PB87-166260

Disjunctive Kriging
3. Cokriging

(U.S.) Robert S. Kerr Environmental
Research Lab., Ada, OK

1986

U.S. Department of Commerce
National Technical Information Service
NTIS

TECHNICAL REPORT DATA *(Please read Instructions on the reverse before complet*)		
1. REPORT NO. EPA/600/J-86/232	2.	3. PB87-166260
4. TITLE AND SUBTITLE DISJUNCTIVE KRIGING 3. Cokriging (Journal Version)		5. REPORT DATE 1986 6. PERFORMING ORGANIZATION CODE
7. AUTHOR(S) S. R. Yates		8. PERFORMING ORGANIZATION REPORT NO.
9. PERFORMING ORGANIZATION NAME AND ADDRESS Robert S. Kerr Environmental Research Laboratory U.S. Environmental Protection Agency Post Office Box 1198 Ada, Oklahoma 74820		10. PROGRAM ELEMENT NO. ABWD1A 11. CONTRACT/GRANT NO. In-House
12. SPONSORING AGENCY NAME AND ADDRESS Robert S. Kerr Environmental Research Lab. - Ada, OK Post Office Box 1198 Ada, Oklahoma 74820		13. TYPE OF REPORT AND PERIOD COVERED Journal Article 14. SPONSORING AGENCY CODE EPA/600/15
15. SUPPLEMENTARY NOTES Published in: WATER RESOURCES RESEARCH, 22(10):1371-1376, September 1986.		

16. ABSTRACT

The disjunctive kriging (DK) method described in the first paper of this series is extended to account for more than one random function. In the derivation contained herein, two random functions are considered, but this is easily generalized to any number. An example is presented using disjunctive cokriging (DCK) where the surface gravimetric moisture content is estimated using the bare soil temperature as an auxiliary random function. The results indicate that the DCK procedure produces a better estimator than ordinary cokriging in terms of reduced variance of errors and exactness of estimation. Also, using DCK, an estimate of the conditional probability that the level of a property is greater than a known cutoff value can be obtained. In general, this conditional probability is better than the DK probability by virtue of the additional information contained in the second, auxiliary random function.

17. KEY WORDS AND DOCUMENT ANALYSIS		
a. DESCRIPTORS	b. IDENTIFIERS/OPEN ENDED TERMS	c. COSATI Field/Group

18. DISTRIBUTION STATEMENT	19. SECURITY CLASS *(This Report)*	21. NO. OF PAGES
RELEASE TO PUBLIC.	UNCLASSIFIED	7
	20. SECURITY CLASS *(This page)* UNCLASSIFIED	22. PRICE

EPA/600/J-86/232
JOURNAL ARTICLE

WATER RESOURCES RESEARCH, VOL. 22, NO. 10, PAGES 1371-1376, SEPTEMBER 1986

Disjunctive Kriging 3. Cokriging

S. R. YATES

R. S. Kerr Environmental Research Laboratory, Ada, Oklahoma

The disjunctive kriging (DK) method described in the first paper of this series is extended to account for more than one random function. In the derivation contained herein, two random functions are considered, but this is easily generalized to any number. An example is presented using disjunctive cokriging (DCK) where the surface gravimetric moisture content is estimated using the bare soil temperature as an auxiliary random function. The results indicate that the DCK procedure produces a better estimator than ordinary cokriging in terms of reduced variance of errors and exactness of estimation. Also, using DCK, an estimate of the conditional probability that the level of a property is greater than a known cutoff value can be obtained. In general, this conditional probability is better than the DK probability by virtue of the additional information contained in the second, auxiliary random function.

INTRODUCTION

In a recent paper [*Yates et al.*, 1986*a*] the theoretical basis for the disjunctive kriging (DK) method was reviewed. In a following paper [*Yates et al.*, 1986*b*] an example of the DK method was presented using the electrical conductivity data collected by *Al-Sanabani* [1982]. The example and review material assumes the presence of only one random function. At times, however, it may be desirable to utilize an additional variable(s) in the estimation process. This might be necessary to fill in missing information or to increase the accuracy of the estimation. When two or more random functions are used for kriging it is termed cokriging [*Journel and Huijbregts*, 1978; *Vauclin et al.*, 1983; *Carr et al.*, 1985; *Myers*, 1982, 1984].

Few examples of cokriging using linear estimators exist in the soil water literature. Exceptions include *Vauclin et al.* [1983], who used ordinary cokriging (OCK) to estimate the available water content where the sand content was included in the estimation process. S. R. Yates and A. W. Warrick (unpublished manuscript, 1986) used cokriging to estimate the surface gravimetric moisture content (GMC) in a 1-ha field. The auxiliary random functions used along with the GMC were the bare soil surface temperature, which was collected with an infrared thermometer, and percent sand content. Their results indicate that the ordinary cokriging estimator is superior to the ordinary kriging estimator in terms of reduced kriging variance as well as reduced sum of squares error.

The purpose of this paper is to extend the disjunctive kriging method to allow more than one random function to be used in the estimation process. In the first part of the paper the theory will be discussed; this is followed by an example. The example will include a comparison between the disjunctive cokriging (DCK) and OCK estimators in terms of variance of errors and exactness of estimation. The conditional probability distributions will be calculated for DCK and compared to the DK estimator of one random function.

THEORY

Consider two second-order stationary random functions sampled on point support in two dimensions where a value of the property of primary interest is obtained at p locations and is denoted as $Z(x_{1i})$, $i = 1, 2, \ldots, p$. Further, an auxiliary random function, $V(x)$, is sampled at q locations: $V(x_{2j})$, $j = 1, 2, \ldots, q$. The 1 and 2 in the subscripts for x are used to identify with which random function the x location is associated. The sampling positions of $Z(x)$ need not be the same as $V(x)$, but it is assumed that at sufficiently many locations both $Z(x)$ and $V(x)$ are sampled so that the cross-correlation structure can be determined.

The DK method utilizes transformed variables which are assumed to be uni- and bivariate normally distributed. For DCK, two such transform functions must be defined, one for each random function (i.e., $Z(x)$ and $V(x)$):

$$Z(x_{1i}) = \phi[Y(x_{1i})] = \sum_{k=0}^{\infty} C_k H_k[Y(x_{1i})] \quad (1)$$

$$V(x_{2j}) = \psi[U(x_{2j})] = \sum_{k=0}^{\infty} D_k H_k[U(x_{2j})] \quad (2)$$

where H_k is a Hermite polynomial of order k, and $Y(x_{1i})$ and $U(x_{2j})$ are random variables which designate the transformed data. The random variables $Y(x_{1i})$ and $U(x_{2j})$ have standard normal distributions which are obtained from the transforms, ϕ and ψ, respectively. Although $Y(x_{1i})$ and $U(x_{2j})$ have the same distribution, they are given different symbols in order to delineate from which data set they are derived. The same is true for the transform relationships.

Disjunctive Cokriging Estimator

The disjunctive cokriging estimator $Z_{DCK}{}^*$ is defined as

$$Z_{DCK}{}^*(x_0) = \sum_{i=1}^{m} f_i[Y(x_{1i})] + \sum_{j=1}^{n} h_j[U(x_{2j})] \quad (3)$$

where f_i and h_j are unknown functions to be determined; m and n are the number of data points (nearest neighbors) used in the estimation; and, in general, m need not equal n.

Expanding the unknown functions in a series of Hermite polynomials gives

$$Z_{DCK}{}^*(x_0) = \sum_{k=0}^{\infty} \sum_{i=1}^{m} f_{ik} H_k[Y(x_{1i})] + \sum_{k=0}^{\infty} \sum_{j=1}^{n} h_{jk} H_k[U(x_{2j})] \quad (4)$$

where the f_{ik} and h_{jk} are the coefficients of the Hermite expansion.

Paper number 6W4396.
0043-1397/86/006W-4396$05.00

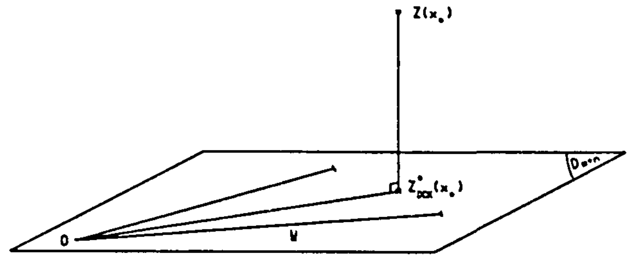

Fig. 1. Schematic of the space D_{m+n}, where $Z_{DCK}{}^*(x_0)$ is the projection of $Z(x_0)$ onto D_{m+n}.

Disjunctive Cokriging Equations

The derivation contained in the next few paragraphs follows the same approach that was used by *Journel and Huijbregts* [1978] in deriving the disjunctive kriging equations for one random variable. The symbolism and many of the definitions that will be used are given in the work by *Yates et al.* [1986*a*].

The first step is to define the vector space D_{m+n}, shown diagrammatically in Figure 1, as the space spanned by the $(m + n)$ single-variable measurable functions $f_i[Y(x_{1i})]$ and $h_j[U(x_{2j})]$. That is, $D_{m+n} = \{f_1[Y(x_{11})] + \cdots + f_m[Y(x_{1m})] + h_1[U(x_{21})] + \cdots + h_n[U(x_{2n})]$. Using the perpendicular projection of $Z(x_0)$ onto D_{m+n} will produce the estimate with the minimum error. Two requirements which assure that the best estimator will be produced are that $Z_{DCK}{}^*$ lies in D_{m+n} and that $Z(x_0) - Z_{DCK}{}^*(x_0)$ must be orthogonal to any vector in D_{m+n} (i.e., W in Figure 1). The latter condition can be written as the scalar product

$$\langle Z(x_0) - Z_{DCK}{}^*(x_0), W\rangle = 0 \tag{5}$$

or [*Journel and Huijbregts*, 1978, p. 569]

$$E[Z(x_0) \mid Y(x_\alpha)] = E[Z_{DCK}{}^*(x_0) \mid Y(x_\alpha)] \tag{6}$$

$$E[Z(x_0) \mid U(x_\beta)] = E[Z_{DCK}{}^*(x_0) \mid U(x_\beta)] \tag{7}$$

where $x_\alpha = x_{1i}$; $\alpha, i = 1, 2, \ldots, m$, and $x_\beta = x_{2j}$; $\beta, j = 1, 2, \ldots, n$. It is evident from (6) and (7) that disjunctive cokriging, like linear cokriging, requires the solution of two simultaneous equations.

The remainder of the solution is straightforward. Incorporating (1) and (3) into both (6) and (7) gives

$$E[\phi[Y(x_0)] \mid Y(x_\alpha)] = \sum_{i=1}^{m} E[f_i[Y(x_{1i})] \mid Y(x_\alpha)] + \sum_{j=1}^{n} E[h_j[U(x_{2j})] \mid Y(x_\alpha)] \tag{8}$$

$$E[\phi[Y(x_0)] \mid U(x_\beta)] = \sum_{i=1}^{m} E[f_i[Y(x_{1i})] \mid U(x_\beta)] + \sum_{j=1}^{n} E[h_j[U(x_{2j})] \mid U(x_\beta)]$$

Incorporating (15) from *Yates et al.* [1986*a*] into (8) and rearranging gives

$$\sum_{k=0}^{\infty} H_k[Y(x_\alpha)]C_k\left\{\rho_{11}{}^k(x_0 - x_\alpha) - \sum_{i=1}^{m} f_{ik}\rho_{11}{}^k(x_{1i} - x_\alpha)/C_k - \sum_{j=1}^{n} h_{jk}\rho_{21}{}^k(x_{2j} - x_\alpha)/C_k\right\} = 0 \tag{9}$$

$$\sum_{k=0}^{\infty} H_k[U(x_\beta)]C_k\left\{\rho_{12}{}^k(x_0 - x_\beta) - \sum_{i=1}^{m} f_{ik}\rho_{12}{}^k(x_{1i} - x_\beta)/C_k - \sum_{j=1}^{n} h_{jk}\rho_{22}{}^k(x_{2j} - x_\beta)/C_k\right\} = 0 \tag{10}$$

In (9) and (10), $\rho_{ii}(x_0 - x_\xi)$ and $\rho_{ij}(x_0 - x_\xi)$ are the correlation and cross-correlation functions for the separation distance $x_0 - x_\xi$.

Since both (9) and (10) must be satisfied for all k the disjunctive cokriging equations can be written as

$$\begin{aligned} \sum_{i=1}^{m} b_{ik}\rho_{11}{}^k(x_{1i}-x_\alpha)+\sum_{j=1}^{n} a_{jk}\rho_{21}{}^k(x_{2j}-x_\alpha)&=\rho_{11}{}^k(x_0-x_\alpha) \\ \sum_{i=1}^{m} b_{ik}\rho_{12}{}^k(x_{1i}-x_\beta)+\sum_{j=1}^{n} a_{jk}\rho_{22}{}^k(x_{2j}-x_\beta)&=\rho_{12}{}^k(x_0-x_\beta) \end{aligned} \tag{11}$$

where a_{jk} and b_{ik} are defined as h_{jk}/C_k and f_{ik}/C_k, respectively. Incorporating these definitions for a_{jk} and b_{ik} into (4) gives

$$Z_{DCK}{}^*(x_0) = \sum_{k=0}^{\infty} C_k\left\{\sum_{i=1}^{m} b_{ik}H_k[Y(x_{1i})] + \sum_{j=1}^{n} a_{jk}H_k[U(x_{2j})]\right\} \tag{12}$$

Since the C_k's are associated with the random function $Z(x)$ (and its transform $Y(x)$), the two terms inside the brackets of (12) can be thought of as a cokriging estimator for $H_k[Y(x_0)]$ at x_0. Therefore (12) can be written as

$$Z_{DCK}{}^*(x_0) = \sum_{k=0}^{\infty} C_k H_k{}^*[Y(x_0)] \tag{13}$$

where $H_k{}^*[Y(x_0)]$ depends on two random variables $Y(x_{1i})$ and $U(x_{2j})$. This result will also be used in the discussion on the conditional probability.

Block Disjunctive Kriging

If block-averaged estimates are desired (assuming that the samples are obtained on point support) the only modification to the above discussion is to use the block-averaged correlation and cross-correlation functions on the right-hand side of (11)

$$\bar{\rho}_{1j}{}^k(x_0 - x_\gamma) = 1/V \int_V \rho_{1j}{}^k(x - x_\gamma)\, dx \tag{14}$$

where γ indicates α or β when $j = 1$ or $j = 2$, respectively; V is the block area (or volume); and x describes the interior of V. Using (14) in (11) and solving the system of equations produces corresponding block-averaged weights $\bar{a}_{jk}$ and $\bar{b}_{ik}$, which are then used in (12).

Disjunctive Cokriging Variance

The disjunctive cokriging variance is

$$\sigma_{DCK}{}^2 = E[\{Z(x_0) - Z_{DCK}{}^*(x_0)\}^2] = E[Z(x_0)^2] - E[Z(x_0)Z_{DCK}{}^*(x_0)] \tag{15}$$

by virtue of the orthogonality condition $E[\{Z(x_0) - Z_{DCK}{}^*(x_0)\}Z_{DCK}{}^*(x_0)] = 0$. From (10) from *Yates et al.* [1986*a*] the first term on the right-most part of (15) is

$$E[Z(x_0)^2] = \sum_{k=1}^{\infty} k!C_k{}^2 + E[Z(x_0)]^2 = \sum_{k=0}^{\infty} k!C_k{}^2 \tag{16}$$

Solving for $E[Z(x_0)Z_{DCK}^*(x_0)]$ gives [see *Journel and Huijbregts*, 1978, especially p. 576]

$$E[Z(x_0)Z_{DCK}^*(x_0)] = \sum_{k=0}^{\infty} k!C_k^2\left\{\sum_{i=1}^{m} b_{ik}\rho_{11}^k(x_0 - x_{1i}) + \sum_{j=1}^{n} a_{jk}\rho_{12}^k(x_0 - x_{2j})\right\} \quad (17)$$

Incorporating (16) and (17) into (15) and noting that for $k = 0$, $E[Z(x_0)Z_{DCK}^*(x_0)]$ is equal to $E[Z(x_0)^2]$ (i.e., $b_{i0} = a_{j0} = 1/(n + m)$) gives the punctual disjunctive cokriging variance

$$\sigma_{DCK}^2 = \sum_{k=1}^{\infty} k!C_k^2\left[1 - \sum_{i=1}^{m} b_{ik}\rho_{11}^k(x_0 - x_{1i}) - \sum_{j=1}^{n} a_{jk}\rho_{12}^k(x_0 - x_{2j})\right] \quad (18)$$

which has a similar form to the linear cokriging variance.

Conditional Probability

The conditional probability (CP) that the unknown value at a randomly located point is above a prescribed cutoff level can be estimated using DCK. The derivation of the CP for cokriging is similar to the disjunctive kriging case, but due to the added information from the second random function, it will, in general, be an improved estimator. For purposes described herein, the point CP will be discussed. However, if a block-averaged value is required, it is only necessary to replace a_{jk} and b_{ik} with $\bar{a}_{jk}$ and $\bar{b}_{ik}$ in the following discussion. The conditional probability that the value of $Z(x_0)$ is greater than a cutoff level z_c (or y_c for the transformed variable) conditioned on the available data is

$$P[Z(x_0) \geq z_c \mid Z(x_\alpha), V(x_\beta)] = P[Y(x_0) \geq y_c \mid Y(x_\alpha), U(x_\beta)] \quad (19)$$

In order to determine the CP using the disjunctive kriging method it is necessary to transform the problem into a suitable form. Since disjunctive kriging is an estimator for the conditional expectation, a means whereby (19) can be written as the conditional expectation must be found. To do this an indicator variable $\Theta_{y_c}[Y(x)]$ is defined such that

$$\begin{aligned} \Theta_{y_c}[Y(x)] &= 1 \qquad Y(x) \geq y_c \\ \Theta_{y_c}[Y(x)] &= 0 \qquad Y(x) < y_c \end{aligned} \quad (20)$$

Using this indicator variable in (19) gives

$$P[Y(x_0) \geq y_c \mid Y(x_\alpha), U(x_\beta)] = P[\Theta_{y_c} = 1 \mid Y(x_\alpha), U(x_\beta)] = E[\Theta_{y_c} \mid Y(x_\alpha), U(x_\beta)] \quad (21)$$

The CP and the conditional expectation in (21) are the same because the indicator variable has a value of zero for $-\infty < Y(x) < y_c$ and unity for $y_c \leq Y(x) < \infty$.

Expanding the unknown $\Theta_{y_c}[Y(x)]$ in a series of Hermite polynomials gives

$$\Theta_{y_c}[Y(x)] = \sum_{k=0}^{\infty} \theta_k H_k[Y(x)] = P_{DCK}^*(x) \quad (22)$$

where the θ_k's are the coefficients of the expansion which are determined by using the orthogonality properties [see *Journal and Huijbregts*, 1978; *Yates et al.*, 1986*a*], and P_{DCK}^* is the estimator of the CP. For $k = 0$, θ_k equals $1 - G(y_c)$, and for $k > 0$, θ_k equals $g(y_c)H_{k-1}(y_c)/k!$, where $G(u)$ and $g(u)$ are the gaussian cumulative frequency and density functions, respectively.

Incorporating the coefficients into (22) gives the estimator for the CP at the point x_0:

$$P_{DCK}^*(x_0) = 1 - G(y_c) + g(y_c)\sum_{k=1}^{\infty} H_{k-1}(y_c)H_k[Y(x_0)]/k! \quad (23)$$

The only unknown in (23) is $H_k[Y(x_0)]$, which is estimated using the disjunctive cokriging estimator (see equations (12) and (13)):

$$H_k^*[Y(x_0)] = \sum_{i=1}^{m} b_{ik}H_k[Y(x_{1i})] + \sum_{j=1}^{n} a_{jk}H_k[U(x_{2j})] \quad (24)$$

Comparing (23) with (42) of *Yates et al.* [1986*a*] demonstrates the similarity between the results for one and two variables. The only difference is that for cokriging there is a contribution to the estimate for $H_k[Y(x_0)]$ from the auxiliary random function.

The CP in (23) can be written in terms of a conditional probability density function, $\text{Pdf}^*(u)$

$$P_{DCK}^*(x_0) = \int_{y_c}^{\infty} \text{Pdf}^*(u)\, du \quad (25)$$

where the probability density function, which is evaluated at x_0 is written as

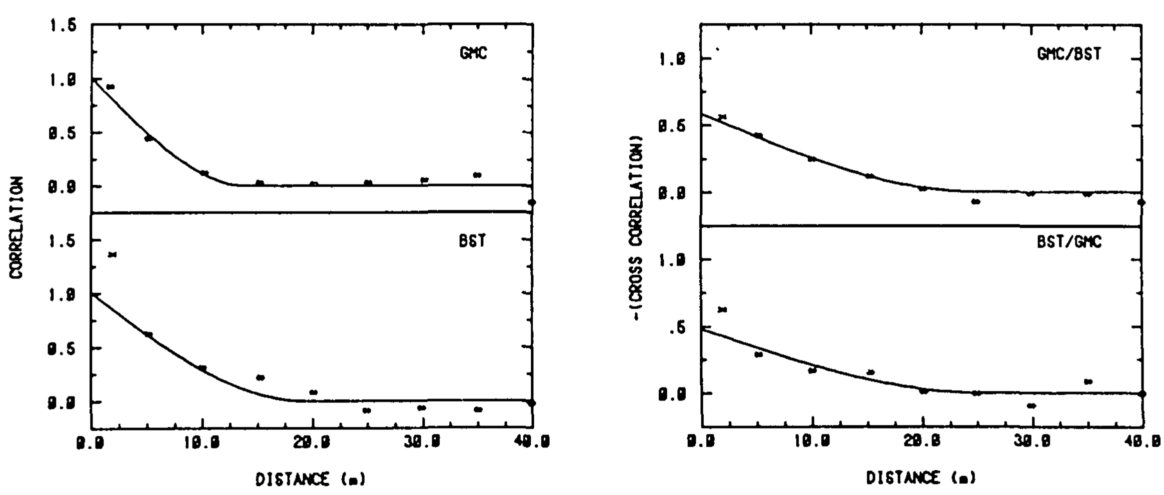

Fig. 2. Global correlation and cross correlation functions. The exes and asterisks indicate that less than 10 and more than 65 pairs, were used to generate a value, respectively.

TABLE 1. Covariance Functions

Covariance* Function	Sill	Range	R_μ	R_{σ^2}
C_{11}	42.5	19.0	−0.010	0.95
C_{22}	33.5	14.0	−0.052	1.01
C_{12}	−18.0	25.0	−0.013	0.84
C_{21}	−22.0	25.0		

Models are of the spherical type. The nugget for each model was found to be zero.

*Definition of subscripts: 1 is for GMC, and 2 is for BST.

$$\mathrm{Pdf}^*(u) = g(u)\left\{1 + \sum_{k=1}^{\infty} H_k(u)H_k[Y(x_0)]/k!\right\} \quad (26)$$

EXAMPLE

The remainder of this paper will illustrate the DCK method with an example. In order to use the cokriging method (whether ordinary or disjunctive), a pair of correlated random variables must be available. The following example uses a data set which contains 71 values of the gravimetric moisture content (GMC) and 148 values of the bare soil surface temperature (BST).

Sample statistics were obtained for the data sets. For the GMC, the mean, variance, skew, and kurtosis are 10.76, 37.06, 0.64, and 7.21, respectively, and 2.14 < GMC (%) < 26.35. For the BST data, the mean, variance, skew, and kurtosis are 38.36, 23.39, 0.39, and 1.70, respectively, and 31.71 < BST (°C) < 47.74. The correlation [*Sokal and Rohlf*, 1981], r, between the GMC and the BST is −0.744.

The hypothesis that the GMC and BST data are either normally or lognormally distributed was tested using the Kolmogorov-Smirnov (KS) test [*Sokal and Rohlf*, 1981; *Rohlf and Sokal*, 1981; *Rao et al.*, 1979]. A probability level of 0.1 was chosen in order to minimize the probability of the type II error [*Rao et al.*, 1979]. The KS critical values for the GMC and BST data are 0.096 and 0.066, respectively. Comparing the KS test values of 0.146 and 0.173, for the GMC and BST data, respectively, to the KS critical values indicates that neither the GMC nor the BST was obtained from a normally distributed random function. Taking the $\log_{10}$ transform of the data and calculating KS test values (0.107 and 0.167, for the GMC and BST, respectively) indicate that neither the GMC nor BST is lognormally distributed.

The spatial correlation functions for the GMC, BST, and their crosscorrelation are plotted in Figure 2, where the solid circles indicate the sample correlation functions and the solid curve the corresponding spherical model. The exes and asterisks indicate that less than 10 and more than 65 pairs of samples fell into the lagged interval, respectively.

The method used to generate the correlation and cross-correlation functions was to calculate the sample covariance and cross covariance [*Journel and Huijbregts*, 1978, especially pp. 40 and 194] from the original data and to fit a spherical model to the sample function. The covariance models were then "tested" for validity using the jackknifing procedure [*Russo*, 1984*a*, *b*; *Vauclin et al.*, 1983] where the reduced mean and variance are calculated using an estimated and actual value at each sample location. Based on the jackknifing method, the spherical models listed in Table 1 were deemed satisfactory representations of the true spatial correlation functions. The reduced mean and variance for the model covariance functions are also given in Table 1. A difficulty in using this method for validating a spatial correlation function is that it lacks an independent method for determining what is an appropriate value for the reduced mean and variance.

The cross-covariance functions were also tested for validity using the jackknifing procedure. The model covariance functions, which were determined and validated prior to the cross-covariance functions, were used in the validation procedure and considered constants.

The correlation and cross-correlation functions were calculated from the covariance by dividing the covariance and cross-covariance functions by $C_{ii}(0)$ and $[C_{11}(0)C_{22}(0)]^{0.5}$, respectively, where C_{11} is used to represent the covariance function for the GMC and C_{22} for the BST. Sample variograms were also calculated (not shown) and verify that the nugget is approximately zero.

The correlation and cross-correlation functions were determined in this manner to facilitate comparison between the DCK and OCK results. An alternate method for calculating the spatial correlation functions directly would be to use the transformed data.

After the spatial correlation functions have been determined, the next step in the disjunctive kriging method is to determine the transform relationships, ϕ and ψ (see equation (1)). These relationships were determined by using 10-term Hermite integration [*Abramowitz and Stegun*, 1965]. For a more detailed description the reader is referred to *Yates et al.* [1986*a*, *b*]. In Figure 3 the empirical distribution for each data set (i.e., GMC, in Figure 3 (left) and BST in Figure 3 (right)) is shown by the dots and the ϕ and ψ transform relationships by the solid and dashed curves, respectively. Ten and 30 coef-

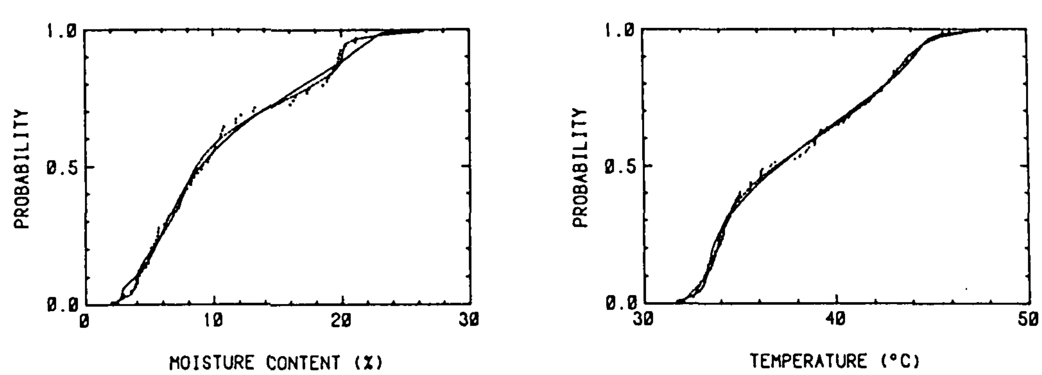

Fig. 3. Comparison between empirical distribution (solid circles) and transform relationship (solid and dashed curves). Ten and 30 coefficients, were used in the calculation for the solid and dashed curves (see equations (1) and (2)), respectively.

ficients (i.e., C_k or D_k) were used in generating the solid and dashed curves in Figure 3. Although 30 coefficients produce a good fit between the data and the transform relationships, the improvement in the disjunctive kriging estimates diminishes rapidly as k increases [*Rendu*, 1980; *Yates et al.*, 1986*b*]. Therefore only 10 coefficients were used in subsequent calculations.

Figures 4 and 5 contain a comparison between OCK (Figure 4) and DCK (Figure 5). A total of 931 estimates on a 3- by 3-m grid system superimposed over the field were used to generate the contour maps. The estimates were calculated using 10 Hermite polynomials and 5 nearest neighbors within a maximum radius of 30 m. It is evident from these figures that zones of high GMC tend to lie near the north and south boundaries of the field and the lower levels of GMC near the middle part of the field. Also, the contour diagram produced using linear cokriging is more continuous when compared to DCK (see, for example, the contour GMC = 7). Visual observation in the field during sampling tends to support the results of DCK over those of ordinary cokriging as giving a more accurate moisture profile, since it was noted that the wetter zones appeared in small patches throughout the field and especially near the north and south borders.

Using the 931 estimates of the GMC, the mean value (of the estimates) was calculated. The DCK method produced a value of 10.38, whereas for OCK the mean was 9.64. For this data set there appears to be more bias in the estimates for OCK compared to DCK. This may be due to using all the available data to determine the covariance functions and then making comparisons to the same data. It would be better if a fraction of the data were used to calculate the covariance function and then the comparisons made to the remaining data. However, adequate data were not available.

The kriging variance was calculated at each of the 931 points using DCK and OCK. From these values the average kriging variance was calculated for each method. The DCK method had an average kriging variance which was 7% lower than for OCK (e.g., 23.53 versus 25.24 for OCK).

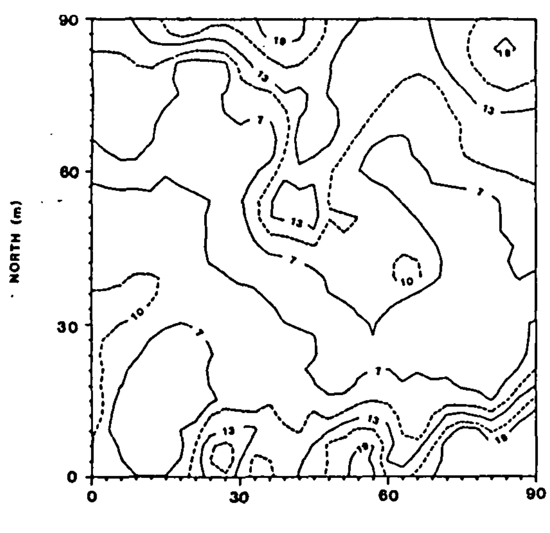

Fig. 4. Contour diagram for the GMC based on ordinary cokriging. Estimates were calculated using five nearest neighbors within a maximum radius of 30 m. The contours 7, 13, and 19% are solid curves and 10 and 16% are dashed curves.

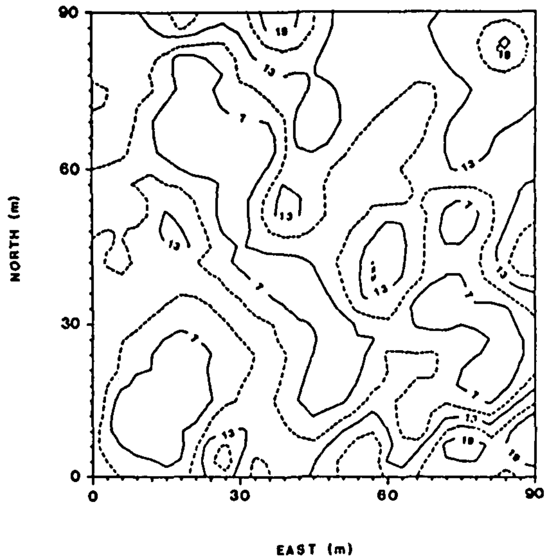

Fig. 5. Contour diagram for the GMC based on disjunctive cokriging. Estimates were calculated using five nearest neighbors within a maximum radius of 30 m. The contours 7, 13, and 19% are solid curves and 10 and 16% are dashed curves.

Another comparison between the two methods is the sum of squares between the actual sample values and the estimated values (based on the jackknifing technique). This method has an advantage over the average kriging variance in that the value of the random variable as well as the spatial correlation enter the calculation. The DCK method produced a value which was 19% lower than for OCK (e.g., 16.3 versus 20.1 for OCK).

The conditional probability that the estimate is greater than a cutoff level can be determined using DCK. Figure 6 contains a comparison of the CP and conditional probability density functions resulting from the DCK and DK methods for two points in the field. The curves marked 1 and 2 indicate the conditional probability distributions for points in the field located at (66 and 39 m) and (81 and 45 m), respectively. The solid and dashed curves represent the results from DCK and DK, respectively. At the point marked 1, the probability distribution resulting from DCK is almost identical to DK. The reason, to the first approximation, is that the GMC is dominating the estimation process, since the GMC data is located nearer to the estimation site than the BST data. The second site (2) is a position in the field where the reverse is true (i.e., the BST is closer to the estimation site). The results for Figure 6 show the advantage of using an auxiliary variable in the estimation process. At site 1 the GMC data supply most of the information and thus the two methods are similar. At site 2 the BST is as important as the GMC data and therefore affects the results more strongly.

CONCLUSIONS

The disjunctive cokriging method described in this paper produces a nonlinear estimator which is better than the linear (ordinary) cokriging estimator in terms of reduced variance of errors and accuracy of estimation. This is expected, since a nonlinear estimator should in general be superior to a linear estimator except when the random variables are bivariate normal. Estimates of the GMC based on 71 samples of GMC and 148 samples of BST were produced using both the DCK

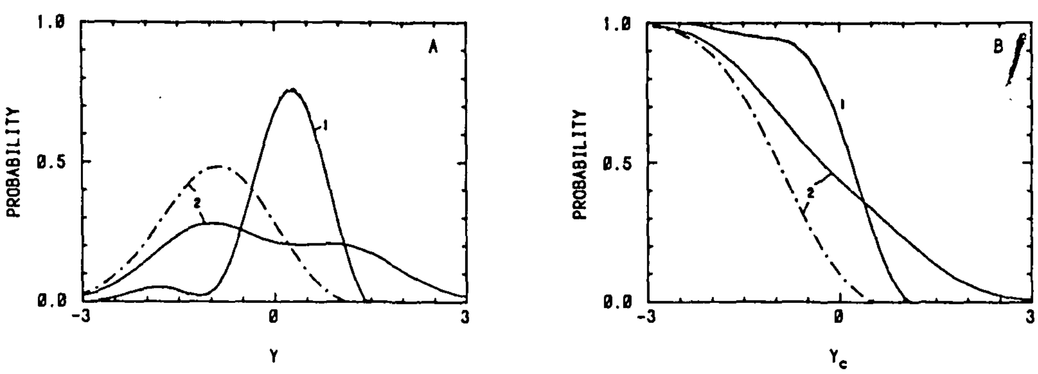

Fig. 6. (*a*) Conditional probability density and (*b*) cumulative distribution functions for two locations in the field. Curves marked 1 and 2 correspond to the locations 66 and 39 m and 81 and 45 m, respectively. Solid and dashed curves indicate the results from disjunctive cokriging and disjunctive kriging, respectively.

and OCK. Visual observation in the field during sampling supports the results produced by the DCK method over those by cokriging.

If in addition to the random function of primary interest there is an additional highly correlated random function, it can be included in the estimation process using the method outlined in this paper. The advantage is that the DCK estimator is a better estimator than the disjunctive kriging estimator (of one variable) by virtue of the information added to the problem by the auxiliary random function provided there is a significant correlation between the two random functions. A situation which lacks such a correlation will produce the same results as the disjunctive kriging method (of one variable) since the cross-correlation terms of (13) will be zero which decouples the equations.

The CP that the value of a property is above a given cutoff level can be based on cokriging. Since a better estimate of the $H_k[Y(x_0)]$ is obtained using DCK, in general, the CP based on DCK will be better than the CP based on disjunctive kriging. Using the DCK conditional probability can improve the sampling efficiency by adding additional highly correlated random functions which are easy to sample, if they exist, into the estimation process.

The CP was calculated at two points in the field. The first point is a location where the auxiliary variable adds little additional information. For this situation, both the DCK and DK methods are virtually identical. The second point is located in an area of the field where there are few GMC data nearby. For this location the DCK method produces different conditional probability distributions compared to the DK.

In general, the disjunctive kriging method requires more computational time compared to ordinary kriging. The same is true for cokriging. Disjunctive kriging (both DK and DCK) becomes an attractive alternative over ordinary kriging in situations where the probability distributions are required at the estimation site. For this situation the extra information which is not readily obtainable using linear kriging methods helps to offset the extra computational costs. However, as the cost of computer time becomes less expensive (i.e., as with personal computers) the additional computational requirements necessary to implement DK will be a less important consideration.

REFERENCES

Abramowitz, M., and A. Stegun, *Handbook of Mathematical Functions*, Dover, New York, 1965.

Al-Sanabani, M., Spatial variability of salinity and sodium adsorption ratio in a Typic Haplargid Soil, M.S. thesis, Univ. of Ariz., Tucson, 1982.

Carr, J. R., D. E. Myers, and C. E. Glass, Co-kriging—A computer program, *Comput. Geosci.*, *11*, 111–128, 1985.

Journel, A. G., and Ch. J. Huijbregts, *Mining Geostatistics*, Academic, Orlando, Fla., 1978.

Myers, D. E., Matrix formulation of co-kriging, *Math. Geol.*, *14*, 249–257, 1982.

Myers, D. E., Co-kriging—New developments, in *Geostatistics for Natural Resources Characterization*, edited by G. Verly et al., pp. 205–305, D. Reidel, Hingham, Mass., 1984.

Rao, P. V., P. S. C. Rao, J. M. Davidson, and L. C. Hammond, Use of goodness-of-fit tests for characterizing the spatial variability of soil properties, *Soil Sci. Soc. Am. J.*, *43*, 274–278, 1979.

Rendu, J. M., Disjunctive kriging: A simplified theory, *Math. Geol.*, *12*, 306–321, 1980.

Rohlf, F. J., and R. R. Sokal, *Statistical Tables*, 2nd ed., 219 pp., W. H. Freeman, New York, 1981.

Russo, D., Statistical analysis of crop yield-soil water relationships in heterogeneous soil under trickle irrigation, *Soil Sci. Soc. Am. J.*, *48*, 1402–1410, 1984*a*.

Russo, D., A geostatistical approach to solute transport in heterogeneous fields and its applications to salinity management, *Water Resour. Res.*, *20*, 1260–1270, 1984*b*.

Sokal, R. R., and F. J. Rohlf, *Biometry*, 2nd ed., 859 pp., W. H. Freeman, New York, 1981.

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

Yates, S. R., A. W. Warrick, and D. E. Myers, Disjunctive kriging, 1, Overview of estimation and conditional probability, *Water Resourc. Res.*, *22*, 615–622, 1986*a*.

Yates, S. R., A. W. Warrick, and D. E. Myers, Disjunctive kriging, 2, Examples, *Water Resourc. Res.*, *22*, 623–630, 1986*b*.

S. R. Yates, R. S. Kerr Environmental Research Laboratory, P. O. Box 1198, Ada, OK 74820.

(Received January 7, 1986;
revised April 21, 1986;
accepted April 28, 1986.)