Mathematical Geology, Vol. 35, No. 1, January 2003 (© 2003)

Successive Nonparametric Estimation of Conditional Distributions[1]

J. Antonio Vargas-Guzmán[2] and Roussos Dimitrakopoulos[2]

Spatial characterization of non-Gaussian attributes in earth sciences and engineering commonly requires the estimation of their conditional distribution. The indicator and probability kriging approaches of current nonparametric geostatistics provide approximations for estimating conditional distributions. They do not, however, provide results similar to those in the cumbersome implementation of simultaneous cokriging of indicators. This paper presents a new formulation termed successive cokriging of indicators that avoids the classic simultaneous solution and related computational problems, while obtaining equivalent results to the impractical simultaneous solution of cokriging of indicators. A successive minimization of the estimation variance of probability estimates is performed, as additional data are successively included into the estimation process. In addition, the approach leads to an efficient nonparametric simulation algorithm for non-Gaussian random functions based on residual probabilities.

KEY WORDS: Non-Gaussian random functions, nonparametric estimation, conditional covariance, cokriging of indicators, indicator simulation.

INTRODUCTION

In geosciences, the modeling of uncertainty and risk in the spatial distribution of non-Gaussian attributes frequently requires the estimation of local conditional distribution functions. Journel (1983) introduced indicator random functions and indicator kriging (IK) for the nonparametric estimation of local conditional distributions of ergodic, non-Gaussian stationary random functions. IK has been widely used in many areas of study including forecasting of recoverable ore reserves (Fytas, Chaouai, and Lavigne, 1990; Lemmer, 1984), risk assessment in polluted soils (Goovaerts, Webster, and Dubois, 1997), modeling of geology (Dimitrakopoulos and Dagbert, 1993; Soares, 1992), mapping (Solow, 1986), characterization of petroleum reservoirs (Data-Gupta, Xue, and Lee, 1999; Hohn

[1]Received 7 December 2001; accepted 5 November 2002.

[2]WH Bryan Mining Geology Research Centre, The University of Queensland, Brisbane, Queensland 4072, Australia; e-mail: roussos@uq.edu.au

and McDowell, 1994), characterization and risk assessment in contaminated sites (Christakos and Hristopoulos, 1998), and analysis of earthquake motions (Carr and Glass, 1985). A related major area of interest and application of nonparametric estimation of conditional distributions is the stochastic simulation of non-Gaussian random functions using, for example, sequential indicator simulation (Journel and Alabert, 1990) or probability field simulation (Srivastava, 1992).

The indicator transform of a stationary and ergodic random function $Z(x)$ in R^n is

$$I(x) = \left\{ \begin{array}{ll} 1 & \text{if } Z(x) \leq z_c \\ 0 & \text{if not} \end{array} \right\} \tag{1}$$

where z_c is a cutoff and x a spatial location in a n-dimensional physical space. In practice, a number of q cutoffs are used; then, the scalar random function $Z(x)$ is transformed into an indicator vector random function as

$$\mathbf{I}(x) = \lfloor I_1(x), \ldots I_t(x), I_u(x), \ldots I_q(x) \rfloor \tag{2}$$

The cross-covariances between the indicator random functions are

$$c_{I_t I_u}(h) = E\{I_t(x+h) I_u(x)\} - E\,[I_t]\,E\,[I_u] \tag{3}$$

where x and $x + h$ are spatial locations, t and u correspond to two categories or cutoffs, and $E[I_t]$ and $E[I_u]$ are the means independent of location. The expected values in Eq. (3) are related to the joint cumulative distribution function $F(Z_t(x + h), Z_u(x))$ and the marginal distributions $F(Z_t)$ and $F(Z_u)$ as

$$c_{I_t I_u}(h) = F(Z_t(x+h), Z_u(x)) - F(Z_t)F(Z_u) \tag{4}$$

IK provides a point estimate $\hat{F}[Z_t(x)]$ of a conditional cumulative distribution function (ccdf) at location x using a single category or cutoff. If $t = u$, in Eq. (4), the covariance for IK is obtained. The IK system does not guarantee that probability estimates satisfy $0 \leq \hat{F}[Z(x)] \leq 1$ or may not follow the order relations of a ccdf. That is, for any two cutoffs $z_1 < z_2$ the estimated probabilities at a location should satisfy $\hat{F}(z_1) < \hat{F}(z_2)$ (Journel and Posa, 1990).

Sullivan (1984) introduced probability kriging (PK) where $Z(x)$ is replaced by a global ccdf such that $U(x) = F(Z(x))$. Then, $U(x)$ is cokriged together with the indicator I(x) of a single cutoff to estimate probabilities. Sullivan's major justification for PK is that the cokriging estimation generates more accurate ccdf's, because of the lower estimation variances when compared to kriging each individual I(x) on its own. Furthermore, Sullivan (1984) suggests that PK is a next

best alternative to the computationally intensive cokriging of $\mathbf{I}(x)$. Although PK was formulated as a practical improvement to IK, the meaning of the estimates becomes complex as the joint distribution between $U(x)$ and $I(x)$ becomes harder to interpret and order relation violations are not eliminated.

It is logical to consider that the inclusion of more information and all cutoffs simultaneously in the estimation of a ccdf should provide more consistent estimates. Suro Perez and Journel (1991) report that the cross-covariances between indicators are not negligible. Chiles and Delfiner (1999) comment that the implementation of cokriging of indicators remains difficult because of the general unavailability of consistent theoretical covariance models for a set of indicators, and the substantial numerical difficulties in solving simultaneous large systems of equations when cokriging indicators. They attribute the above for the practical inferiority of the theoretically better cokriging of indicators, as shown in Goovaerts (1994), when compared to IK. Goovaerts (1997) identifies two major problems in the cokriging of indicators. First, the practical difficulty in the inference of cross-variogram models between indicators, an issue where the contribution by Yao and Journel (1998) may be relevant. Second, the inversion of very large cokriging matrices at each location x in space, where a ccdf is being estimated, is well known for generating numerical instabilities. In addition, one can identify that the inversion of cokriging matrices is linked to scaling problems that may be in part the cause for the inferior results of cokriging of indicators reported in Goovaerts (1997). A notable alternative to avoid scaling problems is the definition of class interval membership indicators in Carle and Fogg (1996), who analyze them with Markovian transition probabilities and formulate an alternative solution that also resorts to the simultaneous solution of cokriging of indicators.

In this paper, the impractical inversion of very large cokriging matrices associated with the classic "simultaneous" solution of the cokriging system is avoided. A new formulation of the complete cokriging of indicators is presented on the basis of a so-called successive approach that allows the use of data in successive steps. The foundation of the approach is based upon conditional covariances of residual probabilities, which are introduced in detail. The successive minimization of the estimation variance for each indicator leads to the development of successive cokriging of indicators. As shown in the subsequent sections, the use of additional information reduces the kriging variance for the estimated probabilities and minimizes the chances of estimated ccdf's outside certain confidence interval. This also means that order relation problems typical in IK formulations are minimized while the succesive use of data allows detection of any data that cause order relation problems. Last but not least, the new formulation of cokriging of indicators leads to a nonparametric stochastic simulation approach that appears promising for the efficient generation of large-size simulations of spatial phenomena.

PROBABILITY AND INDICATORS REVISITED

Indicator Kriging Weights and Probability of Events

Uncertainty analysis in earth sciences phenomena is commonly examined using probabilities of events occurring. An indicator random function $I(x)$ is made of a spatially continuous set of correlated binary random variables (e.g., Cox and Snell, 1989). Recall that the binomial and therefore the Gaussian distributions are elementally made by Bernoulli trials. The function $I(x)$ is stationary and ergodic and has an indicator covariance function $c_I(h)$. Define a datum at location x_i, and a conditional probability to be estimated at location x_k. For a constant cutoff, the mean probability $\bar{F}(z_c)$ independent of location is estimated as a quadratic equation for $h = 0$ from Eq. (4) (Journel, 1983). A residual of the indicator $R(x_i)$ is

$$R(x_i) = I(x_i) - \bar{F}(z_c) \tag{5}$$

When using simple indicator kriging and one sample location, the estimated probability of $I(x_k) = 1$ is

$$\hat{F}(x_k) = \lambda_{ik} R(x_i) + \bar{F}(z_c) \tag{6}$$

where one kriging weight from classic indicator kriging is the ratio of an indicator covariance and the indicator variance. Using the events A is $I(x_i) = 1$ and B is $I(x_k) = 1$, and Eq. (4), where the cumulative distribution is equivalent to the probability of events, this is

$$\lambda_{ik} = \frac{P(A \cap B)(1 - P(A)P(B))}{P(A)(1 - P(A))} \tag{7}$$

To elaborate on Eq. (7) with respect to positive weights, consider the case where both events A and B are low correlated, then the product $P(A)P(B)$ is close to $P(A \cap B)$ and then Eq. (7) yields

$$\lambda_{ik} \approx \frac{P(A \cap B)P(A \cap B)^c}{P(A)P(A^c)} \tag{8}$$

where, the superscript c means the complementary event.

$$\frac{P(A \cap B)^c}{P(A^c)} = \frac{P(A^c \cap B^c)}{P(A^c)} + \frac{P(A \cap B^c)}{P(A^c)} + \frac{P(A^c \cap B)}{P(A^c)} \tag{9}$$

Because of stationarity, the mean probability of an event is $P(A) = P(B)$ and $P(A^c) = P(B^c)$, then

$$\frac{P(A \cap B)^c}{P(A^c)} = P(B^c \mid A^c) + P(A \mid B^c) + P(B \mid A^c) \tag{10}$$

and

$$\lambda_{ik} \approx P(B \mid A)[P(B^c \mid A^c) + P(A \mid B^c) + P(B \mid A^c)] \tag{11}$$

Thus, the kriging weight for this case is related to the product of the conditional probability that the estimate is in the same category as the datum, and the sum of conditional probabilities involving complementary events. The case where $x_i = x_k$ leads to $\lambda_{ik} \approx 1$ because the two last terms in Eq. (11) vanish. If a new datum at point x_m is added to the estimation, conditional covariances will be required as explained in the next section.

Conditional Indicator Covariances From Probabilities

As an example of conditional indicator covariances, consider a single available indicator datum at location x_i is used to estimate x_k. The first step kriging estimator is a prior estimate in Eq. (6) and a weight is computed as the ratio $\lambda_{ik}^{[1]} = c_{ii}^{-1} c_{ik}$. A sampled datum becomes available at x_k. The true indicator value at x_k minus the previous estimate provides a residual probability $R(x_k) = I(x_k) - \hat{F}(x_k)$. Note that an analogous way to probabilities explained in Bartlett (1945), $R(x_i)$ may be positive or negative depending on the true values of $I(x_k)$. True indicator values will be sampled at additional locations x_m and x_r, and they will be known after their estimation. The datum at x_i is used to estimate locations x_k, x_m, x_r, and the rest of the domain. The indicator conditional covariance between residual random variables for a pair of locations x_k and x_m is

$$\xi_{km} = c_{km} - c_{ki} c_{ii}^{-1} c_{im} \tag{12}$$

where the terms are the covariances between pairs of random variables. If $k = m$, the zero lag distance conditional variance ξ_{kk} is

$$\xi_{kk} = c_{kk} - c_{ki} c_{ii}^{-1} c_{ik} \tag{13}$$

Substituting Eq. (3) into (13) yields

$$\xi_{kk} = F(Z)(1 - F(Z)) - [F(Z_i, Z_k) - F(Z_i)F(Z_k)]^2 F(Z)(1 - F(Z))^{-1} \tag{14}$$

where $F(Z)$ is the probability $P(A)$ of an event A and $F(Z_i, Z_k)$ is a probability of a joint event $P(A \cap B)$. Then,

$$\xi_{kk} = P(A)P(A^c) - [P(A \cap B) - P(A)P(B)]^2[P(B)P(B^c)]^{-1} \quad (15)$$

From Eq. (5) the kriging weight for the first step is identified and gives

$$\xi_{kk} = P(A)P(A^c) - \lambda_{ij}^{[1]}[P(A \cap B) - P(A)P(B)]^c \quad (16)$$

Using Eq. (14) in a more general way, the conditional covariance ξ_{km} for $k \neq m$ is

$$\xi_{km} = [F(Z_k, Z_m) - F(Z_k)F(Z_m)] - \lambda_{ik}^{[1]}[F(Z_i, Z_m) - F(Z_i)F(Z_m)] \quad (17)$$

calling C to the event of location $I(x_m) = 1$, Eq. (17) is

$$\xi_{km} = [P(B \cap C) - P(B)\,P(C)] - \lambda_{ik}^{[1]}\,[P(C \cap A) - P(A)\,P(C)] \quad (18)$$

Since the mean probability is constant (i.e., second-order stationary probabilities for a constant cutoff), this yields

$$\xi_{km} = P(B \cap C) - \lambda_{ik}^{[1]}[P(C \cap A)] + [P(A)]^2\left(\lambda_{ik}^{[1]} - 1\right) \quad (19)$$

Note that this relation allows computation of the conditional covariances from probability measurements in a framework of second-order moments. We have explained the successive conditional covariances for residuals in the context of indicators or Bernoulli process and residual probabilities. The above validates the use of successive estimation methods for probability estimates.

SUCCESSIVE INDICATOR ESTIMATION

Estimation With Residual Probabilities

The location x_k was estimated with the prior data. Later, the indicator $I(x_k)$ becomes posterior datum because it is exactly known from sampling, then a residual probability $R_1(x_k)$ may be computed as

$$R_1(x_k) = I(x_k) - \left(\lambda_{ik}^{[1]}R(x_i) + \bar{F}(z_c)\right) \quad (20)$$

where the second term comes from prior estimation as in Eq. (5). This allows for a second step conditional kriging computed with the indicator conditional

covariances of Eqs. (14) and (17) as follows:

$$\lambda_{km}^{[2]} = \xi_{kk}^{-1}\xi_{km} \tag{21}$$

This is used to improve the first estimated probability at x_m, as

$$\hat{F}_2(x_m) = \left(\lambda_{im}^{[1]} R(x_i) + \bar{F}(z_c)\right) + \lambda_{km}^{[2]} R_1(x_k) \tag{22}$$

where $\lambda_{im}^{[1]}$ is computed in the same way as $\lambda_{ik}^{[1]}$. In a next step, $I(x_m)$ becomes known from sampling and a new updated residual probability is

$$R_2(x_m) = I(x_m) - \hat{F}_2(x_m) \tag{23}$$

The probability has also been estimated at point x_r

$$\hat{F}_2(x_r) = \left(\lambda_{ir}^{[1]} R(x_i) + \bar{F}(z_c)\right) + \lambda_{kr}^{[2]} R_1(x_k) \tag{24}$$

This last estimate can be updated by using the residual probability data at x_m as

$$\hat{F}_3(x_r) = \hat{F}_2(x_r) + \lambda_{mr}^{[3]}\left[I(x_m) - \left[\left(\lambda_{im}^{[1]} R(x_i) + \bar{F}(z_c)\right) + \lambda_{km}^{[2]} R_1(x_k)\right]\right] \tag{25}$$

where $\lambda_{mr}^{[3]}$ is a residual conditional kriging weight from

$$\lambda_{mr}^{[3]} = \left[\xi_{mm} - \xi_{mk}\xi_{kk}^{-1}\xi_{km}\right]^{-1}\left[\xi_{mr} - \xi_{mk}\xi_{kk}^{-1}\xi_{kr}\right] \tag{26}$$

The updated indicator conditional covariance in Eq. (26) can be written following Eq. (17) in terms of probabilities, however, that is not necessary here.

Successive Indicator Kriging

The above analysis may proceed for a number of groups of data locations but with the purpose of estimating kriging weights. The approach is for estimation of ccdf's using indicator random functions. The expected value of an estimated ccdf gives an estimate for the original attribute. The expected value of the second moment may allow computation of an estimation variance for the original attribute. The general idea is that for a single cutoff indicator the estimates are conditional probabilities that have an estimation variance (i.e., kriging variance), which drops as more data are used in the estimation. Successive kriging allows the use of more data in a sequential fashion but with numerical results equal to the simultaneous solution.

Successive indicator kriging considers successive subsets of indicator data that are part of a global indicator data set. These are

$$\mathbf{I}_g = \begin{bmatrix} \mathbf{I}_p \\ \mathbf{I}_s \\ \mathbf{I}_r \end{bmatrix} \tag{27}$$

introducing $\mathbf{\Lambda}$ matrices and $\boldsymbol{\lambda}$ vectors of kriging weights that are to estimate a group or a point from data of another group. Generalizing Eq. (25) a successive estimator for the point location x_o is introduced as

$$\begin{aligned} \hat{F}(x_o) = &\left[\boldsymbol{\lambda}^{[1]}_{p(o)}\right]^T \mathbf{R}(x_p) + \bar{F}(z_c) + \left[\boldsymbol{\lambda}^{[2]}_{s(o)}\right]^T \left\lfloor \mathbf{I}(x_s) - \left(\left[\mathbf{\Lambda}^{[1]}_{ps}\right]^T \mathbf{R}(x_p)\right) + \bar{F}(z_c)\right\rfloor \\ &+ \left[\boldsymbol{\lambda}^{[3]}_{r(o)}\right]^T \left[\mathbf{I}(x_r) - \left[\left[\mathbf{\Lambda}^{[1]}_{pr}\right]^T \mathbf{R}(x_p) + \left[\mathbf{\Lambda}^{[2]}_{sr}\right]^T \left(I(x_k)\right.\right.\right. \\ &\left.\left.\left. - \left(\left[\mathbf{\Lambda}^{[1]}_{ps}\right]^T \mathbf{R}(x_p)\right) + \bar{F}(z_c)\right)\right]\right] \end{aligned} \tag{28}$$

where $\mathbf{R}(x)$ are residual probability data matrices. Note this equation represents successive steps. If more new data become available, additional steps can be added, as needed. The true conditional probability is $F(x_o)$, and the estimation error ΔF_o is

$$\Delta F_o = F(x_o) - \hat{F}(x_o) \tag{29}$$

The estimation variance of these probability estimates is

$$E(\Delta F_o)^2 = E(F(x_o) - \hat{F}(x_o))^2 \tag{30}$$

Minimizing Eq. (30) by taking derivatives with respect to the weights yields a succession of steps of conditional indicator kriging equations, these are

$$\begin{aligned} \mathbf{c}^{[1]}_{\text{pp}} \mathbf{\Lambda}^{[1]}_{pD} &= \mathbf{c}^{[1]}_{pD} \\ \boldsymbol{\xi}^{[2]}_{\text{ss}} \mathbf{\Lambda}^{[2]}_{sD} &= \boldsymbol{\xi}^{[2]}_{sD} \\ \boldsymbol{\xi}^{[3]}_{\text{rr}} \mathbf{\Lambda}^{[3]}_{rD} &= \boldsymbol{\xi}^{[3]}_{rD} \end{aligned} \tag{31}$$

where D is the remaining domain including the data not utilized yet and any location x_o. In general, further conditional kriging steps are added, for example $\boldsymbol{\xi}^{[4]}_{(\text{r}+1)(\text{r}+1)} \mathbf{\Lambda}^{[4]}_{(r+1)D} = \boldsymbol{\xi}^{[4]}_{(r+1)D}$. Any conditional residual indicator covariance is a function of the previous step indicator conditional covariances. For example,

this is

$$\xi^{[4]}_{(r+1)D} = \xi^{[3]}_{(r+1)D} - \xi^{[3]}_{Dr}\left[\xi^{[3]}_{rr}\right]^{-1}\xi^{[3]}_{r(r+1)} \tag{32}$$

The approach described here can be applied to a large number of subsets of indicator data. The use of a successive approach allows the use of as much data as available for the estimation of the probabilities. Finally, the simultaneous solutions of indicator kriging with all the data will give the same numerical result as the successive solution.

Cross-Covariances of Indicators

As explained in the Introduction (Chiles and Delfiner, 1999; Sullivan, 1984; Suro Perez and Journel, 1991), the lower estimation variance should, at least in theory, occur when indicator data from all cutoffs are utilized (i.e., cokriging of indicators). Arguably, cokriging of indicators is impractical if a simultaneous solution is sought despite its theoretical advantages. A successive solution can provide a suitable practical alternative.

An important observation is that zero lag distance indicator covariances (variances) have a scaling effect on the cross-covariances and they are exactly predictable for given percentiles, recall the variance is $P(1-P)$. It is well known that the maximum variance is 0.25 and corresponds to the median indicator, $P = 0.5$ (Journel, 1983). Cross-covariances for indicators involving extreme low and high cutoffs will be low in absolute value due to scaling and significant cross-correlation $\rho_{tu}(h)$ may be hidden in the cross-covariance. This is

$$c_{tu}(h) = \rho_{tu}(h)\sqrt{c_{tt}(0)}\sqrt{c_{uu}(0)} \tag{33}$$

Indicators with different variances might not be standardized because the probabilistic meaning of estimates may be distorted. Thus, simultaneous cokriging of indicators may encounter scaling difficulties.

The above analysis leads to an analysis of cross-covariances between indicators for zero lag distance. The cross-covariance between indicators for two cutoffs $z_t < z_u$ is defined in Eq. (4). If A is the event $I_t(x) = 1$ and B is the event $I_u(x+h) = 1$ for $h = 0$, this is if $A \subset B$ and $P(A \cap B) = P(A)$ following definition in Eq. (1). This is if B occurs A also occurs but not vice versa. For $A \subset B$ Eq. (4) gives

$$c_{I_t I_u}(0) \cong F(Z_t(x)) - F\left(Z_t(x)\right) F\left(Z_u(x)\right) \tag{34}$$

In general for $h \neq 0$, $P(A \cap B) = P(A)P(B \mid A)$ is controlled by the correlation $\rho_{tu}(h)$.

Another alternative for cokriging is to formulate indicators that are disjoint sets at shorter and zero lag distance. In this approach, a location only belongs to one class or category. The indicator is defined between intervals $(z_t - z_u)$ as

$$I_{(t-u)}(x,(z_t - z_u)) = \left\{ \begin{array}{ll} 1, & \text{if } z_t \leq z(x) \leq z_u \quad \forall z_u > z_t \\ 0, & \text{if not} \end{array} \right\} \tag{35}$$

This is a membership indicator for a class interval. See also categorical indicators implicit in the computation of transitional probabilities in geology (e.g., Carle and Fogg, 1996). The cross-covariances in this case are also computed following Eqs. (2) and (3). If constant proportions P for class intervals are chosen, the variances $P(1-P)$ are very similar allowing for a cokriging using all class intervals. Since $P(A \cap B) = \phi$ is the empty set for $h = 0$ following Eq. (4) the cross-covariance becomes

$$c_{I_{t-u} I_{l-m}}(0) = -F(z_{t-u})\, F(z_{l-m}) \tag{36}$$

If disjoint sets are used, the estimates are probabilities of falling within an interval. These are the areas within an interval below the conditional probability density function. Adding these areas for each cutoff reproduces the cumulative probabilities.

Successive Cokriging of Indicators

Having the cross-covariances, the successive cokriging of indicators approach proceeds as an extension of successive indicator kriging. The conditional indicator covariance terms are substituted by multivariate matrices of conditional indicator covariances and cross-covariances. Successive indicator cokriging is implemented when using conditional indicator cross-covariances defined below. The constant mean probabilities are handled like in successive indicator kriging. Using matrix notation, at some step r a successive indicator cokriging step is

$$\bar{\xi}^r_{\mathbf{R}^r \mathbf{R}^r} \Lambda^r_{R^r o} = \bar{\xi}^r_{\mathbf{R}^r o} \tag{37}$$

where $\bar{\xi}^r_{\mathbf{R}^r \mathbf{R}^r}$ and $\bar{\xi}^r_{R^r o}$ are the multivariate conditional indicator covariances for all categories or cutoffs, for the $\mathbf{R}^r$ data residual probabilities matrix at several locations, and the data versus estimated locations respectively. The bars on top of the matrices are to explain they are multivariate. The matrix of weights is also multivariate $\Lambda^r_{\mathbf{R}^r o}$. A sequence of conditional cokriging of indicators is solved following the procedure of including a new set of data at each step. Indicator multivariate conditional covariances are needed at each step. For example, if a matrix of conditional covariances between two indicators $I_V(x) = V_x$ and $I_U(y) = U_y$ for two

classes is needed for locations x and y, this is computed in matrix form as follows:

$$\bar{\xi}^{r+1}_{I_x U_y} = \begin{bmatrix} \xi^r_{V_x V_y} & \xi^r_{V_x U_y} \\ \xi^r_{U_x V_y} & \xi^r_{U_x U_y} \end{bmatrix} - \begin{bmatrix} \Lambda^r_{V_d V_y} & \Lambda^r_{V_d U_y} \\ \Lambda^r_{U_d V_y} & \Lambda^r_{U_d U_y} \end{bmatrix}^{\mathrm{T}} \begin{bmatrix} \xi^r_{V_d V_x} & \xi^r_{V_d U_x} \\ \xi^r_{U_d V_x} & \xi^r_{U_d U_x} \end{bmatrix} \tag{38}$$

where V_d and U_d represent indicator data and Λ matrices are the cokriging weights from a previous r step cokriging (Eq. 37). In general, at each step indicator conditional covariances and cross-covariances are updated from previous step as Eq. (38).

Another simpler version of successive cokriging of indicators is generated by estimating probabilities from data of a single cutoff or class first and then improve the estimate with the second cutoff with indicator conditional cross-covariances. Data from a third and so forth cutoffs or classes are included with conditional covariances, updated in similar fashion, as explained. This allows to choose only the significantly correlated classes in the estimation. Successive indicator cokriging as described previously may allow the use of more information than the classic indicator kriging. No large systems of equations are solved and the zero lag cross-covariances and variances are fully computable.

Using a Local Neighborhood With Conditional Covariances

The computation of conditional covariances does not require the use of local neighborhoods. In practice, however, computer time is reduced if updating conditional covariances is made within local neighborhoods. This is further justified considering that a covariance model has a range. For example, if a spherical covariance model is used, the loss of information due to the use of a search neighborhood with a radius slightly larger than the range becomes zero. For other asymptotic models, the loss may be negligible if the neighborhood is sufficiently large. Note that the estimates will be unique and identical to ones generated from a global simultaneous solution. Since the order or sequence in which data are introduced into the successive estimator does not affect the final estimate as implicitly stated, the screen effect has not really theoretical importance for successive kriging. This means the successive solution may indiscriminately incorporate first samples located far or near the estimation point. The final numerical estimate and its kriging variance are the same, although successive weights for each sample are different. This is consistent with the corresponding development for continuous attributes (Vargas-Guzmán and Yeh, 1999; Vargas-Guzmán and Dimitrakopoulos, 2002).

CONDITIONAL SUCCESSIVE COINDICATOR SIMULATION

The successive cokriging of indicators introduced contributes a new tool for estimating ccdf's in the context of indicator simulations. The successive cokriging

of indicators at a single location will produce the numerical estimate of the non-Gaussian ccdf for the continuous attribute $Z(x)$. A Monte Carlo drawing from such numerical ccdf can generate a realization of $Z(x)$. The new value can be transformed into a vector of indicators that are used to compute residual probabilities for updating the remaining ccdf's. In general, the simulation method proceeds step by step as follows:

a) A conditioning data set Ω_d is transformed into indicators following one alternative, cutoffs or class intervals Eq. (1) or (35).
b) The multivariate matrix of indicator covariances functions is computed and modeled from the data. Note that zero lag distance values are from Eqs. (33) to (36).
c) The set of indicator data locations Ω_d is partitioned into a number of small subsets $\{\Omega_p, \Omega_s, \ldots, \Omega_t\}$.
d) Successive indicator cokriging is performed with the indicators for each subset in $\{\Omega_p, \Omega_s, \ldots, \Omega_t\}$ to estimate the entire domain. Conditional covariances are updated at each step. If no more data are available, the simulation is performed at the points of the remaining domain set Ω_g - (see Eq. (37)).
e) Update conditional covariances between the point and the domain - (see Eq. (38)).
f) Following a random path, a location is randomly chosen and a value is drawn at the point from the estimated ccdf and stored as a point realization.
g) The drawn value is converted to indicator. Then, a residual $\Delta\mathbf{z}_g$ for each cutoff or each class is computed between the new value and the previous estimate.
h) Residuals $\Delta\mathbf{z}_j$ are used to update estimated conditional probabilities in the remaining domain with successive indicator cokriging (go to step (f)).

CONCLUSIONS

A new approach for the nonparametric estimation of conditional distributions, termed *successive cokriging of indicators*, was introduced. The approach is based on the successive minimization of the cokriging indicator estimation variance. A key contribution of the approach is that it provides an alternative solution to estimation of multiple indicators that avoids the problem of inversion of large matrices. The successive formulation allows for the use of several cutoffs, class intervals, or categories which are known to have large covariance matrices that produce instabilities in large simultaneous systems for cokriging of indicators. By allowing the partitioned use of large data sets on a successive basis, the proposed successive estimation avoids the simultaneous solution by using conditional indicator covariances for residual probabilities. At each step, residual ccdf's are estimated

with residual data that represent the difference between true indicators and previous probability estimates. Conceptually, the approach considers that previous estimated locations have been sampled for updating.

The successive cokriging of indicators introduces the partitioned use of data sets that facilitates the selection of data, which may affect the quality of the probability estimates generated. In this way, order relation problems can be detected and potentially avoided. The approach also facilitates the incorporation of multiple cutoffs or categories in a sequential fashion, thus improving the quality of the estimates by reducing the estimation variance or the chance of generating estimates outside acceptable ranges of values. The cases of cutoffs as in Eq. (1) or disjoint categories as in Eq. (35) are included in the proposed approach and have the advantage that only significantly correlated categories can be used if desired.

It should be noted that if a new set of data becomes available after the estimation of any ccdf, updating of the ccdf with successive cokriging of indicators is performed without repeating computations. Updating is important for practical and theoretical purposes in both the estimation and simulation of non-Gaussian random functions. In the case of sequential-type simulations, the updating process continues as new values are drawn at previous locations.

Probabilistic analysis of conditional covariances for residuals of indicators develops a concept of residual probabilities which provides the theoretical framework for the successive solution of indicator kriging and cokriging of indicators. Lower estimation variance relates to lower dispersion in probability estimates, which should contribute to substantial reduction of order relation problems as extreme estimates are considered less probable. Expressing the conditional covariances of events in terms of conditional probabilities enables the updating of the probability of locations being above the cutoff based on either prior or analog spatial information. Future research could consider developing a consistent way of generating covariance models for indicators.

ACKNOWLEDGMENTS

Funding was provided by the Sustainable Minerals Institute, The University of Queensland, Australia, Grant No. RERP12, to R. Dimitrakopoulos.

REFERENCES

Bartlett, M. S., 1945, Negative probability: Proc. Cambridge Philos. Soc., v. 41, p. 71–73.

Carle, S., and Fogg, G. E., 1996, Transition probability-based indicator geostatistics: Math. Geol., v. 28, no. 4, p. 453–476.

Carr, J., and Glass, C. E., 1985, Treatment of earthquake ground motion using regionalized variables: Math. Geol., v. 17, no. 3, p. 221–241.

Chiles, J. P., and Delfiner, P., 1999, Geostatistics: Modeling spatial uncertainty: Wiley, New York, 695 p.

Christakos, G., and Hristopulos, D. T., 1998, Stochastic indicator analysis of contaminated sites: J. Appl. Probab., v. 34, no. 4, p. 988–1008.

Cox, D. R., and Snell, E. J., 1989, Analysis of binary data: Chapman and Hall, London, 236 p.

Data-Gupta, A., Xue, G., and Lee, S. H., 1999, Nonparametric transformations for data correlation and integration, theory and practice, *in* Schatzinger, R. A., and Jordan, J. F., eds., Reservoir characterization recent advances, Am. Assoc. Petroleum Geologists, Memoir 71, Tulsa, OK, p. 381–396.

Dimitrakopoulos, R., and Dagbert, M., 1993, Sequential modeling of relative indicator variables: Dealing with multiple lithology types, *in* A. Soares, ed., Geostatistics Tróia '92: Kluwer, Dordrecht, The Netherlands, p. 413–424.

Fytas, K., Chaouai, N.-E., and Lavigne, M., 1990, Gold deposits estimation using indicator kriging: CIM Bull., v. 80, p. 77–83.

Goovaerts, P., 1994, Comparison of CIK, IK and MIK performances for modeling conditional probabilities of categorical variables, *in* Dimitrakopoulos, R., ed., Geostatistics for the next century: Kluwer, Dordrecht, The Netherlands, p. 18–29.

Goovaerts, P., 1997, Geostatistics for natural resources evaluation: Oxford University Press, New York, 483 p.

Goovaerts, P., Webster, R., and Dubois, J. P., 1997, Assessing the risk of soil contamination in the Swiss Jura using indicator geostatistics: Environ. Ecol. Stat., v. 4, p. 31–48.

Hohn, E. M., and McDowell, R. R., 1994, Geostatistical analysis of oil production and potential using indicator kriging, *in* Yarus, J. M., and Chambers, R. L., eds., Stochastic modeling and geostatistics, Am. Assoc. Petroleum Geologists, Computer Applications in Geology, No. 3, Tulsa, OK.

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

Journel, A. G., and Alabert, F. G., 1990, A new method for reservoir mapping: J. Pet. Technol., Feb., 1990, p. 212–218.

Journel, A. G., and Posa, D., 1990, Characteristic behavior and order relations for indicator variograms: Math. Geol., v. 30, no. 8, p. 1011–1025.

Lemmer, I. C., 1984, Estimating local recoverable reserves via indicator kriging, *in* Verly, G., David, M., Journel, A. G., and Marechal, A., eds., Geostatistics for natural resources characterization, Part 1: Reidel, Dordrecht, The Netherlands, p. 365–384.

Soares A., 1992, Geostatistical estimation of multiphase structures: Math. Geol., v. 24, no. 2, p. 149–160.

Solow, A. R., 1986, Mapping by simple indicator kriging: Math. Geol., v. 18, no. 3, p. 335–352.

Srivastava, R. M., 1992, Reservoir characterization with probability field simulation, Society of Petroleum Engineers, Richardson, TX. paper #24753.

Sullivan, J., 1984, Conditional recovery estimation through probability kriging—Theory and practice, *in*, Verly, G., David, M., Journel, A. G., and Marechal, A., eds., Geostatistics for natural resources characterization, Part 1: Reidel, Dordrecht, The Netherlands, p. 365–384.

Suro Pérez, V., and Journel, A. G., 1991, Indicator principal component kriging: Math. Geol., v. 23, no. 5, p. 759–788.

Vargas-Guzmán, J. A., and Yeh, T.-C. J., 1999, Sequential kriging and cokriging, two powerful geostatistical approaches: Stochastic Environ. Res. Risk Assess., v. 13, no. 6, p. 416–435.

Vargas-Guzmán, J. A., and Dimitrakopoclos, R., 2002, Conditional Simulation of random fields by successive residuals: Math. Geol., v. 34, no. 5, p. 597–611.

Yao, T., and Journel, A. G., 1998, Automatic modeling of (cross) covariance tables using fast Fourier transform: Math. Geol., v. 30, no. 6, p. 589–616.