbf{x}) = 1 | \text{surrounding data}\}$ depends on a weighted sum of indicator data values $i_m(\mathbf{x}_\alpha)$ for all categories $m = 1, \ldots, K$ at $\alpha = 1, \ldots, N$ nearby data locations. The simple indicator cokriging estimate $[i_k(\mathbf{x})]^*_{\text{co}SK}$ assumes

$$[i_k(\mathbf{x})]^*_{\text{co}SK} = p_k + \sum_{\alpha=1}^{N} \sum_{m=1}^{K} [i_m(\mathbf{x}_\alpha) - p_m]\, \lambda_{mk,\alpha}$$

where each $\lambda_{mk,\alpha}$ is a weighting coefficient pertaining to data for category m at location $\mathbf{x}_\alpha$. The ordinary indicator cokriging estimate $[i_k(\mathbf{x})]^*_{\text{co}OK}$ assumes

$$[i_k(\mathbf{x})]^*_{\text{co}OK} = \sum_{\alpha=1}^{N} \sum_{m=1}^{K} i_m(\mathbf{x}_\alpha)\, \lambda_{mk,\alpha} \tag{29}$$

subject to the "unbiasedness constraint"

$$\sum_{\alpha=1}^{N} \lambda_{mk,\alpha} = I_{mk} \tag{30}$$

where I_{mk} denotes the identity matrix (Goovaerts, 1994a, 1994b).

Cokriging estimates for the weights $\lambda_{mk,\alpha}$ are formulated by minimizing the expected error variance in the cokriging estimates subject to the unbiasedness constraint (30) for the ordinary cokriging estimate (see Myers, 1982). The resulting simple indicator cokriging system of equations is

$$\begin{bmatrix} C_{ml}(\mathbf{x}_1 - \mathbf{x}_1) & \cdots & C_{ml}(\mathbf{x}_N - \mathbf{x}_1) \\ \vdots & \ddots & \vdots \\ C_{ml}(\mathbf{x}_1 - \mathbf{x}_N) & \cdots & C_{ml}(\mathbf{x}_N - \mathbf{x}_N) \end{bmatrix} \begin{bmatrix} \lambda_{lk,1} \\ \vdots \\ \lambda_{lk,N} \end{bmatrix} = \begin{bmatrix} C_{mk}(\mathbf{x} - \mathbf{x}_1) \\ \vdots \\ C_{mk}(\mathbf{x} - \mathbf{x}_N) \end{bmatrix} \tag{31}$$

where

$$C_{ml}(\mathbf{h}),\ C_{mk}(\mathbf{h}) = \begin{bmatrix} C_{11}(\mathbf{h}) & \cdots & C_{1K}(\mathbf{h}) \\ \vdots & \ddots & \vdots \\ C_{K1}(\mathbf{h}) & \cdots & C_{KK}(\mathbf{h}) \end{bmatrix}$$

The resulting ordinary cokriging system of equations is

$$\begin{bmatrix} C_{ml}(\mathbf{x}_1 - \mathbf{x}_1) & \cdots & C_{ml}(\mathbf{x}_N - \mathbf{x}_1) & I_{ml} \\ \vdots & \ddots & & \vdots \\ C_{ml}(\mathbf{x}_1 - \mathbf{x}_N) & \cdots & C_{ml}(\mathbf{x}_N - \mathbf{x}_N) & I_{ml} \\ I_{ml} & \cdots & I_{ml} & 0 \end{bmatrix} \begin{bmatrix} \lambda_{lk,1} \\ \vdots \\ \lambda_{lk,N} \\ \mu_{lk} \end{bmatrix} = \begin{bmatrix} C_{mk}(\mathbf{x} - \mathbf{x}_1) \\ \vdots \\ C_{mk}(\mathbf{x} - \mathbf{x}_N) \\ I_{mk} \end{bmatrix} \tag{32}$$

where μ_{lk} denotes a $K \times K$ matrix of Lagrange parameters.

If the indicator cross-covariance models assume stationary proportions, then the simple cokriging system of equations for indicator variables can be transformed to

$$\begin{bmatrix} t_{ml}(\mathbf{x}_1 - \mathbf{x}_1) - p_l & \cdots & t_{ml}(\mathbf{x}_N - \mathbf{x}_1) - p_l \\ \vdots & \ddots & \vdots \\ t_{ml}(\mathbf{x}_1 - \mathbf{x}_N) - p_l & \cdots & t_{ml}(\mathbf{x}_N - \mathbf{x}_N) - p_l \end{bmatrix} \begin{bmatrix} \lambda_{lk,1} \\ \vdots \\ \lambda_{lk,N} \end{bmatrix} = \begin{bmatrix} t_{mk}(\mathbf{x} - \mathbf{x}_1) - p_k \\ \vdots \\ t_{mk}(\mathbf{x} - \mathbf{x}_N) - p_k \end{bmatrix} \tag{33}$$

by applying (12) and dividing each row by p_m, where

$$t_{ml}(\mathbf{h}), t_{mk}(\mathbf{h}) = \begin{bmatrix} t_{11}(\mathbf{h}) & \cdots & t_{1K}(\mathbf{h}) \\ \vdots & \ddots & \vdots \\ t_{K1}(\mathbf{h}) & \cdots & t_{KK}(\mathbf{h}) \end{bmatrix}$$

Under the unbiasedness constraint (30), the relation

$$\sum_{\alpha=1}^{N} \sum_{l=1}^{K} p_l \lambda_{lk,\alpha} = p_k \tag{34}$$

holds, so that the ordinary indicator cokriging system of equations can be reformulated as

$$\begin{bmatrix} t_{ml}(\mathbf{x}_1 - \mathbf{x}_1) & \cdots & t_{ml}(\mathbf{x}_N - \mathbf{x}_1) & I_{ml} \\ \vdots & \ddots & \vdots & \vdots \\ t_{ml}(\mathbf{x}_1 - \mathbf{x}_N) & \cdots & t_{ml}(\mathbf{x}_N - \mathbf{x}_N) & I_{ml} \\ I_{ml} & \cdots & I_{ml} & 0 \end{bmatrix} \begin{bmatrix} \lambda_{lk,1} \\ \vdots \\ \lambda_{lk,N} \\ \mu_{lk} \end{bmatrix} = \begin{bmatrix} t_{mk}(\mathbf{x} - \mathbf{x}_1) \\ \vdots \\ t_{mk}(\mathbf{x} - \mathbf{x}_N) \\ I_{mk} \end{bmatrix} \tag{35}$$

by applying (12), dividing each row by p_m, and applying (34) to cancel terms. The systems of Equations (33) and (35) are algebraically equivalent to (31) and (32), respectively, demonstrating that indicator cokriging can be formulated easily with transition probabilities. In Appendix B, the ordinary indicator cokriging equations are derived in terms of transition probabilities directly from the indicator cokriging estimate.

Nonnegative Definiteness and Singularity

Nonnegative definite (cross-) covariance models are required to ensure the existence and uniqueness of solutions to the (co)kriging systems of equations, as well as the nonnegative variance of the (co)kriging estimates (Christakos, 1984). Recall from the previous section that the transition probability-based indicator (co)kriging formulations replace $C_{ml}(\mathbf{h})$ with $t_{ml}(\mathbf{h}) - p_l = C_{ml}(\mathbf{h})/p_m$ in comparison to the traditional formulations. Considering that the p_l are positive constants, a (co)kriging matrix formulated with $C_{ml}(\mathbf{h})/p_m$ will be nonnegative definite if a (co)kriging matrix formulated with $C_{ml}(\mathbf{h})$ also is nonnegative definite (Horn and Johnson, 1991, p. 309). Thus, nonnegative definiteness requirements for solving transition probability-based indicator (co)kriging system of equations (26), (28), (33), and (35) will be satisfied if $t_{ml}(\mathbf{h})$ is modeled by the sum of p_l and the same nonnegative definite functions traditionally used to model $C_{ml}(\mathbf{h})$.

Constant sum requirements may be inherent to the definition of indicator variables; for example, $\Sigma_{k=1}^{K} I_k(\mathbf{x}) = 1$ is required for indicator variables of mutually exclusive, exhaustively defined categories such as geologic units. Constant sum requirements will lead to singularities in the indicator cokriging matrix, whether formulated by $C_{ml}(\mathbf{h})$ or $t_{ml}(\mathbf{h}) - p_l$. Such singularity problems fall under the general case of cokriging variables of a constant sum, referred to as a "regionalized composition" by Pawlowsky (1989) and Pawlowsky, Olea, and Davis (1995). To avoid spurious estimates of the weighting coefficients because of singularities, the indicator cokriging systems of equations can be solved by singular value decomposition techniques (Press and others, 1992, p. 51–63) rather than Gaussian elimination, Gauss–Jordan elimination, Cholesky decomposition, or LU decomposition techniques usually employed in geostatistics (see Deutsch and Journel, 1992, p. 217–219).

An Order Relation Problem

In indicator kriging of cumulatively defined categories, violations of the order relation conditions (18) can occur because of the implicit approximation that the estimate depends on indicator data for category k only (not categories $m \neq k$). The indicator kriging estimate $[i_k(\mathbf{x})]_K^*$ denotes the conditional probability

$$[i_k(\mathbf{x})]_K^* = \Pr\{I_k(\mathbf{x}) = 1 | I_k(\mathbf{x}_\alpha) = i_k(\mathbf{x}_\alpha);\ \alpha = 1, \ldots, N\}$$

Now consider the $N = 1$ (one datum) indicator kriging estimation case with $i_k(\mathbf{x}_1) = 1$ as examined by Solow (1986) and Journel (1993)

$$[i_k(\mathbf{x})]_K^* = \Pr\{I_k(\mathbf{x}) = 1 | I_k(\mathbf{x}_1) = 1\} \tag{36a}$$

Applying the definition of the transition probability (5) to (36a) yields

$$[i_k(\mathbf{x})]_K^* = t_{kk}(\mathbf{x} - \mathbf{x}_1) \quad (36b)$$

for the one-datum indicator kriging estimate with $i_k(\mathbf{x}_1) = 1$.

Applying (36b) to (18), order relation conditions for the one-datum indicator kriging *estimate* (36a) are

$$0 \leq t_{(k-1)(k-1)}(\mathbf{h}) \leq t_{kk}(\mathbf{h}) \leq t_{(k+1)(k+1)}(\mathbf{h}) \leq 1 \quad (37)$$

Applying (20a) and (20b) in sequence, order relation conditions for the autotransition probability *models* are

$$0 \leq p_{(k-1)}t_{(k-1)(k-1)}(\mathbf{h}) \leq p_k t_{kk}(\mathbf{h}) \leq p_{(k+1)}t_{(k+1)(k+1)}(\mathbf{h}) \leq 1 \quad (38)$$

Obviously, the order relation conditions inherent to the one-datum indicator kriging estimate (37) are not consistent with the order relation conditions required of the indicator kriging autotransition probability models (38). Thus, even if the spatial continuity models obey order relation conditions, order relation violations can occur in the indicator kriging estimates.

For example, Deutsch and Journel (1992, p. 262–265) present a suite of standardized (normalized by the variance), omnidirectional $\gamma_{kk}(\mathbf{h})$ models fit to 140 data at nine cutoffs based on cumulative probabilities with multiples of 0.1, shown in Figure 6. We transformed these indicator variogram models to the corresponding $t_{kk}(\mathbf{h})$ models, shown in Figure 7. These spatial continuity models do not violate order relation conditions as specified by Journel and Posa (1990) for $\gamma_{kk}(\mathbf{h})$ and, correspondingly, by (38) for $t_{kk}(\mathbf{h})$. However, these models are not consistent with the order relation conditions in (37) inherent to the one-datum indicator kriging estimates. As apparent from the crossovers that occur in the $t_{kk}(\mathbf{h})$ models in Figure 7, the order relation conditions in (37) would be violated for indicator kriging estimates involving data from the 0.1, 0.8, and

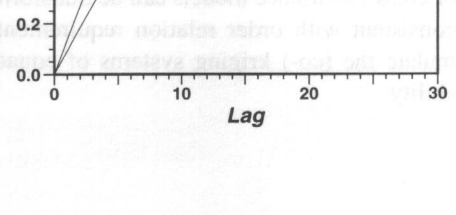

Figure 6. Suite of standardized indicator variogram models described in Deutsch and Journel (1992, p. 262–265).

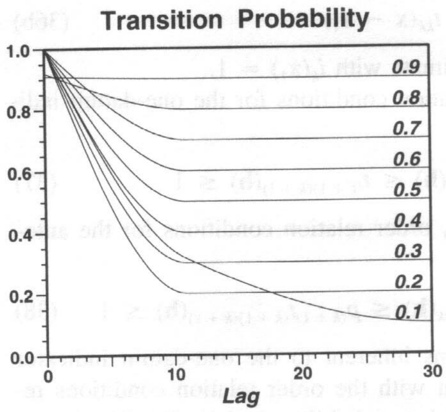

Figure 7. Suite of autotransition probability models corresponding to variogram models in Figure 6.

0.9 cutoffs at $\mathbf{h} < 17.66$, $\mathbf{h} < 2.40$, and $\mathbf{h} < 3.35$, respectively. Such cross-overs do not appear in the $\gamma_{kk}(\mathbf{h})$ models in Figure 6. It is important to realize that such order relation problems arise not from variogram data quality, but from the indicator kriging approximation that estimates depend only on data for the same category.

Note that the median indicator kriging estimate, which uses the same standardized indicator variogram for each cutoff, would not violate (37) or (38). Not surprisingly, "... median indicator kriging drastically reduces the number of order relation deviations ..." (Deutsch and Journel, 1992, p. 80).

CONCLUSIONS

Indicator geostatistics can be implemented entirely in terms of the transition probability instead of the indicator (cross-) variogram or covariance. The transition probability offers advantages where integration of indirect, subjective, or conceptual information is desired. By simple graphical observation, features of the transition probability can be related precisely to geological attributes such as proportion, mean length, and juxtapositioning patterns. Such interpretability not only improves understanding of spatial continuity measurements, but facilitates construction of geologically plausible spatial continuity models.

Moreover, the transition probability readily infuses order relation conditions required by probability law. Corresponding indicator (cross-) variogram or cross-covariance models can be transformed from transition probability models consistent with order relation requirements. Conversely, it is possible to formulate the (co-) kriging systems of equations in terms of the transition probability.

ACKNOWLEDGMENTS

This work was supported by the Lawrence Livermore National Laboratory, N.I.E.H.S. Superfund Grant (ES-04699), U.S.G.S. Water Resources Research Grant (14-18-001-61909), and the U.S. EPA (R819658) Center for Ecological Health Research at UC Davis. Although the information in this document has been funded in part by the United States Environmental Protection Agency, it may not necessarily reflect the views of the Agency and no official endorsement should be inferred.

REFERENCES

Allen, J. R. L., 1970, Studies in fluviatile sediments: a comparison of fining-upward cyclothems with special reference to coarse-member composition and interpretation: Jour. Sed. Pet., v. 40, no. 1, p. 298–323.

Armstrong, M., 1992, Positive definiteness is not enough: Math. Geology, v. 24, no. 1, p. 135–143.

Christakos, G., 1984, On the problem of permissible covariance and variogram models: Water Resources Research, v. 20, no. 2, p. 251–265.

Dagan, G., 1989, Flow and transport in porous formations: Springer-Verlag, New York, 465 p.

Deutsch, C. V., and Journel, A. G., 1992, GSLIB: Geostatistical Software Library: Oxford Univ. Press, New York, 340 p.

Goovaerts, P., 1994a, Comparison of CoIK, IK, mIK performances for modeling conditional probabilities of categorical variables, *in* Dimitrakopoulos, R., ed., Geostatistics for the next century: Kluwer Academic Publ., Dordrecht, p. 18–29.

Goovaerts, P., 1994b, Comparative performance of indicator algorithms for modeling conditional probability distribution functions: Math. Geology, v. 26, no. 3, p. 389–411.

Horn, R. A., and Johnson, C. R., 1991, Topics in matrix analysis: Cambridge Univ. Press, New York, 607 p.

Journel, A. G., 1983, Nonparametric estimation of spatial distributions: Math. Geology, v. 15, no. 3, p. 445–468.

Journel, A. G., 1993, Modeling uncertainty: some conceptual thoughts, *in* Dimitrakopoulos, R., ed., Geostatistics for the next century: Kluwer Academic Publ., Dordrecht, p. 30–43.

Journel, A. G., and Alabert, F., 1989, Non-Gaussian data expansion in the earth sciences: Terra Nova, v. 1, no. 2, p. 123–134.

Journel, A. G., and Huijbregts, 1978, Mining geostatistics: Academic Press, San Diego, California, 600 p.

Journel, A. G., and Posa, D., 1990, Characteristic behavior and order relations for indicator variograms: Math. Geology, v. 22, no. 8, p. 1011–1025.

Leeder, M. R., 1982, Sedimentology: George Allen & Unwin Ltd, London, 344 p.

Miall, A. D., 1973, Markov chain analysis applied to an ancient alluvial plain succession: Sedimentology, v. 20, no. 3, p. 347–364.

Myers, D. E., 1982, Matrix formulation of co-kriging: Math. Geology, v. 14, no. 3, p. 249–257.

Pawlowsky, V., 1989, Cokriging of regionalized compositions: Math. Geology, v. 21, no. 5, p. 513–521.

Pawlowsky, V., Olea, R. A., and Davis, J. C., 1995, Estimation of regionalized compositions: a comparison of three methods: Math. Geology, v. 27, no. 1, p. 105–127.

Press, W. H., Teukolsky, S. A., Vetterling, W. T., and Flannery, B. P., 1992, Numerical recipes in FORTRAN: Cambridge Univ. Press, New York, 963 p.
Ross, S., 1988, A first course in probability: Macmillan Publ. Co., New York, 420 p.
Solow, A. R., 1986, Mapping by simple indicator kriging: Math. Geology, v. 18, no. 3, p. 335–352.

APPENDIX A. DERIVATION OF THE RELATIONSHIP BETWEEN MEAN LENGTH AND THE TRANSITION PROBABILITY

Figure A.1 shows a hypothetical one-dimensional "column" of data for two categories, say k = black and k' = white. (For the general situation of $m = 1, \ldots, K \geq 3$ mutually exclusive categories, k' could represent the union of several categories $m \neq k$). Assume the column is oriented in the direction ϕ and sampled on a regular interval h_ϕ at locations ih_ϕ for $i = 0, \ldots, N(h_\phi)$. Let $T_{kk}(h_\phi)$ and $T_{kk'}(h_\phi)$ denote "transition counts," i.e., the number of transitions from k to k and k to k', respectively, for the lag h_ϕ.

Obviously, the mean length $\bar{l}_{k,\phi}$ of an embedded occurrence of k in the direction ϕ, could be calculated from these continuous data by

$$\bar{l}_{k,\phi} = \frac{\text{total length of } k \text{ in direction } \phi}{\text{number of embedded occurrences of } k}$$

For the example given in Figure A.1, the number of embedded occurrences (layers) of k = black is six. If h_ϕ is made sufficiently small (in theory, smaller than the smallest lengths of the embedded occurrences of k and k'), then

$$p_k N(h_\phi) h_\phi = \text{total length of } k \text{ in direction } \phi$$

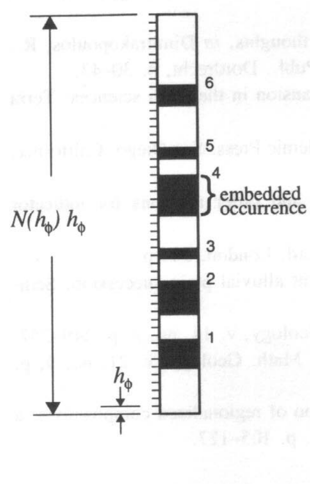

Figure A.1. Diagram showing hypothetical "column" of categorical data.

and

$$T_{kk'}(h_\phi) = \text{number of embedded occurrences of } k$$

so that $\bar{l}_{k,\phi}$ becomes

$$\bar{l}_{k,\phi} = \frac{p_k N(h_\phi) h_\phi}{T_{kk'}(h_\phi)} \tag{A1}$$

where p_k is the proportion of k in the column.

The joint probability $p_{kk'}(h_\phi)$ is computed by

$$p_{kk'}(h_\phi) = \frac{T_{kk'}(h_\phi)}{N(h_\phi)}$$

and consequently by (10) the transition probability $t_{kk'}(h_\phi)$ is computed by

$$t_{kk'}(h_\phi) = \frac{T_{kk'}(h_\phi)}{p_k N(h_\phi)} \tag{A2}$$

Because k and k' are mutually exclusive and exhaustively defined

$$t_{kk}(0) = 1 \tag{A3}$$

and

$$t_{kk'}(h_\phi) = 1 - t_{kk}(h_\phi) \tag{A4}$$

Applying (A3) to (A4) and dividing by h_ϕ, (A2) becomes

$$\frac{t_{kk}(0) - t_{kk}(h_\phi)}{h_\phi} = \frac{T_{kk'}(h_\phi)}{p_k N(h_\phi) h_\phi} \tag{A5}$$

Letting $h_\phi \to 0$ and applying the reciprocal of (A1) to (A5), the slope of $t_{kk}(h_\phi \to 0)$ relates to $\bar{l}_{k,\phi}$, the mean length of category k in the direction ϕ, by

$$-\frac{\partial t_{kk}(h_\phi \to 0)}{\partial h_\phi} = \frac{1}{\bar{l}_{k,\phi}}$$

APPENDIX B. DERIVATION OF THE ORDINARY INDICATOR COKRIGING EQUATIONS IN TERMS OF TRANSITION PROBABILITIES

Applying (29), the error variance $e_k^2(\mathbf{x})$ of the ordinary indicator cokriging estimate is

$$e_k^2(\mathbf{x}) = E\left\{\left[\sum_{\alpha=1}^{N} \sum_{m=1}^{K} i_m(\mathbf{x}_\alpha)\lambda_{mk,\alpha} - i_k(\mathbf{x})\right]^2\right\} \tag{B1}$$

Expanding (B1) becomes

$$e_k^2(\mathbf{x}) = E\left\{ [i_k(\mathbf{x})]^2 + \sum_{\alpha=1}^{N} \sum_{m=1}^{K} \sum_{\beta=1}^{N} \sum_{l=1}^{K} i_m(\mathbf{x}_\alpha)\, \lambda_{ml,\alpha} i_l(\mathbf{x}_\beta) \lambda_{lk,\beta} \right.$$

$$\left. - 2 \sum_{\alpha=1}^{N} \sum_{m=1}^{K} i_m(\mathbf{x}_\alpha) \lambda_{mk,\alpha} i_k(\mathbf{x}) \right\} \tag{B2}$$

Applying (7), dividing by p_m, and applying (10) and (34), (B2) becomes

$$e_k^2(\mathbf{x}) = t_{kk}(0) + \sum_{\alpha=1}^{N} \sum_{m=1}^{K} \sum_{\beta=1}^{N} \sum_{l=1}^{K} t_{ml}(\mathbf{x}_\beta - \mathbf{x}_\alpha)\, \lambda_{ml,\alpha}\, \lambda_{lk,\beta}$$

$$- 2 \sum_{\alpha=1}^{N} \sum_{m=1}^{K} t_{mk}(\mathbf{x} - \mathbf{x}_\alpha)\, \lambda_{mk,\alpha} \tag{B3}$$

Differentiating (B3) with respect to each $\lambda_{mk,\alpha}$ and setting each resulting equation equal to zero yields a system of equations for obtaining the weighting coefficients that minimize $e_k^2(\mathbf{x})$

$$0 = \sum_{\beta=1}^{N} \sum_{l=1}^{K} t_{ml}(\mathbf{x}_\beta - \mathbf{x}_\alpha) \lambda_{lk,\alpha} - t_{mk}(\mathbf{x} - \mathbf{x}_\alpha)$$

$$\text{for } m, k = 1, \ldots K;\ \alpha = 1, \ldots, N \tag{B4}$$

also represented in matrix form as

$$\begin{bmatrix} t_{ml}(\mathbf{x}_1 - \mathbf{x}_1) & \cdots & t_{ml}(\mathbf{x}_N - \mathbf{x}_1) \\ \vdots & \ddots & \vdots \\ t_{ml}(\mathbf{x}_1 - \mathbf{x}_N) & \cdots & t_{ml}(\mathbf{x}_N - \mathbf{x}_N) \end{bmatrix} \begin{bmatrix} \lambda_{lk,1} \\ \vdots \\ \lambda_{lk,N} \end{bmatrix} = \begin{bmatrix} t_{mk}(\mathbf{x} - \mathbf{x}_1) \\ \vdots \\ t_{mk}(\mathbf{x} - \mathbf{x}_N) \end{bmatrix} \tag{B5}$$

The unbiasedness constraint (30) can be imposed using Lagrangian techniques to obtain (35), the ordinary indicator cokriging system of equations formulated in terms of transition probabilities. Equation (33), the simple indicator cokriging system of equations formulated in terms of transition probabilities, can be derived in a similar manner.