AN ABSTRACT OF THE THESIS OF

Jeffrey D. Smyth for the degree of Master of Science in Agricultural Engineering

presented on May 23, 1988.

Title: Multivariate Geostatistical Analysis of Groundwater Contamination by Pesticide

and Nitrate

Abstract approved: _____

Jonathan D. Istok

A field study was conducted to determine the applicability of multivariate geostatistical methods to the problem of estimating and simulating pesticide concentrations in groundwater from measured concentrations of nitrate and pesticide, when pesticide is undersampled. Prior to this study, no published attempt had been made to apply multivariate geostatistics to groundwater contamination.

The study was divided into two complementary aspects of geostatistics: estimation and simulation. The use of kriging and cokriging to estimate nitrate and the herbicide dimethyl tetrachloroterepthalate (DCPA) contaminant densities is described in Chapter I. Measured concentrations of nitrate and the DCPA were obtained for 42 wells in a shallow unconfined alluvial and basin-fill aquifer in a 16.5 km2 agricultural area in eastern Oregon. The correlation coefficient between log(nitrate) and log(DCPA) was 0.74. Isotropic, spherical models were fitted to experimental direct- and cross-semivariograms with correlation ranges and sliding neighborhoods of 4 km. The relative gain for estimates obtained by cokriging ranged from 14 to 34%. Additional sample locations were selected for nitrate and DCPA using the fictitious point method. A simple economic analysis demonstrated that additional nitrate samples would be more beneficial in reducing estimation variances than additional DCPA samples, unless the costs of nitrate and DCPA analysis were identical.

These estimates are by definition, the Best Linear Unbiased Estimates (i.e., the estimates with minimized estimation variance), however the requirement of minimized variance smoothes the variability of contaminant values. The application of conditional simulations to groundwater contamination is described in Chapter II. Conditional simulation allows the degree of fluctuation of nitrate and DCPA between sample points to be assesed. With knowledge of both the 'best' estimates and the of the variability between sample points, nitrate and DCPA groundwater contamination in the study area can be characterized.

Based on the semivariogram models found in Chapter I, univariate and multivariate conditional simulations of nitrate and DCPA were generated using the turning bands method and the kriging or cokriging system. Kriging was used to condition the univariate simulations, while cokriging was used to cross-correlate and condition the multivariate simulations. The mean of 25 conditional and coconditional simulations at 8 different locations in the study area were generated and compared to kriging and cokriging estimates and 95% confidence intervals.

Both conditional and coconditional simulation of the DCPA and nitrate contaminant densities showed large variations when values in different simulations were compared. The fluctuation in values demonstrate the uncertainties in the contaminant distributions when sample sizes are small. As a result of this unkown component, simulated values vary widely. Coconditional simulation displayed the cross-correlation imposed by using the cokriging system to condition the simulations. After 25 simulations, the mean remained unstable indicating that more simulations would be required to enable comparisons with kriging and cokriging estimates.

Multivariate Geostatistical Analysis of Groundwater
Contamination by Pesticide and Nitrate

by

Jeffrey D. Smyth

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Master of Science

Completed May 23, 1988

Commencement June 1989

APPROVED:

_____

Dr. J. D. Istok, Assistant Professor of Agricultural Engineering and Civil Engineering

_____

Dr. A.G. Hashimoto, Head of Department of Agricultural Engineering

_____

Dean of Graduate School

Date thesis is presented_____May 23, 1988_____

Acknowledgement

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

MULTIVARIATE GEOSTATISTICAL ANALYSIS OF GROUNDWATER CONTAMINATION BY PESTICIDE AND NITRATE

INTRODUCTION

Geostatistical methods have been shown to be useful for the estimation and interpolation of regionalized variables (ReVs) i.e., variables that are distributed in space or time. Univariate geostatistical analysis refers to the study of a single ReV. The basic tools used in a univariate geostatistical analysis include the intrinsic hypothesis, the covariance or semivariogram function, and kriging. Multivariate geostatistical analysis refers to the study of two or more coregionalized variables i.e., ReVs that display cross-correlation. The tools used in multivariate geostatistical analysis are analogous to the univariate case and include the intrinsic hypothesis, the cross-covariance or cross-semivariogram function, and cokriging. These will be reviewed in Chapter I.

There have been numerous applications of univariate geostatistical analysis in hydrogeology. Aquifer properties such as porosity, transmissivity, and hydraulic conductivity have been estimated from field measurements [Delhomme, 1979; Van Rooy, 1987]. Myers et al [1982] used kriging to estimate the spatial distribution of geochemicals such as uranium and manganese in the Ogallala formation, while Flaig et al [1986] used kriging to estimate the distribution of nitrate in soil. Kriging has also been used to estimate radionuclide concentrations in soil at the Nevada Test Site [Barnes, 1978; Barnes et al, 1980], dioxin (2,3,7,8-TCDD) contamination in soil on roads in Missouri [Zirschy and Harris, 1986], and dioxin contamination in creek sediments [Zirschy et al, 1985]. Point estimates and estimates of the total amount of the dissolved groundwater contaminants zinc, boron, iron, mangenese, barium, and total volatile organic carbon at the Chem-Dyne Superfund site in Ohio were made using kriging by Cooper and Istok [1988a,b], and Istok and Cooper [1988].

There have been fewer applications of multivariate geostatistical analysis in hydrogeology. Hoeksema and Clapp [1987] studied the cross-correlation between the ReVs water table elevation and ground surface elevation. Vauclin et al [1983] studied the cross-correlation between the ReVs %sand, %silt and %clay and available water content, while Abourifasso and Marino [1984] used cokriging to estimate aquifer transmissivity and specific capacity. No attempt to perform a multivariate geostatistical analysis of groundwater contamination has been previously published.

In many cases of groundwater contamination there is a need to estimate the concentrations of two or more contaminants. This is frequently the case in agricultural areas when groundwater is contaminated by agricultural chemicals (i.e., pesticides and fertilizers). The two agricultural chemicals of interest in this study are Dimethyl tetrachloralterephthalate (DCPA) and nitrate. The cost of measuring DCPA concentration in a groundwater sample in this study was high enough to limit the number of samples that could be analyzed for DCPA. However the low cost of measuring nitrate permited a large number of samples to be analyzed. Earlier studies had indicated that there is a significant correlation between nitrate and DCPA concentrations in groundwater samples from the study area. The overall objective of Chapter I was to determine if it is feasible to use multivariate geostatistical analysis to estimate DCPA concentrations in groundwater from measured DCPA and nitrate concentrations when DCPA is undersampled with respect to nitrate.

Delhomme [1979] notes that the total spatial variability of a ReV can be considered as two parts: the estimate and the uncertaintity of the estimate that exists between sample points. The estimates obtained by kriging or cokriging are, by definition, the estimates with minimized estimation variance and therefore show less fluctuation than the true values [Journel and Huijbregts, 1978]. However, neither the kriging or cokriging system allow the values of the ReVs to vary enough between sample points to reflect the fluctuation which might be possible.

A geostatistical analysis is not complete unless the degree of fluctuation of the ReVs between sample points is assessed. In univariate geostatistics, conditional simulation (CS) fulfills the role of creating realizations of the random function that characterizes the ReV between sample points. The purpose of CS is then to assess the fluctuation of the ReV that is possible in the domain of interest. An additional use of CS is to generate input data for use in stochastic-deterministic groundwater flow and solute transport models.

In hydrogeology CS has been used primarily to generate realizations of transmissivities and hydraulic head [Delhomme, 1979; Dagan, 1982; Clifton and Neumann, 1982; Van Rooy, 1987]. Bryan and Myers [1984] used CS in a case study of lead contamination in soil to generate 'worst-case scenarios' i.e., simulations that show the possible fluctuation in the lead contaminated soil. Technically simulations obtained using CS do not represent worst-case scenarios, but rather a set of possible versions of reality that are consistent with the available data. An application of CS to groundwater contamination has not previously been published. The objective of Chapter II is to extend CS to univariate and multivariate conditional simulation of DCPA and nitrate concentrations in groundwater.

MULTIVARIATE GEOSTATISTICAL ANALYSIS OF GROUNDWATER CONTAMINATION BY PESTICIDE AND NITRATE

I. ESTIMATION

Jeffrey D. Smyth

Hydrology Section, Battelle, Pacific Northwest Laboratories, Richland, Washington 99352 (Formerly at Agricultural Engineering Department, Oregon State University, Corvallis, Oregon 97331)

Jonathan D. Istok

Agricultural Engineering and Civil Engineering Departments, Oregon State University, Corvallis, Oregon 97331

Abstract

A field study was conducted to determine the applicability of multivariate geostatistical methods to the problem of estimating pesticide concentrations in groundwater from measured concentrations of nitrate and pesticide, when pesticide is undersampled. Measured concentrations of nitrate-N and the herbicide Dimethyl tetrachloroterepthalate (DCPA) were obtained for 42 wells in a shallow unconfined alluvial and basin-fill aquifer in a 165 km2 agricultural area of in eastern Oregon. The correlation coefficient between log(nitrate-N) and log(DCPA) was 0.74. Isotropic, spherical models were fitted to experimental direct- and cross-semivariograms with correlation ranges and sliding neighborhoods of 4 km. The relative gain for estimates obtained by cokriging ranged from 14 to 34%. Additional sample locations were selected for nitrate and DCPA using the fictitious point method. A simple economic analysis demonstrated that additional nitrate samples would be more beneficial in reducing estimation variances than additional DCPA samples, unless the costs of nitrate and DCPA analysis were identical.

MULTIVARIATE GEOSTATISTICAL ANALYSIS OF GROUNDWATER CONTAMINATION BY PESTICIDE AND NITRATE

I. ESTIMATION

Jeffrey D. Smyth

Hydrology Section, Battelle, Pacific Northwest Laboratories, Richland, Washington 99352

(Formerly at Agricultural Engineering Department, Oregon State University, Corvallis, Oregon 97331)

Jonathan D. Istok

Agricultural Engineering and Civil Engineering Departments, Oregon State University, Corvallis, Oregon 97331

Introduction

Geostatistical methods have been shown to be useful for the estimation and interpolation of regionalized variables (ReVs) i.e., variables that are distributed in space or time. Univariate geostatistical analysis refers to the study of a single ReV. The basic tools used in a univariate geostatistical analysis include the intrinsic hypothesis, the covariance or semivariogram function, and kriging. Multivariate geostatistical analysis refers to the study of two or more coregionalized variables i.e., ReVs that display cross-correlation. The tools used in multivariate geostatistical analysis are analogous to the univariate case and include the intrinsic hypothesis, the cross-covariance or cross-semivariogram function, and cokriging. These will be reviewed later in the paper.

There have been numerous applications of univariate geostatistical analysis in hydrogeology. Aquifer properties such as porosity, transmissivity, and hydraulic conductivity have been estimated from field measurements [Delhomme, 1979; Van Rooy,

1987]. Myers et al [1982] used kriging to estimate the spatial distribution of geochemicals such as uranium and manganese in the Ogallala formation, while Flaig et al [1986] used kriging to estimate the distribution of nitrate in soil. Kriging has also been used to estimate radionuclide concentrations in soil at the Nevada Test Site [Barnes, 1978; Barnes et al, 1980], dioxin (2,3,7,8-TCDD) contamination in soil on roads in Missouri [Zirschy and Harris, 1986], and dioxin contamination in creek sediments [Zirschy et al, 1985]. Point estimates and estimates of the total amount of the dissolved groundwater contaminants zinc, boron, iron, mangenese, barium, and total volatile organic carbon at the Chem-Dyne Superfund site in Ohio were made using kriging by Cooper and Istok [1988a,b], and Istok and Cooper [1988].

There have been fewer applications of multivariate geostatistical analysis in hydrogeology. Hoeksema and Clapp [1987] studied the cross-correlation between the ReVs water table elevation and ground surface elevation. Vauclin et al [1983] studied the cross-correlation between the ReVs %sand, %silt and %clay and available water content, while Abourifasso and Marino [1984] used cokriging to estimate aquifer transmissivity and specific capacity. No attempt to use cokriging in the analysis of groundwater contaminants has been previously published.

In many cases of groundwater contamination there is a need to estimate the concentrations of two or more contaminants. This is frequently the case in agricultural areas when groundwater is contaminated by agricultural chemicals (i.e., pesticides and fertilizers). The two agricultural chemicals of interest in this study are Dimethyl tetrachloralterephthalate (DCPA) and nitrate. The cost of measuring DCPA concentration in a groundwater sample in this study was high enough to limit the number of samples that could be analyzed for DCPA. However the cost of measuring nitrate was low enough to permit a large number of samples to be analyzed for nitrate. An earlier study had indicated that there is a significant correlation between nitrate and DCPA concentrations in groundwater samples from the study area (Bruck, unpublished data).

The overall objective of this study was to determine if it is feasible to use multivariate geostatistical analysis to estimate DCPA concentrations in groundwater from measured DCPA and nitrate concentrations when DCPA is undersampled with respect to nitrate. The specific objectives of this study were:

1. To estimate concentrations of DCPA and nitrate in an alluvial and basin-fill aquifer.

2. To compare estimates for DCPA and nitrate obtained by univariate and multivariate geostatistical analyses.

3. To compare the reductions in estimation variance obtained by adding DCPA and nitrate samples.

4. To estimate the total aqueous mass of DCPA and nitrate in the portion of aquifer studied and to compute the estimation variances.

Equation Development

The fundamental principles of univariate geostatistical analysis have been well documented previously [David, 1977; Journel and Huijbregts, 1978] and will not be repeated in this paper. In univariate geostatistical analysis, the covariance and direct-semivariogram describe the spatial structure displayed by a ReV. In multivariate geostatistical analysis the spatial structure of a pair of cross-correlated variables is described by the cross-covariance and cross-semivariogram. Assuming second-order stationarity for two ReVs $Z_j(x)$ and $Z_k(x)$ the cross-covariance is defined as

$$C_{jk}(h) = E\{ [Z_j(x+h) - m_j] [Z_k(x) - m_k] \}$$ (1)

where

$$C_{jk}(h) = \text{cross-covariance for ReVs } Z_j(x) \text{ and } Z_k(x),$$

$$h = \text{distance between the two points } x \text{ and } x+h,$$

$$m_j = E[Z_j],$$

$$m_k = E[Z_k],$$

$$\text{and } E[\ ] = \text{expectation operator.}$$

Using the intrinsic hypothesis (i.e., that the expected value and the variance of $Z_j(x)$ and $Z_k(x)$ are not a function of the position in the domain, x) which is implied by the assumption of second-order stationarity, the cross-semivariogram for ReVs $Z_j(x)$ and $Z_k(x)$, $\gamma_{jk}(h)$ is defined as

$$\gamma_{jk}(h) = \frac{1}{2} E \left\{ [Z_j(x + h) - Z_j(x)] [Z_k(x + h) - Z_k(x)] \right\}$$ (2)

Then the relationship between the cross-covariance and cross-semivariogram is

$$2\gamma_{jk}(h) = 2\gamma_{kj}(h) = 2C_{jk}(0) - C_{jk}(h) - C_{kj}(h).$$ (3)

The cross-semivariogram is symmetric (i.e., $\gamma_{jk}(h) = \gamma_{kj}(h)$) and (3) is therefore simplified as

$$\gamma_{jk}(h) = C_{jk}(0) - \frac{1}{2}\left[C_{jk}(h) + C_{kj}(h)\right]. \tag{4}$$

In practice the experimental cross-semivariogram is computed only for pairs of sample points which have data for both ReVs of interest. The experimental cross-semivariogram for ReVs $Z_j(x)$ and $Z_k(x)$ is computed using

$$\overset{*}{\gamma}_{jk}(h) = \frac{1}{2N(h)}\sum_{i=1}^{N(h)}\left\{\left[Z_j(x_i) - Z_j(x_i + h)\right]\left[Z_k(x_i) - Z_k(x_i + h)\right]\right\} \tag{5}$$

where

$N(h)$ = number of pairs of sample points, separated by h, which have measured values of both ReVs $Z_j(x)$ and $Z_k(x)$, and x_i is a sample point, where i= 1,...,n.

Note that when there is only one regionalized variable, (5) reduces to the experimental direct-semivariogram used in univariate geostatistical analysis:

$$\overset{*}{\gamma}_j(h) = \frac{1}{2N(h)}\sum_{i=1}^{N(h)}\left\{\left[Z_j(x_i) - Z_j(x_i + h)\right]^2\right\} \tag{6}$$

where

$N(h)$ = number of pairs of sample points, separated by h, which have measured values of the ReV $Z_j(x)$.

According to Myers [1982] the most general use of cokriging is to reduce estimation variances for ReVs when all ReVs are fully sampled. A more specific use of cokriging is in situations where there is a cross-correlation between ReVs and one or more of the ReVs are undersampled. In this case cokriging is used to estimate the undersampled ReVs and reduce the estimation variance. Following Myers [1982], the development of the cokriging system of equations to estimate an undersampled ReV will be treated as a special case.

A linear estimator for m ReVs is

$$\bar{Z}^*(x) = \sum_{i=1}^{n} \bar{Z}(x_i)\, \Gamma_i \tag{7}$$

where

$$\bar{Z}^*(x) = [\, Z_1^*(x) \;\; Z_2^*(x) \;...\; Z_m^*(x)], \tag{8}$$

are the estimates of the ReVs at point x, and

$$\bar{Z}(x_i) = [\, Z_1(x_i) \;\; Z_2(x_i) \;...\; Z_m(x_i)] \tag{9}$$

are the measured values of the ReVs at the sample points, and

$$\Gamma_i = \begin{bmatrix} \lambda_{11}^i & \cdots & \lambda_{1m}^i \\ \vdots & \vdots & \vdots \\ \lambda_{m1}^i & \cdots & \lambda_{mm}^i \end{bmatrix} \tag{10}$$

are the sample weights, e.g.

λ_{jk}^i = weight attributed to $Z_j(x_i)$ in estimating $Z_k(x_i)$.

Note if x_0 is an unsampled location for the ReV $Z_k(x)$, all weights used to estimate the remaining ReVs from $Z_k(x_0)$ are zero i.e., $\lambda_{k1}^i...\lambda_{km}^i = 0$, and, $\lambda_{1k}^i...\lambda_{mk}^i = 0$. This

requires the kth row and column in (10) to be all zeros for the location i, where $Z_k(x)$ is unsampled.

Assuming the intrinsic hypothesis is valid, the linear estimator (7) is unbiased if

$$\sum_{i=1}^{n} \Gamma_i = I \qquad (11)$$

where

I = identity matrix.

The cokriging estimation variance, σ_{CK}^2 is

$$\sigma_{CK}^2 = \sum_{j=1}^{m} \text{Var} [Z_j(x) - Z_j^{*}(x)]. \qquad (12)$$

The estimation variance can be minimized using Lagrange multipliers [Myers, 1982]. This results in the formulation of the cokriging system of equations,

$$UY = D \qquad (13)$$

where

$$U = \begin{bmatrix} \overline{\gamma}(x_1,x_1) & \cdots & \overline{\gamma}(x_1,x_n) & I \\ \vdots & \vdots & \vdots & \vdots \\ \overline{\gamma}(x_n,x_1) & \cdots & \overline{\gamma}(x_n,x_n) & I \\ I & \cdots & I & 0 \end{bmatrix} \qquad (14)$$

$$Y = \begin{bmatrix} \Gamma_1 \\ \vdots \\ \Gamma_n \\ \overline{\mu} \end{bmatrix} \qquad (15)$$

$$D = \begin{bmatrix} \overline{\gamma}(x_1, x) \\ \vdots \\ \overline{\gamma}(x_n, x) \\ I \end{bmatrix} \qquad (16)$$

$$\overline{\gamma}(x_n, x_n) = \begin{bmatrix} \gamma_{11}(x_n, x_n) & \cdots & \gamma_{1m}(x_n, x_n) \\ \vdots & \vdots & \vdots \\ \gamma_{m1}(x_n, x_n) & \cdots & \gamma_{mm}(x_n, x_n) \end{bmatrix} \qquad (17)$$

and

$\overline{\mu}$ = a matrix of Lagrange multipliers

Equation (13) can be solved for the weights Γ_i for use in (7). The minimized cokriging estimation variance is

$$\sigma_{CK}^2 = Tr \left[\sum_{i=1}^{n} \overline{\gamma}(x, x_i) \, \Gamma_i \right] + Tr \, \overline{\mu} \qquad (18)$$

where

Tr = trace of a matrix.

For the case of the ReVs in this study (5), (8), (12), (14), and (18) can be simplified. Let Z_N and Z_D represent nitrate and DCPA contaminant densities (mass of contaminant per surface area of aquifer, see Methodology), respectively. Assume that n sites have been sampled for DCPA, and m sites for nitrate. Further, assume that $m > n$ (i.e., that DCPA is undersampled). Note that the sample locations for DCPA and nitrate do not necessarily occur at the same points.

Under the intrinsic hypothesis three semivariograms exist: the direct-semivariograms for DCPA and nitrate, $\gamma_{DD}(h)$, $\gamma_{NN}(h)$, and the cross-semivariogram $\gamma_{ND}(h) = \gamma_{DN}(h)$. The three experimental semivariograms are computed using

$$\overset{*}{\gamma}_{DD}(h) = \frac{1}{2n(h)} \sum_{i=1}^{n(h)} \left\{ [Z_D(x_i) - Z_D(x_i + h)]^2 \right\} \tag{19}$$

$$\overset{*}{\gamma}_{NN}(h) = \frac{1}{2m(h)} \sum_{j=1}^{m(h)} \left\{ [Z_N(x_j) - Z_N(x_j + h)]^2 \right\} \tag{20}$$

$$\overset{*}{\gamma}_{DN}(h) = \frac{1}{2N(h)} \sum_{l=1}^{N(h)} \left\{ [Z_D(x_l) - Z_D(x_l + h)][Z_N(x_l) - Z_N(x_l + h)] \right\} \tag{21}$$

where $n(h)$, $m(h)$, and $N(h)$ are the number of pairs of sample points, separated by h, which have measured values of the ReVs $Z_D(x)$, $Z_N(x)$, and both $Z_D(x)$ and $Z_N(x)$, respectively.

The linear estimator for DCPA with nitrate support is

$$Z_D^*(x_0) = \sum_{i=1}^{n} \lambda_D^i Z_D(x_i) + \sum_{j=1}^{m} \lambda_N^j Z_N(x_j) \tag{22}$$

where

x_0 = unsampled location for DCPA,

λ_D^i = weight assigned to a measured value of DCPA at sample point x_i,

and

λ_N^j = weight assigned to a measured value of nitrate at sample point x_j.

Equation (22) becomes an unbiased estimator if the sum of the weights for the DCPA sample points equals one and the sum of the weights for the nitrate sample points equals zero

$$\sum_{i=1}^{n} \lambda_D^i = 1 \quad \text{and} \quad \sum_{j=1}^{m} \lambda_N^j = 0 \tag{23}$$

The linear estimator for nitrate with DCPA support is

$$Z_N^*(x_0) = \sum_{i=1}^{n} \lambda_D^i Z_D(x_i) + \sum_{j=1}^{m} \lambda_N^j Z_N(x_j) \tag{24}$$

with the unbiased condition written as

$$\sum_{j=1}^{m} \lambda_N^j = 1 \quad \text{and} \quad \sum_{i=1}^{n} \lambda_D^i = 0 \tag{25}$$

The cokriging system for (22) is

$$\gamma_{DD}(x_0, x_k) = \sum_{i=1}^{n} \lambda_D^i \gamma_{DD}(x_k, x_i) - \sum_{j=1}^{m} \lambda_N^j \gamma_{DN}(x_k, x_j) - \mu_D \tag{26}$$

$$\text{for } k = 1, \ldots, n.$$

$$\gamma_{DN}(x_0, x_l) = -\sum_{i=1}^{n} \lambda_D^i \gamma_{DD}(x_l, x_i) - \sum_{j=1}^{m} \lambda_N^j \gamma_{NN}(x_l, x_j) - \mu_N \tag{27}$$

$$\text{for } l = 1, \ldots, n.$$

and the minimized estimation variance is

$$\sigma_{CK}^{2*} = -\sum_{i=1}^{n} \lambda_D^i \gamma_{DD}(x_0, x_i) + \sum_{j=1}^{m} \lambda_N^j \gamma_{DN}(x_0, x_j) + \mu_D \tag{28}$$

where μ_D and μ_N are Lagrange multipliers for DCPA and nitrate, respectively.

The 95% confidence interval for DCPA estimates can be written,

$$Z_D^*(x_0)_{95\%} = Z_D^*(x_0) \pm 2(\sigma_{D0}) \tag{29}$$

where, σ_{D0} is the standard deviation of the DCPA estimate, and $Z_D^*(x_0)$ is the estimate of DCPA [Journel and Huijbregts, 1978]. A similiar equation can be written for the 95% confidence interval for nitrate.

From (18) and (26), the estimation variance is computed using only the location of the sample points and not the measured value of the ReVs. Consequently, the computed estimation variances can be used to optimally locate additional samples using the method of 'fictitious points' [Delhomme, 1978]. If the objective of the additional sampling is to reduce the maximum estimation variance, the first additional sample (i.e., the first fictitious point) is placed where the maximum estimation variance occurs. The cokriging system is solved to obtain the new estimation variances, and the next fictitious sample is

placed at the point where the new maximum estimation variance occurs. Additional samples are added until the maximum point estimation variance reaches an acceptable level. If the objective is to reduce the average estimation variance, then the first additional sample is placed in the location that maximizes the reduction in average estimation variance. This location can be determined by assuming a position for the sample, solving the cokriging system, and evaluating the average estimation variance. A new sample position is assumed and the process is repeated until the optimum location is found. Additional samples are added until the average estimation variance reaches an acceptable level. In both cases the relative reduction in estimation variance, $R(x)$, is defined as

$$R(x) = \frac{\sigma_{CK}^2 - \sigma_{CK}'^2}{\sigma_{CK}^2} \qquad (30)$$

where

σ_{CK}^2 = estimation variance for the original sampling

$\sigma_{CK}'^2$ = estimation variance with the additional samples.

By comparing the relative reduction in estimation variance with the cost of obtaining and analyzing additional groundwater samples, the benefit of additional samples can be evaluated. Thus the method of 'fictitious points' could be used to help design an effective and economical sampling program.

Inherent in the use of cokriging to optimize the placement of additional samples is the assumption that a real measurement made at the fictitious sample would not change the type of model or the values of model parameters. The validity of this assumption depends on the number of original sample points used to calculate the experimental semivariograms and on the number of fictitious samples being considered.

Restrictions and Assumptions

Several assumptions are required to apply multivariate geostatistical analysis to problems of groundwater contamination. For a linear geostatistical analysis (i.e., estimators of the form (7)) the ReVs must be normally distributed. In many cases however, the type of distribution is not clear due to the small number of data. Luster [1986] notes that when the type of distribution is not clearly defined an assumption about the underlying distribution must be made from knowledge of the physical, chemical, or biological processes involved. The assumption of a lognormal distribution for groundwater contaminants appears to be valid in some cases [Cooper and Istok, 1988b; Istok and Cooper, 1988; Myers et al, 1982]. Other transformations can be used to improve the fit of the data to a normal distribution as long as the transformation is completely invertible [Journel and Huijbregts, 1978].

Additionally, when using cokriging the ReVs must have joint-normal distributions for h ≥ 0 [Journel and Huijbregts, 1978; Verly, 1984; Luster, 1986]. It is often assumed that if joint-normality is observed at h = 0, then the ReVs are joint-normal at h ≠ 0. This assumption is called the multi-gaussian hypothesis [Verly, 1984; Luster, 1986]. Inherent in the multi-gaussian hypothesis is the assumption that the distributions are stationary. Transforming ReVs can make it difficult to accept the assumption of multi-gaussian behavior, especially when the number of data is small.

The linear model of coregionalization [Journel and Huijbregts, 1978] requires that all direct- and cross-semivariogram models have the same mathematical form and range i.e., the models are of the form

$$\gamma_{\alpha\beta}(h) = \sum_{i=1}^{M} b_{\alpha\beta}^{i} \, \gamma_{i}'(h) \tag{31}$$

where

$$\alpha = Z_N, Z_D$$

$$\beta = Z_N, Z_D$$

$b_{\alpha\beta}^i$ = parameter of the direct– and cross–semivariogram model (e.g., nugget, sill)

for ReV α and β

$\gamma_i'(h)$ = a set of semivariogram models with the same mathematical formula

(e.g., a set of spherical models with identical ranges).

The condition of positive-definiteness is required to insure a unique solution for the cokriging system of equations [Luster, 1986]. For the case of two ReVs, model parameters are choosen so that

$$\det \begin{bmatrix} b_{DD}^i & b_{DN}^i \\ b_{ND}^i & b_{NN}^i \end{bmatrix} > 0 \tag{32}$$

where det[ ] is the determinant,

and

$$b_{DD}^i > 0$$

$$b_{NN}^i > 0$$

The selection of model parameters that insure positive definiteness becomes more difficult as the number of ReVs increases.

Methodology

The study area is an unconfined, basin-fill and alluvial aquifer in eastern Oregon (Figure I.1). The aquifer consists of sand, gravel, and silt, underlain by a clay layer at a depth that ranges from 5 to 20 meters (Gonther, unpublished data, 1985). Average annual precipitation in the region is 304 mm. Irrigation supplies approximately 617 mm/yr on agricultural land. It has been estimated that up to 80% of the recharge to groundwater in this region is from infiltrating surface water.

The aquifer is in an intensive agricultural production region. Principal crops are onions, alfalfa, beans and sugar beets. Soil nutrients are depleted due to intensive cropping practices and soil fertilizers are regularly applied. Application of pesticides to inhibit growth of competing species is also a common practice. In onion production it is typical to apply a pre-emergence broadleaf herbicide known commercially as Dacthal (DCPA). DCPA is incorporated into the soil surface before weed seeds have germinated and onion seedlings have emerged. Recommended DCPA application rates on onions are from 6.7 kg/ha to 10.1 kg/ha [Whitson, et al, 1987].

As part of a statewide drinking water evaluation, the Oregon Department of Environmental Quality (DEQ) and the United States Environmental Protection Agency (EPA) sampled domestic drinking water wells for a variety of chemicals several times from 1983 to 1986 (Figure I.1). 51 of the 108 wells sampled for nitrate exceeded the Oregon planning level of 5.0 ppm for nitrate as nitrogen (nitrate-N, referred to as nitrate in the remainder of this paper) (Pettit, unpublished data, 1987). 54 of the 81 wells sampled for DCPA had concentrations above the detection limit (.05 ppb). Measured concentration of DCPA approached 500 ppb in some wells, but were far below the health advisory level for DCPA of 3500 ppb [EPA, 1987].

The Herbicide Handbook of the Weed Science Society of America [1983] lists DCPA as having relatively low toxicity (oral Ld_{50} of greater than 3000 mg/kg for rats).

Studies of DCPAs' effect on reproductivity, teratogenicity, mutagenicity, carcinogenicity, and organ toxicity in test animals were all negligible [Extension Toxicology Network, 1988]. Aerobic microbial degradation is the main process controlling DCPA persistence in the soil; the half-life of the degradation process is between 45 and 90 days. The solubility of DCPA in water is 0.5 ppm (log octanol/water partition coefficient of 4.15). However, the metabolite, tetrachloroterephthalic acid (TTA), is believed to be much more water soluble than the parent compound. In water, DCPA and its metabolites are resistant to degradation between pH 5 and 9 [Extension Toxicology Network, 1988]. The metabolic pathway of the DCPA molecule is shown in Figure I.2. Apparently TTA is the dominant form of DCPA occuring in groundwater in this study area. No detailed information on the solubility, half-life, or toxicity of TTA was found in the literature.

A thorough site characterization of the study area was performed using soil, geologic, and topographic maps, published reports, well logs, and field reconnaissance surveys. The goal was to characterize groundwater flow patterns, aquifer properties and geometry, and potential contamination sources. Based on the results of the site characterization a smaller portion of the aquifer (16.5 km^2) was selected for detailed analysis (Figure I.3). The eastern borders of the study area are the Snake River and adjacent low-lying swamps, assumed to be groundwater flow divides. The southern and western boundaries of the study area were selected because of the presence of numerous deep drainage ditches thought to interupt groundwater flow in those directions. The Malhuer River forms the northern study area boundary. The natural extents of the aquifer form the remaining boundaries.

For this study, 70 additional groundwater samples were taken from existing domestic drinking water wells during August and September, 1987. The drinking water samples were tested for nitrate and DCPA. The location of wells sampled for this study and for the previous DEQ/EPA study are in Figure I.3. The geographic location, depth of

well and name of the original well owner were recorded and well logs were obtained for some of the wells.

Table I.1 lists the number, coordinates, and measured nitrate and DCPA concentrations for each well sampled. Two separate water samples were taken from each well. Nitrate samples were stored in plastic 100 ml containers and refrigerated at approximately 0° C until analyzed. DCPA samples were stored in 2L glass bottles. The nitrate analysis was conducted using an Alchem RFA autoanalyzer. Calibration was performed using KCI standards of 5, 10, and 20 ppm. Sensitivity of the analyzer is ± 0.2 ppm on 20 ppm full scale, and ± 0.02 on 2 ppm full scale.

All DCPA metabolites in a water sample were recovered as the parent compound and then analyzed using an Electron Capture Gas Chromatograph (ECGC). Percent recovery, measurement error, and measurement repeatibility were recorded in the laboratory procedures. One DCPA and nitrate sample had multiple analyses performed to determine the magnitude of laboratory measurement variance (Table I.2 and Table I.3). To determine the amount of DCPA that was recovered in an analysis, known water samples with concentrations of 1.0 ppb and 10.0 ppb were analyzed repeatedly (Table I.2). Based on average recoveries for ranges of DCPA concentrations, reported laboratory results were adjusted to 100% recovery. No adjustment was required for the nitrate samples.

DCPA and nitrate concentrations were transformed into contaminant densities by multiplying contaminant concentrations by the assumed values for contaminated thickness and aquifer porosity (Table I.1). This transformation is required to preserve the additivity of ReVs required by linear geostatistics [Cooper and Istok, 1988a,b] and is useful for computing global estimates. Well logs provided sufficient information on aquifer depth and screened depth of wells. A contaminated thickness of 3.2 m was assumed for all wells. This was the most common screened depth recorded on the well logs. A standard porosity of 35% was assumed based on the aquifer descriptions in the well logs.

The correlation coefficient, ρ_{ND}, [Journel and Huijbregts, 1978] was computed using

$$\rho_{ND} = \frac{C_{ND}(0)}{\sqrt{C_{NN}(0)\ C_{DD}(0)}} \tag{33}$$

where $C_{NN}(0)$, $C_{DD}(0)$, and $C_{ND}(0)$ are the sample variances for nitrate and DCPA and the sample covariance, respectively. If the value of $\rho_{ND} < 1/2$, then the probability that the two ReVs are not cross-correlated becomes large. A reduction in estimation variance will still be achieved using cokriging [Journel and Huijbregts, 1978], but the reduction in estimation variance would be small and the additional effort required for a multivariate analysis would probably not be justified.

Experimental direct-semivariograms were calculated for log(nitrate), and log(DCPA) and an experimental cross-semivariogram was calculated for log(DCPA) and log(nitrate). Average values of $\gamma^*(h)$ were computed for groups of pairs of sample points, where each group had 30 pairs [Istok et al, 1988]. Using the linear model of coregionalization, spherical models with similar ranges were fit to the experimental semivariograms.

The suitability of model semivariograms was assessed using cross-validation (jack knifing). In cross-validation, measured concentrations are removed one sample point at a time. The concentration at the point is then estimated using kriging or cokriging. The differences between the measured and estimated concentrations at the sample points are used to compute three statistics, the average kriging error (AKE), mean squared error (MSE), and standardized mean squared error (SMSE) [Delhomme, 1978]. Initial estimates for model semivariograms were obtained by visual inspection of the experimental semivariograms. The parameters were then adjusted by trial-and-error to improve the cross-validation statistics. Cross-validation was performed using the method

of sliding neighborhoods with a radius of 4 km. The maximum radius for which cokriging can be performed is defined as one-half the maximum distance between the groups of data pairs i.e., $h_{max}/2$. This was calculated to be 5 km for γ^*_{NN}, 4 km for γ^*_{DD}, and 4 km for γ^*_{DN}.

Results

Histograms of the DCPA and nitrate densities suggested that both variables were lognormally distributed. The probabilities that densities were lognormal was computed using the univariate Shapiro-Wilkes Statistic [SAS, 1985]. The results were 0.932 for log(DCPA) and 0.735 for log(nitrate). Based on these results the univariate distributions for both ReVs were assumed to be lognormal. It was also assumed that the joint distribution between DCPA and nitrate was lognormal. This assumption seems reasonable based on the univariate normality displayed by each transformed contaminant, the correlation of the ReVs, and on earlier studies [Cooper and Istok, 1988 a,b; Myers et al, 1984] that suggested groundwater contaminants are lognormally distributed.

The correlation coefficient for DCPA and nitrate, ρ_{DN} (33), was 0.74, indicating that a reduction in estimation variance would occur if log(DCPA) was estimated with log(DCPA) and log(nitrate) support.

The three experimental semivariograms all exhibited a nested structure consisting of nugget and transition structures. Theoretically, experimental semivariograms should approach zero as the distance between pairs approaches zero. In practice this rarely occurs due to the presence of measurement error and small scale variability. The effect of laboratory measurement error on the magnitude of the nugget was assesed using the results of the laboratory measurement repetitions. Using the natural logarithm of the repeatibility data in Table I.2, the variance of log(nitrate) and log(DCPA) as contaminant densities were 0.009 and 0.166 respectively. This corresponded to 1.1% of the nugget for log(nitrate) and 22.4% for log(DCPA).

The transition structures were fit with spherical models with sills equal to the sample variance or covariance and a range of 4 km (Figure I.4). The existence of a unique solution to the cokriging system was ensured by requiring γ_{DD}, γ_{NN}, and γ_{DN} to

be positive-definite. The determinants computed using (32) were 0.55 and 0.02 for the nuggets and sills, respectively of γ_{NN}, γ_{DD}, and γ_{DN}.

The direct-semivariograms for nitrate and DCPA had AKE's of -0.06, -0.10, MSE's of 1.66, 1.24, and SMSE's of 1.24, and 0.94 respectively (Figure I.4). The values of AKE, MSE, and SMSE for the cross-semivariogram were -0.08, 0.76, and 0.84 for the estimation of log(DCPA), and -0.03, 1.36, and 1.15 for the estimation of log(nitrate). The AKE and MSE should have values close to zero and one [Delhomme, 1978]. When fitting a model, model parameters should be be selected to minimize the MSE while constraining the SMSE to be within the interval $1 \pm 2\{[2/(2n)]^{1/2}\}$ [Delhomme, 1978]. For this study, the SMSE range was 1.00±0.31 for nitrate, and 1.00 ± 0.35 for DCPA.

Point estimates and estimation variances for log(nitrate) and log(DCPA) were obtained by kriging (Figures I.5 and I.6). The inverse-transform was performed on the estimates correcting for bias using procedures in Journel and Huijbregts [1978, p 468-471]. The boundaries on the contour maps are straight-line approximations of the actual study area boundaries (Figure I.3). The maximum point estimate for nitrate densitiy was 2.87×10^4 mg/m^2 (equivalent to a concentration of 26.9 ppm) and the maximum estimation variance was 1.2×10^8 (mg/m^2)^2 (equivalent to a standard deviation of 10.26 ppm). The maximum estimate for DCPA density was 310 mg/m^2 (290.5 ppb). The maximum point estimation variance was 5550.0 (mg/m^2)^2 (74.5 ppb).

The location of the drinking water standard of 10 ppm (10670 mg/m^2) and the Oregon planning level of 5 ppm (5335 mg/m^2) for nitrate can be estimated from Figure I.5. The shaded region of Figure I.6 identifies the area that has 95% probability of exceeding the planning level for nitrate. The minimum 95% confidence interval for point estimates is calculated using (29) [Journel and Huijbregts, 1978].

Cokriging was performed on DCPA with nitrate support (Figure I.8). The maximum estimate was 358 mg/m^2 (355.5 ppb). The maximum estimation variance was

4690 $(mg/m^2)^2$ (64.1 ppb). None of the estimates or the 95% confidence intervals calculated using (29), exceeded the health advisory level of 3500 ppb.

The largest estimation variances for nitrate, kriged DCPA, and cokriged DCPA were consistently located along the boundaries of the study area, resulting from the small number of sample data in these locations. These locations also correspond to the location of additional samples identified by the method of fictitious points selected (see below). The maximum estimates for nitrate, and DCPA occured west of the city of Ontario. This region is an area of heavy nitrate and DCPA use, and can be viewed as a concentrated, non-point source.

Cokriging had varying effects on the DCPA estimates (Figure I.9a). Estimates were reduced up to 97 mg/m^2 (90 ppb) and increased up to 103 mg/m^2 (96 ppb), however estimation variances were reduced for the entire study area (Figure I.9b). The minimum relative gain , as defined in (30) was 14% and the maximum relative gain was 34% (Figure I.10).

Global estimates were made for dissolved nitrate and DCPA in the portion of the aquifer studied using the procedure described in Istok and Cooper [1988]. The global estimate was 2.478 x 10^6 kg for nitrate and 21962 kg for DCPA. The EPA Health Advisory for DCPA lists a calculated partition coefficient for octanol/water of $10^{4.5}$. If the organic content of the soil and the percent of maximum solubility of DCPA in the aquifer is known, the equilibrium partition coefficient and the distribution coefficient could be calculated [de Marsily, 1986]. The distribution coefficient could be used to estimate the sorbed nitrate and DCPA in the aquifer. However due to the limited amount of information on organic matter content and extent of maximum solubility, these global estimates only reflect the amount of DCPA and nitrate dissolved in the aquifer. Based on recommended DCPA application rates of 10 kg/ha [Whitson et al, 1987] and assuming that 30% of the study area (4.95 km^2) has DCPA applied in any year (i.e., is used to produce onions), the

global estimate of 21962 kg is equivalent to 4.4 years of DCPA application. A similiar estimate for nitrate was not possible due to widely varying application rates.

Optimum locations for nitrate and DCPA samples were selected using the cokriging estimation variances and the fictitious point method. Maximum estimation variance was used as a basis for the selection of additional sample locations in Table I.4. The location of the selected fictitious points are shown in Figure I.8b as open circles. The fictitious nitrate and DCPA sample points had similar effects on reducing maximum estimation variances (Figure I.11a). Additional fictitious DCPA samples reduced average estimation variances more effectively than additional fictitious nitrate points (Figure I.11b). This trend is relatively constant through the first 5 additional fictitious sample points.

The benefit of additional DCPA and nitrate samples can be expressed as $R(x)/CR$ where $R(x)$ is the relative gain (30) and CR is the DCPA : nitrate measurement cost ratio. As the cost increases to 2.5:1 (DCPA : nitrate) the additional fictitious nitrate samples were more beneficial than the additional fictitious DCPA samples until the fifth sample (Figure I.11c). At cost ratios of 5:1 and 10:1 additional fictitious nitrate samples were more beneficial in reducing estimation variances than additional fictitious DCPA samples through the fifth sample.

Conclusions

Previous sampling of domestic water supply wells in eastern Oregon identified an area near Ontario with nitrate concentrations exceeding the Oregon state planning level (5 ppm) and the federal drinking water standard (10 ppm). A pesticide, DCPA, was also found in the drinking water samples at levels approaching 500 ppb, but not exceeding the EPA health advisory level of 3500 ppb. For this study a 16.5 km^2 portion of the area near Ontario was selected. Geographical features believed to affect the extent of groundwater contamination (rivers, drainage ditches, aquifer boundaries) were used to define the extents of the study area.

Within the study area, DCPA and nitrate concentrations were measured in 70 wells and the concentrations were converted to contaminant densities (mg/m^2) assuming a constant porosity of 0.35 and an average contaminated aquifer thickness of 3.2 m. Porosity and contaminated thickness could also be considered. However, due to a lack of data for porosity and contaminated aquifer thickness, constant values were assumed. The result of assuming constant values for porosity and contaminated aquifer thickness is that the contaminant concentration and the contaminant density distributions are equivalent. If non-constant values were used for either parameter it is possible that the contaminant density distribution might not be lognormally distributed.

The small size of the data sets did not allow the type of underlying distribution for contaminant densities to be clearly distinguished, however, the data appeared to be lognormally distributed. A natural logarithmic transformation was performed on the contaminant densities and assumed to yield univariate normal distributions. The log(nitrate) and log(DCPA) distributions were also assumed to be joint-normal. The assumptions of univariate normal distributions and joint-normal distributions for the transformed data are required for calculating the estimation variance [Journel and Huijbregts, 1978]. Consequently, if these assumptions were not valid, the calculated

95% confidence limits and the estimation variances were not valid. The restriction of having a small number of sample data, assuming constant porosity, and contaminated aquifer thickness severely limits the ability to distingush how nitrate and DCPA contaminant densities are distributed in groundwater.

Many combinations of model direct- and cross-semivariograms had acceptable cross-validation statistics, making it difficult to identify the 'best' model to represent the spatial distribution for nitrate and DCPA. The computed estimates and the estimation variances will vary with different models. The cross-validation statistics are only guidelines for determining the suitability of the model for kriging and cokriging. In order to make the procedure less subjective more work is needed to develope methods for assessing the quality of semivariogram models..

The kriging estimates for nitrate indicated that a substantial portion of the study area exceeds the 5 ppm nitrate planning level. Using the estimation variances, an area that has 95% probability of exceeding 5 ppm nitrate level was approximately 2.9 km^2 (17.5%) of the study area. As estimation variances are used to calulate the 95% probability, the size of this area is a function of the validity of the assumption that the transformed contaminant densities have normal and joint-normal distributions and on the choice of the direct semivariogram model for nitrate.

Point estimates for DCPA obtained by kriging and cokriging were substantially different on the outer regions of the study area. These regions had several nitrate samples but few DCPA samples. DCPA point estimates were reduced up to 103 mg/m^2 by cokriging. The relative gain from cokriging over kriging was computed using estimation variances. The minimum and maximum relative gains were 14% and 34%, respectively.

Global estimates were made of the dissolved portion of DCPA and nitrates in the study area. The global estimate was 21962 kg for DCPA and 2.478 x 10^6 kg for nitrate. Assuming that 30% of the study area has DCPA applied at a rate of 10.1 kg/ha each year, the global estimate reflects 4.4 years of DCPA application that is dissolved in the aquifer.

Using the method of fictitious points, the contribution of additional nitrate and DCPA samples to decreasing cokriging estimation variances were compared. In this study there was relatively little difference in the reduction in the maximum DCPA estimation variances obtained with additional nitrate or DCPA samples. However, additional DCPA samples reduced the average DCPA estimation variance more than additional nitrate samples.

When the benefit of an additional sample was considered, additional nitrate samples were superior to additional DCPA samples at a DCPA : nitrate cost ratio of 2.5 : 1 and greater. This relationship held for the first four additional fictitious samples at a ratio cost of 2.5 : 1, and for the first five additional fictitious samples at cost ratios above 2.5 : 1.

This case history has demonstrated that multivariate geostatistical analysis is an applicable and useful tool in estimating undersampled DCPA using nitrate and DCPA support. It does not however show that the cokriging system can be applied to every groundwater contamination case. This study describes the assumptions necessary, in this case, to apply multivariate geostatistics to groundwater contamination. Clearly each groundwater contamination case should be thouroughly evaluated to see if it meets the special criteria that allow geostatistics to be applied. A future possibility that could lead to extended application of geostatistics to groundwater contamination cases is the recently developed theory of non-parameteric geostatistics [Journel and Issaks, 1984]. Until more information on this procedure becomes available standard multivariate geostatistics are the recommended method.

TABLE I.1. Well data from the study area

Well	Coordinates		Concentrations		Densities	
	X	Y	nitrate*	DCPA	nitrate*	DCPA
	(km)		(ppm)	(ppb)	(mg/m²)	
1	22.36	20.97	13.60	93.60	14511.20	99.87
2	22.92	21.04	17.70	24.40	18885.90	26.03
3	22.65	25.56	24.50	400.00	26141.50	426.80
4	23.23	25.46	3.20	10.50	3414.40	11.20
5	24.10	25.65	12.60	18.90	13444.20	20.17
6	24.49	25.45	0.20	10.60	231.40	11.31
7	23.29	24.04	13.60	41.00	14511.20	43.75
8	19.92	27.54	1.00	4.40	1067.00	4.69
9	19.51	25.84	10.40	107.00	11096.80	114.17
10	20.04	24.00	19.90	114.00	21233.30	121.64
11	19.52	24.02	18.00	50.50	19206.00	53.88
12	21.92	23.10	20.80	169.00	22193.60	180.32
13	21.79	22.24	9.90	55.80	10563.30	59.54
14	20.13	22.58	21.60	103.00	23047.20	109.90
15	20.17	23.22	14.40	126.00	15364.80	134.44
16	20.96	21.96	9.30	-999.00†	9923.10	-999.00
17	22.72	19.45	0.41	-999.00	437.47	-999.00
18	24.55	29.69	11.80	-999.00	12590.60	-999.00
19	20.29	28.38	12.00	-999.00	12804.00	-999.00
20	20.04	24.44	12.00	31.00	12804.00	33.08
21	26.29	26.89	11.00	5.10	11737.00	5.44
22	24.60	28.73	0.24	0.48	256.08	0.51
23	24.91	28.91	6.20	-999.00	6615.40	-999.00
24	24.01	27.79	12.60	-999.00	13444.20	-999.00
25	23.62	28.94	17.20	-999.00	18352.40	-999.00
26	23.41	28.78	16.00	39.00	17072.00	41.61
27	23.26	28.04	25.30	431.00	26995.10	459.88
28	23.69	27.73	16.60	150.00	17712.20	160.05
29	23.89	27.23	9.80	-999.00	10456.60	-999.00
30	21.84	27.72	22.20	180.00	23687.40	192.06
31	21.57	27.47	23.30	188.00	24861.10	200.60
32	22.08	26.79	32.00	245.00	34144.00	261.41
33	21.79	25.97	26.50	194.00	28275.50	207.00
34	21.46	26.33	21.50	316.00	22940.50	337.17
35	20.71	26.28	27.40	213.00	29235.80	227.27
36	23.77	25.99	4.40	-999.00	4694.80	-999.00
37	23.81	26.86	12.30	121.00	13124.10	129.11
38	22.33	25.20	30.60	236.00	32650.20	251.81
39	23.64	25.41	0.36	8.00	384.12	8.54
40	20.77	25.45	11.00	65.00	11737.00	69.35
41	21.59	23.48	16.50	46.00	17605.50	49.08
42	19.42	26.25	1.30	-999.00	1387.10	-999.00

* Nitrate expressed as nitrogen
† Missing samples identified by -999.00

TABLE I.2. Recovery and repeatability of DCPA laboratory analysis

	Repeatability (ppb)	Recovery	
		1.0 (ppb) analysis results (%)	10.0 (ppb) analysis results (%)
	8.2	91.0	87.0
	7.0	67.0	87.0
	6.5	78.0	86.0
	7.6	85.0	96.0
	7.1	104.0	97.0
	6.3	111.0	
mean	7.1	89.3	89.1
variance	0.49	144.0	16.0

Table I.3 Repeatability of nitrate* laboratory analysis

	Repeatability (ppm)
	1.0
	2.4
	2.7
	2.4
	2.4
mean	2.18
std.dev	0.47

* Nitrate expressed as nitrogen

TABLE I.4 Locations of additional fictitious samples and their effect on estimation variance

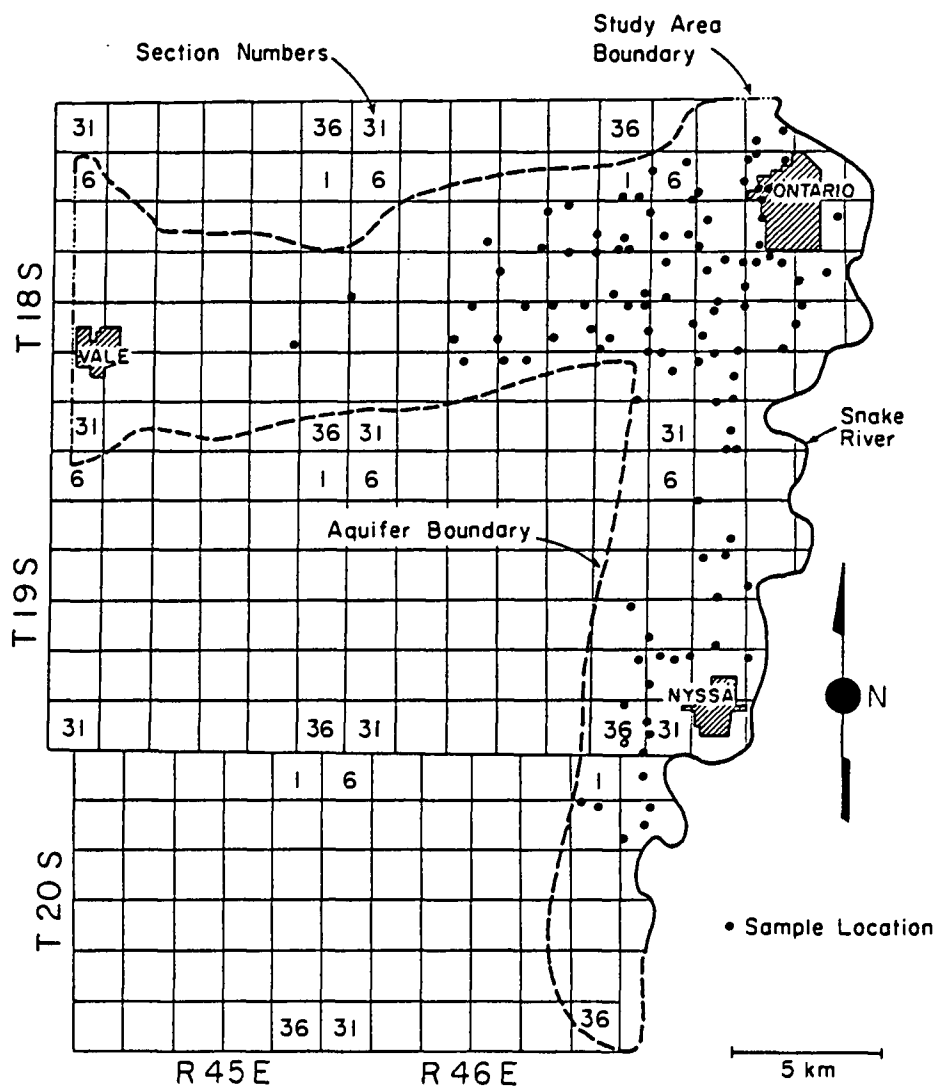| fictitious samples	X (km)	Y (km)	nitrate* σ_{max}^2 $\log((mg/m^2)^2)$	σ_{avg}^2	DCPA σ_{max}^2 $\log((mg/m^2)^2)$	σ_{avg}^2
0			1.524	1.100	1.524	1.100
	1	26.19	25.60	1.400	1.094	1.400
1.090						
2	21.09	28.90	1.265	1.087	1.359	1.079
3	24.39	24.10	1.345	1.083	1.345	1.072
4	21.09	21.10	1.325	1.080	1.296	1.065
5	22.59	19.60	1.270	1.079	1.252	1.061

* Nitrate expressed as nitrogen

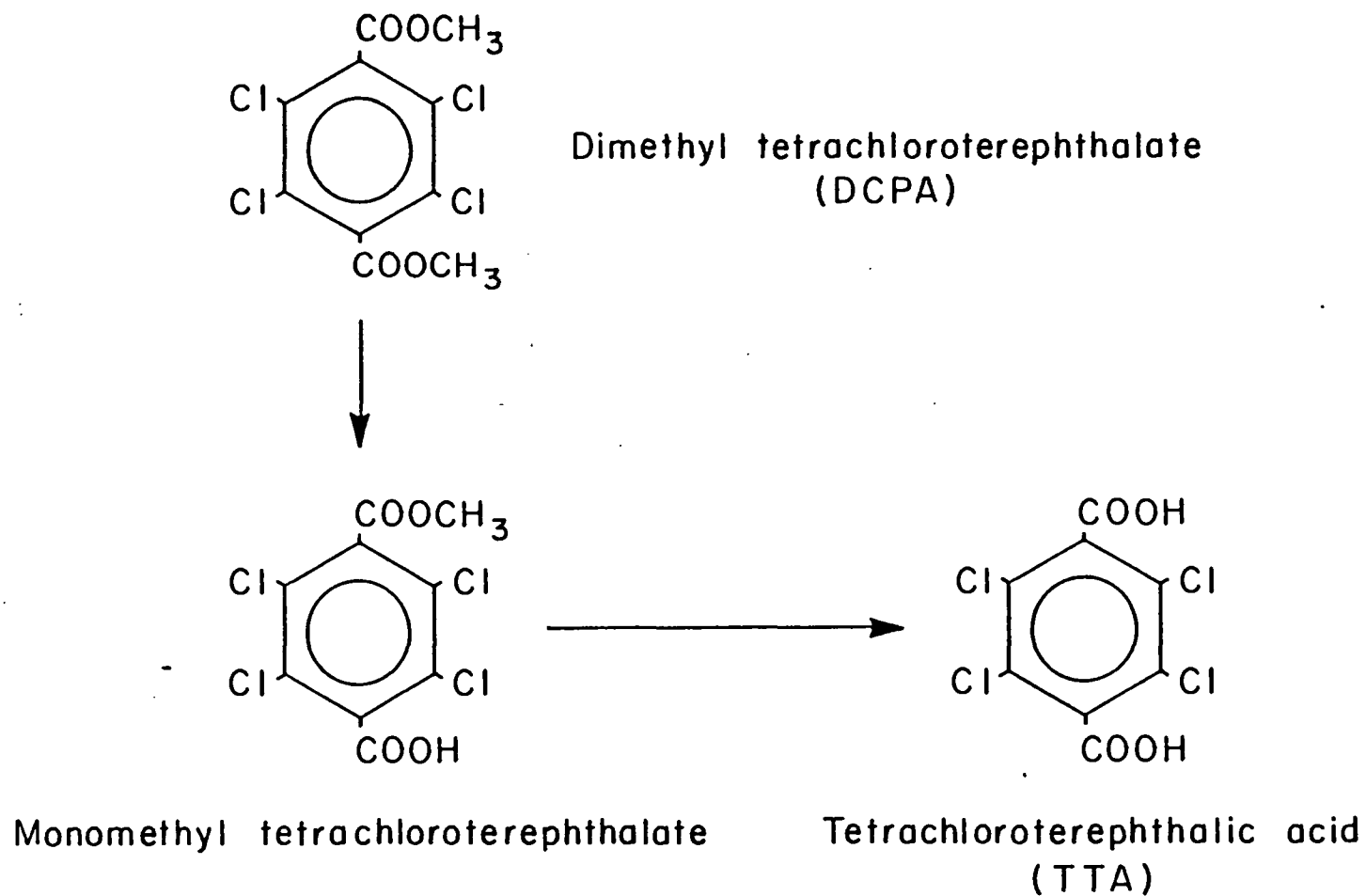Figure I.1. Location of previously sampled wells.

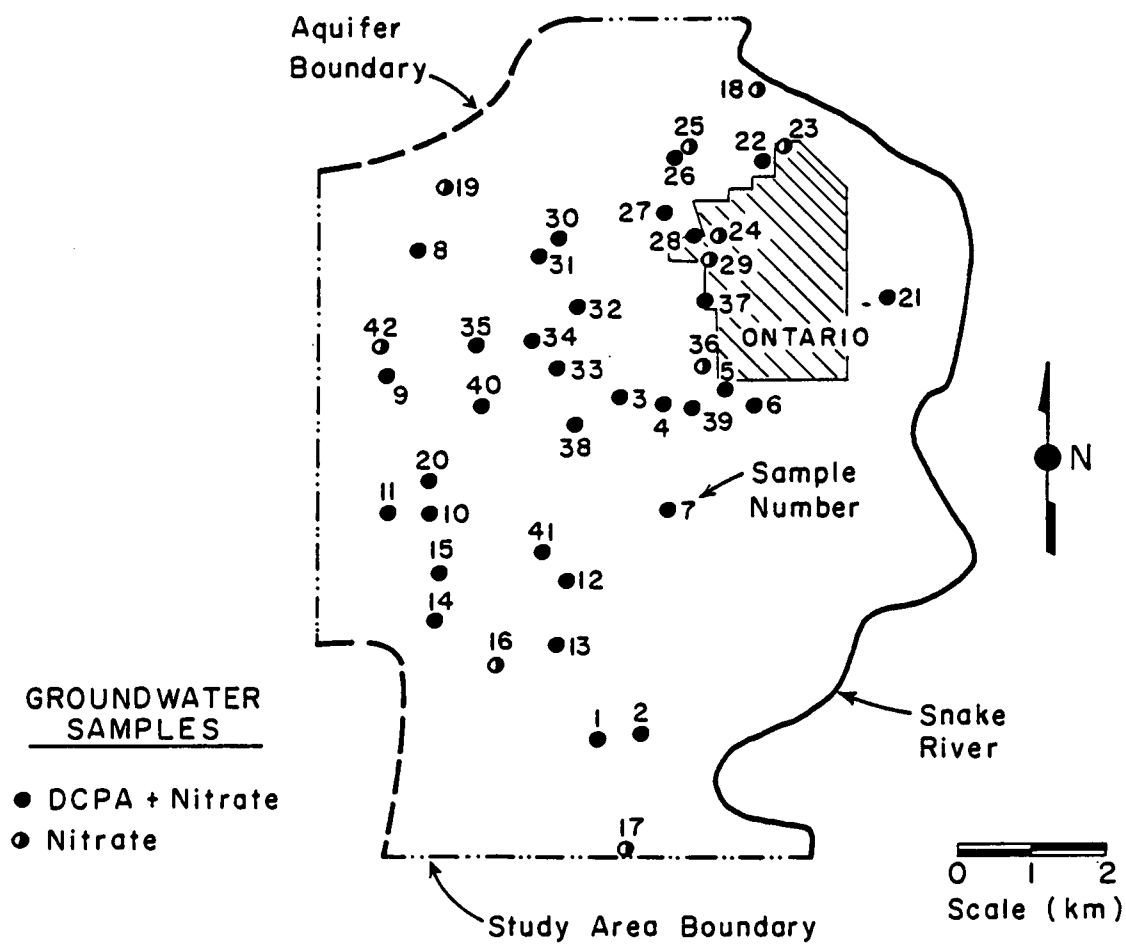Figure I.2. Partial degradation pathway for the DCPA molecule.
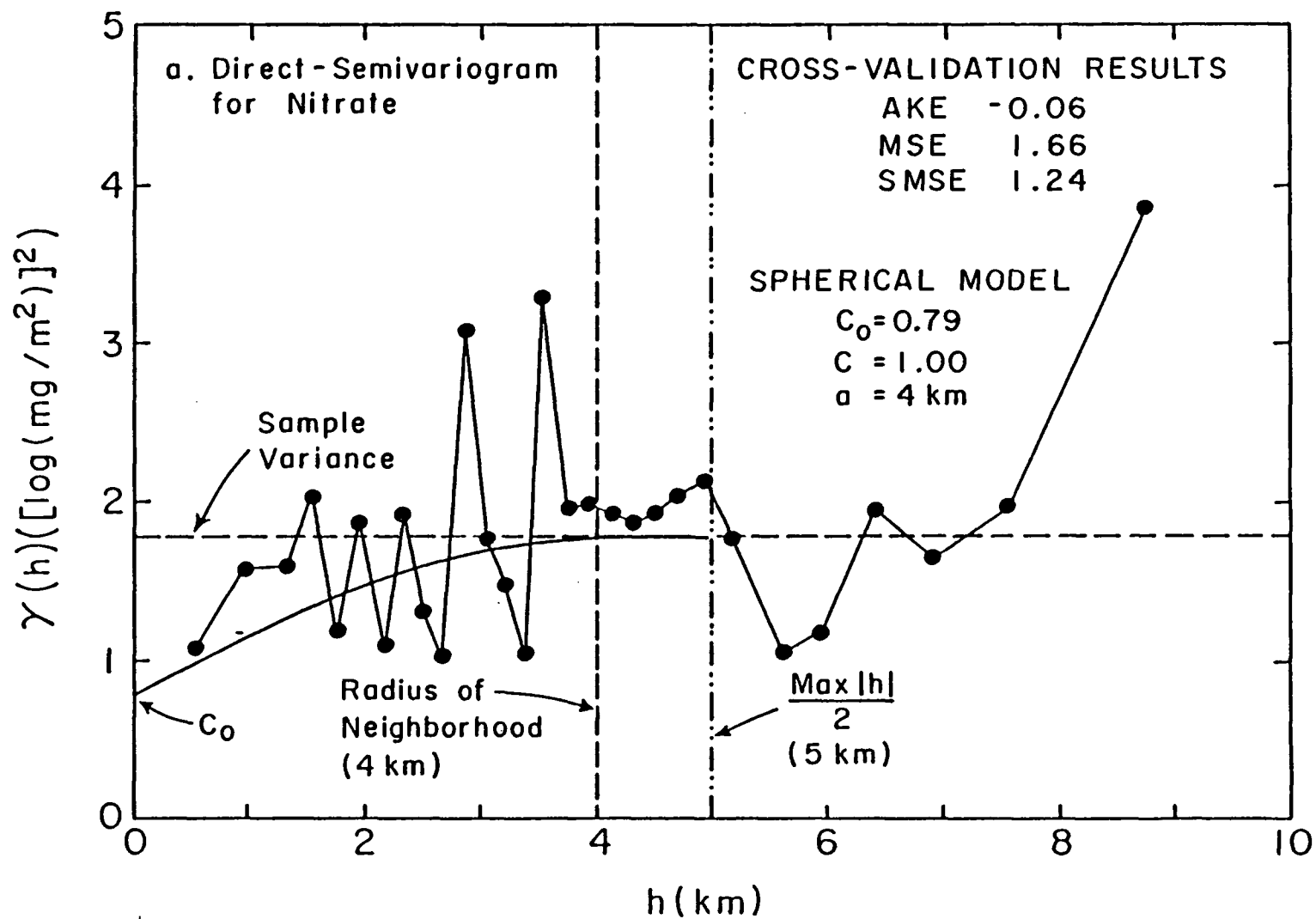
Figure I.3. The study area.

Figure I.4. Experimental semivariograms, fitted models, and cross-validation results.
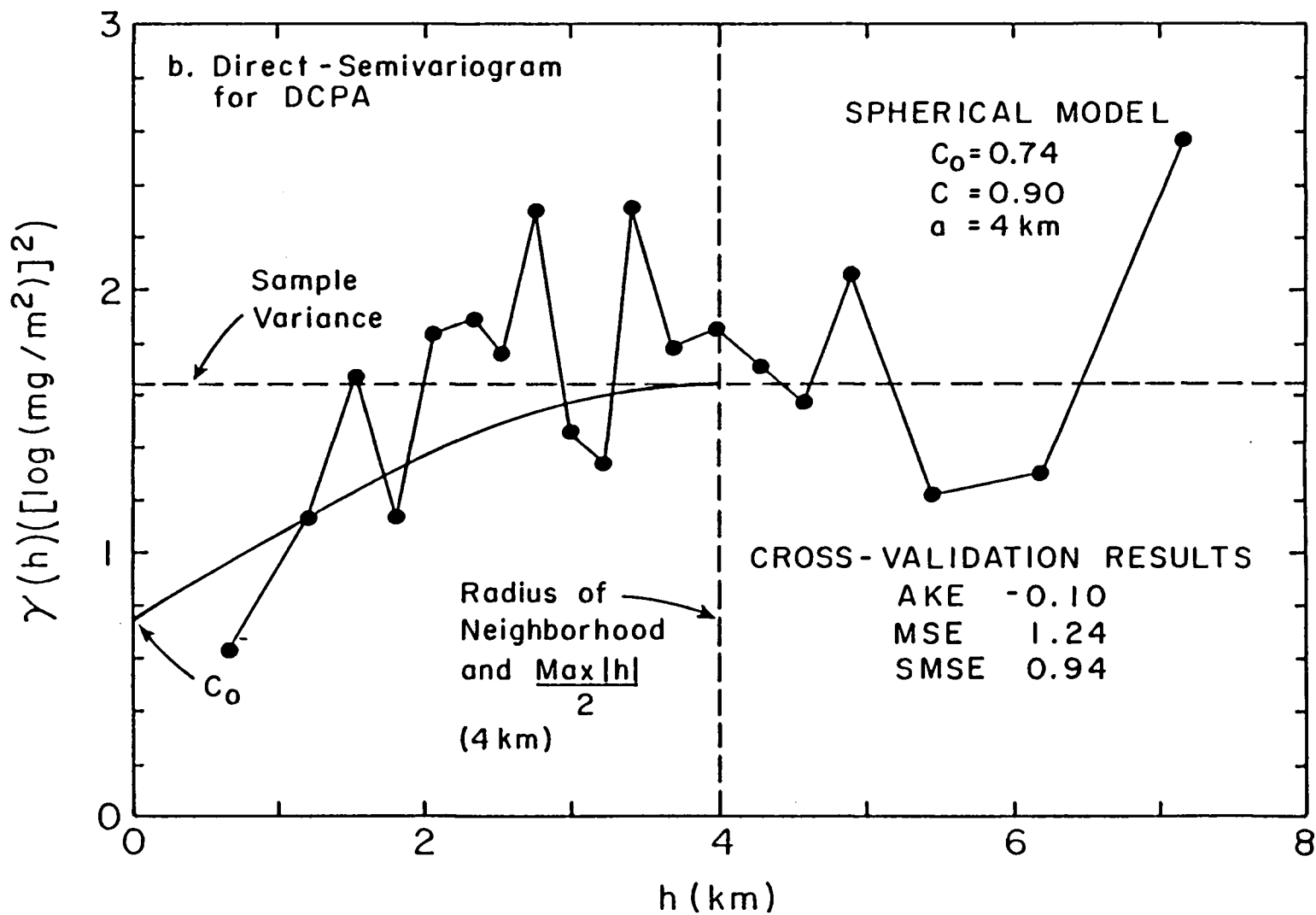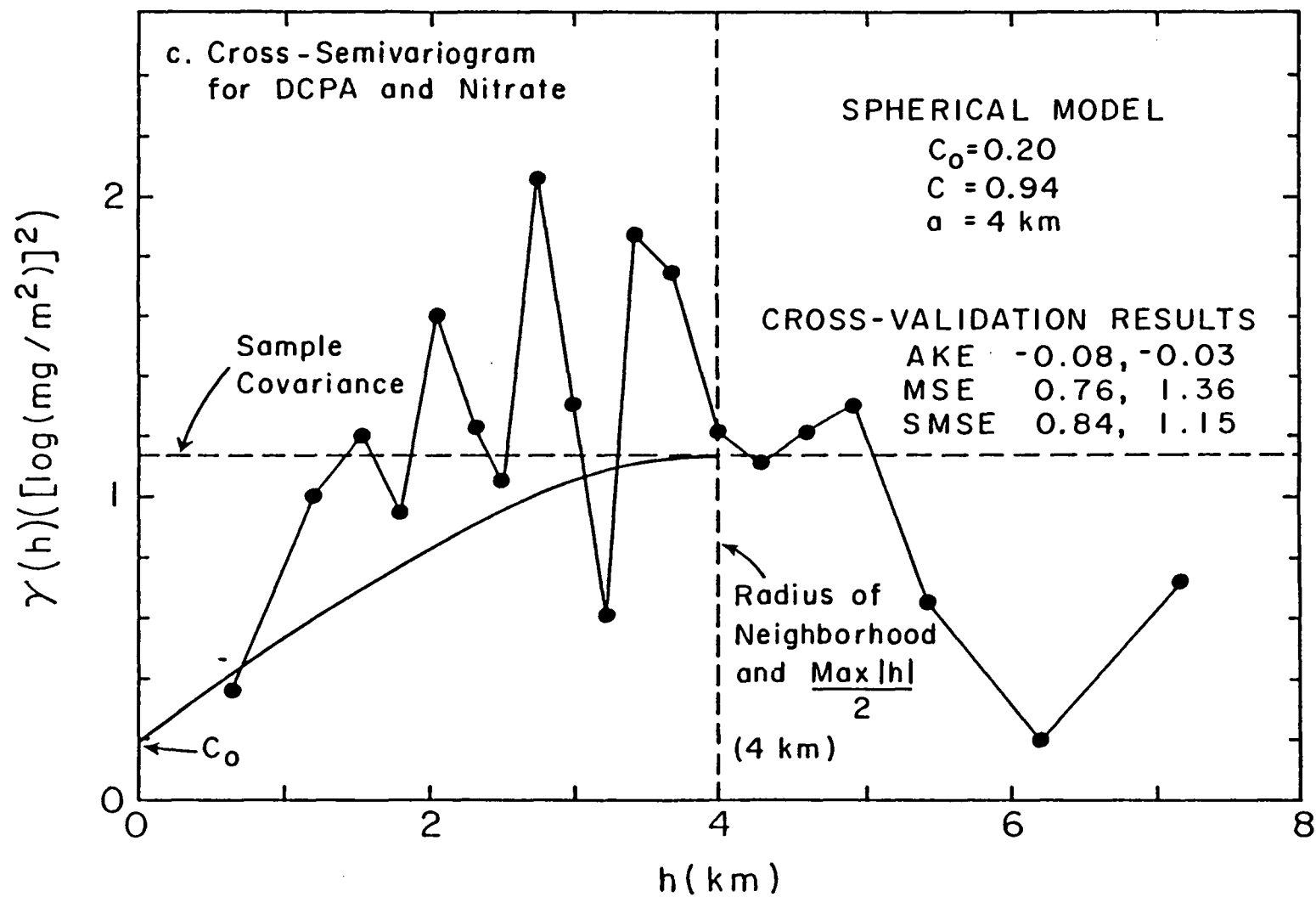
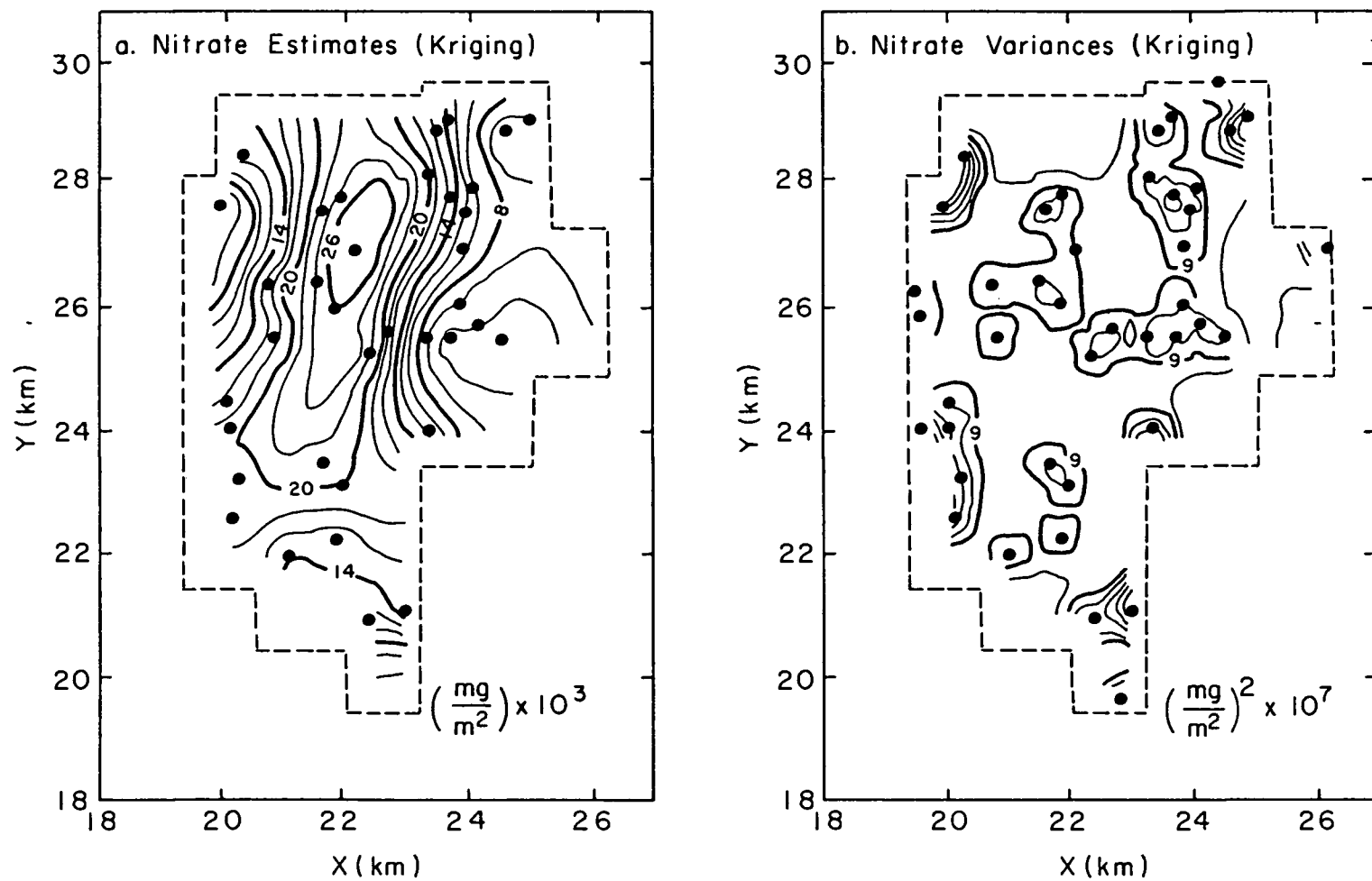Figure I.4. (continued)

Figure I.4. (continued)

Figure I.5. Contour map of nitrate kriging estimates, estimation variances, and sample locations.
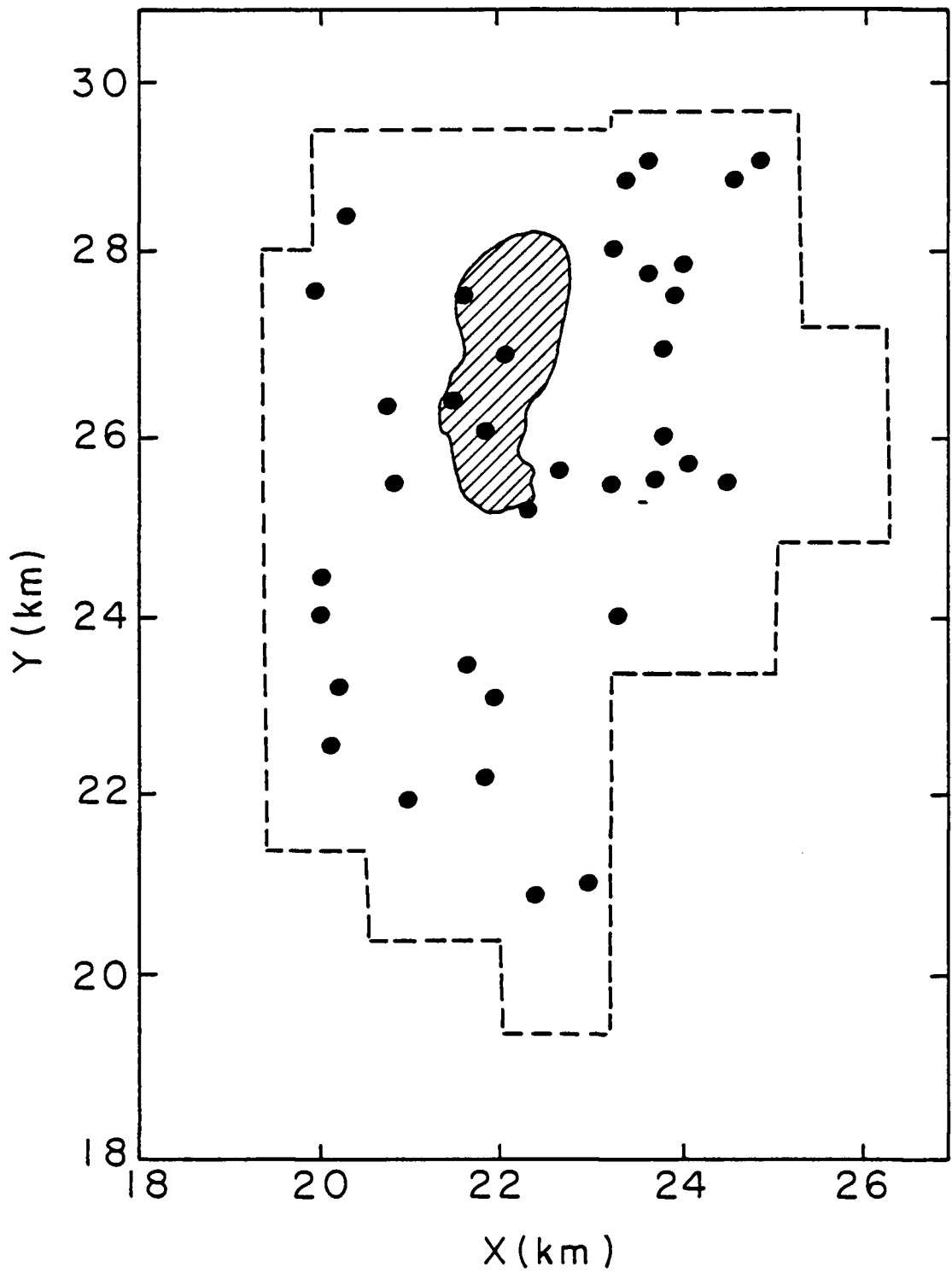
Figure I.6. Area exceeding Oregon nitrate planning level of 5 ppm (5335 mg/m^2) with 95% probability.
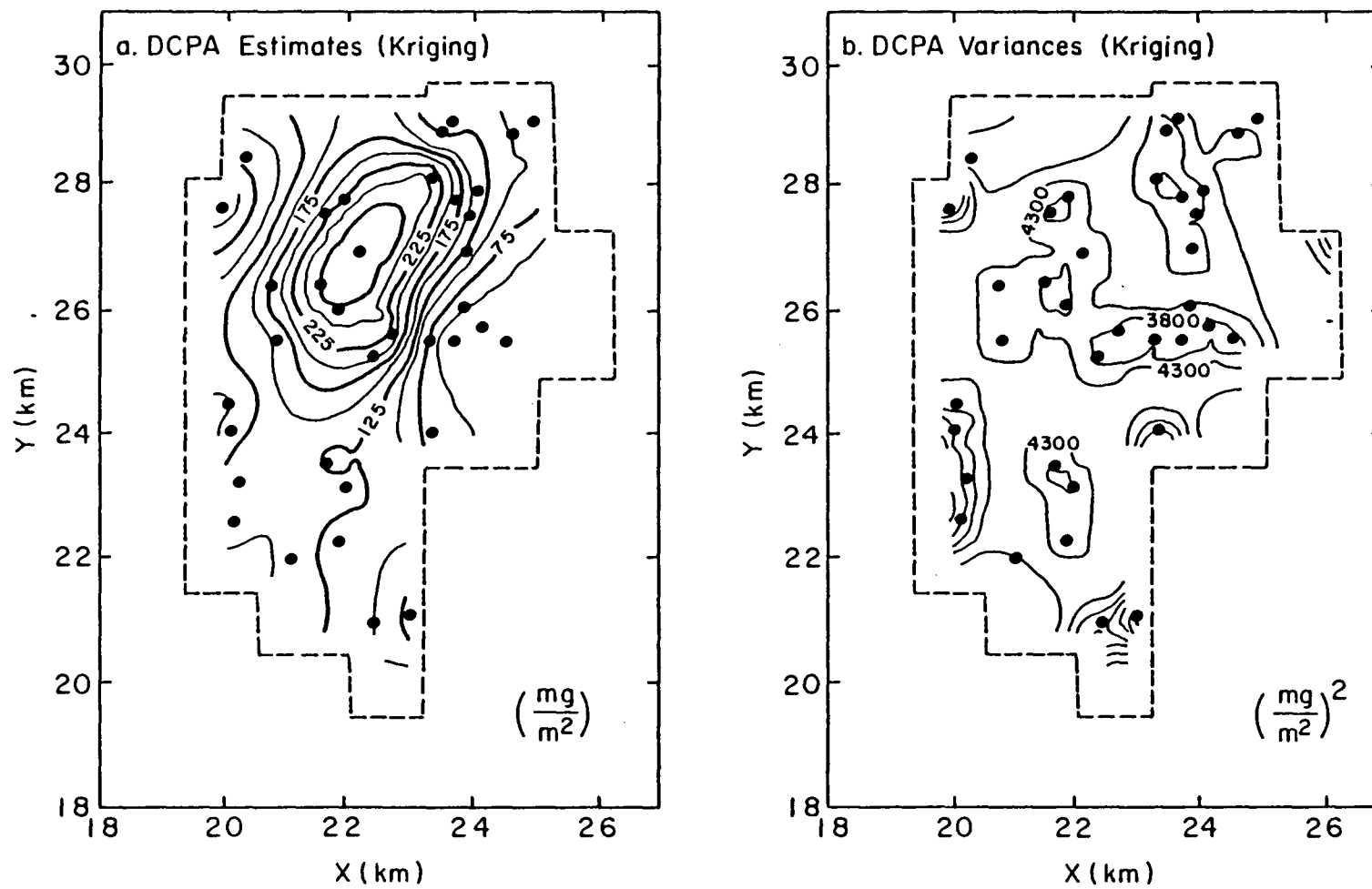
Figure I.7. Contour map of DCPA kriging estimates, estimation variances, and sample locations.
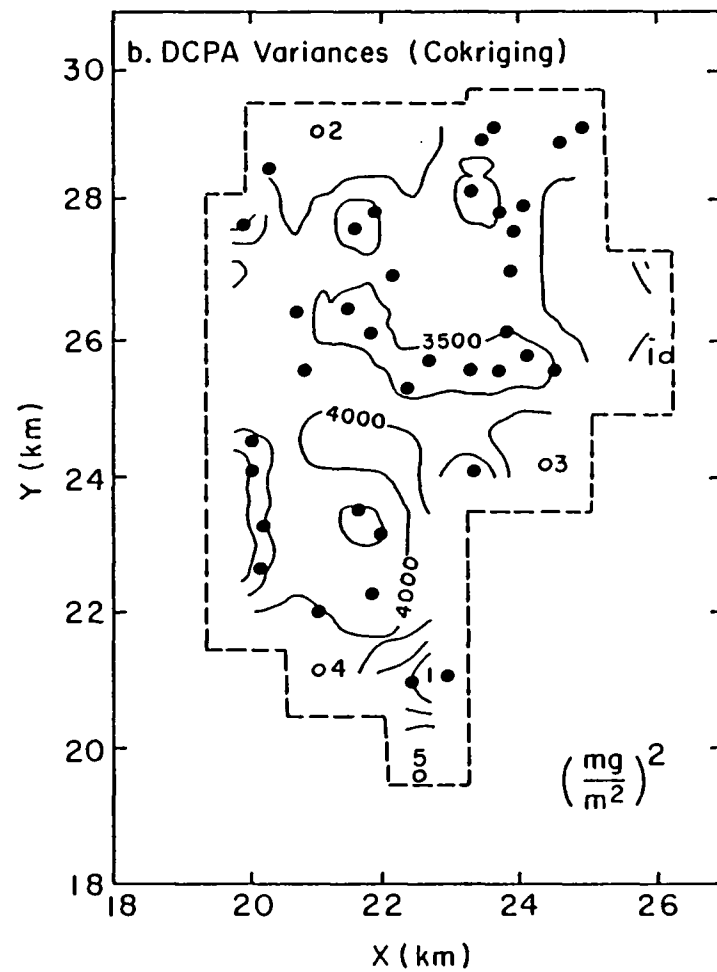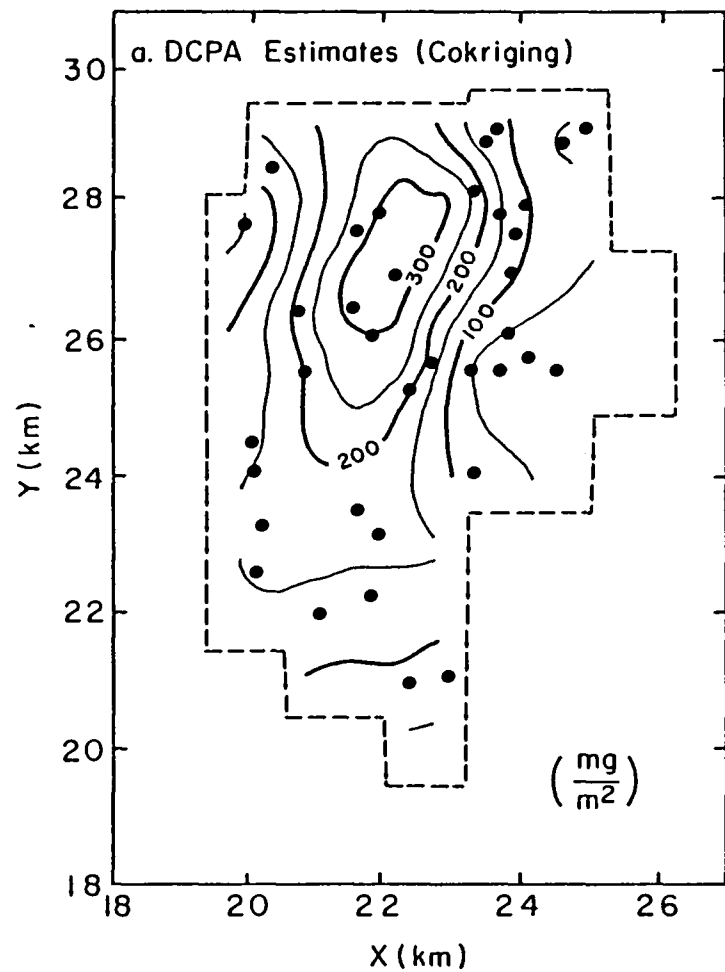
Figure I.8. Contour map of DCPA cokriging estimates, estimation variances and fictitious sample locations.
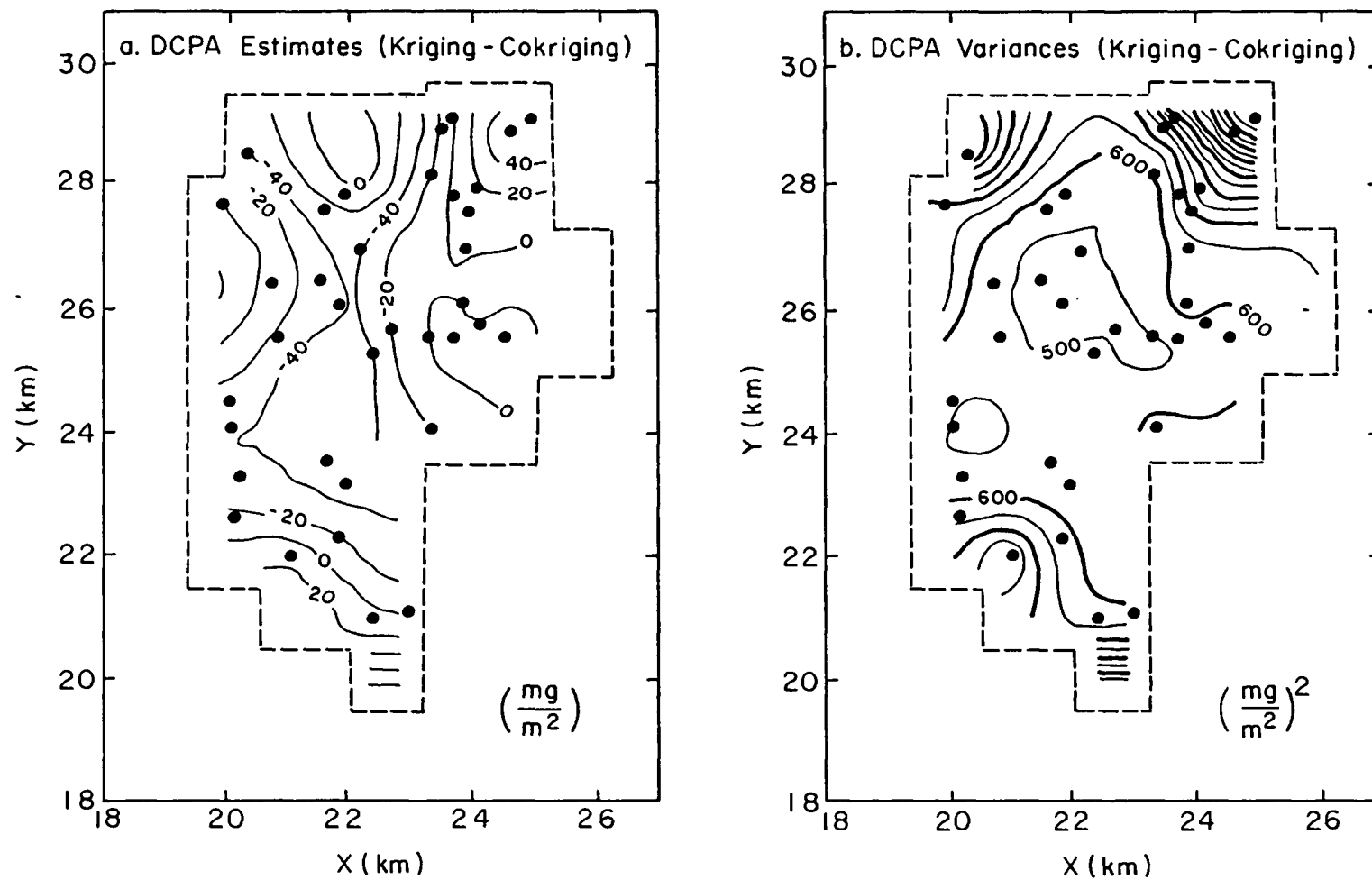
Figure I.9. Contour map of the diffference between DCPA kriging and cokriging estimates and estimation variances.

Figure I.10. Contour map of relative gain from cokriging for DCPA.

Figure I.11. Reduction in the maximum estimation error (a), average estimation error (b), and benefit from additional fictitious samples (c).

Figure I.11. (continued)
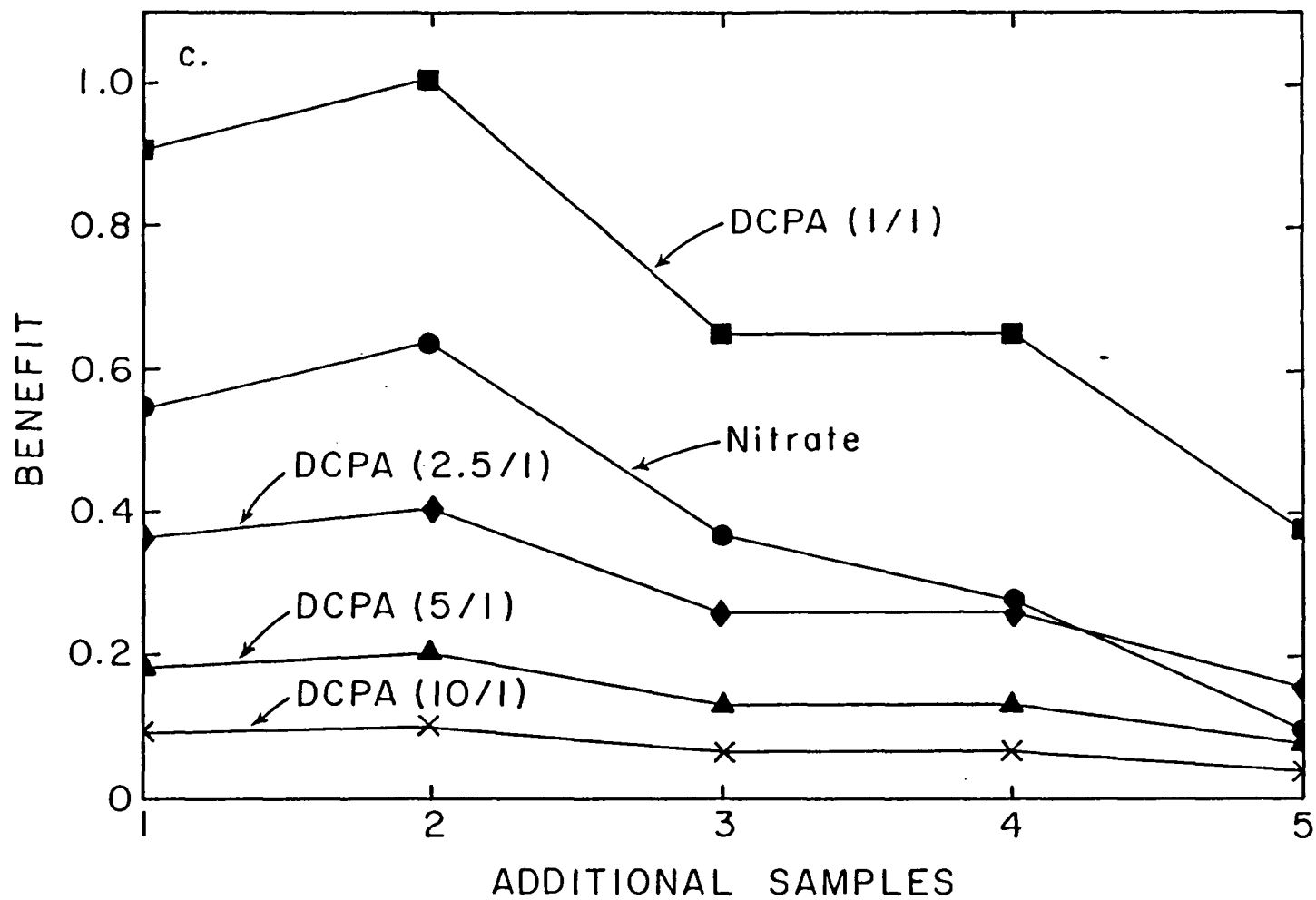
Figure I.11. (continued)

Notation

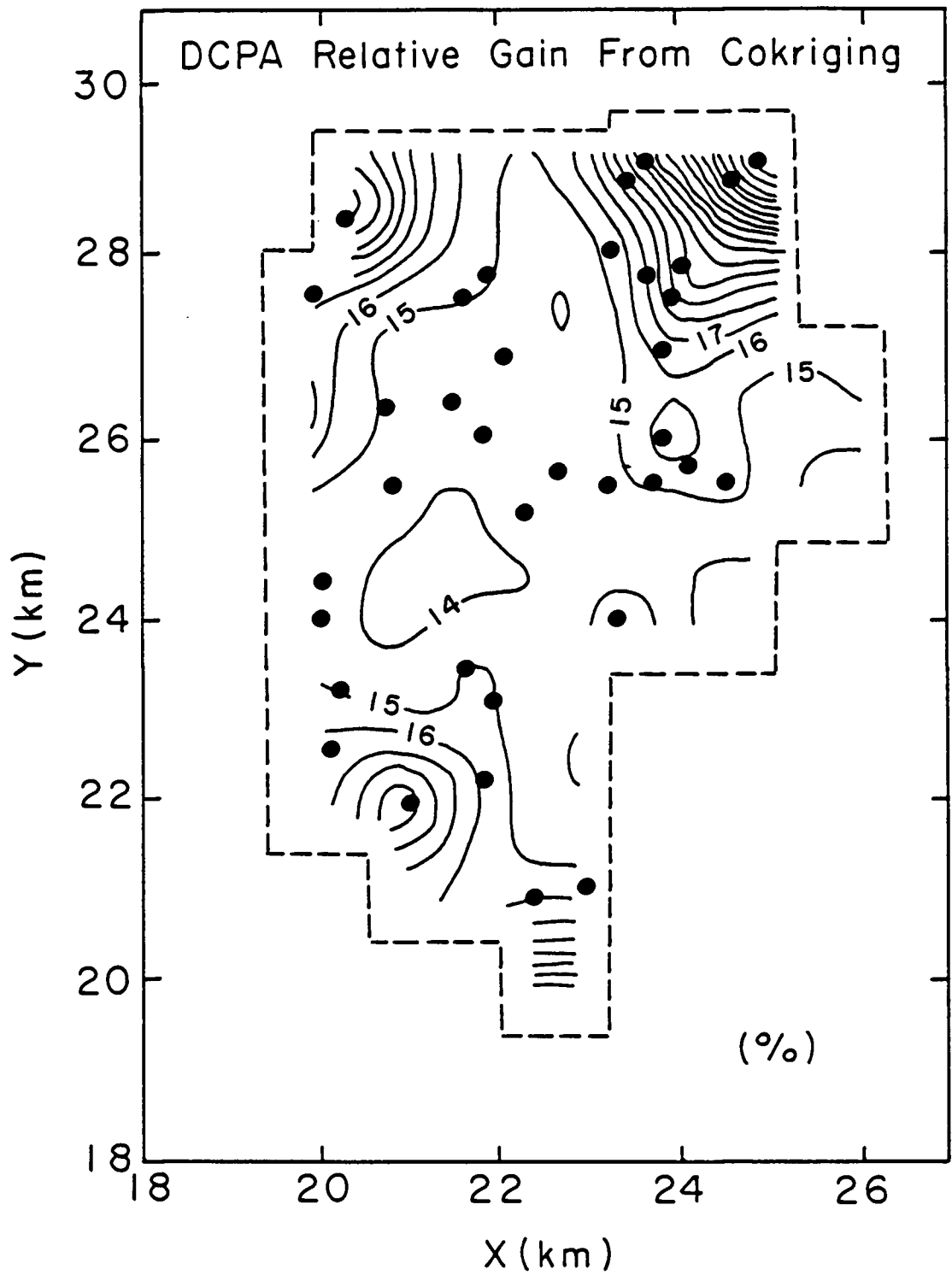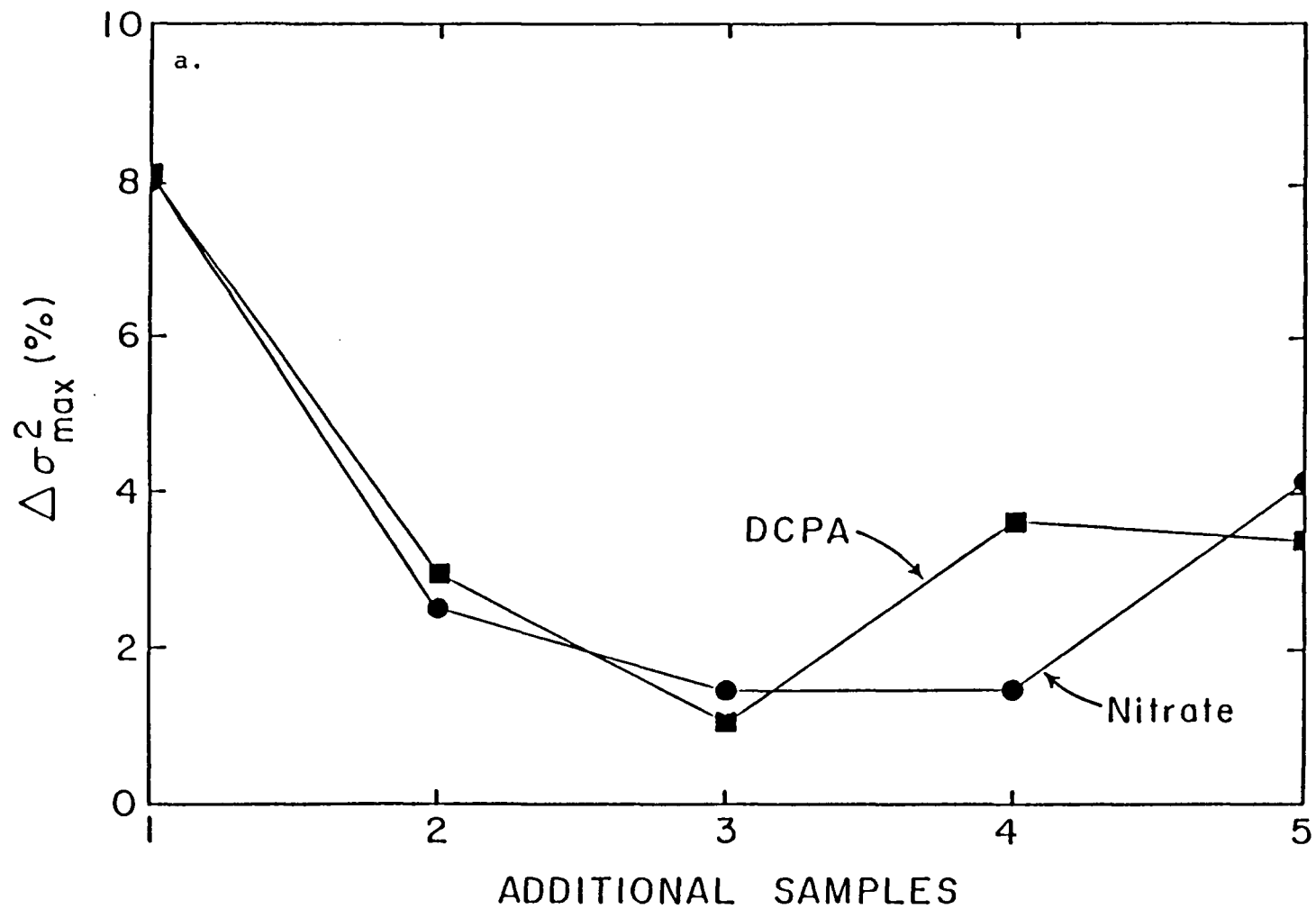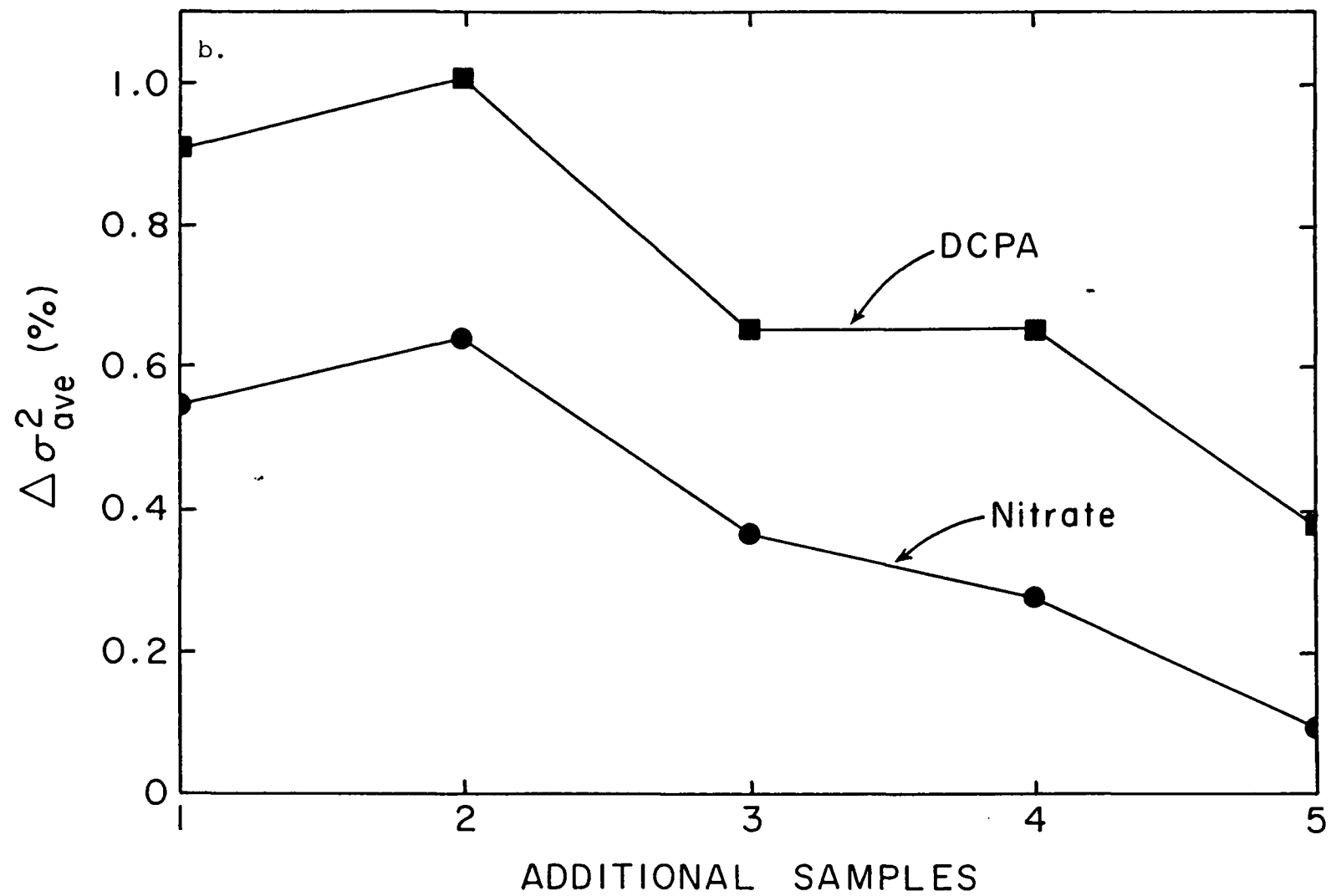|	
$b_{NN}{}^i$	structure i (sill or nugget) of model semivariogram, NN
$C_{jk}(h)$	cross-covariance for Revs Z_k and Z_j
D	matrix of semivariogram vectors and a matrix of identity vectors
$\gamma_{jk}(h)$	cross-semivariogram for Revs Z_k and Z_j separated by the distance h
$\gamma_{jk}{}^*(h)$	experimental cross-semivariogram for ReVs Z_k and Z_j
$\gamma_{jj}{}^*(h)$	experimental direct-semivariogram for ReV Z_j
$\gamma_{jj}(h)$	direct-semivariogram for ReV Z_j
$\gamma(x_n,x_n)$	semivariogram matrices for the distance separating points x_n and x_n
Γ_i	matrix of $\lambda_{jk}{}^i$ vectors
I	identity matrix
k	number of sampled locations for DCPA
l	number of sampled locations for nitrate
$\lambda_{jk}{}^i$	weight in Γ_i attributed to Z_j in estimating Z_k at location x_i
λii	weight attributed to Z_j in estimating Z_j at location x_i
m_j	mean of regionalized variable Z_j
m_k	mean of regionalized variable Z_k
m	total number of ReVs
μ_D	Lagrange multiplier for DCPA estimation
μ_N	Lagrange multiplier for nitrate estimation
μ	matrix of lagrange multipliers
n	total number of sampled locations, x_i
$N(h)$	number of sampled points separated by the distance h
ReV	regionalized variable
$R(x)$	relative reduction in variance
ρ_{jk}	correlation coefficient for Z_j and Z_k

σ_{ck}^2	cokriging variance or estimation error
$\sigma_{ck}^{2'}$	cokriging variance with fictitious point
Tr	the trace operation of a matrix
U	matrix of semivariogram matrices and identity matrices
Y	vector of Γ_i and μ matrices
Z*(x)	vector of estimated values for a ReV
Z(x$_i$)	vector of sampled ReV values at point x_i
Z$_D$(x$_i$)	sampled value of DCPA at x_i
Z$_D$(x$_i$+h)	sampled value of DCPA at distance h from x_i
Z$_N$(x$_i$)	sampled value of nitrate at x_i
Z$_N$(x$_i$+h)	sampled value of nitrate at distance h from x_i

References

Barnes, M.G., Statistical design and analysis in the cleanup of environmental radionuclide contamination, DRI publication no. 45012, Desert Research Institute, University of Nevada, Reno NV, 1978.

Barnes, M.G., Giacomini, J.J., Reiman, R.T., and Elliot, B., NTS Radiological assesment program: Results for Frenchman Lake Area 5, DOE/DP/01253-17, Desert Research Institute, University of Nevada, Reno NV, 1980.

Beste, C.E., Chairman of the Herbicide Handbook Committee, *Herbicide Handbook of the Weed Science Society of America 5 th Edition*, 492 pp., Weed Science Society of America, Chicago IL, 1983.

Cooper, R.M., and Istok, J.D., Geostatistics applied to Groundwater Contamination, 1. Methodology, *J. Envir. Engrg., ASCE, 114(2)*, 270-286, 1988a.

Cooper, R.M., and Istok, J.D., Geostatistics applied to Groundwater Contamination, 2. Application, *J. Envir. Engrg., ASCE, 114(2)*, 287-299, 1988b.

David, M., *Geostatistical Ore Reserve Estimation*, 364 pp., Elsevier Scientific Publishing Company, New York NY, 1977.

Delhomme, J.P., Kriging in the hydrosciences, *Advances in Water Resources, 1(5)*, 475-499, 1978.

Flaig, E.G., Bottcher, A.B., and Cambell, K.L., Estimatation of nitrate leaching using geostatistics, Presented at the ASAE Summer Meeting , San Luis Obispo CA, June 29-July 2, 1986.

Hoeksema, R.J. and Clapp, R.B., Estimating water table elevations using cokriging, (abstract), *EOS Trans. AGU, 68*, 44, 1987.

Istok, J.D., and Cooper, R.M., Geostatistics applied to Groundwater Contamination, 3. Global Estimation, *J. Envir. Engrg., ASCE*, in press, 1988.

Istok, J.D., Cooper, R.M., and Flint, A.L., Three-dimensional cross-semivariogram calculations for hydrogeologic data, *Groundwater*, in press, 1988.

Journel, A.G., and Huijbregts, C.J., *Mining Geostatistics* , 600 pp., Academic Press, New York NY, 1978.

Journel, A.G., and Issaaks, E.H., Conditional indicator simulation: Application to a Saskatchewan uranium deposit, *Jour Math Geol*, 16(7), 685-718, 1984.

Luster, G.R., Raw materials for Portland cement: Applications of conditional simulation of coregionalization, Ph. D. dissertation, 531 pp., Stanford University, 1986.

Marsily, de Ghislain, *Quantitative Hydrogeology, Groundwater Hydrology for Engineers*, 439 pp., Academic Press, New York NY, 1986.

Myers, D.E., Matrix formulation of cokriging, *Jour. Math. Geol., 14(3)*, 249-257, 1982.

Myers, D.E., Begovich, C.L., Butz, T.R, and Kane, V.E., Variogram models for regional groundwater data, *Int. Ass. Math. Geol. Jour., 14(6)*, 629-644, 1982.

Myers, J.C., and Bryan, R.C., Geostatistics applied to toxic waste - A case study, in *Geostatistics for Natural Resource Characterization, Part 2*, G. Verely, M. David, A.G. Journel, and A. Marechal, Eds., 893-901, D. Reidel Publishing Company, Dordrecht, Holland, 1984.

SAS/Institute Inc., SAS/STAT *Guide for Personal Computers, Version 6 Edition*, 378 pp., Cary NC: SAS Institute Inc., 1985.

United States Environmental Protection Agency, *Dacthal Draft Health Advisory*, 13 pp., Office of Drinking Water, August, 1987.

Van Rooy, D., Conditional stochastic simulation of groundwater contamination, A case study, Ph.D. disertation, 154 pp., Institute of Hydrodynamics and Hydraulic Engineering, Technical University of Denmark, DK-Lyngby, 1986.

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

Verly, G.W, The block distribution given a point multivariate normal distribution, *Geostatistics for Natural Resource Characterization, Part 1* , G. Verely, M. David, A.G. Journel, and A. Marechal, Eds., 495-515, D. Reidel Publishing Company, Dordrecht, Holland, 1984.

Verly, G.W., David, M., Journel, A.G., and Marechal, A., eds., *Geostatistics for Natural Resources Characterization*, Proceedings of the NATO Advanced Study Institute, South Lake Tahoe CA, Sept 6-17, D. Reidel, Dordrecht, Holland, 1092 pp, 1983.

Whitson, T.D., William, R.D., Parker, R., Swan, D.G., Dewey, S., Eds., *Pacific Northwest Weed Control Handbook*, 213 pp., Extension Service, Oregon State University, Wahington State University, and the University of Idaho, 1987.

Zirschky, J.H., Keary, G.P., Gilbert, R.O., and Middlebrooks, E.J., Spatial estimation of hazardous waste site data, *J. Envir. Engrg., ASCE, Vol 111*, 777-789, 1985.

Zirschky, J.H., and Harris, D.J., Geostatistical analysis of hazardous waste data, *J. Envir. Engrg., ASCE, Vol 112*, 770-785, 1986.

MULTIVARIATE GEOSTATISTICAL ANALYSIS OF GROUNDWATER CONTAMINATION BY PESTICIDE AND NITRATE

II. SIMULATION

Jeffrey D. Smyth

Hydrology Section, Battelle Pacific Northwest Laboratories, Richland, Washington

99352 (Formerly at Agricultural Engineering Department, Oregon State University,

Corvallis, Oregon 97331)

Jonathan D. Istok

Agricultural Engineering and Civil Engineering Departments, Oregon State University,

Corvallis, Oregon 97331

Abstract

Based on the semivariogram models found in Chapter I, univariate and multivariate conditional simulations of nitrate and DCPA were generated using the turning bands method and the kriging and cokriging system. Kriging was used to condition the univariate simulations, while cokriging was used to cross-correlate and condition the multivariate simulations. The mean of 25 conditional and coconditional simulations at 8 different locations in the study area were generated and compared to kriging and cokriging estimates and 95% confidence intervals. The results indicated that there was a large component of randomness in the simulations due to the small sample data set. Contour maps of conditional simulations also displayed fluctuation due to the random component of the contaminant densities. Coconditional simulation appeared to display the cross-correlation imposed by using the cokriging system to condition the simulations. This case study demonstrates the importance of supplementing geostatisical estimation methods with simulation in a multivariate geostatistical analysis

MULTIVARIATE GEOSTATISTICAL ANALYSIS OF GROUNDWATER CONTAMINATION BY PESTICIDE AND NITRATE

II. SIMULATION

Jeffrey D. Smyth

Hydrology Section, Battelle, Pacific Northwest Laboratories, Richland, Washington 99352 (Formerly at Agricultural Engineering Department, Oregon State University, Corvallis, Oregon 97331)

Jonathan D. Istok

Agricultural Engineering and Civil Engineering Departments, Oregon State University, Corvallis, Oregon 97331

Introduction

The estimates obtained by the kriging or cokriging system are, by definition, the estimates with minimized estimation variance and therefore show less fluctuation than the actual, unknown values [Journel and Huijbregts, 1978]. As shown in Smyth and Istok [1988], estimation variances obtained from a multivariate geostatistical analysis are useful for assessing the quality of the estimates, and for determining the optimal location for additional samples. However, because the estimation variances obtained by kriging or cokriging are minimized they do not reflect the entire range of fluctuation that is possible for a regionalized variable (ReV) between sample points. Delhomme [1979] notes that the total spatial variability of a ReV can be considered as two parts: the estimate and the uncertaintity of the estimate that exists between sample points. A geostatistical analysis is not complete unless the degree of fluctuation of the ReVs is assessed. In univariate geostatistics, conditional simulation (CS) can be used to simulate additional realizations

from the random function that characterizes the ReV. The variance for the realizations simulated using CS is twice the kriging or cokriging estimation variance [Journel and Huijbregts, 1978].

In hydrogeology CS has been used primarily to generate realizations of transmissivities, and hydraulic head [Delhomme, 1979; Dagan, 1982; Clifton and Neumann, 1982; Van Rooy, 1987]. An example use for the realizations obtained by CS is to generate input data for use in stochastic-deterministic groundwater flow and solute transport models. Bryan and Myers [1984] used CS in a case study of lead contamination in soil to generate 'worst-case scenarios' i.e., simulations that show the maximum posiible fluctuation of lead in the soil . Technically simulations obtained using CS do not represent worst-case scenarios, but rather possible versions of reality that are consistent with the available data.

An application of CS to groundwater contamination has not previously been published. The objective of this paper is to extend CS to univariate and multivariate conditional simulation of two agricultural groundwater contaminants, dimethyl tetrachloroterephthalate (DCPA) and nitrate. The specific objectives are to generate univariate CSs for DCPA and nitrate, and to use the cokriging system to cross-correlate CSs for DCPA and nitrate. CSs will be generated in each case to demonstrate the possible fluctuation of the ReVs. The mean and variance of multiple CSs will be calculated and compared with the kriging and cokriging estimates and estimation variances.

Equation Development

One method for obtaining a CS, developed by Matheron in the early 1960's, is the turning bands method (TBM). Compared to the other methods for CS, TBM is the most efficient and accurate in terms of computer time [Mantoglou and Wilson,1981]. The theory and method of generating a CS using the TBM and the process of moving-averages on which it is based is in Journel [1974], Journel and Huijbregts [1978], Mantoglou and Wilson [1981], and Luster [1986]. Only the equations that were used in this study to generate univariate and multivariate CSs of DCPA and nitrate will be presented.

The first step in an unconditional simulation (i.e., simulations that have not been adjusted to reflect the measured values at sample points) is to generate one-dimensional simulations of a random function (RF) with a specified variance and mean. For the spherical covariance model or semivariogram, if second-order stationairity is assumed, the moving-average process used to correlate the independent random numbers into one-dimensional simulations is given by

$$y_i = \sum_{w=-R}^{w=R} \left((t_{i+w}) \left(\sqrt{\frac{12K}{a^3}} \right) wb \right) \tag{1}$$

where y_i is a realization at the ith point along a one-dimensional line with a one-dimensional spherical covariance given by

$$C^1(s) = K \left[1 - \frac{3s}{a} + \frac{2s^3}{a^3} \right], 0 \leq s \leq a \tag{2}$$

and

$$C^1(s) = 0, \ s \geq a \tag{3}$$

where t_{i+w} is a uniformly distributed, random [0,1] number at the i+w point, b is the length of interval between the points 't', and K and a are the sill and the range corresponding to the model semivariogram $\gamma(h)=C(0)-C(h)$. Algebraically, b=a/(2R), and R is accepted in practice as being an integer less than or equal to 20 [Journel and Huijbregts, 1978] (Figure II.1).

By orienting the set of one-dimensional lines calculated by (1) through a central point to represent a spherical configuration, the one-dimensional simulations can be extended to three-dimensional space using the TBM. In practice, 15 different lines are used and the spherical configuration is approximated as the lines connecting the midpoints of opposite sides of a regular icosohedron [Journel, 1974]. The three dimensional simulation at a point is calculated by,

$$Z_0(x_i) = \frac{1}{\sqrt{15}} \sum_{j=1}^{15} (y_i)_j \qquad (4)$$

where $Z_0(x_i)$ is the simulated point in three dimensional space, and $(y_i)_j$ is the one-dimensional simulated point on the jth line [Journel and Huijbregts, 1978].

The next step is to condition the simulations to observed sample data. The equation representing this is

$$Z_{CS}(x) = Z_{UC}(x) - Z_{UC}^*(x) + Z_{OB}^*(x)$$

$$(5)$$

where $Z_{CS}(x)$ and $Z_{NC}(x)$ are the conditioned and unconditioned simulations, respectively, $Z_{UC}^*(x)$ are the kriged estimates using the unconditioned simulations as data, and $Z_{OB}^*(x)$ are the kriged estimates using the observed sample values. The result

is a univariate conditional simulation in three-dimensional space with a spatial distribution represented by a spherical semivariogram and in which the simulated value equals the sample value at the sample points.

In multivariate conditional simulation, two general approaches are available that differ in the method used to ensure that the spatial cross-corelation between ReVs is preserved. One method is to force the simulations to be cross-correlated during the generation of the simulations, and uses kriging for conditioning (Journel and Huijbregts, 1978; Luster, 1986). The more general method presented by Carr and Myers [1985] uses solutions to the cokriging system to establish the cross-correlation between the ReVs and to condition the simulations to the sample values. The advantage of the first method is that the spatial cross-correlation between simulated ReVs is obtained without the use of cokriging. However, only a single pair of ReVs can be simulated at one time. Complete details of multivariate simulation using the linear model of coregionalization are in Luster [1986]. The advantage of the method used by Carr and Myers [1985] is that it can be applied to simulations of more than two ReVs at the same time (i.e., for the case of coconditional simulation).

Empirical results indicate that cokriging is successful in cross-correlating simulated ReVs [Carr and Myers, 1985]. However Luster [1986] notes that a lack of cross-correlation between the simulated ReVs can occur in areas with an insufficient number of sample points.

The coconditional simulation procedure uses the same method to generate unconditioned three-dimensional simulations described in (1), (2), (3), and (4). However the conditioning step, Equation (5), is modified for the multivariate case as,

$$\bar{Z}_{CS}(x) = \bar{Z}_{UC}(x) - \bar{Z}_{UC}^{*}(x) + \bar{Z}_{OB}^{*}(x) \qquad (6)$$

where

$$\bar{Z}_{CS}(x) = \begin{bmatrix} Z_{CS,1}(x) \\ Z_{CS,2}(x) \\ \vdots \\ Z_{CS,m}(x) \end{bmatrix}, \text{ or the conditional simulation of m ReVs,}$$

$$\bar{Z}_{UC}(x) = \begin{bmatrix} Z_{UC,1}(x) \\ Z_{UC,2}(x) \\ \vdots \\ Z_{UC,m}(x) \end{bmatrix}, \text{ the uncondional simulation of m ReVs ,}$$

$$\bar{Z}^*_{UC}(x) = \begin{bmatrix} Z^*_{UC,1}(x) \\ Z^*_{UC,2}(x) \\ \vdots \\ Z^*_{UC,m}(x) \end{bmatrix}, \text{ the cokriging estimates from the unconditional simulations,}$$

$$\bar{Z}^*_{OB}(x) = \begin{bmatrix} Z^*_{OB,1}(x) \\ Z^*_{OB,2}(x) \\ \vdots \\ Z^*_{OB,m}(x) \end{bmatrix}, \text{ the cokriging estimates from the sample values.}$$

In the case of groundwater contamination by nitrate and DCPA, (6) is written,

$$\begin{bmatrix} Z_{CS,D}(x) \\ Z_{CS,N}(x) \end{bmatrix} = \begin{bmatrix} Z_{UC,D}(x) \\ Z_{UC,N}(x) \end{bmatrix} - \begin{bmatrix} Z^*_{UC,D}(x) \\ Z^*_{UC,N}(x) \end{bmatrix} + \begin{bmatrix} Z^*_{OB,D}(x) \\ Z^*_{OB,N}(x) \end{bmatrix} \tag{7}$$

with N and D representing nitrate and DCPA, resectively. The cokriging procedure used to calculate $Z_{UC,D}^*(x)$, $Z_{UC,N}^*(x)$, $Z_{OB,D}^*(x)$, and $Z_{OB,N}^*(x)$ is described in the first paper of this series [Smyth and Istok, 1988].

Methodology

Several assumptions were made in estimating the contaminant densities of nitrate and DCPA in the alluvial and basin-fill aquifer near Ontario, Oregon. The 42 nitrate and 32 DCPA domestic drinking water samples were shown to have univariate lognormal distributions [Smyth and Istok, 1988]. It was assumed that log(nitrate) and log(DCPA) had univariate normal distributions, and additionally that the joint distributions were normal. A constant porosity of 35% and contaminated aquifer thickness of 3.2 m was also assumed for the study area. The contaminant densities were assumed to have second-order stationarity within sliding neighborhoods of 4 km. All geostatistical analysis was performed on log(nitrate) and log(DCPA). Results were then reported as nitrate and DCPA contaminant densities. Isotropic, spherical semivariogram models with correlation ranges of 4 km were found to represent the spatial distribution of log(nitrate) and log(DCPA) (Table II.1).

The FORTRAN program COSIM [Carr and Meyrs, 1985] was used to produce conditional and coconditional simulations of log(nitrate) and log(DCPA). For each case two unconditional and conditional simulations were generated. The simulated values were inverted into their original lognormal distributions, $Z_s(x)$, using

$$Z_s(x) = \exp[Y_s(x)] \tag{8}$$

where $Y_s(x)$ are the simulated log(nitrate) and log(DCPA) values.

Contour maps were prepared for the simulated values $Z_s(x)$. Contour maps were also prepared for unconditioned and conditioned nitrate simulations.

Kriging and cokriging estimates were computed (Figure II.8) with 25 conditional simulations and cosimulations of nitrate and DCPA generated at each point. The mean of the simulated values were compared to the kriging and cokriging estimates and 95% confidence intervals (C.I.) to assess the fluctuation of the simulated values.

Results

Contour maps for the results of unconditional nitrate simulations show a large degree of fluctuation in the simulated values and in the location of maximum values (Figures II.2a, and II.3a). After conditioning the maximum values of the nitrate simulation occur in the same locations in both simulations (Figures II.2b, and II.3b). The maximum contoured values of nitrate densities are 100000 mg/m^2 (93.7 ppm) and 150000 mg/m^2 (140.6 ppm). The maximum sampled value of nitrate in the aquifer was 34144 mg/m^2. The large difference between the maximum sample value and the conditional simulation values and the fluctuation of the simulated values demonstrates the influence of the unknown component of nitrate distribution in the study area.

The conditional DCPA simulation contour maps reflect the same fluctuation as the nitrate simulations (Figures II.4 and II.5). The range of simulated values do not vary as widely as the nitrate simulations and the maximum values occur in different locations in each figure. Maximum contoured values of DCPA are 600 mg/m^2 (562 ppb) and 400 mg/m^2 (375 ppb) (Figures II.4 and II.5). That the location of the 600 mg/m^2 contour in Figure II.4 corresponds to a 20 mg/m^2 contour in Figure II.5 demonstrates the variability of the simulation that is possible when a small data set (i.e., small number of DCPA samples) is used to condition the simulations.

Coconditional simulation uses the cross-correlation between nitrate and DCPA to condition the simulated data. This procedure uses nitrate sample data to help condition the undersampled DCPA data, and further to ensure that the cross-correlation between nitrate and DCPA, as defined by the model cross-semivariogram $\gamma_{DN}(h)$ is duplicated by the simulations. The coconditional nitrate and DCPA simulations, Figures II.6 and II.7 respectively, illustrate the effects of imposing the cross-correlation restriction on the simulations. In general the nitrate and DCPA values were lower in the center of the study area, with increasing values near the borders. This trend is especially evident in the

southern portion of the study area where the maximum contoured values of nitrate (100000 mg/m^2) and DCPA (600 mg/m^2) contaminant densities occurred (Figures II.6 and II.7).

Conditional and coconditional simulations are basically random simulations that have a specified variance and covariance, spatial distribution, and are conditioned to equal sample values at sample locations. In addition the mean of several CS at a point should be equivalent to the kriging or cokriging estimate [Delhomme, 1979]. These properties of conditional simulation were used to assess the quality of the simulation (i.e., the simulation method and the assumptions). Eight locations were randomly selected in the study area (Figure II.8). 25 conditional and coconditional simulations of nitrate and DCPA were produced at each point, and the mean calculated. None of the locations corresponded exactly to a sample location, however point 5 is within 500 m of well 9 (Table I.1). The calculated mean had not stabilized after 25 simulations. To better evaluate the accuracy of the simulations, 95% confidence intervals (C.I.s) were calculated for the kriging and cokriging estimates using

$$Z_D(x_0)_{95\%} = Z_D^*(x_0) \pm 2(\sigma_{D0}) \tag{9}$$

where, σ_{D0} is the standard deviation of the DCPA estimate, and $Z_D^*(x_0)$ is the estimate of DCPA [Journel and Huijbregts, 1978]. A similiar equation can be written for the 95% confidence interval for nitrate. Simulations should have a variance that is twice the kriging or cokriging estimation variance, and therefore are not required to fall within the 95% C.I., however the mean of 25 simulations should occur near the 95% C.I..

The conditional and coconditonal nitrate simulations were the least uniform and the calculated means typically fell outside the 95% C.I. (Table II.2). While not a sample location, point 5 should be heavily influenced by well 9 due to its close proximity. The value of nitrate contaminant density at well 9 was 11096 mg/m^2, the kriged value 11300

mg/m^2, and the conditional and coconditional simulations were 17498 mg/m^2 and 69684 mg/m^2, respectively.

The DCPA conditional and coconditional simulation calculated means fell within their 95% C.I. more regularly; 6 of 8 conditional simulations, and 5 of 8 coconditional simulations (Tables II.3 and II.4). Point 5 fell within the 95% C.I. for the conditional simulation, but not for the coconditional simulation.

The 95% C.I. evaluations demonstrate that at 25 simulations the calculated mean is heavily enfluenced by an abnormally large or small simulation value. The lack of uniformity of the calculated means occurring within their 95% C.I. also shows the large component of randomness in the simulations, that was indicated by the contour maps. The negative values of 95% C.I. limits that occur for some points (e.g., point 4) are due to an estimate near zero, and a large estimation variance. Since a negative contaminant value is meaningless for analysis, it should be discarded.

Conclusions

The direct- and cross-semivariogram models $\gamma_{NN}(h)$, $\gamma_{DD}(h)$, and $\gamma_{DN}(h)$ were assumed to represent the spatial distribution of log(nitrate), log(DCPA) and the cross-correlation between log(nitrate) and log(DCPA) respectively. The FORTRAN program COSIM written by Carr and Myers [1985] was used to produce conditional and coconditional simulations of log(nitrate) and log(DCPA). The simulated values were then inverted into lognormal distributions using (8).

Contour maps of conditional simulations showed a large degree of fluctuation in simulated nitrate and DCPA values. The fluctuation was assumed to be due to the small number of data (42 nitrate and 32 DCPA samples). The contour maps of the coconditional simulation also displayed large fluctuations of values, but the cross-corelation between the nitrate and DCPA was preserved.

The calculated mean's of 25 conditional and coconditional simulations at eight different locations in the study area also were influenced by the small data set. Calculation of the 95% C.I'.s were performed to see if the calculated mean fell within the interval. The mean's did not typically show a set pattern of falling within the 95% C.I., however this was assumed to be due to the instability of the calculated mean's with only 25 simulations.

Large fluctuation are possible in the distribution of nitrate and DCPA contaminant densities as demonstrated by the conditional and coconditional simulation analysis. The fluctuation in the simulated values of nitrate and DCPA in this study are due to the small number of sample data used in this case study. This emphasizes the importance of including conditional simulation in a geostatistic analysis. Alone, the estimates of the contaminant densities do not reveal the possible fluctuation in values. Together, the estimates and the simulations suggest that more samples should be used to better characterize the nitrate and DCPA groundwater contamination.

To use the simulations as input data in solute-transport models, it would be necesarry to preserve the cross-correlation that was found to exist (0.74) between nitrate and DCPA contaminant densities Coconditional simulation of nitrate and DCPA contaminant densities resulted in the cross-correlation of the ReVs as expected. However the instability of the calculated mean of the conditional and coconditional simulations is evidence that more than 25 simulations are required to use the simulations as input data for a deterministic-stochastic solute-transport model.

Table II.1 Direct- and cross-semivariogram models for log(nitrate) and log(DCPA)

| MODEL | TYPE | NUGGET | SILL | RANGE |
		$(\log((mg/m^2)^2))$		(km)
$\gamma_{NN}(h)$	spherical	0.79	1.00	4.0
$\gamma_{DD}(h)$	spherical	0.74	0.90	4.0
$\gamma_{DN}(h)$	spherical	0.20	0.94	4.0

Table II.2 Conditional and coconditional nitrate simulations compared with kriging estimates

| POINT | X | Y | CONDITIONAL SIMULATIONS | COCONDITIONAL SIMULATIONS | KRIGING ESTIMATES | 95% C.I. |
	(km)		(mg/m^2)	(mg/m^2)	(mg/m^2)	(mg/m^2)
1	22.59	28.61	45813.	44310.	22800.	2800.,42800.
2	22.29	26.81	99850.	35041.	35041.	8234.,47365.
3	22.89	25.01	131180.	41454.	10400.	-9257.,30057.
4	21.09	28.61	43307.	54361.	17300.	-3676.,38276.
5	19.59	25.31	17498.	69684.	11300.	-8700.,31300.
6	22.29	27.41	74487.	24380.	28700.	8951.,47156.
7	21.69	25.91	114012.	26843.	26000.	6681.,45318.
8	21.09	21.11	15505.	97505.	12600.	-8376.,33576.

Table II.3 Conditional DCPA simulations compared with kriging estimates

| POINT | X | Y | CONDITIONAL SIMULATIONS | KRIGING ESTIMATES | 95% C.I. |
	(km)		(mg/m^2)	(mg/m^2)	(mg/m^2)
1	22.59	28.61	509.	183.	44.,321.
2	22.29	26.81	258.	302.	169.,434.
3	22.89	25.01	359.	91.	-41.,224.
4	21.09	28.61	128.	118.	-25.,261.
5	19.59	25.31	244.	112.	-24.,248.
6	22.29	27.41	158.	293.	158.,426.
7	21.69	25.91	205.	277.	146.,407.
8	21.09	21.11	130.	135.	-8.,278.

Table II.4 Coconditional DCPA simulations compared with cokriging estimates

POINT	X	Y	COCONDITIONAL SIMULATIONS	COKRIGING ESTIMATES	95% C.I.
		(km)	(mg/m^2)	(mg/m^2)	(mg/m^2)
1	22.59	28.61	321.	264.	134.,393.
2	22.29	26.81	454.	346.	222.,469.
3	22.89	25.01	311.	100.	-24.,224.
4	21.09	28.61	193.	187.	54.,319.
5	19.59	25.31	600.	107.	-20.,233.
6	22.29	27.41	354.	353.	228.,478.
7	21.69	25.91	397.	308.	186.,429.
8	21.09	21.11	276.	96.	-37.,229.

Figure II.1. Definition sketch for the moving average process used in TBM.
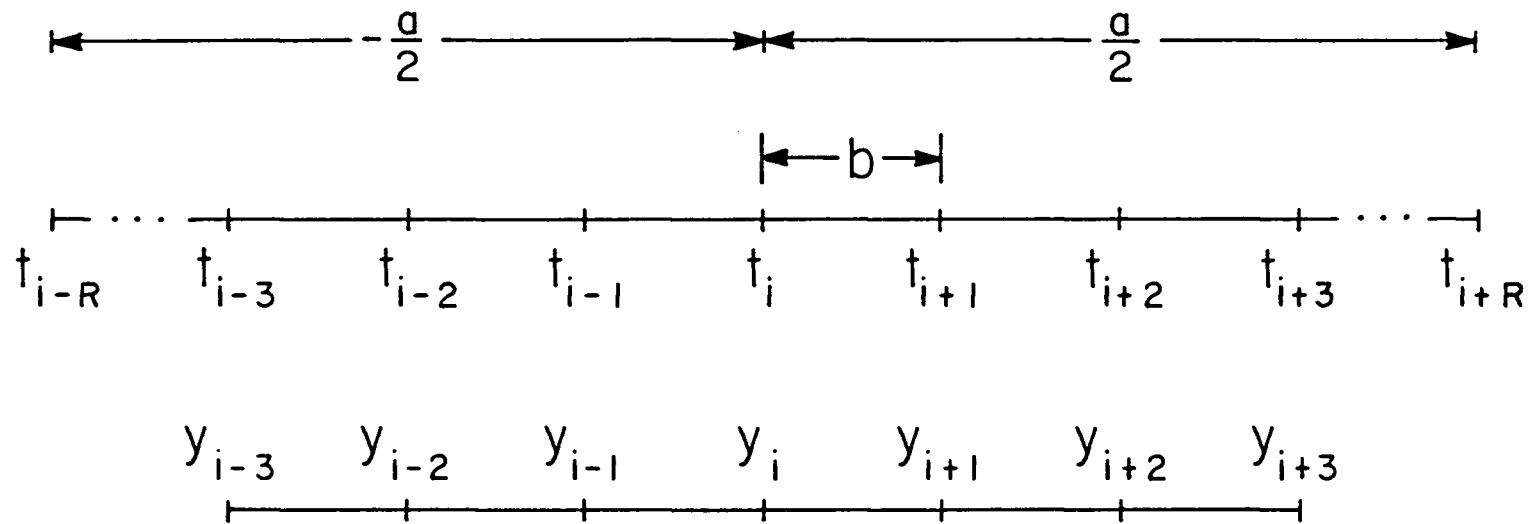
$$y_i = \sum_{w=-R}^{w=R} \left[(t_{i+w}) \sqrt{\frac{12K}{a^3}} \cdot wb \right]$$
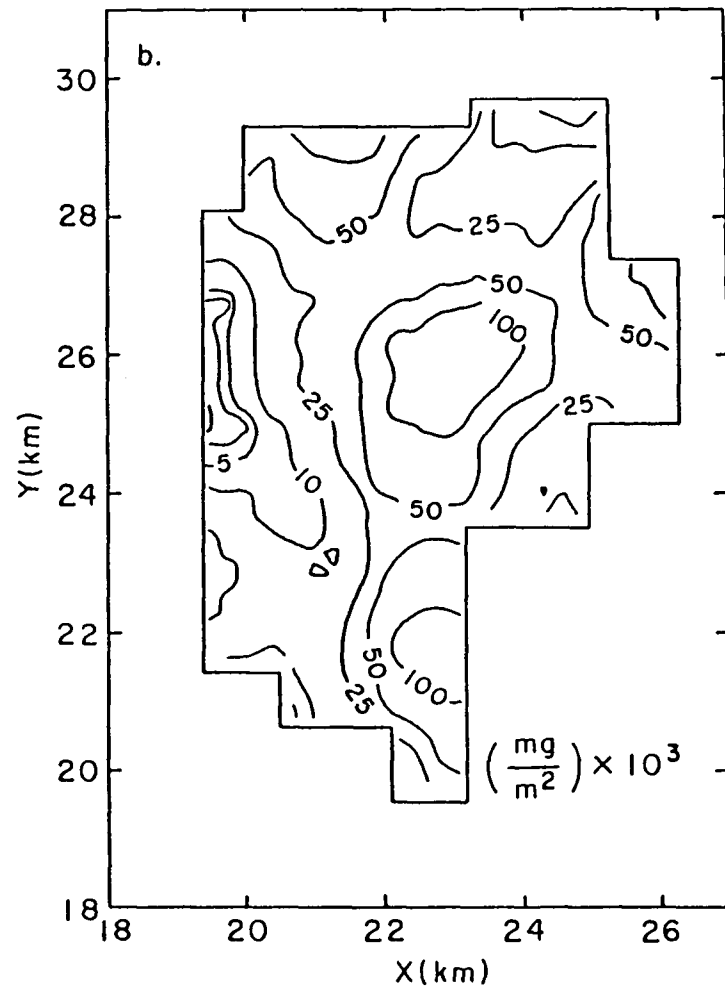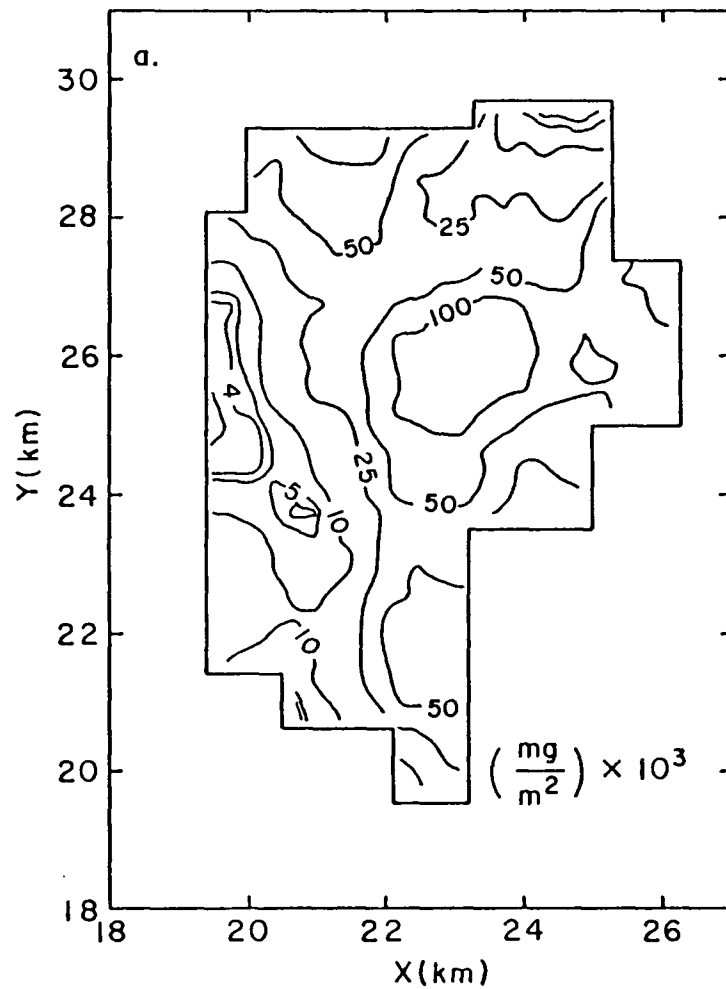
Figure II.2. Contour map of nitrate simulation 1, (a) unconditioned, (b) conditioned.
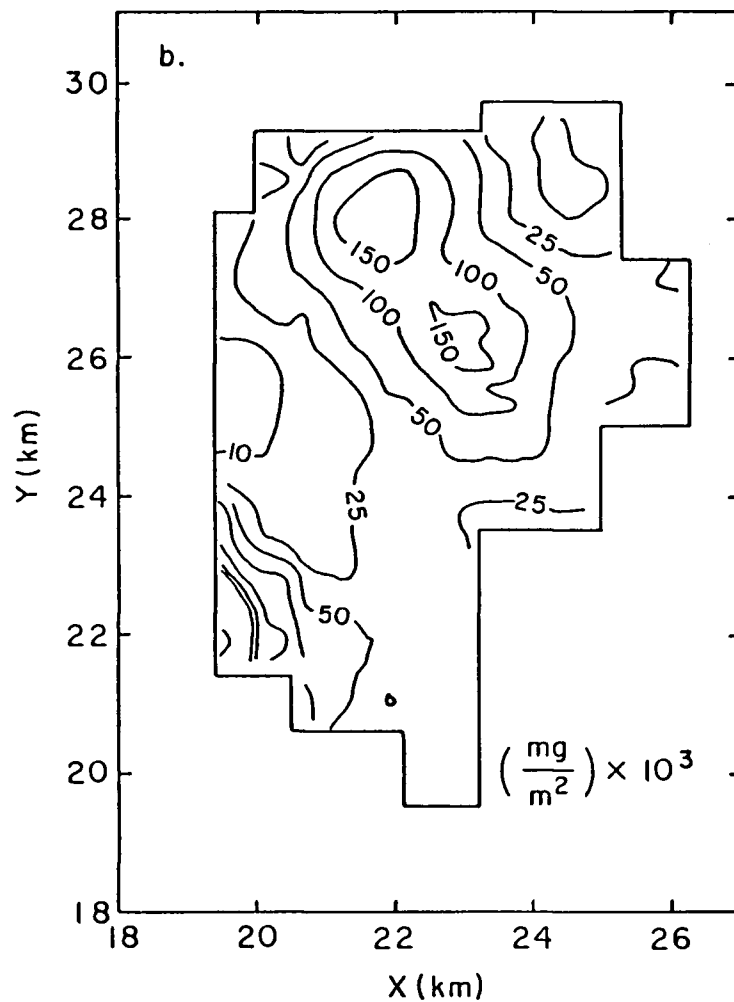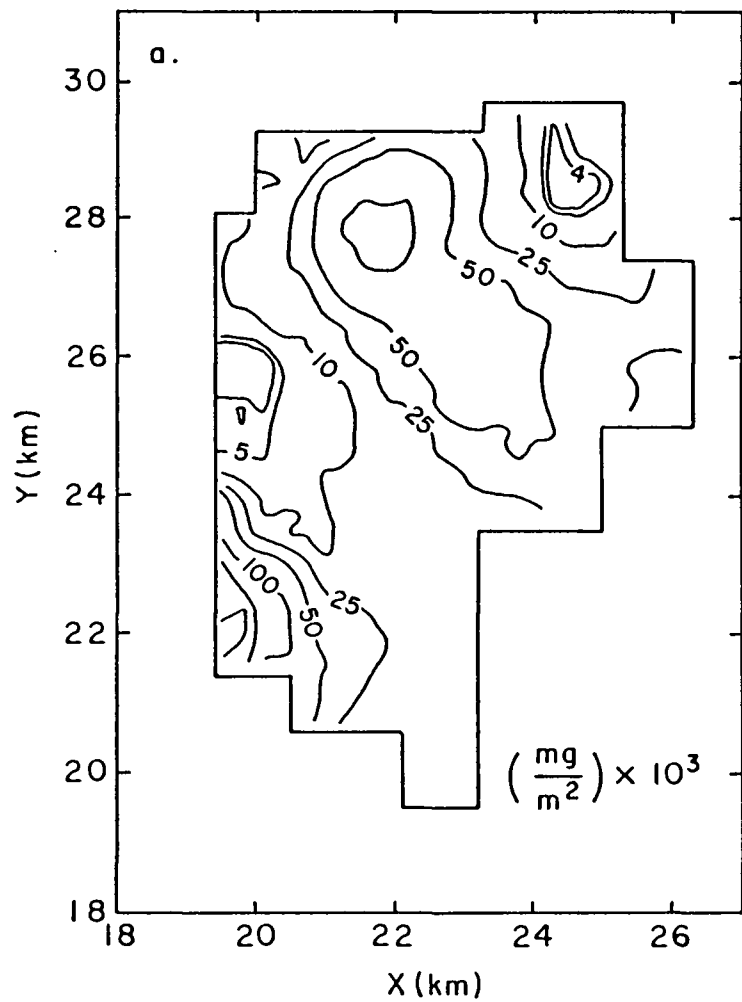
Figure II.3. Contour map of nitrate simulation 2, (a) unconditioned, (b) conditioned.
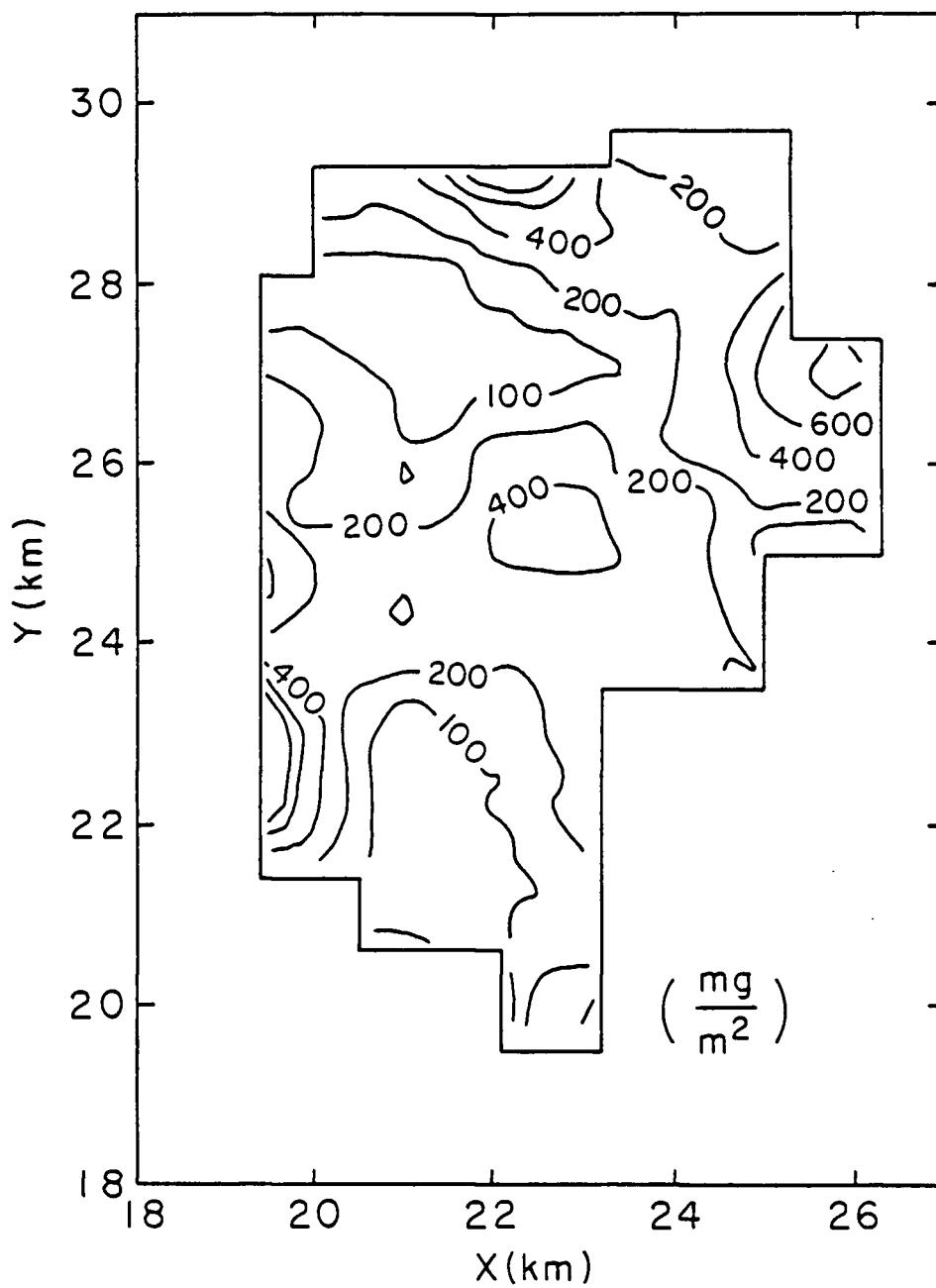
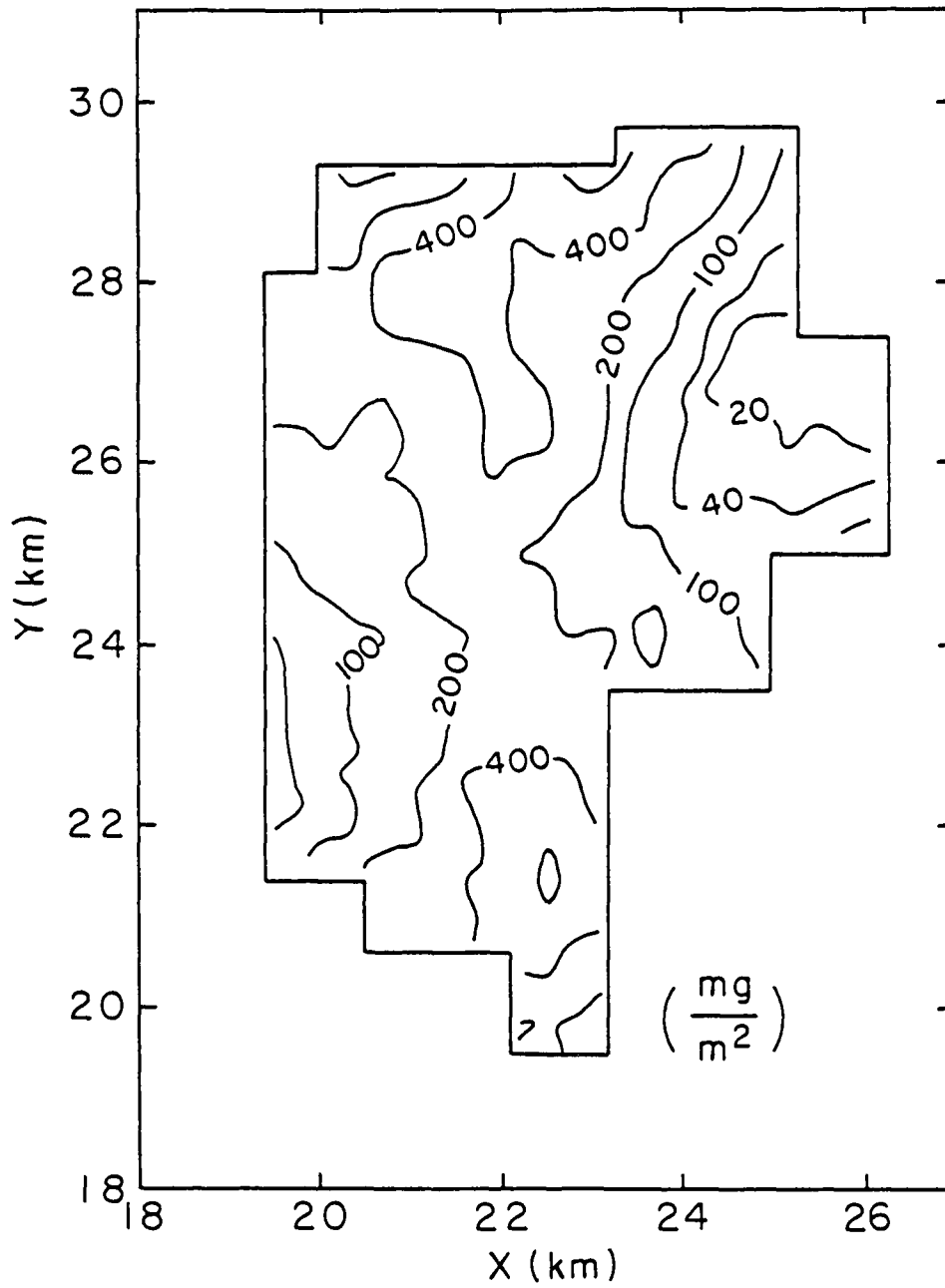Figure II.4. Contour map of conditional DCPA simulation 1.

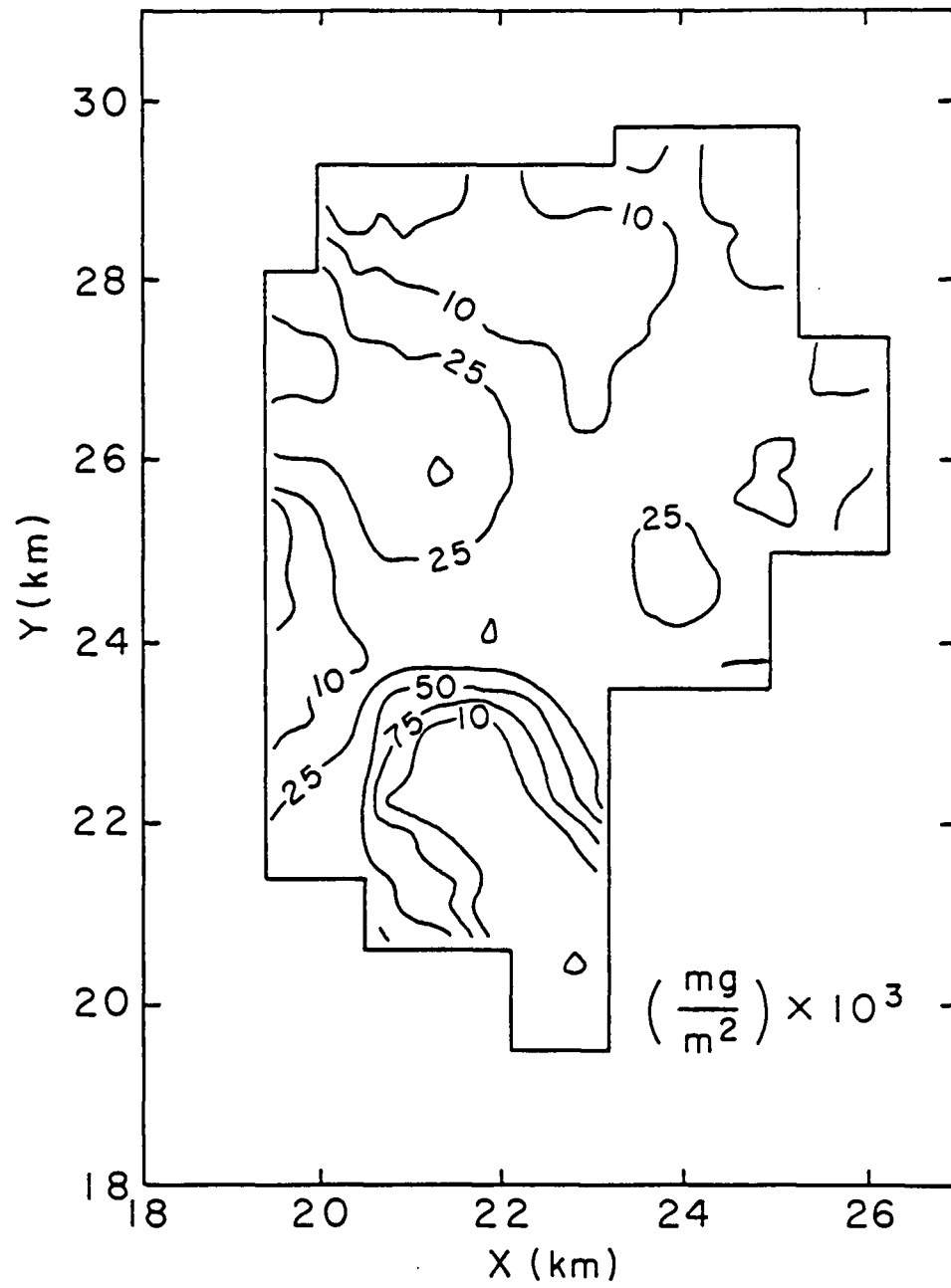Figure II.5. Contour map of conditional DCPA simulation 2.

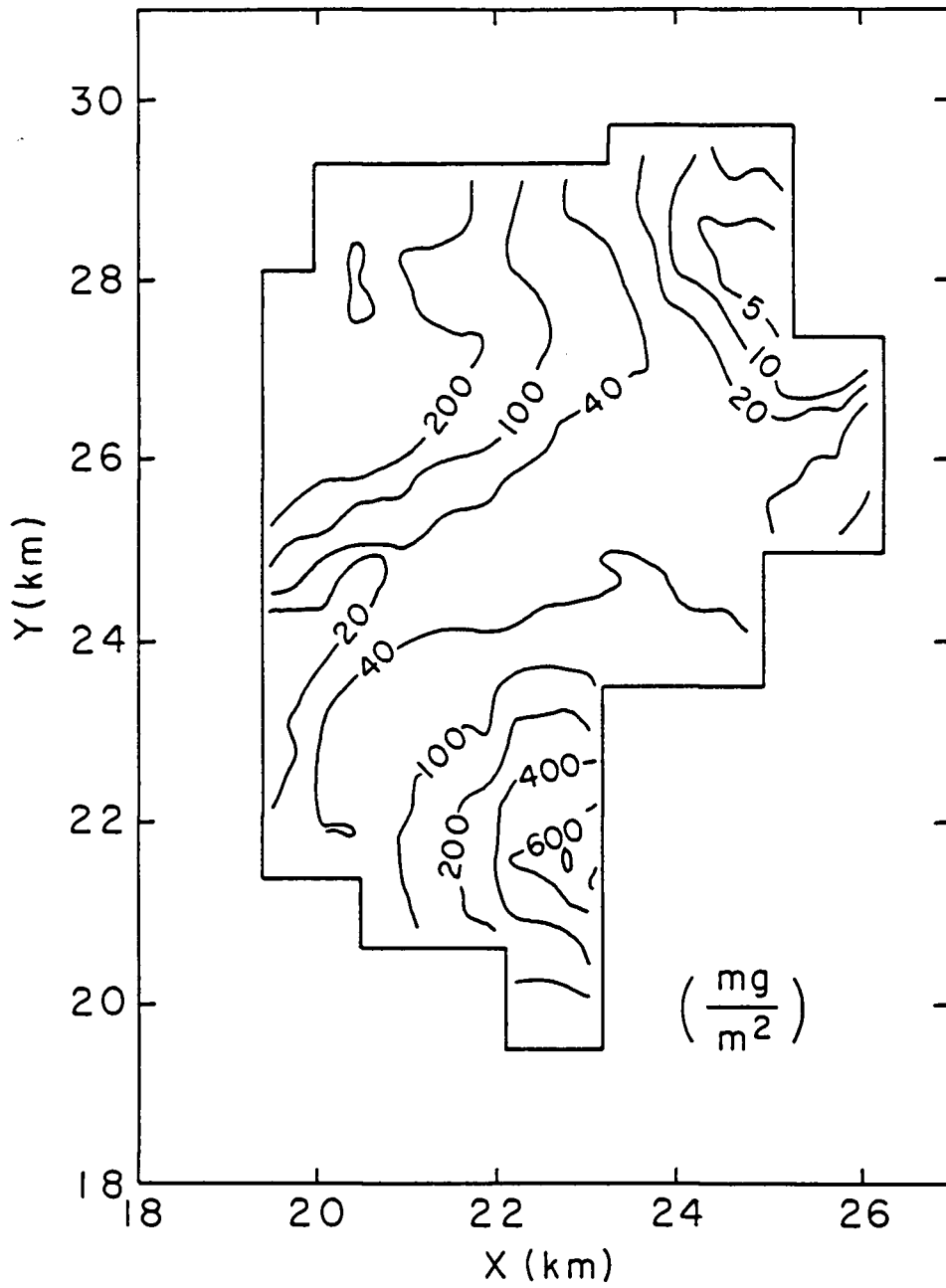Figure II.6. Contour map of coconditional nitrate simulation.

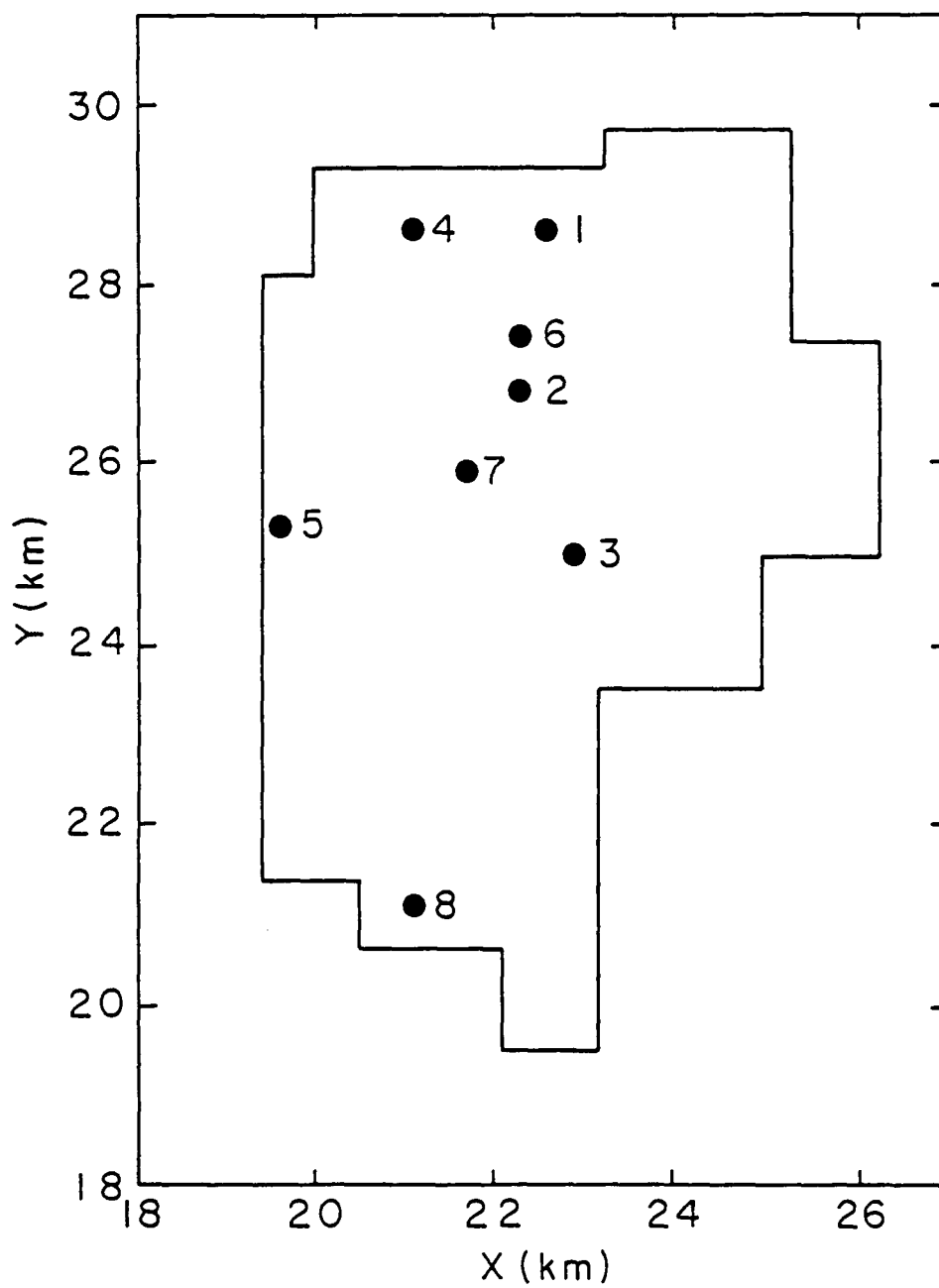Figure II.7. Contour map of coconditional DCPA simulation.

Figure II.8. Location map of eight random simulation points.

Notation

a	the one-dimensional range of influence
b	length of an interval between y-values
$C^I(s)$	one-dimensional spherical covariance at the distance s
K	the sill of the model semivariogram
R	integer count of the number of t-values included in a/2
s	(w) * (b)
t	a random number from a uniform [0,1] distribution
w	index count of R
y_i	simulation number with a one-dimensional specified spherical covariance
$Y_s(x)$	simulated value in log units
$Z_0(x_i)$	unconditioned simulation value in three-dimensions at x_i
$Z_{CS}(x)$	conditioned simulation value
$\overline{Z_{CS}}(x)$	vector of conditioned simulation values
$Z_{CS,D}(x)$	conditioned DCPA simulation value
$Z_{CS,N}(x)$	conditioned nitrate simulation value
$Z_{OB}^*(x)$	kriging or cokriging estimate of sample values
$\overline{Z_{OB}}^*(x)$	vector of kriging or cokriging estimates using sample values
$Z_{OB,D}^*(x)$	kriging or cokriging estimate using DCPA sample values
$Z_{OB,N}^*(x)$	kriging or cokriging estimate using nitrate sample values
$Z_s(x)$	simulated value
$Z_{UC}(x)$	unconditioned simulation value
$\overline{Z_{UC}}(x)$	vector of unconditioned simulation values
$Z_{UC}^*(x)$	kriging or cokriging estimate of unconditioned simulation

$\vec{Z}_{UC}^{*}(x)$ vector of kriging or cokriging estimates using unconditioned simulation values

$Z_{UC,D}(x)$ unconditioned DCPA simulation value

$Z_{UC,D}^{*}(x)$ kriging or cokriging estimate using unconditioned DCPA simulation values

$Z_{UC,N}(x)$ unconditioned nitrate simulation value

$Z_{UC,N}^{*}(x)$ kriging or cokriging estimate using unconditioned nitrate simulation values

References

Carr, J.R., and Myers, D.E., COSIM: A FORTRAN IV program for coconditional simulation, *Computers & Geosciences, 11(6)*, 675-705, 1985.

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

Dagan, G., Stochastic modeling of groundwater flow by unconditional and conditional probabilities, 1. Conditional simulation and the direct problem, *Water Resources Research, 18(4)*, 813-833, 1982.

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

Journel, A.G., Geostatistics for conditional simulation of ore bodies, *Economic Geology, Vol. 69*, 673-687, 1974.

Journel, A.G., and Huijbregts, C.J., *Mining Geostatistics* , 600 pp., Academic Press, New York NY, 1978.

Luster, G.R., Raw materials for Portland cement: Applications of conditional simulation of coregionalization, Ph. D. dissertation, 531 pp., Stanford University, 1986.

Mantoglou, A., and Wilson, J.L., *Simulation of Random Fields With The Turning Bands Method,* Report No. 264, Ralph M. Parsons Laboratory, Department of Civil Engineering, Massachusetts Institute of Technology, Cambridge, 199pp, 1981.

Mantoglou, A., and Wilson, J.L, The turning bands method for simulation of random fields using line generation by a spectral method, *Water Resources Research, 18(5)*, 1379-1394, 1982.

Myers, J.C., and Bryan, R.C., Geostatistics applied to toxic waste - A case study, in *Geostatistics for Natural Resource Characterization, Part 2*, G. Verely, M. David, A.G. Journel, and A. Marechal, Eds., 893-901, D. Reidel Publishing Company, Dordrecht, Holland, 1984.

Smyth, J.D., and Istok, J.D., Multivariate geostatistical analysis of groundwater contamination by pesticide and nitrate, submitted for publication, 1988.

Van Rooy, D., Conditional stochastic simulation of groundwater contamination, A case study, Ph.D. disertation, 154 pp., Institute of Hydrodynamics and Hydraulic Engineering, Technical University of Denmark, DK-Lyngby, 1986.

Verly, G.W., David, M., Journel, A.G., and Marechal, A., eds., *Geostatistics for Natural Resources Characterization*, Proceedings of the NATO Advanced Study Institute, South Lake Tahoe CA, Sept 6-17, D. Reidel, Dordrecht, Holland, 1092 pp, 1983.

BIBLIOGRAPHY

Barnes, M.G., Statistical design and analysis in the cleanup of environmental radionuclide contamination, DRI publication no. 45012, Desert Research Institute, University of Nevada, Reno NV, 1978.

Barnes, M.G., Giacomini, J.J., Reiman, R.T., and Elliot, B., NTS Radiological assesment program: Results for Frenchman Lake Area 5, DOE/DP/01253-17, Desert Research Institute, University of Nevada, Reno NV, 1980.

Beste, C.E., Chairman of the Herbicide Handbook Committee, *Herbicide Handbook of the Weed Science Society of America 5 th Edition,* 492 pp., Weed Science Society of America, Chicago IL, 1983.

Carr, J.R., and Myers, D.E., COSIM: A FORTRAN IV program for coconditional simulation, *Computers & Geosciences, 11(6)*, 675-705, 1985.

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

Cooper, R.M., and Istok, J.D., Geostatistics applied to Groundwater Contamination, 1. Methodology, *J. Envir. Engrg., ASCE, 114(2)*, 270-286, 1988a.

Cooper, R.M., and Istok, J.D., Geostatistics applied to Groundwater Contamination, 2. Application, *J. Envir. Engrg., ASCE, 114(2)*, 287-299, 1988b.

Dagan, G., Stochastic modeling of groundwater flow by unconditional and conditional probabilities, 1. Conditional simulation and the direct problem, *Water Resources Research, 18(4)*, 813-833, 1982.

David, M., *Geostatistical Ore Reserve Estimation*, 364 pp., Elsevier Scientific Publishing Company, New York NY, 1977.

Delhomme, J.P., Kriging in the hydrosciences, *Advances in Water Resources, 1(5)*, 475-499, 1978.

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

Flaig, E.G., Bottcher, A.B., and Cambell, K.L., Estimatation of nitrate leaching using geostatistics, Presented at the ASAE Summer Meeting , San Luis Obispo CA, June 29-July 2, 1986.

Hoeksema, R.J. and Clapp, R.B., Estimating water table elevations using cokriging, (abstract), *EOS Trans. AGU, 68*, 44, 1987.

Istok, J.D., and Cooper, R.M., Geostatistics applied to Groundwater Contamination, 3. Global Estimation, *J. Envir. Engrg., ASCE*, in press, 1988.

Istok, J.D., Cooper, R.M., and Flint, A.L., Three-dimensional cross-semivariogram calculations for hydrogeologic data, *Groundwater*, in press, 1988.

Journel, A.G., Geostatistics for conditional simulation of ore bodies, *Economic Geology, Vol. 69*, 673-687, 1974.

Journel, A.G., and Huijbregts, C.J., *Mining Geostatistics* , 600 pp., Academic Press, New York NY, 1978.

Journel, A.G., and Issaaks, E.H., Conditional indicator simulation: Application to a Saskatchewan uranium deposit, *Jour Math Geol*, 16(7), 685-718, 1984.

Luster, G.R., Raw materials for Portland cement: Applications of conditional simulation of coregionalization, Ph. D. dissertation, 531 pp., Stanford University, 1986.

Mantoglou, A., and Wilson, J.L., *Simulation of Random Fields With The Turning Bands Method,* Report No. 264, Ralph M. Parsons Laboratory, Department of Civil Engineering, Massachusetts Institute of Technology, Cambridge, 199pp, 1981.

Mantoglou, A., and Wilson, J.L, The turning bands method for simulation of random fields using line generation by a spectral method, *Water Resources Research, 18(5),* 1379-1394, 1982.

Marsily, de Ghislain, *Quantitative Hydrogeology, Groundwater Hydrology for Engineers,* 439 pp., Academic Press, New York NY, 1986.

Myers, D.E., Matrix formulation of cokriging, *Jour. Math. Geol., 14(3),* 249-257, 1982.

Myers, D.E., Begovich, C.L., Butz, T.R, and Kane, V.E., Variogram models for regional groundwater data, *Int. Ass. Math. Geol. Jour., 14(6),* 629-644, 1982.

Myers, J.C., and Bryan, R.C., Geostatistics applied to toxic waste - A case study, in *Geostatistics for Natural Resource Characterization, Part 2,* G. Verely, M. David, A.G. Journel, and A. Marechal, Eds., 893-901, D. Reidel Publishing Company, Dordrecht, Holland, 1984.

SAS/Institute Inc., SAS/STAT *Guide for Personal Computers, Version 6 Edition,* 378 pp., Cary NC: SAS Institute Inc., 1985.

United States Environmental Protection Agency, *Dacthal Draft Health Advisory,* 13 pp., Office of Drinking Water, August, 1987.

Van Rooy, D., Conditional stochastic simulation of groundwater contamination, A case study, Ph.D. disertation, 154 pp., Institute of Hydrodynamics and Hydraulic Engineering, Technical University of Denmark, DK-Lyngby, 1986.

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

Verly, G.W, The block distribution given a point multivariate normal distribution, *Geostatistics for Natural Resource Characterization, Part 1* , G. Verely, M. David, A.G. Journel, and A. Marechal, Eds., 495-515, D. Reidel Publishing Company, Dordrecht, Holland, 1984.

Verly, G.W., David, M., Journel, A.G., and Marechal, A., eds., *Geostatistics for Natural Resources Characterization*, Proceedings of the NATO Advanced Study Institute, South Lake Tahoe CA, Sept 6-17, D. Reidel, Dordrecht, Holland, 1092 pp, 1983.

Whitson, T.D., William, R.D., Parker, R., Swan, D.G., Dewey, S., Eds., *Pacific Northwest Weed Control Handbook*, 213 pp., Extension Service, Oregon State University, Wahington State University, and the University of Idaho, 1987.

Zirschky, J.H., Keary, G.P., Gilbert, R.O., and Middlebrooks, E.J., Spatial estimation of hazardous waste site data, *J. Envir. Engrg., ASCE, Vol 111*, 777-789, 1985.

Zirschky, J.H., and Harris, D.J., Geostatistical analysis of hazardous waste data, *J. Envir. Engrg., ASCE, Vol 112*, 770-785, 1986.