R82-51 SP 89253

MIT

A COMPARISON OF LINEAR AND NONLINEAR RANDOM FIELD ESTIMATORS

by

CARLOS ENRIQUE PUENTE ANGULO
and
RAFAEL L. BRAS

RALPH M. PARSONS LABORATORY
HYDROLOGY AND WATER RESOURCES SYSTEMS

Department of Civil Engineering
Massachusetts Institute of Technology

DEPARTMENT
OF
CIVIL
ENGINEERING

SCHOOL OF ENGINEERING
MASSACHUSETTS INSTITUTE OF TECHNOLOGY
Cambridge, Massachusetts 02139

Report No. 287

Prepared with support of National Science Foundation and Office of Surface Mining through Grants CME-7919836 and G5105071

November 1982

R82-51

A COMPARISON OF LINEAR AND NONLINEAR RANDOM FIELD ESTIMATORS

by

CARLOS ENRIQUE PUENTE ANGULO

and

RAFAEL L. BRAS

RALPH M. PARSONS LABORATORY

HYDROLOGY AND WATER RESOURCES SYSTEMS

Department of Civil Engineering

Massachusetts Institute of Technology

Report No. 287

Prepared with support of National Science Foundation and Office of Surface Mining through Grants CME-7919836 and G5105071

November 1982

A COMPARISON OF LINEAR AND NONLINEAR RANDOM FIELD ESTIMATORS

by

CARLOS ENRIQUE PUENTE ANGULO

Licenciado en Matemáticas, Universidad de los Andes
Bogotá, Colombia
(1978)

Ingeniero Civil, Universidad de los Andes, Bogotá, Colombia
(1979)

Submitted to the Department of Civil Engineering on November 5, 1982 in partial fulfillment of the requirements for the degree of Master of Science in Civil Engineering.

ABSTRACT

The estimation of random fields from limited samples is an important issue in most fields of geophysics, such as Hydrology and Meteorology. Work by Matheron and others at the Paris School of Mines has popularized Kriging techniques to estimate random fields at specified locations or to get areal averages.

This work presents the theoretical and practical aspects of both Linear and Nonlinear (Disjunctive) Kriging estimators, and provides a comparison of their performance in estimating point and areal values of generated fields. The experiments performed were designed to closely resemble actual and practical situations.

The results show that small sample based inconsistencies lead to a Disjunctive Kriging solution which does not give more accurate estimates than the theoretically less precise Linear Kriging estimator. The results also suggest the use of a multi-realization approach when using these techniques in network design problems.

Thesis Supervisor: Rafael L. Bras
Title: Associate Professor of Civil Engineering

ACKNOWLEDGEMENTS

This report essentially constitutes the Masters of Science degree thesis of Carlos E. Puente Angulo under the supervision of Professor Rafael L. Bras. The work was sponsored by the National Science Foundation and Office of Surface Mining through Grants CME-7919836 and G5105071. All the computer work was performed at the MIT Information Processing Center.

Discussions with Aris Georgakakos, Telis Mantoglou, Stefanos Andreous, Diana Kirshen, Scott Miller, Jorge Ramirez, Pedro Restrepo, Mark Alexandridis, Konstantine Georgakakos, and Mario Diaz-Granados have been useful and they are rightfully acknowledged.

The fine typing of this report is due to Suzanne Shanahan and Antoinette DiRenzo.

TABLE OF CONTENTS

Page

LIST OF TABLES

LIST OF FIGURES

LIST OF PRINCIPAL SYMBOLS

Symbols	Title
A	Domain of the field $Z(\underline{u})$ and $Y(\underline{u})$, $\underline{u} \varepsilon A$
a_*, a	Constants
a_o	Range of the spherical semivariogram
a_k	Constant on the linear interpolation of the anamorphosis, $k = 1, \ldots, N-1$
b	A constant, spatial correlation parameter of exponential covariance
b_k	Constant on the linear interpolation of the anamorphosis, $k = 1, \ldots, N-1$
C	Coefficient of nugget effect on generalized covariance function
C_o	Sill of the spherical semivariogram
CC	Correlation coefficient between true and estimated values
CD	Correlation distance at which the correlation function has dropped to a half
$Cov(\underline{u}_1, \underline{u}_2)$	Covariance function of Z at points $\underline{u}_1$ and $\underline{u}_2$
$Cov_Y(\cdot)$	Covariance function of the gaussian data $Y(\underline{u})$
c	A constant
D	Generalized covariance function results estimated by Delfiner's approach
DK_s	Disjunctive Kriging estimator calculated using a spherical variogram of the gaussian data
$E(\cdot)$	Expectation operator
$F(\cdot)$	Univariate distribution function of the field $Z(\underline{u})$
$F(\cdot,\cdot)$	Bivariate distribution function of the field $Z(\underline{u})$
$f(\cdot,\cdot,\ldots,\cdot)$	N-dimensional measurable function
$f_i(\cdot)$	1-dimensional measurable function, $i = 1, \ldots, N$

Symbols	Title
TD	Mean distance of two points taken randomly and uniformly on a figure (Typical Distance)
T_i^P	Constanst in regression equations in the calculation of generalized covariance functions, $i = 1,\ldots, N$, $P = 0,1,\ldots, k$
$\underline{u}$	Location in region A, $\underline{u} = (u,v)$
$\underline{u}_1,\underline{u}_2$	Location in region A
$\underline{u}_o$	Location at the site of interest
$\underline{u}_*$	Arbitrary location
$V, V_{\underline{u}_o}$	Block to be estimated, $V \subseteq A$
$Var(\cdot)$	Variance function of the field Z
$Y(\underline{u})$	Gaussian variate at location $\underline{u} \varepsilon A$
$Z(\underline{u})$	Field of interest at location $\underline{u} \varepsilon A$
$Z(\underline{u}_i), Z_i$	Observations of the field at points $\underline{u}_i$, $i = 1,\ldots, N$
$\hat{Z}_i$	Estimate of the field at point $\underline{u}_i$, $i = 1,\ldots, N_1$
$Z(\lambda_i)$	Constructed generalzied increment at point $\underline{u}_i$
$Z_k(\underline{u})$	Field obtained after $k^{\underline{th}}$ order increments
α_j	Constants on generalized covariance function, $j = 1,3,5$
$\gamma(\underline{u}_1,\underline{u}_2)$	Semivariogram function at locations $\underline{u}_1$ and $\underline{u}_2$
$\delta_{\underline{u}_i}$	Indicator function of field at location $\underline{u}_1$
δ_ℓ	Coefficients in the drift, $\ell = 1,\ldots, kl$
$\psi(u,v)$	Bivariate gaussian density function of $Y(\underline{u})$ and $Y(\underline{v})$
ξ_k	Constants, $k = 0, 1,\ldots, ml$
$\eta_k(\cdot)$	$k^{\underline{th}}$ standardized Hermite polynomial
λ_i	Kriging weights, $i = 1,\ldots,N$
$\lambda_i{}^*$	Optimal Kriging weights, $i = 1,\ldots,N$
$\lambda(d\underline{u})$	Measure that allow the definition of extended combinations

Symbol	Title
$f_{i,k}$	Hermite expansion coefficients of the function $f_i(\cdot)$, $i = 1,\ldots, N$, $k = 0,1,\ldots$
$f_{i,k}^*$	Optimal coefficients of $f_i(\cdot)$, solution of the Disjunctive Kriging systems, $i = 0,1\ldots$
$f_\ell(\cdot)$	Functions that determine the drift form, $\ell = 1,\ldots,kl$
GC	Generalized covariance function
$G(du)$	Univariate Gaussian distribution function
$G(du,dv)$	Bivariate Gaussian distribution function
$G_{u/v}(du)$	Conditional distribution of $Y(\underline{u})$ given $Y(\underline{v})$
$g(\cdot)$	Univariate Gaussian density function
$\underline{h}$	Distance under stationarity assumptions on the covariance, $\underline{h} = \underline{u}_1 - \underline{u}_2$
h	Distance under isotropic assumptions on the covariance and generalized covariances, $h = \lvert\underline{h}\rvert$
IRF	Intrinsic Random Function
J_1,J_2	Disjoint groups in which the data is splitted in the calculation of the jacknife estimator $\hat{\rho}$
j_i	Non-negative integers
$K(\cdot)$	Generalized covariance function
K_M	Universal Kriging estimator calculated using the least mean square error procedure on the estimation of the generalized covariance function
KT	Universal Kriging estimator of Gaussian variables and transformed via the anamorphosis function
k	Intrinsic Random Functions order, $k = 0,1,2$
kl	Number of functions used in the drift form
L	Disjunctive kriging estimator calculated using a linear variogram of the gaussian data
$L(\cdot,\cdot,\ldots,\cdot)$	Lagrangian function
$\mathscr{L}(Z_o)$	Quantity at point $\underline{u}_0$ to be estimated.

Symbols	Title
$\mathscr{L}(Z_o)^*_K$	Linear Kriging estimator of the unknown
$\mathscr{L}(Z_o)^*_{DK}$	Disjunctive Kriging estimator of the unknown
M	Universal Kriging estimator calculated using the least mean square error procedure on the estimation of the generalized covariance function.
M_5	Local mean estimator calculated with five neighboring points
MaPV	Maximum predicted variance over the estimated grid used in the comparisons
MPV	Mean predicted variance over the estimated grid used in the comparisons
MPV_1	Mean predicted variance predicted by the jacknife estimator
MSE	Mean square error of estimation over the grid used in the comparisons
m	Constant mean
m1	Cutoff on the Hermite polynomial expansions
m2	Number of feasible generalized covariance models
$m(\underline{u})$	Drift at location $\underline{u} \varepsilon A$
N	Number of points used in the estimation (Number of observations)
N_0	Neighboring points used in the calculation of generalized increments
N_1	Number of points used in the comparisons
$P_k(\cdot)$	Polynomial of degree less than or equal to k
Q	Square error on regressions in the calculation of generalized covariance functions
R	Universal Kriging estimator calculated using the ranking procedure on the estimation of the generalized covariance function
r	Theoretical ratio of the true over the predicted variances
S	Disjunctive Kriging estimator calculated using a spherical variogram of the gaussian data

Symbols	Title
λ_{ij}	Weights in the calculation of $Z(\underline{u}_i)$ from neighboring points, $i = 1,\ldots, N$, $j = 1,\ldots, N_0$
$\phi(\cdot)$	Anamorphosis function
ϕ_k	Hermite expansion coefficients of the anamorphosis function $k = 0,1,\ldots$
ρ	Estimator of r, the ratio of true over predicted variance
$\hat{\rho}$	Jackknife estimator of r, the ratio of true over predicted variance
$\hat{\rho}_1$	Consistency parameter, ratio of MSE over MPV
ρ_{uv}	Correlation between $Y(\underline{u})$ and $Y(\underline{v})$
μD	Mean of the true minus the estimated values over the grid
μ_ℓ	Lagrange multipliers, $\ell = 1,\ldots, kl$
σ^2	Variance of exponential covariance function
σ^2_{CE}	Conditional expectation variance
σ^2_{DK}	Disjunctive Kriging variance
σ^2_K	Linear Kriging variance
σ^{2*}_{DK}	Optimal minimum Disjunctive Kriging variance
σ^{2*}_K	Optimal miminum Linear Kriging variance
σ^2_i	Estimated variance at point $\underline{u}_i$, $i = 1,\ldots, N$
σD	Standard deviation of the true minus estimated values over the grid
$\%\mu$	Percent error in fitting the mean of Z by using an anamorphosis finite expansion and a given interpolation
$\%V$	Percent error in fitting the variance of Z by using an anamorphosis finite expansion and a given interpolation

Chapter 1

INTRODUCTION

1.1 Motivation and Objectives

Sampling in the physical sciences is by necessity finite, both in the time and space domains. Nevertheless, it is often necessary to obtain descriptions of the physical processes in points at regions where data is not fully available. Such is commonly the case in measuring precipitation, piezometric heads, soil properties, wind fields, and other processes of interest to hydrologists. The theory of random fields, and specifically the Linear and Disjunctive Kriging techniques introduced by Matheron, 1971, 1973; 1976a, are designed to allow a systematic solution to the above problems, using the spatial structure of the quantity of interest.

Many authors have applied the Linear Kriging technique in Hydrology. Applications include groundwater modelling as in Delhomme, 1979, Gambolati and Volpi, 1979, Volpi et al. 1979, Rouhani, 1981; precipitation modelling as in Delfiner and Delhomme, 1973, Delhomme and Delfiner, 1973, Delhomme 1976, 1978, Shaw and O'Connell, 1976, Shaake, 1978, Chua and Bras, 1980; and water quality modelling as in Hughes and Lettenmaier, 1980, Doctor and Nelson, 1981. Disjunctive Kriging, developed after the Linear Kriging estimator, has not yet been applied, although as it will be seen its potential is similar to that of the linear estimator.

Although the theory of both estimators is well documented, comparisons of the Disjunctive and Linear Kriging estimators in practice have been few. Among them the following works do compare them in mining applications: Marechal, 1976a, 1976b, Rendu, 1980. In all of them, however, many more data points than usually found in Hydrology were employed.

The goal of this work is to explore the advantages and disadvantages of the Linear and Disjunctive Kriging estimators when they are calculated using the number of data points and spatial structures found in hydrological applications.

1.2 Report Outline

The theoretical basis and definitions of the Linear Kriging estimator are given in Chapter 2 under different assumptions about the data sets. The theory of Intrinsic Random Functions and the practical estimation of generalized covariances are also reviewed there. In Chapter 3, the theoretical and practical aspects of the Disjunctive Kriging estimator are presented. A description of the experiments that were performed together with the results of the point estimation comparisons are given in Chapter 4. In Chapter 5, the results of the experiments made to compare performance on block estimation are given. Results and conclusions, as well as recommendations for further research, are given in Chapter 6.

Chapter 2

OPTIMAL LINEAR ESTIMATION: LINEAR KRIGING

In this chapter the theoretical basis and definitions of the Linear Kriging estimator are given, under various assumptions. The theory of Intrinsic Random Functions and the practical application of generalized covariances are reviewed. Treatment follows Delfiner, 1976; Chua and Bras, 1980; and Kafritsas and Bras, 1981.

2.1 Basic Definitions

Of interest is the modelling of a random field $Z(\underline{u})$, given by a family of random variables defined over points $\underline{u}$ in a region A of the Eucledian space $\mathbb{R}^n$. Associated to the random field there is a family of cumulative distribution functions $F(\underline{u})$ that allow the definition of the mean or drift of the field as:

$$m(\underline{u}) = E[Z(\underline{u})] = \int_A Z(\underline{u})dF(\underline{u}) \tag{2.1}$$

If the mean does not depend on the vector location $\underline{u}$, it is said that the field is stationary in the mean.

The covariance function of the field is defined by:

$$\mathrm{Cov}(\underline{u}_1,\underline{u}_2) = E(Z(\underline{u}_1) \cdot Z(\underline{u}_2)) - m(\underline{u}_1)m(\underline{u}_2) \tag{2.2}$$

where the expected value of the product of $Z(\underline{u}_1)$ and $Z(\underline{u}_2)$, or correlation of the field, is given by:

$$E[Z(\underline{u}_1)Z(\underline{u}_2)] \quad = \quad \int_A \int_A Z(\underline{u}_1)Z(\underline{u}_2)\; dF(\underline{u}_1,\underline{u}_2) \qquad (2.3)$$

being $F(\underline{u}_1,\underline{u}_2)$ the bivariate distribution of $Z(\underline{u}_1)$ and $Z(\underline{u}_2)$.

When $\underline{u}_1 = \underline{u}_2$, equation (2.2) yields the variance:

$$\mathrm{Var}(\underline{u}) \quad = \quad \mathrm{Cov}(\underline{u},\underline{u}) \qquad (2.4)$$

If the covariance function depends only on the distance between the points $\underline{u}_1$ and $\underline{u}_2$, the field is said to be stationary in the covariance, and then can be written as:

$$\mathrm{Cov}(\underline{u}_1,\underline{u}_2) \quad = \quad \mathrm{Cov}(\underline{h}) \qquad (2.5)$$

where $\underline{h} = \underline{u}_1 - \underline{u}_2$.

If a field is stationary on the mean and in the covariance, it is called second-order stationary.

If the covariance function is the same in every direction, the field is called isotropic, which gives:

$$\mathrm{Cov}(\underline{h}) \quad = \quad \mathrm{Cov}(|\underline{h}|) \quad = \quad \mathrm{Cov}(h) \qquad (2.6)$$

where $h = |\underline{h}|$.

The semivariogram function is defined as:

$$\gamma(\underline{u}_1,\underline{u}_2) = \frac{1}{2} E\{(Z(\underline{u}_1) - Z(\underline{u}_2))^2\} - \frac{1}{2}[m(\underline{u}_1) - m(\underline{u}_2)]^2 \tag{2.7a}$$

$$= \frac{1}{2} \operatorname{Var}\{Z(\underline{u}_1) - Z(\underline{u}_2)\} \tag{2.7b}$$

Since,

$$\frac{1}{2} \operatorname{Var}\{Z(\underline{u}_1) - Z(\underline{u}_2)\} = \frac{1}{2} \operatorname{Var}\{Z(\underline{u}_1)\} + \frac{1}{2} \operatorname{Var}\{Z(\underline{u}_2)\} - \operatorname{Cov}(\underline{u}_1,\underline{u}_2)$$

under stationarity on the covariance function, the semivariogram and covariance function are related by:

$$\gamma(\underline{u}_1 - \underline{u}_2) = \operatorname{Cov}(\underline{0}) - \operatorname{Cov}(\underline{u}_1 - \underline{u}_2) \tag{2.8}$$

The random field $Z(\underline{u})$ is said to satisfy the intrinsic hypothesis if its first order differences $Z(\underline{u}_1) - Z(\underline{u}_2)$ are stationary in the mean and in the variance, i.e.:

$$E[Z(\underline{u}_1) - Z(\underline{u}_2)] = m(\underline{h}) \tag{2.9a}$$

and

$$\operatorname{Var}[Z(\underline{u}_1) - Z(\underline{u}_2)] = 2\gamma(\underline{h}) \tag{2.9b}$$

for any $\underline{u}_1 - \underline{u}_2 = \underline{h}$. Clearly, second-order stationarity implies the intrinsic hypothesis, but not vice-versa.

In later sections, the Linear Kriging estimator will be given under second-order stationarity and intrinsic hypothesis assumptions.

2.2 Problem Definition

Given a realization of a random field Z defined over a region A, say $Z(\underline{u}_1)$, $Z(\underline{u}_2)$ ..., $Z(\underline{u}_N)$, or simply Z_1, Z_2 ..., Z_N, the purpose in this work is:

<u>Point Estimation</u>: determine $Z(\underline{u}_0)$, at an arbitrary point $\underline{u}_0$ on the region A.

<u>Block Estimation</u>: determine the mean value of the field over an area included in A.

$$\frac{1}{V_{\underline{u}_0}} \int_{V_{\underline{u}_0}} Z(\underline{u})\, d\underline{u} \tag{2.10}$$

where $V_{\underline{u}_0}$, denoted simply V, is for example a rectangle with $\underline{u}_0$ as its middle point.

The general unknown will be denoted $\mathscr{L}(Z(\underline{u}_0))$ or $\mathscr{L}(Z_0)$, because the procedures, to be explained, can be applied not only to point and block estimation, but also to more general linear functions of the field.

2.3 Characteristics of the Linear Kriging Estimator

The usual Kriging estimator is characterized by the following conditions:

1. Linearity: The estimator, $\mathscr{L}(Z_0)^*_K$, is a linear combination of the observed variables, Z_i, i=1, ..., N:

$$\mathscr{L}(Z_0)^*_K = \sum_{i=1}^{N} \lambda_i Z_i \tag{2.11}$$

where the weights λ_i, i=1, ..., N are chosen such that the next two conditions are satisfied.

2. Unbiasedness: The combination, $\mathscr{L}(Z_0)^*_K$, must be an unbiased estimator of the unknown $\mathscr{L}(Z_0)$; this means:

$$E[\mathscr{L}(Z_0)^*_K] = E[\mathscr{L}(Z_0)] \tag{2.12}$$

which gives the following condition on the weights:

$$\sum_{i=1}^{N} \lambda_i \, m(\underline{u}_i) = E[\mathscr{L}(Z_0)] \tag{2.13}$$

3. Minimum variance: The estimator, $\mathscr{L}(Z_0)^*_K$, is such that its variance of estimation:

$$\sigma_K^2 = \mathrm{Var}\{\mathscr{L}(Z_0) - \mathscr{L}(Z_0)^*_K\} \tag{2.14}$$

is minimized over all linear combinations that satisfy equation (2.13).

If the variance of estimation is expanded and the unbiasedness condition employed, it can be shown that σ_K^2 is also the mean square error of estimation, MSE, or:

$$\sigma_K^2 = E\{(\mathscr{L}(Z_0) - \mathscr{L}(Z_0)_K^*)^2\} \tag{2.15}$$

This can be easily shown to be:

$$\sigma_K^2 = E[\mathscr{L}(Z_0)^2] - 2\sum_{i=1}^{N} \lambda_i E[\mathscr{L}(Z_0)Z_i]$$

$$+ \sum_{i=1}^{N}\sum_{j=1}^{N} \lambda_i\lambda_j\ E[Z_iZ_j] \tag{2.16a}$$

or equivalently in terms of the covariance function:

$$\sigma_K^2 = \mathrm{Cov}(\mathscr{L}(Z_0),\mathscr{L}(Z_0)) - 2\sum_{i=1}^{N} \lambda_i \mathrm{Cov}(\mathscr{L}(Z_0),Z_i)$$

$$+ \sum_{i=1}^{N}\sum_{j=1}^{N} \lambda_i\lambda_j\ \mathrm{Cov}(Z_i,Z_j) \tag{2.16b}$$

Then, in summary, the Linear Kriging estimator is found from the minimization problem:

$$\min_{\lambda_1,\ldots,\lambda_N} \quad \sigma_K^2 = E[\mathscr{L}(Z_0)^2] - 2\sum_{i=1}^{N} \lambda_i E[\mathscr{L}(Z_0)Z_i]$$

$$+ \sum_{i=1}^{N}\sum_{j=1}^{N} \lambda_i\lambda_j \; E[Z_iZ_j] \tag{2.17}$$

$$\text{s.t.} \quad \sum_{i=1}^{N} \lambda_i \; m(\underline{u}_i) = E[\mathscr{L}(Z_0)]$$

In the following sections, the Linear Kriging estimator will be given, for point and block estimation, under different assumptions on the field.

2.4 Kriging Under Second-Order Stationarity Assumptions

In this section, the equations that determine the optimal weights, λ_i, i=1, ..., N, are given for the case in which $\mathrm{Cov}(\underline{h})$ and the constant mean m, are known.

The unbiasedness condition, equation (2.13), reduces to the condition:

$$\sum_{i=1}^{N} m \cdot \lambda_i = E[\mathscr{L}(Z_0)] \tag{2.18}$$

The variance of estimation can be written in terms of the covariance function as follows:

$$\sigma_K^2 = E[\mathscr{L}(Z_0)^2] - 2 \sum_{i=1}^{N} \lambda_i E[\mathscr{L}(Z_0) Z_i]$$

$$+ \sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_i \lambda_j \{\mathrm{Cov}(\underline{u}_i - \underline{u}_j) + m^2\} \tag{2.19}$$

In the following subsections, these previous two equations will be specialized for the interesting cases of point and block estimation.

2.4.1 Point Kriging Under Second-Order Stationarity

For point estimation, the unknown is simply $\mathscr{L}(Z_0) = Z_0$, which reduces equation (2.18) to:

$$\sum_{i=1}^{N} \lambda_i = 1 \tag{2.20}$$

The variance of estimation can be written as:

$$\sigma_K^2 = \mathrm{Cov}(\underline{0}) + m^2 - 2 \sum_{i=1}^{N} \lambda_i \{\mathrm{Cov}(\underline{u}_0 - \underline{u}_i) + m^2\}$$

$$+ \sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_i \lambda_j \{\mathrm{Cov}(\underline{u}_i - \underline{u}_j) + m^2\} \tag{2.21}$$

which gives, replacing the unbiasedness condition, the expression:

$$\sigma_K^2 = \mathrm{Cov}(\underline{0}) - 2 \sum_{i=1}^{N} \lambda_i \, \mathrm{Cov}(\underline{u}_0 - \underline{u}_i)$$

$$+ \sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_i \lambda_j \, \mathrm{Cov}(\underline{u}_i - \underline{u}_j) \tag{2.22}$$

Recall that the Kriging estimator is found when σ_K^2 is minimized subject to the unbiasedness condition (see equation (2.17)). This minimization problem can be solved using the Lagrange multipliers technique.

The Lagrangian function becomes:

$$L(\lambda_1, \ldots, \lambda_N, \mu) = \sigma_K^2 + 2\mu \left(\sum_{i=1}^{N} \lambda_i - 1 \right) \tag{2.23}$$

which gives as the necessary conditions:

$$\frac{\partial L}{\partial \lambda_i} = -2\ \mathrm{Cov}(\underline{u}_0 - \underline{u}_i) + 2 \sum_{j=1}^{N} \lambda_j\ \mathrm{Cov}(\underline{u}_i - \underline{u}_j) + 2\mu = 0 \qquad (2.24a)$$

$$i=1,\ ..,\ N$$

$$\frac{\partial L}{\partial \mu} = 2\left(\sum_{i=1}^{N} \lambda_i - 1\right) = 0 \qquad (2.24b)$$

This simultaneous set of equations, on the unknown weights $\lambda_1, \ldots, \lambda_N$ and the Lagrange multiplier μ, is called the Kriging system. It can be written in matrix form as:

$$\begin{bmatrix} \mathrm{Cov}(\underline{0}) & \mathrm{Cov}(\underline{u}_1-\underline{u}_2) & \ldots & \mathrm{Cov}(\underline{u}_1-\underline{u}_N) & 1 \\ \mathrm{Cov}(\underline{u}_2-\underline{u}_1) & \mathrm{Cov}(\underline{0}) & \ldots & \mathrm{Cov}(\underline{u}_2-\underline{u}_N) & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ \mathrm{Cov}(\underline{u}_N-\underline{u}_1) & \mathrm{Cov}(\underline{u}_N-\underline{u}_2) & \ldots & \mathrm{Cov}(\underline{0}) & 1 \\ 1 & 1 & \ldots & 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_N \\ \mu \end{bmatrix} = \begin{bmatrix} \mathrm{Cov}(\underline{u}_0-\underline{u}_1) \\ \mathrm{Cov}(\underline{u}_0-\underline{u}_2) \\ \vdots \\ \mathrm{Cov}(\underline{u}_0-\underline{u}_N) \\ 1 \end{bmatrix}$$

(2.25)

This system will give a unique solution $\lambda_1^*, \lambda_2^*, \ldots, \lambda_N^*, \mu^*$ if the left-hand matrix is non-singular, which clearly correspond to the minimum.

If the equations (2.24a) are multiplied by the respective optimal λ_i^*, and then added over values of i, the following expression is reached:

$$\sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_i^* \lambda_j^* \, Cov(\underline{u}_i - \underline{u}_j) = \sum_{i=1}^{N} \lambda_i^* \, Cov(\underline{u}_0 - \underline{u}_i) - \sum_{i=1}^{N} \lambda_i^* \mu^*$$

If this is replaced back on equation (2.22), the optimal minimum variance of estimation is found to be:

$$\sigma_K^{2^*} = Cov(\underline{0}) - \sum_{i=1}^{N} \lambda_i^* \, Cov(\underline{u}_0 - \underline{u}_i) - \mu^* \qquad (2.26)$$

Observe that if the point $\underline{u}_0$ coincides with some of the $\underline{u}_i$, the estimated value will be exactly the true value, with estimation variance zero. The solution to the Kriging system will be zero for μ and all λ's except the one corresponding to the point in question, which takes the value 1. This will be true for any Linear Kriging estimator.

Note that the Kriging system and the optimal variance of estimation can also be expressed in terms of the semivariogran function, using equation (2.8), valid under second-order stationarity assumptions. This readily gives the system and estimation variance as follows:

$$\begin{bmatrix} 0 & -\gamma(\underline{u}_1-\underline{u}_2) & \cdots & -\gamma(\underline{u}_1-\underline{u}_N) & 1 \\ -\gamma(\underline{u}_2-\underline{u}_1) & 0 & \cdots & -\gamma(\underline{u}_2-\underline{u}_N) & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ -\gamma(\underline{u}_N-\underline{u}_1) & -\gamma(\underline{u}_N-\underline{u}_2) & \cdots & 0 & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_N \\ \mu \end{bmatrix} = \begin{bmatrix} -\gamma(\underline{u}_0-\underline{u}_1) \\ -\gamma(\underline{u}_0-\underline{u}_2) \\ \vdots \\ -\gamma(\underline{u}_0-\underline{u}_N) \\ 1 \end{bmatrix} \qquad (2.27)$$

and

$$\sigma_K^{2^*} = \sum_{i=1}^{N} \lambda_i^* \gamma(\underline{u}_0 - \underline{u}_i) - \mu^* \tag{2.28}$$

It is important to note that under second-order stationarity conditions, the use of the semivariogram avoids the need of estimating the actual value of the mean, which in any case would introduce bias on the estimated covariance function. In the case of stationarity:

$$\gamma(\underline{u}_1 - \underline{u}_2) = \frac{1}{2} E\{(Z(\underline{u}_1) - Z(\underline{u}_2))^2\} \tag{2.29}$$

which indicates that the semivariogram can be estimated from the differences of the field, without the knowledge of the constant mean that has been filtered out.

2.4.2 Block Kriging Under Second-Order Stationarity

For block estimation, the unknown, $\mathscr{L}(Z_0)$, is the average of the field over a region V, having $\underline{u}_0$ as its middle point, i.e.:

$$\mathscr{L}(Z_0) = \frac{1}{V} \int_V Z(\underline{u}) \, d\underline{u}$$

This replaced on equation (2.18) gives the unbiasedness condition:

$$\sum_{i=1}^{N} m \lambda_i = \frac{1}{V} \int_V m \, d\underline{u}$$

which is clearly the same as:

$$\sum_{i=1}^{N} \lambda_i = 1 \tag{2.30}$$

Replacing the expression of $\mathscr{L}(Z_0)$ on the variance of estimation, equation (2.19)[1], and using equation (2.30), the quantity σ_K^2 can be written in terms of the covariance function, as:

$$\sigma_K^2 = \frac{1}{V^2} \int_V \int_V \mathrm{Cov}(\underline{u} - \underline{v}) d\underline{u}\, d\underline{v} - 2 \sum_{i=1}^{N} \lambda_i \frac{1}{V} \int_V \mathrm{Cov}(\underline{u}_i - \underline{u})\, d\underline{u}$$

$$+ \sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_i \lambda_j \mathrm{Cov}(\underline{u}_i - \underline{u}_j) \tag{2.31}$$

The optimal weights are found using, as before, the Lagrange multipliers technique. This readily gives the objective function:

$$L(\lambda_1, \ldots, \lambda_N, \mu) = \sigma_K^2 + 2\mu \left(\sum_{i=1}^{N} \lambda_i - 1 \right) \tag{2.32}$$

The conditions that should be satisfied by the optimal solution of the minimization problem are in this case:

---
[1] Or Equation (2.16b)

$$\frac{\partial L}{\partial \lambda_i} = -2 \cdot \frac{1}{V} \int_V \mathrm{Cov}(\underline{u}_i - \underline{u})\, d\underline{u} + 2 \sum_{j=1}^{N} \lambda_j \, \mathrm{Cov}(\underline{u}_i - \underline{u}_j) + 2\mu = 0$$

$$i = 1, \ldots, N \qquad (2.33a)$$

$$\frac{\partial L}{\partial \mu} = 2 \left(\sum_{i=1}^{N} \lambda_i - 1 \right) = 0 \qquad (2.33b)$$

which represent the Kriging system of (N+1) equations on the (N+1) variables $\lambda_1, \ldots, \lambda_N$ and μ. This system in matrix form is:

$$\begin{bmatrix} \mathrm{Cov}(\underline{0}) & \mathrm{Cov}(\underline{u}_1-\underline{u}_2) & \cdots & \mathrm{Cov}(\underline{u}_1-\underline{u}_N) & 1 \\ \mathrm{Cov}(\underline{u}_2-\underline{u}_1) & \mathrm{Cov}(\underline{0}) & \cdots & \mathrm{Cov}(\underline{u}_2-\underline{u}_N) & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ \mathrm{Cov}(\underline{u}_N-\underline{u}_1) & \mathrm{Cov}(\underline{u}_N-\underline{u}_2) & & \mathrm{Cov}\ (\underline{0}) & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_N \\ \mu \end{bmatrix} = \begin{bmatrix} \frac{1}{V}\int_V \mathrm{Cov}(\underline{u}_1-\underline{u})\,d\underline{u} \\ \frac{1}{V}\int_V \mathrm{Cov}(\underline{u}_2-\underline{u})\,d\underline{u} \\ \vdots \\ \frac{1}{V}\int_V \mathrm{Cov}(\underline{u}_N-\underline{u})\,d\underline{u} \\ 1 \end{bmatrix}$$

(2.34)

Notice that the left-hand matrix is the same previously found for point estimation, equation (2.25), and that the solution gives a unique minimum if that matrix is non-singular.

The optimal variance of estimation can be found in terms of the optimal weights $\lambda_1^*, \ldots, \lambda_N^*$ and Lagrange multipler μ^* in the same way that was done for the case of point estimation, i.e., taking equations (2.33a) multiplied by its respective λ_i^* and then adding over all i gives:

$$\sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_i^* \lambda_j^* \, Cov(\underline{u}_i - \underline{u}_j) = \sum_{i=1}^{N} \lambda_i^* \frac{1}{V} \int_V Cov(\underline{u}_i - \underline{u}) du - \mu^*$$

which replaced on the expression for σ_K^2, equation (2.31), gives as the optimal minimum variance:

$$\sigma_K^{2*} = \frac{1}{V^2} \int_V \int_V Cov(\underline{u} - \underline{v}) \, d\underline{u} \, d\underline{v} - \sum_{i=1}^{N} \lambda_i^* \frac{1}{V} \int_V Cov(\underline{u}_i - \underline{u}) d\underline{u} - \mu^* \tag{2.35}$$

The Kriging system for block estimation and the optimal variance of estimation can also be written in terms of the semivariogram function, using the equation (2.8), that links the semivariogram and covariance function under second-order stationarity assumptions. As in the previous case of point estimation, it is advantageous to use the semivariogram because the constant value of the mean need not be known in the estimation.

The equations (2.34) and (2.35) are transformed into:

$$\begin{bmatrix} 0 & -\gamma(\underline{u}_1-\underline{u}_2) & \cdots & -\gamma(\underline{u}_1-\underline{u}_N) & 1 \\ -\gamma(\underline{u}_2-\underline{u}_1) & 0 & \cdots & -\gamma(\underline{u}_2-\underline{u}_N) & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ -\gamma(\underline{u}_N-\underline{u}_1) & -\gamma(\underline{u}_N-\underline{u}_2) & \cdots & 0 & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_N \\ \mu \end{bmatrix} = \begin{bmatrix} -\frac{1}{V}\int_V \gamma(\underline{u}_1-\underline{u})d\underline{u} \\ -\frac{1}{V}\int_V \gamma(\underline{u}_2-\underline{u})d\underline{u} \\ \vdots \\ -\frac{1}{V}\int_V \gamma(\underline{u}_N-\underline{u})d\underline{u} \\ 1 \end{bmatrix} \tag{2.36}$$

and

$$\sigma_K^{2*} = -\frac{1}{V^2}\int_V\int_V \gamma(\underline{u} - \underline{v})d\underline{u}d\underline{v} + \sum_{i=1}^{N} \lambda_i^* \frac{1}{V}\int_V \gamma(\underline{u}_i-\underline{u})d\underline{u} - \mu^* \tag{2.37}$$

2.5 Kriging Under the Intrinsic Hypothesis and Constant Mean Assumptions

In this section the equations that determine the optimal weights of the Kriging estimator are given for the case in which the mean is constant and the variance of the first-order differences of the field is stationary.

The intrinsic hypothesis alone implies a linear drift, see Chua and Bras (1980), that gives:

$$E\{Z(\underline{u}_1) - Z(\underline{u}_2)\} = a_*(\underline{u}_1 - \underline{u}_2) \qquad (2.38)$$

This means the actual value of the constant a_* is needed in the estimation of the semivariogram function, see equation (2.7a). The additional constant mean condition, which is clearly equivalent as having a_* equal to zero, is assumed to avoid the bias introduced in the estimation of the semivariogram by the use of an estimated a_*.

The intrinsic hypothesis does not quarantee the existence of $\mathrm{Var}[Z(\underline{u})]$ and $\mathrm{Cov}(\underline{u}_i,\underline{u}_j)$. However, as it will be shown later, the variance of estimation can be calculated considering differences with respect to any reference value, $Z(\underline{u}^*)$, using the following equation (see Chua and Bras, 1980):

$$\begin{aligned}2\gamma(\underline{u}_i - \underline{u}_j) &= \mathrm{Var}[Z(\underline{u}_i) - Z(\underline{u}^*) - Z(\underline{u}_j) + Z(\underline{u}^*)] \\ &= \mathrm{Var}[Z(\underline{u}_i) - Z(\underline{u}^*)] + \mathrm{Var}[Z(\underline{u}_j) - Z(\underline{u}^*)] \\ &\quad - 2\,\mathrm{Cov}[\{Z(\underline{u}_i) - Z(\underline{u}^*)\}\{Z(\underline{u}_j) - Z(\underline{u}^*)\}]\end{aligned}$$

This gives:

$$\mathrm{Cov}(\underline{u}_i - \underline{u}^*, \underline{u}_j - \underline{u}^*) = \gamma(\underline{u}_i - \underline{u}^*) + \gamma(\underline{u}_j - \underline{u}^*) - \gamma(\underline{u}_i - \underline{u}_j) \tag{2.39}$$

Notice that the above equation is valid because all the terms on the right-hand side, being variances of differences, exist.

In the following subsections, the equations for point and block estimation will be written in terms of the existing stationary semi-variogram.

2.5.1 Point Kriging Under the Intrinsic Hypothesis and Constant Mean

In this case $\mathscr{L}(Z_0) = Z_0$, and then the unbiasedness condition gives, as before:

$$\sum_{i=1}^{N} \lambda_i = 1 \tag{2.40}$$

Defining $\lambda_0 = -1$, this above condition can be rewritten as:

$$\sum_{i=0}^{N} \lambda_i = 0 \tag{2.41}$$

and the variance of estimation, equation (2.14), can be expressed as:

$$\sigma_K^2 = \mathrm{Var}\left\{\sum_{i=0}^{N} \lambda_i Z_i\right\} \tag{2.42}$$

Subtracting $\sum_{i=0}^{N} \lambda_i Z(\underline{u}^*) = 0$, for an arbitrary location $\underline{u}^*$, gives:

$$\sigma_K^2 = \mathrm{Var}\left\{\sum_{i=0}^{N} \lambda_i [Z_i - Z(\underline{u}^*)]\right\} \tag{2.43}$$

which can be expanded in terms of existing covariance terms as:

$$\sigma_K^2 = \sum_{i=0}^{N} \sum_{j=0}^{N} \lambda_i \lambda_j \mathrm{Cov}(\underline{u}_i - \underline{u}^*, \underline{u}_j - \underline{u}^*) \tag{2.44}$$

Replacing the expression of the covariance in terms of the semi-variogram, equation (2.39), gives for the variance:

$$\sigma_K^2 = \sum_{i=0}^{N} \sum_{j=0}^{N} \lambda_i \lambda_j \gamma(\underline{u}_i - \underline{u}^*) + \sum_{i=0}^{N} \sum_{j=0}^{N} \lambda_i \lambda_j \gamma(\underline{u}_j - \underline{u}^*)$$

$$- \sum_{i=0}^{N} \sum_{j=0}^{N} \lambda_i \lambda_j \gamma(\underline{u}_i - \underline{u}_j) \tag{2.45}$$

which, using equation (2.41) on the first two terms, and expanding the last one, clearly gives the following equation independent of the field at the location, $\underline{u}^*$:

$$\sigma_K^2 = 2 \sum_{i=1}^{N} \lambda_i \, \gamma(\underline{u}_0 - \underline{u}_i) - \sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_i \lambda_j \, \gamma(\underline{u}_i - \underline{u}_j) \tag{2.46}$$

The above derivation shows that linear combinations of the form $\sum_{i=0}^{N} \lambda_i Z_i$ have finite variance if the weights λ_i sum to zero. Such linear combinations are termed authorized linear combinations and under the intrinsic hypothesis assumption are the only ones that have a finite variance, Matheron (1971). In following sections the concept of authorized linear combinations will be extended to the idea of Intrinsic Random Functions.

The optimal weights are found, as before, using the Lagrange multipliers technique. The objective function is:

$$L(\lambda_1, \ldots, \lambda_N, \mu) = \sigma_K^2 + 2\mu \left(\sum_{i=1}^{N} \lambda_i - 1 \right) \tag{2.47}$$

and the conditions that the weights and Lagrange multiplier should satisfy are:

$$\frac{\partial L}{\partial \lambda_i} = 2\gamma(\underline{u}_0 - \underline{u}_i) - 2 \sum_{j=1}^{N} \lambda_j \, \gamma(\underline{u}_i - \underline{u}_j) + 2\mu = 0 \tag{2.48a}$$

$$i=1, \ldots, N$$

$$\frac{\partial L}{\partial \mu} = 2 \left(\sum_{i=1}^{N} \lambda_i - 1 \right) = 0 \tag{2.48b}$$

These equations written in matrix form give the same Kriging system previously found under second-order stationary assumptions, or:

$$\begin{bmatrix} 0 & -\gamma(\underline{u}_1-\underline{u}_2) & \cdots & -\gamma(\underline{u}_1-\underline{u}_N) & 1 \\ -\gamma(\underline{u}_2-\underline{u}_1) & 0 & \cdots & -\gamma(\underline{u}_2-\underline{u}_N) & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ -\gamma(\underline{u}_N-\underline{u}_1) & -\gamma(\underline{u}_N-\underline{u}_2) & \cdots & 0 & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_N \\ \mu \end{bmatrix} \begin{bmatrix} -\gamma(\underline{u}_0-\underline{u}_1) \\ -\gamma(\underline{u}_0-\underline{u}_2) \\ \vdots \\ -\gamma(\underline{u}_0-\underline{u}_N) \\ 1 \end{bmatrix} \tag{2.49}$$

The optimal variance of estimation is as before:

$$\sigma_K^{2*} = \sum_{i=1}^{N} \lambda_i^* \gamma(\underline{u}_0 - \underline{u}_i) - \mu^* \tag{2.50}$$

where $\lambda_1^*, \ldots, \lambda_N^*$ are the optimal weights and μ^* is the optimal Lagrange multiplier.

2.5.2 Block Kriging Under the Intrinsic Hypothesis and Constant Mean

Using the constant mean assumption, the unbiasedness condition gives for block estimation:

$$\sum_{i=1}^{N} \lambda_i = \frac{1}{V} \int_V d\underline{u} = 1 \tag{2.51}$$

Define a measure, $\lambda(d\underline{u})$, over the region A, where the field is defined, as that which gives the operator:

$$\int_A \lambda(d\underline{u}) = \sum_{i=1}^{N} \lambda_i \, \delta_{\underline{u}_i} - \frac{1}{V} \int_V d\underline{u} \tag{2.52}$$

with $\delta_{\underline{u}_i}$ the indicator function of the field at location $\underline{u}_i$, i.e.:

$$\delta_{\underline{u}_i}(Z(\underline{u})) = \begin{cases} Z_i & \text{if } \underline{u} = \underline{u}_i \\ 0 & \text{otherwise} \end{cases} \tag{2.53}$$

This allows the unbiasedness condition to be written in terms of the extended combination:

$$\int_A Z(\underline{u}) \; \lambda(d\underline{u}) = \sum_{i=1}^{N} \lambda_i \, Z_i - \frac{1}{V} \int_V Z(\underline{u}) d\underline{u} \tag{2.54}$$

as:

$$E\left\{ \int_A Z(\underline{u}) \; \lambda(d\underline{u}) \right\} = m \int_A \lambda(d\underline{u}) = 0 \tag{2.55}$$

and the variance of estimation as:

$$\sigma_K^2 = \text{Var}\left\{ \int_A Z(\underline{u})\ \lambda(d\underline{u}) \right\} \tag{2.56}$$

Subtracting $\int_A Z(\underline{u}^*)\ \lambda(d\underline{u}) = 0$, for an arbitrary location $\underline{u}^*$, gives for the variance:

$$\sigma_K^2 = \text{Var}\left\{ \int_A (Z(\underline{u}) - Z(\underline{u}^*))\ \lambda(d\underline{u}) \right\} \tag{2.57}$$

which can be expanded in terms of the existing covariance function of the terms $Z(\underline{u}) - Z(\underline{u}^*)$ and $Z(\underline{v}) - Z(\underline{u}^*)$, as:

$$\sigma_K^2 = \int_A \int_A \text{Cov}(\underline{u} - \underline{u}^*, \underline{v} - \underline{u}^*)\ \lambda(d\underline{u})\ \lambda(d\underline{v}) \tag{2.58}$$

Replacing the expression that gives the above finite covariances in terms of the existing stationary semivariogram, equation (2.39), gives:

$$\sigma_K^2 = 2 \int_A \int_A \gamma(\underline{u} - \underline{u}^*)\ \lambda(d\underline{u})\ \lambda(d\underline{v})$$

$$- \int_A \int_A \gamma(\underline{u} - \underline{v})\ \lambda(d\underline{u})\ \lambda(d\underline{v}) \tag{2.59}$$

which readily reduces, as in the previously considered case of point estimation, to:

$$\sigma_K^2 = - \int_A \int_A \gamma(\underline{u} - \underline{v}) \; \lambda(d\underline{u}) \; \lambda(d\underline{v}) \tag{2.60}$$

or

$$\sigma_K^2 = - \sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_i \lambda_j \; \gamma(\underline{u}_i - \underline{u}_j) + 2 \sum_{i=1}^{N} \lambda_i \; \frac{1}{V} \int_V \gamma(\underline{u}_i - \underline{u}) d\underline{u}$$

$$- \frac{1}{V^2} \int_A \int_A \gamma(\underline{u} - \underline{v}) d\underline{u} \; d\underline{v} \tag{2.61}$$

As in the case of point estimation, the only extended combinations, equation (2.54), that give finite variance are the ones that satisfy $\int_A \lambda(d\underline{u}) = 0$, Matheron (1971). These combinations are called <u>extended authorized combinations</u>.

The minimization problem is solved as before using the Lagrange multipliers technique. The objective function is:

$$L(\lambda_1, \ldots, \lambda_N, \mu) = \sigma_K^2 + 2\mu \left(\sum_{i=1}^{N} \lambda_i - 1 \right) \tag{2.62}$$

and the conditions that should be satisfied by the weights and Lagrange multiplier are:

$$\frac{\partial L}{\partial \lambda_i} = -2 \sum_{j=1}^{N} \lambda_j \, \gamma(\underline{u}_i - \underline{u}_j) + 2 \cdot \frac{1}{V} \int_V \gamma(\underline{u}_i - \underline{u}) d\underline{u} + 2\mu = 0$$

$$i=1,\ldots, N \tag{2.63a}$$

$$\frac{\partial L}{\partial \mu} = 2\left(\sum_{i=1}^{N} \lambda_i = 1\right) = 0 \tag{2.63b}$$

These equations written in matrix form give, as in the case of point estimation, the same Kriging system previously found under second-order stationarity assumptions, or:

$$\begin{bmatrix} 0 & -\gamma(\underline{u}_1-\underline{u}_2) & \cdots & -\gamma(\underline{u}_1-\underline{u}_N) & 1 \\ -\gamma(\underline{u}_2-\underline{u}_1) & 0 & \cdots & -\gamma(\underline{u}_2-\underline{u}_N) & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ -\gamma(\underline{u}_N-\underline{u}_1) & -\gamma(\underline{u}_N-\underline{u}_2) & \cdots & 0 & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_N \\ \mu \end{bmatrix} = \begin{bmatrix} -\frac{1}{V} \int_V \gamma(\underline{u}_1-\underline{u}) d\underline{u} \\ -\frac{1}{V} \int_V \gamma(\underline{u}_2-\underline{u}) d\underline{u} \\ \vdots \\ -\frac{1}{V} \int_V \gamma(\underline{u}_N-\underline{u}) d\underline{u} \\ 1 \end{bmatrix} \tag{2.64}$$

The optimal minimum variance of estimation is also the same as found under second-order stationarity assumptions. It is:

$$\sigma_K^{2*} = -\frac{1}{V^2}\int_V\int_V \gamma(\underline{u} - \underline{v})d\underline{u}\; d\underline{v} + \sum_{i=1}^{N} \lambda_i^* \; \frac{1}{V}\int_V \gamma(\underline{u}_i - \underline{u})d\underline{u} - \mu^* \tag{2.65}$$

where $\lambda_1^*, \ldots, \lambda_N^*$ are the optimal weights and μ^* is the optimal Lagrange multiplier.

2.6 Kriging Under a Stationary Covariance and Known Form of the Drift: Universal Kriging

In this section, the equations that determine the optimal weights of the Kriging estimator are given for the case in which the covariance of the field is stationary and the mean is known to be of the form:

$$m(\underline{u}) = \sum_{\ell=1}^{kl} \delta_\ell \, f_\ell(\underline{u}) \tag{2.66}$$

where the functions f_ℓ are known and linearly independent, usually taken as monomials.

It will be shown that the equations do not depend explicitly on the values of the coefficients δ_ℓ. But certainly, the covariance function depends on those coefficients, which must be estimated prior to the estimation of the covariance.

2.6.1 Point Universal Kriging Under a Stationary Covariance

Under the stated conditions, the unbiasedness condition for point estimation at the point $\underline{u}_0$, see equation (2.13), reduces to:

$$\sum_{i=1}^{N} \lambda_i \sum_{\ell=1}^{kl} \delta_\ell \, f_\ell \, (\underline{u}_i) = \sum_{\ell=1}^{kl} \delta_\ell \, f_\ell \, (\underline{u}_0) \tag{2.67}$$

Interchanging the order of the summations on the left-hand side and using the linearly independent assumption on the functions f_ℓ gives the following kl conditions that are equivalent to the above equation:

$$\sum_{i=1}^{N} \lambda_i \; f_\ell \; (\underline{u}_i) \;=\; f_\ell \; (\underline{u}_0) \quad , \quad \ell=1,\ldots,kl \tag{2.68}$$

As in previous cases, the minimization problem that determines the optimal weights can be solved using the Lagrange multipliers technique.

The kl conditions are taken into the objective function through the use of kl Lagrange multipliers, as follows:

$$L(\lambda_1,\ldots,\lambda_N, \mu_1,\ldots,\mu_{kl}) = \sigma_K^2 + 2 \sum_{\ell=1}^{kl} \mu_\ell \left\{ \sum_{i=1}^{N} \lambda_i f_\ell (\underline{u}_i) - f_\ell (\underline{u}_0) \right\} \tag{2.69}$$

where σ_K^2 can be calculated in terms of the covariance function from equation (2.16b) as:

$$\sigma_K^2 = \mathrm{Cov}(\underline{0}) - 2 \sum_{i=1}^{N} \lambda_i \; \mathrm{Cov}(\underline{u}_0 - \underline{u}_i)$$

$$+ \sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_i \lambda_j \; \mathrm{Cov}(\underline{u}_i - \underline{u}_j) \tag{2.70}$$

The derivatives of the Lagrangian function, L, with respect to the unknowns set to zero give the necessary conditions for the unknowns, namely:

$$\frac{\partial L}{\partial \lambda_i} = -2\ \mathrm{Cov}(\underline{u}_0-\underline{u}_i) + 2\sum_{j=1}^{N} \lambda_j\ \mathrm{Cov}(\underline{u}_i-\underline{u}_j) + 2\sum_{\ell=1}^{k1} \mu_\ell\ f_\ell(\underline{u}_i) = 0$$

$$i=1,\ldots,N$$

(2.71a)

$$\frac{\partial L}{\partial \mu_\ell} = 2\left(\sum_{i=1}^{N} \lambda_i\ f_\ell(\underline{u}_i) - f_\ell(\underline{u}_0)\right) = 0\ ,\ \ell=1,\ldots,k1 \qquad (2.71b)$$

These equations lead to the Universal Kriging system of (N+ k1) equations on the (N+k1) unknowns $\lambda_1,\ldots,\lambda_N, \mu_1,\ldots,\mu_{k1}$:

$$\begin{bmatrix}
\mathrm{Cov}(\underline{0}) & \mathrm{Cov}(\underline{u}_1-\underline{u}_2) & \cdots & \mathrm{Cov}(\underline{u}_1-\underline{u}_N) & f_1(\underline{u}_1) & f_2(\underline{u}_1) & \cdots & f_{k1}(\underline{u}_1) \\
\mathrm{Cov}(\underline{u}_2-\underline{u}_1) & \mathrm{Cov}(\underline{0}) & \cdots & \mathrm{Cov}(\underline{u}_2-\underline{u}_N) & f_1(\underline{u}_2) & f_2(\underline{u}_2) & \cdots & f_{k1}(\underline{u}_2) \\
\vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\
\mathrm{Cov}(\underline{u}_N-\underline{u}_1) & \mathrm{Cov}(\underline{u}_N-\underline{u}_2) & \cdots & \mathrm{Cov}(\underline{0}) & f_1(\underline{u}_N) & f_2(\underline{u}_N) & \cdots & f_{k1}(\underline{u}_N) \\
f_1(\underline{u}_1) & f_1(\underline{u}_2) & \cdots & f_1(\underline{u}_N) & 0 & 0 & \cdots & 0 \\
f_2(\underline{u}_1) & f_2(\underline{u}_2) & \cdots & f_2(\underline{u}_N) & 0 & 0 & \cdots & 0 \\
\vdots & \vdots & & \vdots & \vdots & \vdots & & \vdots \\
f_{k1}(\underline{u}_1) & f_{k1}(\underline{u}_2) & \cdots & f_{k1}(\underline{u}_N) & 0 & 0 & \cdots & 0
\end{bmatrix}
\begin{bmatrix} \lambda_1 \\ \lambda_2 \\ \vdots \\ \lambda_N \\ \mu_1 \\ \mu_2 \\ \vdots \\ \mu_{k1} \end{bmatrix}
=
\begin{bmatrix} \mathrm{Cov}(\underline{u}_0-\underline{u}_1) \\ \mathrm{Cov}(\underline{u}_0-\underline{u}_2) \\ \vdots \\ \mathrm{Cov}(\underline{u}_0-\underline{u}_N) \\ f_1(\underline{u}_0) \\ f_2(\underline{u}_0) \\ \vdots \\ f_{k1}(\underline{u}_0) \end{bmatrix}$$

(2.72)

Multiplying equations (2.71a) by its respective λ_i^* and adding on i gives:

$$\sum_{i=1}^{N}\sum_{j=1}^{N} \lambda_i^* \lambda_j^* \operatorname{Cov}(\underline{u}_i-\underline{u}_j) = \sum_{i=1}^{N} \lambda_i^* \operatorname{Cov}(\underline{u}_0-\underline{u}_i) - \sum_{\ell=1}^{k1}\sum_{i=1}^{N} \lambda_i^* \mu_\ell^* \; f_\ell(\underline{u}_i) \tag{2.73}$$

Replacing this expression back on equation (2.70), and using the unbiasedness conditions, equations (2.68), gives the optimal minimum variance of estimation in terms of the optimal weights $\lambda_1^*,\ldots,\lambda_N^*$ and optimal Lagrange multipliers $\mu_1^*,\ldots,\mu_{k1}^*$, as

$$\sigma_K^{2*} = \operatorname{Cov}(\underline{0}) - \sum_{i=1}^{N} \lambda_i^* \operatorname{Cov}(\underline{u}_0-\underline{u}_i) - \sum_{\ell=1}^{k1} \mu_\ell^* \, f_\ell(\underline{u}_0) \tag{2.74}$$

Note that the Universal Kriging system and the optimal variance of estimation can be written in terms of the semivariogram function, using the relation that gives the semivariogram in terms of the covariance function under the assumption of stationarity on the covariance. The expression are completely analogous to equations (2.72) and (2.74) with the covariance replaced by the negative of the semivariogram.

2.6.2 Block Universal Kriging Under a Stationary Covariance

The unbiasedness condition for block estimation over an area V are for this case, see equation (2.13):

$$\sum_{i=1}^{N} \lambda_i \sum_{\ell=1}^{k1} \delta_\ell \ f_\ell(\underline{u}_i) = \frac{1}{V}\int_V \sum_{\ell=1}^{k1} \delta_\ell \ f_\ell(\underline{u})d\underline{u} \tag{2.75}$$

Interchanging the order of the summations on the left-hand side and integration and summation on the right-hand side, and using the linearly independent assumption on the functions f_ℓ gives the following k1 conditions:

$$\sum_{i=1}^{N} \lambda_i \ f_\ell(\underline{u}_i) = \frac{1}{V}\int_V f_\ell(\underline{u})d\underline{u} \ , \ \ell=1,\ldots,k1 \tag{2.76}$$

As in previous cases, the optimal weights are found using the Lagrange multipliers technique. The k1 conditions are taken into the objective function using Lagrange multipliers $\mu_1,\ldots,\mu_{k1}$, to get:

$$L(\lambda_1,\ldots,\lambda_N, \mu_1,\ldots,\mu_{k1}) = \sigma_K^2 + 2\sum_{\ell=1}^{k1} \mu_\ell\left(\sum_{i=1}^{N} \lambda_i f_\ell(\underline{u}_i) - \frac{1}{V}\int_V f_\ell(\underline{u})d\underline{u}\right) \tag{2.77}$$

where the variance of estimation in terms of the stationary covariance function is found, using equation (2.16b), as:

$$\sigma_K^2 = \frac{1}{V^2}\int_V\int_V \mathrm{Cov}(\underline{u} - \underline{v})d\underline{u}\, d\underline{v} - 2\sum_{i=1}^{N} \lambda_i \frac{1}{V}\int_V \mathrm{Cov}(\underline{u}_i-\underline{u})d\underline{u}$$

$$+ \sum_{i=1}^{N}\sum_{j=1}^{N} \lambda_i\lambda_j\, \mathrm{Cov}(\underline{u}_i-\underline{u}_j) \tag{2.78}$$

The necessary conditions of the minimization problem readily give the equations:

$$\frac{\partial L}{\partial \lambda_i} = -2\,\frac{1}{V}\int_V \mathrm{Cov}(\underline{u}_i-\underline{u})d\underline{u} + 2\sum_{j=1}^{N} \lambda_j\, \mathrm{Cov}(\underline{u}_i-\underline{u}_j) + 2\sum_{\ell=1}^{k1} \mu_\ell\, f_\ell(\underline{u}_i) = 0$$

$$i=1,\ldots,N \tag{2.79a}$$

$$\frac{\partial L}{\partial \mu_\ell} = 2\left(\sum_{i=1}^{N} \lambda_i\, f_\ell(\underline{u}_i) - \frac{1}{V}\int_V f_\ell(\underline{u})d\underline{u}\right) = 0\ ,\ \ell=1,\ldots,k1 \tag{2.79b}$$

which gives a Universal Kriging system with the left-hand side exactly as in equation (2.72), but with the right-hand side replaced by:

$$
\begin{bmatrix}
\frac{1}{V}\int_V \mathrm{Cov}(\underline{u}_i-\underline{u})\,d\underline{u} \\
\frac{1}{V}\int_V \mathrm{Cov}(\underline{u}_2-\underline{u})\,d\underline{u} \\
\vdots \\
\frac{1}{V}\int_V \mathrm{Cov}(\underline{u}_N-\underline{u})\,d\underline{u} \\
\frac{1}{V}\int_V f_1(\underline{u})\,d\underline{u} \\
\frac{1}{V}\int_V f_2(\underline{u})\,d\underline{u} \\
\vdots \\
\frac{1}{V}\int_V f_{k1}(\underline{u})\,d\underline{u}
\end{bmatrix}
\tag{2.80}
$$

The optimal variance of estimation, in terms of the optimal weights $\lambda_1^*, \ldots, \lambda_N^*$ and optimal Lagrange multipliers $\mu_1^*, \ldots, \mu_{k1}^*$, can be found using the procedure previously employed. It is:

$$
\sigma_K^{2*} = \frac{1}{V^2}\int_V\int_V \mathrm{Cov}(\underline{u} - \underline{v})\,d\underline{u}\,d\underline{v} - \sum_{i=1}^{N} \lambda_i^* \frac{1}{V}\int_V \mathrm{Cov}(\underline{u}_i-\underline{u})\,d\underline{u}
$$

$$
- \sum_{\ell=1}^{k1} \mu_\ell^* \frac{1}{V}\int_V f_\ell(\underline{u})\,d\underline{u} \tag{2.81}
$$

Once again, the Universal Kriging system and the optimal variance of estimation can be written in terms of the semivariogram function. The resulting expressions are the same with the covariance replaced by the negative of the semivariogram.

2.7 Intrinsic Random Functions of Order k

As was seen in the previous section, knowledge of the covariance or semivariogram function and the form of the drift are needed in order to calculate the Universal Kriging estimator. A practical contradiction appears and is that of estimating the covariance or semivariogram from the data points. Although in the Universal Kriging equations the exact coefficients of the drift do not appear, they are needed to calculate the covariance or semivariogram (except if $\bar{k}1 = 1$) function.

Even if the drift can be estimated without bias, Olea (1975a) has shown that the resulting estimated covariance or semivariogram function, after replacing the estimated drift, will be biased. In fact, if the drift is linear, the bias will be quadratic in distance $|h|$; if the drift is quadratic, the bias will be of fourth-order in $|h|$, etc.

The Intrinsic Random Functions of order k theory, introduced by Matheron (1973) and denoted IRF-k, provides a way of performing the Universal Kriging estimator without calculating the drift and the covariance function. This is done introducing the concept of generalized increments and replacing the covariance function by the generalized covariance function.

Instead of working with the random field $Z(\underline{u})$, the attention is focused on increments of $Z(\underline{u})$. To illustrate the general idea, assume the mean of the field to be a constant, say m. Then from equation (2.7a):

$$E\{(Z(\underline{u} + \underline{h}) - Z(\underline{u}))^2\} = \mathrm{Var}\{Z(\underline{u} + \underline{h}) - Z(\underline{u})\} = 2\,\gamma(\underline{h})$$

which means that the semivariogram can be inferred without bias[1] from the sample differences squared, $(Z(\underline{u}_i + \underline{h}) - Z(\underline{u}_i))^2$, without knowing the value of m. On the other hand, the estimated covariance function will be biased because an estimate m^* will be introduced.

In the above example, the field was replaced by its 0^{th} order increment (first-order difference as in the intrinsic hypothesis) and because the constant m was not needed, it is said that 0^{th} order increments filter out a constant mean, i.e., $E\{Z(\underline{u} + \underline{h}) - Z(\underline{u})\} = 0$.

A similar approach will be followed for more general forms of the drift, the idea being to estimate a certain function, called generalized covariance function, such that the drift is filtered out and then not needed in the estimation.

The formal definitions of the above concepts are as follows.

For a random field defined in the Euclidean space $\mathbb{R}^n$, a linear combination $\sum_{i=1}^{N} \lambda_i\, Z(\underline{u}_i)$, $\underline{u}_i=(u_{i1},\ldots, u_{in})$, is a generalized increment of order k if:

$$\sum_{i=1}^{N} \lambda_i\, u_{i1}^{j_1}\, u_{i2}^{j_2} \cdots u_{in}^{j_n} = 0 \qquad (2.82)$$

[1]Note that this is no longer true if the drift is not constant.

for all non-negative integers $j_1, j_2, \ldots, j_n$ such that:

$$j_1 + j_2 + \ldots + j_n \leq k$$

In particular in $\mathbb{R}^2$, denoting the coordinates $\underline{u}_i = (u_i, v_i)$, gives the following conditions:

$$k = 0: \quad \sum_{i=1}^{N} \lambda_i = 0 \tag{2.83}$$

$$k = 1: \quad \sum_{i=1}^{N} \lambda_i = 0$$

$$\sum_{i=1}^{N} \lambda_i u_i = 0 \quad ; \quad \sum_{i=1}^{N} \lambda_i v_i = 0 \tag{2.84}$$

$$k = 2: \quad \sum_{i=1}^{N} \lambda_i = 0 \quad ; \quad \sum_{i=1}^{N} \lambda_i u_i = 0$$

$$\sum_{i=1}^{N} \lambda_i v_i = 0 \quad ; \quad \sum_{i=1}^{N} \lambda_i u_i v_i = 0$$

$$\sum_{i=1}^{N} \lambda_i u_i^2 = 0 \quad ; \quad \sum_{i=1}^{N} \lambda_i v_i^2 = 0 \tag{2.85}$$

It is easily seen that an increment of order 0 filters out a constant drift, an increment of order 1 filters out a linear drift, an increment of order 2 filters out a quadratic drift, and so on. As an example, consider a generalized increment of order 1 and a linear drift $m(\underline{u}) = a + bu + cv$, then:

$$E\left\{\sum_{i=1}^{N} \lambda_i \; Z(\underline{u}_i)\right\} = \sum_{i=1}^{N} \lambda_i \; [a + bu_i + cv_i]$$

$$= a \sum_{i=1}^{N} \lambda_i + b \sum_{i=1}^{N} \lambda_i \; u_i + c \sum_{i=1}^{N} \lambda_i \; v_i$$

$$= 0$$

for any values of the coefficients a, b, and c. Note however, that if the drift is quadratic it is not filtered out by a first-order increment.

An <u>Intrinsic Random Function of order k</u> (IRF-k) is a random field whose generalized increments of order k are second-order stationary. This means that for all sets of weights λ_i that satisfy equations (2.82), $\sum_{i=1}^{N} \lambda_i \; Z(\underline{u}_i + \underline{h})$ has a mean and variance which do not depend on $\underline{h}$, and a covariance function depending only in the distance $\underline{h}$.

Suppose that the k^{th} increments of $Z(\underline{u})$ produce a second-order stationary field $Z_k(\underline{u})$. Since the coefficients of the drift could be anything, the IRF-k $Z(\underline{u}) + P_k(\underline{u})$, with $P_k(\underline{u})$ any polynomial of degree less than or equal to k, will also result in $Z_k(\underline{u})$ after increments of order k. This means, when working with IRF-k one is dealing with a class of fields and not with a unique field.

It can be shown (Matheron, 1973) that if $Z(\underline{u})$ is an IRF-k, there exists a function $K(\underline{h})$ such that for any generalized increment of order k:

$$\mathrm{Var}\left(\sum_{i=1}^{N} \lambda_i \, Z(\underline{u}_i)\right) = \sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_i \lambda_j \, K(\underline{u}_i - \underline{u}_j) \tag{2.86}$$

Because this expression has exactly the same form as if the function $K(\underline{h})$ were an ordinary covariance function, it is called a generalized covariance function, GC. It plays the role, but is not the covariance function of the IRF-k. Being $K(\underline{h})$ characteristic of a class of random fields, it is not unique but is defined up to an even polynomial of degree less than or equal to 2k, Matheron (1973).

Note that IRF-k generalize the previously defined intrinsic hypothesis: clearly generalized increments of order 0 are authorized linear combinations and the generalized covariance function for the IRF-0 reduces to minus the semivariogram function (see equations (2.42) through (2.46)).

In the next sections, it will be seen how to apply the IRF-k theory to find the Universal Kriging estimator for points and blocks, and how to deal with generalized covariance functions in practice.

2.8 Universal Kriging for Intrinsic Random Functions

In this section the Universal Kriging equations are rewritten in terms of the generalized covariance function, for the cases of point and block estimation.

2.8.1 Point Universal Kriging for Intrinsic Random Functions

The purpose is the estimation of the field at location $\underline{u}_0$, from available information at points $\underline{u}_1, \ldots, \underline{u}_N$. Recall that the estimator is given by the linear combination of the data points:

$$Z_{0_K}^* = \sum_{i=1}^{N} \lambda_i \, Z_i \tag{2.87}$$

such that the following two requirements are satisfied. First the unbiasedness conditions must hold, i.e.:

$$\sum_{j=1}^{N} \lambda_j \, f_\ell(\underline{u}_j) = f_\ell(\underline{u}_0) \;, \; \ell=1,\ldots,k1 \tag{2.88}$$

where the k1 functions f_ℓ determine the form of the drift. And second, the variance of estimation must be minimized over all possible combinations that satisfy equation (2.88).

If the functions f_ℓ are taken as monomials, for example in $\mathbb{R}^2$, $\underline{u}=(u,v)$, as:

$$
\begin{aligned}
f_1(\underline{u}) &= 1 \\
f_2(\underline{u}) &= u \\
f_3(\underline{u}) &= v \\
f_4(\underline{u}) &= uv \\
f_5(\underline{u}) &= u^2 \\
f_6(\underline{u}) &= v^2
\end{aligned}
\tag{2.89}
$$

and so on, and if λ_0 is defined as -1, then the unbiasedness conditions (2.88) take the form:

$$
\sum_{j=0}^{N} \lambda_i \, f_\ell(\underline{u}_j) \;=\; 0 \;,\; \ell=1,\ldots,k1 \tag{2.90}
$$

which say that the linear combination, $\sum_{j=0}^{N} \lambda_j \, Z(\underline{u}_j)$, is a generalized increment of a certain order k (see equations (2.83), (2.84), (2.85)) with k dependent on the number k1 of unbiasedness conditions considered, i.e., in $\mathbb{R}^2$ (see equations (2.89)) if k1=1 then k=1; if k1=3 then k=2; if k1=6 then k=3, and so on.

But then the variance of the above linear combination can be expanded in terms of an appropriate generalized covariance function $K(\underline{h})$, which gives for the variance of estimation:

$$
\sigma_K^2 = \mathrm{Var}\left\{ \sum_{j=0}^{N} \lambda_j \, Z(\underline{u}) \right\} \;=\; \sum_{i=0}^{N} \sum_{j=0}^{N} \lambda_i \lambda_j \, K(\underline{u}_i - \underline{u}_j) \tag{2.91}
$$

or the expansion in terms of unknown weights:

$$\sigma_K^2 = \sum_{i=1}^{N} \sum_{j=1}^{N} \lambda_i \lambda_j \, K(\underline{u}_i - \underline{u}_j) + K(\underline{0})$$

$$- 2 \sum_{i=1}^{N} \lambda_i \, K(\underline{u}_i - \underline{u}_0) \qquad (2.92)$$

Using Lagrange multipliers to impose the unbiasedness conditions, the optimal weights and multipliers that minimize σ_K^2 should satisfy:

$$\sum_{j=1}^{N} \lambda_j \, K(\underline{u}_i - \underline{u}_j) = K(\underline{u}_0 - \underline{u}_i) - \sum_{\ell=1}^{k1} \mu_\ell \, f_\ell(\underline{u}_i) \,,$$

$$i=1,\ldots,N \qquad (2.93a)$$

$$\sum_{j=1}^{N} \lambda_j \, f_\ell(\underline{u}_j) = f_\ell(\underline{u}_0) \,, \quad \ell=1,\ldots,k1 \qquad (2.93b)$$

Note that this has exactly the same form as the previously found Universal Kriging system, equation (2.72), with the covariance function replaced by the generalized covariance function.

The optimal variance of estimation, once the optimal weights $\lambda_1^*, \ldots, \lambda_N^*$ and optimal Lagrange multipliers $\mu_1^*, \ldots, \mu_{k1}^*$ have been found, is:

$$\sigma_K^{2*} = K(\underline{0}) - \sum_{i=1}^{N} \lambda_i^* \, K(\underline{u}_i - \underline{u}_0) - \sum_{\ell=1}^{k1} \mu_\ell^* \, f_\ell(\underline{u}_0) \qquad (2.94)$$

2.8.2 Block Universal Kriging for Intrinsic Random Functions

Define a measure, $\lambda(d\underline{u})$ as in section 2.5.2, such that the following operator is defined over the region A, the domain of the field:

$$\int_A \lambda(d\underline{u}) = \sum_{i=1}^{N} \lambda_i \, \delta_{\underline{u}_i} - \frac{1}{V}\int_V d\underline{u} \tag{2.95}$$

with $\delta_{\underline{u}_i}$ the indicator function at location $\underline{u}_i$, i.e.:

$$\delta_{\underline{u}_i}(Z(\underline{u})) = \begin{cases} Z(\underline{u}_i) & \text{if } \underline{u} = \underline{u}_i \\ 0 & \text{otherwise} \end{cases} \tag{2.96}$$

Then the unbiasedness conditions for block estimation, equations (2.76), can be written in terms of the above operator as:

$$\int_A f_\ell(\underline{u}) \, \lambda(d\underline{u}) = 0 \ , \ \ell=1,\ldots, k1 \tag{2.97}$$

which if the functions f_ℓ are monomials, says that the extended combination:

$$\int_A Z(\underline{u}) \, \lambda(d\underline{u}) = \sum_{i=1}^{N} \lambda_i \, Z_i - \frac{1}{V}\int_V Z(\underline{u}) d\underline{u} \tag{2.98}$$

is an extended generalized increment of order k (see equations (2.82)), with the order k, dependent on the number k1 of unbiasedness conditions considered.

The extension is indeed complete because the variance of an extended generalized increment enjoys an expansion in terms of the generalized covariance function (Matheron, 1973), which in particular gives for the variance of estimation:

$$\sigma_K^2 = \mathrm{Var}\left\{\int_A Z(\underline{u})\ \lambda(d\underline{u})\right\} = \int_A\int_A K(\underline{u} - \underline{v})\ \lambda(d\underline{u})\ \lambda(d\underline{v}) \tag{2.99}$$

or simply:

$$\sigma_K^2 = \sum_{i=1}^{N}\sum_{j=1}^{N} \lambda_i\lambda_j\ K(\underline{u}_i - \underline{u}_j) - 2\sum_{i=1}^{N} \lambda_i\ \frac{1}{V}\int_V K(\underline{u}_i - \underline{u})d\underline{u}$$

$$+ \frac{1}{V^2}\int_V\int_V K(\underline{u} - \underline{v})d\underline{u}\ d\underline{v} \tag{2.100}$$

Minimizing this previous equation subject to the unbiasedness conditions results in the system:

$$\sum_{j=1}^{N} \lambda_j\ K(\underline{u}_i - \underline{u}_j) = \frac{1}{V}\int_V K(\underline{u}_i - \underline{u})d\underline{u} - \sum_{\ell=1}^{k1} \mu_\ell\ f_\ell(\underline{u}_i)\ , \quad i=1,\ldots,\ N \tag{2.101a}$$

$$\sum_{j=1}^{N} \lambda_j\ f_\ell(\underline{u}_j) = \frac{1}{V}\int_V f_\ell(\underline{u})d\underline{u}\ ,\ \ell=1,\ldots,\ k1 \tag{2.101b}$$

that has exactly the same form as equations (2.79a) and (2.79b), with the covariance function replaced by the generalized covariance function.

The optimal minimum variance of estimation can be readily found in terms of the optimal weights $\lambda_1^*, \ldots, \lambda_N^*$ and optimal multipliers $\mu_1^*, \ldots, \mu_{k1}^*$ as:

$$\sigma_K^{2*} = \frac{1}{V^2} \int_V \int_V K(\underline{u} - \underline{v}) d\underline{u}\; d\underline{v} - \sum_{i=1}^{N} \lambda_i^* \frac{1}{V} \int_V K(\underline{u}_i - \underline{u}) d\underline{u}$$

$$- \sum_{\ell=1}^{k1} \mu_\ell^* \frac{1}{V} \int_V f_\ell(\underline{u}) d\underline{u} \qquad (2.102)$$

2.9 Generalized Covariance Functions in Practice

In the previous section it was seen how the Universal Kriging equations can be found in terms of the generalized covariance functions. This section deals with some generalized covariance models and their estimation.

2.9.1 Polynomial Models of Generalized Covariance Functions

With the class of generalized covariance functions, it is possible to model fields for which a covariance function do not exist. For example, models with $\gamma(h) = |h|^{\alpha}$, $0 < \alpha < 2$, are such that a valid covariance function cannot be found; nevertheless, their generalized covariance is minus the variogram. Optimal estimation is then possible.

Among the generalized covariance functions, the polynomial models are convenient in applications. For an IRF-k, the function (Matheron, 1973).

$$K(h) = \sum_{P=0}^{k} (-1)^{P+1} \frac{\alpha_{2P+1}}{(2P+1)!} \frac{\Gamma(\frac{n}{2})\, P!}{\sqrt{\pi}\,\Gamma(\frac{2P+n+1}{2})} |h|^{2P+1} \tag{2.103}$$

is a valid isotropic generalized covariance function in $\mathbb{R}^n$ provided the coefficients α_p satisfy a conditionally positive definite condition, i.e.:

$$\sum_{P=0}^{k} \alpha_{2P+1} X^{k-P} \geq 0 \qquad \forall \qquad X \geq 0 \tag{2.104}$$

Since in applications one assumes the field to be locally isotropic, the drift of interest is a slowly varying function, and then quadratic approximations of it are usually sufficient. This results in Table 2.1 of the common polynomial models, where the term $C\delta(h)$ stands for the nugget effect and $\delta(h)$ takes the value of 1 at $h = 0$ and 0 everywhere else.

Estimation of parameters in these polynomial models will be explained in the following subsection.

2.9.2 Estimation of Polynomial Generalized Covariance Functions

For IRF-k with $k \leq 2$ there are twenty-five possible polynomial generalized covariance model forms: three for IRF-0, i.e.: $K(h) = C\delta(h)$, $K(h) = \alpha_1|h|$, and $K(h) = C\delta(h) + \alpha_1|h|$; seven for IRF-1; and fifteen for IRF-2. This section explains a procedure, proposed by Delfiner (1976), and implemented in the computer package AKRIP by Kafritsas and Bras (1981), to select the best model for a given data set.

First to be discussed is how to find the optimal coefficients for any of the possible model forms. Then follows a procedure of selection among the different admissible models, i.e., those that satisfy the conditionally positive definite conditions, equation (2.104).

Drift	k	f_ℓ in $\mathbb{R}^2$	f_ℓ in $\mathbb{R}^3$	Models for GC
Constant	0	1	1	$K(h) = C\delta(h) + \alpha_1 \lvert h \rvert$
Linear	1	1, u, v	1, u, v, w	$K(h) = C\delta(h) + \alpha_1 \lvert h \rvert + \alpha_3 \lvert h \rvert^3$
Quadratic	2	1, u, v, uv, u^2, v^2	1, u, v, w, uv, uw, vw, u^2, v^2, w^2	$K(h) = C\delta(h) + \alpha_1 \lvert h \rvert + \alpha_3 \lvert h \rvert^3 + \alpha_5 \lvert h \rvert^5$

Constraints on the Coefficients

In $\mathbb{R}^2$: $\alpha_1 \leq 0$, $\alpha_5 \leq 0$

$$\alpha_3 \geq -\frac{10}{3}\sqrt{\alpha_1 \alpha_5}$$

In $\mathbb{R}^3$: $\alpha_1 \leq 0$, $\alpha_5 \leq 0$

$$\alpha_3 \geq -\sqrt{10}\,\sqrt{\alpha_1 \alpha_5}$$

Table 2.1

Polynomial Models of GC in $\mathbb{R}^2$ and $\mathbb{R}^3$

(after Delfiner, 1976)

2.9.2.1 Parameter Estimation for a Given Polynomial GC Model Form

Recall that if the data set comes from an IRF-k, with generalized covariance function K(h), any generalized increment of order k, denoted $Z(\lambda)$, satisfies:

$$\mathrm{Var}\left(Z(\lambda)\right) = \mathrm{Var}\left(\sum_{i=1}^{N} \lambda_i Z_i\right) = \sum_{i=1}^{N}\sum_{j=1}^{N} \lambda_i\lambda_j K(\underline{u}_i-\underline{u}_j) \tag{2.105}$$

To estimate parameters of K(h) the idea is to construct generalized increments of order k, and select the coefficients C and α_j's such that the difference between the left and right-hand sides of the above equation is minimized.

The procedure starts giving an initial value to the parameters. (Delfiner (1976) suggests the use of $C = \alpha_3 = \alpha_5 = 0$ and $\alpha_1 = -1$ in all the cases.) Then, if all the data points are estimated using the Universal Kriging system with the initial K(h) and appropriate k, employing $N_0 < N$ neighboring points[1], and if Z_i^* is the estimated value at point $\underline{u}_i$, i.e.:

$$Z_i^* = \sum_{j=1}^{N_0} \lambda_{ij} Z_j \ , \ i=1,\dots, N \tag{2.106}$$

then the linear combination of the data:

[1]By definition, point $\underline{u}_i$ is excluded when $Z(\underline{u}_i)$ is estimated.

$$Z(\lambda_i) = \sum_{j=1}^{N_0} \lambda_{ij} Z_j - Z_i \; , \; i=1,\ldots, N \tag{2.107}$$

is by construction (see equations (2.82) and (2.93)), a generalized increment of order k.

Replacing the form of K(h), equation (2.105) can be evaluated at the i^{th} constructed generalized increment as follows:

$$\begin{aligned} \operatorname{Var}\left(Z(\lambda_i)\right) &= E\left[Z(\lambda_i)^2\right] \\ &= C \sum_{\alpha=1}^{N_0} (\lambda_{i\alpha})^2 + \sum_{P=0}^{k} \alpha_{2P+1} \sum_{\alpha=1}^{N_0} \sum_{\beta=1}^{N_0} \lambda_{i\alpha}\lambda_{i\beta} |\underline{u}_\alpha - \underline{u}_\beta|^{2P+1} \end{aligned} \tag{2.108}$$

Defining $T_i^0 = \sum_{\alpha=1}^{N_0} (\lambda_{i\alpha})^2$ and

$$T_i^{2P+1} = \sum_{\alpha=1}^{N_0} \sum_{\beta=1}^{N_0} \lambda_{i\alpha}\lambda_{i\beta} |\underline{u}_\alpha - \underline{u}_\beta|^{2P+1} \; , \; P=0,1,\ldots, k$$

equation (2.108) reduces to:

$$E\left[Z(\lambda_i)^2\right] = C\, T_i^0 + \sum_{P=0}^{k} \alpha_{2P+1} T_i^{2P+1} \tag{2.109}$$

which is a regression equation of $Z(\lambda_i)^2$ on the unknown parameters. Then the parameters can be evaluated using the N constructed generalized increments solving the problem:

$$Q = \min_{C,\ \alpha_P's} \sum_{i=1}^{N} \left[Z(\lambda_i)^2 - C\ T_i^0 - \sum_{P=0}^{k} \alpha_{2P+1} T_i^{2P+1} \right]^2 \qquad (2.110)$$

The calculated regression coefficients will be, in general, different from the initial coefficients used to calculate the generalized increments. However, new generalized increments can be constructed using the new K(h), which in turn will produce other regression coefficients. The optimal K(h) will be found if the estimated coefficients by the regression match the coefficients of K(h) used to create the generalized increments.

2.9.2.2 Selection of the Best Model

To avoid the computation of all the twenty-five models of polynomial generalized covariance functions, the order k is first selected as follows (Delfiner, 1976):

i) Estimate all the data points from its neighbors, using the Universal Kriging system for the three models $K(h) = -|h|$, $k = 0$, 1 and 2.

ii) Compute, at each point, the absolute value of the differences between the actual and the estimated value given by each of the above models.

iii) Find the rank of the models at each point, assigning a value of 0 to the model with lowest absolute difference at that point, 1 to the following model and 2 to the model with the biggest absolute difference at that point.

iv) Find the rank of the above three models adding their point ranks over all the data points.

v) Select the order k according to the model which gives the lowest rank.

Once the order k is fixed, all the respective models are estimated using the regression procedure explained in the previous subsection. Note that for the optimal values of the coefficients, the variance of estimation at each point, σ_i^2, can be calculated from equation (2.94). Because $Z(\lambda_i)^2$ is also a measure of the variance at point i, the best model will be the one that has:

$$r = \frac{\sum_{i=1}^{N} Z(\lambda_i)^2}{\sum_{i=1}^{N} \sigma_i^2} \tag{2.111}$$

closest to 1.

Because r is a biased estimator of:

$$\rho = \frac{E\left[\sum_{i=1}^{N} Z(\lambda_i)^2\right]}{E\left[\sum_{i=1}^{N} \sigma_i^2\right]} \tag{2.112}$$

Delfiner suggests to split the data set in two subgroups, J_1 and J_2 of lengths N_1 and N_2, $N_1 + N_2 = N$, and use the <u>jackknife</u> estimator:

$$\hat{\rho} = 2r - \frac{N_1 r_1 + N_2 r_2}{N} \tag{2.113}$$

where:

$$r_1 = \frac{\sum_{j\varepsilon J_1} Z(\lambda_j)^2}{\sum_{j\varepsilon J_1} \sigma_j^2} \quad , \quad r_2 = \frac{\sum_{j\varepsilon J_2} Z(\lambda_j)^2}{\sum_{j\varepsilon J_2} \sigma_j^2} \tag{2.114}$$

Then, in summary the procedure is:

1. Rank the models $K(h) = -|h|$ for $k = 0$, 1, and 2 and select k according to the model which has the lowest rank.

2. Estimate the optimal models of the respective k and choose the feasible one that has jackknife estimator closest to 1.

2.10 Summary

In this chapter the Linear Kriging equations for point and block estimation of a random field have been presented under different underlying assumptions. It has been shown how the theory of intrinsic random functions generalize the intrinsic hypothesis, and how polynomial models of generalized covariance functions should be selected for fields with given forms of the drift.

Chapter 3

OPTIMAL SEPARABLE ESTIMATION: DISJUNCTIVE KRIGING

In this chapter the theoretical and practical aspects of the Disjunctive Kriging estimator are presented. The treatment follows the original work by Matheron (1976a), and the lucid book by Journel and Huijbregts (1978).

3.1 Characteristics of the Disjunctive Kriging Estimator

The Disjunctive Kriging estimator is a generalization of the previously presented Linear Kriging estimator. It was proposed by Matheron (1976a), as a practical substitute of the conditional expectation. The estimator is characterized by the following conditions:

1) Separability - The estimator, $\mathscr{L}(Z_0)^*_{DK}$ is a combination of measurable single variable functions, f_i, of the observed variables, Z_i:

$$\mathscr{L}(Z_0)^*_{DK} = \sum_{i=1}^{N} f_i(Z_i) \qquad (3.1)$$

where the functions f_i are found such that the next two conditions are satisfied.

2) Unbiasedness - The estimator, $\mathscr{L}(Z_0)^*_{DK}$, must be an unbiased estimator of the unknown $\mathscr{L}(Z_0)$, i.e.:

$$E\left[\mathscr{L}(Z_0)^*_{DK}\right] = E\left[\mathscr{L}(Z_0)\right] \tag{3.2}$$

which gives the following condition on the functions f_i:

$$\sum_{i=1}^{N} E\left[f_i(Z_i)\right] = E\left[\mathscr{L}(Z_0)\right] \tag{3.3}$$

3) Minimum Variance - The estimator, $\mathscr{L}(Z_0)^*_{DK}$, must have minimum variance of estimation, σ^2_{DK}, over all the unbiased estimators of the form given by equation (3.1).

$$\sigma^2_{DK} = \mathrm{Var}\{\mathscr{L}(Z_0) - \mathscr{L}(Z_0)^*_{DK}\} \tag{3.4}$$

Using the unbiasedness condition and expanding gives the variance to be minimized as:

$$\begin{aligned}\sigma^2_{DK} &= E\left\{\left[\mathscr{L}(Z_0) - \mathscr{L}(Z_0)^*_{DK}\right]^2\right\} \\ &= E\left[\mathscr{L}(Z_0)^2\right] - 2\sum_{i=1}^{N} E\left[\mathscr{L}(Z_0)f_i(Z_i)\right] \\ &\quad + \sum_{i=1}^{N}\sum_{j=1}^{N} E\left[f_i(Z_i)f_j(Z_j)\right]\end{aligned} \tag{3.5}$$

Notice that σ^2_{DK} can be calculated if the bivariate distributions of the pairs (Z_i, Z_j) and $(Z_i, \mathscr{L}(Z_0))$ are known. This is considerably more information than the required semivariogram or generalized covariance function used in the linear Kriging estimator.

In fact, Linear Kriging is a particular case of Disjunctive Kriging. in which $f_i(Z_i) = \lambda_i Z_i$. Due to the added information, Disjunctive Kriging has a theoretical increase in accuracy, i.e.:

$$\sigma^2_{DK} \leq \sigma^2_K \tag{3.6}$$

In summary the Disjunctive Kriging estimator is found solving the minimization problem:

$$\min_{f_1,\ldots,f_n} \sigma^2_{DK} = E[\mathscr{L}(Z_0)^2] - 2 \sum_{i=1}^{N} E[\mathscr{L}(Z_0) f_i(Z_i)]$$

$$+ \sum_{i=1}^{N} \sum_{j=1}^{N} E[f_i(Z_i) f_j(Z_j)] \tag{3.7}$$

$$\text{s.t.} \quad \sum_{i=1}^{N} E[f_i(Z_i)] = E[\mathscr{L}(Z_0)]$$

This problem can be solved if further assumptions are made on the structure of the random field. In the next section a model will be given that will allow an approximated computation of the optimal functions.

3.2 The Hermitian Model

It is assumed that the random field under consideration, $Z(\underline{u})$, can be obtained as a transformation of a second-order stationary field, $Y(\underline{u})$, which has univariate standard Gaussian distribution.

$$Z(\underline{u}) = \phi(Y(\underline{u})) \tag{3.8}$$

The function ϕ is called the anamorphosis and has to be estimated in practice. Section 3.5 explains how to do this task.

Consider the Disjunctive Kriging estimation of the unknown $\mathscr{L}(Z_0)$ in terms of Gaussian variables, $Y_i = \phi^{-1}(Z_i)$. As was seen in the previous section, the estimator is:

$$\mathscr{L}(Z_0)^*_{DK} = \sum_{i=1}^{N} f_i(Y_i) \tag{3.9}$$

with the functions f_i, i=1,...N, being the solution of the problem:

$$\min_{f_1,\dots,f_N} \sigma^2_{DK} = E[\mathscr{L}(Z_0)^2] - 2\sum_{i=1}^{N} E[\mathscr{L}(Z_0) f_i(Y_i)]$$

$$+ \sum_{i=1}^{N}\sum_{j=1}^{N} E[f_i(Y_i) f_j(Y_j)] \tag{3.10}$$

$$\text{s.t.} \quad \sum_{i=1}^{N} E[f_i(Y_i)] = E[\mathscr{L}(Z_0)]$$

Assuming that the bivariate distributions of the pairs $(Y(\underline{u}),Y(\underline{v}))$ are also Gaussian, the above problem can be rewritten in a way that will permit the calculation of the optimal functions f_i, $i=1,\ldots,N$.

Let G(du) be the distribution of a standard Gaussian variable, i.e.:

$$G(du) = \frac{1}{\sqrt{2\pi}} e^{-\frac{u^2}{2}} du \tag{3.11}$$

Assume the functions f_i, $i=1,\ldots,N$ satisfy the condition:

$$\int_{-\infty}^{\infty} f_i^2(u)\, G(du) < \infty \quad , \; i=1,\ldots,N$$

Then they can be expanded in a series of Hermite polynomials as:

$$f_i(u) = \sum_{k=0}^{\infty} f_{i,k}\, \eta_k(u) \quad , \; i=1,\ldots,N \tag{3.12}$$

where η_k is the k^{th} standardized Hermite polynomial defined by:

$$\eta_k(u) = \frac{1}{\sqrt{k!}} e^{\frac{u^2}{2}} \frac{d^k}{du^k} e^{-\frac{u^2}{2}} \tag{3.13}$$

and the series coefficients, $f_{i,k}$, are given by:

$$f_{i,k} = \int_{-\infty}^{\infty} f_i(u)\, \eta_k(u)\, G(du) \quad , \quad \begin{array}{l} i=1,\ldots,N \\ k=0,1,\ldots \end{array} \tag{3.14}$$

Similarly, assume the anamorphosis function satisfies:

$$\int_{-\infty}^{\infty} \phi^2(u)\; G(du) < \infty$$

then it can be expanded as:

$$\phi(u) = \sum_{k=0}^{\infty} \phi_k\, \eta_k(u) \tag{3.15}$$

with the coefficients ϕ_k given by:

$$\phi_k = \int_{-\infty}^{\infty} \phi(u)\; \eta_k(u)\; G(du) \quad , \quad k=0,1,\ldots \tag{3.16}$$

Replacing the Hermite expansions into the defining expression of the Disjunctive Kriging estimator, equation (3.9), gives the following equivalent estimator:

$$\mathscr{L}(Z_0)^*_{DK} = \sum_{i=1}^{N} \sum_{k=0}^{\infty} f_{i,k}\; \eta_k(Y_i) \tag{3.17}$$

Similarly, using the linearity of the function $\mathscr{L}$ and of the expectation operator, the minimization problem can be rewritten in terms of the coefficients $f_{i,k}$ as:

$$\min_{\substack{f_{i,k} \\ i=1,..,N \\ k=0,1,..}} \sigma^2_{DK} = \sum_{k=0}^{\infty} \sum_{k'=0}^{\infty} \phi_k \phi_{k'} E\{\mathscr{L}(\eta_k(Y_0)) \cdot \mathscr{L}(\eta_{k'}(Y_0))\}$$

$$+ \sum_{i=1}^{N} \sum_{k=0}^{\infty} \sum_{k'=0}^{\infty} \phi_k f_{i,k'} E\{\mathscr{L}(\eta_k(Y_0))\eta_{k'}(Y_i)\}$$

$$+ \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{k=0}^{\infty} \sum_{k'=0}^{\infty} f_{i,k} f_{j,k'} E\{\eta_k(Y_i)\eta_{k'}(Y_j)\} \tag{3.18}$$

$$\text{s.t.} \quad \sum_{i=1}^{N} \sum_{k=0}^{\infty} f_{i,k} E\{\eta_k(Y_i)\} = \sum_{k=0}^{\infty} \phi_k E\{\mathscr{L}(\eta_k(Y_0))\}$$

The expectations above can be explicitly evaluated using the factor representation of the bivariate Gaussian distribution, G(du,dv), Matheron (1976a):

$$G(du,dv) = \psi(u,v)\ G(du)\ G(dv) \tag{3.19}$$

with G(du) the univariate standard Gaussian distribution and the joint density $\psi(u,v)$ given by:

$$\psi(u,v) = \sum_{k=0}^{\infty} \rho^k_{uv}\ \eta_k(Y(\underline{u}))\ \eta_k(Y(\underline{v})) \tag{3.20}$$

where ρ_{uv}^{k} is the correlation between the pair $(Y(\underline{u}),Y(\underline{v}))$ raised to the k^{th} power, i.e.:

$$\rho_{uv}^{k} = E[Y(\underline{u})\ Y(\underline{v})]^{k} \tag{3.21}$$

Using the orthonormality properties of the standardized Hermite polynomials:

$$\int_{-\infty}^{\infty} \eta_k(u)\ \eta_{k'}(u)\ G(du) = \begin{cases} 1 \text{ if } k=k' \\ 0 \text{ otherwise} \end{cases} \tag{3.22}$$

the expectations in σ_{DK}^{2} are found as follows:

$$E[\eta_k(Y_i)\eta_{k'}(Y_j)] = \int_{\infty}^{\infty}\int_{-\infty}^{\infty} \eta_k(Y_i)\eta_{k'}(Y_j)\ \psi(i,j)\ G(\mathbf{du}_i)\ G(du_j)$$

$$= \int_{\infty}^{\infty}\int_{-\infty}^{\infty} \eta_k(Y_i)\eta_{k'}(Y_j) \sum_{\ell=0}^{\infty} \rho_{ij}^{\ell}\ \eta_{\ell}(Y_i)\eta_{\ell}(Y_j)\quad G(du_i)\ G(du_j)$$

$$= \begin{cases} \rho_{ij}^{k} & \text{if } k=k' \\ 0 & \text{otherwise} \end{cases} \tag{3.23}$$

Replacing this back on equation (3.18), gives the estimation variance of the estimator as:

$$\sigma_{DK}^2 = \sum_{k=0}^{\infty} \phi_k^2 E\{\mathscr{L}(\eta_k(Y_0))^2\} - 2 \sum_{i=1}^{N} \sum_{k=0}^{\infty} \phi_k \, f_{i,k} E\{\mathscr{L}(\eta_k(Y_0))\eta_k(Y_i)\}$$

$$+ \sum_{i=1}^{N} \sum_{j=1}^{N} \sum_{k=0}^{\infty} f_{i,k} \, f_{j,k} \, \rho_{ij}^k \tag{3.24}$$

The unbiasedness conditions written in terms of the coefficients $f_{i,k}$ and ϕ_k give the simple condition:

$$\sum_{i=1}^{N} f_{i,0} = \phi_0 \tag{3.25}$$

The above results from $\eta_0(u) = 1$, $\mathscr{L}(c) = c$ for any constant c for either point and block estimation, and the fact that equations (3.14) and (3.16) yield:

$$E\left[\mathscr{L}(Z_0)^*_{DK}\right] = \sum_{i=1}^{N} \int_{-\infty}^{\infty} f_i(Y_i)\eta_0(Y_i) \; G(du_i) = \sum_{i=1}^{N} f_{i,0}$$

and

$$E[\mathscr{L}(Z_0)] = \mathscr{L}\{ \int_{-\infty}^{\infty} \phi(Y_0)\eta_0(Y_0) \; G(du_0)\} = \mathscr{L}(\phi_0) = \phi_0$$

Substituting the unbiasedness condition into the expression for σ_{DK}^2, equation (3.24), cancels all the terms with k=0. The weights, $f_{i,k}$, i=1,..,N, k=1,2,..., are then found solving the unconstrained problem:

$$\min_{\substack{f_{i,k}\\ i=1,\dots,N\\ k=1,2,\dots}} \sigma^2_{DK} = \sum_{k=1}^{\infty} \phi_k^2 E\{\mathscr{L}(\eta_k(Y_0))^2\} - 2\sum_{i=1}^{N}\sum_{k=1}^{\infty} \phi_k f_{i,k} E\{\mathscr{L}(\eta_k(Y_0))\eta_k(Y_i)\}$$

$$+ \sum_{i=1}^{N}\sum_{j=1}^{N}\sum_{k=1}^{\infty} f_{i,k} f_{j,k}\, \rho_{ij}^{k} \tag{3.26}$$

The solution is found setting the derivatives of σ^2_{DK} with respect to the unknowns $f_{i,k}$ equal to zero:

$$\frac{\partial \sigma^2_{DK}}{\partial f_{i,k}} = -2\,\phi_k E\{\mathscr{L}(\eta_k(Y_0))\eta_k(Y_i)\} + 2\sum_{j=1}^{N} f_{j,k}\,\rho_{ij}^{k} = 0, \qquad \begin{array}{l} i=1,\dots,N \\ k=1,2,\dots \end{array} \tag{3.27}$$

These equations written in matrix form give the infinite set of Disjunctive Kriging systems of N equations with N unknowns,

$$\begin{bmatrix} \rho_{11}^{k} & \rho_{12}^{k} & \cdots & \rho_{1N}^{k} \\ \rho_{21}^{k} & \rho_{22}^{k} & \cdots & \rho_{2N}^{k} \\ \vdots & \vdots & & \vdots \\ \rho_{N1}^{k} & \rho_{N2}^{k} & \cdots & \rho_{NN}^{k} \end{bmatrix} \begin{bmatrix} f_{1,k} \\ f_{2,k} \\ \vdots \\ f_{N,k} \end{bmatrix} = \begin{bmatrix} \phi_k E\{\mathscr{L}(\eta_k(Y_0))\eta_k(Y_1)\} \\ \phi_k E\{\mathscr{L}(\eta_k(Y_0))\eta_k(Y_2)\} \\ \vdots \\ \phi_k E\{\mathscr{L}(\eta_k(Y_0))\eta_k(Y_N)\} \end{bmatrix} \tag{3.28}$$

for every integer $k \geq 1$.

Substituting the optimality conditions, equation (3.27), into the expression for σ^2_{DK}, equation (3.26), gives the optimal variance of estimation as:

$$\sigma^{2*}_{DK} = \sum_{k=1}^{\infty} \phi_k^2 \, E\{\mathscr{L}(\eta_k(Y_0))^2\} - \sum_{i=1}^{N} \sum_{k=1}^{\infty} \phi_k f^*_{i,k} \, E\{\mathscr{L}(\eta_k(Y_0))\eta_k(Y_i)\} \qquad (3.29)$$

where $f^*_{i,k}$ represent the optimal values of the coefficients.

Notice that the conditional distribution of $Y(\underline{u})$ given $Y(\underline{v})$ can be found from the factor representation of the bivariate Gaussian distribution as:

$$G_{u|v}(du) = \psi(u,v) \; G(du) \qquad (3.30)$$

This allows the calculation of the conditional expectation of the Disjunctive Kriging estimator given a Gaussian variable as:

$$E\left[\sum_{i=1}^{N} f_i(Y_i) \,|\, Y_j\right] = \int_{-\infty}^{\infty} \sum_{i=1}^{N} \sum_{k=0}^{\infty} f_{i,k} \; \eta_k(Y_i) \; \psi(u_i,u_j) \; G(du_i)$$

which, replacing the expansion for the density function and using the orthonormality properties of the standardized Hermite polynomials, gives:

$$E\left[\sum_{i=1}^{N} f_i(Y_i) \,|\, Y_j\right] = \sum_{k=0}^{\infty} \sum_{i=1}^{N} f_{i,k} \; \rho^k_{ij} \; \eta_k(Y_j) \qquad (3.31)$$

Replacing the unbiasedness and optimality conditions, equations (3.25) and (3.27), give the above conditional expectation in terms of the coefficients of the anamorphosis expansion as:

$$E\left[\sum_{i=1}^{N} f_i(Y_i)\,|\,Y_j\right] = \sum_{k=0}^{\infty} \phi_k E\{\mathscr{L}(\eta_k(Y_0))\eta_k(Y_i)\}\ \eta_k(Y_j)$$

The term on the right-hand side is easily seen to be exactly, in the same way as above, the conditional expectation of the unknown $\mathscr{L}(\phi(Y_0))$ given the Gaussian variable Y_i. It then follows that the following relationship is valid under the Hermitian model assumption:

$$E\left[\mathscr{L}(Z_0)^{*}_{DK}\,|\,Y_i\right] = E\left[\mathscr{L}(Z_0)\,|\,Y_i\right] \quad , \quad i=1,\ldots,N \tag{3.32}$$

In fact, it is true that these conditions are equivalent to the requirements of unbiasedness and minimum variance imposed on the Disjunctive Kriging estimator (Matheron, 1976a). This means that the functions f_i are characterized by being those such that the estimator preserves the conditional expectations of the unknown given the Gaussian variables.

In the following sections, the generic Disjunctive Kriging system and optimal variance of estimation, equations (3.28) and (3.29), will be specialized for the cases of point and block estimation.

3.3 Point Disjunctive Kriging with the Hermitian Model

For point estimation, the unknown is simply $\mathscr{L}(Z_0) = Z_0$. which gives the estimator:

$$Z_{0\,DK}^{*} = \sum_{i=1}^{N} \sum_{k=0}^{\infty} f_{i,k}\, \eta_k(Y_i) \tag{3.33}$$

with the coefficients, $f_{i,k}$, satisfying the unbiasedness condition:

$$\sum_{i=1}^{N} f_{i,0} = \phi_0 \tag{3.34}$$

and the optimality conditions:

$$\sum_{j=1}^{N} f_{j,k}\, \rho_{ij}^{k} = \phi_k\, \rho_{0i}^{k}\ , \quad \begin{array}{l} i=1,\dots,N \\ k=1,2,\dots \end{array} \tag{3.35}$$

where $\rho_{ij} = E\{Y_i\ Y_j\}$.

Because the mean of the univariate distributions of the Gaussian variables is zero, the optimality conditions can be written in matrix form in terms of the stationary covariance function of the Gaussian field, as:

$$
\begin{bmatrix}
Cov_Y(\underline{0})^k & Cov_Y(\underline{u}_1-\underline{u}_2)^k & \cdots & Cov_Y(\underline{u}_1-\underline{u}_N)^k \\
Cov_Y(\underline{u}_2-\underline{u}_1)^k & Cov_Y(\underline{0})^k & \cdots & Cov_Y(\underline{u}_2-\underline{u}_N)^k \\
\vdots & \vdots & & \vdots \\
Cov_Y(\underline{u}_N-\underline{u}_1)^k & Cov_Y(\underline{u}_N-\underline{u}_2)^k & \cdots & Cov_Y(\underline{0})^k
\end{bmatrix}
\begin{bmatrix} f_{1,k} \\ f_{2,k} \\ \vdots \\ f_{N,k} \end{bmatrix}
=
\begin{bmatrix} \phi_k Cov_Y(\underline{u}_0-\underline{u}_1)^k \\ \phi_k Cov_Y(\underline{u}_0-\underline{u}_2)^k \\ \vdots \\ \phi_k Cov_Y(\underline{u}_0-\underline{u}_N)^k \end{bmatrix}
$$

$$k=1,2,\ldots \qquad (3.36)$$

In practice only a finite number of systems of equations are solved. As will be seen later, the total number to be used depends on the accuracy of the Hermitian fit of the anamorphosis function.

The optimal variance of estimation is found from equation (3.29) using the fact that Y_0 is standard Gaussian. It is:

$$\sigma_{DK}^2 = \sum_{k=1}^{\infty} \phi_k^2 - \sum_{i=1}^{N} \sum_{k=1}^{\infty} \phi_k \, f_{i,k} \, Cov_Y(\underline{u}_0-\underline{u}_i)^k \qquad (3.37)$$

Note that if $\underline{u}_0$ coincides with some observation at $\underline{u}_j$, the estimated value will be the observation, with estimation variance equal to zero, because the solution of the system of equations clearly reduces in that case to:

$$f_{i,k} = \begin{cases} \phi_k & i=j \\ 0 & \text{otherwise} \end{cases}$$

3.4 Block Disjunctive Kriging with the Hermitian Model

For the case of block estimation, the unknown is $\frac{1}{V}\int_V Z(\underline{u})d\underline{u}$, which gives the estimator:

$$\mathscr{L}(Z_0)^*_{DK} = \sum_{i=1}^{N}\sum_{k=0}^{\infty} f_{i,k}\,\eta_k(Y_i) \tag{3.38}$$

with the coefficients, $f_{i,k}$, satisfying the unbiasedness condition:

$$\sum_{i=1}^{N} f_{i,0} = \phi_0 \tag{3.39}$$

and the optimality conditions:

$$\sum_{j=1}^{N} f_{k,k}\,\rho_{ij}^{k} = \phi_k \frac{1}{V}\int_V \rho_{\underline{u}i}^{k}\ d\underline{u} \qquad \begin{array}{l} i=1,\ldots,N \\ k=1,2,\ldots \end{array} \tag{3.40}$$

where $\rho_{\underline{u}i}^{k} = E\{Y(\underline{u})Y_i\}^k$.

Using the standard Gaussian assumption on the variables Y_i's, the optimality conditions can be written in matrix form in terms of the stationary covariance function of the Gaussian variables as:

$$
\begin{bmatrix}
\mathrm{Cov}_Y(\underline{0})^k & \mathrm{Cov}_Y(\underline{u}_1-\underline{u}_2)^k & \cdots & \mathrm{Cov}_Y(\underline{u}_1-\underline{u}_2)^k \\
\mathrm{Cov}_Y(\underline{u}_2-\underline{u}_1)^k & \mathrm{Cov}_Y(\underline{0})^k & \cdots & \mathrm{Cov}_Y(\underline{u}_2-\underline{u}_N)^k \\
\vdots & \vdots & & \vdots \\
\mathrm{Cov}_Y(\underline{u}_N-\underline{u}_1)^k & \mathrm{Cov}_Y(\underline{u}_N-\underline{u}_2)^k & \cdots & \mathrm{Cov}_Y(\underline{0})
\end{bmatrix}
\begin{bmatrix}
f_{1,k} \\ f_{2,k} \\ \vdots \\ f_{N,k}
\end{bmatrix}
=
\begin{bmatrix}
\phi_k \frac{1}{V}\int_V \mathrm{Cov}_Y(\underline{u}-\underline{u}_1)^k d\underline{u} \\
\phi_k \frac{1}{V}\int_V \mathrm{Cov}_Y(\underline{u}-\underline{u}_2)^k d\underline{u} \\
\vdots \\
\phi_k \frac{1}{V}\int_V \mathrm{Cov}_Y(\underline{u}-\underline{u}_N)^k d\underline{u}
\end{bmatrix}
$$

$$k \geq 1 \qquad (3.41)$$

The optimal variance of estimation can be written in terms of the optimal coefficients $f^*_{i,k}$, using the standard Gaussian assumption on the univariate distribution of the field:

$$
\sigma^{2*}_{DK} = \sum_{k=1}^{\infty} \phi_k^2 \frac{1}{V^2} \int_V \int_V \mathrm{Cov}_Y(\underline{u}-\underline{v})^k d\underline{u} d\underline{v}
$$

$$
- \sum_{k=1}^{\infty} \sum_{j=1}^{N} \phi_k f^*_{j,k} \frac{1}{V} \int_V \mathrm{Cov}_Y(\underline{u}-\underline{u}_j)^k d\underline{u} \qquad (3.42)
$$

3.5 Disjunctive Kriging in Practice: Determination of the Anamorphosis

The Hermitian model, described in this chapter, assumed the field $Z(\underline{u})$ came from a stationary field $Y(\underline{u})$ with univariate standard Gaussian distribution, through the anamorphosis function, i.e., $Z(\underline{u}) = \phi(Y(\underline{u}))$. Because there is a one-to-one relationship between the univariate distribution of the field $Z(\underline{u})$ and the univariate standard Gaussian distribution, the anamorphosis can be found as follows.

1. Calculate the univariate distribution function of the field, and then find the correspondent Gaussian values, solving the equation:

$$\text{Prob}\{Z \leq Z_i\} = \text{Prob}\{Y \leq Y_i\} \tag{3.43}$$

   where the observation values have been ordered, i.e., $Z_1 \leq Z_2 \leq \cdots \leq Z_N$. In Figure 3.1 this step is explained graphically.

2. Take the pairs (Y_i, Z_i) and plot the anamorphosis, interpolating between known points as in Figure 3.2.

The quality of interpolation can be corroborated from the first term of the Hermitian expansion of the anamorphosis since:

$$E(Z) = E(\phi(Y)) = \int_{-\infty}^{\infty} \phi(Y)\, \eta_0(Y)\, G(dY) = \phi_0 \tag{3.44}$$

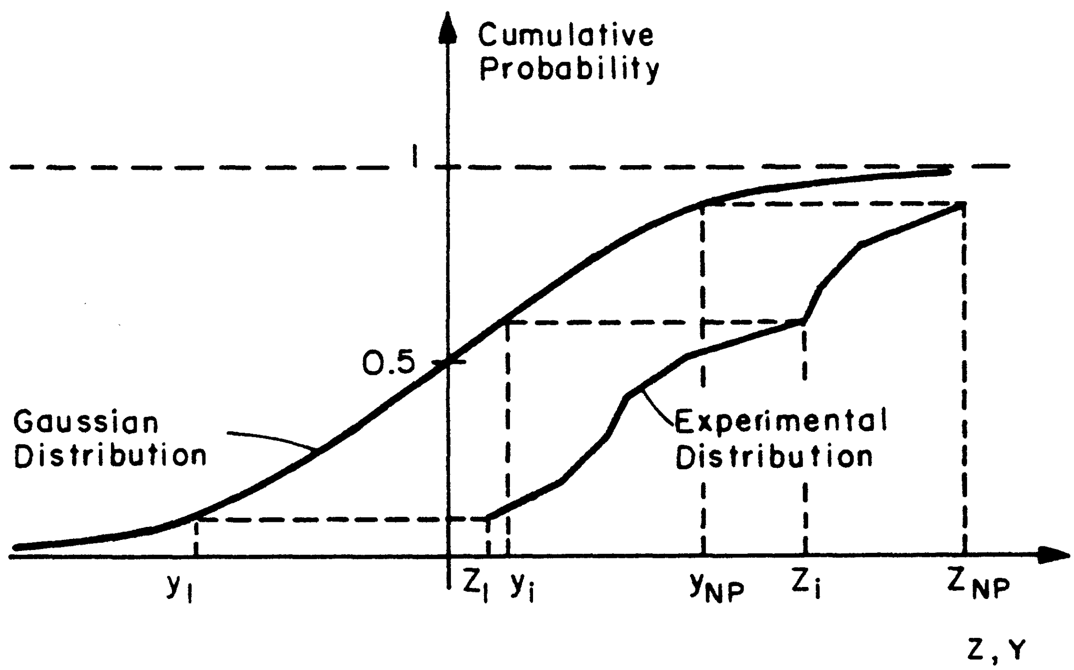


Figure 3.1

Graphical Determination of the Gaussian Variables

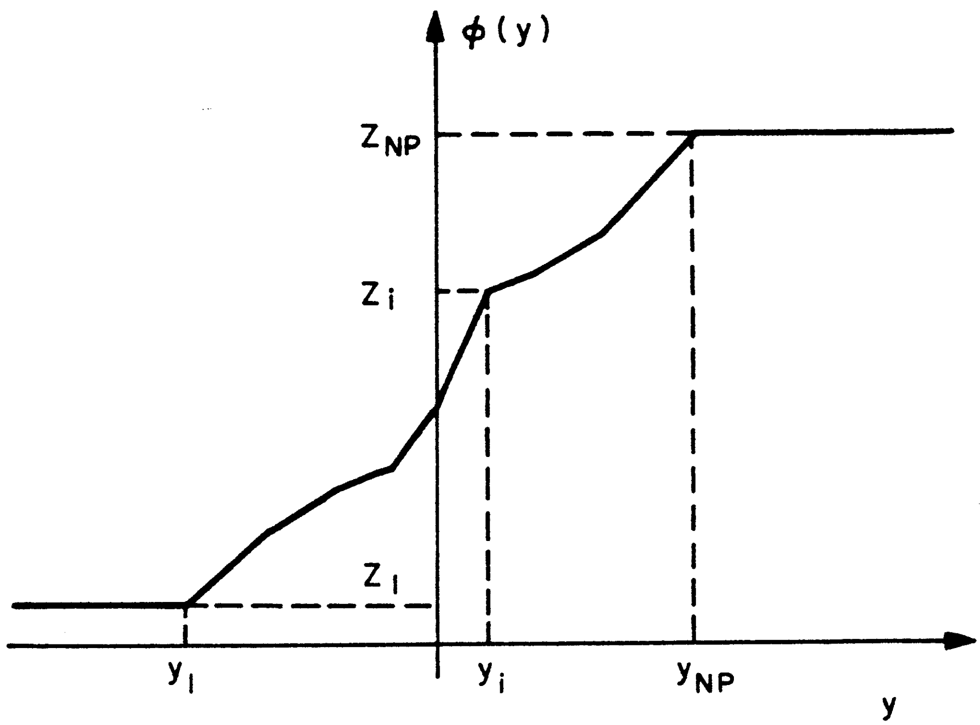


Figure 3.2

Linear Interpolation of the Anamorphosis Function

If the mean of the field is not near the value ϕ_0 calculated from the interpolation of the anamorphosis, then the interpolation used was not appropriate, and another one should be employed. Appendix A gives a simple form of the coefficients ϕ_k when linear interpolation of the anamorphosis is used.

The Disjunctive Kriging estimator is approximated in practice by solving a finite number of systems of equations:

$$\mathscr{L}(Z_0)^*_{DK} = \sum_{\alpha=1}^{N} \sum_{k=0}^{m1} f_{\alpha,k}\, \eta_k(Y_\alpha) \tag{3.45}$$

It is assumed that the order of approximation needed to expand the unknown functions f_α is the same that is needed to represent ϕ. Because the variance of the field Z admits the following representation:

$$\mathrm{Var}(Z) = \int_{-\infty}^{\infty} \left[\sum_{k=1}^{\infty} \phi_k\, \eta_k(Y) \right]^2 G(dY) = \sum_{k=1}^{\infty} \phi_k^2 \tag{3.46}$$

and the contributions ϕ_k^2 decrease rapidly as k increases, the value of m1 can be taken such that $\sum_{k=1}^{m1} \phi_k^2$ is sufficiently near to Var(Z).

Although a finite expansion is used, it is important to realize that the Hermitian finite representation is the best possible polynomial approximation of order m1 of a function, in the sense that:

$$I = \int_{-\infty}^{\infty} \left(f(x) - \sum_{k=0}^{m1} \xi_k\, x^k\right)^2 G(dx) \tag{3.47}$$

is minimized precisely when:

$$\sum_{k=0}^{ml} \xi_k \, x^k = \sum_{k=0}^{ml} f_k \eta_k(x) \tag{3.48}$$

3.6 Hierarchy of the Kriging Estimators

As has been seen in previous sections, the Linear and Disjunctive Kriging estimators are linear and separable combinations of the data variables, such that unbiasedness is satisfied and such that the estimation variance is minimized.

If the form of the estimator is taken as N variable measurable functions of the observations, $f(Z_1,\ldots,Z_N)$, under the unbiasedness and minimum variance conditions the conditional expectation $E(\mathscr{L}(Z_0)|Z_1,\ldots,Z_N)$ is reached. Because the linear and separable forms of estimators are particular cases of the general N-variable measurable functions, the accuracy of the conditional expectation is better than that of Disjunctive and Linear Kriging, i.e.:

$$\sigma^2_K \geq \sigma^2_{DK} \geq \sigma^2_{CE} \tag{3.49}$$

with:

$$\sigma^2_{CE} = E\{[\mathscr{L}(Z_0) - E(\mathscr{L}(Z_0)|Z_1,\ldots,Z_N)]^2\} \tag{3.50}$$

The accuracy of the conditional expectation is based on the knowledge of the N+1-variate distribution function of $(\mathscr{L}(Z_0), Z_1,\ldots,Z_N)$. Because this distribution function cannot be adequately estimated from the observations of the field, the conditional expectation has no practical applicability, and simpler, less accurate estimators, such as the ones presented in this work, should be employed.

Table 3.1 summarizes the different estimators together with their requirements.

	Linear Kriging	Disjunctive Kriging	Conditional Expectation
Form	$\mathscr{L}(Z_0)^*_k = \sum_{i=1}^{N} \lambda_i Z_i$	$\mathscr{L}(Z_0)^*_{DK} = \sum_{i=1}^{N} f_i(Z_i)$	$\mathscr{L}(Z_0)^*_{CE} = f(Z_1,\ldots,Z_N)$
Requisites	Drift Form Covariance, Semivariogram or Generalized Covariance Function	Bivariate Distributions of (Z_i,Z_j) and $(\mathscr{L}(Z_0), Z_i)$	(N+1) Variate Distribution of $(\mathscr{L}(Z_0), Z_1,\ldots,Z_N)$
Estimation Variance	$\sigma^2_K \geq$	$\sigma^2_{DK} \geq$	σ^2_{CE}

Table 3.1

Hierarchy of Estimators of a Random Field

(after Huijbregts and Journel, 1978)

3.7 Summary

In this chapter the Disjunctive Kriging equations for point and block estimation of a random field have been presented. Assuming that the data comes from a second-order stationary Gaussian field, via an anamorphosis function, the need of bivariate distributions has been reduced to the knowledge of the covariance structure of the Gaussian variables. Using Hermite expansions of the unknown functions as well as the anamorphosis function, it has been seen that an infinite number of simultaneous equations characterize the estimator. It has also been shown how the Hermitian fit of the anamorphosis function provides a way to determine a finite approximation of the estimator.

Chapter 4

POINT ESTIMATION COMPARISONS

This chapter discusses and compares the performance of Universal and linear Kriging at point estimation.

4.1 Experiments Description

In order to make an appropriate comparison of the different estimators, a simulated field was used as the reality. The turning bands method, introduced by Matheron 1973, was used to generate isotropic fields as well as intrinsic random functions of different orders.

The fields were generated using the computer program TUBA, developed by Mantoglou and Wilson, 1981, at points on a rectangular area of 30,000 Km^2 with sides on proportion 2 and 1, on a 21x11 regular grid, which gives a distance between points of 12.25 Km; see Figure 4.1.

Three different types of random fields were studied. Considered first where isotropic fields with an exponential covariance structure of the form:

$$Cov(h) = \sigma^2 \exp(-bh) \tag{4.1}$$

where σ^2 and b represent the variance and correlation parameter of the field, respectively. Correlation increases as b decreases. The second set are fields obtained by squaring isotropic random fields with the above covariance function. Finally, intrinsic random functions with generalized covariance functions of the form:

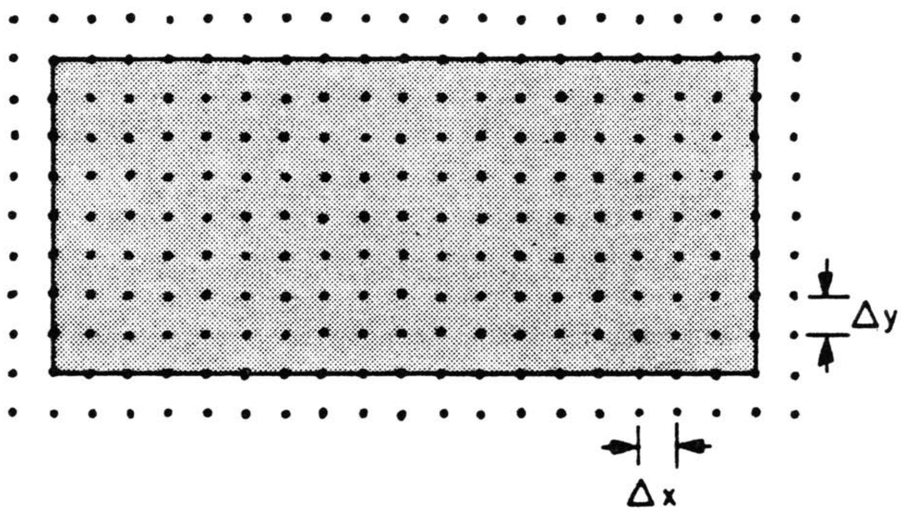


Figure 4.1

Grid Used in the Point Estimation Comparisons

$\Delta X = \Delta Y = 12.25$ Km

$$K(h) = \alpha_1 |h| + \alpha_3 |h|^3 + \alpha_5 |h|^5 \tag{4.2}$$

were considered. Appendix C includes plots of these fields.

Prior to the description of the different cases considered, the following definitions and remarks are in order. Following Matern, 1960, define the typical distance of a plane figure as the mean distance between two points chosen randomly and uniformly in the figure. It can be shown, Rodriguez-Iturbe and Mejia (1973), that for the previously defined rectangle the typical distance, TD, is about 98.6 Kms. Define also the correlation distance of a covariance function, CD, as the length at which the covariance function has dropped to half of its value at the origin, i.e.,

$$\text{Cov}(CD) = \frac{1}{2}\,\text{Cov}(0)$$

For the isotropic fields, six different combinations on the parameters σ^2 and b were studied: three different cases of spatial correlation, and two levels of variability. The three cases of correlations were taken such that CD = 2TD, CD = TD, and CD = $\frac{1}{2}$ TD, which gives values of the parameter b of 0.0035, 0.007, and 0.0141, respectively. For each correlation structure the point variability parameter σ was set to values of 10 and 20; with the stationary mean kept constant in all the cases at a value of 10. Table 4.1 summarizes these fields.

Zero mean fields, $Z_1(\underline{u})$, were generated using the same correlation parameters discussed in the previous paragraph, and then were transformed by:

b	σ
0.0035	10
0.007	10
0.0141	10
0.0035	20
0.007	20
0.0141	20

Table 4.1

Structure of the Generated Isotropic Fields

$$Cov(h) = \sigma^2 \exp(-bh)$$

Stationary mean = m = 10

$$Z(\underline{u}) = \frac{1}{30} Z_1^2(\underline{u}) \tag{4.3}$$

to get fields summarized on Table 4.2. The constant 30 was used to have the transformed fields and the previously defined isotropic fields of about the same order to magnitude.

Eight different Intrinsic Random Functions, with the previously described polynominal generalized covariance structure, were considered. The parameters were taken such that these generated fields are of about the same order of magnitude as the previously outlined isotropic and transformed isotropic fields. Table 4.3 details the cases considered.

For each of the twenty fields, six sets of points were chosen randomly from the 21x11 grid. Three of these sets of about fifty points[1] and the other three of about thirty points. The values of the generated fields at these points defined the observations network from which structure identification for the different estimators was made. This subsequently allowed the estimation of the complete inner 19x9 grid, from which comparisons of the different estimators were made by studying the differences between the true (generated) and estimated values.

The boundaries of 21x11 grid were not considered to avoid biased estimates. The fifty and thirty points taken, represent precipitation networks of 1.6 to 1 station per 1000 Km^2, respectively. It should be pointed out, however, that only the isotropic correlation parameter, b, resemble observed precipitation records, Rodriguez-Iturbe and Mejia, 1973, and that negative generated values were permitted.

---

[1] In general less due to repetition during sampling

b	σ
0.0035	10
0.007	10
0.0141	10
0.0035	20
0.007	20
0.0141	20

Table 4.2

Structure of the Generated Transformed Isotropic Fields

$$\mathrm{Cov}_{Z_1}(h) = \sigma^2 \exp(-bh)$$

$$\text{Stationary mean} = m_{Z_1} = 0$$

$$\text{Transformation: } Z(\underline{u}) = \frac{1}{30} Z_1^2(\underline{u})$$

IRF Order	α_1	α_3	α_5
0	-1	0	0
0	-3	0	0
1	0	0.005	0
1	-0.005	0.005	0
2	0	0	$-0.5\text{x}10^{-8}$
2	0	0.005	$-0.5\text{x}10^{-8}$
2	-1	0	$-0.5\text{x}10^{-9}$
2	-2	0.05	$-1\text{x}10^{-9}$

Table 4.3

Structure of the Generated Intrinsic Random Functions

$$K(h) = \alpha_1|h| + \alpha_3|h|^3 + \alpha_5|h|^5$$

In the following sections, the different estimators and the measures used in the comparisons are presented.

4.2 Description of the Estimators

Not only the techniques but also different ways of estimating their parameters were compared. This section discusses the various parameter identification methods used with each technique.

4.2.1 Universal Kriging Estimators

Three methods for identifying and estimating the polynomial generalized covariance from limited data were investigated.

The first method was previously described in Section 2.9.2 and was introduced by Delfiner, 1976. The optimal polynomial model of generalized covariance function is found by first finding the IRF order k by a ranking scheme, then estimating all the models corresponding to that k by an iterative procedure, and finally selecting the model with closest to one jackknife estimator of the ratio of sampled and theoretical mean square error.

The second and third methods considered all the twenty five possible models applicable to IRF's up to order 2. Once all the models have been estimated as in Section 2.9.2.1, the feasible ones, say m2 of the twenty five, are ranked as in Section 2.9.2.1, letting the point ranks go from 0 for the best model up to m2-1 for the worst model. The second procedure selects the model with lowest rank.

When the optimal parameters of the generalized covariance models are calculated, the optimal measure of the fit is also found, i.e., the mean square error of the optimal regression, Q/N, with Q given by Equation (2.110). The third procedure selects among the feasible models the one that has lowest mean square error.

In all the cases, the identification and estimation of the optimal models was done using $N_0 = 16$ neighboring points. Actual estimation of field values in the 19x9 inner grid was done from the N = 8 closest points.

4.2.2 Disjunctive Kriging Estimators

Two different models for the semivariogram function of the Gaussian data were considered.

First a spherical semivariogram of the form:

$$\gamma(h) = \begin{cases} C_o\left\{\frac{2}{3}\frac{h}{a_o} - \frac{1}{2}\frac{h^3}{a_o^3}\right\} & , \; h < a_o \\ C_o & , \; h > a_o \end{cases} \tag{4.4}$$

in which C_o represents the <u>sill</u> and a_o the <u>range</u>. This model was fitted using the computer progam MAREC given in David 1977, and a weighted least squares procedure with the sill constrained to be the sample variance of the Gaussian data.

The second model is a linear semivariogram:

$$\gamma(h) = C\delta(h) + \alpha_1|h| \tag{4.5}$$

This model, being the negative of a generalized covariance function of an IRF-0, can be estimated using the iterative procedure explained on Section 2.9.2.1. All the three alternative models given by Equation (4.5) were estimated, and the best mean square model, i.e., the one that minimizes Q given by Equation (2.110), was chosen.

For both cases, once the semivariogram has been estimated, the covariance function, of the Gaussian data can be readily estimated from Equation (2.8). As in the Universal Kriging cases, the field values in the 19x9 inner grid were estimated using the N = 8 closest data points. The anamorphosis function was interpolated linearly (see Appendix A) and fitted using the finite Hermite expansions given by the first five standardized Hermite polynominals (see Equation (3.35)).

4.2.3 Universal Kriging of Transformed Data

The Universal Kriging estimator of the Gaussian variates was also calculated, and then transformed through the anamorphosis function to get an estimator of the actual field:

$$\hat{Z}(\underline{u}) = \phi(\hat{Y}(\underline{u})) \tag{4.6}$$

with

$$\hat{Y}(\underline{u}) = \sum_{i=1}^{N} \lambda_i \; Y_i \tag{4.7}$$

and the function ϕ approximated by an order ml expansion:

$$\phi(u) = \sum_{k=o}^{ml} \phi_k \; \eta_k(u) \tag{4.8}$$

All the 25 models of generalized covariance functions of the Gaussian data were considered, and the optimal one was assumed given by the least mean square error criterion.

The generalized covariance functions were estimated using $N_0 = 16$ neighboring points, while estimation of the inner grid used the $N = 8$ closest points. The approximation order, ml, was taken as in the Disjunctive Kriging estimator at a value of 4.

Notice that although an estimate of the predicted variance of the Gaussian values can be found at any point in the 19x9 grid, the predicted variance of the actual field cannot be found, due to the non-linearity of the anamorphosis function.

4.2.4 Local Mean Estimator

A simple local mean estimator was also considered. The estimate at a point was taken as the average of the 5 closest data points.

4.3 Description of the Measures of Comparison

In this section the different measures used to compare the performance of the different estimators are described. The following attributes were calculated for every case:

i) Mean Square Error, MSE, defined as

$$\frac{1}{N_1} \sum_{j=1}^{N_1} (Z_j - \hat{Z}_j)^2 \tag{4.9}$$

where Z_j represents actual values (generated using Turning bands), and $\hat{Z}_j$ are the estimated values. The number of points averaged, N_1, was not necessarily 19x9 = 171 because unusual outliers were not taken into consideration. Outliers were defined as those such that:

$$\hat{Z}_j \geq \max_i Z_i + 2(\max_i Z_i - \min_i Z_i)$$

or

$$\hat{Z}_j \leq \min_i Z_i - 2(\max_i Z_i - \min_i Z_i)$$

They were not considered because their magnitude would give unrealistic estimates of the overall performance of the estimators. Notice that in any case N_1 is also a measure of how the estimators work, with better results if N_1 is near 171.

ii) Correlation Coefficient, CC, between the actual and estimated values, calculated using the same N_1 points as above. Values closest to one are indicative of good agreement.

iii) Mean of Deviations, μD, and the standard deviation of the deviations, σD, where the deviations are $[Z_j - \hat{Z}_j]$.

iv) Mean Predicted Variance, MPV, over the grid - N_1 points - and the Maximum predicted variance, MaPV on the 19x9 grid.

v) Consistency Parameter, $\hat{\rho}_1$, defined as the ratio of the MSE ~~over the MPV, a generalization of~~ Equation (2.111). Notice that values of $\hat{\rho}_1$ near one indicate good agreement between what happens in reality and what the employed estimator predicts.

Graphical representations of typical cases are also given in the section that compares the different estimators, Section 4.6. They are, mainly, plots of real versus estimated values that provide a visual idea of how well the estimators work on a given field for a given set of historical points. Points on a 45° line would indicate perfect agreement.

4.4 Comparison of the Universal Kriging Estimators

In this section the different methodologies used in the determination of the optimal generalized covariance function are compared according to the measures described in the previous section. The experiments were performed using the computer package AKRIP, developed by Kafritsas and Bras, 1981.

4.4.1 Isotropic Fields

The results of the thirty six isotropic field cases (six different fields sampled six times each) are summarized on Table 4.4.

First, the three methods (R for ranking procedure, M for least mean square error procedure, and D for Delfiner's methodology) are compared quantitatively according to the different measures, at different levels. For example, in 29 of the 36 cases the ranking procedure gave a mean square error that was either the lowest of the three methods or had a value that differed from the lowest by less than 10 percent of that value.

Following the comparative measures, there is a summary of the selected orders of IRF given by the three procedures, as well as the number of cases in which the methods selected the same order and the number of cases in which models given by different procedures coincide.

Finally, the table includes information concerning the consistency parameter, $\hat{\rho}_1$, i.e., the minimum and maximum values over all the cases; the number of cases less than or equal to one; the number of cases

	MSE		MPV		MaPV	CC	μD	σD	MPV_1	
	(10%)	(25%)	(10%)	(25%)	(25%)	(5%)	($\pm$0.5)	(10%)	(10%)	(25%)
R	29		30		20	32	32	30	27	
M	32	34	29	35	27	35	33	33	30	35
D	6	9	1	4	13	21	25	11	1	4

	Selected IRF 0	1	2
R	9	15	12
M	13	16	7
D	16	10	10

Same Order	Same Values
RMD-15	RMD- 0
RM - 8	RM -20
RD -11	RD - 2
DM - 1	DM - 0

CONSISTENCY

$$(\hat{\rho}_1 = \frac{MSE}{MPV})\ (PRIOR)$$

	Min	Max	#$\leq$1	#>1	# [0.8,1.2]	μ	σ
R	0.23(0.73)	22.7(2.4)	10(11)	26(24)	8 (19)	2.05(1.2)	2.0(0.41)
M	0.26(0.73)	19.5(3.2)	11(11)	25(25)	9 (22)	1.59(1.15)	1.21(0.31)
D	0.13(0.76)	1.14(1.2)	34(28)	2(8)	10(34)	0.62(0.99)	0.23(0.06)

Table 4.4

Comparison of Universal Kriging Estimators

Isotropic Fields (36 cases) Point Results greater than one; the number of cases between 0.8 and 1.2; and the mean and standard deviation of $\hat{\rho}_1$ over all the cases. The prior information of $\hat{\rho}_1$, given by the jackknife estimator as in Equation (2.113) using the historical N points, is also given in parentheses.

As is shown the three methods did not prefer any IRF order and although they selected the same order in 15 cases, there was not three way agreement on the same model. Ranking and least mean square error procedures coincided 20 times in selecting the generalized covariance model. Only in two cases the Delfiner's methodology coincided with the ranking or least mean square error procedures.

As is shown in Appendix B, when $k = 0$ the distance-independent model $K(h) = C\delta(h)$, gives a jacknife estimate value of one. That is why in 16 cases the Delfiner's methodology selected that model. Since in this case the Kriging estimator weights are all equal to one over the number of points used in the estimation, i.e., 1/8, the use of that model gives a conservative estimate that will not capture the extreme values of the field and will overestimate the variance of estimation. Figures 4.2 b, 4.3 b, and 4.6 b show typical outcomes using the above model.

Table 4.4 indicates that the ranking and least mean square error procedures performed better than the Delfiner's methodology, as expressed by lower MSE, MPV, MaPV, and σD, closer to one CC and closer to zero μD[1]. This is confirmed when the least mean square error and Delfiner's method are compared at a 25 percent level on MSE and MPV: the first

[1] This attribute was calculated defining an interval of equally good performance by adding and subtracting 0.5 to the μD of the method that deviated less from zero.

method gives as good or better results in almost all the cases. Figures 4.2 a, b; 4.3 a, b; 4.4 a, b, c; 4.5 a, b; 4.6 a, b; 4.7 a, b; and 4.8 a, b, c show comparisons of real versus estimated points for the Universal Kriging Estimators for typical cases of isotropic fields and various model identification approaches.

The consistency parameter table shows that the actual performance deviates more from 1 than the prior jackknife estimator, for all the three methods. The Delfiner's approach tends to have values of $\hat{\rho}_1$ less than 1 which means that the model given by this method overestimates the variance of estimation. On the other hand the ranking and least mean square error procedures select models that tend to have values of $\hat{\rho}_1$ greater than one, which indicates that these procedures underestimate the variance of estimation. It should be pointed out that very extreme values of $\hat{\rho}_1$ (below 0.15 and above 6) were not considered in the calculation of the mean and standard deviation of the consistency parameter.

The MPV was corrected for all experiments using the prior estimator of $\hat{\rho}_1$, i.e., the jackknife estimator. This new index, MPV_1, (see Table 4.4) does not reveal additional insight and gives very close results to those of the MPV: according to the MPV_1 the performance of Delfiner's method is not better than that of the other methods.

4.4.2 Transformed isotropic Fields

Table 4.5 summarizes the results of the 36 cases of transformed isotropic fields considered. The general results are similar to those of previously described isotropic field cases.

	MSE		MPV		MaPV	CC	μD	σD	MPV_1	
	(10%)	(25%)	(10%)	(25%)	(25%)	(5%)	(± 0.5)	(10%)	(10%)	(25%)
R	31		24		20	33	30	34	28	
M	27	35	33	36	32	28	32	30	31	36
D	5	13	3	5	15	4	27	8	4	6

Selected IRF

	0	1	2
R	22	9	5
M	28	3	5
D	29	3	4

Same Orders
RMD-22
RM - 3
RD - 5
DM - 4

Same Values
RMD- 0
RM -17
RD - 1
DM - 5

CONSISTENCY

$(\hat{\rho}_1 = \frac{MSE}{MPV})$ (PRIOR)

	Min	Max	$\# \leq 1$	$\# > 1$	# [0.8,1.2]	μ	σ
R	0.33(0.66)	14.5(2.4)	13(13)	23(23)	11(26)	1.81(1.15)	1.67(0.34)
M	0.37(0.83)	14.5(2.4)	11(12)	25(24)	6(27)	2.23(1.2)	2.07(0.36)
D	0.18(0.83)	3.2(1.08)	26(34)	10(2)	10(36)	0.92(0.99)	0.65(0.03)

Table 4.5

Comparisons of Universal Kriging Estimators

Transformed Isotropic Fields (36 cases) Point Results

The three identification methods tended to select an IRF of order 0 as the structure that best explained the given fields. Never did all the methods give the same generalized covariance model. The ranking and least mean square error procedures coincided 17 times. Again, Delfiner's approach does not give better results than the other two methods, which cannot be distinguished.

Delfiner's approach shows a slight tendency of overestimating the predicted variance, while the other methods again tend to underestimate it. Again, some very extreme values of $\hat{\rho}_1$ were obtained indicating inconsistent performance. However, plots of real versus estimated values can still be very satisfactory as shown in Figures 4.9 a, and 4.11 a.

Figures 4.9 a, b; 4.10 a, b; c; 4.11 a, b; and 4.12 a, b show typical Universal Kriging outcomes for the experiments with Transformed Isotropic fields.

4.4.3 Intrinsic Random Functions

Tables 4.6, 4.7, and 4.8 summarize the results found with intrinsic random functions of orders 0, 1, and 2, respectively. The results follow the same tendencies previously encountered on the isotropic and transformed isotropic fields.

Although in 31 of the 48 cases (8 fields each sampled 6 times) all the three procedures selected the same IRF order, only in 5 of them all the methods coincided in the estimation of the model. As in previous cases, the ranking and least mean square error procedures were the combination that most often gave the same generalized covariance model.

	MSE (10%)	MSE (25%)	MPV (10%)	MPV (25%)	MaPV (25%)	CC (5%)	μD ($\pm$0.5)	σD (10%)	MPV_1 (10%)	MPV_1 (25%)
R	0		10		8	10	12	11	10	
M	11	12	10	11	9	12	11	11	11	11
D	4	4	2	3	10	3	11	4	4	5

Selected IRF

	0	1	2
R	5	6	1
M	8	3	1
D	5	6	1

<u>Same Order</u>

RMD- 8
RM - 0
RD - 4
DM - 0

<u>Same Values</u>

RMD- 1
RM - 7
RD - 1
DM - 0

CONSISTENCY

$(\hat{\rho}_1 = \frac{MSE}{MPV})$ (PRIOR)

	Min	Max	#$\leq$1	#>1	# [0.8,1.2]	μ	σ
R	0.6 (0.85)	3.28(1.87)	5(5)	7(7)	4(10)	1.29(1.11)	0.86(0.27)
M	0.62(0.85)	2.39(1.29)	4(5)	8(7)	5(9)	1.07(1.17)	0.57(0.14)
D	0.4 (0.99)	0.89(1.03)	9(6)	3(6)	3(12)	0.89(1.03)	0.49(0.06)

Table 4.6

Comparisons of Universal Kriging Estimators

Intrinsic Random Functions of Order 0 (12 cases) Point Results

	MSE		MPV		MaPV	CC	μD	σD	MPV_1	
	(10%)	(25%)	(10%)	(25%)	(25%)	(5%)	(±0.5)	(10%)	(10%)	(25%)
R	9		10		10	12	12	11	9	
M	11	12	12	12	11	12	12	11	12	12
D	2	2	3	2	1	12	12	4	2	2

Selected IRF

	0	1	2
R	0	7	5
M	0	8	4
D	0	3	9

<u>Same Order</u>

RMD- 7
RM - 4
RD - 1
DM - 0

<u>Same Values</u>

RMD- 1
RM - 8
RD - 1
DM - 1

CONSISTENCY

$(\hat{\rho}_1 = \frac{MSE}{MPV})$ (PRIOR)

	Min	Max	#≤1	#>1	# [0.8,1.2]	μ	σ
R	0.44(0.84)	43.2 (3.19)	7(5)	5(7)	5(9)	1.2(1.26)	1.3(0.67)
M	0.48(0.85)	43.2 (3.19)	5(4)	7(8)	5(7)	1.45(1.37)	1.26(0.68)
D	0.18(0.86)	1.14(1.07)	9(8)	3(4)	4(12)	0.63(0.98)	0.34(0.06)

Table 4.7

Comparison of Universal Kriging Estimators

Intrinsic Random Functions of Order 1 (12 cases) Point Results

	MSE		MPV		MaPV	CC	μD	σD	MPV_1	
	(10%)	(25%)	(10%)	(25%)	(25%)	(5%)	($\pm$0.5)	(10%)	(10%)	(25%)
R	18		14		16	24	23	20	17	
M	22	24	20	23	21	24	24	24	18	24
D	10	13	4	6	5	24	20	12	6	6

Selected IRF

	0	1	2
R	0	7	17
M	0	8	16
D	0	2	22

Same Order

RMD-16
RM - 3
RD - 3
DM - 2

Same Values

RMD- 3
RM - 8
RD - 5
DM - 1

CONSISTENCY

$(\hat{\rho}_1 = \frac{MSE}{MPV})$ (PRIOR)

	Min	Max	#≤1	#>1	# [0.8,1.2]	μ	σ
R	0.12(0.58)	7.7(2.31)	9 (9)	15(15)	5(12)	1.6 (1.2)	1.19(0.47)
M	0.16(0.24)	9.65(3.58)	9 (8)	15(16)	5(9)	1.63(1.4)	1.24(0.78)
D	0.10(0.66)	7.7(1.84)	21(12)	3(12)	6(20)	0.65(1.0)	0.37(0.21)

Table 4.8

Comparison of Universal Kriging Estimators

Intrinsic Random Functions of Order 2 (24 cases) Point Results

As is shown the procedures tended to select the IRF-order that was used to generate the historical data. The least mean square method was the most consistent in this respect. When the generated IRF-order was 1 or 2 none of the method gave models with IRF-orders 0.

The overall performance of the ranking and least mean square error procedure is better than that of Delfiner's methodology. For example, the least mean square error method gave results that were best in MSE at 10 percent in 44 of the 48 cases, while Delfiner's approach only gives good results in 16 of the 48 cases. At 25 percent level, the least mean square error procedure is best in 48 cases versus 19 for Delfiner's method.

Fields corresponding to the two intrinsic random functions of order 1 and the first two and last of order 2 (see Table 4.3), were very well estimated as can be seen in Figures 4.15 a, b; 4.16 a, b, c; and 4.18 a, b. The correlation coefficient was on the order of 0.99 for all the methods and cases. The figures also show that the model selected by Delfiner's methodology gives a wider band around the 45° line, increasing the MSE.

The consistency parameter again shows the tendency of underestimating the variance of estimation by the ranking and least mean square error procedures, with the actual values having a wider range than the prior jacknife estimates. Delfiner's methodology again tends to overestimate the predicted variance.

Typical results for the Universal Kriging estimators are shown in Figures 4.13 a, b, c and 4.14 a, b for IRF-0, in Figures 4.15 a, b and 4.16 a, b, c for IRF-1 and in Figures 4.17 a, b and 4.18 a, b for IRF-2.

It should be pointed out that in none of the IRF cases nor in the other random fields considered, a generalized covariance model with more than two parameters was chosen by any of the methodologies. This suggests the use in practice of fewer than 25 models, discarding from the beginning those that have three or four parameters. Of course, since the 120 cases considered (36 + 36 + 48) do not include all the possible random fields, the above point should be investigated further.

4.5 Comparison of the Disjunctive Kriging Estimators

In this section the two methods for estimating the semivariogram function of the Gaussian data to be used in the Disjunctive Kriging estimator are compared. In Tables 4.9 to 4.13 the final results are summarized; L stands for the linear semivariogram and S for the spherical semivariogram.

4.5.1 Isotropic Fields

As is shown in Table 4.9, the spherical semivariogram gave better results in MSE, but the linear semivariogram did better on MPV. A further check with the consistency parameter reveals a more coherent Disjunctive Kriging estimator if the spherical semivariogram is used. The MPV given by the linear semivariogram is lower than it should be in reality which explains the apparent advantage it has according to that measure.

In Figures 4.2 c, d; 4.3 c, d; 4.4 d, e; 4.5 c; 4.6 c, d; 4.7 c; and 4.8 d some Disjunctive Kriging results for isotropic fields are shown.

4.5.2 Transformed Isotropic Fields

As in the previously described isotropic field cases, the Disjunctive Kriging estimator using the spherical semivariogram gave better overall results than using a linear semivariogram. Both semivariograms underestimated variances of estimation as seen in Table 4.10. The spherical semivariogram not only performed better with respect to the MSE but also with respect to MPV, even though the use of the linear semivariogram led to more serious underestimation of the predicted variances.

	MSE (10%)	MSE (25%)	MPV (10%)	MPV (25%)	MaPV (25%)	CC (5%)	μD (± 0.5)	σD (10%)
L	21	23	27	29	22	28	32	22
S	32	33	17	21	21	35	32	33

CONSISTENCY

$$\left(\hat{\rho}_1 = \frac{MSE}{MPV}\right)$$

	Min	Max	$\# \leq 1$	$\# > 1$	# [0.8,1.2]	μ	σ
L	0.49	3.95	13	23	12	1.43	0.89
S	0.22	3.72	20	16	12	1.11	0.74

Table 4.9

Comparisons of Disjunctive Kriging Estimators

Isotropic Fields (36 cases) Point Results

	MSE		MPV		MaPV	CC	μD	σD
	(10%)	(25%)	(10%)	(25%)	(25%)	(5%)	($\pm$0.5)	(10%)
L	16	22	20	23	15	17	32	20
S	35	36	22	26	28	35	33	35

CONSISTENCY

$$\left(\hat{\rho}_1 = \frac{MSE}{MPV}\right)$$

	Min	Max	$\#\leq 1$	$\#>1$	[0.8,1.2]	μ	σ
L	0.17	5.79	9	27	5	2.09	1.62
S	0.21	12.16	13	23	5	1.8	1.6

Table 4.10

Comparison of Disjunctive Kriging Estimators

Transformed Isotropic Fields (36 cases) Point Results

Figures 4.9 c, d; 4.10 d, e; 4.11 c, d; and 4.12 c, d show some of the Disjunctive Kriging results for the transformed isotropic fields.

4.5.3 Intrinsic Random Functions

The results of estimating intrinsic random functions using Disjunctive Kriging are presented in Tables 4.11, 4.12, and 4.13 for orders 0, 1, and 2, respectively.

When the order is 0, the spherical variogram gives better results in all the attributes, as can be seen in Table 4.11. As in the previously described isotropic and transformed isotropic cases, the Disjunctive Kriging estimator tends to underestimate the variance of estimation with both semivariograms, with the spherical one being slightly more consistent than the linear semivariogram. Figures 4.13 d, e and 4.14 c show typical results of these cases.

The linear semivariogram gave better results, according to MSE, MPV, and MaPV for intrinsic random functions of orders 1 or 2. As can be seen in Table 4.12 and 4.13, both methods overestimated the variance of estimation, with the linear semivariogram doing better. Figures 4.15 c, d; 4.16 d, e; 4.17 c, d; and 4.18 c, d show some of these results.

Notice from Equation (2.8) that using the linear semivariogram may lead to negative covariance estimates if a neighbor point is far enough such that the value of the semivariogram is greater than the sampled

	MSE		MPV		MaPV	CC	μD	σD
	(10%)	(25%)	(10%)	(25%)	(25%)	(5%)	($\pm$0.5)	(10%)
L	1	2	5	5	2	4	9	2
S	12	12	10	10	11	11	10	12

CONSISTENCY

$$\left(\hat{\rho}_1 = \frac{MSE}{MPV}\right)$$

	Min	Max	$\# \leq 1$	$\# > 1$	# [0.8,1.2]	μ	σ
L	0.85	5.84	4	8	4	1.96	1.49
S	0.77	3.18	4	8	3	1.53	0.87

Table 4.11

Comparisons of Disjunctive Kriging Estimators

Intrinsic Random Functions of Order 0 (12 cases) Point Results

	MSE		MPV		MaPV	CC	μD	σD
	(10%)	(25%)	(10%)	(25%)	(25%)	(5%)	(± 0.5)	(10%)
L	11	12	11	11	10	12	10	11
S	6	10	1	1	2	12	12	9

CONSISTENCY

$$(\hat{\rho}_1 = \frac{MSE}{MPV})$$

	Min	Max	$\# \leq 1$	$\# > 1$	# [0.8,1.2]	μ	σ
L	0.17	1.42	8	4	4	0.78	0.42
S	0.07	0.7	12	0	0	0.35	0.2

Table 4.12

Comparisons of Disjunctive Kriging Estimators

Intrinsic Random Functions of Order 1 (12 cases) Point Results

	MSE		MPV		MaPV	CC	μD	σD
	(10%)	(25%)	(10%)	(25%)	(25%)	(5%)	(± 0.5)	(10%)
L	19	19	21	21	19	22	21	20
S	15	17	6	6	7	21	21	18

CONSISTENCY

$$\left(\hat{\rho}_1 = \frac{MSE}{MPV}\right)$$

	Min	Max	$\# \leq 1$	$\# > 1$	# [0.8,1.2]	μ	σ
L	0.07	8.38	17	7	4	1.0	1.1
S	0.07	10.86	19	5	1	0.58	0.78

Table 4.13

Comparisons of Disjunctive Kriging Estimators

Intrinsic Random Functions of Order 2 (24 cases) Point Results

variance of the Gaussian data. In fact, the linear semivariogram works very well when many historical data points are available, but as the number of points decreases the spherical semivariogram will give better results, because by definition (see Equation (4.4)) it will produce non-negative values of the covariance function.

4.6 Comparison of the Different Estimators

The different estimators are compared in this section. The Universal Kriging estimator used is the one obtained by estimating the generalized covariance function employing the slightly better least mean square error procedure. It was also chosen because its use requires less computational effort than the ranking procedure. The spherical semivariogram of the Gaussian data was used to compute the Disjunctive Kriging estimator.

4.6.1 Isotropic Fields

Table 4.14 summarizes the results of the comparisons when estimating isotropic fields. K_M stands for Universal Kriging with its generalized covariance function estimated using the least mean square error methodology. DK_s is the Disjunctive Kriging estimator with its covariance function obtained via a spherical semivariogram. KT stands for the Universal Kriging estimation of Gaussian data and back transformation using the anamorphosis function. And M_5 is the local mean estimator calculated by averaging the closest five data points. Only the MSE and MPV are given in the table, since they are the most sensitive attributes.

Table 4.14 shows that Universal Kriging gave better results than Disjunctive Kriging in both MSE and MPV at a 10 percent level. When the level is increased Universal Kriging still performs better in MPV but Disjunctive Kriging improves significantly. The consistency para-

	MSE			MPV		
	(10%)	(25%)	(50%)	(10%)	(25%)	(50%)
K_M	29	30	35	30	30	32
DK_s	18	27	31	7	10	14
KT	16	24	32			
M_5	3	5	11			

CONSISTENCY

$$(\hat{\rho}_1 = \frac{MSE}{MPV})$$

	Min	Max	$\# \leq 1$	$\# > 1$	# [0.8,1.2]	μ	σ
K_M	0.26	19.5	11	25	9	1.59	1.21
DK_s	0.22	3.72	20	16	12	1.11	0.74

Table 4.14

Comparisons of the Different Estimators
Isotropic Fields (36 cases) Point Results

meter indicates an underestimated variance of estimation for the Universal Kriging procedure but this fact does not help to explain the advantage this estimator has over the theoretically more accurate Disjunctive Kriging estimator at a 50 percent level on MPV.

It is important to keep in mind that a finite Hermite expansion is being used in the Disjunctive Kriging estimator, but even more important is the fact that the use of a finite amount of data results in inconsistencies in the Hermitian model that permits a solution for the optimal functions. Recall the procedure to obtain the Gaussian data from the original observations, see Figure 3.1. Because of the symmetry of the Gaussian distribution, the Gaussian data will be composed by plus and minus values. This means that when the sampled mean of the Gaussian data is calculated, it gives exactly a value of zero. However, because the tails of the Gaussian distribution are sampled up to a finite value, the sampled variance of the Gaussian data will be less than one, in contradiction with what the Hermitian model assumes. The actual variance of the Gaussian data depends only on the number of observations. Some key values are given on Table 4.15. Notice that this inconsistency will be present even if the anamorphosis function is perfectly fitted.

With about 30 and 50 data points, the Disjunctive Kriging estimator leads to values of the sampled variance of the Gaussian data of about 0.84 and 0.88, underestimating the theoretical value of one. This inconsistency may explain why the estimated variance for Disjunctive Kriging was not found smaller than that of the simpler and theoretically less accurate Universal Kriging estimator.

Number of Points	Sampled Variance
10	0.690
15	0.755
20	0.794
25	0.821
30	0.841
35	0.856
40	0.869
50	0.888
70	0.912
100	0.932
120	0.940
150	0.950
200	0.959
250	0.966
300	0.971

Table 4.15

Sampled Variance of Gaussian
Data as Function of the Number of Points

The linear intropolation of the anamorphosis function as well as the use of the five Hermite polynomials were found adequate in all the cases. Less than 5 percent errors in the fitting of the mean, $\%\mu$, and in the variance, %V, of the actual data were found in 35 of the 36 cases, see Equations (3.34) and (3.36). This is confirmed in the Table 4.14 by the performance of the KT procedure; i.e., this method gives very close results to that of the Disjunctive Kriging estimator. See Figures 4.2 d, f; 4.4 e, g; 4.5 d, f; and 4.8 g, f. In Figure 4.8 g, a typical fitted anamorphosis function is shown.

It was found that when the number of historical points increases, all the estimators perform better. As the field is more correlated (b decreases) and the dispersion σ decreases, also the estimators perform better as should be expected. Figures 4.2 to 4.8 illustrate these remarks.

Universal Kriging, Disjunctive Kriging and Kriging of Transformed data estimators gave a deviations histogram more concentrated around the mean than the corresponding Gaussian density curve; however, the less accurate local mean estimator tended to have deviation histograms of Gaussian shape. This is illustrated in Figures 4.8 h, i, j.

4.6.2 Transformed Isotropic Fields

The results of the 36 cases of transformed isotropic fields are summarized in Table 4.16. As can be seen both Disjunctive Kriging and Universal Kriging gave better results than the local mean estimator, but no apparent differences between them can be detected from the results.

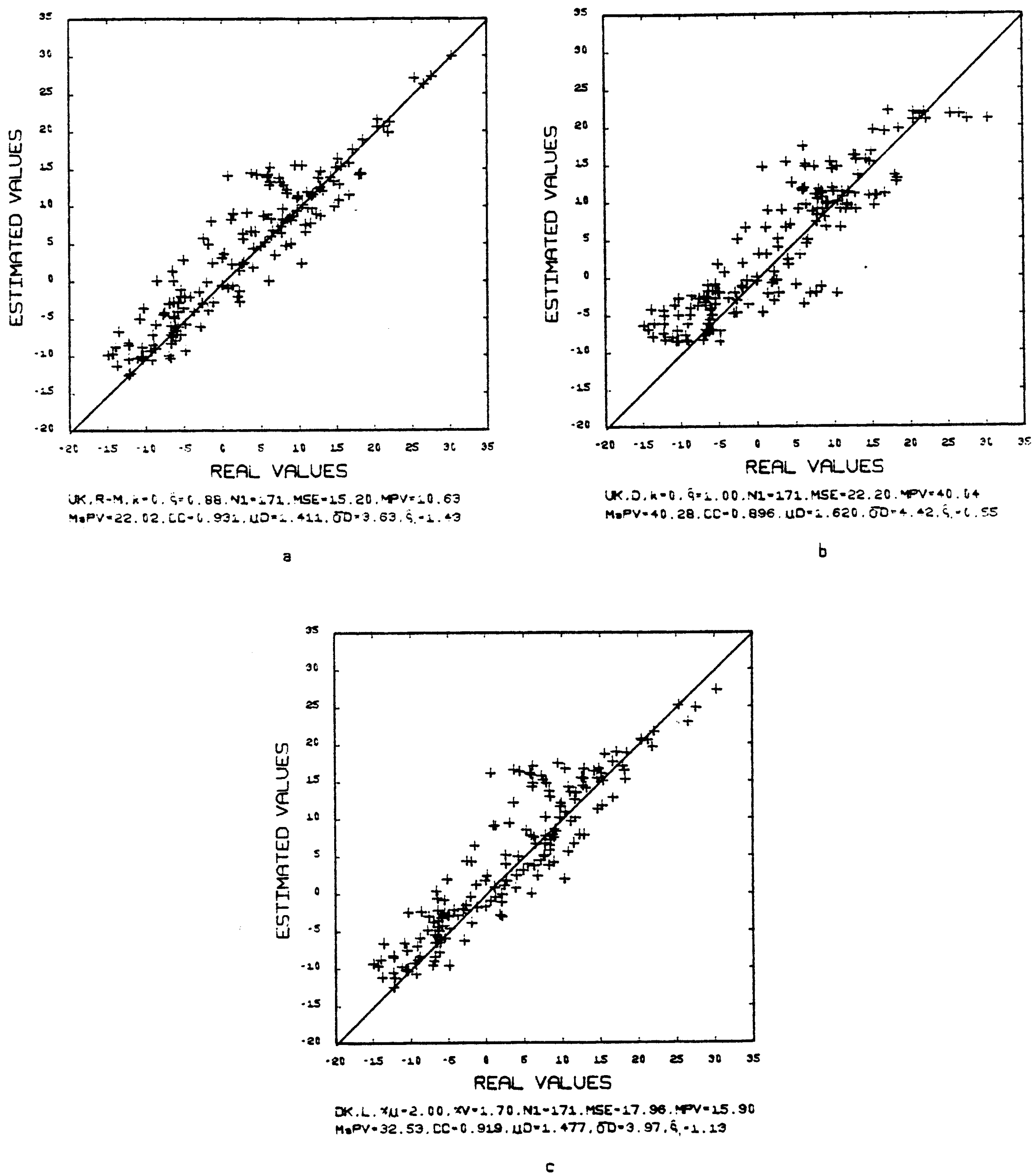
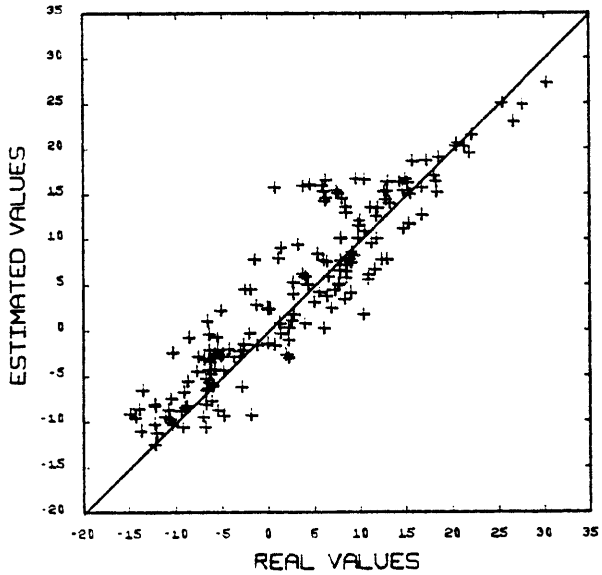


Figure 4.2

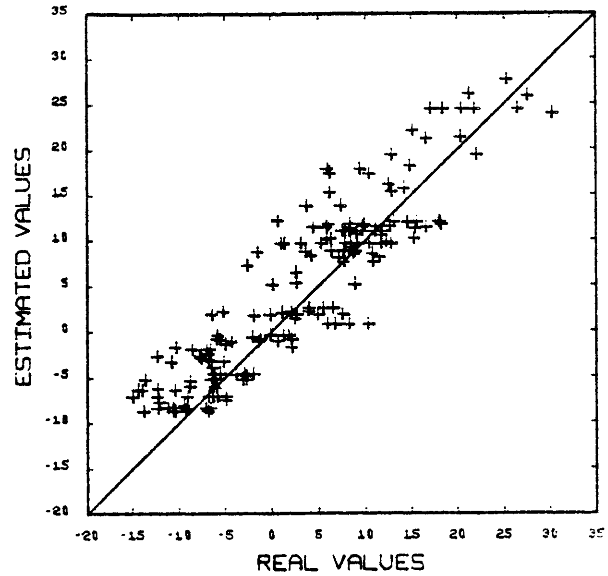
Point Estimation Comparisons. Isotropic Field

($b = 0.007$, $\sigma = 10$, $m = 10$, $N = 46$)



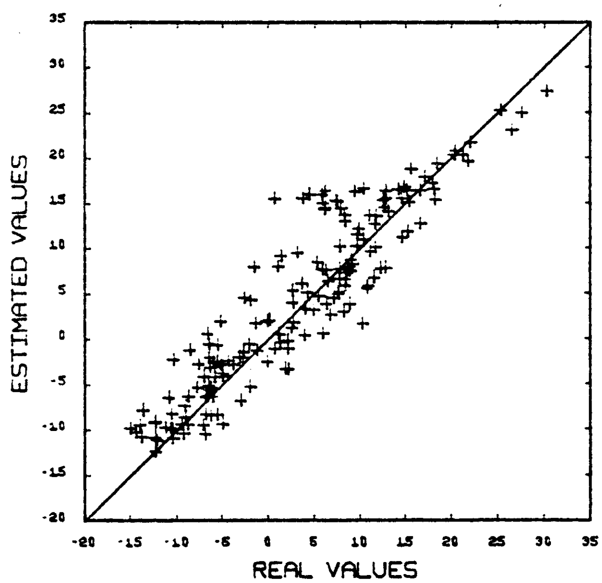
DK.S.%μ=2.00.%V=1.70.N1=171.MSE=17.95.MPV=8.505
M=PV=17.55.CC=0.917.μD=1.375.σD=4.01.δ=-2.11

d



Ms.N1=171.MSE=21.01.CC=0.906.μD=1.766.σD=4.23

e



KT.M.N1=171.MSE=16.72.CC=0.921.μD=1.183.σD=3.91

f

Figure 4.2

Point Estimation Comparisons. Isotropic Field

(b = 0.007, σ = 10, m = 10, N = 46)

(continued)

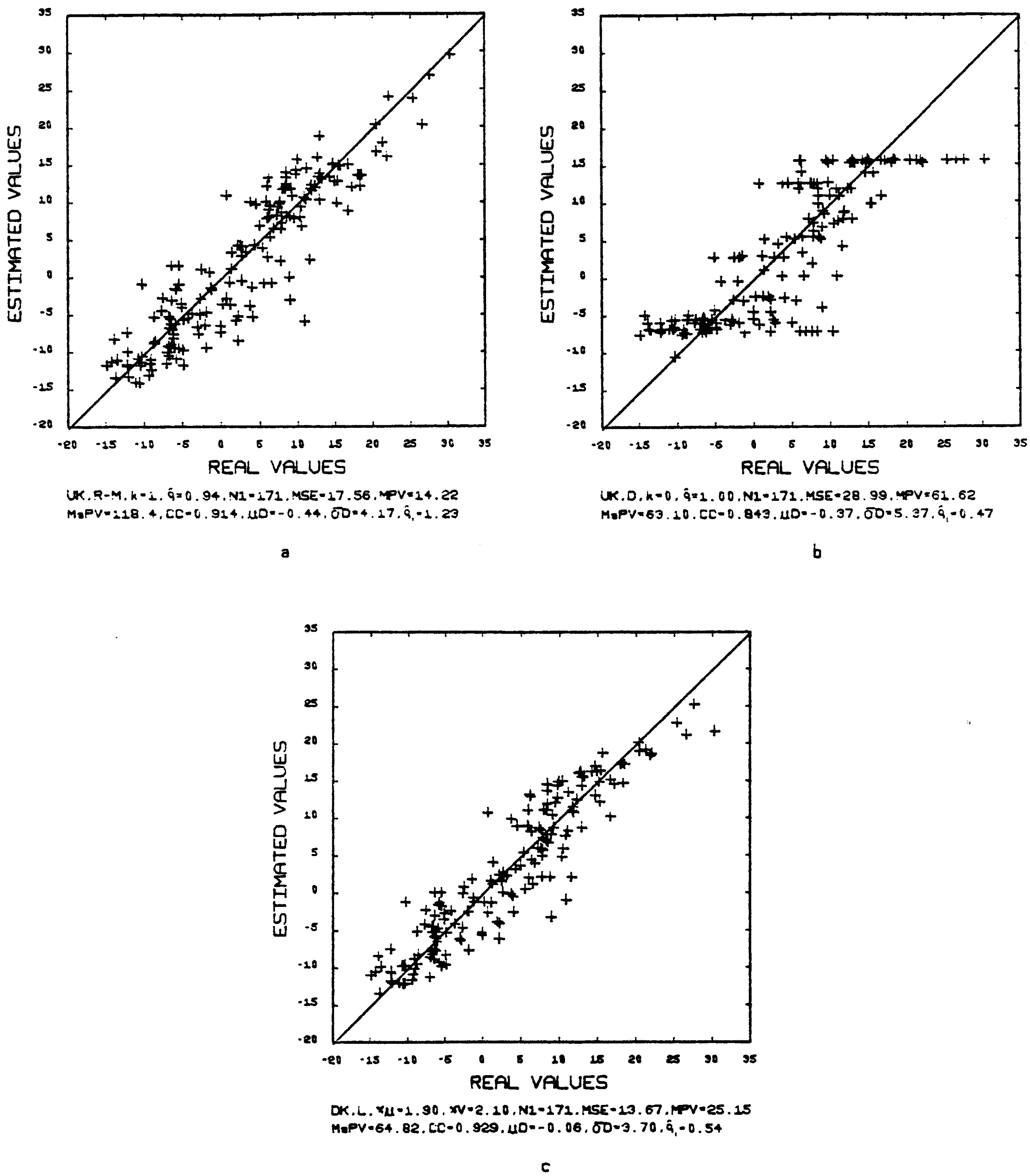
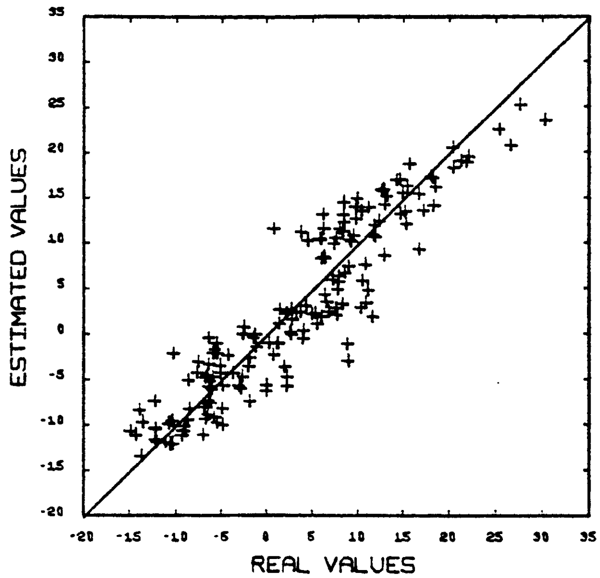


Figure 4.3

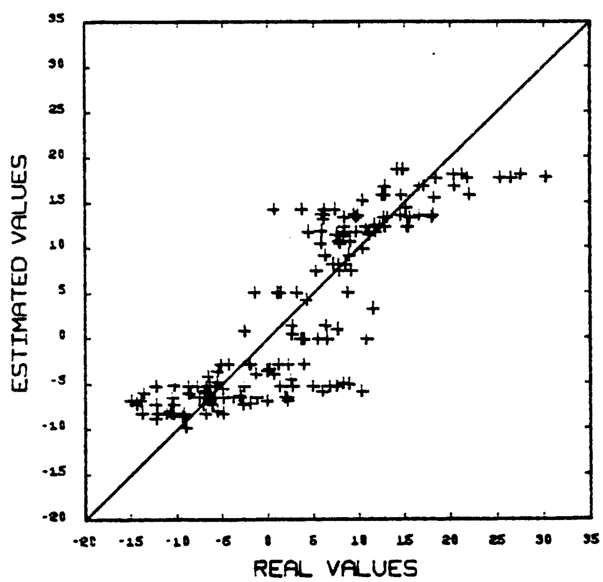
Point Estimation Comparisons. Isotropic Field

(b = 0.007, σ = 10, m = 10, N = 26)



DK.5.%μ=1.90.%V=2.10.N1=171.MSE=13.72.MPV=6.509
MePV=19.67.CC=0.929.μD=-0.19.σD=3.70.$\hat{q}_1$=2.11

d



M_5.N1=171.MSE=25.85.CC=0.865.μD=-0.42.σD=5.07

e

Figure 4.3

Point Estimation Comparisons. Isotropic Field

(b = 0.007, σ = 10, m = 10, N = 26)

(continued)

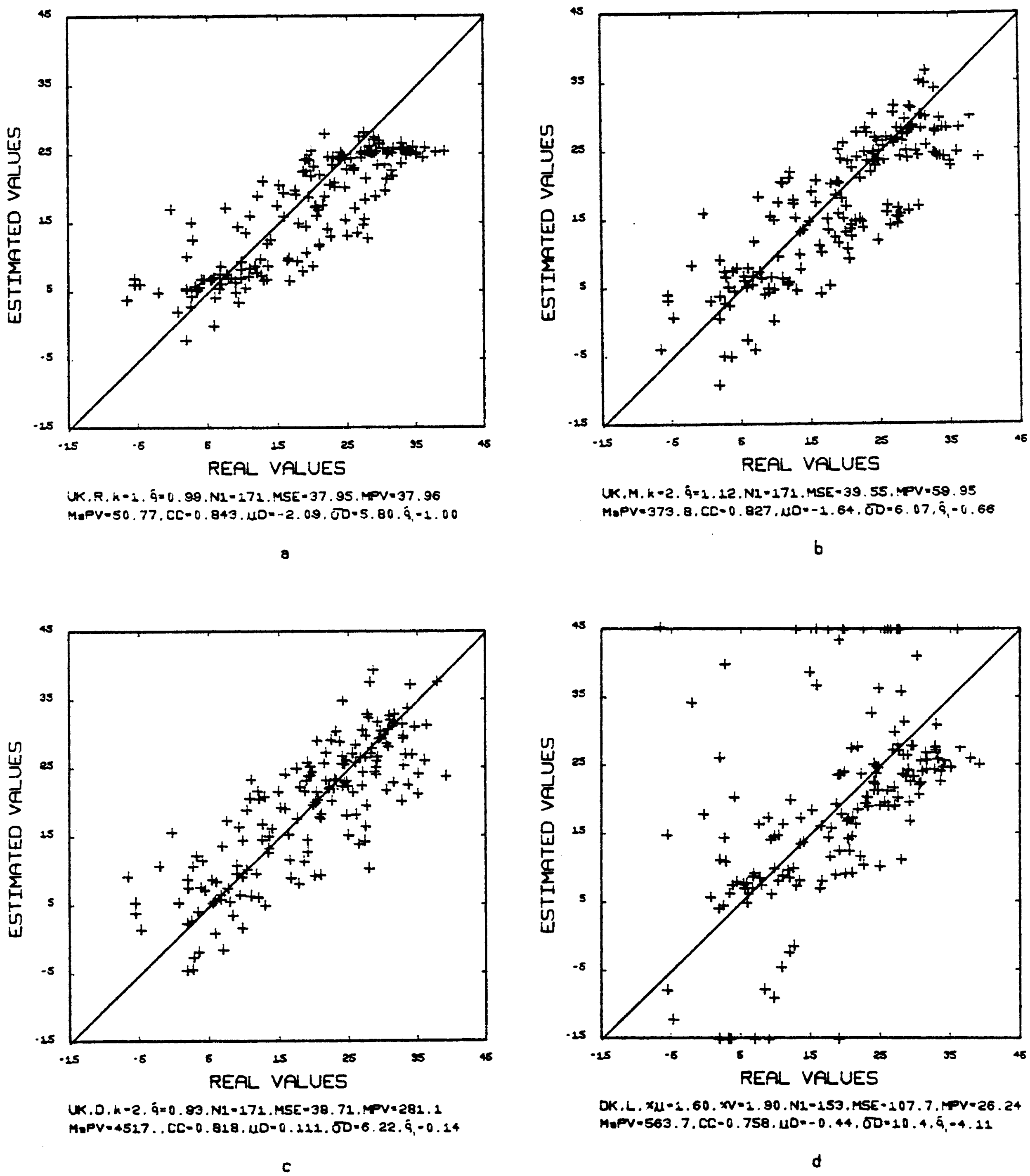
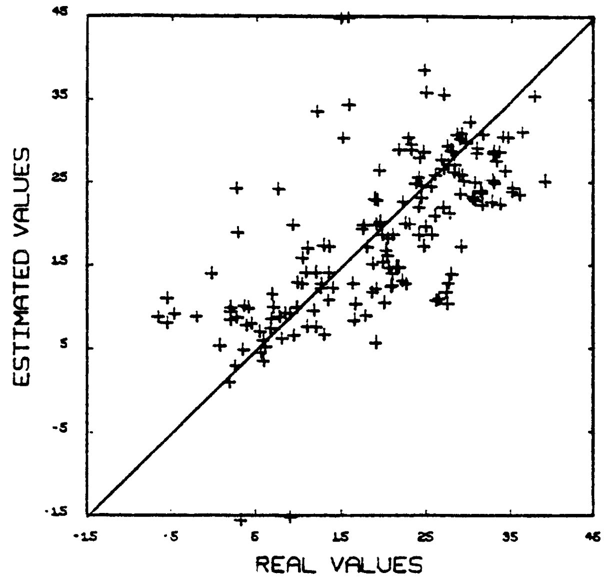


Figure 4.4

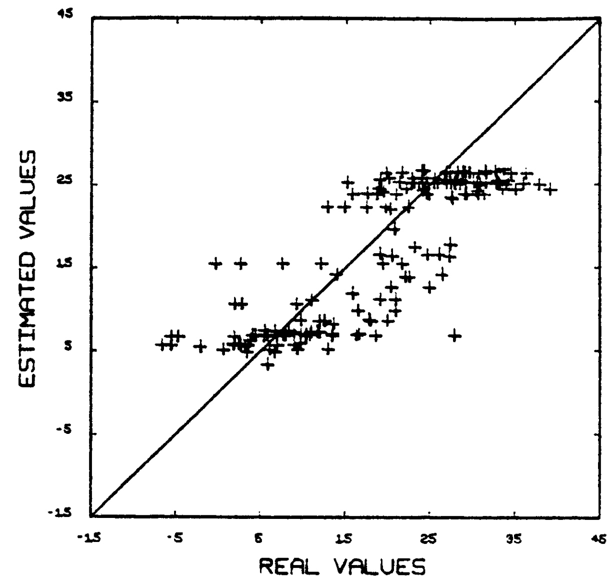
Point Estimation Comparisons. Isotropic Field

($b = 0.0141$, $\sigma = 10$, $m = 10$, $N = 24$)



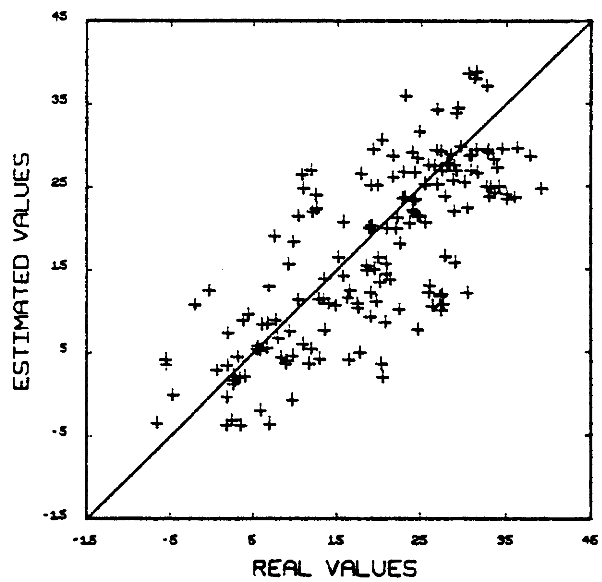
DK.S.%μ=1.60.%V=1.90.N1=169.MSE=61.79.MPV=17.12
MaPV=194.1.CC=0.713.μD=-0.51.σD=7.84.ĉ₁=3.61

e



Ms.N1=171.MSE=40.48.CC=0.822.μD=-1.91.σD=6.07

f



KT.M.N1=171.MSE=57.16.CC=0.762.μD=-1.91.σD=7.32

g

Figure 4.4

Point Estimation Comparisons. Isotropic Field

($b = 0.0141$, $\sigma = 10$, $m = 10$, $N = 24$)

(continued)

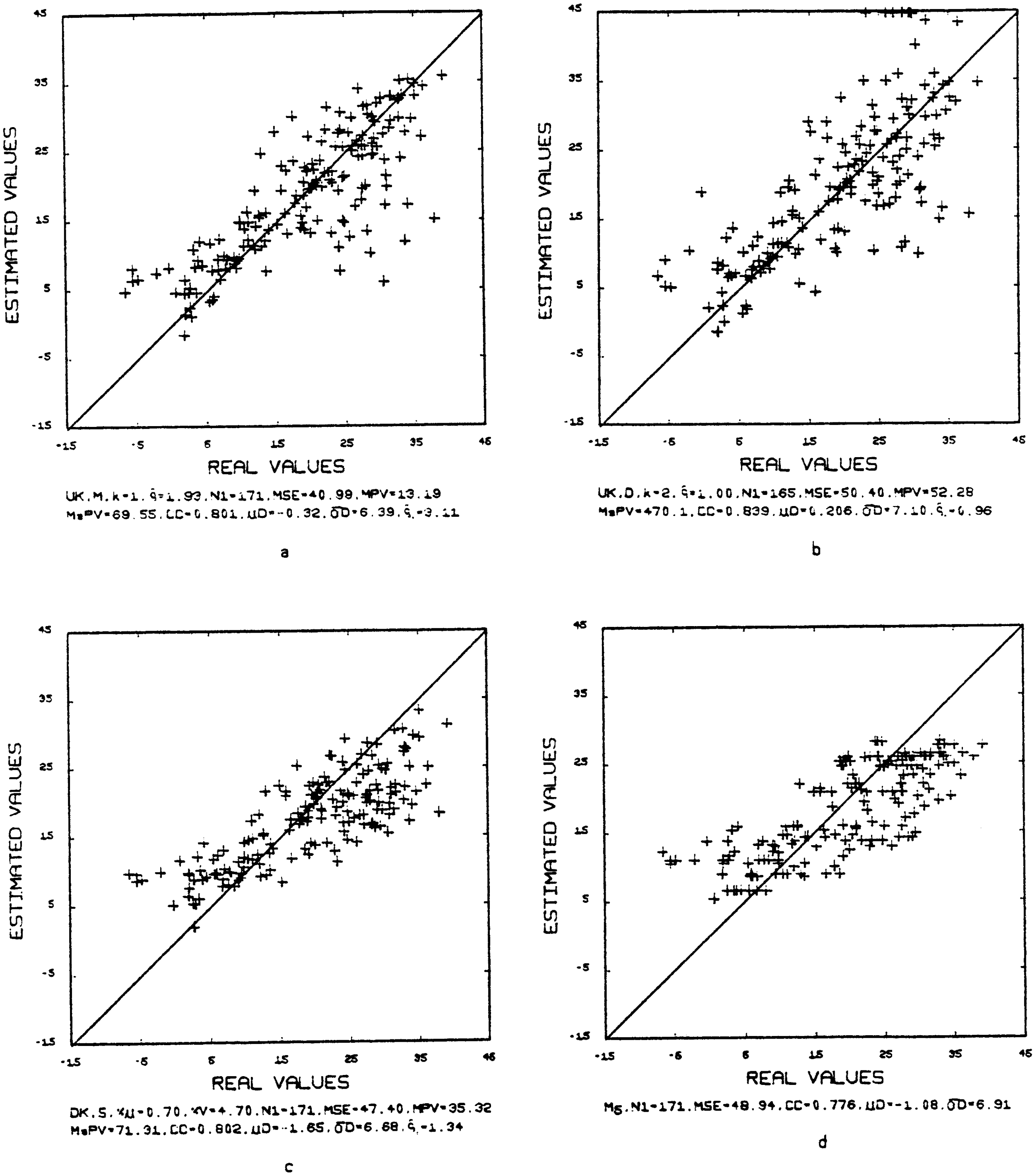
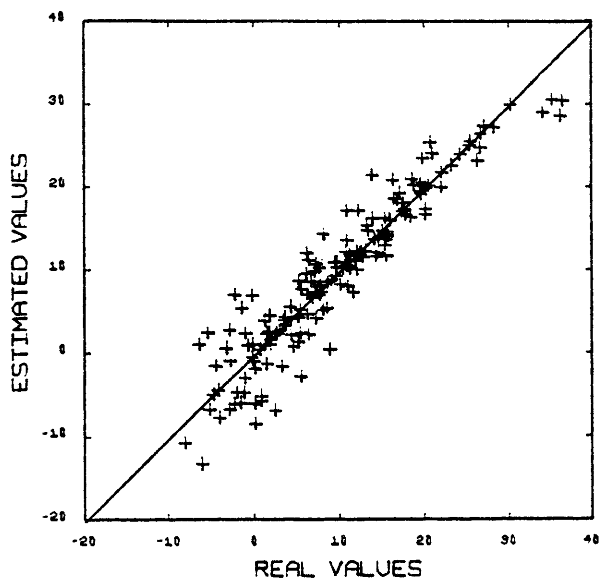


Figure 4.5

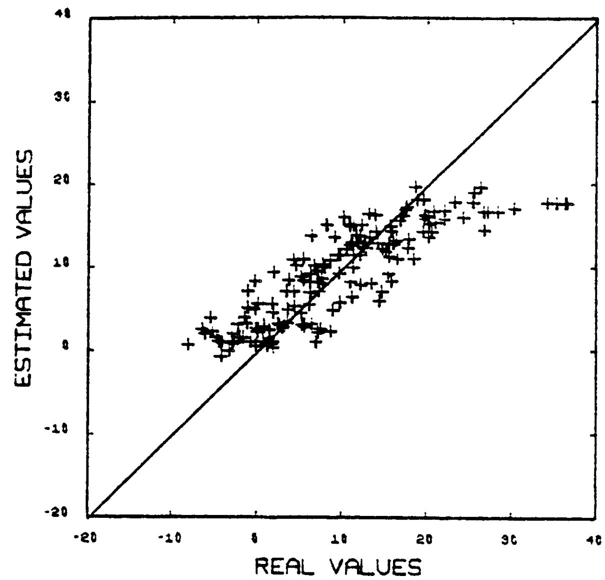
Point Estimation Comparisons. Isotropic Field

(b = 0.0141, σ = 10, m = 10, N = 29)



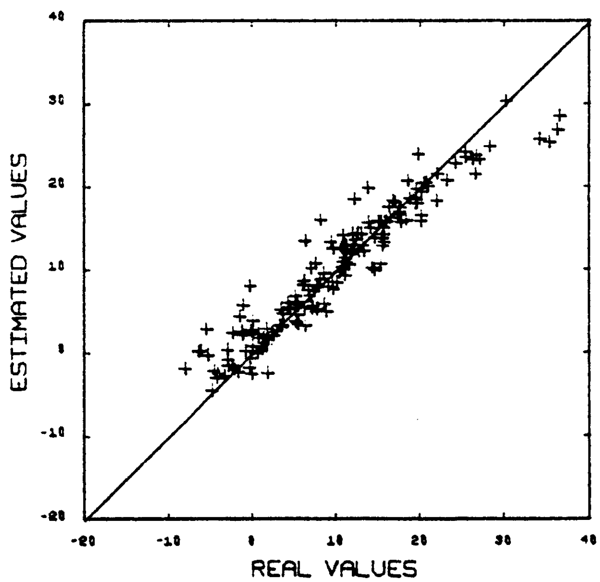
UK,R-M,k=2,q̂=2.25,N1=171,MSE=10.45,MPV=2.263
MsPV=27.66,CC=0.941,μD=-0.01,σD=3.23,q̂₁=-4.62

a



UK,D,k=0,q̂=1.00,N1=171,MSE=27.50,MPV=59.69
MsPV=60.40,CC=0.855,μD=-0.14,σD=5.24,q̂₁=0.46

b



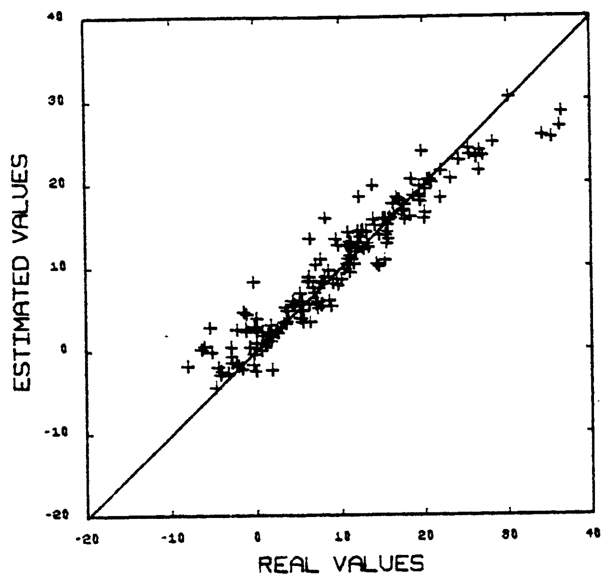
DK,L,%μ=7.50,%V=2.60,N1=171,MSE=8.539,MPV=13.04
MsPV=26.19,CC=0.954,μD=0.283,σD=2.91,q̂₁=0.66

c

Figure 4.6

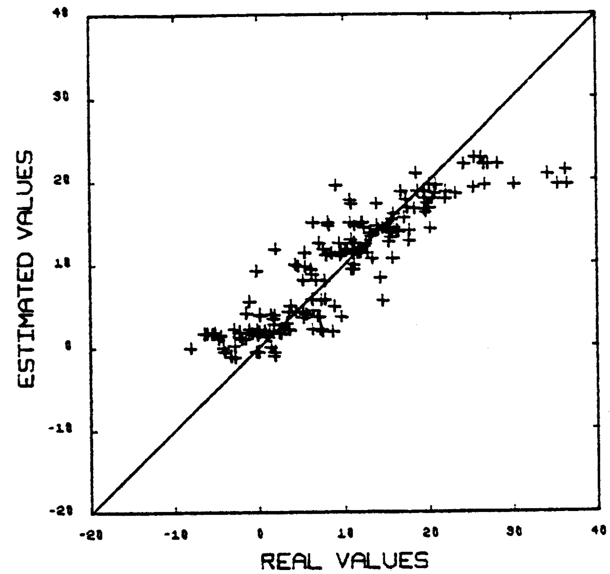
Point Estimation Comparisons. Isotropic Field

($b = 0.0035$, $\sigma = 20$, $m = 10$, $N = 47$)



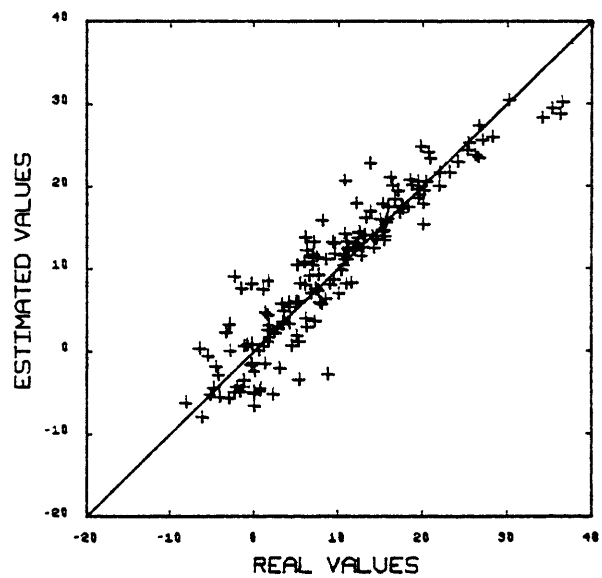
DK.S.%μ=7.50.%V=2.60.N1=171.MSE=8.367.MPV=13.08
MaPV=26.80.CC=0.955.μD=0.283.σD=2.88.$\hat{q}_1$=0.64

d



Mg.N1=171.MSE=19.75.CC=0.888.μD=0.071.σD=4.44

e



KT.M.N1=171.MSE=12.68.CC=0.926.μD=0.328.σD=3.55

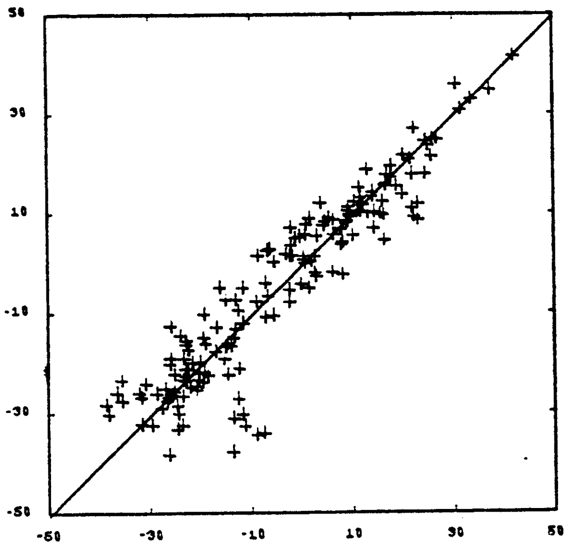
f

Figure 4.6

Point Estimation Comparisons. Isotropic Field

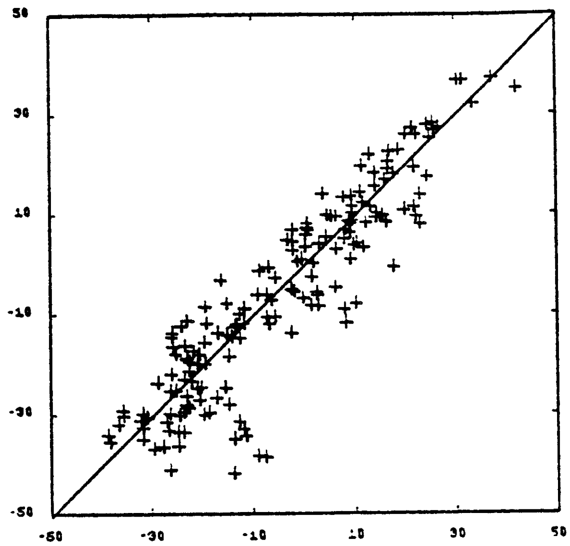
(b = 0.0035, σ = 20, m = 10, N = 47)

(continued)



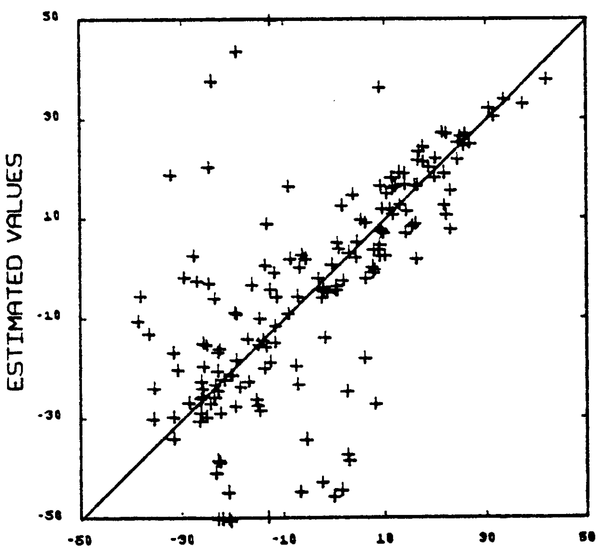
UK.R-M.k=1. $\hat{q}$=1.09.N1=171.MSE=45.03.MPV=41.19
M=PV=245.4.CC=0.933.μD=-0.09.σD=6.71.$\hat{q}_1$=1.09

a



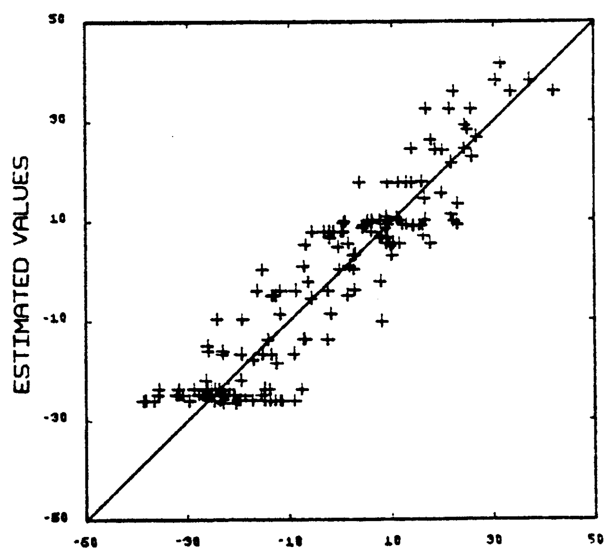
UK.D.k=1. $\hat{q}$=1.01.N1=171.MSE=68.06.MPV=79.22
M=PV=188.8.CC=0.910.μD=-1.24.σD=8.16.$\hat{q}_1$=0.86

b



OK.S.%μ=3.10.%V=0.30.N1=167.MSE=256.3.MPV=80.46
M=PV=294.3.CC=0.710.μD=0.179.σD=16.0.$\hat{q}_1$=3.19

c



Ms.N1=171.MSE=54.33.CC=0.923.μD=0.655.σD=7.34

d

Figure 4.7

Point Estimation Comparisons. Isotropic Field

(b = 0.007, σ = 20, m = 10, N = 29)

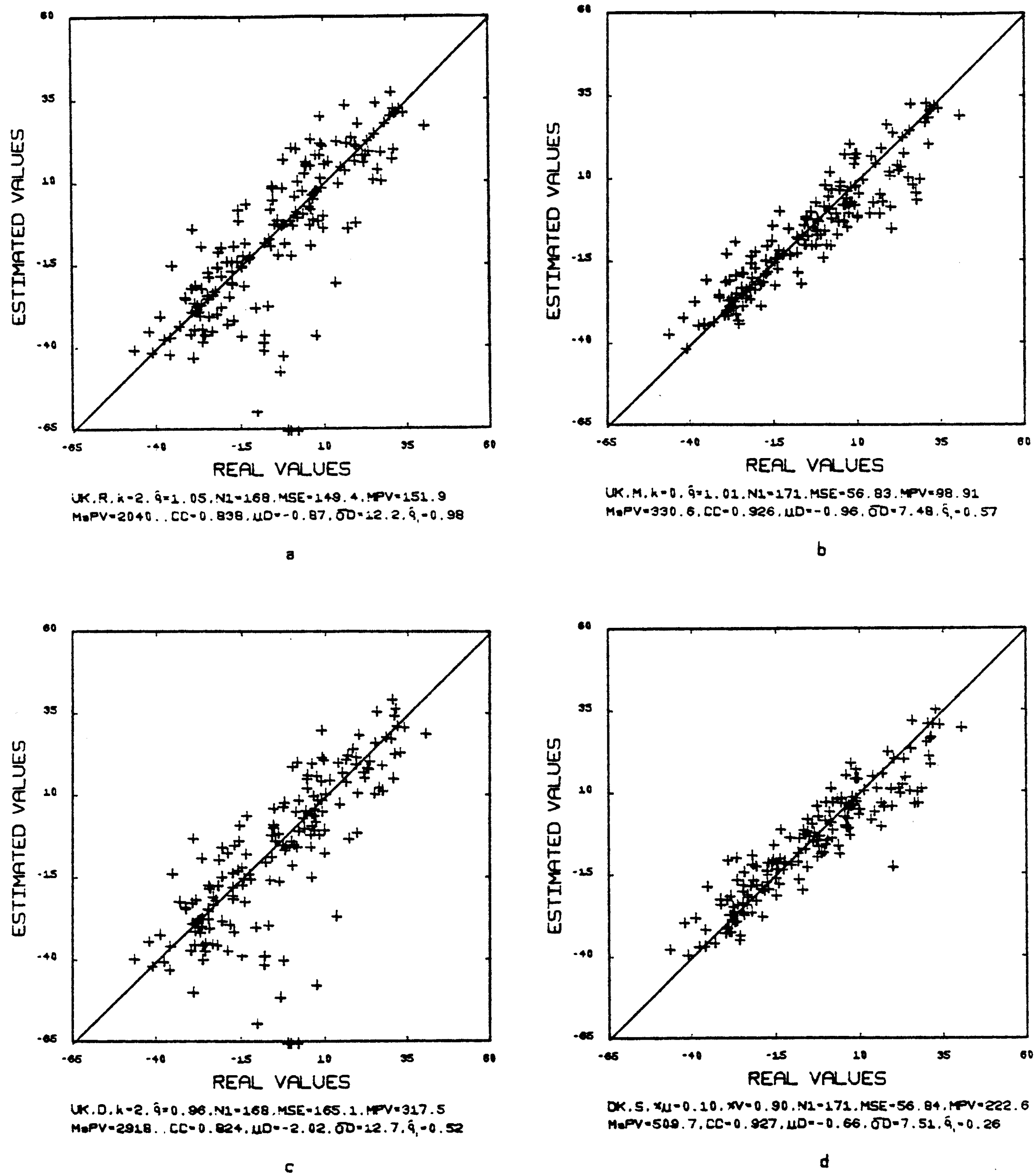
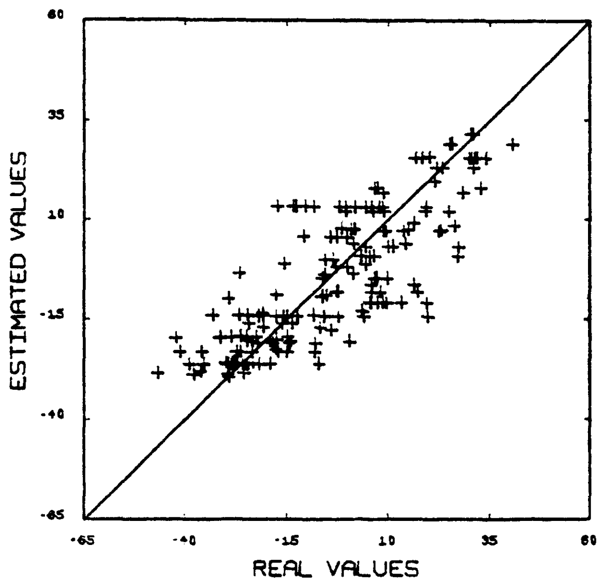


Figure 4.8

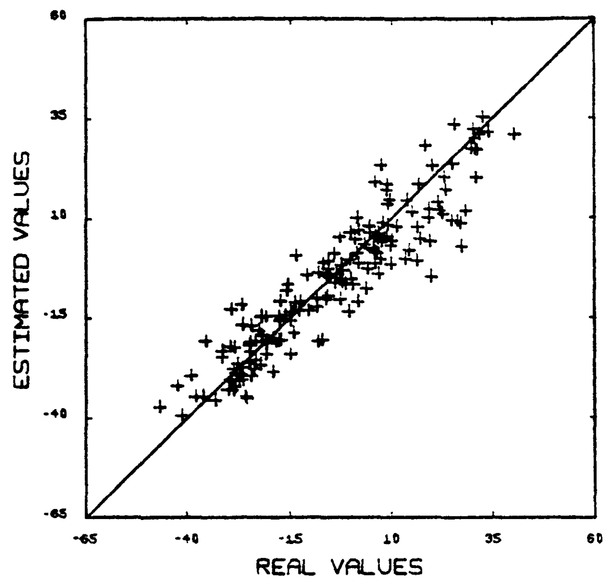
Point Estimation Comparisons. Isotropic Field

($b = 0.0141$, $\sigma = 20$, $m = 10$, $N = 28$)



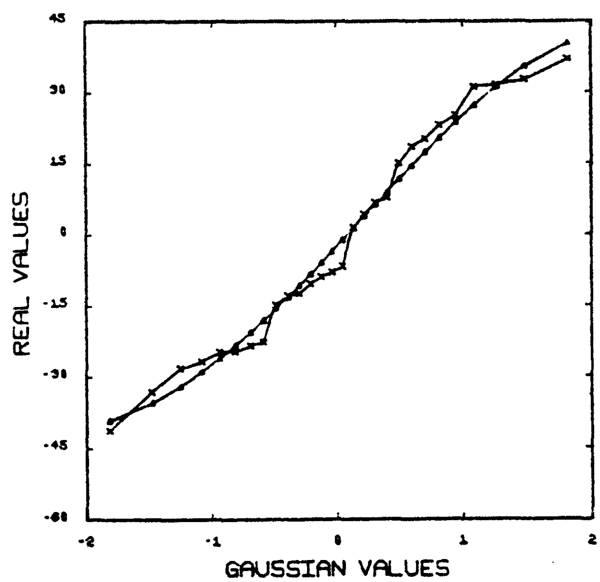
Ms.N1=171.MSE=133.8.CC=0.813.μD=-0.68.σD=11.5

e



KT.M.N1=171.MSE=56.17.CC=0.928.μD=-1.36.σD=7.37

f



ANAMORPHOSIS FUNCTION. X - SAMPLED. ▲ - FITTED

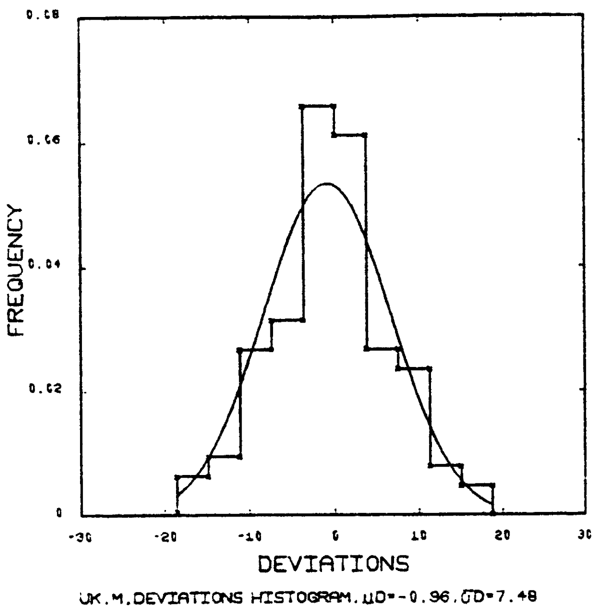
g

Figure 4.8

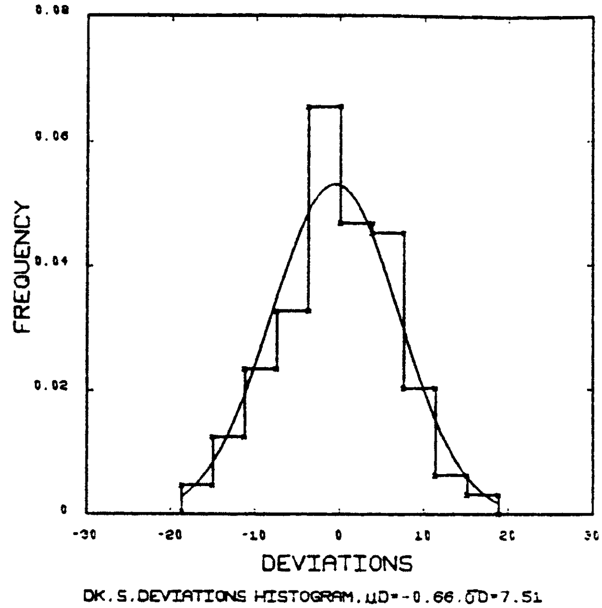
Point Estimation Comparisons. Isotropic Field

(b = 0.0141, σ = 20, m = 10, N = 28)

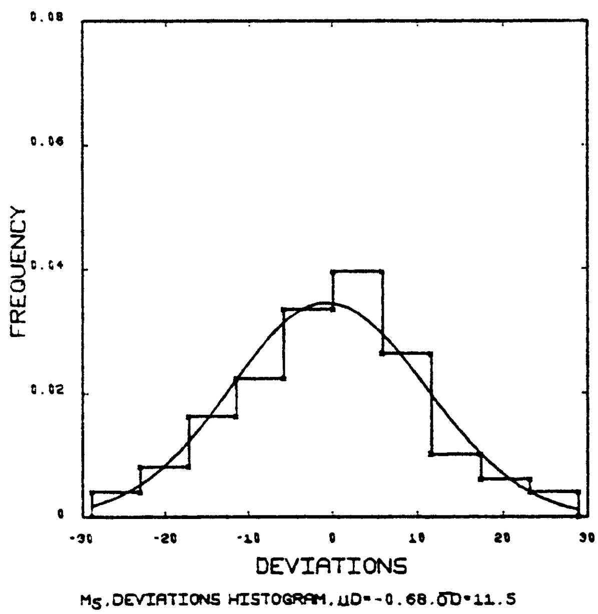
(continued)



h



i



j

Figure 4.8

Point Estimation Comparisons. Isotropic Field

($b = 0.0141$, $\sigma = 20$, $m = 10$, $N = 28$)

(continued)

	MSE			MPV		
	(10%)	(25%)	(50%)	(10%)	(25%)	(50%)
K_M	24	26	31	22	26	26
DK_s	24	32	33	20	21	23
M_5	2	5	15			

CONSISTENCY

$$(\hat{\rho}_1 \doteq \frac{MSE}{MPV})$$

	Min	Max	$\# \leq 1$	$\# > 1$	# [0.8,1.2]	μ	σ
K_M	0.37	14.5	11	25	6	2.23	2.07
DK_s	0.21	12.16	13	23	5	1.8	1.6

Table 4.16

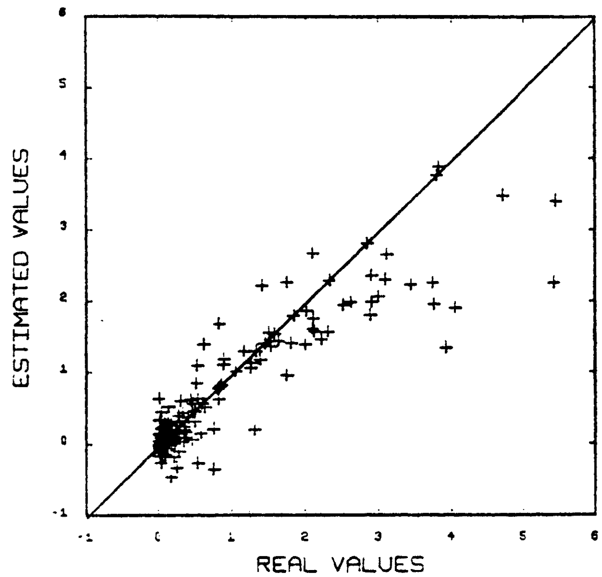
Comparisons of the Different Estimators
Transformed Isotropic Fields (36 cases) Point Results

Both estimators tended to underestimate the variance of estimation as expressed by 25 (for K_M) and 23 (for DK_s) of the 36 cases with consistency parameter greater than one. Once again the Universal Kriging estimator gave as good or better results than Disjunctive Kriging.

Figures 4.9 a, d, e; 4.10 b, e, f; 4.11 a, d, e; and 4.12 b, d, e give a visual image of how the different estimators work with the transformed isotropic fields as the reality. Because these fields have more local variability than the isotropic fields, it is not expected to obtain very good performance of the estimators at the extreme values. In fact, the figures show that the estimators did not perform as well as with the isotropic fields. Also, the anamorphosis fit was less adequate than in the isotropic field cases, with errors in the mean and variance of the data that were less than 11 percent; see Figure 4.12 f.

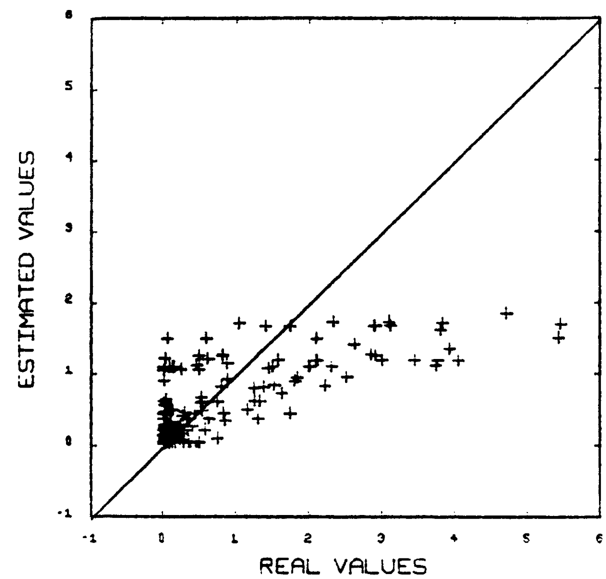
4.6.3 Intrinsic Random Functions

Table 4.17 summarizes the results found with the 12 cases of IRF of order 0. The Disjunctive Kriging and Universal Kriging estimators performance was about the same with respect to MSE, but Disjunctive Kriging gave better results with respect to MPV. The consistency parameter, however, indicates better behavior of the Universal Kriging estimator, i.e., Disjunctive Kriging underestimated values of the predicted variance more often. Figures 4.13 b, e, f and 4.14 a, c, d show how similar the performance of the Kriging estimators was, and the advantange the two of them have over the local mean estimator.



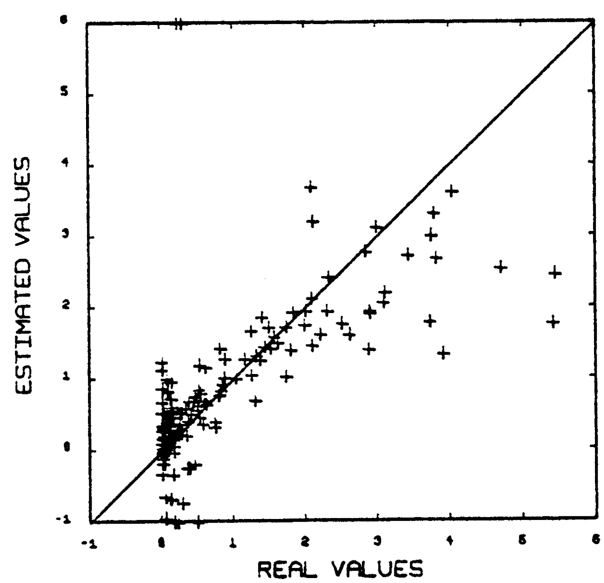
UK.R-M.k=1.$\hat{q}$=2.40.N1=171.MSE=0.293.MPV=0.019
MaPV=0.100.CC=0.907.μD=-0.14.$\bar{\sigma}$D=0.52.$\hat{q}_1$=15.4

a



UK.D.k=0.$\hat{q}$=1.00.N1=171.MSE=0.783.MPV=0.707
MaPV=0.707.CC=0.734.μD=-0.16.$\bar{\sigma}$D=0.87.$\hat{q}_1$=1.11

b



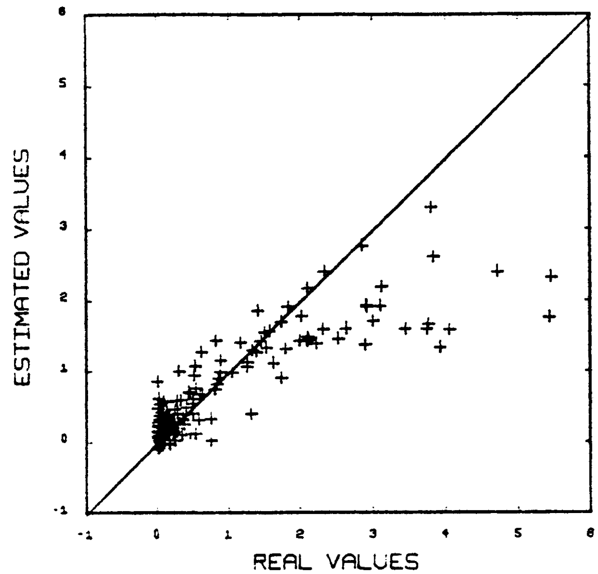
DK.L.$\kappa\mu$=3.90.κV=7.30.N1=166.MSE=0.422.MPV=0.236
MaPV=0.732.CC=0.828.μD=-0.08.$\bar{\sigma}$D=0.64.$\hat{q}_1$=1.79

c

Figure 4.9

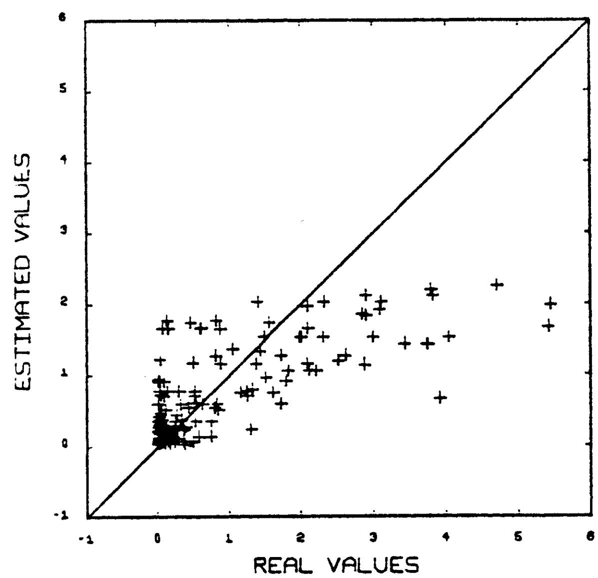
Point Estimation Comparisons. Transformed Isotropic Field

(b = 0.0035, σ = 10, N = 47)



DK.S.%μ=3.90.%V=7.30.N1=171.MSE=0.450.MPV=0.334
MaPV=0.618.CC=0.873.μD=-0.14.σD=0.66.$\hat{q}$=1.35

d



M_S.N1=171.MSE=0.696.CC=0.743.μD=-0.14.σD=0.82

e

Figure 4.9
Point Estimation Comparisons. Transformed Isotropic Field
(b = 0.0035, σ = 10, N = 47)
(continued)

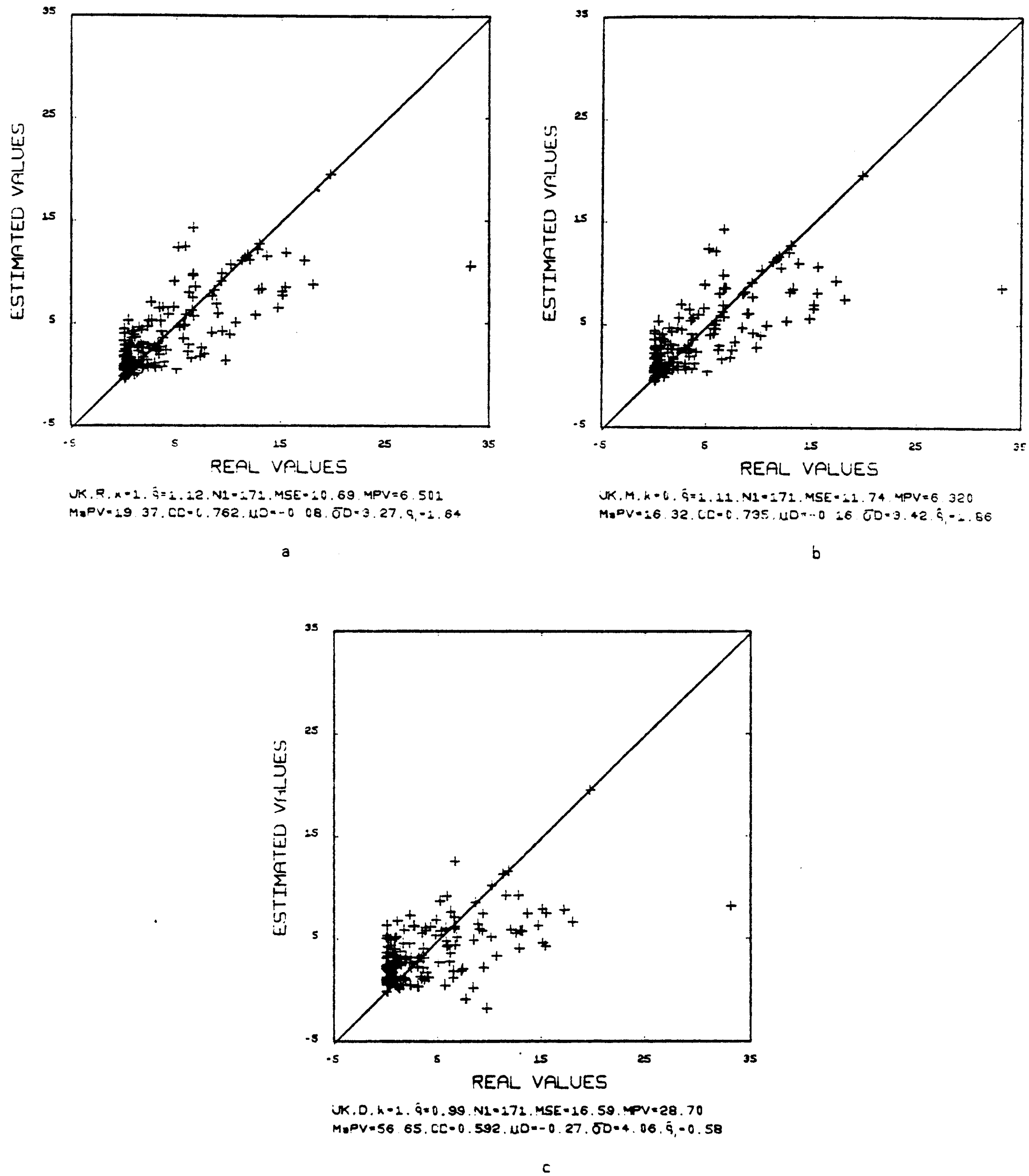
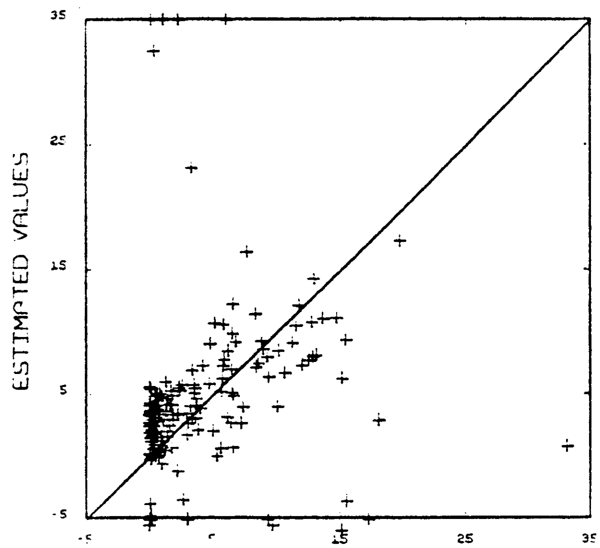


Figure 4.10

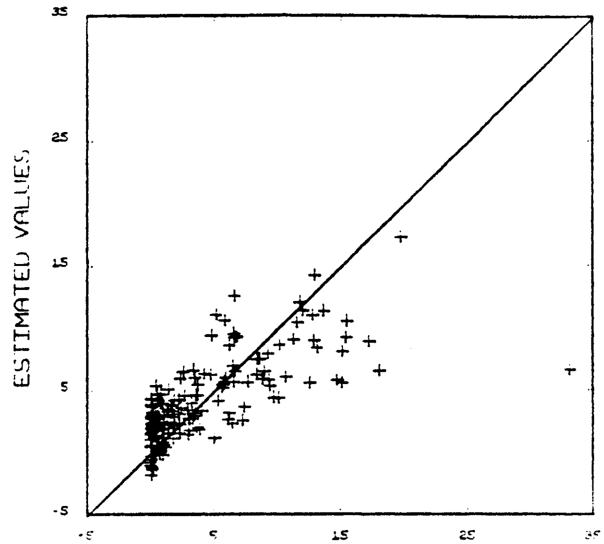
Point Estimation Comparisons. Transformed Isotropic Field

($b = 0.0141$, $\sigma = 10$, $N = 29$)



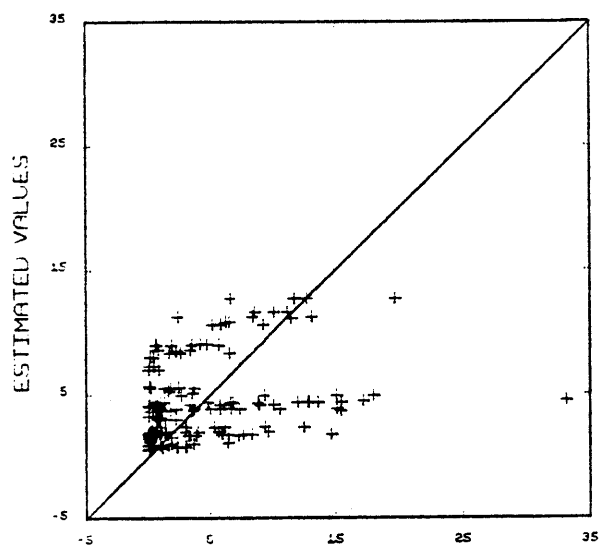
DK.L.%μ=4.50.%V=5.00.N1=159.MSE=32.22.MPV=15.46
MsPV=42.08.CC=0.215.μD=0.320.σD=5.67.ρ=2.08

d



DK.S.%μ=4.50.%V=5.00.N1=171.MSE=12.32.MPV=12.19
MsPV=24.21.CC=0.722.μD=0.096 σD=3.51.ρ=1.12

e



Ms.N1=171.MSE=24.88.CC=0.335.μD=0.179.σD=4.99

f

Figure 4.10
Point Estimation Comparisons. Transformed Isotropic Field
(b = 0.0141, σ = 10, N = 29)
(continued)

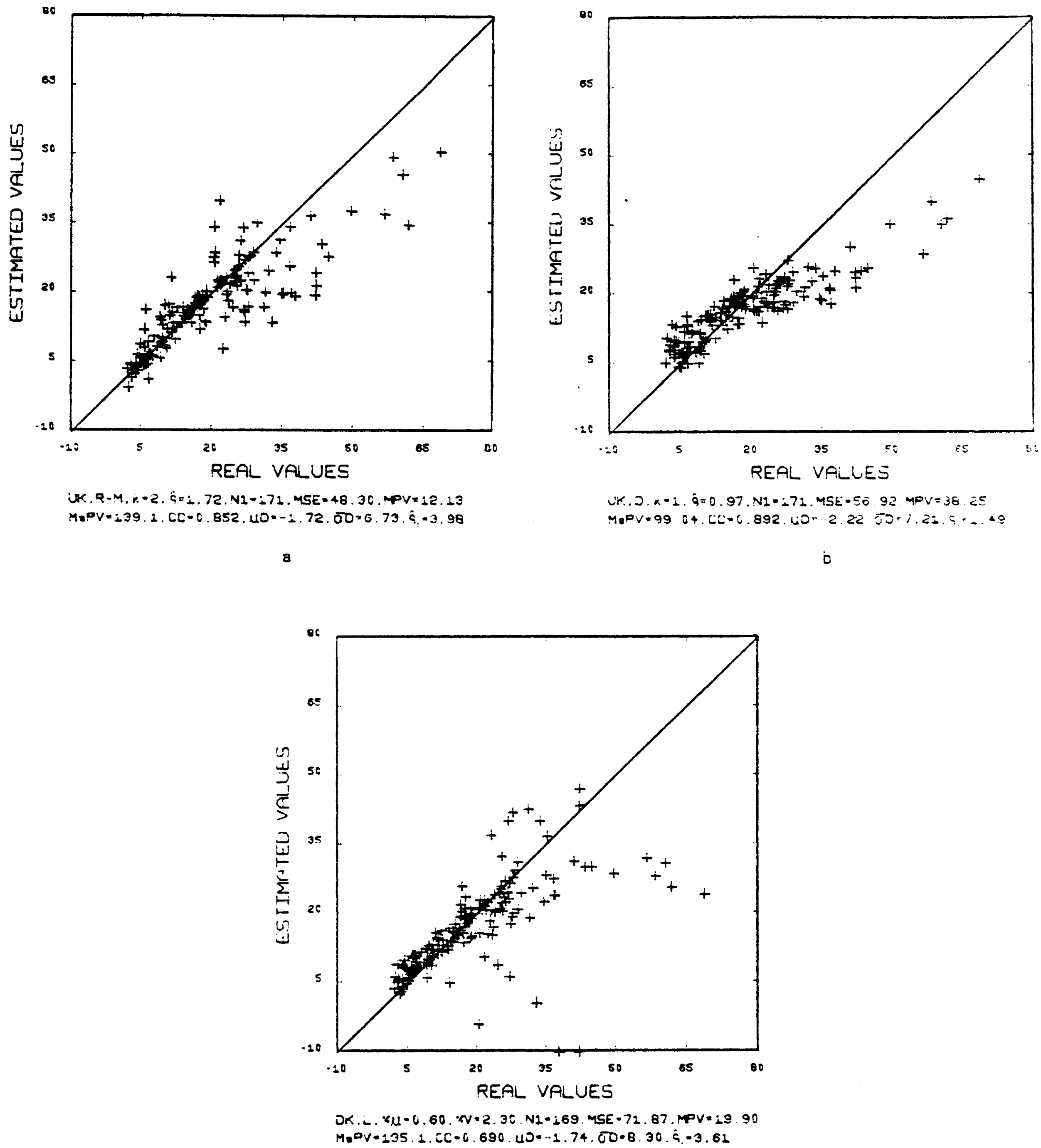
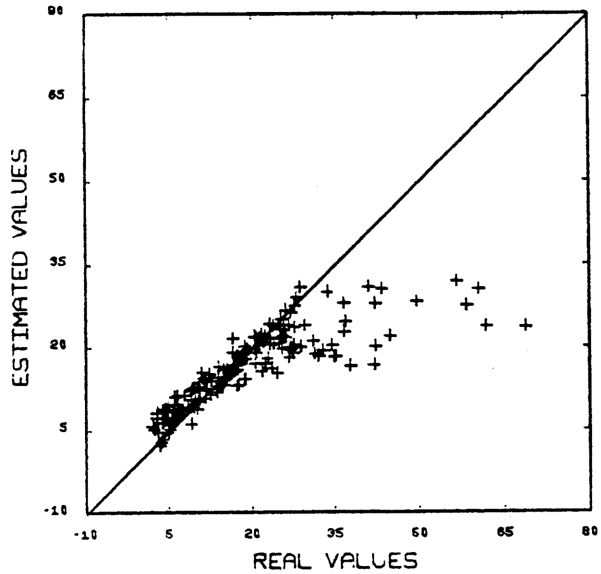


Figure 4.11

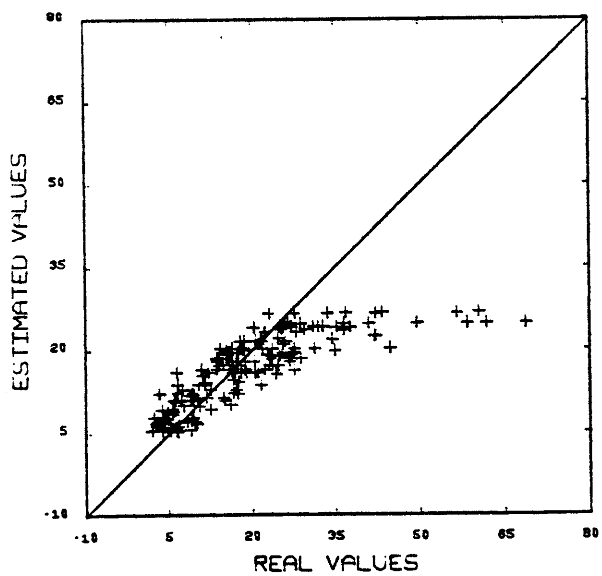
Point Estimation Comparisons. Transformed Isotropic Field

(b = 0.0035, σ = 20, N = 42)



DK,S,%μ=0.60,%V=2.30,N1=171,MSE=70.13,MPV=22.39
MaPV=52.42,CC=0.836,μD=-2.69,σD=7.93,$\hat{q}_1$=3.13

d



M_S,N1=171,MSE=77.06,CC=0.816,μD=-2.72,σD=8.35

e

Figure 4.11

Point Estimation Comparisons. Transformed Isotropic Field

(b = 0.0035, σ = 20, N = 42)

(continued)

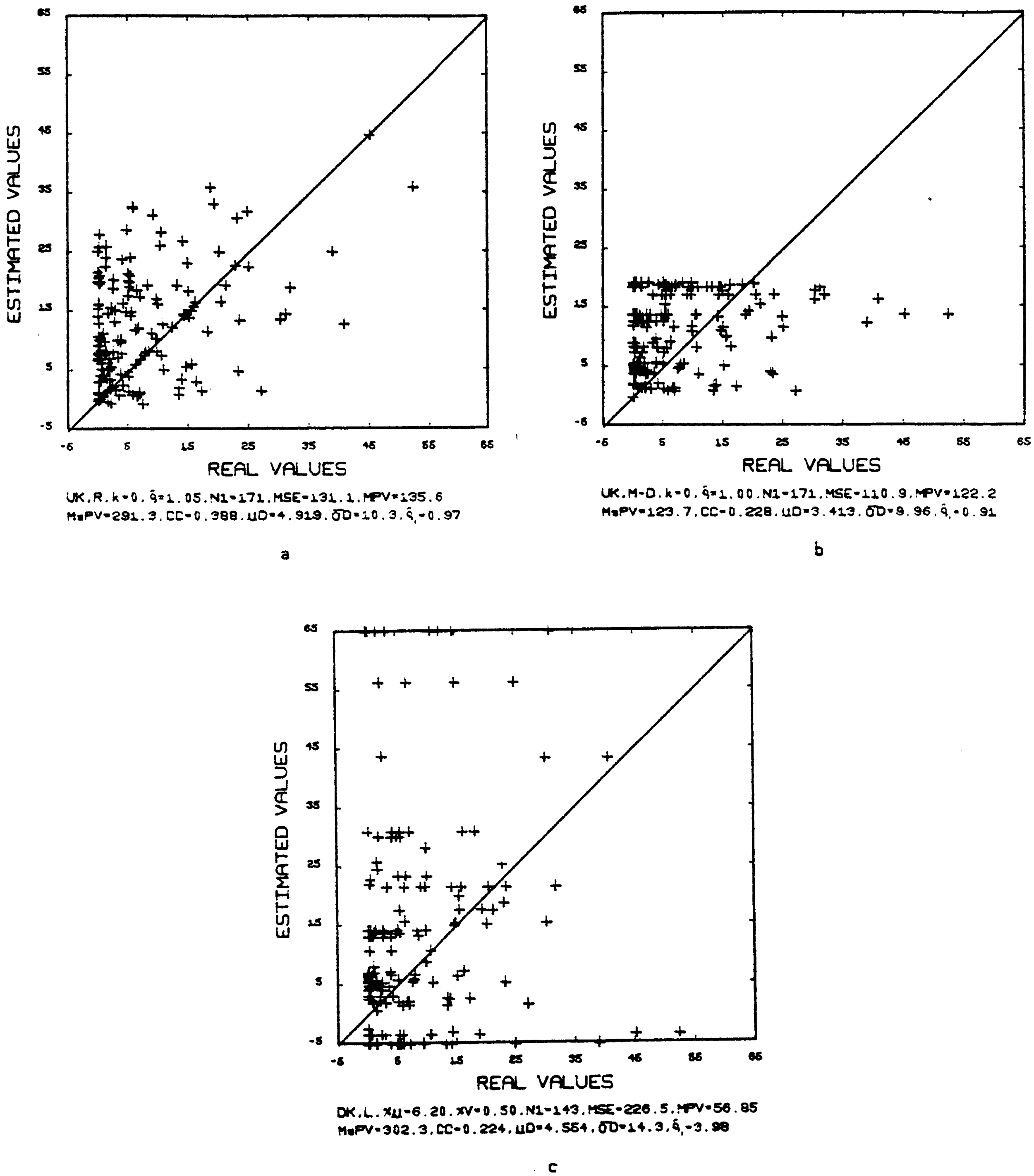
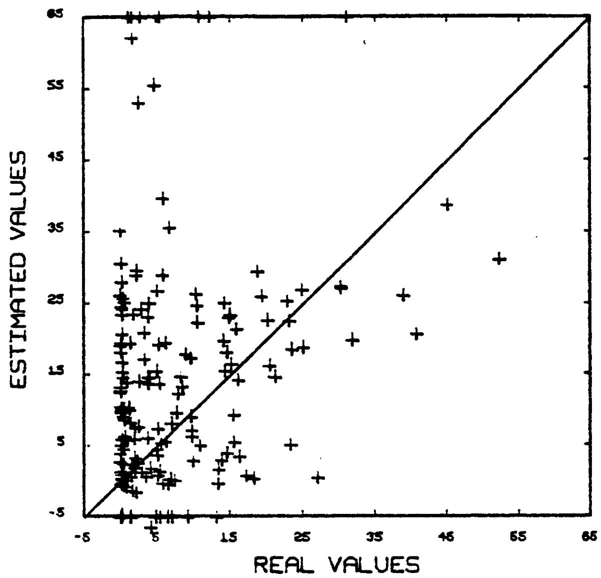


Figure 4.12

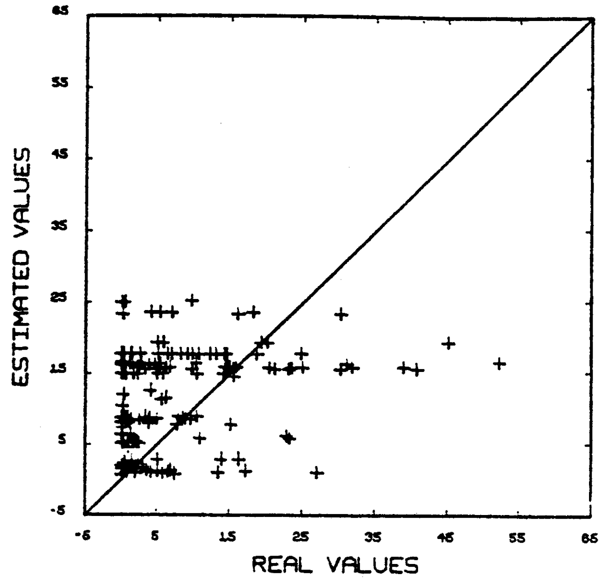
Point Estimation Comparisons. Transformed Isotropic Field

(b = 0.0141, σ = 20, N = 30)



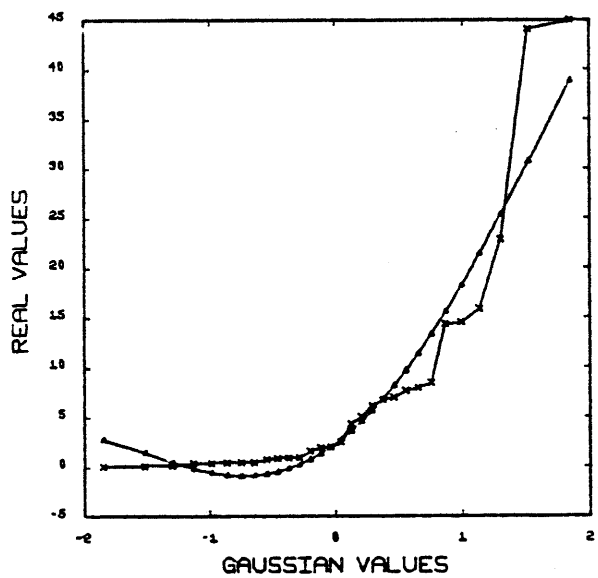
DK,S,%μ=6.20,%V=0.50,N1=153,MSE=215.3,MPV=17.93
MaPV=150.0,CC=0.282,μD=5.878,$\bar{\sigma}$D=13.4,$\hat{q}_1$=12.0

d



M5,N1=171,MSE=123.2,CC=0.235,μD=4.370,$\bar{\sigma}$D=10.2

e



ANAMORPHOSIS FUNCTION, X - SAMPLED, Δ - FITTED

f

Figure 4.12

Point Estimation Comparisons. Transformed Isotropic Field

(b = 0.0141, σ = 20, N = 30)

(continued)

	MSE			MPV		
	(10%)	(25%)	(50%)	(10%)	(25%)	(50%)
K_M	9	10	11	6	7	7
DK_s	9	11	12	9	11	12
M_5	0	2	8			

CONSISTENCY

$$(\hat{\rho}_1 = \frac{MSE}{MPV})$$

	Min	Max	$\# \le 1$	$\# > 1$	# [0.8,1.2]	μ	σ
K_M	0.62	2.39	4	8	5	1.07	0.57
DK_s	0.77	3.18	4	8	3	1.53	0.87

Table 4.17

Comparisons of the Different Estimators

Intrinsic Random Functions of Order 0 (12 cases) Point Results

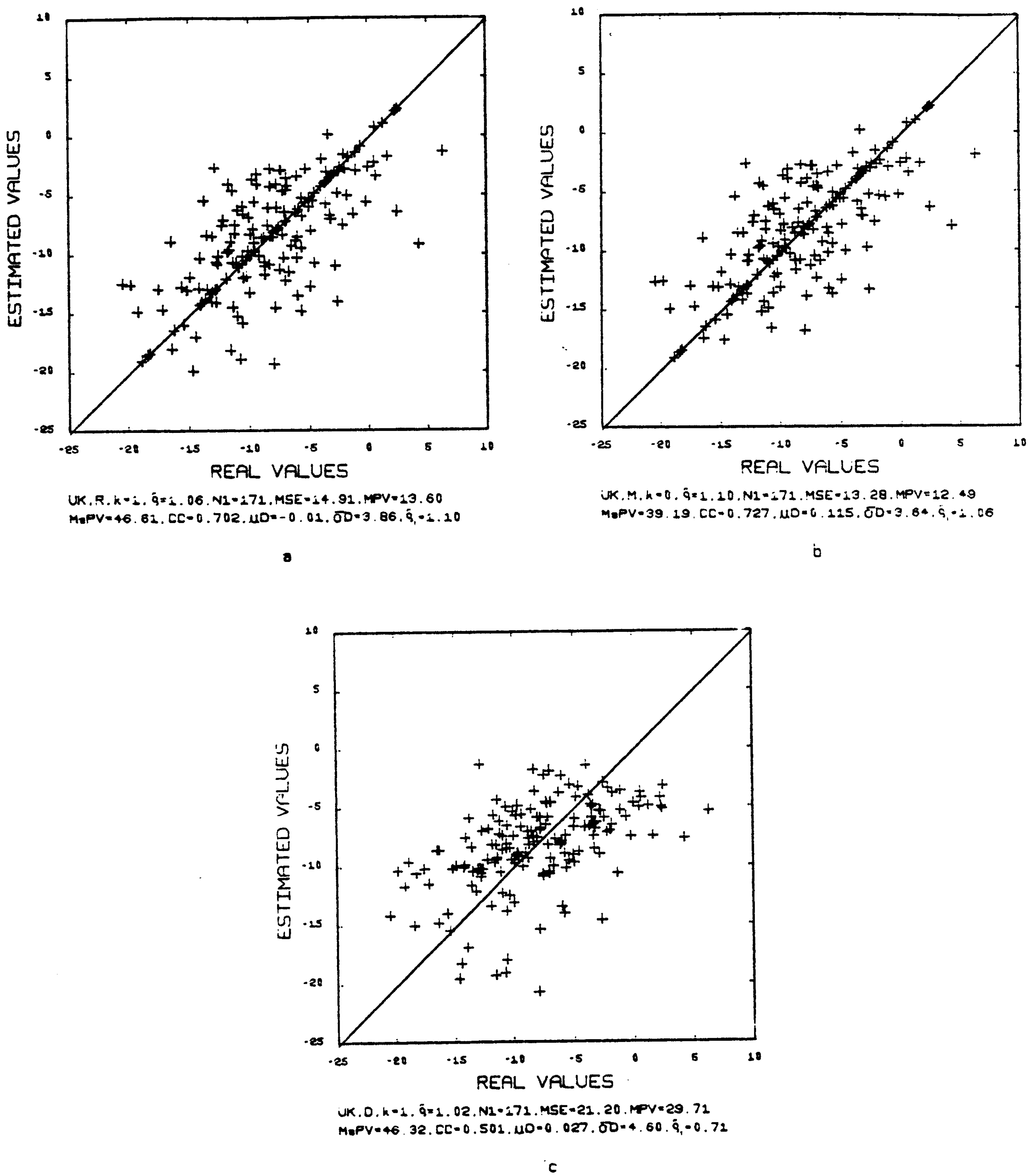
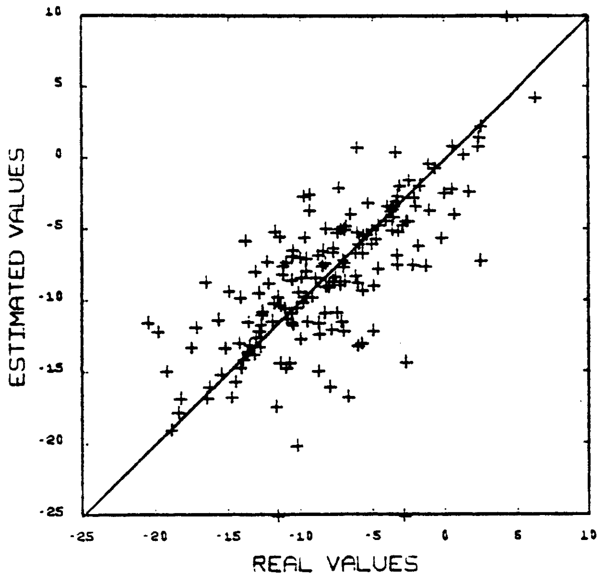


Figure 4.13

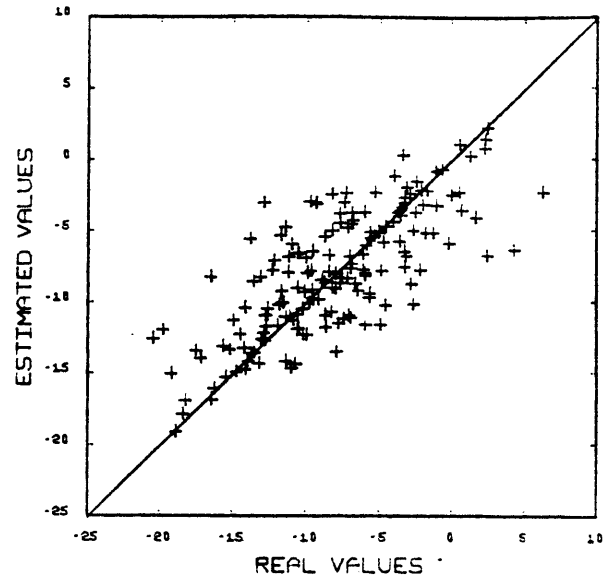
Point Estimation Comparisons. Intrinsic Random Function

($k = 0$, $\alpha_1 = -1$, $\alpha_3 = \alpha_5 = 0$, $N = 48$)



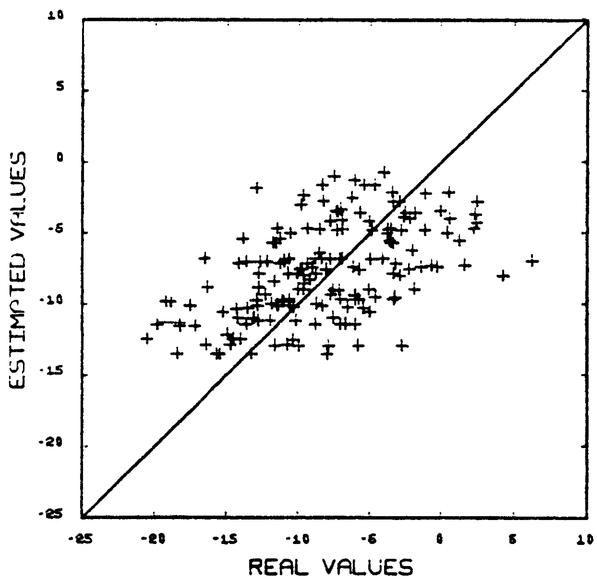
DK.L.%μ=0.20.%V=1.90.N1=168.MSE=12.66.MPV=14.76
MsPV=230.9.CC=0.786.μD=-0.03.σD=3.56.q̂=0.86

d



DK.S.%μ=0.20.%V=1.90.N1=171.MSE=11.78.MPV=11.59
MsPV=31.88.CC=0.752.μD=0.267.σD=3.42.q̂=1.02

e



Ms.N1=171.MSE=20.24.CC=0.510.μD=0.410.σD=4.48

f

Figure 4.13

Point Estimation Comparisons. Intrinsic Random Function

($k = 0$, $\alpha_1 = -1$, $\alpha_3 = \alpha_5 = 0$, $N = 48$)

(continued)

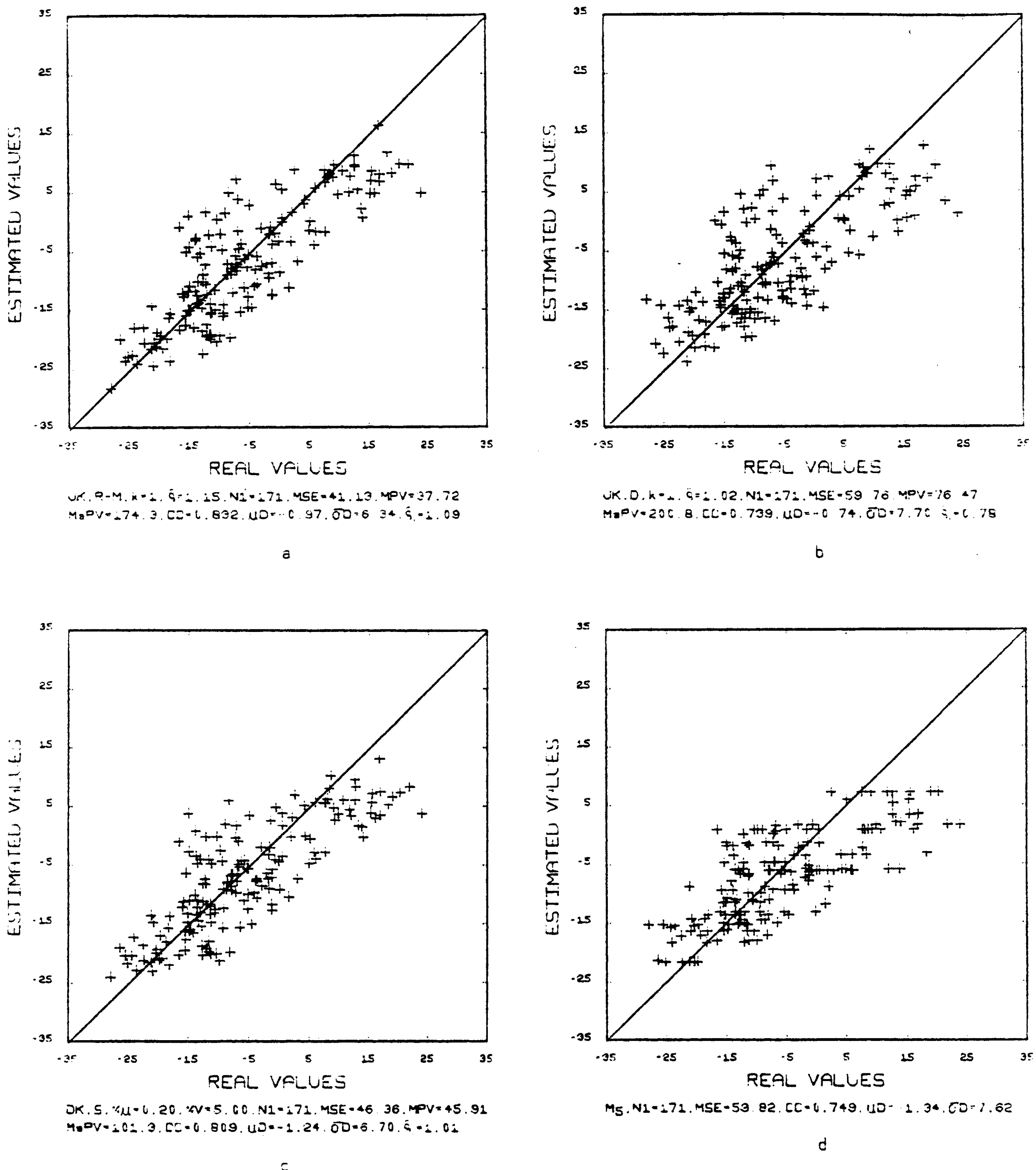


Figure 4.14

Point Estimation Comparisons. Intrinsic Random Function

$(k = 0,\ \alpha_1 = -3,\ \alpha_3 = \alpha_5 = 0,\ N = 26)$

Tables 4.18 and 4.19 show the results when the IRF had orders 1 and 2, respectively. As is seen, the Universal Kriging estimator gave as good or better results in MSE and MPV than the Disjunctive Kriging estimator. While the Universal Kriging estimator tends to underestimate the variance of estimation, the Disjunctive Kriging tends to overestimate it. Figures 4.15 a, d; 4.16 b, e; 4.17 a, d; and 4.18 a, d shows the very accurate performance of the Universal Kriging estimator and the good looking approximation the Disjunctive Kriging estimator provides. However, there is no doubt that the Universal Kriging estimator, which adequately selected the IRF order in these cases (see Section 4.4.3), best captured the structure of the given fields. The anamorphosis fit of the Gaussian data used in the Disjunctive Kriging was found adequate with fitting errors in the mean and in the variance of the field below 5 percent in 45 of the 48 cases. See Figure 4.18 f.

Figures 4.15 f, g, h show typical deviations histograms, which show again the Gaussian like shape for the local mean estimator and the more concentrated densities for the Kriging estimators.

	MSE (10%)	MSE (25%)	MSE (50%)	MPV (10%)	MPV (25%)	MPV (50%)
K_M	12	12	12	12	12	12
DK_L	0	0	0	0	0	0
M_5	0	0	0			

CONSISTENCY

$$(\hat{\rho}_1 = \frac{MSE}{MPV})$$

	Min	Max	$\# \leq 1$	$\# > 1$	# [0.8,1.2]	μ	σ
K_M	0.48	43.2	5	7	5	1.45	1.26
DK_L	0.17	1.42	8	4	4	0.78	0.42

Table 4.18

Comparisons of the Different Estimators
Intrinsic Random Functions of Order 1 (12 cases) Point Results

	MSE			MPV		
	(10%)	(25%)	(50%)	(10%)	(25%)	(50%)
K_M	24	24	24	24	24	24
DK_s	0	2	3	0	0	0
M_5	0	0	0			

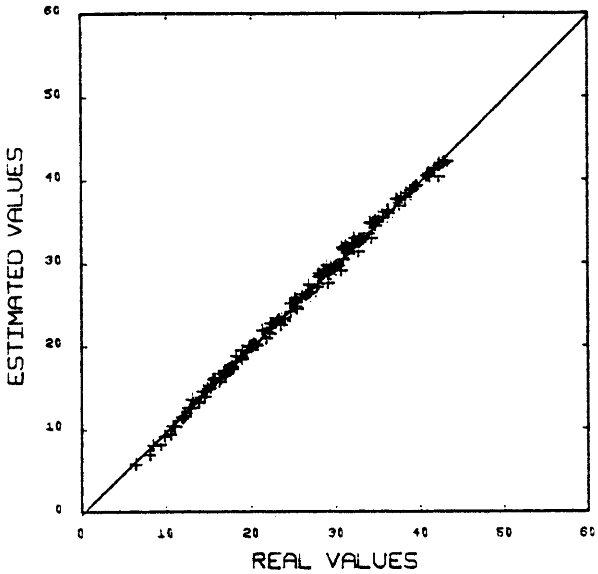
CONSISTENCY

$$(\hat{\rho}_1 = \frac{MSE}{MPV})$$

	Min	Max	#<1	#>1	# [0.8,1.2]	μ	σ
K_M	0.16	9.65	9	15	5	1.63	1.24
DK_s	0.07	10.86	19	5	1	0.58	0.78

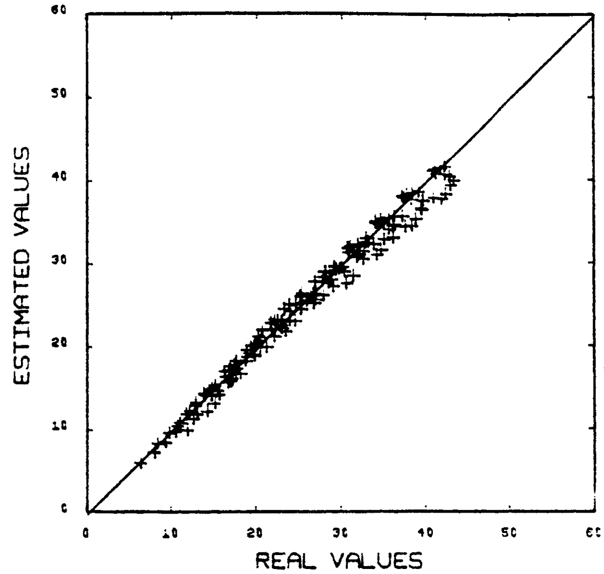
Table 4.19

Comparisons of the Different Estimators
Intrinsic Random Functions of Order 2 (24 cases)



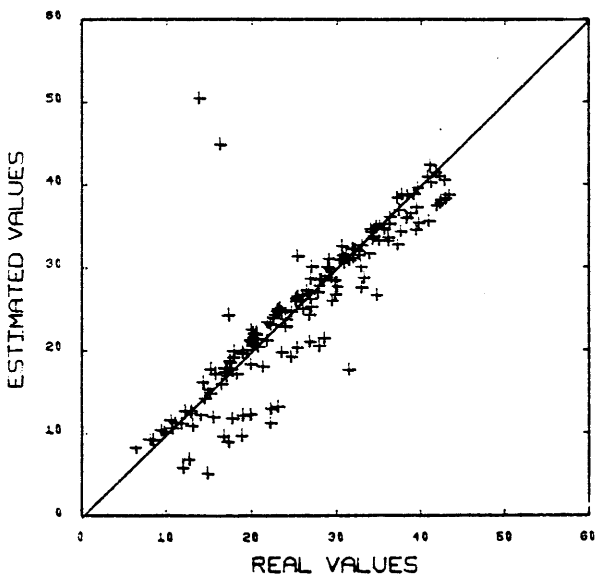
UK.R-M.k=1.q̂=0.91.N1=171.MSE=0.269.MPV=0.327
MsPV=2.258.CC=0.998.μD=0.089.σ̄D=0.51.q̂1=0.82

a



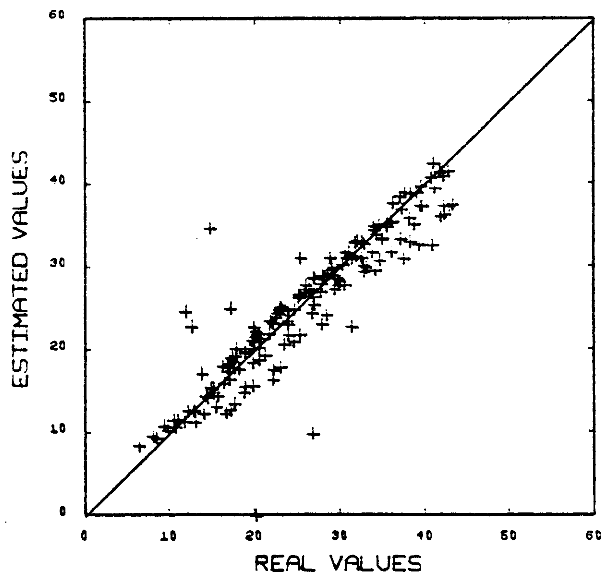
UK.D.k=1.q̂=0.98.N1=171.MSE=1.475.MPV=6.753
MsPV=9.091.CC=0.992.μD=-0.27.σ̄D=1.18.q̂1=0.22

b



DK.L.%μ=0.10.%V=1.00.N1=171.MSE=23.25.MPV=12.88
MsPV=65.80.CC=0.866.μD=-0.58.σ̄D=4.79.q̂1=1.81

c



DK.S.%μ=0.10.%V=1.00.N1=170.MSE=12.01.MPV=19.24
MsPV=447.0.CC=0.938.μD=-0.48.σ̄D=3.43.q̂1=0.62

d

Figure 4.15

Point Estimation Comparisons. Intrinsic Random Function

$(k = 1, \alpha_3 = 0.005, \alpha_1 = \alpha_5 = 0, N = 24)$

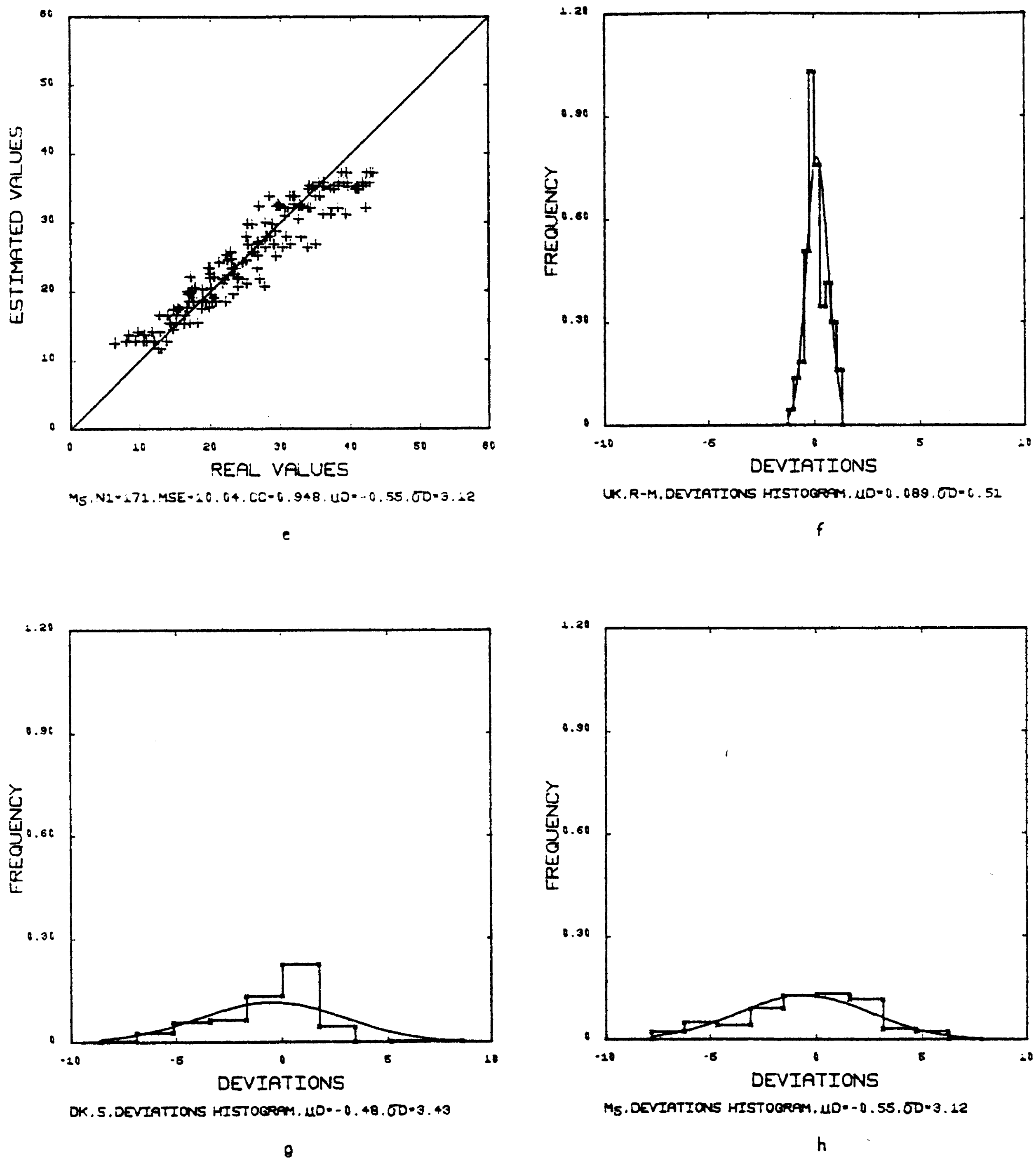


Figure 4.15

Point Estimation Comparisons. Intrinsic Random Function

($k = 1$, $\alpha_3 = 0.005$, $\alpha_1 = \alpha_5 = 0$, $N = 24$)

(continued)

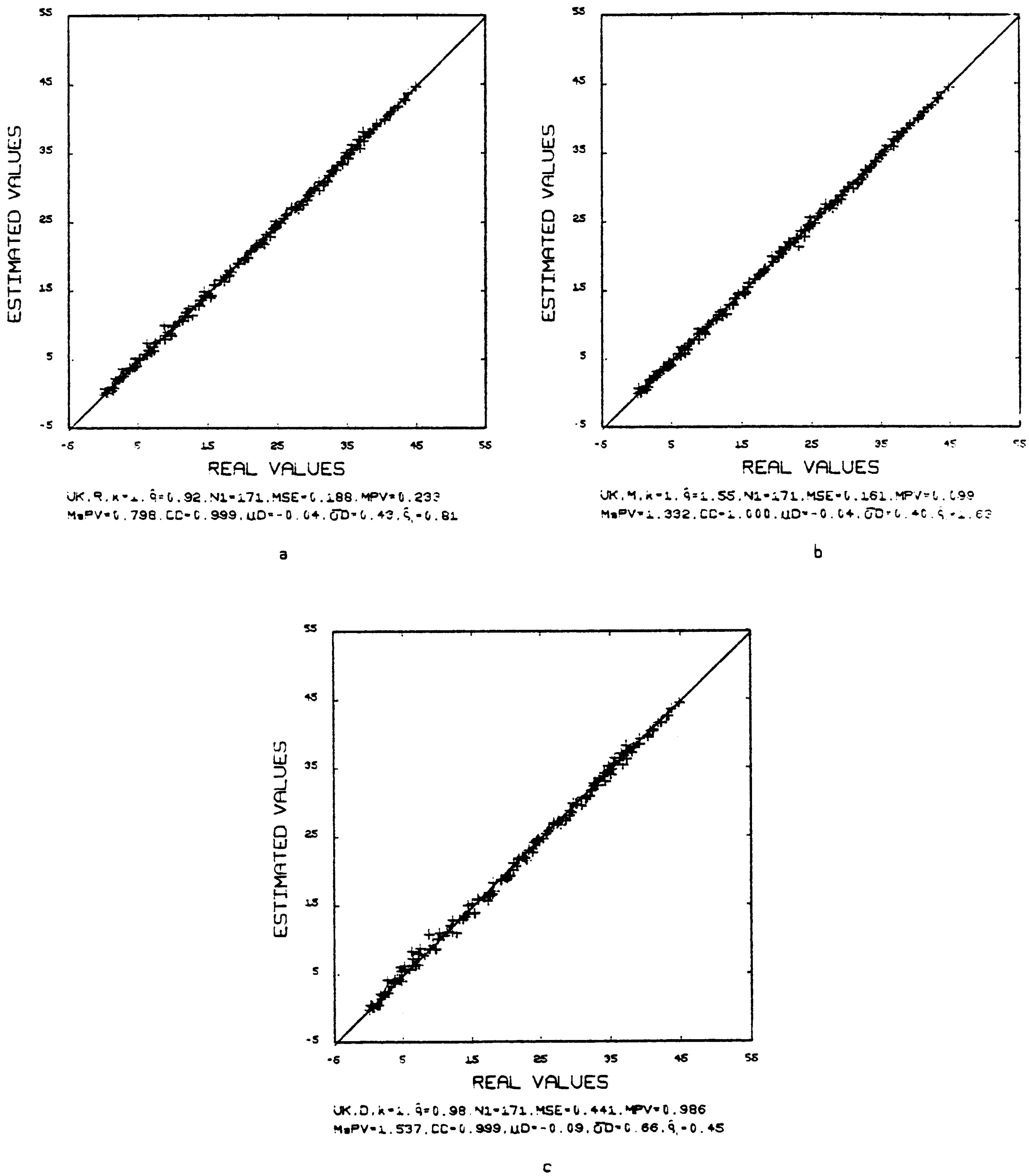
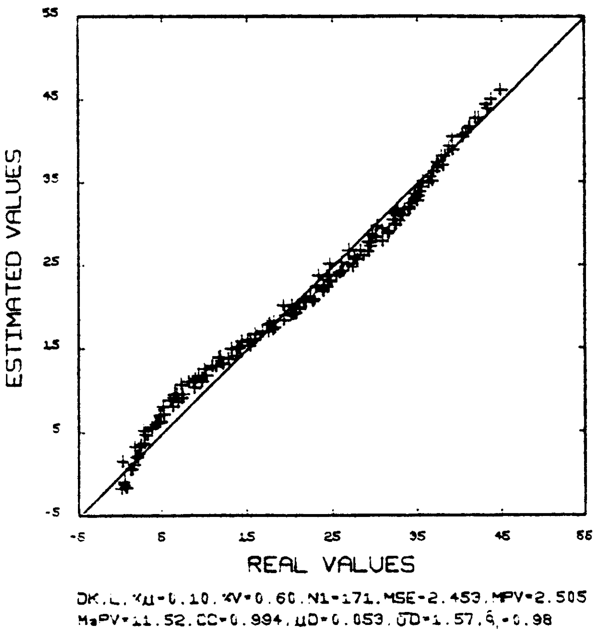


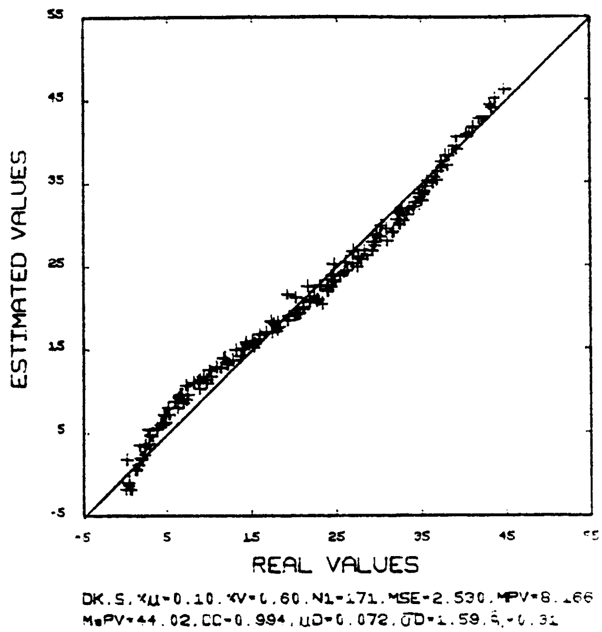
Figure 4.16

Point Estimation Comparisons. Intrinsic Random Function

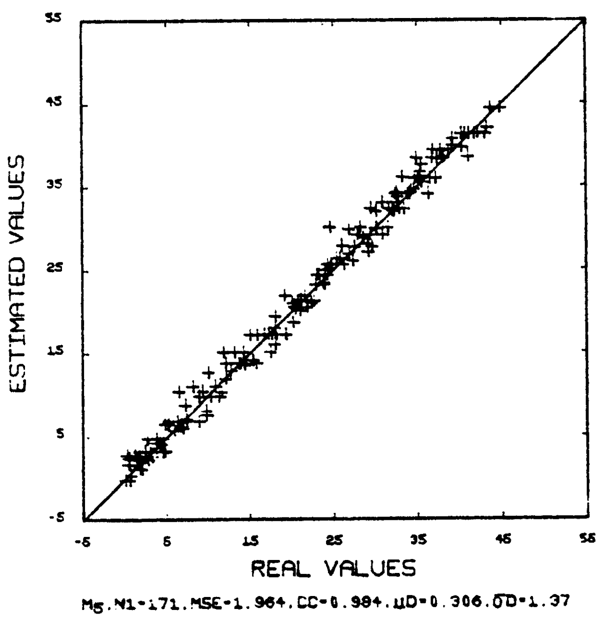
(k = 1, α_1 = -0.005, α_3 = 0.005, α_5 = 0, N = 48)



d



e



f

Figure 4.16

Point Estimation Comparisons. Intrinsic Random Function

($k = 1$, $\alpha_1 = -0.005$, $\alpha_3 = 0.005$, $\alpha_5 = 0$, $N = 48$)

(continued)

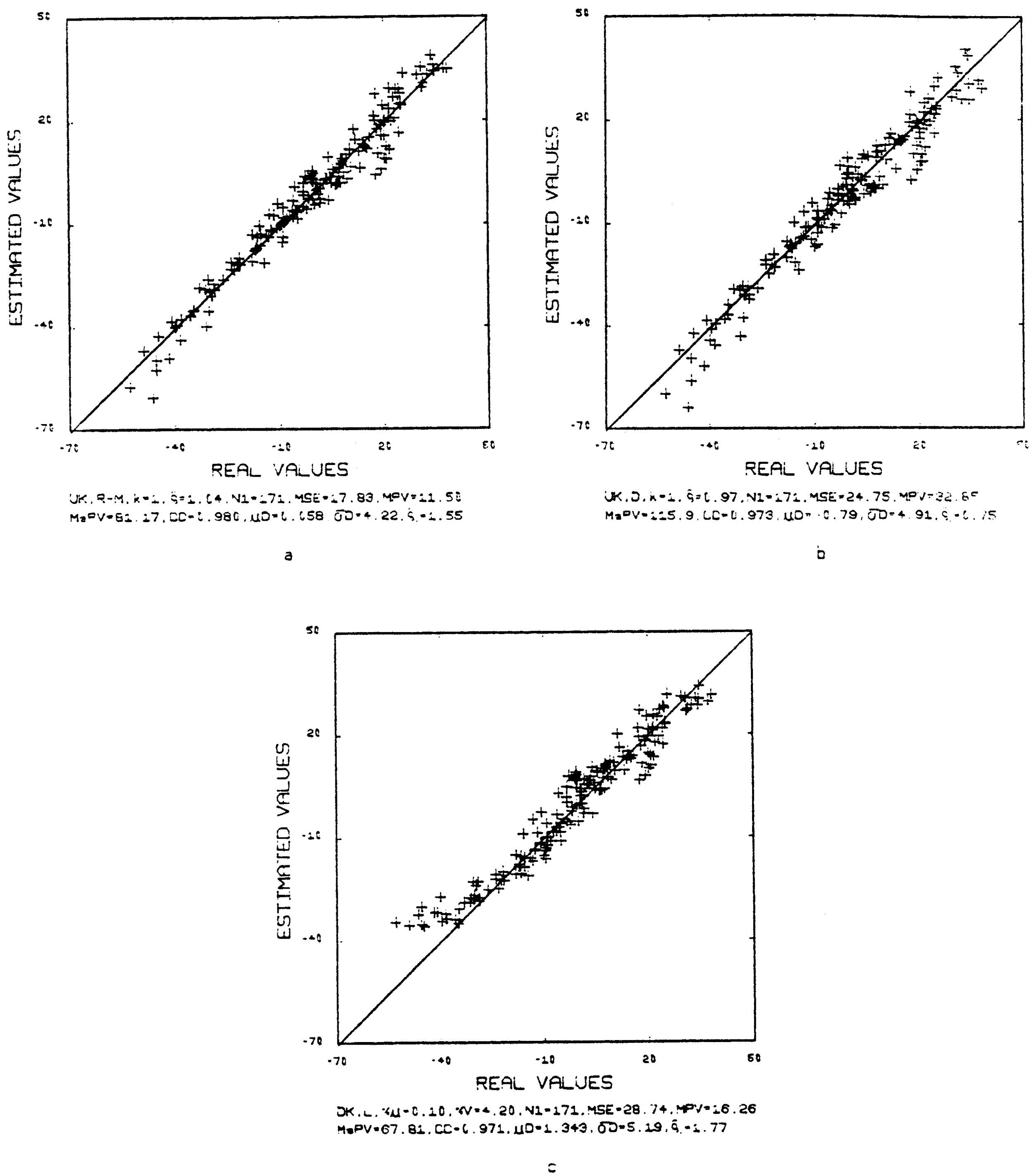
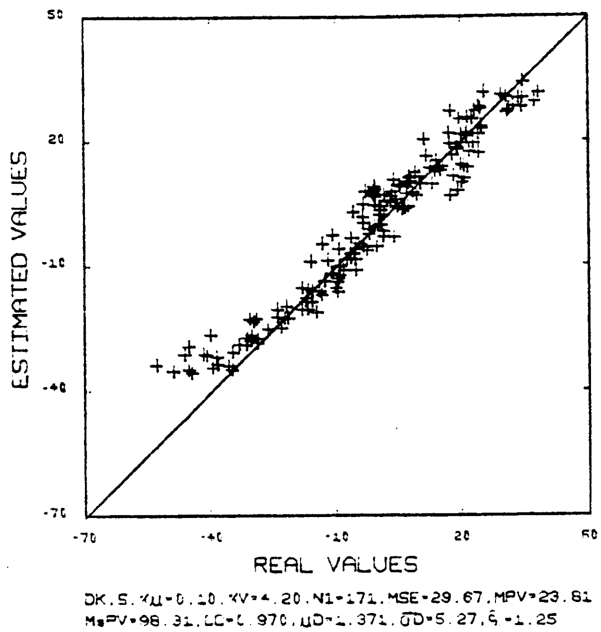


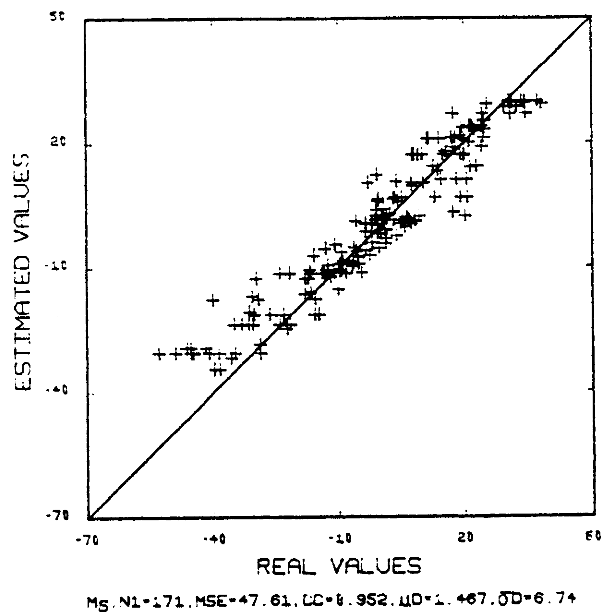
Figure 4.17

Point Estimation Comparisons. Intrinsic Random Function

($k = 2$, $\alpha_1 = -1$, $\alpha_3 = 0$, $\alpha_5 = -0.5 \times 10^{-9}$, $N = 44$)



d



e

Figure 4.17

Point Estimation Comparisons. Intrinsic Random Function

($k = 2$, $\alpha_1 = -1$, $\alpha_3 = 0$, $\alpha_5 = -0.5\text{x}10^{-9}$, $N = 44$)

(continued)

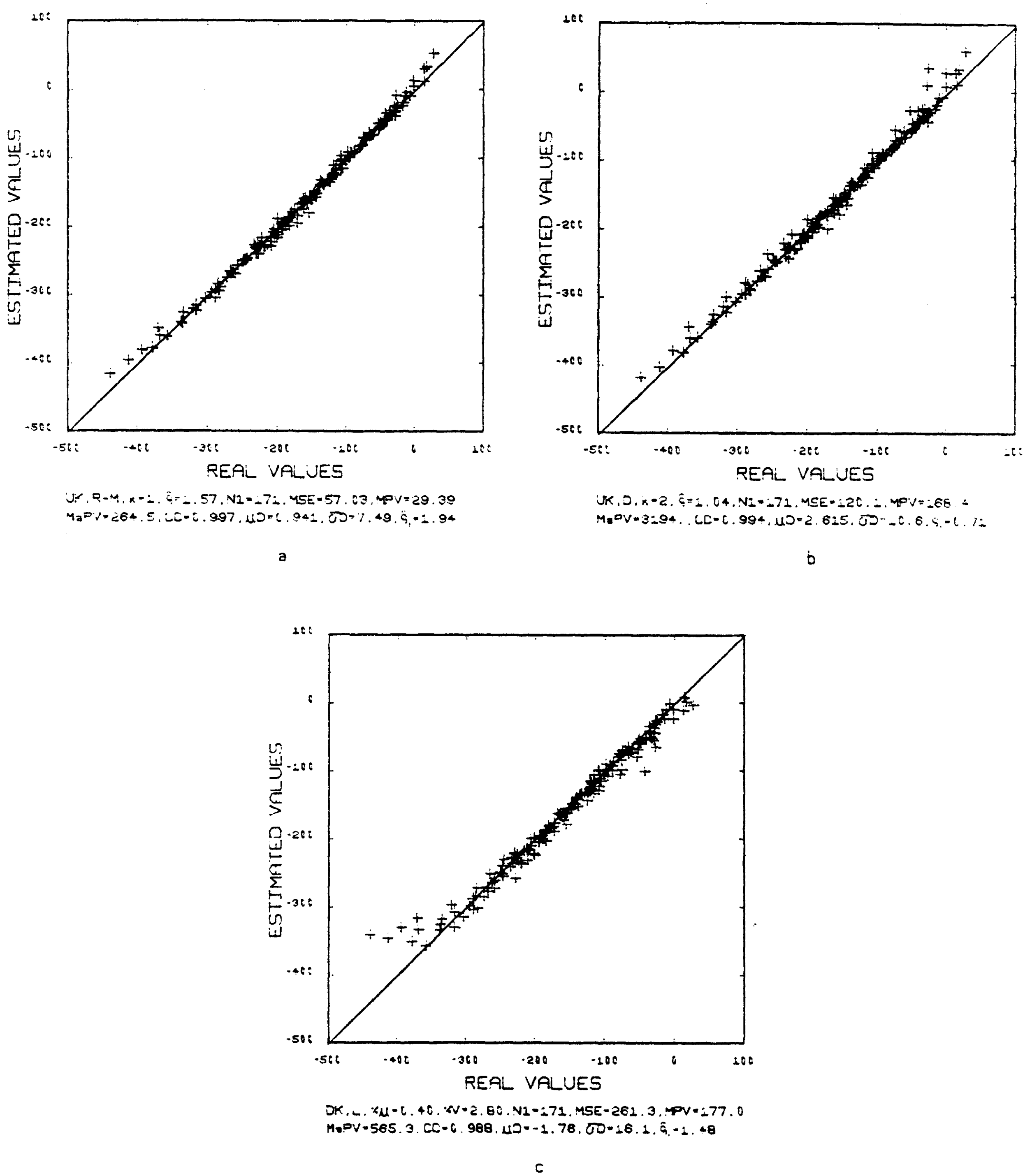
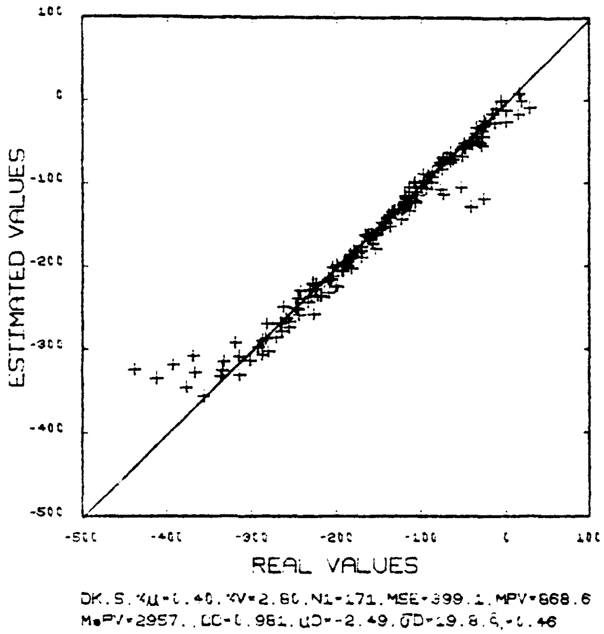


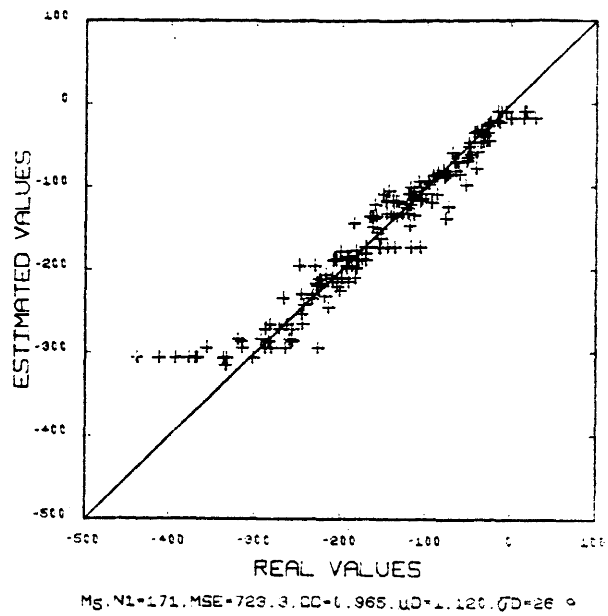
Figure 4.18

Point Estimation Comparisons. Intrinsic Random Function

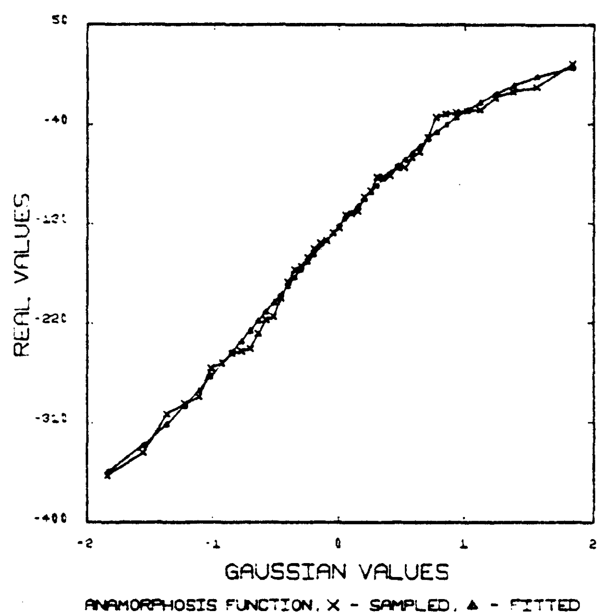
($k = 2$, $\alpha_1 = -2$, $\alpha_3 = 0.05$, $\alpha_5 = -1x10^{-9}$, $N = 45$)



d



e



f

Figure 4.18

Point Estimation Comparisons. Intrinsic Random Function

($k = 2$, $\alpha_1 = -2$, $\alpha_3 = 0.05$, $\alpha_5 = -1x10^{-9}$, $N = 45$)

(continued)

4.7 Summary

Using generated fields as the reality, it has been shown the least mean square error and ranking procedures gave better results than the methodology proposed by Delfiner (1976) in identifying and estimating the generalized covariance function of the fields. It was also found that the spherical semivariogram gave better results than the linear semivariogram when they were used in the calculation of the Disjunctive Kriging estimators.

Although it was consistently found that the Universal Kriging estimator underestimated the variance of estimation, its performance was as good or better than the theoretically more accurate Disjunctive Kriging estimator. This is mainly attributed to small sample induced biases in the estimation of the process variance.

Chapter 5

BLOCK ESTIMATION COMPARISONS

The results of the experiments made to compare the block estimation performance of Universal and Disjunctive Kriging are presented in this chapter.

5.1 Experiments Description

As in the case of point estimation, generated fields were used as the reality. The turning bands method was employed to generate values on a 61x31 grid on the previously defined rectangular area of 30.000 Km^2. The dark area on Figure 4.1 was divided in thirty six squares of side 24.5 Kms and its average (block) value was calculated employing the fourty nine generated values which lie inside or in the boundary of each block, weighting each value by its respective area, see Figure 5.1

The three different types of random fields considered in point estimation comparisons, see Tables 4.1, 4.2 and 4.3, were also used in block estimation, but for every generated field not six but two cases were studied: one with about 50 historical data points and the other with 30 historical data points. To have these points spreaded evenly over the area, they were sampled not from the 61x31 grid but from the subgrid of 21x11 points as previously used in point estimation[1].

[1] Notice that these values do not necessarily are equal to the generated grid on Chapter 4, because the values given by the Turning Bands method depend on the total number of points to be generated. See Mantoglou and Wilson (1981).

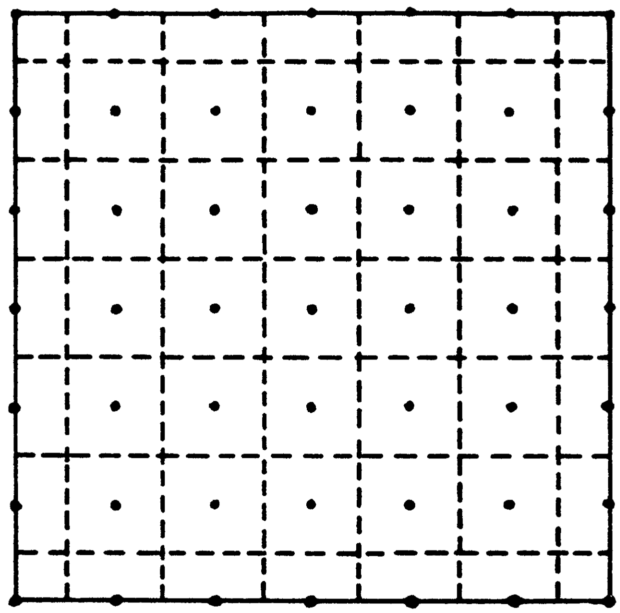


Figure 5.1
Calculation of the True Block Values
Sides = 24.5 Kms.

As in point estimation, the historical points were used to identify and estimate the structures of the different estimators, which were then employed to estimate areal averages of the fields on the 36(4x9) blocks. Once the block values were found, comparisons with the true (constructed from the generated results) values were made.

In the following sections, the results of comparing the different estimators according to the measures defined on Section 4.3 are presented.

5.2 Comparisons of the Universal Kriging Estimators

In this section the performances of the Universal Kriging estimators, using the three previously described methodologies of selection of the optimal generalized covariance function, are discussed.

The block estimator was calculated employing the eight nearest points to the center of each block, approximating the required integrals over each block (see Chapter 2) as summations over 25 equally spaced points.

5.2.1 Isotropic Fields

Table 5.1 summarizes the results of the 12 cases (six fields each, sampled twice) of isotropic fields. As is can be seen, the ranking, R, and least mean square error, M, procedures have an advantage over Delfiner's methodology, M, as expressed by lower values of MSE and MPV over the thirty six considered blocks. As previoulsy found in point estimation comparisons, generalized covariance models obtained using Delfiner's method tend to overestimate the variance of estimation, although prior point values of the consistency parameter $\hat{\rho}_1$ lie close to 1.

The consistency parameter does not show; however, the previously found trend of underestimating the predicted variance with the generalized covariance models given by the ranking and least mean square error procedures. On the contrary, now the tendency is toward overestimation, with values of $\hat{\rho}_1$ greater on the average than the values of $\hat{\rho}_1$ given using Delfiner's approach.

	MSE (10%)	MSE (25%)	MPV (10%)	MPV (25%)	MaPV (25%)	CC (5%)	μD ($\pm$0.5)	σD (25%)
R	8	8	10	10	9	11	9	10
M	11	11	9	9	8	11	11	11
D	1	2	1	3	4	8	8	5

Selected IRF

	0	1	2
R	3	6	3
M	3	8	1
D	3	7	2

Same Orders

RMD- 5
RM - 2
RD - 4
DM - 1

Same Values

RMD- 0
RM - 7
RD - 0
DM - 0

CONSISTENCY

$$(\hat{\rho}_1 = \frac{MSE}{MPV}) \text{ (POINT PRIOR)}$$

	Min	Max	#≤1	#>1	# [0.8,1.2]	μ	σ
R	0.22(0.78)	9.07(2.18)	9(4)	3(8)	1(9)	0.75(1.13)	0.53(0.36)
M	0.22(0.92)	3.2 (1.65)	10(3)	2(9)	2(10)	0.86(1.13)	0.83(0.21)
D	0.3 (0.97)	1.56(1.10)	11(8)	1(4)	0(12)	0.58(1.01)	0.33(0.03)

Table 5.1

Comparisons of Universal Kriging Estimators

Isotropic Fields (12 cases) Block Results

As in point comparisons, the three methods selected any of the IRF orders, with Delfiner's approach not coinciding with the other methods in any of the cases.

As should be expected, the block values give less extreme points than the point values. This fact explains why the model selected by Delfiner's approach gives a block estimate that performs reasonably in the extremes, in contrast with the conservative point estimates that result with this method. Figures 5.2 a, b, c; 5.3 a, b, c; and 5.4 a, b show typical results of the estimation of isotropic fields.

5.2.2 Transformed Isotropic Fields

Table 5.2 summarizes the results of the 12(6x2) cases of transformed isotropic fields. The results follow the same pattern found with the isotropic fields. Again, Delfiner's approach does not give better results than the other two methods.

The three identification methods were inclined to select IRF of orders 0 and 1. Again, Delfiner's model did not coincide with models given by the other methods which, one the other hand, resulted in the same structure in five of the twelve cases.

The consistency parameter show a tendency toward overestimation for all the three methods, with more overestimation with Delfiner's approach. The ranking procedure gave better results than the least mean square error method in terms of $\hat{\rho}_1$. Overall, though, no significant difference between those two methods can be inferred.

	MSE		MPV		MaPV	CC	μD	σD
	(10%)	(25%)	(10%)	(25%)	(25%)	(5%)	(±0.5)	(25%)
R	11	11	7	8	8	12	11	12
M	7	10	12	12	12	11	11	11
D	1	4	0	2	3	1	10	6

Selected IRF

	0	1	2
R	10	1	1
M	7	5	0
D	10	1	1

Same Orders
RMD- 5
RM - 0
RD - 7
DM - 0

Same Values
RMD- 0
RM - 5
RD - 0
DM - 0

CONSISTENCY

$(\hat{\rho}_1 = \frac{MSE}{MPV})$ (POINT PRIOR)

	Min	Max	#≤1	#>1	# [0.8,1.2]	μ	σ
R	0.19(0.64)	2.02(1.23)	9(3)	3(9)	2(10)	0.66(1.04)	0.54(0.16)
M	0.32(0.82)	17.9 (3.03)	8(2)	4(10)	0(8)	1.37(1.34)	1.4 (0.63)
D	0.20(0.68)	0.77(1.0)	12(12)	0(0)	0(11)	0.42(0.97)	0.17(0.09)

Table 5.2

Comparisons of Universal Kriging Estimators

Transformed Isotropic Fields (12 cases) Block Results

Examples of the estimation of the transformed isotropic fields are given in Figures 5.5 a, b, c; 5.6 a, c, b; and 5.7 a, b, c.

5.2.3 Intrinsic Random Functions

Table 5.3 summarizes the comparisons of the Universal Kriging estimators when the original fields were intrinsic random functions. As can be seen there is again an advantage of the ranking and least mean square error procedures over Delfiner's approach, as expressed by lower MSE, MPV, and MaPV.

The consistency parameter again shows a greater mean predicted variance than the true mean square error for all the methodologies of estimation of the generalized covariance function. Again, the lower values of $\hat{\rho}_1$ were found with Delfiner's model which as before had better prior consistency.

Figures 5.8 a, b; 5.9 a, b; 5.10 a; 5.11 a, b; and 5.12 a, b show estimation results of intrinsic random functions. As is seen the performances of the ranking and least mean square error procedures are very similar, although Table 5.3 suggests better results with the last method. As found in point comparisons, when the IRF order is one or two, the agreement reached is remarkable with correlation coefficients between true and estimated values of 0.99.

	MSE (10%)	MSE (25%)	MPV (10%)	MPV (25%)	MaPV (25%)	CC (5%)	μD (±0.5)	σD (25%)
R	10	11	9	9	12	15	14	15
M	14	15	13	13	15	16	16	16
D	8	8	3	3	5	14	11	11

Selected IRF

	0	1	2
R	1	5	10
M	2	6	8
D	1	4	11

Same Orders
RMD-11
RM - 1
RD - 4
DM - 0

Same Values
RMD- 2
RM - 6
RD - 4
DM - 0

CONSISTENCY

$$(\hat{\rho}_1 = \frac{MSE}{MPV})\text{ (PRIOR POINT)}$$

	Min	Max	#≤1	#>1	# [0.8,1.2]	μ	σ
R	0.16(0.58)	41.5(1.86)	11(7)	5(9)	3(12)	0.77(1.1)	0.68(0.38)
M	0.11(0.64)	41.5(2.03)	10(7)	6(9)	2(8)	1.2 (1.22)	1.47(0.41)
D	0.05(0.73)	1.13(1.3)	13(7)	3(9)	4(14)	0.59(1.0)	0.29(0.13)

Table 5.3

Comparisons of Universal Kriging Estimators

Intrinsic Random Functions (16 cases) Block Results

5.3 Comparisons of the Disjunctive Kriging Estimators

Following is the study of the Disjunctive Kriging estimator when calculated with linear and spherical semivariograms.

As done with the Universal Kriging estimators, the calculations were made using the eight neighboring points of the center of each block. The necessary integrals over each block (see Chapter 3) were approximated from twenty five equally spaced points. As in point estimation calculations, Disjunctive Kriging was performed using the first five Hermite polynomials and assuming a linearly interpolated anamorphosis function.

5.3.1 Isotropic Fields

Table 5.4 summarizes the results of the Disjunctive Kriging estimators when the generated fields were isotropic. There is a slight advantage of the spherical semivariogram with respect to the mean square error over the thirty six blocks, while the use of the linear semivariogram gave better results in the mean and maximum predicted variance.

The consistency parameter shows that both the linear and spherical semivariograms tend to overestimate the predicted variance, with better results for the linear semivariogram. See Figures 5.2 d, e; 5.3 d, e; and 5.4 c, d.

The use of linear interpolation and finite Hermite expansion of the anamorphosis led to less than 6.5 percent error in fitting the mean, $\%\mu$, and variance, $\%V$, of the actual data, see Equations (3.34) and (3.36).

	MSE		MPV		MaPV	CC	μD	σD
	(10%)	(25%)	(10%)	(25%)	(25%)	(5%)	(± 0.5)	(25%)
L	7	9	10	11	11	10	10	9
S	12	12	3	6	7	12	12	12

CONSISTENCY

$$(\hat{\rho}_1 = \frac{MSE}{MPV})$$

	Min	Max	$\# \leq 1$	$\# > 1$	# [0.8,1.2]	μ	σ
L	0.20	1.80	9	3	0	0.77	0.6
S	0.17	1.05	11	1	2	0.43	0.29

Table 5.4

Comparisons of Disjunctive Kriging Estimators
Isotropic Fields (12 cases) Block Results

5.3.2 Transformed Isotropic Fields

The results for the twelve transformed isotropic fields are given in Table 5.5. Disjunctive Kriging gave better results when it was calculated with the spherical semivariogram, as expressed by lower values of mean square error, mean and maximum predicted variance and closest to one correlation coefficient.

The consistency parameter shows that both semivariograms give results that tend to overestimate the predicted variance, with more consistent results given by the spherical semivariogram. Figures 5.5 d, e; 5.6 d, e; and 5.7 d, e show typical outcomes of these cases.

Again, the linearly interpolated and finite Hermite expansion of the anamorphosis gave adequate fits of the actual data, as expressed by values of %μ and %V smaller than 6 percent in all twelve cases.

5.3.3 Intrinsic Random Functions

The comparisons of the the two Disjunctive Kriging estimators when generated fields were given by intrinsic random functions are summarized in Table 5.6. The spherical semivariogram gave better results with respect to the mean square error, while the linear semivariogram gave lower predicted variances.

The consistency parameter shows that both semivariograms gave mean predicted variances that were larger than the mean square errors on almost all the 16 cases. The linear semivariogram, however, is more consistent as expressed by more values on the interval [0.8,1.2] and greater mean value of the parameter $\hat{\rho}_1$.

	MSE (10%)	MSE (25%)	MPV (10%)	MPV (25%)	MaPV (25%)	CC (5%)	μD (± 0.5)	σD (25%)
L	5	6	4	4	4	7	10	7
S	12	12	9	9	9	12	12	12

CONSISTENCY

$$(\hat{\rho}_1 = \frac{MSE}{MPV})$$

	Min	Max	$\# \leq 1$	$\# > 1$	# [0.8,1.2]	μ	σ
L	0.14	1.82	8	4	1	0.87	0.55
S	0.08	1.55	8	4	4	0.85	0.4

Table 5.5

Comparisons of Disjunctive Kriging Estimators
Transformed Isotropic Fields (12 cases) Block Results

	MSE		MPV		MaPV	CC	μD	σD
	(10%)	(25%)	(10%)	(25%)	(25%)	(5%)	(± 0.5)	(25%)
L	9	11	13	14	10	13	13	14
S	13	15	6	7	8	16	15	15

CONSISTENCY

$$(\hat{\rho}_1 = \frac{MSE}{MPV})$$

	Min	Max	$\# \leq 1$	$\# > 1$	# [0.8,1.2]	μ	σ
L	0.12	1.13	15	1	5	0.55	0.31
S	0.04	1.08	15	1	1	0.33	0.3

Table 5.6

Comparisons of the Disjunctive Kriging Estimators
Intrinsic Random Functions (16 cases) Block Results

Figures 5.8 c, d; 5.9 c, d; 5.10 b, c; 5.11 c, d; and 5.13 c, d show typical outcomes of these cases.

5.4 Comparison of the Different Estimators

This section compares the Universal and Disjunctive Kriging estimators. As in point estimation comparisons, the Universal Kriging estimator used was the one given by the least mean square error procedure, although certainly the ranking procedure gave as good results. Similarly, the Disjunctive Kriging estimator chosen was the one calculated using a spherical semivariogram of the Gaussian data.

A local mean estimator was also compared to the Kriging estimators. It was calculated by estimating twenty five equally spaced points inside each block by averaging their nearest five observations, and then averaging over the twenty five estimated values to get the block estimate.

5.4.1 Isotropic Fields

The more relevant results of the twelve cases of Isotropic fields are summarized in Table 5.7. The table reveals better performance for the Universal Kriging estimator in both mean square error and specially in mean predicted variance. The local mean estimator performed much worse than the others as can be seen in Figures 5.2 b, e, f; 5.3 b, e, f; and 5.4 a, d, e.

The consistency parameter, $\hat{\rho}_1$, shows a tendency of overestimating the predicted variance by both estimators. The table suggests more consistent results for the Universal Kriging estimator, however, a higher standard deviation of $\hat{\rho}_1$, exists.

	MSE (10%)	MSE (25%)	MSE (50%)	MPV (10%)	MPV (25%)	MPV (50%)
K_M	11	11	11	11	11	11
DK_s	7	8	8	1	2	3
M_5	1	3	3			

CONSISTENCY

$$(\hat{\rho}_1 = \frac{MSE}{MPV})$$

	Min	Max	$\# \leq 1$	$\# > 1$	# [0.8,1.2]	μ	σ
K_M	0.22	3.2	10	2	2	0.86	0.83
DK_s	0.17	1.05	11	1	2	0.43	0.29

Table 5.7

Comparisons of the Different Estimators
Isotropic Fields (12 cases) Block Results

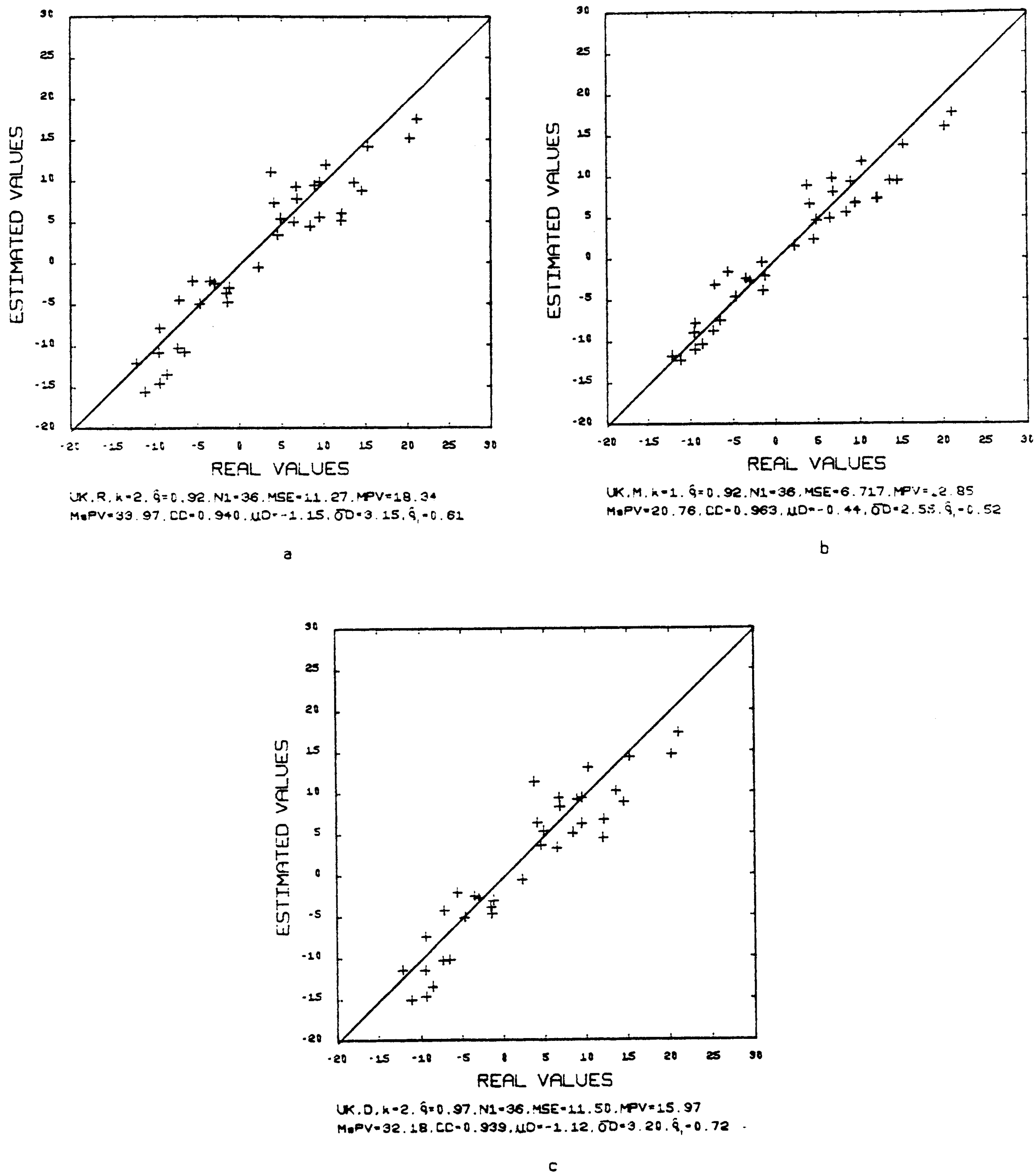
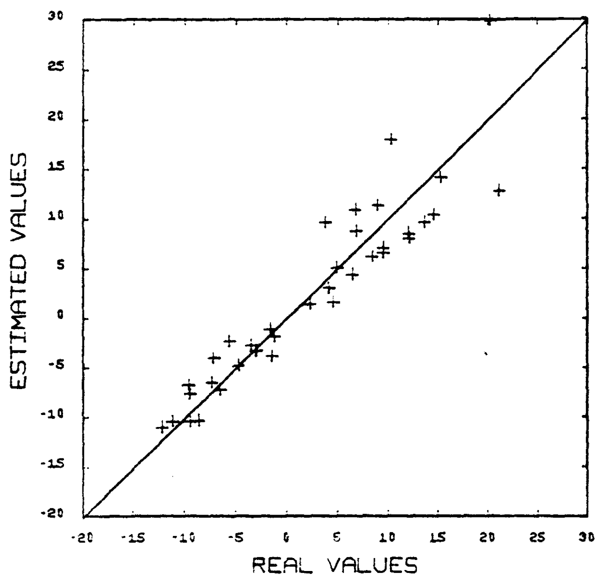


Figure 5.2

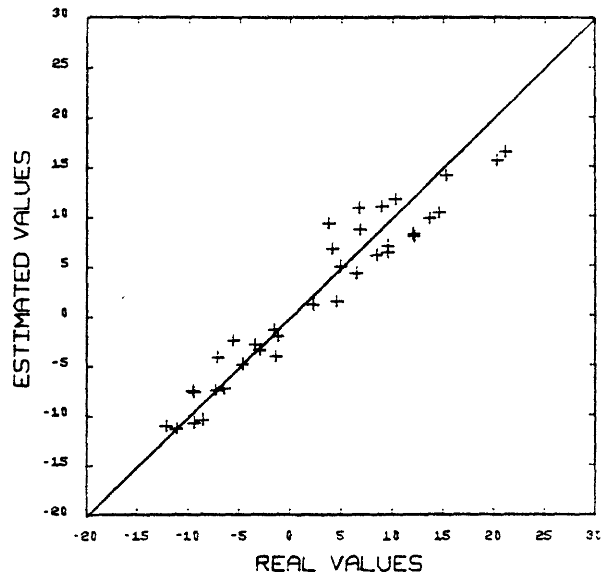
Block Estimation Comparisons. Isotropic Field

($b = 0.007$, $\sigma = 10$, $m = 10$, $N = 28$)



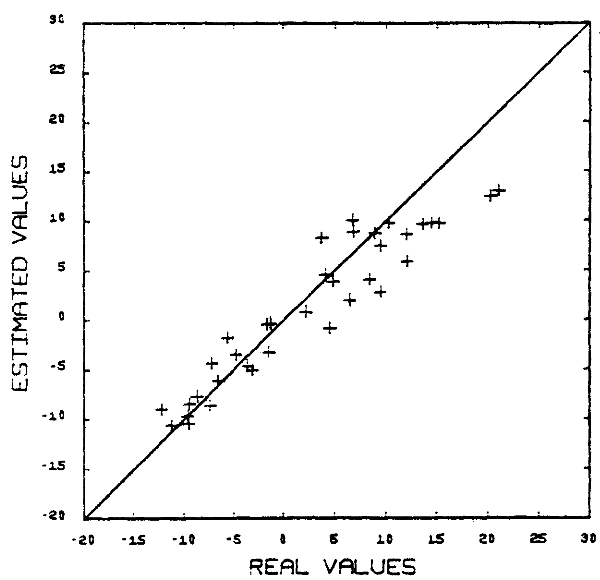
DK.L. %μ=4.10. %V=5.00. N1=35. MSE=9.456. MPV=14.52
MaPV=23.63. CC=0.938. μD=-0.17. σD=3.07. $\hat{\rho}_1$=0.85

d



DK.S. %μ=4.10. %V=5.00. N1=36. MSE=6.797. MPV=22.06
MaPV=36.33. CC=0.962. μD=-0.36. σD=2.58. $\hat{\rho}_1$=0.31

e



M_5. N1=36. MSE=12.46. CC=0.945. μD=-1.28. σD=3.29

f

Figure 5.2

Block Estimation Comparisons. Isotropic Field

(b = 0.007, σ = 10, m = 10, N = 28)

(continued)

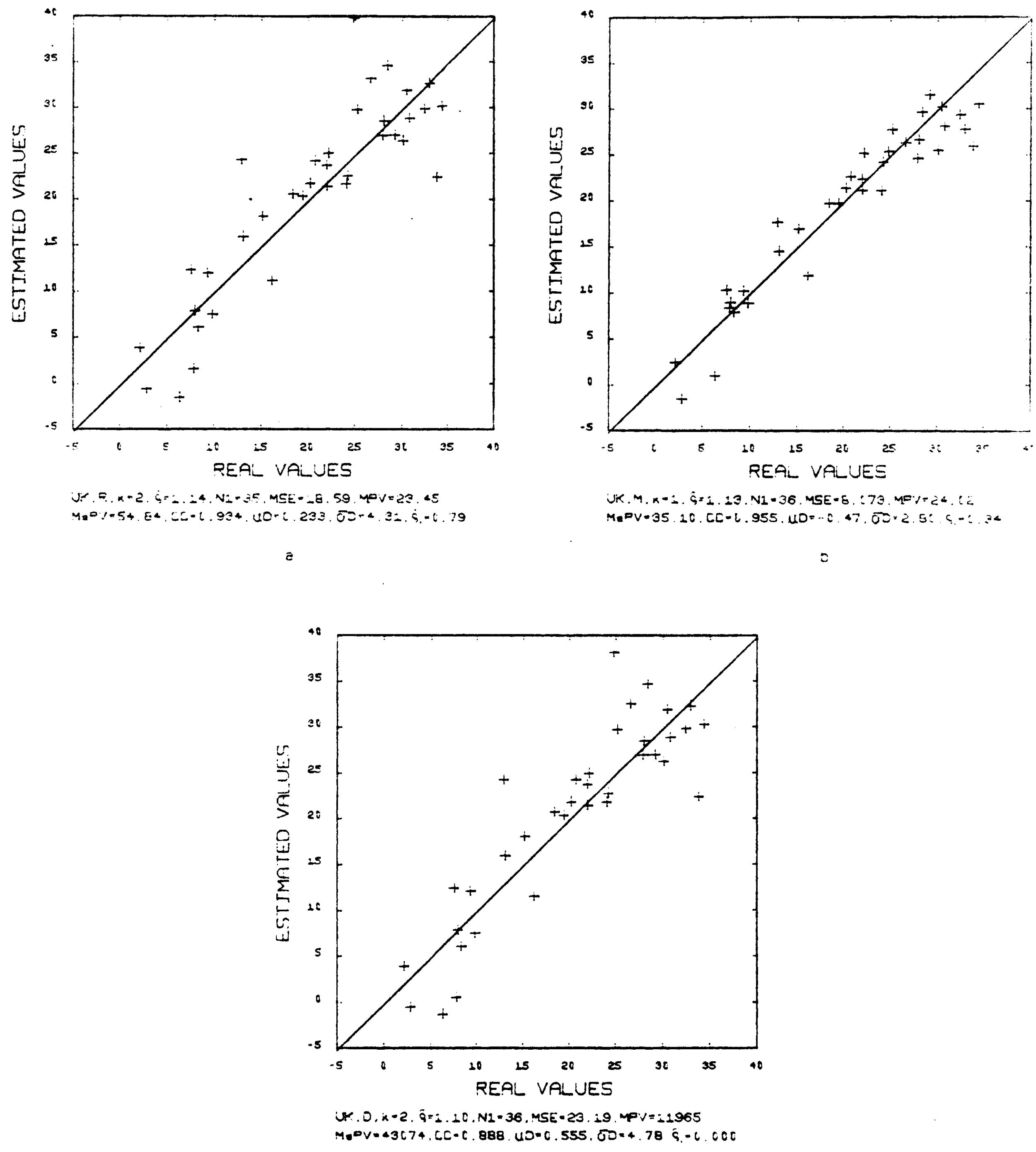
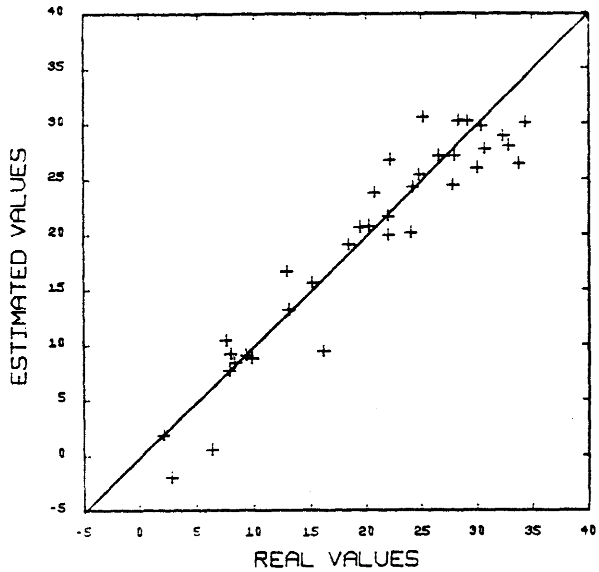


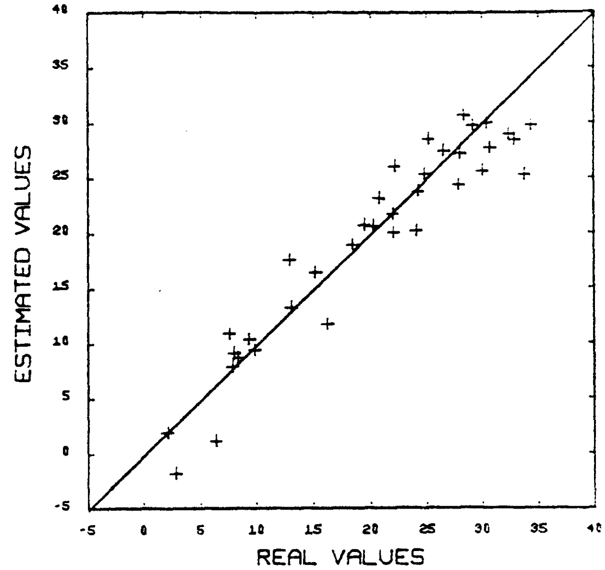
Figure 5.3

Block Estimation Comparisons. Isotropic Field

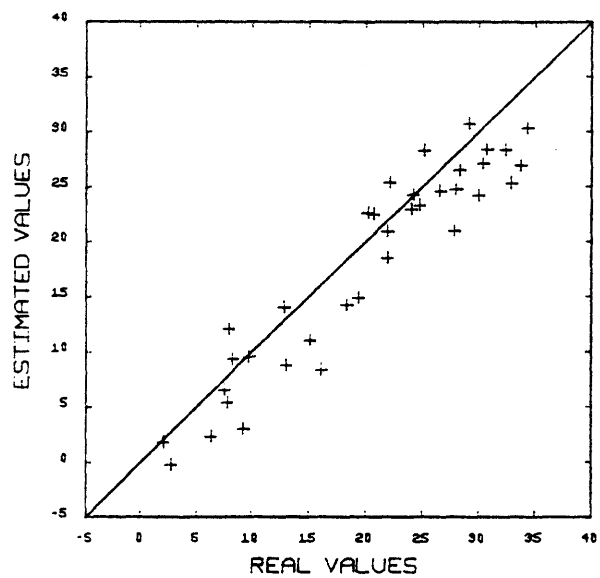
(b = 0.0141, σ = 10, m = 10, N = 45)



d



e



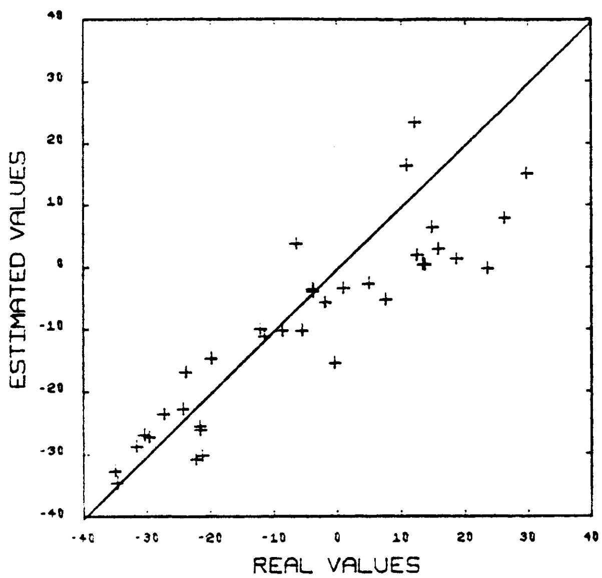
f

Figure 5.3

Block Estimation Comparisons. Isotropic Field

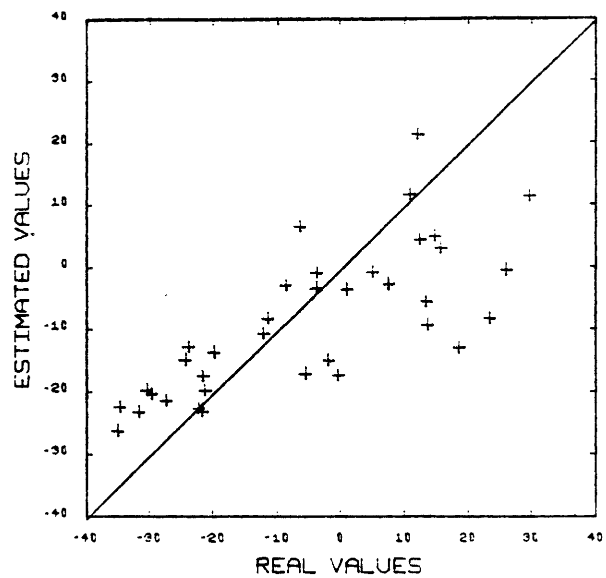
(b = 0.0141, σ = 10, m = 10, N = 45)

(continued)



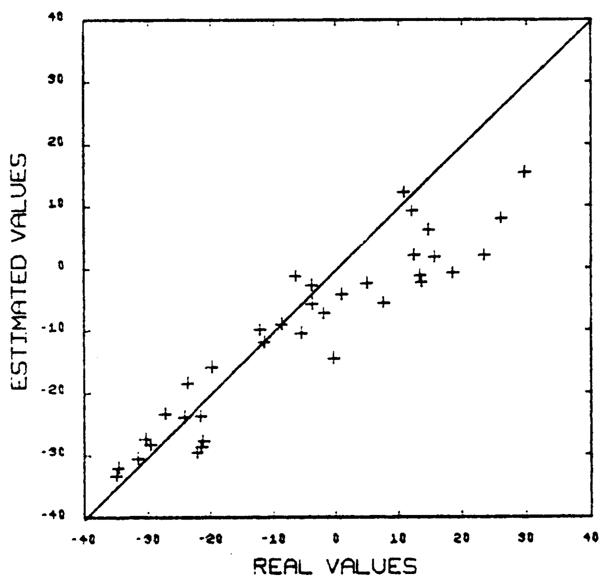
UK.R-M.k=1.$\hat{q}$=1.43.N1=36.MSE=85.22.MPV=49.37
MaPV=119.3.CC=0.898.μD=-3.73.$\bar{\sigma} D$=8.44.$\hat{q}_{\cdot}$=1.73

a



UK.D.k=1.$\hat{q}$=1.00.N1=36.MSE=166.9.MPV=232.1
MaPV=300.1.CC=0.753.μD=-2.91.$\bar{\sigma} D$=12.6 $\hat{q}_{\cdot}$=0.72

b



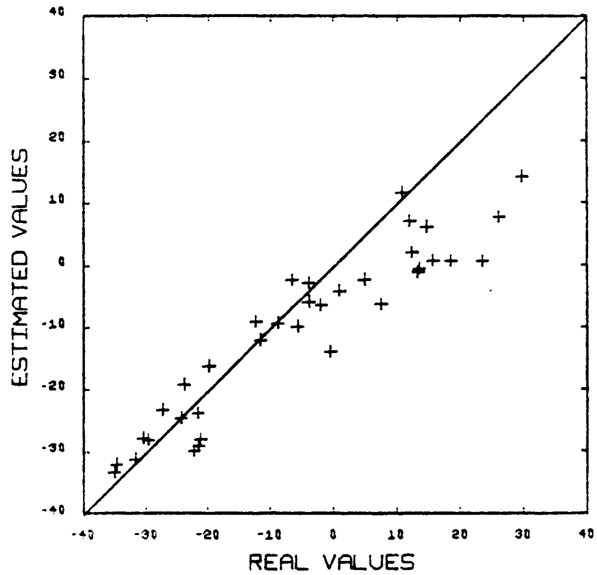
OK.L.%μ=0.80.%V=4.20.N1=36.MSE=80.35.MPV=65.75
MaPV=113.9.CC=0.936.μD=-4.77.$\bar{\sigma} D$=7.59.$\hat{q}_{\cdot}$=1.22

c

Figure 5.4

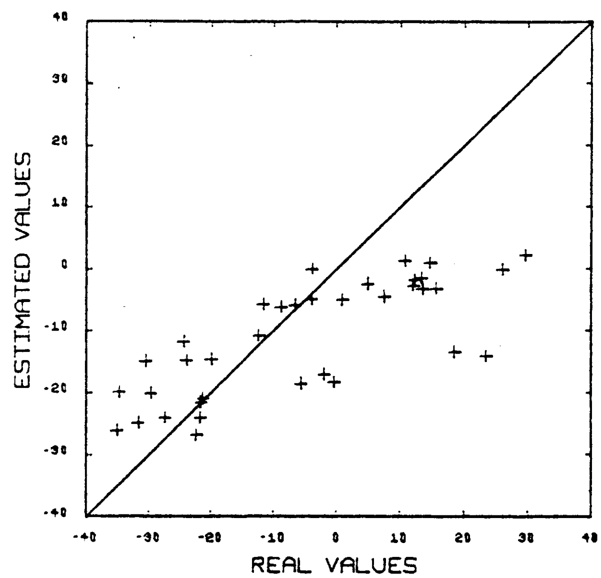
Block Estimation Comparisons. Isotropic Field

(b = 0.0141, σ = 20, m = 10, N = 28)



DK.S, %μ=0.80, %V=4.20, N1=36, MSE=81.32, MPV=77.40
MsPV=139.1, CC=0.939, μD=-4.95, σD=7.54, $\hat{\delta}_1$=-1.05

d



M_S, N1=36, MSE=205.9, CC=0.764, μD=-5.62, σD=13.2

e

Figure 5.4

Block Estimation Comparisons. Isotropic Field

(b = 0.0141, σ = 20, m = 10, N = 28)

(continued)

5.4.2 Transformed Isotropic Fields

Performance in estimating the transformed isotropic fields is summarized in Table 5.8. As is seen the Disjunctive Kriging estimator gave as good or better results than the Universal Kriging technique: they have about equal performance with respect to mean square error, while Disjunctive Kriging holds some what of an edge on mean predicted variance.

The consistency parameter also favors results of the Disjunctive Kriging estimator, with both estimators showing a tendency toward overestimation.

Figures 5.5 b, e, f; 5.6 b, e, f; and 5.7 b, e, f show typical results. Both Kriging estimators clearly do better than the local mean estimator.

5.4.3 Intrinsic Random Functions

Table 5.9 summarizes the results of the different estimators when the true fields were generated by intrinsic random functions. Again, the local mean estimator performed worse than the Kriging estimators.

When the IRF order was zero, Disjunctive Kriging was found as good or better than Universal Kriging. The good scores of DK on Table 5.9 correspond to these cases, see Figures 5.8 a, d, e. However, when the IRF order was one or two, the Universal Kriging technique gave better results in both mean square error and mean predicted variance. The excellence of the estimation is shown in Figures 5.9 a; 5.10 a; 5.11 a;

	MSE (10%)	(25%)	(50%)	MPV (10%)	(25%)	(50%)
K_M	8	8	9	5	5	5
DK_s	7	9	10	7	7	7
M_5	1	3	3			

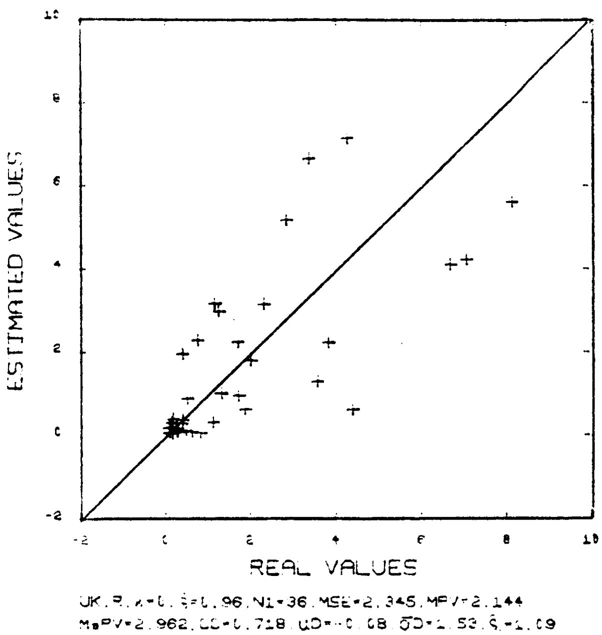
CONSISTENCY

$$(\hat{\rho}_1 = \frac{MSE}{MPV})$$

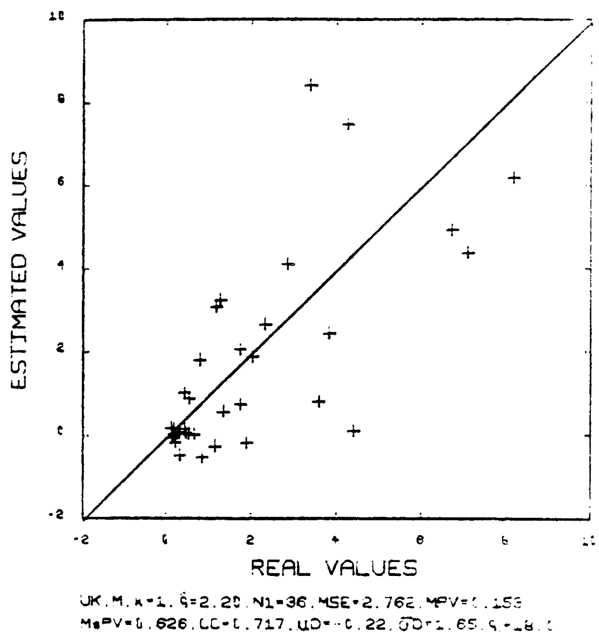
	Min	Max	$\# \leq 1$	$\# > 1$	# [0.8,1.2]	μ	σ
K_M	0.32	17.9	8	4	0	1.37	1.4
DK_s	0.08	1.55	8	4	4	0.85	0.4

Table 5.8

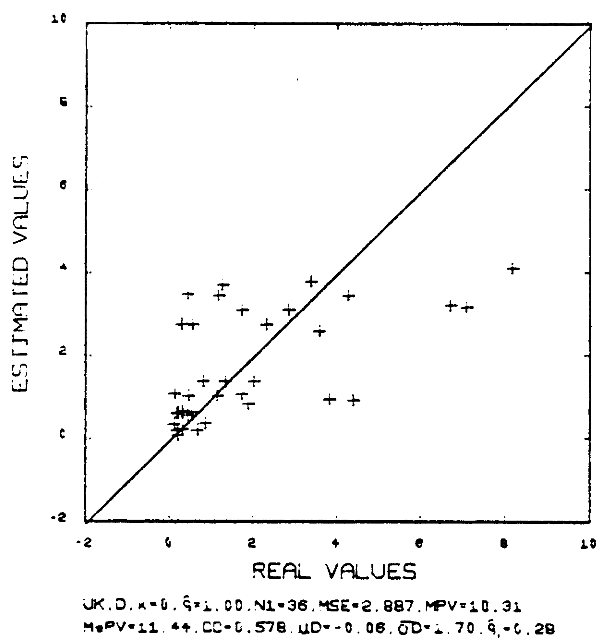
Comparisons of the Different Estimators
Transformed Isotropic Fields (12 cases) Block Results



a



b

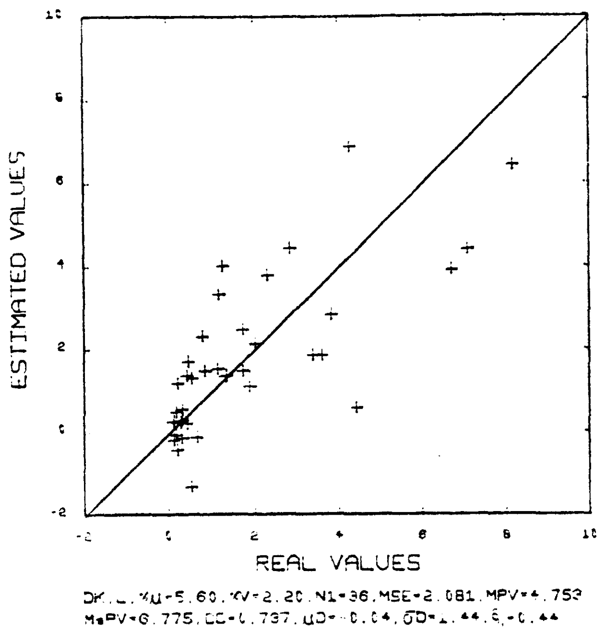


c

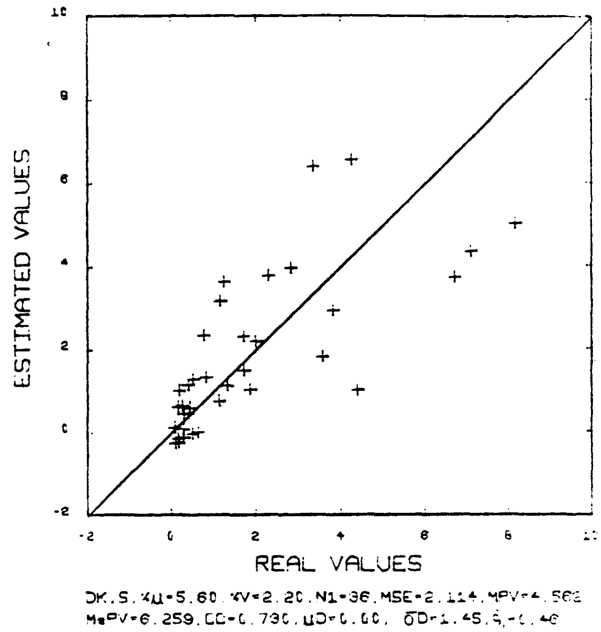
Figure 5.5

Block Estimation Comparisons. Transformed Isotropic Field

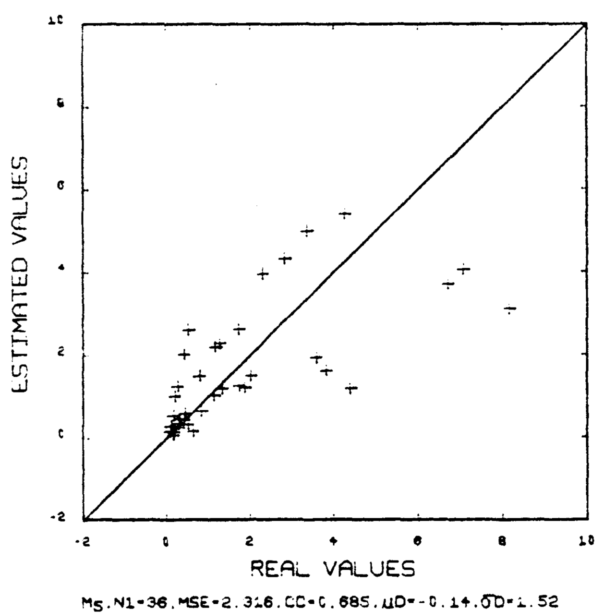
(b = 0.007, σ = 10, N = 27)



d



e



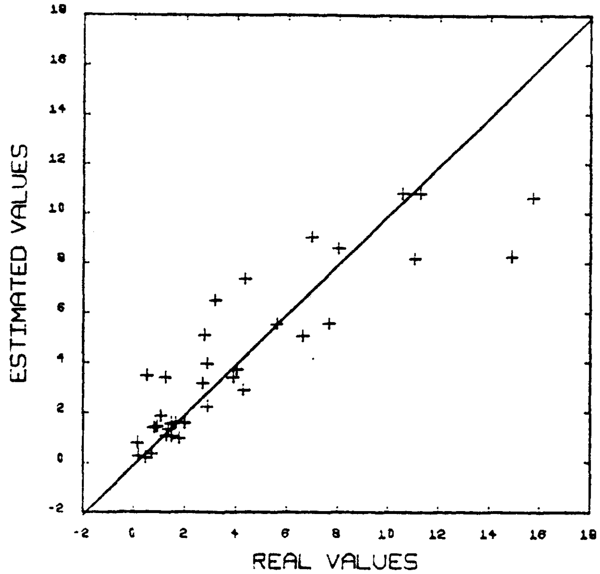
f

Figure 5.5

Block Estimation Comparisons. Transformed Isotropic Field

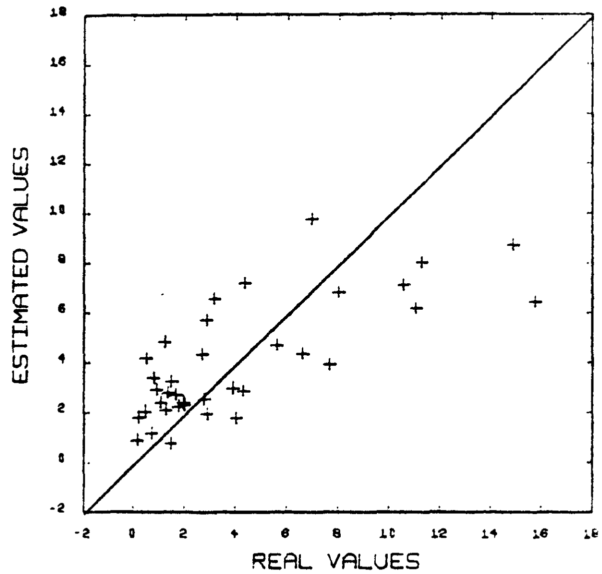
(b = 0.007, σ = 10, N = 27)

(continued)



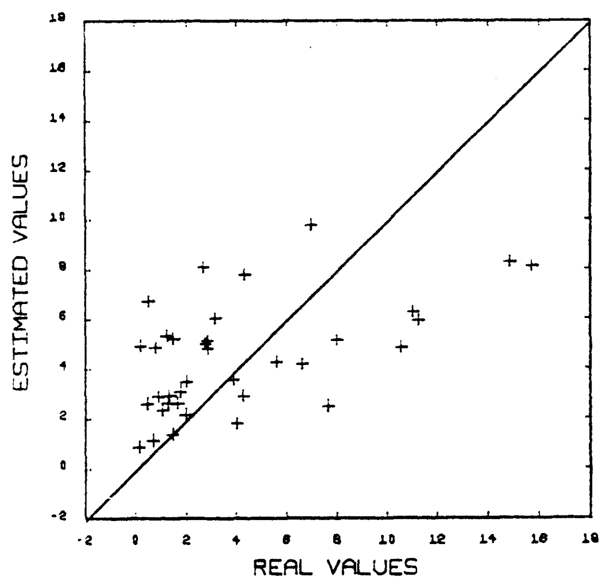
UK.R.k=0.$\hat{\varsigma}$=1.19.N1=36.MSE=3.742.MPV=15.34
MePV=25.01.CC=0.886.μD=0.004.$\bar{\sigma}$D=1.93.$\hat{\varsigma}_1$=0.24

a



UK.M.k=1.$\hat{\varsigma}$=1.02.N1=36.MSE=8.070.MPV=16.89
MePV=22.21.CC=0.740.μD=0.00. $\bar{\sigma}$D=2.84.$\hat{\varsigma}_1$=0.48

b



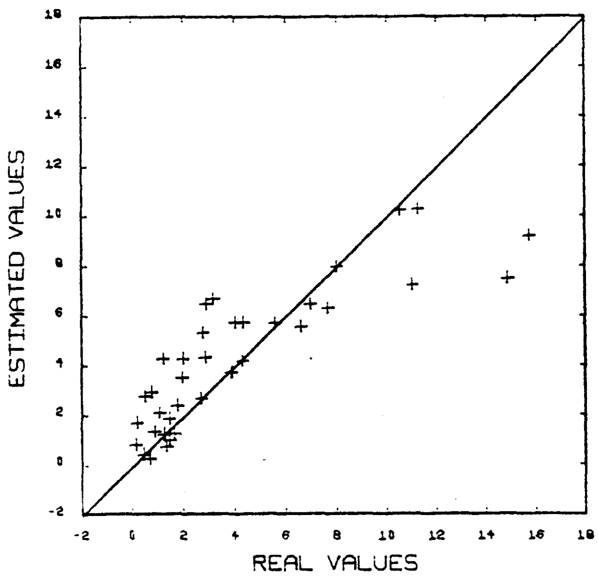
UK.D.k=0.$\hat{\varsigma}$=1.00.N1=36.MSE=12.09.MPV=18.33
MePV=20.34.CC=0.536.μD=0.421.$\bar{\sigma}$D=3.45.$\hat{\varsigma}_1$=0.66

c

Figure 5.6

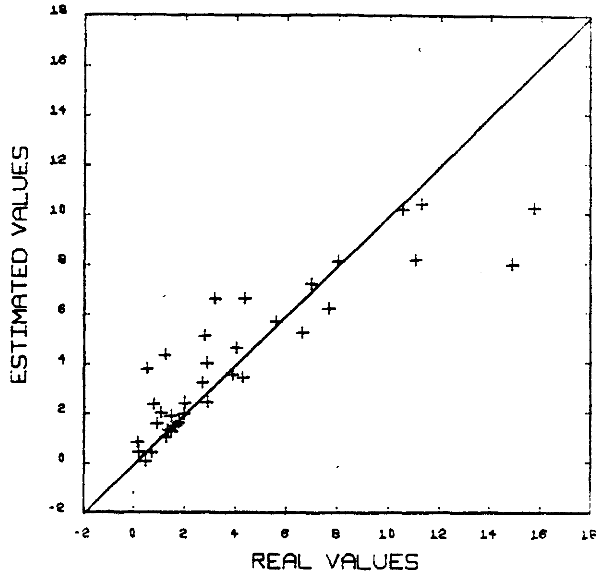
Block Estimation Comparisons. Transformed Isotropic Field

(b = 0.0141, σ = 10, N = 48)



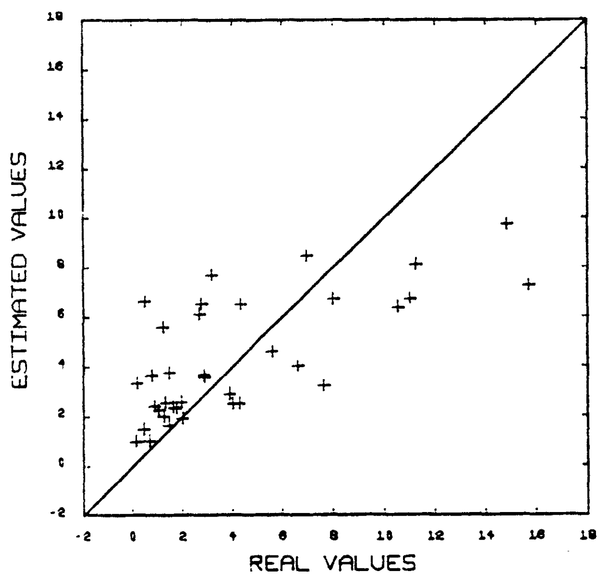
DK.L.%μ=3.20.%V=5.20.N1=36.MSE=5.176.MPV=14.78
M_sPV=22.31.CC=0.850.μ_D=0.229.σ_D=2.26.$\hat{q}_1$=0.35

d



DK.S.%μ=3.20.%V=5.20.N1=36.MSE=3.940.MPV=9.253
M_sPV=15.11.CC=0.889.μ_D=0.073.σ_D=1.98.$\hat{q}_1$=0.43

e



M_5.N1=36.MSE=8.923.CC=0.694.μ_D=0.173.σ_D=2.98

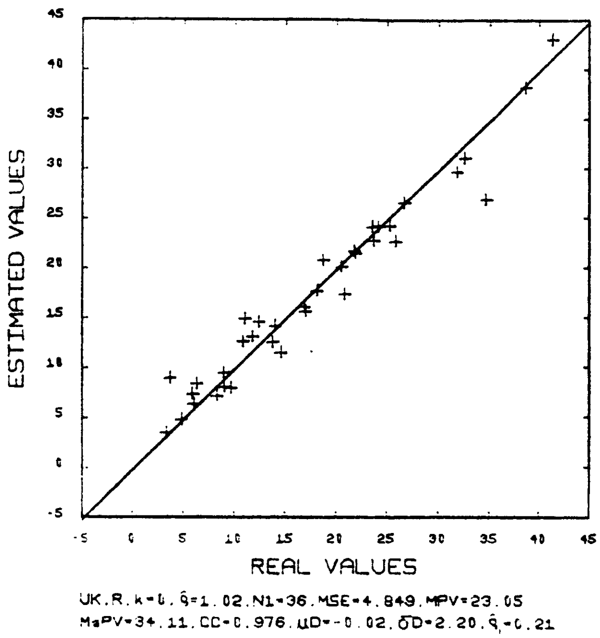
f

Figure 5.6

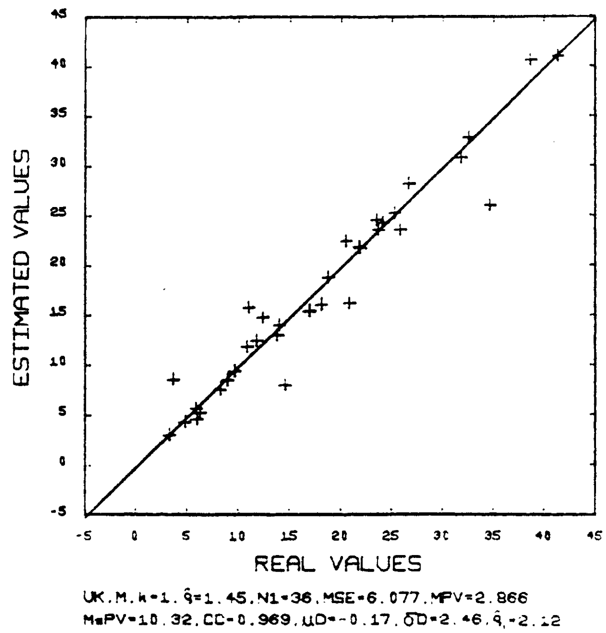
Block Estimation Comparisons. Transformed Isotropic Field

(b = 0.0141, σ = 10, N = 48)

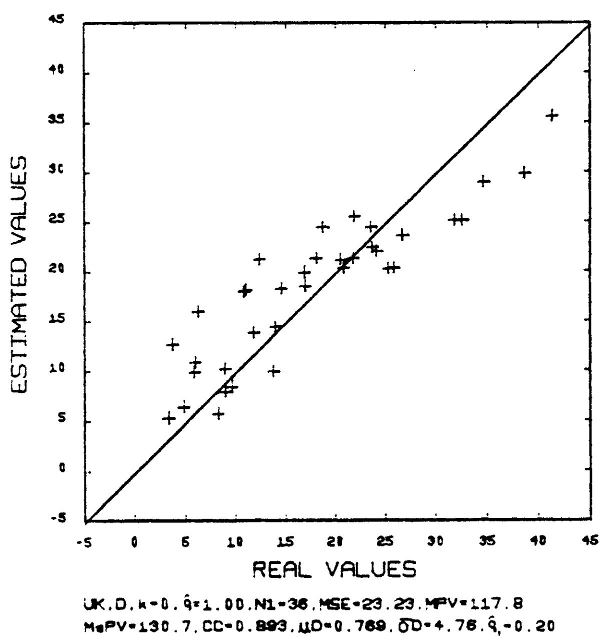
(continued)



a



b

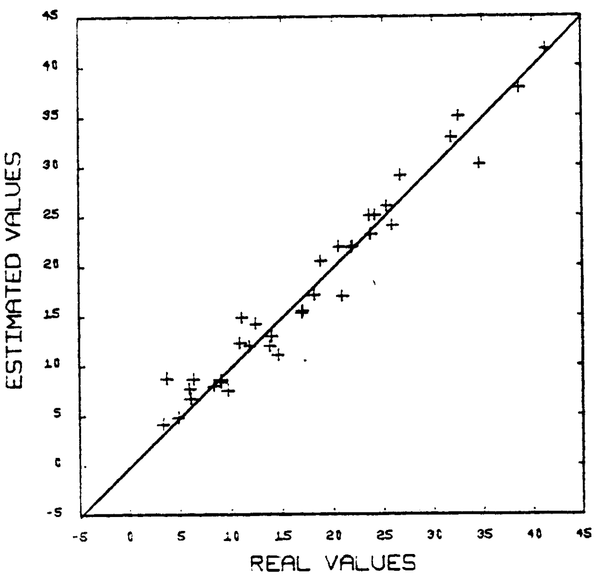


c

Figure 5.7

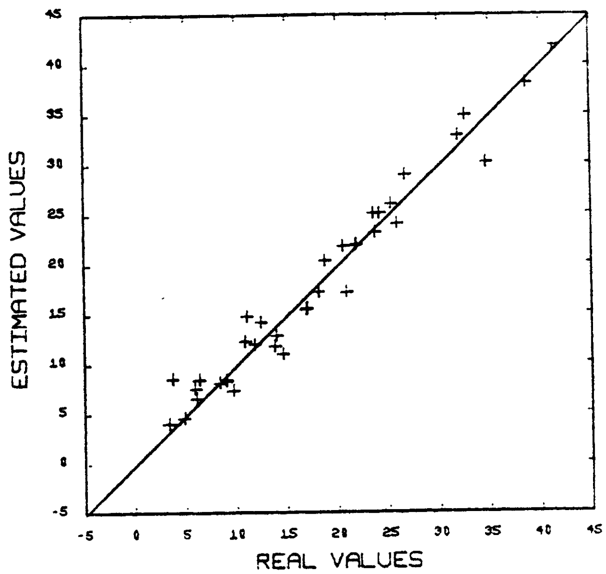
Block Estimation Comparisons. Transformed Isotropic Field

(b = 0.0035, σ = 20, N = 47)



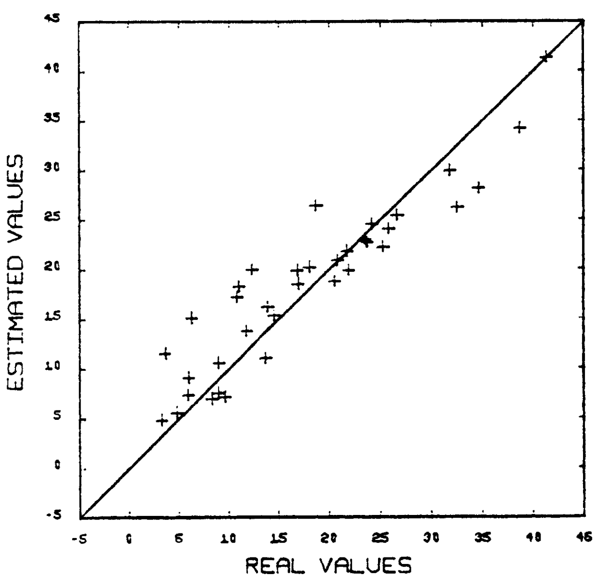
DK.L.%μ=1.50.%V=5.10.N1=36.MSE=4.016.MPV=28.57
MaPV=42.37.CC=0.980.μD=0.286.σD=1.98.$\hat{\varsigma}$=-0.14

d



DK.S.%μ=1.50.%V=5.10.N1=36.MSE=3.850.MPV=46.28
MaPV=69.07.CC=0.981.μD=0.273.σD=1.94.$\hat{\varsigma}$=-0.083

e



M_5.N1=36.MSE=14.94.CC=0.930.μD=0.796.σD=3.78

f

Figure 5.7

Block Estimation Comparisons. Transformed Isotropic Field

(b = 0.0035, σ = 20, N = 47)

(continued)

	MSE			MPV		
	(10%)	(25%)	(50%)	(10%)	(25%)	(50%)
K_M	15	15	15	14	15	15
DK_s	4	4	6	3	4	4
M_5	0	2	3			

CONSISTENCY

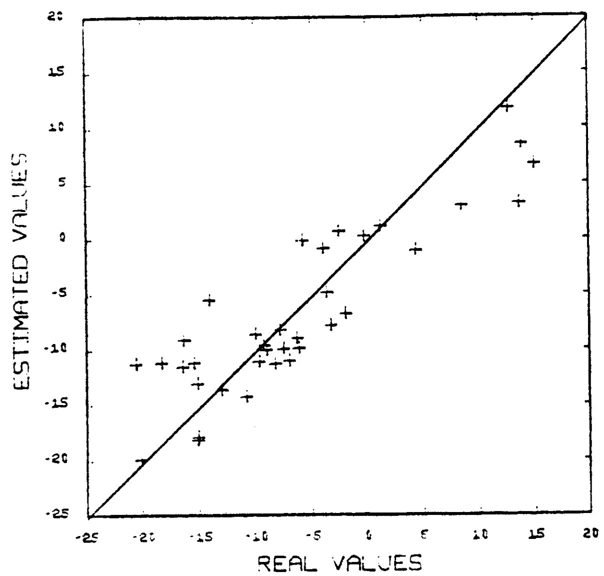
$$(\hat{\rho}_1 = \frac{MSE}{MPV})$$

	Min	Max	$\# \leq 1$	$\# > 1$	# [0.8,1.2]	μ	σ
K_M	0.11	41.5	10	6	2	1.2	1.47
DK_s	0.04	1.08	15	1	1	0.33	0.3

Table 5.9

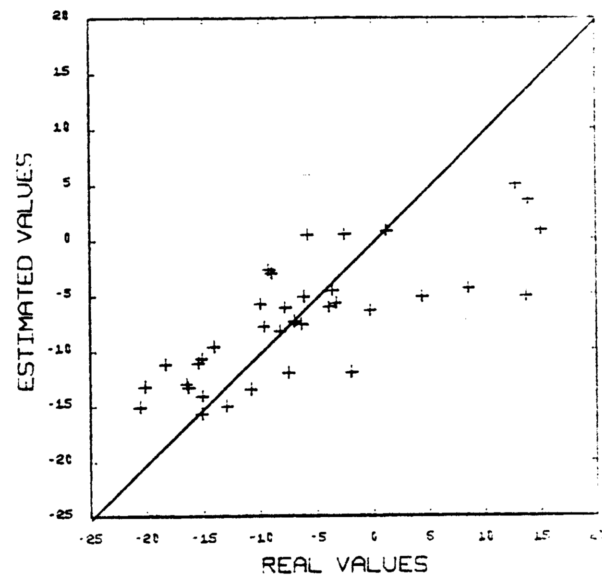
Comparisons of the Different Estimators

Intrinsic Random Functions (16 cases) Block Results



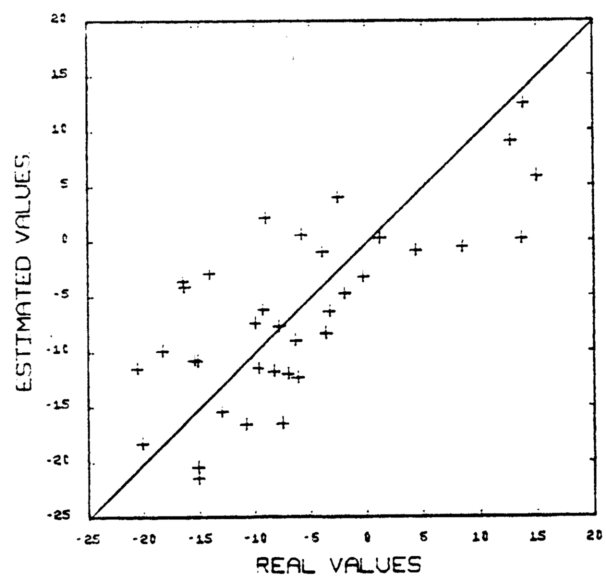
UK,R-M,k=0,ĝ=0.98,N1=36,MSE=20.96,MPV=46.10
MaPV=73.69,CC=0.887,μD=-0.26,σD=4.57,ĝ=0.45

a



UK,D,k=0,ĝ=1.00,N1=36,MSE=41.85,MPV=86.21
MaPV=95.66,CC=0.773,μD=-0.74,σD=6.43,ĝ=0.49

b



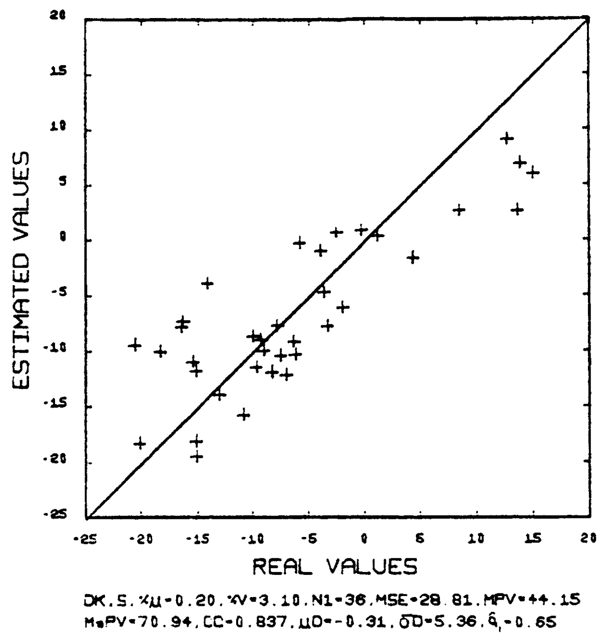
OK,L,%μ=0.20,%V=3.10,N1=36,MSE=43.79,MPV=44.66
MaPV=78.94,CC=0.732,μD=-0.06,σD=6.62,ĝ=-0.98

c

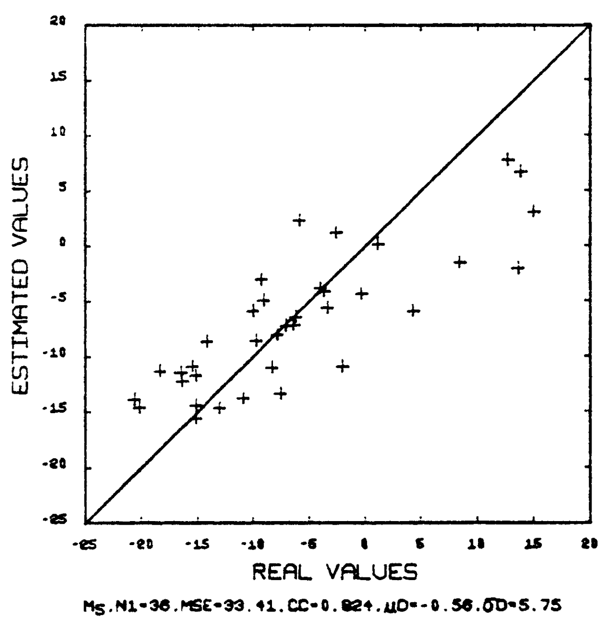
Figure 5.8

Block Estimation Comparisons. Intrinsic Random Function

($k = 0$, $\alpha_1 = -3$, $\alpha_3 = \alpha_5 = 0$, $N = 45$)



d



e

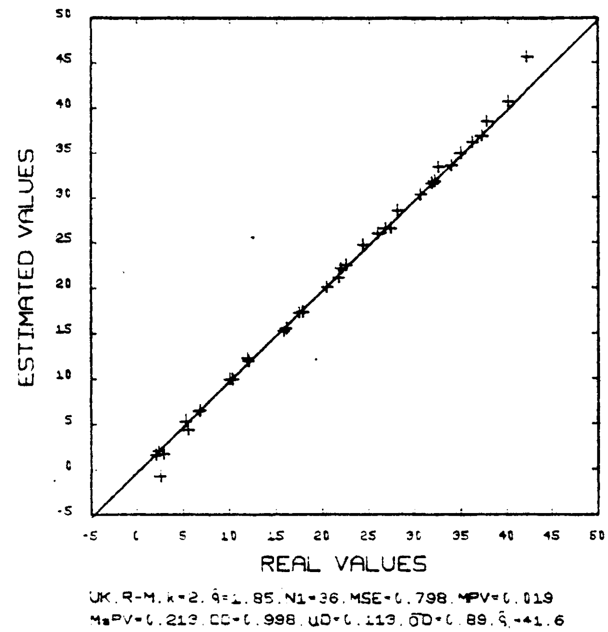
Figure 5.8

Block Estimation Comparisons. Intrinsic Random Function

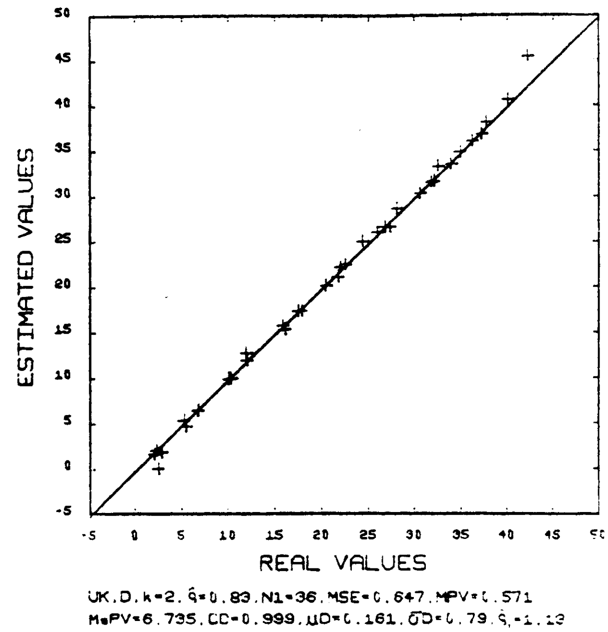
($k = 0$, $\alpha_1 = -3$, $\alpha_3 = \alpha_5 = 0$, $N = 45$)

(continued)

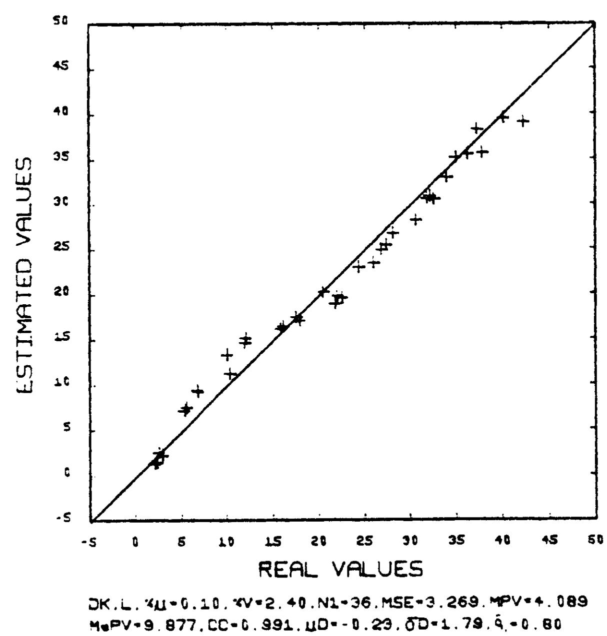
and 5.12 b. The consistency parameter show the tendency of overestimating the predicted variance by both methods, with Disjunctive Kriging overestimating more than Universal Kriging. See Figures 5.9 a, d, e; 5.10 a, c, d; 5.11 a, d, e; and 5.12 b, d, e.



a



b

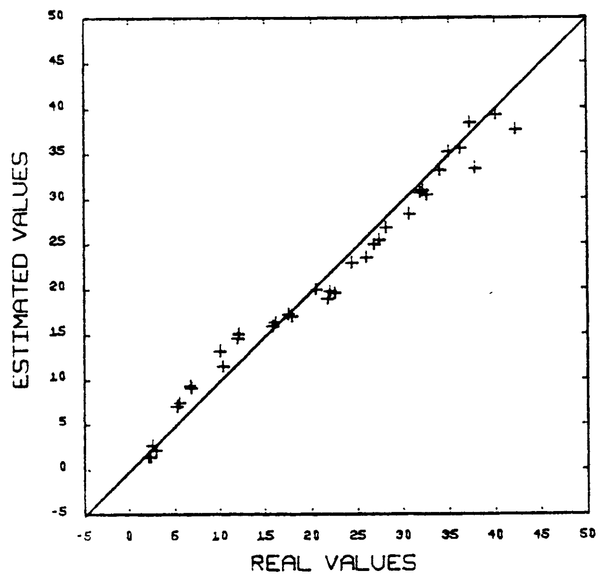


c

Figure 5.9

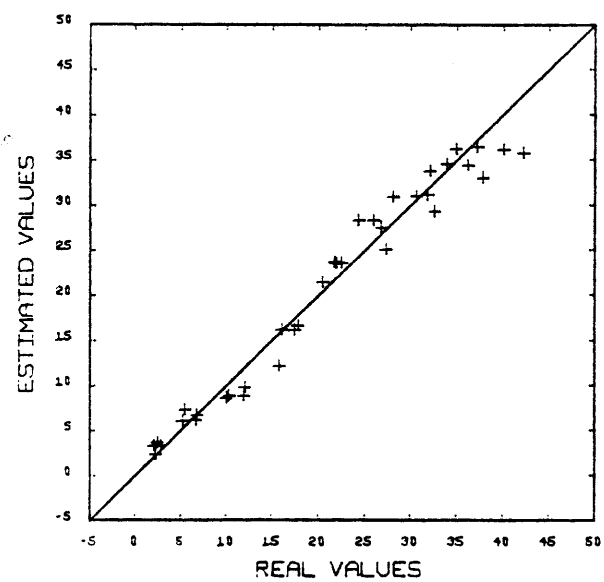
Block Estimation Comparisons. Intrinsic Random Function

($k = 1$, $\alpha_1 = -0.005$, $\alpha_3 = 0.005$, $\alpha_5 = 0$, $N = 28$)



DK.S.%μ=0.10.%V=2.40.N1=36.MSE=3.950.MPV=15.34
M.PV=35.56.CC=0.990.μD=-0.36.σD=1.95.$\hat{\sigma}_1$=0.26

d



M_5.N1=36.MSE=5.144.CC=0.983.μD=-0.34.σD=2.24

e

Figure 5.9

Block Estimation Comparisons. Intrinsic Random Function

($k = 1$, $\alpha_1 = -0.005$, $\alpha_3 = 0.005$, $\alpha_5 = 0$, $N = 28$)

(continued)

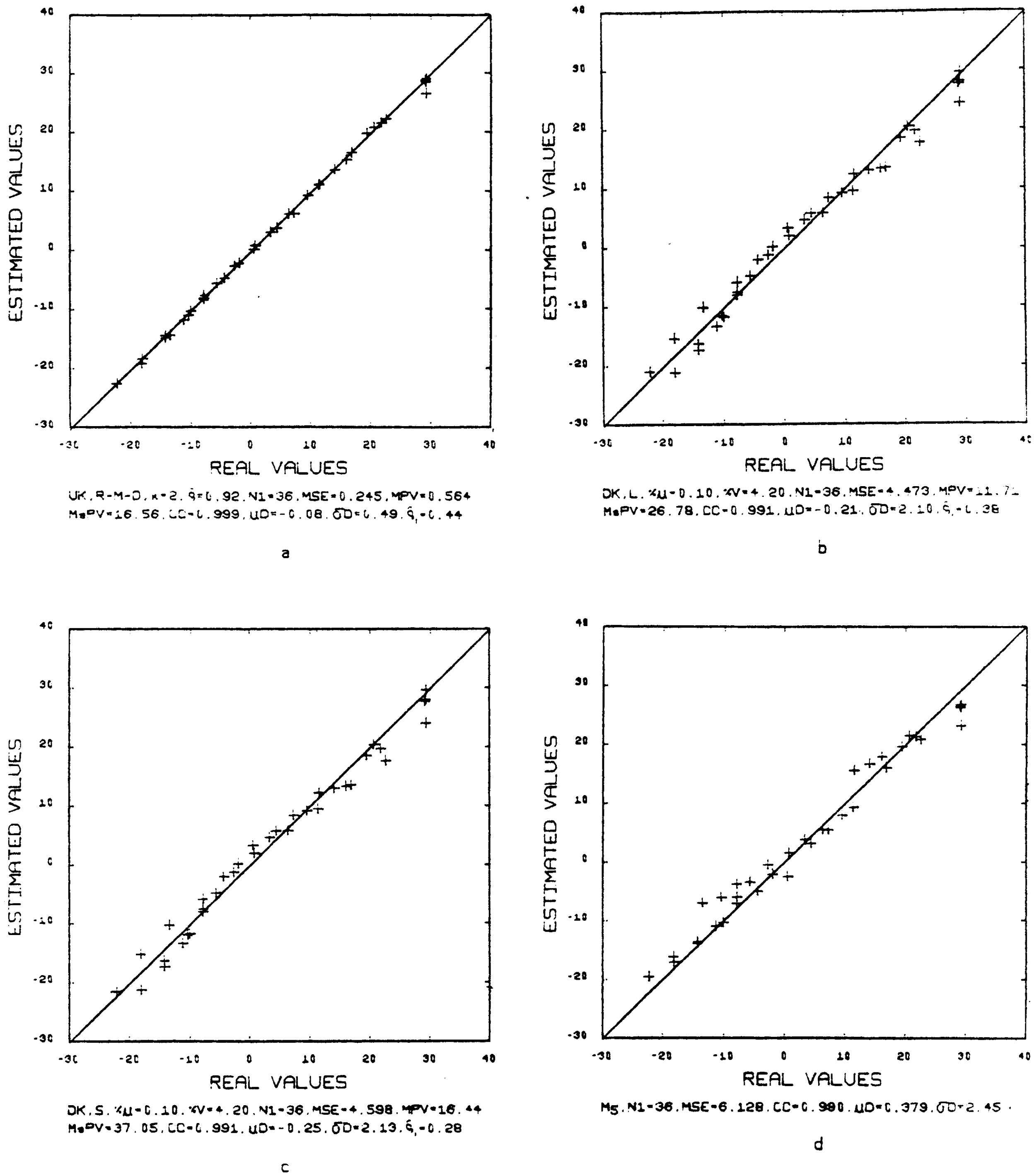
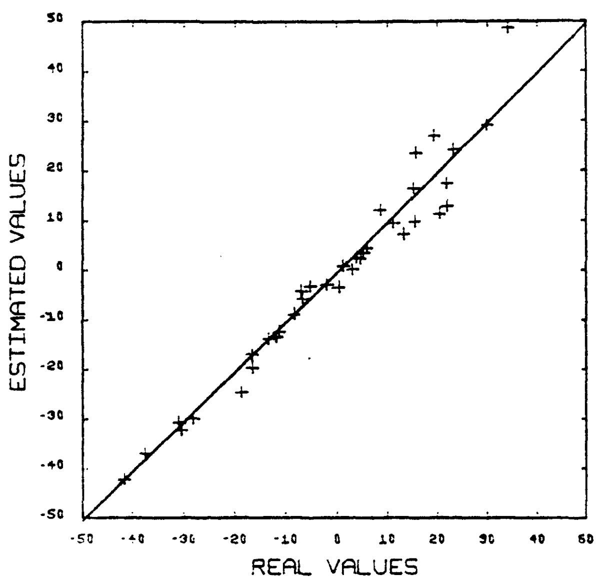


Figure 5.10

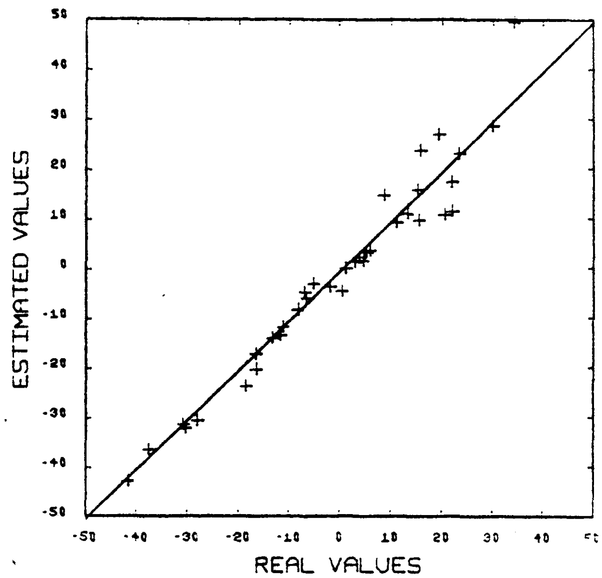
Block Estimation Comparisons. Intrinsic Random Function

($k = 2$, $\alpha_1 = 0$, $\alpha_3 = 0.005$, $\alpha_5 = -0.5\times10^{-8}$, $N = 28$)



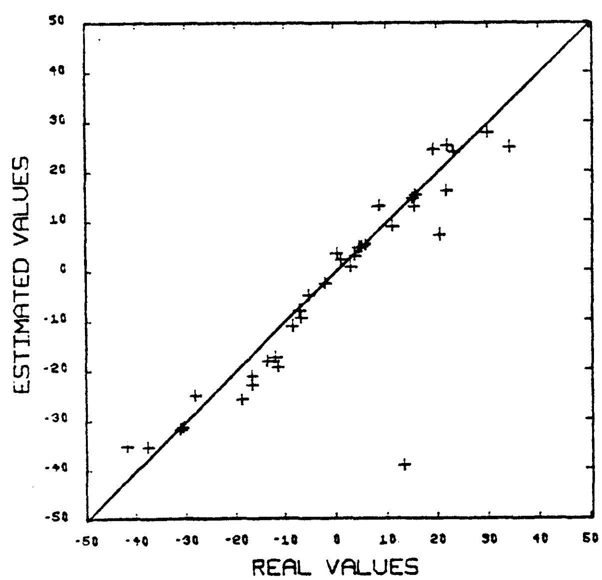
UK.R-M.k=2.$\hat{q}$=1.78.N1=36.MSE=19.05.MPV=17.08
MsPV=40.37.CC=0.976.μD=-0.28.$\bar{\sigma}$D=4.36.$\hat{q}_1$=1.12

a



UK.D.k=2.$\hat{q}$=1.30.N1=35.MSE=14.19.MPV=34.96
MsPV=109.2.CC=0.980.μD=-0.67.$\bar{\sigma}$D=3.71.$\hat{q}_1$=0.41

b



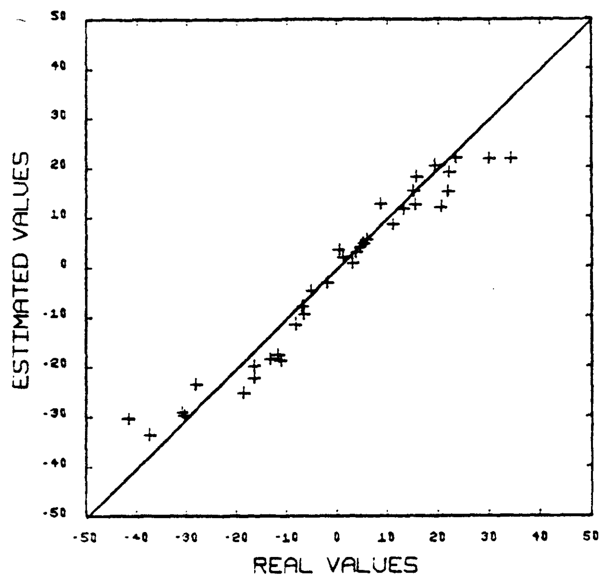
DK.L.*μ=2.90.*V=1.90.N1=36.MSE=93.83.MPV=82.84
MsPV=656.8.CC=0.883.μD=-2.72.$\bar{\sigma}$D=9.30.$\hat{q}_1$=1.13

c

Figure 5.11

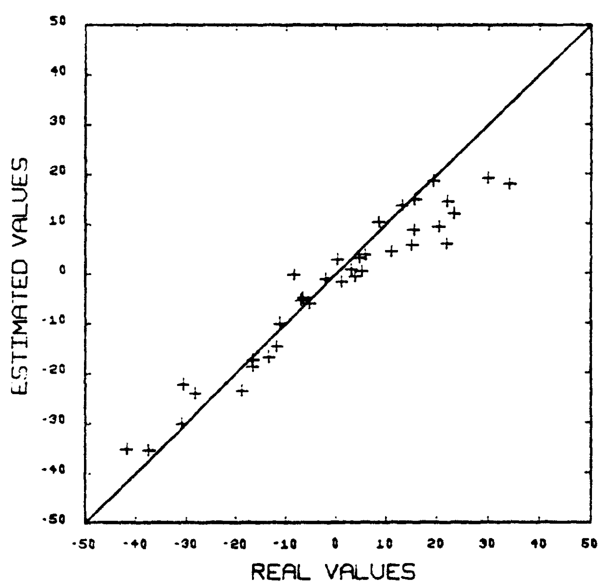
Block Estimation Comparisons. Intrinsic Random Function

($k = 2$, $\alpha_1 = -1$, $\alpha_3 = 0$, $\alpha_5 = -0.5 \times 10^{-9}$, $N = 29$)



DK.S.%μ=2.90.%V=1.90.N1=36.MSE=21.07.MPV=183.6
MsPV=302.8.CC=0.975.μD=-1.33.σD=4.39.ρ=-0.11

d



Ms.N1=36.MSE=39.35.CC=0.966.μD=-2.38.σD=5.80

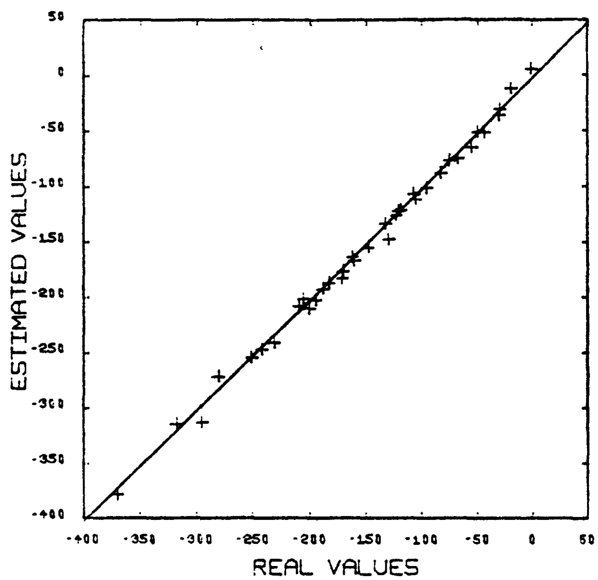
e

Figure 5.11

Block Estimation Comparisons. Intrinsic Random Function

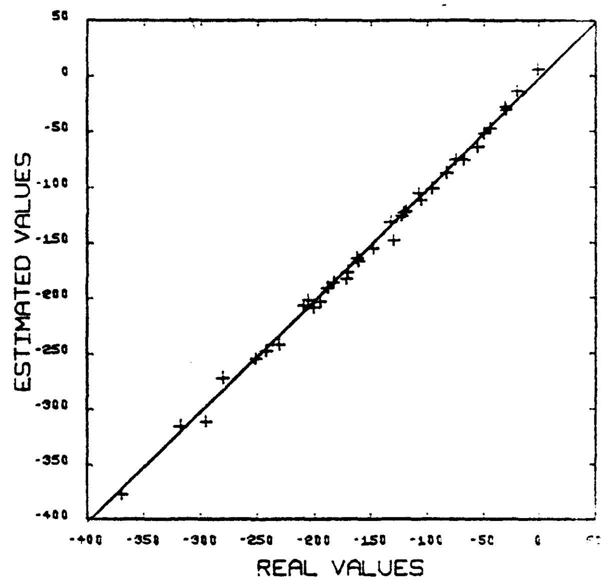
($k = 2$, $\alpha_1 = -1$, $\alpha_3 = 0$, $\alpha_5 = -0.5\text{x}10^{-9}$, $N = 29$)

(continued)



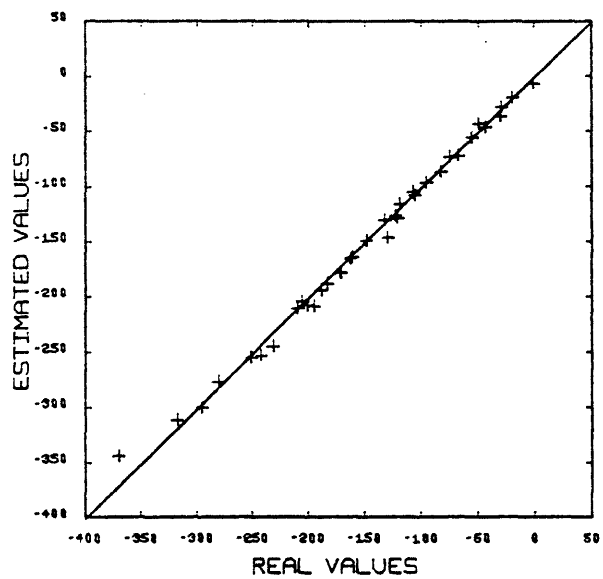
UK.R-D.k=2.$\hat{q}$=0.95.N1=36.MSE=38.56.MPV=48.61
MePV=248.3.CC=0.998.μD=-1.74.$\bar{\sigma}$D=5.96.$\hat{q}_1$=0.79

a



UK.M.k=2.$\hat{q}$=2.03.N1=36.MSE=35.08.MPV=7.483
MePV=143.1.CC=0.998.μD=-1.26.$\bar{\sigma}$D=5.79.$\hat{q}_1$=4.69

b



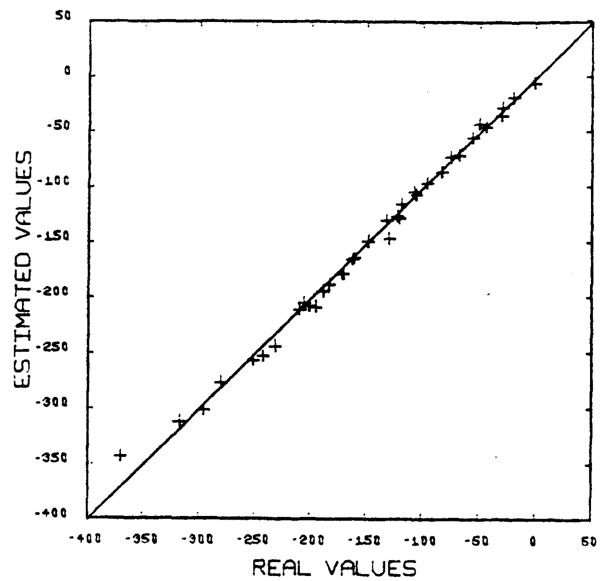
DK.L.%μ=0.20.%V=3.60.N1=36.MSE=52.23.MPV=256.4
MePV=370.0.CC=0.997.μD=-1.07.$\bar{\sigma}$D=7.15.$\hat{q}_1$=0.20

c

Figure 5.12

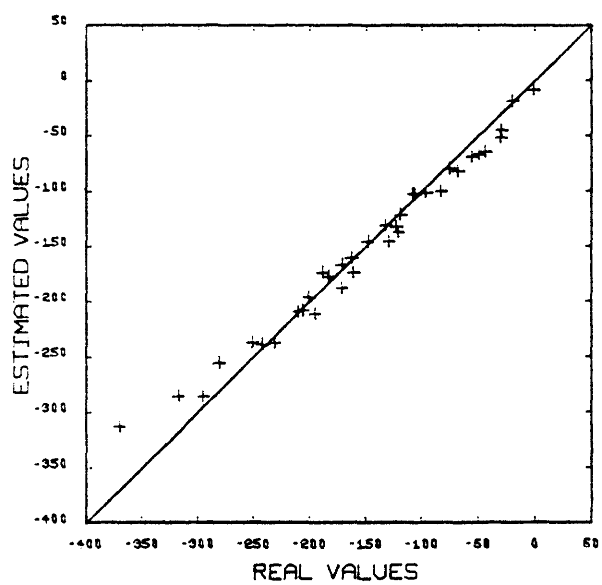
Block Estimation Comparisons. Intrinsic Random Function

($k = 2$, $\alpha_1 = -2$, $\alpha_3 = 0.05$, $\alpha_5 = -1\text{x}10^{-9}$, $N = 45$)



DK.S.%μ=0.20.%V=3.60.N1=36.MSE=52.46.MPV=865.9
M=PV=1252. CC=0.997.μD=-1.03.σD=7.17.ς=0.061

d



M5.N1=36.MSE=253.8.CC=0.991.μD=-1.08.σD=15.9

e

Figure 5.12

Block Estimation Comparisons. Intrinsic Random Function

($k = 2$, $\alpha_1 = -2$, $\alpha_3 = 0.05$, $\alpha_5 = -1 \times 10^{-9}$, $N = 45$)

(continued)

5.5 Summary

It has been shown that the Universal Kriging estimator gives better results when the generalized covariance function is identified using the least mean square error or the ranking procedures.

Disjunctive Kriging did not show big differences when calculated from spherical or linear semivariograms of the Gaussian data. For the sake of comparison, the spherical semivariogram was used in Disjunctive Kriging and the least mean square error procedure in Universal Kriging.

The Disjunctive Kriging estimator was found as good or better than the Universal Kriging technique only in estimating transformed isotropic fields. In all other cases Universal Kriging gave better results both in mean square error and mean predicted variance, contradicting the theoretical expectations.

In contrast with point estimation comparisons, both Kriging techniques tended to overestimate the predicted block variance consistently. Some experiments were made calculating the block integrals from summations of different sizes, but no different tendencies were found. The contrast could be explained by the fact that using the eight neighboring points of the center of each block does not capture well the field, especially near the block boundaries where other neighboring points are expected to perform better.

Chapter 6

SUMMARY, CONCLUSIONS, AND FURTHER RESEARCH

6.1 Summary and Conclusions

This work deals with the Linear and Disjunctive Kriging estimators of a random phenomena. Taking into account the spatial characteristics of a given realization, both techniques provide systematic procedures to find estimates at specified locations (point values), as well as areal averages (block values).

The relevant theory of stochastic processes and the Linear Kriging estimator characteristics are given in Chapter 2. The Kriging equations together with the optimal estimation variance are given for both the point and block cases under different assumptions about the underlying field. It is shown how the intrinsic random functions theory generalize the intrinsic hypothesis and how the Universal Kriging equations can be written in terms of the generalized covariance function, which has the advantage over the covariance function of filtering out the form of the drift. Practical models of generalized covariances as well as their estimation procedures are also presented in Chapter 2.

In Chapter 3 the Disjunctive Kriging equations for both point and block estimation are presented. From the assumption that the field comes from a second-order stationary Gaussian field, via an anamorphosis function, it is shown how the original need of bivariate distributions of the field is reduced to the covariance structure of the Gaussian variables. It is seen how Hermite expansions of the anamorphosis and unknown functions give an infinite set of simultaneous equations for the optimal estimate, and how this infinite set can be reduced to a finite one by checking at finite expansion characteristics of the anamorphosis function. The chapter ends with an overview of the possible estimators and discusses the theoretical higher accuracy of the Disjunctive Kriging estimator.

Chapter 4 compares the performance of the various point estimators under "real world" situations. It is shown that the Universal Kriging estimator gives better results when its generalized covariance function is calculated from least mean square error and ranking procedures in contrast with the less expensive Delfiner's methodology (1976). The Disjunctive Kriging estimator is better when a spherical semivariogram, versus a linear one is used to model the Gaussian data. Both Kriging estimators give better point estimates than a local mean obtained with the five nearest data points to a given location.

The Universal Kriging estimator gives as good or better results than the Disjunctive Kriging estimator which contradicts theoretical expectations. The reasons could be the use of finite expansions in the estimation of the unknown and anamorphosis functions, and the fact that small samples (as the ones usually found in hydrological applications) lead to inconsistencies in the Hermitian model, necessary to have a Disjunctive Kriging solution. Universal Kriging gave particularly better results when the data came from Intrinsic Random Functions of order 1 and 2, this is explained because Universal Kriging explicitly tracks the spatial varying drift while Disjunctive Kriging does not. Although better, it was found that Universal Kriging consistently underestimated the predicted variance of estimation; while Disjunctive Kriging, which tended to overestimate with Intrinsic Random Functions of orders 1 and 2, gave slightly more consistent agreement between true and predicted variances.

The results of comparing the block estimation performance of the different estimators are presented in Chapter 5. As in point estimation comparisons, Universal Kriging gives better results when its generalized covariance function is calculated using the least mean square error and ranking procedures. Disjunctive Kriging gives about the same results when the spherical or linear semivariograms of the Gaussian data are used. As previously found, both Kriging estimators give better results than a local mean estimator. A comparison of Universal Kriging with Disjunctive Kriging show as good or better results with the former. Contrasting with point estimation results, both block Kriging estimators tend to overestimate the predicted variances. This was found independent of the number of points used in the approximation of the block integrals and could be explained by the lack of accuracy the estimates of points in the boundaries of the blocks have.

6.2 Possible Further Research

As the experiments show, good estimated values do not necessarily imply correct predicted variances. Then, networks designed with these variances may give more or less stations than needed in reality. This stresses the multi-realization approach when performing network design as opposed to a single-realization approach. In any case, research should move toward the improvement of the Kriging estimators, under small samples at hand, to get more consistent results from which properly based station placement decisions could be made.

The Kriging estimators were calculated in this work under fixed conditions such as five Hermite polynomials, eight neighboring points, fixed block sizes, and so on. Research should be directed to study the sensitivity in the estimators due to changes in those parameters.

The issue of how big the data set should be to avoid in practice the small sample inconsistencies of the Hermitian model, used in the Disjunctive Kriging estimator, should be also further investigated.

REFERENCES

Armstrong, M. and R. Jabin, "Variogram Models Must be Positive-Definite", Journal of Mathematical Geology, 13(5), pp. 455-459, 1981.

Cressie, N. and D. M. Hawkins, "Robust Estimation of the Variogram: I", Journal of Mathematical Geology, 12(2), pp. 115-125, 1980.

Chua, S. H. and R. L. Bras, "Estimation of Stationary and Non-Stationary Fields: Kriging in the Analysis of Orographic Precipitation", Ralph M. Parsons Laboratory for Water Resources and Hydrodynamics, MIT, Report No. 255, 171 pp., 1980.

David, M., "The Practice of Kriging", in Advanced Geostatistics in the Mining Industry, edited by M. Guarascio et al., pp. 31-48, D. Reidel, Dordrecht-Holland/Boston - USA, 1976.

David, M., "Geostatistical Ore Resource Estimation", Elsevier Scientific Publishing Company, New York, 364 pp., 1977.

Davis, M. and M. David, "Automatic Kriging and Contouring in the Presence of Trends (Universal Kriging Made Simple)", Journal of Canadian Petroleum Technology, pp. 90-98, 1978.

Delfiner, P. and J. P. Delhomme, "Optimum Interpolation by Kriging", in Display and Analysis of spatial data, J. C. Davis and M. J. McCullagh, eds., pp. 96-114, John Wiley & Sons, 1973.

Delfiner, P., "Linear Estimation of Non Stationary Spatial Phenomena", in Advanced Geostatistics in the Mining Industry, edited by M. Guarascio et al., pp. 49-68, D. Reidel, Dordrecht-Holland/Boston - USA, 1976.

Delfiner, P. and J. P. Chiles, "Conditional Simulations: A New Monte Carlo Approach to Probabilistic Evaluation of Hydrocarbon in Place", Ecole des Mines de Paris, Fontainebleau, 1978.

Delhomme, J. P. and P. Delfiner, "Application du Krigeage à l'optimisation d'une Campagne Pluviométrique en Zone Aride", in Proceedings of the Symposium of the Design of Water Resource Projects with Inadequate Data, (2), pp. 191-210, Unesco, Madrid, 1973.

Delhomme, J. P., "Applications de la Théorie des Variables Regionalisées dans les Sciences de l'eau", Doctoral thesis, 161 pp., Ecole des Mines, Fontainebleau, France, 1976.

Delhomme, J. P., "Kriging in the Hydrosciences", Advances in Water Resources, 1(5), pp. 251-256, 1978.

Delhomme, J. P., Spatial Variability and Uncertainty in Groundwater Flow Parameters: A Geostatistical Approach", Water Resources Research, 15(2), pp. 269-280, 1979.

Doctor, P. G., "An Evaluation of Kriging Techniques for High Level Radioactive Waste Repository Site Characterization", Pacific Northwest Laboratory, PNL-2903, UC-70, 67 pp., 1979.

Doctor, P.G. and R. W. Nelson, "Geostatistical Estimation of Parameters for Hydrological Transport Modeling", Journal of Mathematical Geology, 13(5), pp. 415-428, 1981.

Gambolati, G. and G. Volpi, "Groundwater Contour Mapping in Venice by Stochastic Interpolators. 1. Theory", Water Resources Research, 15 (2), pp. 281-290, 1979a.

Gambolati, G. and G. Volpi, "A Conceptual Deterministic Analysis of the Kriging Technique in Hydrology", Water Resources Research, 15(3), pp. 625-629, 1979b.

Gandin, L. S., "The Planning of Meteorological Station Newtorks", in Distribution of Precipitation in Montainous Areas, Vol. I, Proceedings Of Geilo Symposium, WMO Technical Note 111, 25 pp., Geneva, 1970.

Hannan, E. J., "Multiple Time Series", John Wiley & Sons, New York, 1970.

Huijbregts, Ch. and G. Matheron, "Universal Kriging: An Optimum Method for Estimating and Contouring in Trend-Surface Analysis", CIMM International Symposium, pp. 159-160, Montreal, 1970.

Huijbregts, Ch. J., "Regionalized Variables and Application to Quantitative Analysis of Spatial Data", in Display and Analysis of Spatial data, J. C. Davis and M. J. McCullagh, eds., pp. 38-53, John Wiley & Sons, 1973.

Hughes, J. P. and D. P. Lettenmaier, "Aquatic Monitoring: Data Analysis and Network Design Using Regionalized Variable Theory", Charles W. Harris Hydraulics Laboratory, University of Washington, Technical Report No. 65, 182 pp., 1980.

Hughes, J. P. and D. P. Lettenmaier, "Data Requirements for Kriging Estimation and Network Design", Water Resources Research, 17(6), pp. 1641-1650, 1981.

Journel, A. G ., "Geostatistics for Conditional Simulation of Ore Bodies", Economic Geology, 69, pp. 673-687, 1974.

Journel, A. G., "Kriging in Terms of Projections", Journal of Mathematical Geology, 9(6), pp. 563-586, 1977.

Journel, A. G. and Ch. J. Huijbregts, "Mining Geostatistics", Academic Press, London, 600 pp., 1978.

Journel, A. G., "The Lognormal Approach to Predicting Local Distributions of Selective Mining Unit Grades", Journal of Mathematical Geology, 12(4), pp. 285-303, 1980.

Kafritsas, J. and R. L. Bras, "The Practice of Kriging", Ralph M. Parsons Laboratory for Water Resources and Hydrodynamics, MIT, Report No. 263, 107 pp., 1981.

Karlinger, M. and J. Shrivan, "An Application of Kriging to a Groundwater Quality Network", Paper Presented at American Geophysical Union Chapman Conference on Network Design, Tucson, 1978.

Lenton, R. and I. Rodriguez-Iturbe, "On the Collection, the Analysis, and the Synthesis of Spatial Rainfall Data", Ralph M. Parsons Laboratory for Water Resources and Hydrodynamics, MIT, Report No. 194, 218 pp., 1974.

Mantoglou, A. and J. L. Wilson, "Simulation of Random Fields with the Turning Bands Method", Ralph M. Parsons Laboratory for Water Resources and Hydrodynamics, MIT, Report No. 264, 199 pp., 1981.

Maréchal, A., "Selecting Mineable Blocks: Experimental Results Observed on a Simulated Orebody", in Advanced Geostatistics in the Mining Industry, edited by M. Guarascio et al., pp. 137-161, D. Reidel, Dordrecht-Holland/Boston - USA, 1976a.

Maréchal, A., "The Practice of Transfer Functions: Numerical Methods and Their Application", in Advanced Geostatistics in the Mining Industry, edited by M. Guarascio et al., pp. 253-276, D. Reidel, Dordrecht-Holland/Boston - USA, 1976b.

Matern, B., "Spatial Variation", Almaenna Föerlaget, Stockholm, 1960.

Matheron, G., "The Theory of Regionalized Variables and its Applications", Les Cahiers du Centre de Morphologie Mathématique, Fasc. 5, Centre de Géostatistique, Fontainebleau, 1971.

Matheron, G., "The Intrinsic Random Functions and Their Applications", Advanced in Applied Probability, 5, pp. 439-468, 1973.

Matheron, G., "A Simple Substitute for Conditional Expectation: The Disjunctive Kriging", in Advanced Geostatistics in the Mining Industry, edited by M. Guarascio et al., pp. 221-236, D. Reidel, Dordrecht-Holland/Boston - USA, 1976a.

Matheron, G., "Forecasting Block Grade Distributions: The Transfer Functions", in Advanced Geostatistics in the Mining Industry, edited by M. Guarascio et al., pp. 237-251, D. Reidel, Dordrecht-Holland/Boston - USA, 1976b.

McLaughlin, D. B., "Hanford Groundwater Modeling - Statistical Methods for Evaluating Uncertainty and Assessing Sampling Effectiveness", Rockwell Hanford Operations, RHO-C-18, 75 pp., 1979.

Olea, R. A., "Optimum Mapping Techniques Using Regionalized Variable Theory", Series on Spatial Analysis, No. 2, Kansas Geological Survey, 137 pp., 1975a.

Olea, R. A., "Measuring Spatial Dependence with Semivariograms", Series on Spatial Analysis, No. 3, Kansas Geological Survey, 26 pp, 1975b.

Puente, C. E. and R. L. Bras, "Disjunctive Kriging, Universal Kriging or no Kriging: What are the Answers", paper presented at International Conference on Time Series Methods in Hydrosciences, Burlington, Ontario, October 1981.

Rendu, J. M., "Disjunctive Kriging: Comparison of Theory with Actual Results", Journal of Mathematical Geology, 12(4), pp. 305-320, 1980.

Rodriguez-Iturbe, I. and J. M. Mejia, "The Design of Rainfall Networks in Time and Space", Ralph M. Parsons Laboratory for Water Resources and Hydrodynamics, MIT, Report No. 176, 77 pp., 1973.

Rouhani, S., "Application of the Variance Reduction to the Universal Kriging", Division of Applied Sciences, Harvard University, 1981.

Schaake, J., "Accuracy of Point and MAP Estimates from Point Precipitation Data (Draft)", NOAA Technical Memo, NWS-Hydro, Washington, D.C., 1978.

Shaw, E. M. and P. E. O'Connell, "Design of Networks and Data Transfer for Precipitation", Hydrological Network Design and Information Transfer, WMO Operational Hydrology Report No. 8, WMO Report No. 433, Geneva, 1976.

Starks, T. H. and J. H. Fang, "On the Estimation of the Generalized Covariance Function", Journal of Mathematical Geology, 14(1), pp. 57-64, 1982.

Veneziano, D. A., "Random Processes for Engineering Applications", Course Notes, MIT, 1978.

Villeneuve, J. P., Morin G., Bobee B., and D. Leblanc, "Kriging in the Design of Streamflow Sampling Networks", Water Resources Research, 15(6), pp. 1833-1840, 1979.

Volpi, G. and G. Gambolati, "On the Use of a Main Trend for the Kriging Technique in Hydrology", Advances in Water Resources, 1(6), pp. 345-349, 1978.

Volpi, G., Gambolati, G., Carbognin L., Gatto P., and G. Mozzi, "Groundwater Contour wapping in Venice by Stochastic Interpolators. 2. Results", Water Resources Research, 15(2), pp. 291-297, 1979.

Appendix A

HERMITIAN EXPANSION OF THE LINEARLY INTERPOLATED ANAMORPHOSIS

Let $Z_1, Z_2, \ldots, Z_N$ be the ordered data observations and $y_1, y_2, \ldots, y_N$ the correspondent standard Gaussian values. Assume the anamorphosis is linearly interpolated as in Figure 3.2, i.e.:

$$\phi(y) = \begin{cases} Z_1 & , \quad y \le y_1 \\ a_k\, y + b_k & , \quad y_k \le y \le y_{k+1} \quad , \quad k=1,\ldots,N-1 \\ Z_N & , \quad y \ge y_N \end{cases} \tag{A.1}$$

with

$$a_k = \frac{Z_{k+1} - Z_k}{y_{k+1} - y_k} \tag{A.2}$$

and

$$b_k = \frac{Z_k\, y_{k+1} - Z_{k+1}\, y_k}{y_{k+1} - y_k} \tag{A.3}$$

Recall that the Hermitian expansion is:

$$\phi(y) = \sum_{k=0}^{\infty} \phi_k\, \eta_k(y) \tag{A.4}$$

with

$$\phi_k = \int_{-\infty}^{\infty} \phi(y)\eta_k(y)\ G(dy) \tag{A.5}$$

The coefficients ϕ_k are found using the equations:

$$\int_{-\infty}^{\infty} \eta_i(u)\ G(du) = \frac{1}{\sqrt{i}}\ \eta_{i-1}(u)\ g(u) \qquad i \geq 1 \tag{A.6}$$

and

$$\int_{-\infty}^{\infty} u\ \eta_i(u)\ G(du) = \frac{1}{\sqrt{i}}\ \eta_{i-1}(u)\ g(u) - \frac{1}{\sqrt{i(i+1)}}\ \eta_{i-2}(u)\ g(u) \qquad i \geq 2 \tag{A.7}$$

where g(u) is the standard Gaussian density function, and G(u) the standard Gaussian distribution function evaluated at u, i.e.:

$$G(u) = \int_{-\infty}^{u} g(x)dx \tag{A.8}$$

The final expressions are:

$$\phi_0 = Z_1 G(y_1) + \sum_{k=1}^{N-1} \left(G(x)\Big|_{y_k}^{y_{k+1}} \cdot b_k + g(x)\Big|_{y_k}^{y_{k+1}} \cdot a_k \right) + Z_N\ \{1 - G(y_N)\} \tag{A.9}$$

$$\phi_1 = Z_1 g(y_1) + \sum_{k=1}^{N-1} \left(g(x) \Big|_{y_k}^{y_{k+1}} b_k + \left[x \cdot g(x) - G(x) \right] \Big|_{y_k}^{y_{k+1}} a_k \right) - Z_N \, g(y_N)$$

(A.10)

$$\phi_j = Z_1 \frac{1}{\sqrt{j}} \eta_{j-1}(y_1) \, g(y_1) - Z_N \frac{1}{\sqrt{j}} \eta_{j-1}(y_N) \, g(y_N)$$

$$+ \sum_{k=1}^{N-1} \left\{ g(x) \frac{1}{\sqrt{j}} \eta_{j-1}(x) \Big|_{y_k}^{y_{k+1}} b_k + \left[x \, g(x) \frac{1}{\sqrt{j}} \eta_{j-1}(x) - g(x) \frac{1}{\sqrt{j(j-1)}} \eta_{j-2}(x) \right] \Big|_{y_k}^{y_{k+1}} a_k \right\}$$

$j \geq 2$ (A.11)

with

$$g(x) \Big|_{a}^{b} = g(b) - g(a)$$

Appendix B

JACKNIFE ESTIMATOR FOR THE IRF-0 MODEL $K(h) = C\delta(h)$

Under the hypothesis stated, the Universal Kriging system, equation (2.93) reduces to:

$$\begin{bmatrix} C & 0 & \cdots & 0 & 1 \\ 0 & C & \cdots & 0 & 1 \\ \vdots & \vdots & & \vdots & \vdots \\ 0 & 0 & \cdots & C & 1 \\ 1 & 1 & \cdots & 1 & 0 \end{bmatrix} \begin{bmatrix} \lambda_{i1} \\ \lambda_{i2} \\ \vdots \\ \lambda_{iN_0} \\ \mu \end{bmatrix} = \begin{bmatrix} 0 \\ 0 \\ \vdots \\ 0 \\ 1 \end{bmatrix} \tag{B.1}$$

which clearly has as solution: $\lambda_{ij} = \frac{1}{N_0}$, $j = 1, 2, \ldots, N_0$ and $\mu = -\frac{C}{N_0}$, for any i.

Note that generalized increments $Z(\lambda_i)$, equation (2.107), produced by an IRF-0 with a generalized covariance function $K(h) = C\delta(h)$ do not depend on the value of C, i.e., the solution of equation (B.1) gives the weights $\lambda_i = \frac{1}{N_0}$ independent of C. Therefore, the optimal value of C can be found solving the problem, see equation (2.10):

$$Q = \min_{C} \sum_{i=1}^{N} \left\{ Z(\lambda_i) - C\, T_i^{\,0} \right\}^2 \tag{B.2}$$

The optimal constant C should then satisfy:

$$\frac{\partial Q}{\partial C} = 0 = -2 \sum_{i=1}^{N} Z(\lambda_i)^2 \, T_i^{\,0} + 2C \sum_{i=1}^{N} (T_i^{\,0})^2 \tag{B.3}$$

with $T_i^{\,0}$, see equation (2.108), given by:

$$T_i^{\,0} = \sum_{\alpha=0}^{N_0} (\lambda_{i_\alpha})^2 = 1 + \frac{1}{N_0} \tag{B.4}$$

This together with equation (B.3) gives:

$$\sum_{i=1}^{N} Z(\lambda_i)^2 = N \left(1 + \frac{1}{N_0}\right) C \tag{B.5}$$

The predicted variance of estimation for each point, see equation (2.94), reduces in the present case to:

$$\sigma_k^{\,2} = C - \mu = \left(1 + \frac{1}{N_0}\right) C \tag{B.6}$$

which then gives together with equation (B.5), see equation (2.111):

$$r = \frac{N(1 + \frac{1}{N_0})\, C}{N(1 + \frac{1}{N_0})\, C} = 1$$

From equations (2.113) and (2.114) it is easily seen that the jackknife estimator takes also the values of 1 because:

$$N_1 r_1 + N_2 r_2 = \frac{N_1 \sum\limits_{j \varepsilon J_1} Z(\lambda_j)^2}{N_1(1 + \frac{1}{N_0})\ C} + \frac{N_2 \sum\limits_{j \varepsilon J_2} Z(\lambda_j)^2}{N_2(1 + \frac{1}{N_0})\ C}$$

$$= \frac{1}{(1 + \frac{1}{N_0})\ C} \left[\sum_{j=1}^{N} Z(\lambda_i)^2 \right] = N$$

Appendix C

CONTOUR PLOTS OF THE FIELDS STUDIED

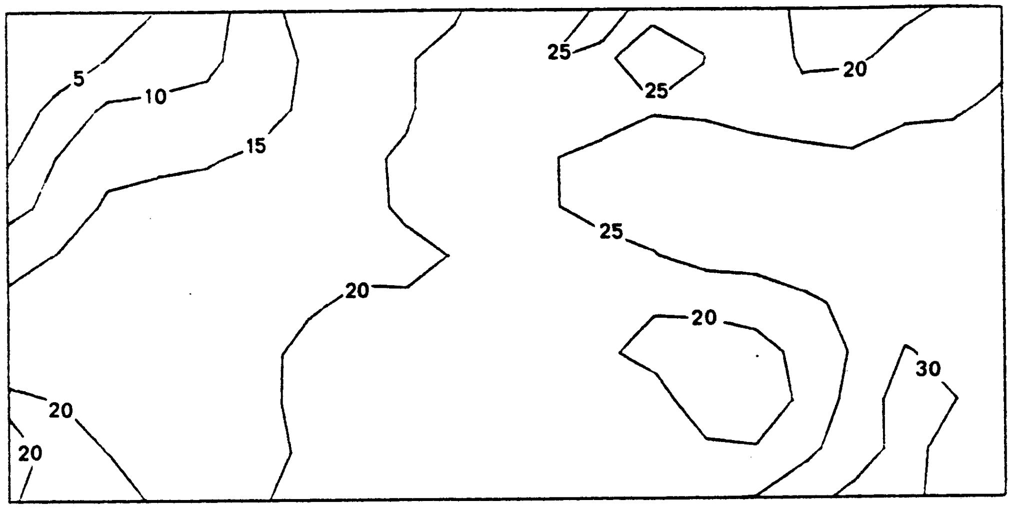


Figure C.1

Isotropic Field. (b = 0.0035, σ = 10, m = 10)

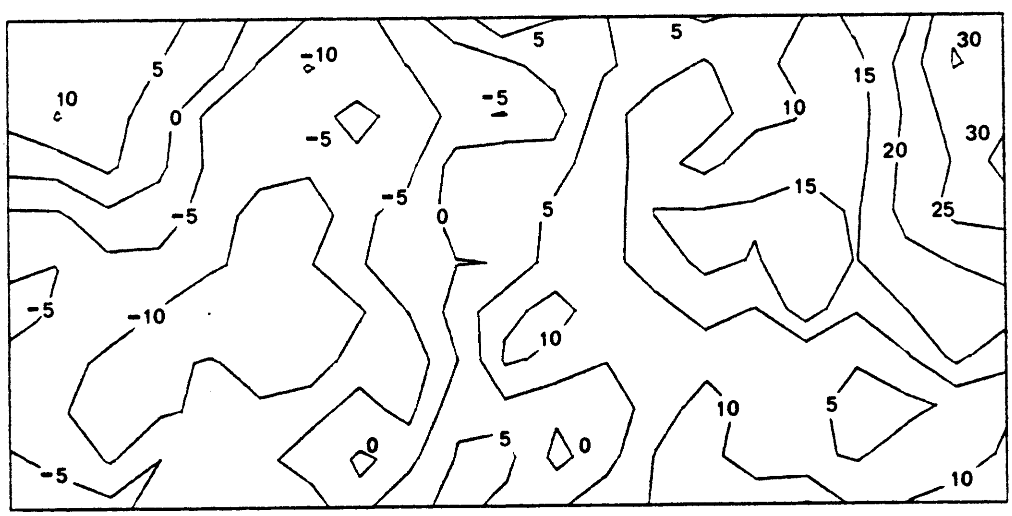


Figure C.2

Isotropic Field. (b = 0.007, σ = 10, m = 10)

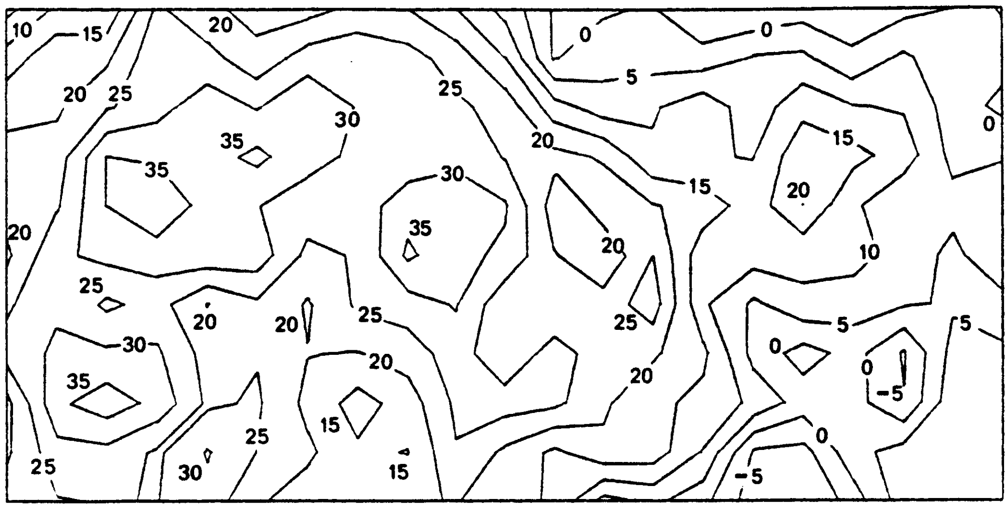


Figure C.3

Isotropic Field. (b = 0.0141, σ = 10, m = 10)

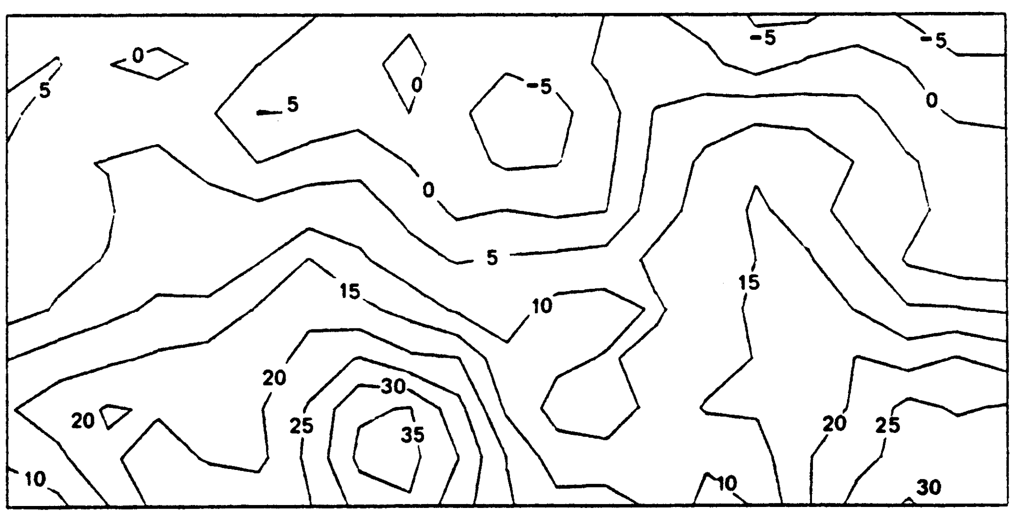


Figure C.4

Isotropic Field. (b = 0.0035, σ = 20, m = 10)

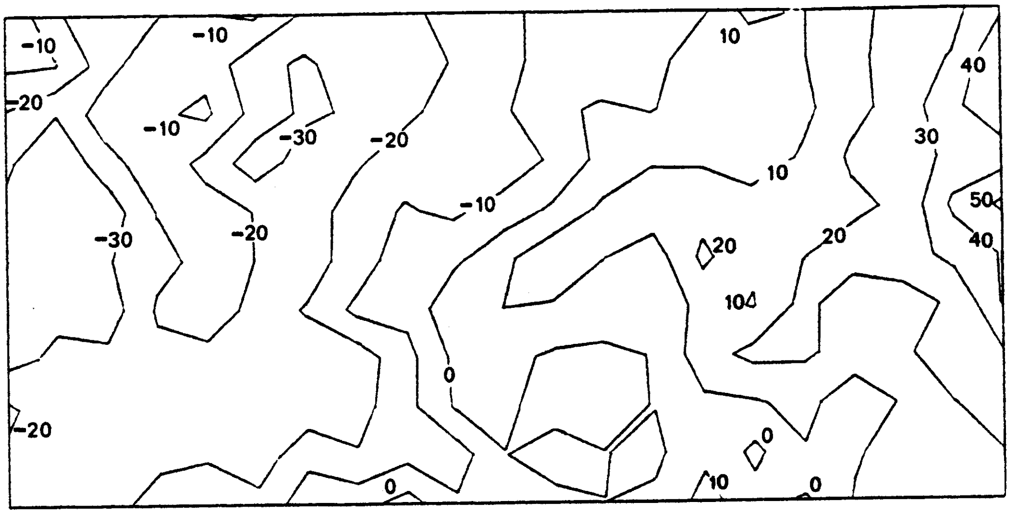


Figure C.5
Isotropic Field. (b = 0.007, σ = 20, m = 10)

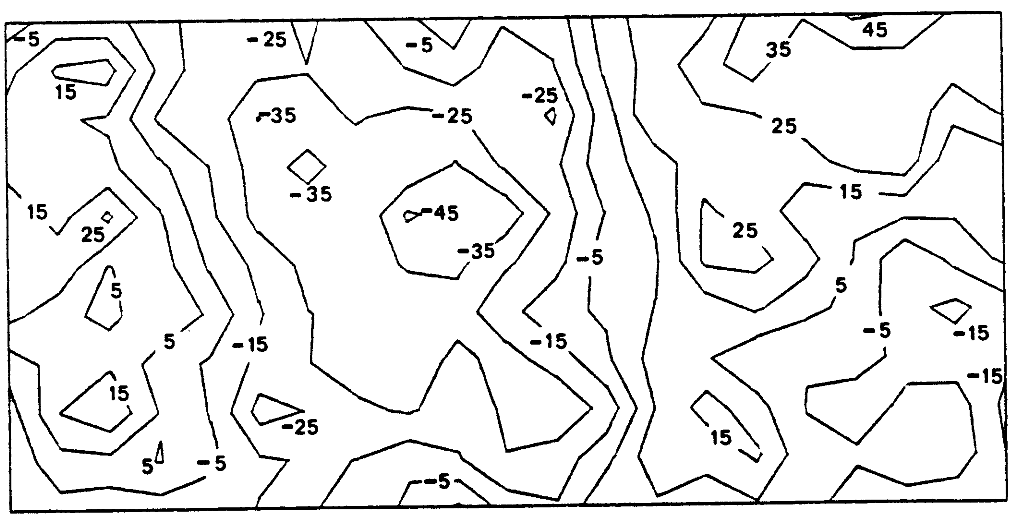


Figure C.6
Isotropic Field. (b = 0.0141, σ = 20, m = 10)

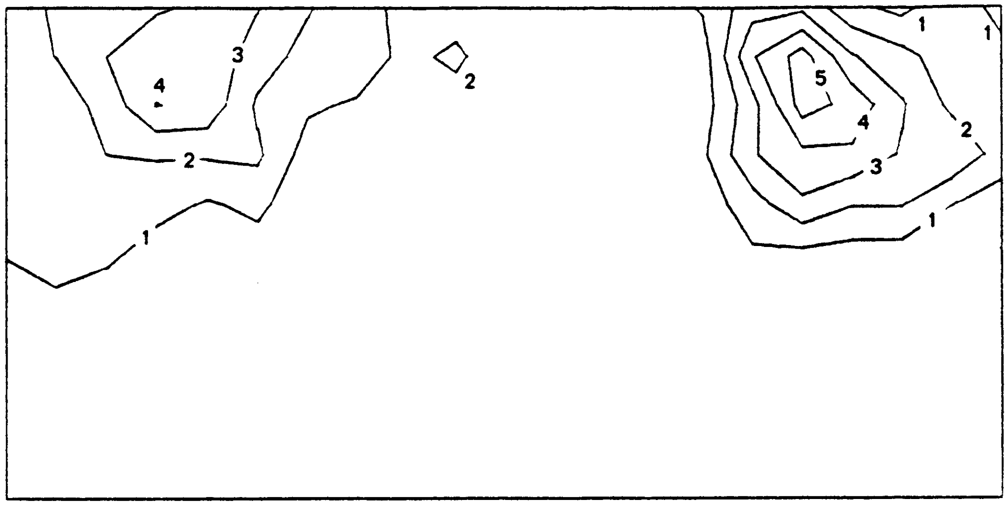


Figure C.7
Transformed Isotropic Field. (b = 0.0035, σ = 10)

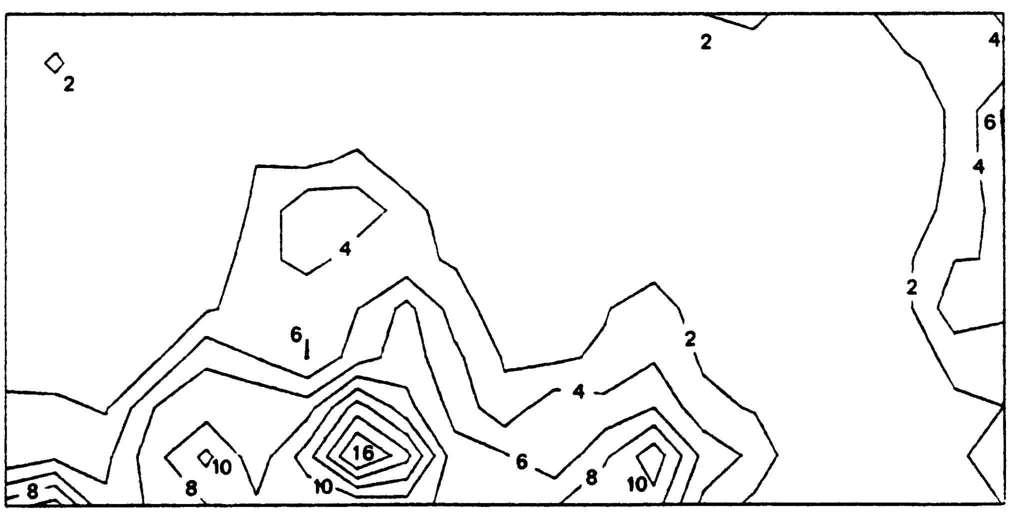


Figure C.8
Transformed Isotropic Field. (b = 0.007, σ = 10)

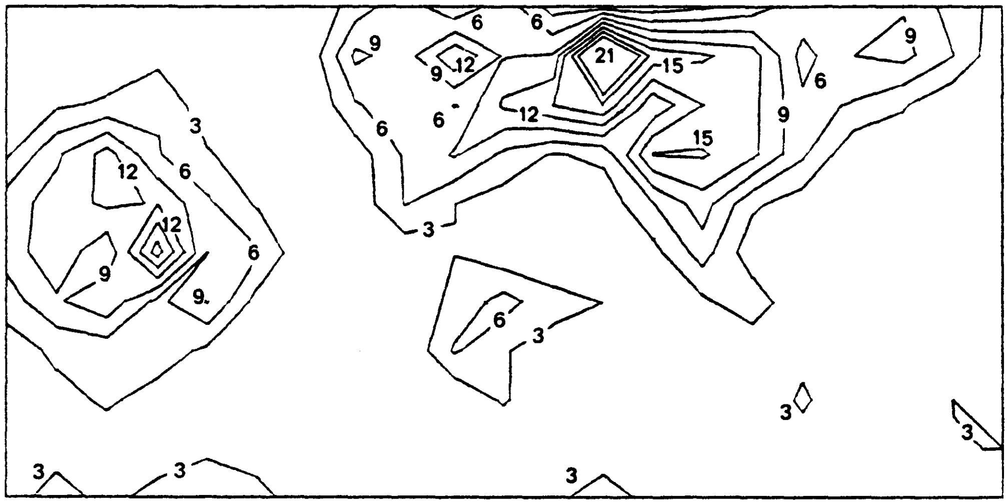


Figure C.9
Transformed Isotropic Field. (b = 0.0141, σ = 10)

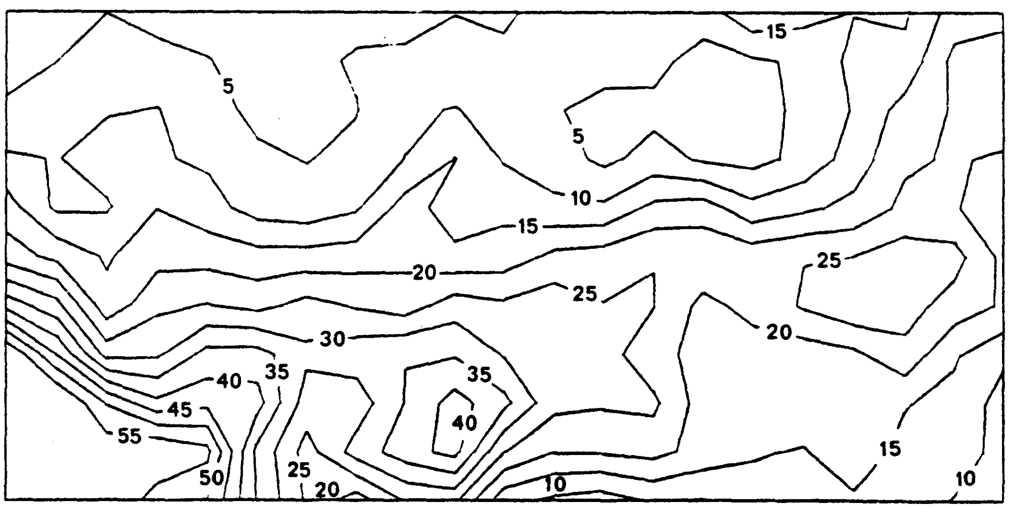


Figure C.10
Transformed Isotropic Field. (b = 0.0035, σ = 20)

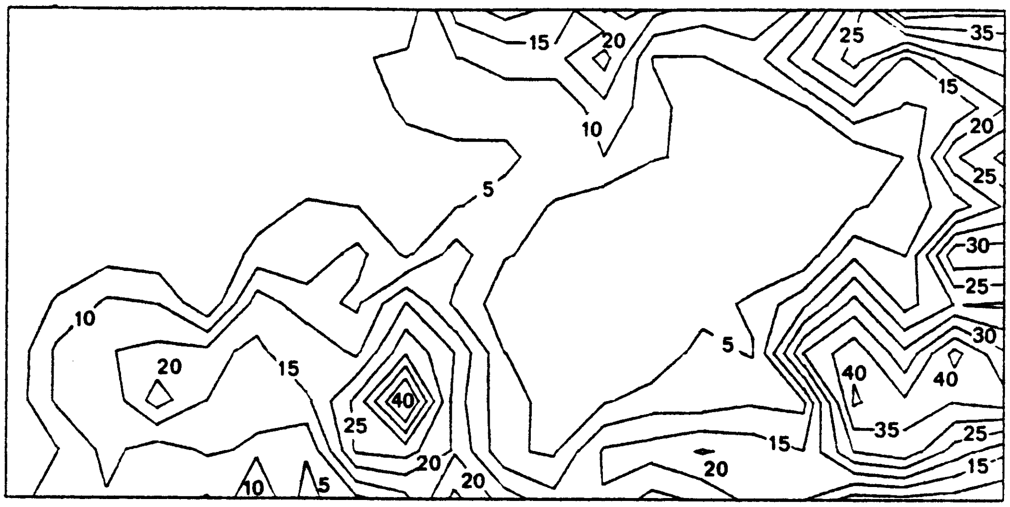


Figure C.11

Transformed Isotropic Field. (b = 0.007, σ = 20)

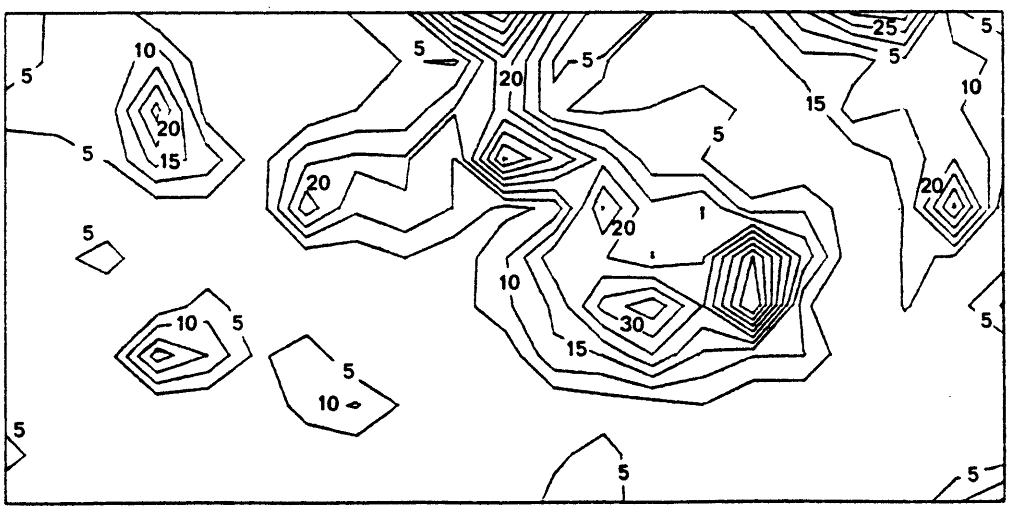


Figure C.12

Transformed Isotropic Field. (b = 0.0141, σ = 20)

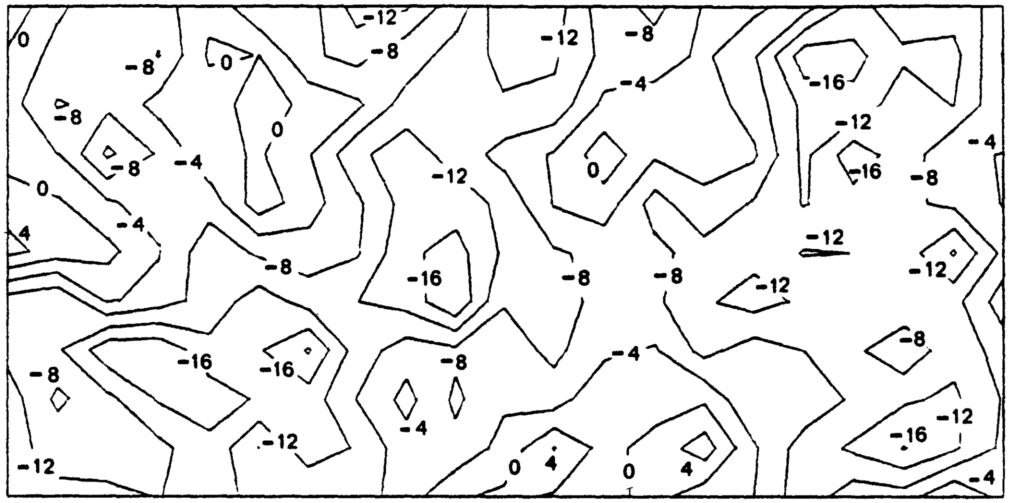


Figure C.13

Intrinsic Random Function. (k = 0, α_1 = -1, α_3 = α_5 = 0)

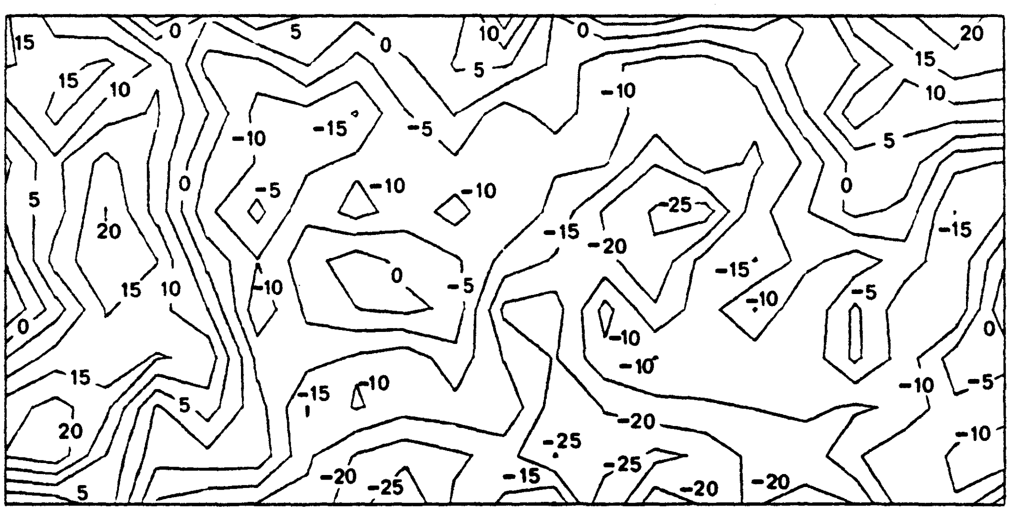


Figure C.14

Intrinsic Random Function. (k = 0, α_1 = -3, α_3 = α_5 = 0)

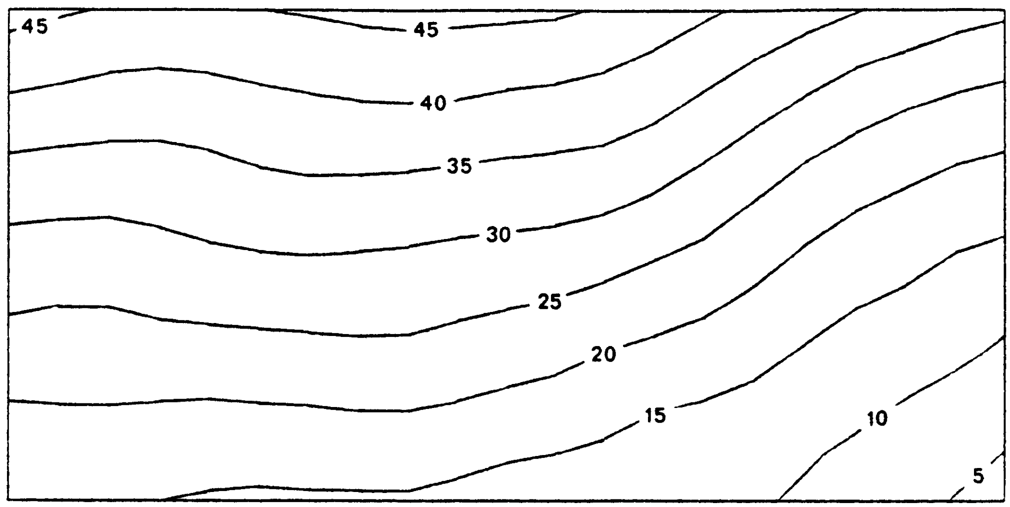


Figure C.15

Intrinsic Random Function. $(k = 1, \alpha_3 = 0.005, \alpha_1 = \alpha_5 = 0)$

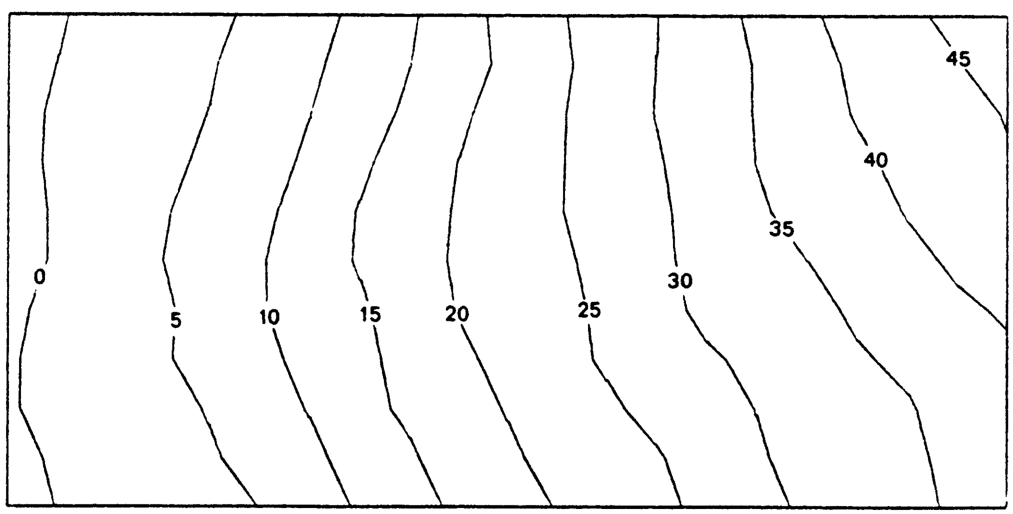


Figure C.16

Intrinsic Random Function. $(k = 1, \alpha_1 = -0.005, \alpha_3 = 0.005, \alpha_5 = 0)$

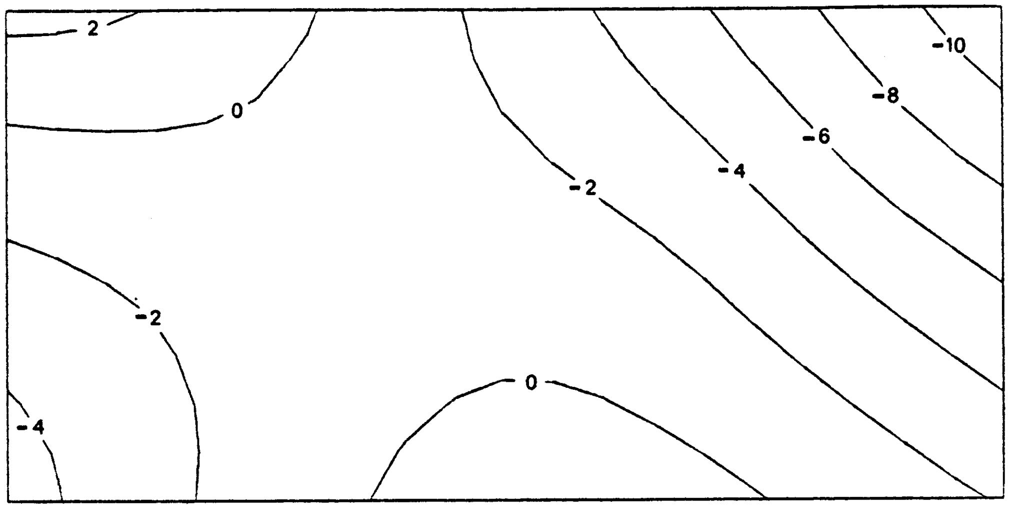


Figure C.17

Intrinsic Random Function. $(k = 2, \alpha_1 = 0, \alpha_3 = 0, \alpha_5 = -0.5x10^{-8})$

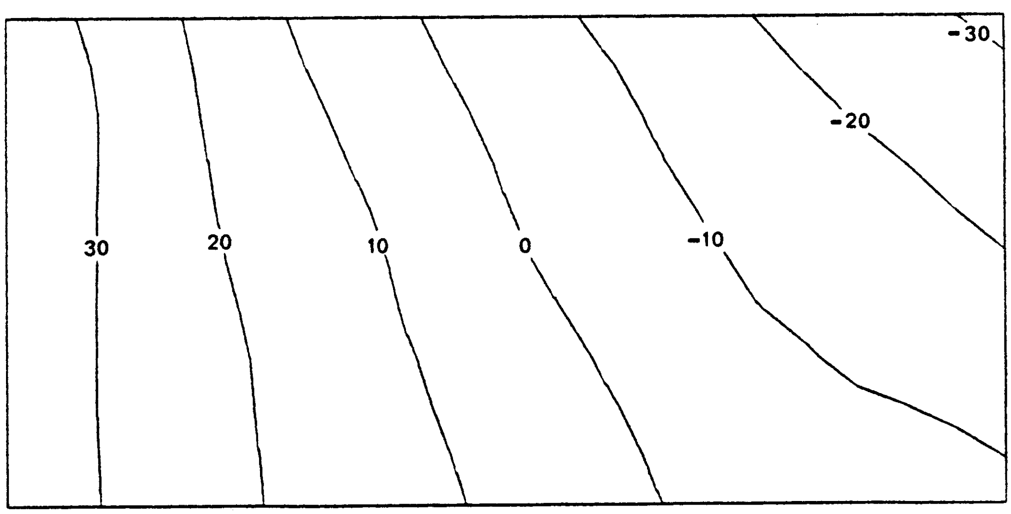


Figure C.18

Intrinsic Random Function. $(k = 2, \alpha_1 = 0, \alpha_3 = 0.005, \alpha_5 = -0.5x10^{-8})$

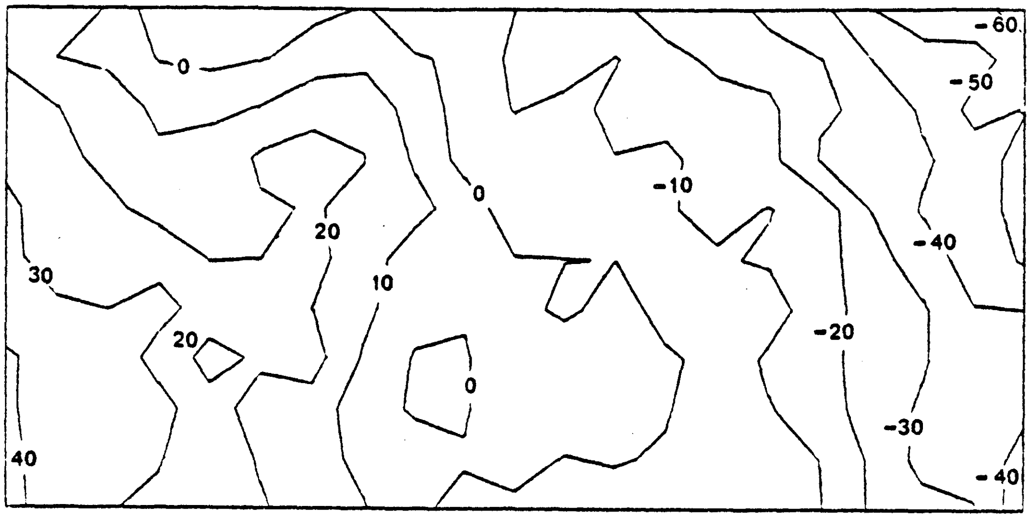


Figure C.19

Intrinsic Random Function. $(k = 2, \alpha_1 = -1, \alpha_3 = 0, \alpha_5 = -0.5x10^{-9})$

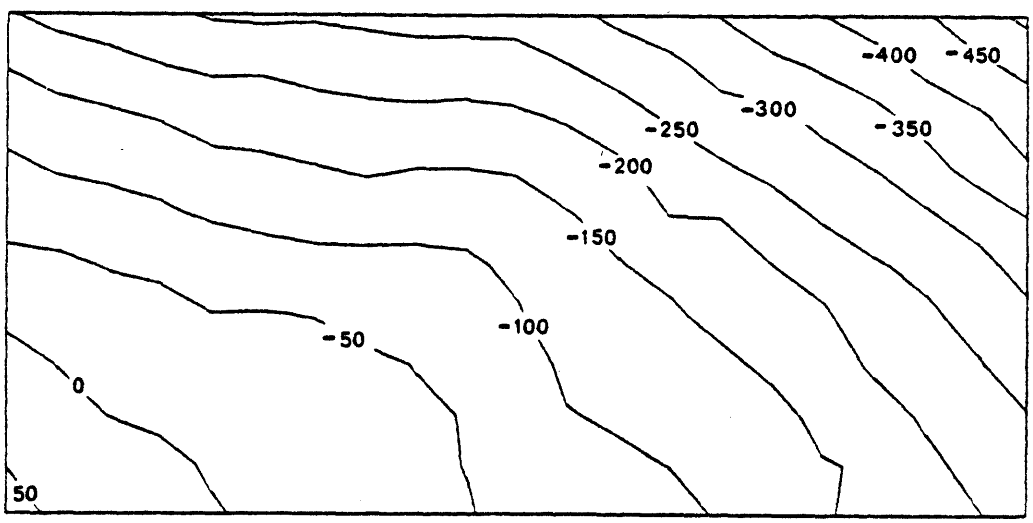


Figure C.20

Intrinsic Random Function. $(k = 2, \alpha_1 = -2, \alpha_3 = 0.05, \alpha_5 = -1x10^{-9})$