ROBUSTNESS TO MISSPECIFICATIONS OF CORRELATION MODELS

Ruth Sheshinski

TECHNICAL REPORT NO. 2
July 22, 1976PREPARED UNDER THE AUSPICES OF
SIAM INSTITUTE FOR MATHEMATICS AND SOCIETY DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA **MASTER**

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

Prepared For
THE U.S. ENERGY RESEARCH AND DEVELOPMENT ADMINISTRATION
UNDER CONTRACT NO. EY-76-S-02-2874.*000

DISCLAIMER

This report was prepared as an account of work sponsored by an agency of the United States Government. Neither the United States Government nor any agency Thereof, nor any of their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness, or usefulness of any information, apparatus, product, or process disclosed, or represents that its use would not infringe privately owned rights. Reference herein to any specific commercial product, process, or service by trade name, trademark, manufacturer, or otherwise does not necessarily constitute or imply its endorsement, recommendation, or favoring by the United States Government or any agency thereof. The views and opinions of authors expressed herein do not necessarily state or reflect those of the United States Government or any agency thereof.

DISCLAIMER

Portions of this document may be illegible in electronic image products. Images are produced from the best available original document.

ROBUSTNESS TO MISSPECIFICATIONS OF CORRELATION MODELS

by

RUTH SHESHINSKI

Hebrew University, Jerusalem
and Stanford University

TECHNICAL REPORT NO. 2

JULY 22, 1976

STUDY ON STATISTICS AND ENVIRONMENTAL
FACTORS IN HEALTH

PREPARED UNDER SUPPORT TO SIMS FROM
ENERGY RESEARCH AND DEVELOPMENT ADMINISTRATION (ERDA)
ROCKEFELLER FOUNDATION

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

NOTICE

This report was prepared as an account of work sponsored by the United States Government. Neither the United States nor the United States Energy Research and Development Administration, nor any of their employees, nor any of their contractors, subcontractors, or their employees, makes any warranty, express or implied, or assumes any legal liability or responsibility for the accuracy, completeness or usefulness of any information, apparatus, product or process disclosed, or represents that its use would not infringe privately owned rights.

MASTER

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

ACKNOWLEDGEMENTS

This work is part of a Ph.D. dissertation written for the Hebrew University, Jerusalem. The work was conducted at the Department of Statistics at Stanford University. I am most grateful to Professor Ingram Olkin for enabling me to be associated with Stanford University.

I wish to take this opportunity to thank my teacher Professor Reuben Gabriel, who introduced me to the field of regional variables.

My greatest intellectual and personal debt I owe to Professor Paul Switzer, without whom this work would never have been done. His extensive involvement, helpful discussions and "open door" policy, guided me all along. Professor Switzer read some earlier drafts and suggested numerous improvements. Needless to say, any remaining errors are solely mine.

The excellent typing and patience of Ms. Karola Lof are highly appreciated. And last, but not least, to my dearest, who constantly encouraged me not to give up professional work, but nevertheless suffered the most from it, Eytan, Yael, Hadar, Michal and Noa.

Abstract

Optimal linear interpolation between discrete observations in R^2 depends on the random function model assumed to generate the two-dimensional field. Sensitivity of the optimal solution to errors in specification of the model are studied. An illustration using second-order stationary models with a misspecified autocorrelation function is examined in detail. We also explore, by examples, the sensitivity procedure when more distant data points are omitted from the calculations.

We have found that the optimal linear estimator is grossly insensitive to misspecifications of the autocorrelation function. However, if the misspecified autocorrelation function is used in the calculation of the squared interpolation error - then such error calculations may be grossly misleading.

Key Words: Discrete observations in R^2 , isopleths - mapping, interpolation in R^n , misspecification of correlation models, robustness of estimation procedure, robustness of error estimator, isotropic correlation, anisotropic correlation, best linear unbiased estimator, covariogram - correlogram.

Robustness to Misspecifications of Correlation Models

by

Ruth Sheshinski

Hebrew University, Jerusalem
and Stanford University

Introduction.

In the analysis of air pollution data the problem of interpolation between monitoring stations occurs in the construction of pollution isopleth maps. Interpolation has been carried out using many different techniques which may be divided into two main groups.

(1) Weighted averages of the observations, where the weights are selected according to different rules. For illustration see Behar (1972) who uses as weights the inverse of the squared distance from the interpolated point to the monitoring station. Such algorithms for calculating the weighting coefficients are usually arbitrary and have no common theoretical basis. See also references for other interpolation algorithms in Switzer, Mohr and Heitman (1964). Other ad hoc techniques use equal weighting of the nearest observations. Tinkler (1971) also used nearest neighbor considerations in choosing weight functions. Burnaby (1970) used the concept of information to choose the weighting coefficients. Some interpolation methods also depend on fixed grid patterns of observations, which are inappropriate for direct applications to monitoring station data.

(2) Polynomial surface fitting. The polynomial fitting procedures to data - sometimes referred to as "trend surface analysis" has also been used for interpolation. Whitten (1970) makes direct use of orthogonal polynomials. Additional references to polynomial fits are found in Whitten's paper. Mandelbaum (1963) makes use of nonorthogonal polynomials for interpolation. Geologists, such as Krumbein (1959) and Robinson and Merriam (1971), use the trend surface analysis to map geological patterns (i.e., to interpolate between given observations). Polynomial fitting techniques fail to reproduce the monitoring station data, when the interpolation point is an observation point. For more objections to polynomial interpolation see Mathéron (1967), Olea (1972), Norcliffe (1969), and Whitten (1974). Rao and Rao (1970) used averages of nearest neighbors as well as polynomial trend surface analysis as a composite analysis of the two above mentioned methods.

The interpolation techniques listed do not explicitly use a model of autocorrelation between the station observations. An interpolation theory based on an autocorrelation model has been developed in detail by the French geomathematician Georges Mathéron. He has calculated weighted average interpolation rules whose estimation variance is minimal for a specified correlation model. See Mathéron (1965, 1969, 1972, 1973) and Huijbregts and Mathéron (1970). Each interpolated location is accompanied by an interpolation variance indicating the precision of the estimation according to the autocorrelation model. By this theory Blais and Carlier (1967) estimated ore grade distribution isopleths.

However, the model of autocorrelation of spatially distributed pollutant concentrations is never known, and as such, has to be estimated. We do have to assume a certain model of correlation (its functional form, etc.). The question is: How sensitive is the interpolated value and the precision of interpolation w.r.t. assumptions (or misspecifications) on the correlation model of the process.

Misspecifications of correlograms are possible in many different ways, as for example: misspecification of the parameter in a given function, or misspecification of the form of the function, or, choosing an anisotropic process (with correlations depending on direction) rather than a correlation model which is a function of the distance solely.

The observations we had in mind are the measurements of air pollution at various stations. The following study deals with interpolation between stations. It seems reasonable to engage in the process of interpolation using only a few stations that geographically are close to the interpolation point. (Unless deletion of data points effects severely the precision and outcome of the estimate, which has been studied for a few configurations in a separate chapter.) Knowing that the stations are not independent stations, it seems important to study the variability of the estimation precision with respect to different assumptions for correlation models between the stations, hoping that our estimate is robust to misspecifications of correlation models.

1. Misspecification of a Parameter in the Correlation Model.

a. General Notes and Notation.

Starting with 3 observations $Z(\underline{X}_1), Z(\underline{X}_2), Z(\underline{X}_3)$ at locations $\underline{X}_1, \underline{X}_2, \underline{X}_3, \underline{X}_i \in \mathbb{R}^2$, located conveniently at the vertices of an equilateral triangle, we want to estimate - interpolate - the value of $Z(\underline{X}_0)$ at a point $\underline{X}_0$ inside the triangle formed by the data. We model $\{Z(\underline{X}), \underline{X} \in \mathbb{R}^2\}$ to be a random function which, for the time being, will be considered an isotropic process. An isotropic process is a stationary process with the property that its autocorrelation function is a function only of distance in the plane.

In general, a linear estimate Z^* at $(\underline{X}_0)$ (linear combination of the observations) is given by

$$(1.1) \quad Z^*(\underline{X}_0) = \sum_{i=1}^n c_i Z(\underline{X}_i)$$

where we want: 1) Z^* to be an unbiased estimator of Z and
2) C_i 's are such that the estimation variance of Z^* is minimal w.r.t. C_i 's. The estimation squared error averaged over the process, for a linear estimate Z^* (function of n observations) is

$$(1.2) \quad E[Z^*(\underline{X}_0) - Z(\underline{X}_0)]^2 = \text{Cov}[Z(\underline{X}_0)Z(\underline{X}_0)] + \sum_{i=1}^n \sum_{j=1}^n c_i c_j \text{cov}(Z(\underline{X}_i)Z(\underline{X}_j)) - 2 \sum_{i=1}^n c_i \text{cov}(Z(\underline{X}_i)Z(\underline{X}_0)).$$

Without loss of generality, we may rewrite the variance-covariance matrix in terms of the isotropic correlations ρ_{ij} , (or $\rho(h)$, where h means distance in R^2)

$$\rho_{ij} = \rho(|\underline{X}_i - \underline{X}_j|)$$

$$\rho_{io} = \rho(|\underline{X}_i - \underline{X}_o|)$$

and (2) is written as

$$(1.3) \quad \text{Var}[Z^*(\underline{X}_o) - Z(\underline{X}_o)] = E[Z^*(\underline{X}_o) - Z(\underline{X}_o)]^2 \\ = \left(1 + \sum_{i=1}^n \sum_{j=1}^n c_i c_j \rho_{ij} - 2 \sum_{i=1}^n c_i \rho_{io} \right) \cdot \sigma^2 .$$

σ^2 is the stationary variance of the process.

($\text{Var}[Z^*(\underline{X}_o) - Z(\underline{X}_o)]$ will also be denoted $\sigma_{\underline{X}_o}^2$.) WLOG we shall assume $\sigma^2 = 1$ in the following work. Under the condition that Z^* be an unbiased estimate of Z we have the linear constraint $\sum_{i=1}^n c_i = 1$. ($c_n = 1 - \sum_{i=1}^{n-1} c_i$). We look for

$$(1.4) \quad \min_{\text{w.r.t. } c_i \text{'s}} \sigma_{\underline{X}_o}^2 = \min_{c_i \text{'s}} \left[1 + \sum_{i=1}^n \sum_{j=1}^n c_i c_j \rho_{ij} - 2 \sum_{i=1}^n c_i \rho_{io} \right],$$

under the above mentioned constraint. The reduced problem (searching only for $n-1$ optimal coefficients) may be written in the following form:

$$\begin{aligned} \min_{\text{w.r.t. } C_1, \dots, C_{n-1}} \sigma_{X_o}^2 = \min_{C_1, \dots, C_{n-1}} & \left[1 + \sum_i^{n-1} \sum_j^{n-1} C_i C_j \rho_{ij} - 2 \sum_i^{n-1} C_i \rho_{io} \right. \\ & \left. + \left(1 - \sum_i^{n-1} C_i \right)^2 + 2 \left(1 - \sum_i^{n-1} C_i \right) \sum_i^{n-1} C_i \rho_{in} - 2 \left(1 - \sum_i^{n-1} C_i \right) \rho_{no} \right]. \end{aligned}$$

By derivation w.r.t. $C_i, \dots, C_{n-1}$

$$\begin{aligned} \frac{\partial \sigma_{X_o}^2}{\partial C_k} = & 2 \sum_{j=1}^n C_j \rho_{jk} - 2 \rho_{ko} + 2 \left(1 - \sum_j^{n-1} C_j \right) (-1) + 2 \rho_{kn} - 2 \sum_j^{n-1} C_j \rho_{jn} \\ & - 2 \rho_{kn} \sum_j^{n-1} C_j + 2 \rho_{no}; \quad k = 1, 2, \dots, (n-1) \end{aligned}$$

and equating to zero, we get $(n-1)$ linear equations:

$$\begin{aligned} (1.5) \quad \sum_{j=1}^{n-1} C_j [\rho_{kj} + 1 - \rho_{jn} - \rho_{kn}] &= [\rho_{ko} + 1 - \rho_{kn} - \rho_{no}]; \\ k &= 1, 2, \dots, (n-1). \end{aligned}$$

Rewriting the above $n-1$ equations in matrix notation defining:

$$R_{ji} = R_{ij} \equiv (1 + \rho_{ij} - \rho_{jn} - \rho_{in})$$

$$\text{Matrix}_{n-1}^{n-1} R_{n-1} = ((R_{ij}))$$

$$\text{Matrix}_{1 \text{ } n-1}^R = ((R_{io}))$$

$$\text{Matrix}_{1 \text{ } n-1}^{C_{n-1}} = ((C_1, C_2, \dots, C_{n-1}))$$

results in:

$$C^T = R^{-1} \cdot R_0^T \quad (n-1 \text{ coefficients}), \quad T \text{ denoting transpose,}$$

$$C_n = 1 - \sum_{i=1}^{n-1} C_i.$$

To simplify the presentation, we shall estimate at points $\underline{X}_0$ along the line $\overline{AD}$ in figure 1.

Figure 1: Geometrical setup for three observations.

We shall indicate the estimated point by $\underline{X}_0$. We assume that the observations are one unit distance apart.

We shall be looking into robustness of linear estimation and error estimation for interpolation of nonisotropic processes (i.e., where the correlation is a function, depending also on direction) in a separate chapter. For illustration we use the parametrized family of correlation functions $\rho(h) = \exp\{-u \cdot h^2\}$, $u > 0$. The larger the u value, the faster the correlation function decays to zero.

Figure 2: $\rho(h) = e^{-uh^2}$

We may also express the family as $\rho(h) = [\rho(1)]^{h^2}$, with $u = -\ln \rho(1)$. For $u \rightarrow \infty$ we get uncorrelated observations. The correlogram has the following form:

$$(1.6) \quad \rho(h) = \begin{cases} 0 & h > 0 \\ 1 & h = 0 \end{cases}.$$

This situation is the worst from the point of view of relevant information that we are able to extract and use in estimation out of the given observations. Having less information on hand (no correlation between points, not even between points close to each other) must result in a larger error for the estimate. From equation (3)

$$\lim_{u \rightarrow \infty} \text{Var}[Z^*(\underline{X}_0) - Z(\underline{X}_0)] = \frac{4}{3} \quad (\text{for } n = 3)$$

as expected for uncorrelated models.

In general, for correlated models the closer $\underline{X}_0$ is to an observation point - the more weight should be given to this observation point (i.e., the larger the corresponding coefficient) and the smaller will be the error of estimation. To demonstrate this point, let us take 5 different geometrical configurations, denoted $\underline{X}_1 - \underline{X}_5$ in Figure 1, calculate their optimal coefficients C_i (to ensure Z^* unbiased and $\min \text{Var}[Z^* - Z]$), and their corresponding estimation variance denoted $\text{opt } \sigma_0^2$, for using $Z^*(\underline{X}_0)$ instead of $Z(\underline{X}_0)$. See table 1.

b. Suboptimality Resulting From a Misspecified Parameter Value.

Complete knowledge of the true correlation function would result in having the optimal estimation variance for linear estimates. This optimal estimate will be denoted $Z_{u_0}^*(\underline{X}_0)$, (u_0 being the true parameter) and its estimation variance will be denoted $\text{Var}_{u_0} Z_{u_0}^*$. On the other hand, if we use an incorrect value of $u \neq u_0$ to calculate the weights $\{C_i\}$, then we denote the sub-optimal estimator by $Z_u^*(\cdot)$. The variance calculation is done with correlations using the true parameter u_0 . We shall denote this suboptimal variance as $\text{Var}_{u_0} Z_u^*(\cdot)$.

If misspecification of correlation model in the procedure of estimation does not change the error variance substantially, we may conclude that the above linear estimation procedure is robust under

Table 1: Optimal coefficients and variance using correlation model

$$\rho_u(1) = 0.6067 \quad (\text{i.e. } u = -\ln \rho(1) = 0.5)$$

$\underline{x}_0$ According to Figure 1	(0., 0.)	(0., 0.134)	(0., $\sqrt{3}/6$) center	(0., 0.6)	(0., $\sqrt{3}/2$) observation
configuration
optimal coefficients c_1	0.0026	0.1475	1/3	0.7198	1
c_2	0.4987	0.4262	1/3	0.1401	0
c_3	0.4987	0.4262	1/3	0.1401	0
optimal $\sigma_{\underline{x}_0}^2$	0.0383	0.0412	0.0447	0.0232	0 (for any $\rho_u(1)$)

misspecification of the parameter of the correlation model.

The interesting ratio to look at would be

$$(1.7) \quad \sqrt{\frac{\text{Var}_{u_0} Z_u^*(\cdot)}{\text{Var}_{u_0} Z_{u_0}^*(\cdot)}} \geq 1.$$

The underlying correlation models we consider, range from

$$\rho_{u_0}(1) = e^{-u_0} = e^{-0.1} = 0.905 \quad \text{till}$$

$$\rho_{u_0}(1) = e^{-u_0} = 0 \quad (\text{for } u_0 \uparrow \infty)$$

The same range is used for misspecified models, calculated with $u \neq u_0$.

Under the limit condition of uncorrelated observations, i.e.,

$\rho_{u_0}(1) = 0$, ($u_0 = \infty$), and/or misspecified correlation model

$\rho_u(1) = 0$, ($u = \infty$), we calculate the optimal and suboptimal var-

iances directly from equation (1.3) with reference to the correlo-

gram given in (1.6). The optimal coefficients calculated under true

parameter $u_0 = \infty$, for n observations is: $C_i(u_0) = 1/n$, $\forall i$, and

under misspecified parameter $u = \infty$: $C_i(u) = 1/n$, $\forall i$. As a result:

$$(1.8) \quad \text{Var}_{u_0 \rightarrow \infty} Z_{u_0}^* = 1 + \sum_{i=1}^n C_i^2(u_0) = 1 + \frac{1}{n}$$

and

$$(1.9) \quad \text{Var}_{u \rightarrow \infty} Z_u^* = 1 + \frac{1}{n^2} \sum_{i=1}^n \sum_{j=1}^n \rho_{u_0}(|X_i - X_j|) - \frac{2}{n} \sum_{i=1}^n \rho_{u_0}(|X_i - X_0|).$$

The points at which we interpolate are those shown in Figure 1.

$\underline{X}_0 = (0., 0.)$ and $(0., 0.6)$. Tables 2, A and B show the suboptimal variance $\text{Var}_{u_0} Z_u^*(.)$ and the square root ratio of suboptimal to optimal variance. It seems evident from these tables that the linear interpolation procedure is not very sensitive to choice of coefficients. Even if coefficients are calculated using a grossly misspecified parameter in the correlation model, the estimation error is not severely influenced.

By and large if the misspecified model is of stronger underlying correlation than the true model, the suboptimal coefficients are almost unchanged in comparison to the optimal ones. The errors are more noticeable (but still small) when the estimation point $\underline{X}_0$ is close to an observation point. Compare Table 2_A at $\underline{X}_0 = (0., 0.)$ with Table 2_B at $\underline{X}_0 = (0., 0.6)$.

The worst error in the table appears when we in fact have a very high correlation model $\rho_{u_0}(1) = 0.9048$ and misspecify it by $\rho_u(1) = 0$ in the estimate Z^* . (Misspecifying correlated observations by uncorrelated ones.) For misspecified $\rho_u(1) = 0.05$ at $\underline{X}_0 = (0., 0.)$ the performance ratio given by formula (1.7) is 1.20, where at $\underline{X}_0' = (0., 0.6)$ (location close to an observation), the same ratio is 1.42. The most extreme case of misspecification is when $\rho_u(1) = 0$ ($u = \infty$). Then for $\underline{X}_0 = (0., 0.)$, the performance ratio is 3.13 and for $\underline{X}_0 = (0., 0.6)$ the same ratio is 4.23.

This again illuminates the well known fact that in the case where we deal with correlated observations, we should by no means treat

TABLE 2_A 1. $\text{Var}_{u_0} Z_u^*$; 2. $(\text{Var}_{u_0} Z_u^*(\cdot) / \text{Var}_{u_0} Z_{u_0}^*(\cdot))^{1/2}$

$n=3$; $\underline{X}_0=(0.,0.)$; $\rho_u(h)=e^{-uh^2}$ (h =distance between points)

True $\rho_{u_0}(1)$

Misspecified $\rho_u(1)$	0.9048 $u_0=0.1$	0.7788 $u_0=0.25$	0.6065 $u_0=0.5$	0.4724 $u_0=0.75$	0.3679 $u_0=1.$	0.1353 $u_0=2.$	0.0498 $u_0=3.$	0.0 $u_0=\infty$
0.9048 $u=.1$	0.001799 1	0.01057 1.0000	0.0383 1.0000	0.07812 1.0002	0.1263 1.0004	0.3546 1.0026	0.5801 1.0071	1.4999 1.0606
0.7788 $u=.25$	0.001799 1.0000	0.01057 1	0.0383 1.0000	0.07812 1.0001	0.1263 1.0003	0.3545 1.0025	0.5800 1.0070	1.4993 1.0604
0.6065 $u=.5$	0.001800 1.0002	0.01057 1.0000	0.0383 1	0.07811 1.0000	0.1263 1.0002	0.3544 1.0023	0.5796 1.0067	1.4974 1.0597
0.4724 $u=.75$	0.001803 1.0013	0.01058 1.0004	0.0383 1.0000	0.07810 1	0.12626 1.0001	0.3541 1.0019	0.5789 1.0061	1.4943 1.0586
0.3679 $u=1.$	0.001813 1.0040	0.01060 1.0014	0.0383 1.0004	0.07811 1.0000	0.1262 1	0.3537 1.0014	0.5781 1.0054	1.4900 1.0570
0.1353 $u=2.$	0.0020 1.0547	0.0110 1.0213	0.0390 1.0095	0.07891 1.0051	0.1270 1.0028	0.3528 1	0.5740 1.0018	1.4644 1.0480
0.0498 $u=3.$	0.0026 1.2064	0.0124 1.0851	0.0414 1.0406	0.0820 1.0246	0.1303 1.0161	0.3546 1.0026	0.5719 1	1.4328 1.0366
0.0 $u=\infty$	0.0176 3.1323	0.0473 2.1149	0.1028 1.6392	0.1630 1.4447	0.2253 1.3358	0.4661 1.1495	0.6664 1.0790	$\frac{4}{3}$ 1

TABLE 2_B1. $\text{Var}_{u_0} Z_u^*$; 2. $(\text{Var}_{u_0} Z_u^* / \text{Var}_{u_0} Z_{u_0}^*)^{1/2}$ $n=3$ $\underline{X}_0 = (0., 0.6)$ $\rho_u(h) = e^{-uh^2}$ True $\rho_{u_0}(1)$

Misspecified $\rho_u(1)$		0.9048 $u_0=0.1$	0.7788 $u_0=0.25$	0.6065 $u_0=0.50$	0.4724 $u_0=0.75$	0.3679 $u_0=1$	0.1353 $u_0=2$	0.0498 $u_0=3$	0.0 $u_0=\infty$
0.9048	.1	0.00113	0.00657	0.0235	0.0474	0.0758	0.2058	0.3301	1.5333
		1	1.0017	1.0058	1.0094	1.0125	1.0189	1.0178	1.0724
0.7788	.25	0.00114	0.00654	0.0233	0.0470	0.0752	0.2043	0.3281	1.5425
		1.0043	1	1.0022	1.0054	1.0083	1.0151	1.0147	1.0756
0.6065	.50	0.00119	0.00660	0.0232	0.0466	0.0745	0.2022	0.3253	1.5574
		1.0284	1.0043	1	1.0012	1.0034	1.0099	1.0104	1.0808
0.4724	.75	0.00129	0.00675	0.0233	0.0465	0.0741	0.2007	0.3231	1.5716
		1.0683	1.0159	1.0018	1	1.0008	1.0060	1.0071	1.0856
0.3679	1	0.00141	0.00698	0.0235	0.0466	0.0740	0.1996	0.3215	1.5849
		1.1177	1.0325	1.0067	1.0010	1	1.0034	1.0046	1.0903
0.1353	2	0.00196	0.00809	0.0250	0.0480	0.0749	0.1983	0.3189	1.6259
		1.3191	1.1117	1.0379	1.0157	1.0066	1	1.0004	1.1043
0.0498	3	0.00227	0.00874	0.0260	0.0490	0.0759	0.1985	0.3186	1.6437
		1.4205	1.1559	1.0577	1.0266	1.0131	1.0006	1	1.1103
0.0	∞	0.0202	0.0528	0.1114	0.1722	0.2330	0.4512	0.6135	$\frac{4}{3}$
		4.2280	2.8410	2.1908	1.9245	1.7749	1.5086	1.3876	1

them as independent ones.

For the more likely to appear situations, (where in practice we know that we deal with correlated observations but we misspecify the parameter of the model either by over - or under estimating it) the errors are much smaller and may be considered in most practical applications as negligible.

If the true situation has independent observations, any misspecified model of correlation does not (for practical purposes) influence the precision of the estimate, even for locations close to an observation point. For $\underline{X}_O = (0., 0.6)$, $\rho_{u_O}(1) = 0.0498$, $\rho_u(1) = 0.9048$, the ratio is less than 1.02. For $\underline{X}_O = (0., 0.)$ the corresponding ratio is less than 1.01.

For the most extreme case, where $\rho_{u_O}(1) = 0$ ($u_O = \infty$) , and misspecification as above, the performance ratio for $\underline{X}_O = (0., 0.)$ is 1.06 and for $\underline{X}_O = (0., 0.6)$ is 1.07.

The performance ratios corresponding to different misspecified correlations $\rho_u(1)$ are not going up monotonically as a function of the parameter u going down.

From looking at Tables 2_A, 2_B and Figure 3_A and 3_B, it is noticeable that in case the direction of the misspecification is to assume weaker association $\rho(1)$, than the true model, the performance ratio is increasing. On the other hand, assuming stronger association $\rho(1)$ than the true one, does not cause any substantial deviation from the optimal estimation variance.

Similar calculations were done with 4 observations located at the 4 corners of a unit square. Figure 4 indicates the geometrical setup of

Figure 3A: Performance Ratio vs. misspecified correlation. $n = 3$

Figure 3B: Performance Ratio vs. misspecified correlation. $n = 3$

the observations. We wish to interpolate the random function $Z(\underline{X})$ at a point $\underline{X} = \underline{X}_0$ in the square.

Figure 4: Geometrical setup for 4 observations

The parametrized family of correlation functions used is $\rho(h) = e^{-uh^2}$, $u > 0$.

The first estimation point $\underline{X}_0$ we consider is the center of the square $\underline{X}_0 = (0.5, 0.5)$. We examine the consequences of misspecifying the parameter of the correlation model. However, any specification of the correlation model gives coefficients C_i which are all equal to $1/4$. So, every specification gives the optimal estimate.

As mentioned before, the location of the estimated point, as well as the direction of the assumed misspecified parametric value for the correlation model influences the performance ratio (1.7). Therefore we shall choose two additional estimation locations $\underline{X}_0 = (0.1, 0.1)$, (close to an observation point) and $\underline{X}_0 = (0.5, 0)$.

By choosing the above three mentioned estimation locations $\underline{X}_O$, under three different "true" correlation models, we tried to get some feeling how a misspecified parameter u would influence the performance ratio (1.7). See Figure 4 for relative locations of the various $\underline{X}_O$, to the observations. See Table 3 for results. We plot the performance ratios as a function of $\rho_u(1)$ in Figures 5_A, 5_B, and 5_C. The range of the misspecified correlation model is:

$$0 \leq \rho_u(1) \leq 0.9048.$$

The three "true" underlying correlation models that we consider are:

$$\rho_{u_O}(1) = 0 \quad (u_O = \infty), \quad \text{Figure 5}_A$$

$$\rho_{u_O}(1) = 0.3679 \quad (u_O = 1), \quad \text{Figure 5}_B$$

$$\rho_{u_O}(1) = 0.9048 \quad (u_O = 0.1), \quad \text{Figure 5}_C$$

When the misspecified u value equals ∞ , we calculate $\text{Var}_{u_O} Z_{u_O}^*$ from (1.9), and in our setup of observations and with the above chosen correlation function:

$$\sum_i \sum_j \rho_{ij} = 4 + 8e^{-u_O} + 4e^{-2u_O}.$$

Calculation from equation (1.8) results in $\text{Var}_{u_O} Z_{u_O}^*$, as $u_O = \infty$,

$$\text{Var}_{u_O \rightarrow \infty} Z_{u_O}^*(\cdot) = \frac{5}{4},$$

Figure 5A,B,C: Performance ratio vs. misspecified correlation. $n = 4$.

for $n = 4$ observations. It follows that in the case of true uncorrelated observations, $\rho_{u_o}(1) \downarrow 0$, ($u_o \uparrow \infty$) when the misspecified correlations vary over the entire range, the suboptimal variances have to vary between the value 2 and 1.25, (for any interpolated location $\underline{X}_o$) and the performance ratio varies between $\sqrt{2/(5/4)} = 1.265$ to 1. See Table 3.

Under $\rho_{u_o}(1) = 0$, ($u_o = \infty$) we encounter a noticeable error by falsely assuming existing correlations. The performance ratio is 1.10 for an interpolation location close to an observation point, ~ 1.20 , for further away locations of $\underline{X}_o$ and $\Xi = 1$, if $\underline{X}_o$ happens to be the center of the square grid. (Figure 5_A.)

Looking at Figure 5_B and Table 3 for "moderate" "true" correlations, i.e., for $\rho_{u_o}(1) = 0.3679$ ($u_o = 1$), the only effective deviation from the optimal standard deviation for the estimate, occurs when the misspecified correlation assumption is in the direction of assuming less correlation than the "true" model. Substantial deviation occurs only if we falsely assume essentially uncorrelated observations instead of the "true" model. The performance ratio plotted versus $\rho_u(1)$ in Figure 5_B may achieve larger values than those achieved for a true uncorrelated model, but for a substantially large range of misspecifications we do not get variances much higher than the optimal solution. At all three estimation points we considered, the worst ratio achieved is 3.8792 at $\underline{X}_o = (0.1, 0.1)$ for the extreme misspecification of $\rho_u(1) = 0$ instead of $\rho_{u_o}(1) = 0.37$.

TABLE 3 Three Different Correlation Models for Three Interpolation Locations.

1. $\text{Var}_{u_0} Z_u^*(\cdot)$; 2. $(\text{Var}_{u_0} Z_u^*(\cdot) / \text{Var}_{u_0} Z_{u_0}^*(\cdot))^{1/2}$; $n=4$; $\rho_u(h) = e^{-uh^2}$, $u > 0$.

Misspecified		True $\rho_{u_0}(1) = 0$ $u_0 = \infty$			True $\rho_{u_0}(1) = 0.3679$ $u_0 = 1$			True $\rho_{u_0}(1) = 0.9048$ $u_0 = 0.1$		
$\rho_u(1)$	u	(0.5,0.5)	(0.1,0.1)	(0.5,0.)	(0.5,0.5)	(0.1,0.1)	(0.5,0.)	(0.5,0.5)	(0.1,0.1)	(0.5,0.)
0.0	∞	5/4 1.	5/4 1.	5/4 1.	0.2547 1.	0.4383 3.8792	0.4025 1.8150	0.0046 1.	0.06161 3.2315	0.7808 6.6418
0.0498	3.	5/4 1.	1.8743 1.2245	1.4525 1.0780	0.2547 1.	0.0310 1.0310	0.1345 1.0493	0.0046 1.	0.00204 1.8660	0.0025 1.1807
0.1353	2.	5/4 1.	1.8582 1.2192	1.5354 1.1083	0.2547 1.	0.0309 1.0180	0.1232 1.0040	0.0046 1.	0.00206 1.8742	0.0019 1.0247
0.3679	1.	5/4 1.	1.7976 1.1992	1.5741 1.1012	0.2547 1.	0.029127 1.	0.122180 1.	0.0046 1.	0.00135 1.5178	0.0024 1.1558
0.4066	0.90	5/4 1.	1.7880 1.1959	1.5723 1.1215	0.2547 1.	0.029155 1.0005	0.122182 1.0000	0.0046 1.	0.00124 1.4534	0.0023 1.1476
0.4724	0.75	5/4 1.	1.7723 1.1907	1.5670 1.1196	0.2547 1.	0.0293 1.0035	0.1222 1.0001	0.0046 1.	0.0011 1.3494	0.0022 1.1252
0.6065	0.50	5/4 1.	1.7427 1.1807	1.5517 1.1142	0.2547 1.	0.0301 1.0172	0.1225 1.0015	0.0046 1.	0.0008 1.1704	0.0020 1.0686
0.7788	0.25	5/4 1.	1.7091 1.1693	1.5289 1.1059	0.2547 1.	0.03188 1.0462	0.1250 1.0116	0.0046 1.	0.00062 1.0300	0.0182 1.0132
0.9048	0.10	5/4 1.	1.6874 1.1619	1.5122 1.0999	0.2547 1.	0.0335 1.0730	0.1250 1.0114	0.0046 1.	0.00059 1.	0.0018 1.

All other ratios have values much lower.

Looking at Figure 5_C and Table 3, where the underlying true correlation assumption is $\rho_{u_0}(1) = 0.9048$, ($u_0 = 0.1$), we realize that any misspecified assumption is attributing less correlation between points. The suboptimality resulting from misspecification of the parameter in the model is now substantial in extreme cases. The maximum value for the performance ratio in the given range of correlations ($0 \leq \rho_u(1) \leq 0.9048$) is 3.2315, ($u = \infty$) at $\underline{X}_0 = (0.1, 0.1)$ and 6.6418, ($u = \infty$), at $\underline{X}_0 = (0.5, 0.)$. The performance ratio is not monotonically dropping to 1, but achieves local maxima, for intermediate misspecified u values. For example: at $\rho_u(1) = 0.1353$, ($u = 2$) the performance ratio is 1.874, at $\underline{X}_0 = (0.1, 0.1)$, gradually dropping till 1.

2. Effects on Variance Calculations of Misspecified Parameter in the Correlation Model.

The variance we would calculate for an unbiased estimate Z^* using a misspecified parameter u in the correlation model $\rho_u(h) = \rho_{ij}$ (between two points $\underline{X}_i, \underline{X}_j$ where $h = |\underline{X}_i - \underline{X}_j|$), is denoted by $\text{Var}_u Z^*$ and referred to as the calculated variance. If Z^* is the estimate which would be optimal (in the sense explained in equation (1.1)) for the misspecified u , the calculated variance is denoted $\text{Var}_u Z_u^*$. If u_0 denotes the true correlation parameter, then we are interested in the ratio

$$(2.1) \quad \sqrt{\frac{\text{Var}_u Z_u^*}{\text{Var}_{u_0} Z_{u_0}^*}} = \sqrt{\frac{\text{Calculated variance of } Z_u^*}{\text{experienced or true variance of } Z_{u_0}^*}}$$

The calculated variance will be larger than the true-experienced variance if $\rho_u(1) < \rho_{u_0}(1)$, ($u > u_0$), i.e., using correlations which are too weak, and smaller - if $\rho_u(1) > \rho_{u_0}(1)$, ($u < u_0$), i.e., using correlations which are too strong.

The magnitude of the ratio (2.1) is illustrated with a correlation model $\rho(h) = e^{-uh^2}$, $u > 0$, h = distance between points, for interpolation locations inside the triangle given in Figure 1. We would hope that the ratio is close to one. However, misspecification of the correlation model in the variance calculation may lead to appreciable errors.

We will list the outcome for two locations $\underline{X}_0 = (0., 0.)$ and $\underline{X}_0 = (0., 0.6)$ in Tables 4_A, 4_B and Figure 6_A and 6_B. For the true

TABLE 4_A 1. $\text{Var}_{u_0} Z_u^*$; 2. $(\text{Var}_{u_0} Z_u^* / \text{Var}_{u_0} Z_u^*)^{1/2}$
 $n = 3$; $\underline{X}_0 = (0., 0.)$; $\rho_u(h) = e^{-uh^2}$.

Misspecified		True $\rho_{u_0}(1)$							
$\rho_u(1)$	u	0.9048 $u_0=0.1$	0.7788 $u_0=0.25$	0.6065 $u_0=0.5$	0.4724 $u_0=0.75$	0.3679 $u_0=1$	0.1353 $u_0=2$	0.0498 $u_0=3$	0.0 $u_0=\infty$
0.9048	0.1	<u>0.001799</u> 1.	0.01057 0.4124	0.0383 0.2168	0.07812 0.1517	0.1263 0.1193	0.3546 0.0712	0.5801 0.0557	1.4999 1/(28.8746)
0.7788	0.25	0.001799 2.4245	<u>0.01057</u> 1.	0.0383 0.5256	0.07812 0.3679	0.1263 0.2893	0.3545 0.1727	0.5800 0.1350	1.4993 1/(11.9099)
0.6065	0.5	0.001800 4.6111	0.01057 1.9022	<u>0.0383</u> 1.	0.07811 0.6999	0.1263 0.5505	0.3544 0.3286	0.5796 0.2569	1.4974 1/(6.2527)
0.4724	0.75	0.001803 6.5806	0.01058 2.7166	0.0383 1.4285	<u>0.07810</u> 1.	0.12626 0.7865	0.3541 0.4696	0.5789 0.3673	1.4943 1/(4.3741)
0.3679	1	0.001813 8.3438	0.01060 3.4503	0.0383 1.8155	0.07811 1.2713	<u>0.1262</u> 1.	0.3537 0.5974	0.5781 0.4673	1.4900 1/(3.4361)
0.1353	2	0.00200 13.2777	0.0110 5.6552	0.0390 3.0076	0.07891 2.1144	0.1270 1.6669	<u>0.3528</u> 1.	0.5740 0.7839	1.4644 1/(2.0373)
0.0498	3	0.002618 14.7799	0.0124 6.7775	0.0414 3.7152	0.0820 2.6412	0.1303 2.0948	0.3546 1.2699	<u>0.5719</u> 1.	1.4328 (1/1.5828)
0.0	∞	0.0176 8.7039	0.0473 5.3093	0.1028 3.6014	0.1630 2.8601	0.2253 2.4327	0.4661 1.6913	0.6664 1.4145	<u>4/3</u> 1.

underlined variances are $\text{Var}_{u_0} Z_u^*$ (nominators for comparison ratio)

TABLE 4_B 1. $\text{Var}_{u_0} Z_u^*$; 2. $(\text{Var}_u Z_u^* / \text{Var}_{u_0} Z_u^*)^{1/2}$
 $n = 3$; $\underline{X}_0 = (0., 0.6)$; $\rho_u(h) = e^{-uh^2}$, $u > 0$.

Misspecified									
$\rho_u(1)$	u	0.9048	0.7788	0.6065	0.4724	0.3679	0.1353	0.0498	0.0
		$u_0=0.1$	$u_0=0.25$	$u_0=0.5$	$u_0=0.75$	$u_0=1$	$u_0=2$	$u_0=3$	$u_0=\infty$
0.9048	0.1	<u>0.00113</u>	0.00657	0.0235	0.0474	0.0758	0.2058	0.3301	1.5333
		1.	0.4144	0.2192	0.1543	0.1219	0.0740	0.0584	1/(36.8361)
0.7788	0.25	0.00114	<u>0.00654</u>	0.0233	0.0470	0.0752	0.2043	0.3281	1.5425
		2.3982	1.	0.5299	0.3731	0.2950	0.1790	0.1412	1/(15.3576)
0.6065	0.5	0.00119	0.00660	<u>0.0232</u>	0.0466	0.0745	0.2022	0.3253	1.5574
		4.4104	1.8750	1.	0.7055	0.5582	0.3387	0.2671	1/(8.1932)
0.4724	0.75	0.00129	0.00675	0.0233	<u>0.0465</u>	0.0741	0.2007	0.3231	1.5716
		6.0104	2.6240	1.4130	1.	0.7922	0.4813	0.3793	1/(5.8136)
0.3679	1	0.00141	0.00698	0.0235	0.0466	<u>0.0740</u>	0.1996	0.3215	1.5849
		7.2449	3.2561	1.7734	1.2599	1.	0.6087	0.4796	1/(4.6279)
0.1353	2	0.00196	0.00809	0.0250	0.0480	0.0749	<u>0.1983</u>	0.3189	1.6259
		10.0514	4.9517	2.8163	2.0331	1.6266	1.	0.7885	1/(2.8634)
0.0498	3	0.00227	0.00874	0.0260	0.0490	0.0759	0.1985	<u>0.3186</u>	1.6437
		11.8318	6.0369	3.5033	2.5499	2.0488	1.2669	1.	1/(2.2714)
0.0	∞	0.0202	0.0528	0.1114	0.1722	0.2330	0.4512	0.6135	<u>4/3</u>
		8.1244	5.0252	3.4596	2.7826	2.3922	1.7190	1.4742	1.

Figure 6A: Ratio (2.1) vs. misspecified correlation at unit distance. $n = 3$.

Figure 6B: Comparison ratio vs. misspecified correlation, $n = 3$

correlation model $\rho_{u_0}(1) = 0$, (i.e., $u_0 = \infty$), the true-experienced variance (with misspecification of the parameter in the optimal estimate Z_u^*) and the falsely calculated variance as $u = \infty$ are derived from equations (1.9) and (1.8) respectively. These limit values have already been calculated for Tables 2_A and 2_B. The ratio (2.1) for extreme cases is: In the case where the true model of correlation is a model of uncorrelated observations, $\rho_{u_0}(1) = 0$, and we falsely specify a strong correlation model in estimation procedure, $\rho_u(1) = 0.9048$, then the ratios (2.1) at the two locations $\underline{x}_0 = (0., 0.)$ and $\underline{x}_0 = (0., 0.6)$ are $1/(28.87)$ and $1/(36.84)$ respectively.

On the other hand, if the true correlation model is $\rho_{u_0}(1) = 0.9048$ and $\rho_u(1) = 0$, then the ratios (2.1) at the above mentioned locations are: 8.7039 and 8.1244 respectively. Even if we misspecify the parameter of the correlation model by a very small amount, for example, if $\rho_{u_0}(1)$ is 0.4724 and the misspecified model has $\rho_u(1) = 0.3679$, then the error in the ratio of the standard deviations from Table 4_B is already 1.26. Once again, if $\rho_{u_0}(1)$ is 0.4724, but now we overestimate the correlations, i.e., $\rho_u(1) = 0.6067$, then the ratio of the standard deviations is 0.7055, i.e., $1/(1.42)$. An appreciable error appears even for modest misspecifications, especially when misspecification is in the direction of overestimating correlations. This conclusion contrasts with the conclusion of chapter 1. There we saw that the

estimate Z_u^* , calculated using a misspecified parameter u , is very robust in the sense that the actual variance of the suboptimal estimate Z_u^* is considerably closer to the optimal variance.

Similar calculations have been done for interpolation between four observations on a square as described in Figure 4. The results are summed up in Table 5_A, and Figure 7, for the interpolation point $\underline{X}_0 = (0.5, 0.5)$. This specific case being the center of the unit square has all four coefficients identical ($\equiv 1/4$) for any underlying correlation model. But yet, using a misspecified parameter in the correlation model when calculating the variance estimate, results in noticeable overestimate of $\text{Var}_u Z_u^*$ if the misspecification was of the form $\rho_u(1) < \rho_{u_0}(1)$ and results in an underestimate of $\text{Var}_u Z_u^*$ if $\rho_u(1) > \rho_{u_0}(1)$. The standard deviation ratio (2.1) for this configuration (in either direction of misspecification, underestimation as well as overestimation) is as large as 16.

For the interpolation points $\underline{X}_0 = (0.1, 0.1)$, and $\underline{X}_0 = (0.5, 0.)$, at three different true underlying models ($\rho_{u_0}(1) = 0.9048, 0.3679$ and 0) we get the results that are summarized in Table 5_B. The graphs of the ratios versus $\rho_u(1)$ are similar in shape to Figure 7. At these locations the discrepancies are even higher, than those achieved at the center of the square.

It is very important to have the precise knowledge of correlations when calculating the estimation variance $\text{Var}_u Z_u^*(\cdot)$, which is heavily dependent on the parameter value. The estimate $Z_u^*(\cdot)$

TABLE 5: 1. $\text{Var}_{u_0} Z_u^*$ (same value for all u values) ; 2. $(\text{Var}_u Z_u^* / \text{Var}_{u_0} Z_u^*)^{1/2}$

$$n = 4 ; \underline{X}_0 = (0.5, 0.5) \quad \rho_u(h) = e^{-uh^2} ; u > 0 .$$

Misspecified		True $\rho_{u_0}(1)$							
$\rho_u(1)$	u	$u_0=0.1$	$u_0=0.25$	$u_0=0.5$	$u_0=0.75$	$u_0=1$	$u_0=2$	$u_0=3$	$u_0=\infty$
0.9048	0.1	<u>0.004642</u> 1.	0.02604 0.4222	0.087633 0.2301	0.167387 0.1665	0.254712 0.1350	0.586488 0.08896	0.829253 0.0748	5/4 0.0609
0.7788	0.25	0.004642 2.3683	<u>0.02604</u> 1.	... 0.4825	... 0.3944	... 0.3197	... 0.2107	... 0.1772	... 0.1443
0.6065	0.5	... 4.3447	... 2.0726	<u>0.087633</u> 1.	... 0.7235	... 0.5865	... 0.3865	... 0.3251	... 0.2648
0.4724	0.75	... 6.0046	... 2.5354	... 1.3821	<u>0.167387</u> 1.	... 0.8106	... 0.5342	... 0.4493	... 0.3659
0.3679	1	... 7.4071	... 3.1275	... 1.7049	... 1.2335	<u>0.254712</u> 1.	... 0.6590	... 0.5542	... 0.4514
0.1353	2	... 11.2396	... 4.7458	... 2.5870	... 1.8718	... 1.5174	<u>0.586488</u> 1.	... 0.8410	... 0.6850
0.0498	3	... 13.3649	... 5.6432	... 3.0762	... 2.2258	... 1.8043	... 1.1891	<u>0.829253</u> 1.	... 0.8145
0.0	∞	0.004642 16.4098	0.02604 6.9284	0.087633 3.7768	0.167387 2.7327	0.254712 2.2153	0.586488 1.4599	0.829253 1.2277	<u>5/4</u> 1.

Figure 7: Comparison ratio (2.1) vs. misspecified correlation model at unit distance. $n = 4$.

TABLE 5_B: 1. $\text{Var}_{u_0} Z_u^*$; 2. $(\text{Var}_{u_0} Z_u^* / \text{Var}_{u_0} Z_u^*)^{1/2}$
 $n = 4$; $\rho(h) = e^{-uh^2}$

Misspecified		True $\rho_{u_0}(1) = 0$ $u_0 = \infty$			True $\rho_{u_0}(1) = 0.3679$ $u_0 = 1$			True $\rho_{u_0}(1) = 0.9048$ $u_0 = 0.1$		
$\rho_u(1)$	u	(0.5,0.5)	(0.1,0.1)	(0.5,0.)	(0.5,0.5)	(0.1,0.1)	(0.5,0.)	(0.5,0.5)	(0.1,0.1)	(0.5,0.)
0.0	∞	5/4 1.	5/4 1.	5/4 1.	0.2547 2.2153	0.4383 1.6888	0.4025 1.7623	0.0046 16.4845	0.06161 4.5043	0.7808 1.2653
0.0498	3	... 0.8145	1.8743 0.2424	1.4525 0.6306	... 1.8043	0.0310 1.8864	0.1345 2.0724	... 13.3649	0.00204 7.3436	0.0025 15.2907
0.1353	2	... 0.6850	1.8582 0.1947	1.5354 0.4798	... 1.5174	0.0309 1.5275	0.1232 1.6935	... 11.2396	0.00206 5.8456	0.0019 13.7772
0.3679	1	... 0.4514	1.7976 0.1273	1.5741 0.2786	0.2547 1.	0.029127 1.	0.122180 1.	... 7.4071	0.00135 4.6419	0.0024 7.1822
0.4066	0.9	... 0.4189	1.7880 0.1185	1.5723 0.2552	... 0.9280	0.029155 0.9284	0.122182 0.9156	... 6.8737	0.00124 4.5026	0.0023 6.6233
0.4724	0.75	... 0.3659	1.7723 0.1043	1.5670 0.2188	... 0.8106	0.0293 0.8111	0.1222 0.7836	... 6.0046	0.0011 4.2497	0.0022 5.7814
0.6065	0.5	... 0.2648	1.7427 0.0770	1.5517 0.1539	... 0.5865	0.0301 0.5853	0.1225 0.5475	... 4.3447	0.0008 3.5840	0.0020 4.2592
0.7788	0.25	... 0.1443	1.7091 0.0432	1.5289 0.0819	... 0.3197	0.03188 0.3162	0.1250 0.2881	... 2.3683	0.00062 2.2630	0.00182 2.3749
0.9048	0.1	5/4 0.0609	1.6874 0.0186	1.5122 0.0342	0.2547 0.135	0.0335 0.1323	0.1250 0.1191	0.0046 1.	0.00059 1.	0.001773 1.

is not very sensitive to misspecified models, but the variance estimate is.

To sum up the results of the two previous chapters, we can conclude that:

where u_o is the "true" parameter of the model of correlation and u is the misspecified parameter of the model of correlation and $Z^*(\cdot)$ is the linear estimate (interpolated value) for $Z(\cdot)$.

3. Taylor Approximation for the Variance of a Suboptimal Linear

Estimation Procedure.

a. Derivation of Approximation.

In chapter 1 we constructed suboptimal unbiased linear interpolation schemes through misspecification of the parameter u of the autocorrelation model. The variance of these suboptimal schemes (1.3) was denoted $\text{Var}_{u_0} Z_u^*$, where u_0 is the true parameter value, and u the misspecified parameter under which suboptimal coefficients for estimator $Z^*(\underline{X}_0)$ were derived.

We shall denote $\text{Var}_{u_0} Z_u^*$ as $\sigma^2(u; u_0)$, bearing in mind that σ^2 is a function of the misspecified value u , as well as the true parameter u_0 , which actually brings in the influence of the interpolation location $\underline{X}_0$ on $\sigma_{\underline{X}_0}^2$. We are able to express $\sigma^2(u; u_0)$ in its Taylor approximation around u_0 . The first derivative of $\sigma^2(u; u_0)$ achieves its minimum value at $u = u_0$, and we have to express $\sigma^2(u; u_0)$ at least to the second order Taylor approximation.

$$(3.1) \quad \sigma^2(u; u_0) = \sigma^2(u_0; u_0) + \frac{(u-u_0)^2}{2!} (\sigma^2(u; u_0))'' \Big|_{u=u_0} + \dots + R_n$$

R_n is the remainder term. As a simple illustration we shall use the triangular data configuration described in Figure 1. We then compare the Taylor approximation with the calculated estimation variances given in Tables 2_A and 2_B. We shall rewrite the variance equation (1.3), indicating specifically the u dependent variables.

The interpolation point is $\underline{X}_0$, and observation locations

$\underline{X}_i, i = 1, \dots, n$ then

$$(3.2) \quad \sigma^2(u; u_0) = 1 + \sum_{i=1}^n \sum_{j=1}^n c_i^*(u) c_j^*(u) \rho_{u_0}(|\underline{X}_i - \underline{X}_j|) - 2 \sum_{i=1}^n c_i^*(u) \rho_{u_0}(|\underline{X}_i - \underline{X}_0|).$$

$\rho_{u_0}(|\underline{X}_i - \underline{X}_j|)$ is the correlation (function of u_0) between observation $Z_i = Z(\underline{X}_i)$ and $Z_j = Z(\underline{X}_j)$, which are $|\underline{X}_i - \underline{X}_j| = h_{ij}$ distance units apart. Let us denote $\rho_u(|\underline{X}_i - \underline{X}_j|) = \rho_u(h_{ij}) = \rho_u(i, j)$.

$\{c_i^*(u)\}$ is the set of optimal constrained coefficients derived from minimizing the estimation variance of an unbiased estimator

$Z^*(\underline{X}_0)$ of $Z(\underline{X}_0)$. The constraint on the coefficients is

$\sum_{i=1}^n c_i^*(u) = 1$. As mentioned before, to be able to approximate $\sigma^2(u; u_0)$ by a Taylor series, we need to get an expression for the second term of the Taylor series.

Theorem.

For $n = 3$ observation points in the configuration given in Figure 1, and evaluating at point $u = u_0$ we get:

$$(3.3) \quad \left(\sigma^2(u_0; u_0) \right)' = \frac{4}{3} [\phi_1^2 + \phi_2^2 + \phi_3^2 - \phi_1\phi_2 - \phi_1\phi_3 - \phi_2\phi_3]$$

where

$$(3.4) \quad \phi_i = \frac{1}{\sqrt{1-b(u)}} \left(\rho_u'(i, 0) + \frac{b'(u) \rho_u(i, 0)}{1-b(u)} \right) \Big|_{u=u_0}, \quad \forall i.$$

In the above observation configuration
 $\rho_u(1,2) = \rho_u(1,3) = \rho_u(2,3) = b(u) = b$. ϕ_i depends on the estimation point $\underline{X}_0$ and the true parameter u_0 for the underlying correlation model. The derivatives are with respect to u .

Proof:

Differentiating once $\sigma^2(u;u_0)$ which is given in (3.1) results in:

$$(3.5) \quad (\sigma^2(u;u_0))' = \sum_i \sum_j C_i^{*'}(u) \left[\sum_{j=1}^n (C_j^*(u) \cdot \rho_u(i,j)) - \rho_u(i,0) \right] .$$

Let us rewrite (3.5) as a function of unconstrained coefficients $\{C_i\}$, where the relationship between $\{C_i^*\}$ and $\{C_i\}$ is:

$$(3.6) \quad C_i^*(u) = C_i(u) + \frac{1 - \sum_{j=1}^n C_j(u)}{n} .$$

This relationship holds as long as the sum of each row (or column) in the correlation matrix is a constant $\forall i$. Differentiating (3.6) we get:

$$(3.7) \quad C_i^{*'}(u) = C_i'(u) - \frac{1}{n} \sum_{j=1}^n C_j'(u) .$$

Inserting (3.6) (3.7) into (3.5) results in:

(3.8)

$$\begin{aligned} (\sigma^2(u; u_o))' &= 2 \sum_{i=1}^n \left\{ \left[C_i'(u) - \frac{1}{n} \sum_{m=1}^n C_m'(u) \right] \times \right. \\ &\quad \times \left. \left[\sum_{j=1}^n C_j(u) \rho_u(i, j) + \frac{1}{n} \sum_{j=1}^n \rho_u(i, j) - \frac{\sum_{m=1}^n C_m(u)}{n} \cdot \sum_{j=1}^n \rho_u(i, j) - \rho_u(i, o) \right] \right\} \end{aligned}$$

Equation (3.6) holds under the condition that the correlation matrix, denoted γ_u , is invariant under rotation of the observation points.

As a result

$$\begin{aligned} (3.9) \quad \sum_{i=1}^n \rho_u(i, j) &= \text{constant} = k, \quad \forall j \quad (\text{or equivalently} \\ \sum_{j=1}^n \rho_u(i, j) &= k, \quad \forall i) . \end{aligned}$$

In the above mentioned configuration with n observations

$k = 1 + (n-1)b(u)$. Inserting k into (3.8) we get:

$$\begin{aligned} (3.10) \quad (\sigma^2(u; u_o))' &= 2 \sum_{i=1}^n \left\{ \left[C_i'(u) - \frac{1}{n} \sum_{m=1}^n C_m'(u) \right] \times \right. \\ &\quad \times \left. \left[\frac{k}{n} \left(1 - \sum_{m=1}^n C_m(u) \right) + \left(\sum_{j=1}^n C_j(u) \rho_u(i, j) - \rho_u(i, o) \right) \right] \right\} \end{aligned}$$

Denoting $L_u = \frac{k}{n} \left(1 - \sum_{m=1}^n C_m(u) \right)$,

$$W_i(u) = \left(\sum_{j=1}^n C_j(u) \rho_u(i, j) - \rho_u(i, o) \right)$$

$$\bar{W}(u) = \frac{1}{n} \sum_{i=1}^n W_i(u)$$

and inserting these in (3.10) results in:

$$(3.11) \quad (\sigma^2(u; u_0))' = 2 \sum c_i'(u) (w_i(u) - \bar{w}(u))$$

Differentiating (3.11) once more results in:

$$(3.12) \quad (\sigma^2(u; u_0))'' = 2 \sum_{i=1}^n \left[c_i''(u) \cdot (w_i(u) - \bar{w}(u)) + c_i'(u) (w_i(u) - \bar{w}(u))' \right]$$

$\uparrow$
 at $u = u_0$ this term equals zero,

therefore does not have to be evaluated any further. The second term on the R.H.S. is

$$(3.13) \quad [w_i(u) - \bar{w}(u)]' = \sum_{j=1}^n c_j'(u) \left(\rho_{ij}(u) - \frac{k}{n} \right)$$

and hence, when evaluating the second derivative at $u = u_0$ we get:

$$(3.14) \quad (\sigma^2(u_0; u_0))'' = 2 \sum_{i=1}^n c_i'(u_0) \cdot \sum_{j=1}^n c_j'(u_0) \cdot \left(\rho_{ij}(u_0) - \frac{k}{n} \right)$$

For $n = 3$, $k = 1+2b(u)$, (3.14) may be written

$$(3.15) \quad (\sigma^2(u_0; u_0))'' = \frac{4}{3}(1-b(u_0)) \left[\sum_{i=1}^3 (c_i'(u_0))^2 - c_1'(u_0) \cdot c_2'(u_0) - c_1'(u_0) c_3'(u_0) - c_2'(u_0) c_3'(u_0) \right]$$

Minimizing $\sigma^2(u, u_0)$ without any constraint on the coefficients results in n linear equations w.r.t. the unconstrained coefficients $\{c_i\}$.

$$(3.16) \quad \sum_{i=1}^n c_i(u) \rho_u(k, i) - \rho_u(k, 0) = 0, \quad \forall k, \quad k = 1, 2, \dots, n.$$

In matrix notation:

$$Z_u \underline{c}(u) = \underline{\rho}_0,$$

$$\text{where } \underline{\rho}_0 = (\rho_u(1, 0), \rho_u(2, 0), \dots, \rho_u(n, 0))^T$$

$$\underline{c}(u) = (c_1(u), \dots, c_n(u))^T$$

and T denotes transpose.

$$(3.17) \quad \text{Det } Z_u = (1-b)(1+b-2b^2)$$

and the optimal solution w.r.t. the coefficients is

$$(3.18) \quad \underline{c}(u) = Z_u^{-1} \underline{\rho}_0 \quad \text{where}$$

$$(3.19) \quad Z_u^{-1} = \frac{1}{\text{Det } Z_u} = \begin{pmatrix} (1-b^2)\rho_u(1,0) + (b^2-b)\rho_u(2,0) + (b^2-b)\rho_u(3,0) \\ (b^2-b)\rho_u(1,0) + (1-b^2)\rho_u(2,0) + (b^2-b)\rho_u(3,0) \\ (b^2-b)\rho_u(1,0) + (b^2-b)\rho_u(2,0) + (1-b^2)\rho_u(3,0) \end{pmatrix}.$$

Hence

$$(3.20) \quad C_i(u) = \frac{\rho_u(i,0)}{1-b(u)} + \frac{b}{1+b-2b^2} \sum_{j=1}^3 \rho_u(j,0)$$

deriving and inserting $u = u_0$ implies:

(3.21)

$$C'_i(u_0) = \frac{1}{(1-b(u_0))^2} \left[\rho'_{u_0}(i,0)(1-b(u_0)) + b'(u_0)\rho_{u_0}(i,0) \right] + \Delta(u_0)$$

where $\Delta(u) = \frac{b(u)}{1+b(u)-2b^2(u)} \sum_{j=1}^3 \rho_u(j,0)$. $\Delta(u_0)$ needs no further

evaluation in calculating (3.15) as all the terms (squares and cross product) derived from $\Delta^2(u_0)$ appear in $(C'_i(u_0))^2$, and $C'_i(u_0)C'_j(u_0)$ with opposite signs. Calculating equation (3.15), using (3.20), (3.21) and denoting ϕ_i as in the conditions of the theorem, i.e., as given in (3.4), we obtain (3.3).

Q.E.D.

Finally, we get the Taylor approximation of $\sigma^2(u;u_0)$ around u_0 in the following form:

$$(3.22) \quad \sigma^2(u;u_0) \doteq \sigma^2(u_0,u_0) + \frac{(u-u_0)^2}{2!} \cdot \frac{4}{3} \cdot \left[\sum_{i=1}^3 \phi_i^2 - \phi_1\phi_2 - \phi_1\phi_3 - \phi_2\phi_3 \right]$$

with ϕ_1, ϕ_2, ϕ_3 , given by equation (3.4), and as stated before,

depending on the estimation location $\underline{X}_0$ and the true value of the correlation model parameter u_0 .

b. Examples Applying Taylor Expansion for Evaluating the Flatness of the Optimal Estimation Variance.

Let us continue our illustration with the family of isotropic correlation functions used in chapters 1 and 2, $\rho_u(h) = e^{-uh^2}$, $u > 0$. Then, for the triangular data configuration of Figure 1, substituting into equation (3.3) gives:

$$(3.23) \quad (\sigma^2(u_0; u_0))'' = \frac{8}{3} \left[e^{-u_0 h_3^2} \cdot (1+h_3^2) - e^{-u_0 h_1^2} \cdot (1+h_1^2) \right]^2$$

where $h_i = |\underline{X}_i - \underline{X}_0|$ and $\underline{X}_0$ is the interpolation location.

The Taylor expansion of $\sigma^2(u; u_0)$ around u_0 to second order is:

$$\sigma^2(u; u_0) = \sigma^2(u_0; u_0) + \frac{(u-u_0)^2}{2!} (\sigma^2(u_0; u_0))'' ;$$

$\sigma^2(u_0; u_0)$ is the optimal estimation variance obtained when the correlation parameter is correctly specified for calculating the estimate; $\sigma^2(u; u_0)$ is the estimation variance obtained using suboptimal coefficients based on a misspecified correlation parameter $u \neq u_0$.

The flatness of the optimum variance may be gauged by the ratio

$$(3.24) \quad (\sigma^2(u_0; u_0))'' / \sigma^2(u_0; u_0) .$$

Tables 6_A and 6_B - last row - show calculated values for this ratio at two estimation locations in the triangular data configuration of Figure 1. ($\underline{X}_0 = (0., 0.)$ and $\underline{X}_0 = (0., 0.6)$). The optimum is extremely flat when the correlation at unit distance is in the vicinity of 0.5. For example: at $\underline{X}_0 = (0., 0.)$, for $\rho_{u_0}(1) = 0.4724$, ratio (3.24) is 0.0524. For strong underlying correlation models, the optimum is not flat and is very sensitive to the specification. The ratio (3.24) in Table 6_A (estimation at location $\underline{X}_0 = (0., 0.)$) is as high as 242 at $\rho_{u_0}(1) = 0.9$. For weak underlying correlation $\rho_{u_0}(1) = 0.05$, the flatness indicator is 0.7686. In Table 6_B results are given for estimation at location $\underline{X} = (0., 0.6)$, which is closer to a data point. We again experience a sharp deviation from flatness of the optimum for the extreme case, with strongly correlated observations. For $\rho_{u_0}(1) = 0.9$, we get the convexity ratio (3.24) as high as 481, compared to the ratio for the almost uncorrelated true model - $\rho_{u_0}(1) = 0.05$ - , where the corresponding ratio is 3.09. Once again, the optimum is very flat for intermediate correlation models.

Comparing the last row of Table 6_A to the last row of Table 6_B, indicates that for an interpolation location close to a data point ($\underline{X}_0 = (0., 0.6)$) the curvature of the variance function at the true parameter value is greater than the corresponding curvature for an estimation location farther away from a data point, for very strong or very weak underlying correlation models.

TABLE 6_A: Variance of Suboptimal Estimates (VSE) Versus Taylor Expansion

$n = 3$; $\underline{X}_0 = (0., 0.)$

Misspecified $\rho_u(1)$ u		$\rho_{u_0}(1) = 0.9048$ $u_0 = 0.1$		$\rho_{u_0}(1) = 0.7788$ $u_0 = 0.25$		$\rho_{u_0}(1) = 0.6065$ $u_0 = 0.5$		$\rho_{u_0}(1) = 0.4724$ $u_0 = 0.75$		$\rho_{u_0}(1) = 0.3679$ $u_0 = 1$		$\rho_{u_0}(1) = 0.1353$ $u_0 = 2$		$\rho_{u_0}(1) = 0.0498$ $u_0 = 3$	
		VSE	Taylor	VSE	Taylor	VSE	Taylor	VSE	Taylor	VSE	Taylor	VSE	Taylor	VSE	Taylor
0.9048	0.1	<u>0.001799</u>	<u>0.001799</u>	0.01057	0.01286										
0.7788	0.25	0.001799	0.006705	<u>0.01057</u>	<u>0.01057</u>	0.0383	0.039127	0.07812	0.07861						
0.6065	0.5			0.01057	0.01694	<u>0.0383</u>	<u>0.0383</u>	0.07811	0.07823	0.1263	0.13389				
0.4724	0.75					0.0383	0.039127	<u>0.07810</u>	<u>0.07810</u>	0.12626	0.12710				
0.3679	1					0.0383	0.041609	0.07811	0.07823	<u>0.1262</u>	<u>0.1262</u>	0.3537	0.5330		
0.1353	2							0.07891	0.08130	0.1270	0.15496	<u>0.3528</u>	<u>0.3528</u>	0.5740	0.7917
0.0498	3											0.3546	0.5330	<u>0.5719</u>	<u>0.5719</u>
$(\sigma^2(u_0, u_0))''$			0.43613		0.20392		0.02647		0.00409		0.0575		0.36051		0.43958
$\frac{(\sigma^2(u_0, u_0))''}{\sigma^2(u_0, u_0)}$			242.4291		19.2923		0.6911		0.0524		0.4556		1.0218		0.7686

TABLE 6_B: Variance of Suboptimal Estimates (VSE) Versus Taylor Expansion

$n = 3$; $\underline{X}_0 = (0., 0.6)$

Misspecified		$\rho_{u_0}(1) = 0.9048$ $u_0 = 0.1$		$\rho_{u_0}(1) = 0.7788$ $u_0 = 0.25$		$\rho_{u_0}(1) = 0.6065$ $u_0 = 0.5$		$\rho_{u_0}(1) = 0.4724$ $u_0 = 0.75$		$\rho_{u_0}(1) = 0.3679$ $u_0 = 1$		$\rho_{u_0}(1) = 0.1353$ $u_0 = 2$		$\rho_{u_0}(1) = 0.0498$ $u_0 = 3$	
$\rho_u(1)$	u	VSE	Taylor	VSE	Taylor	VSE	Taylor	VSE	Taylor	VSE	Taylor	VSE	Taylor	VSE	Taylor
0.9048	0.1	<u>0.00113</u>	<u>0.00113</u>	0.00657	0.009813										
0.7788	0.25	0.00114	0.00724	<u>0.00654</u>	<u>0.00654</u>	0.0233	0.02515	0.0470	0.046504						
0.6065	0.5			0.00660	0.01563	<u>0.0232</u>	<u>0.0232</u>	0.0466	0.046501	0.0745	0.07903				
0.4724	0.75					0.0233	0.02516	<u>0.0465</u>	<u>0.0465</u>	0.0741	0.07525				
0.3679	1					0.0235	0.03103	0.0466	0.046501	<u>0.0740</u>	<u>0.0740</u>	0.1996	0.4733		
0.1353	2							0.0480	0.046525	0.0749	0.09411	<u>0.1983</u>	<u>0.1983</u>	0.3189	0.81097
0.0498	3											0.1985	0.4733	<u>0.3186</u>	<u>0.3186</u>
$(\sigma^2(u_0, u_0))''$			0.54362		0.290906		0.062608		0.0000325		0.040223		0.549958		0.98474
$\frac{(\sigma^2(u_0, u_0))''}{\sigma^2(u_0, u_0)}$			481.0796		44.4810		2.6986		0.0007		0.5435		2.7734		3.0908

But, on the other hand, once we deal with intermediate correlation values, the interpolation point close to a data point shows much flatter optimality of the variance than an estimation point further away from a data point. For example, if $\rho_{u_0}(1) = 0.4724$ the flatness of the optimum as measured by (3.24) at $\underline{x}_0 = (0., 0.6)$ close to a data point $\underline{x}_1 = (0., \frac{\sqrt{3}}{2})$, is 0.0007 compared to 0.0524 for the same underlying correlation, at $\underline{x}_0 = (0., 0.)$. The flatness of the optimum may also be judged directly by inspecting the variances in Tables 2_A and 2_B. The conclusions are similar to the extent that the second order Taylor expansion for $\sigma^2(u; u_0)$ is a good approximation for the variance of the suboptimal estimators.

A few examples of this approximation are shown in the upper part of Tables 6_A and 6_B for interpolation location $\underline{x}_0 = (0., 0.)$ and $\underline{x}_0 = (0., 0.6)$ respectively. Approximation is done for values near u_0 . Under very weak or very strong correlation models, such as $\rho_{u_0}(1) = 0.0498$ or $\rho_{u_0}(1) = 0.9048$, the Taylor approximation may not be accurate. On the other hand, for moderate underlying correlation models, the Taylor approximation is extremely close to the true value.

4. Robustness of the Linear Estimate under Misspecification of the Functional Form of the Isotropic Correlation Model.

To illustrate we shall choose four data points on a square and interpolate between them as in Figure 4. Further, we take the true correlation model to be $f_0 = e^{-uh^2} (= \rho_0(h))$, where u is a parameter and h is distance. The misspecified model of the correlation function we take to be $f = 1/(1+wh^2) (= \rho(h))$, where w is a parameter.

From a given realization of data on a unit grid it might be a fairly easy task to estimate $\rho(1)$, (correlation at unit distance). It makes sense, therefore, to adjust the parameters of the two above mentioned correlation functions so that $\rho_0(1) = \rho(1)$. See Figure 11 for the above condition. This adjustment results in:

$$(4.1) \quad w = e^u - 1.$$

From now on we shall refer to the true and false functions as

$$(4.2) \quad \begin{aligned} f_0 &= e^{-uh^2} = \rho_0(h) \quad \text{and} \\ f &= 1/[1 + (e^u - 1)h^2] = \rho(h) \end{aligned}$$

respectively. Let $Z_{f_0}^*(\underline{X}_0)$, $Z_f^*(\underline{X}_0)$ denote linear estimators at an interpolation point $\underline{X}_0$, calculated according to (1.1), using respectively f_0 and f (4.2) for the correlation functions. If f_0 is taken to be the true correlation function, then the two estimators

are respectively, optimal and suboptimal. If the variances of these estimates are denoted $\text{Var}_{f_o} Z_{f_o}^*$ and $\text{Var}_{f_o} Z_f^*$, then we are interested in the ratio

$$(4.3) \quad R1 = \sqrt{\frac{\text{Var}_{f_o} Z_f^*}{\text{Var}_{f_o} Z_{f_o}^*}}$$

as a measure of robustness to misspecification of the correlation functions f_o in the computation of the interpolated estimates. The ratio of the incorrectly calculated variance $\text{Var}_f Z_f^*$ versus the true estimation variance of the suboptimal estimate is

$$(4.4) \quad R2 = \sqrt{\frac{\text{Var}_f Z_f^*}{\text{Var}_{f_o} Z_{f_o}^*}}$$

Table 7 shows the relevant variances needed in $R1$ and $R2$. Table 8, Figure 8 and Figure 9 show the above defined ratios and their graphs plotted versus the correlation at unit distance.

Misspecifying the functional form in the calculation of estimation variances may result in a considerably large error, but not as severe as the error derived in chapter 2, when the misspecification occurred in the parameter of the given family of correlation functions. (See Tables 4_A , 4_B , 5_A , 5_B .) The above statement holds even for interpolating at the center of the square when the weighting coefficients are identically equal to $1/4$ whatever correlation function is used. The ratio $R2$ is greatest ($= 1.4$) for strong correlation models, and decreases to 1.0 for models with weak correlation.

TABLE 7: 1. $\text{Var}_{f_o} Z_f^*$ 2. $\text{Var}_{f_o} Z_{f_o}^*$ 3. $\text{Var}_f Z_f^*$

$$\left. \begin{aligned} f_o &= e^{-uh^2} \\ ff &= 1/(1+wh^2) \end{aligned} \right\} w = e^u - 1$$

$\rho_{f_o}(1)=\rho_f(1)$	u	$\underline{X}_o=(0.5,0.5)$	$\underline{X}_o=(0.5,0.4)$	$(\underline{X}_o=(0.5,0.))$	$\underline{X}_o=(0.2,0.2)$	$\underline{X}_o=(0.1,0.1)$
0.9048	0.1	4.642×10^{-3} 4.642×10^{-3} 8.888×10^{-3}	4.4594×10^{-3} 4.4563×10^{-3} 8.53×10^{-3}	1.7878×10^{-3} 1.7727×10^{-3} 3.5792×10^{-3}	1.8987×10^{-3} 1.8749×10^{-3} 3.5672×10^{-3}	5.9977×10^{-4} 5.8675×10^{-4} 1.1097×10^{-3}
0.7788	0.25	0.02604 0.02604 0.04754	0.02499 0.02497 0.04562	0.01032 0.01026 0.02141	0.01052 0.01033 0.01888	3.2942×10^{-3} 3.1876×10^{-3} 5.7864×10^{-3}
0.6065	0.5	0.0876 0.0876 0.1519	0.0840 0.0840 0.1462	0.0367 0.0367 0.0812	0.0346 0.03405 0.06153	0.0107 0.0103 0.0187
0.4723	0.75	0.1674 0.1674 0.2802	0.1603 0.1603 0.2708	0.0760 0.0750 0.1722	0.0650 0.06438 0.11926	0.0198 0.0193 0.0365
0.3678	1	0.2547 0.2547 0.4145	0.24416 0.2440354 0.4028508	0.1272 0.1222 0.2849	0.0986 0.09786 0.18900	0.0297 0.0291 0.0591
0.1353	2	0.5865 0.5865 0.8590	0.5699 0.5651 0.8497	0.4096 0.3535 0.7587	0.2562 0.23717 0.54958	0.0734 0.0704 0.2109
0.0498	3	0.8293 0.8293 1.0916	0.8148 0.8054 1.0877	0.6757 0.5777 1.0496	0.4643 0.36481 0.90399	0.1426 0.1102 0.4888
0.0	∞	1.25 1.25 1.25	1.25 1.25 1.25	1.25 1.25 1.25	1.25 1.25 1.25	1.25 1.25 1.25

TABLE 8: $R1 = (\text{Var}_{f_o} Z_f^* / \text{Var}_{f_o} Z_{f_o}^*)^{1/2}$; $R2 = (\text{Var}_f Z_f^* / \text{Var}_{f_o} Z_f^*)^{1/2}$

$\rho_{f_o}(1) = \rho_f(1)$	u	$\underline{X}_o = (0.5, 0.5)$		$\underline{X}_o = (0.5, 0.4)$		$\underline{X}_o = (0.5, 0.)$		$\underline{X}_o = (0.2, 0.2)$		$\underline{X}_o = (0.1, 0.1)$	
		R1	R2	R1	R2	R1	R2	R1	R2	R1	R2
0.9048	0.1	1	1.3837	1.00036	1.3830	1.0042	1.4149	1.0063	1.3707	1.0110	1.3602
0.7788	0.25	1	1.3512	1.0005	1.3509	0.0028	1.4403	1.0093	1.3395	1.0166	1.3253
0.6065	0.5	1	1.3167	1.00022	1.3191	1.00013	1.4863	1.0074	1.3342	1.0158	1.3252
0.4724	0.75	1	1.2938	1.0000	1.2997	1.0065	1.5052	1.0044	1.3549	1.0125	1.3587
0.3679	1	1	1.2757	1.0002	1.2845	1.0204	1.4964	1.0037	1.3845	1.0095	1.4105
0.1353	2	1	1.2102	1.0043	1.2210	1.0764	1.3611	1.0394	1.4645	1.0209	1.6949
0.0498	3	1	1.1473	1.0058	1.1554	1.0816	1.2463	1.1281	1.3953	1.1378	1.8512
0.0067	5	1	1.0616	1.0027	1.0676	1.0385	1.1321	1.2014	1.2092	1.6519	1.4670
$4.5 \cdot 10^{-5}$	10	1	1.0053	1.0001	1.0075	1.0029	1.0344	1.0827	1.1048	1.5746	1.2210
0.0	∞	1	1	1	1	1	1	1	1	1	1

Figure 8: Performance Ratio $R1$ versus correlation model at unit distance.

Figure 9: Ratio $R2$ versus $\rho_{u_0}(1) = \rho_u(1) = e^{-u}$.

Table 8 contains five different estimation locations in the unit square to indicate the clear dependence of the distance from the center as opposed to the distance from an observation point.

In chapter 1, when dealing with the variance $\text{Var}_{u_o} Z_u^*$ of a suboptimal estimator versus the optimal variance $\text{Var}_{u_o} Z_u^*$, the suboptimality was due to the misspecification of the parameter u in the given correlation function family. We found that suboptimality resulting from this misspecification was inconsequential.

Here as we observe R_1 , we reach a similar conclusion for a misspecified functional form. The suboptimal variance, $\text{Var}_{f_o} Z_f^*$ does not - by and large - differ much from the optimal variance.

There is little sense trying to fit a correlation function based on an estimated $\rho(1)$, when there is no additional information available about the correlation function at distances $h < 1$. Any interpolation procedure in the unit square should be based on a simple average. If at $h = 1$ the correlation is practically zero, then the only practical distances for which $f_o \neq f$ are $h \ll 1$. This means, that for interpolation close to a data point (where $h_i = |X_o - X_i|$ are relatively short) the correlation $\rho(i,o)$ are susceptible to large discrepancies resulting from misspecifying f_o by f . (Figure 10 - (4), (5).) Not only can $\rho(i,o)$ be very discrepant, but also the corresponding coefficient c_i may have much more weight than the other coefficients, as the distance h_i is small. At distances $h_i \ll 1$, the differences between f_o and f

Diagram of Relevant
distances engaged in
computing the estimate Z^*

- (a) $h_4 = h_2 \sim 1$
 $h_3 \sim 1.3 > 1$
 $h_1 = 0.14 \ll 1$

- (b) $h_1 = h_2 = 0.5 < 1$
 $h_3 = h_4 = \frac{5}{4} > 1$

Figure 10: Correlation function f and f_0 w.r.t. different underlying correlation models.

are huge. The derivatives f'_0 and f' are large, which results in a large $\Delta\rho$ for a small Δh , and causes huge errors of misspecification even if we deal with, practically, uncorrelated data.

In Table 8, we notice the large performance ratios for very weak correlation models. (R1 as well as R2.) The ratios are indicating noticeably larger errors for estimation locations close to data points, for weak correlation models, a phenomena which may be explained by the large differences between f_0 and f for short distances, as indicated in Figure 10-(5).

5. Misspecification of an Anisotropic Correlation Function.

A stochastic process $\{Z(\underline{X}), \underline{X} \in R^2\}$ is called an isotropic process if it is stationary and its autocorrelation function ρ is independent of direction. Let $\underline{X}_j$ and $\underline{X}_i$ be two points in R^2 ; then the correlation between $Z(\underline{X}_i)$ and $Z(\underline{X}_j)$ may be given by a function $\rho(h)$ where $h = |\underline{X}_j - \underline{X}_i|$. For illustration of comparisons with isotropic assumptions we shall use the following parametric family of anisotropic autocorrelation functions (dependent on direction):

$$(5.1) \quad \text{correlation } (Z(\underline{X}_i), Z(\underline{X}_j)) = \rho(\underline{h}) = e^{-u^*(s^2 + \alpha t^2)}$$

where $u^* > 1$ and $0 < \alpha < 1$ are parameters, and $\underline{h} = (s, t) = (\underline{X}_i - \underline{X}_j)$.

If $\alpha = 1$, we have an isotropic process. Strong anisotropy corresponds to small values of α . For example at $\alpha = 1/4$, the correlation at unit distance varies from .37 to .78 when $u^* = 1$. At $\alpha = 1/2$ the correlation at unit distance varies from .37 to .61 when $u^* = 1$. See Figure 11 for differences between isotropic (circular) and anisotropic (in our special case-elliptical) correlations.

It seems logical that in applications we may tend to simplify the correlation model, while the actual true model is more complex. Therefore, we shall use the anisotropic model of correlation as the true model with parameters u^* and α and the isotropic one - as the misspecified correlation functional form.

Let $Z_{f_0}^*$ denote the best linear unbiased estimator at an unobserved point X_0 , using the optimal data weighting coefficients derived from the underlying anisotropic correlation function. Its estimation variance, the optimal variance, is denoted by $\text{Var}_{f_0} Z_{f_0}^*$. Let Z_f^* denote the best linear unbiased estimator at an unobserved point X_0 using the suboptimal data weighting coefficients derived from the incorrectly specified isotropic correlation function. The variance of this suboptimal estimator is denoted $\text{Var}_f Z_f^*$. The presumed variance of the suboptimal estimator calculated from the incorrect isotropic model is denoted $\text{Var}_f Z_f^*$.

To compare the two correlation models $e^{-u^*(s^2 + \alpha t^2)}$; $0 < \alpha < 1$, and $e^{-u(s^2 + t^2)}$, we have to find a corresponding u value for every fixed u^* value in the anisotropic model. We do this by finding the average correlation (using the anisotropic function) on a circle of radius h and determining the u value for an isotropic correlation function of the form e^{-uh^2} that would result in the same average correlation. This average correlation is given by

$$(5.2) \quad \frac{1}{2\pi} \int_0^{2\pi} \exp[-u^* h^2 (\cos^2 \theta + \alpha \sin^2 \theta)] d\theta.$$

Hence the corresponding value of u will have e^{-uh^2} equal to expression (5.2). In general, this will depend on h but, as we shall see, the dependence on h is slight.

The expression (5.2) may be approximated numerically or by a convergent series (see footnote below). Instead, we have chosen an approximation which is the geometric mean of the maximum and minimum values of the integrand of (5.2), viz.

$$(5.3) \quad \left[e^{-u^* h^2} \cdot e^{-u^* h^2 \alpha} \right]^{\frac{1}{2}} = e^{-u^* h^2 (1+\alpha)/2} .$$

This has the advantage that the corresponding u value for the isotropic model does not depend on h , i.e.

$$(5.4) \quad u = u^* (1+\alpha)/2 .$$

To check whether $e^{-u^* h^2 (1+\alpha)/2}$ is a good approximation to the integral (5.2), we performed several numerical integrations of the integral for parameter values

$$\alpha = 0.1, 0.25, 0.5 ,$$

$$u^* = 0.1, 1, 3 ,$$

$$h = 0.5, 0.75, 1, 2 .$$

Footnote: The elliptical integral (5.2) may be approximated by a converging infinite series

$$e^{-u^* h^2} \left[1 + \frac{1}{\pi^{1/2}} \sum_{i=1}^{\infty} \frac{\sqrt{(i+\frac{1}{2})k^i}}{\sqrt{(i+1)^2}} \right]$$

where $k = u^* h^2 (1-\alpha)$.

The results are summarized in Table 9, for numerical integration with $N = 91$ points. These values are compared to the geometric mean of maximal and minimal values of the integrand.

Figure 11: Isotropic and Anisotropic Contours of Constant Correlation

Table 9: Numerical Integration of Elliptical Integral vs. Geometric Mean of Extreme Values

Aniso- tropic u *	h	$\alpha = 0.25$				$\alpha = 0.50$			
		Numerical Integration 1)	Isotropic u 2)	Geometric mean of max. & min. 3)	Isotropic u from Geom. mean 4)	N.I.	Isotropic u	Geometric mean of max. & min.	Isotropic u from Geom. mean
0.1	0.5	0.984	0.62	0.984	0.0625	0.981	0.075	0.981	0.075
	0.75	0.966	0.62	0.965	0.0625	0.959	0.075	0.959	0.075
	1.	0.940	0.62	0.939	0.0625	0.928	0.075	0.928	0.075
	2.	0.783	0.61	0.779	0.0625	0.743	0.074	0.741	0.075
1.	0.5	0.857	0.62	0.855	0.625	0.830	0.75	0.829	0.750
	0.75	0.711	0.60	0.704	0.625	0.659	0.74	0.656	0.750
	1.	0.554	0.59	0.535	0.625	0.480	0.73	0.472	0.750
	2.	0.136	0.50	0.082	0.625	0.063	0.69	0.050	0.750
3.	0.5	0.638	1.79	0.626	1.875	0.575	2.21	0.570	2.25
	0.75	0.384	1.70	0.348	1.875	0.295	2.17	0.282	2.25
	1.	0.206	1.58	0.153	1.875	0.121	2.11	0.105	2.25
	2.	0.010	1.15	0.0005	1.875	0.0006	1.85	0.0001	2.25

1) Integration carried out with $N = 91$ integration points in the first quadrant.

2) $u = [-\ln p(h)]/h^2$

3) $G.M. = [e^{-u^*h^2} \cdot e^{-u^*h^2\alpha}]^{1/2} = e^{-u^*h^2(1+\alpha)/2}$

4) $u = u^*(1+\alpha)/2$

The effect of incorrectly specifying an isotropic correlation function will be measured by the ratios

$$R1 = \sqrt{\frac{\text{Var}_f Z_o^*}{\text{Var}_f Z_o^*}} ; R2 = \sqrt{\frac{\text{Var}_f Z_f^*}{\text{Var}_f Z_f^*}}$$

For illustration we use the geometric configuration and coordinate systems of four observation points on a square (Figure 4).

The ratios R_1 and R_2 were calculated for four different estimation points inside the data square: (0.5,0.5) center, (0.1,0.1) a point close to observation point, (0.5,0.) and (0.,0.5). See Figure 4. The underlying anisotropic correlation parameters of the model (5.1) used in our illustrations were $\alpha = .1, .25, .5$ and $u^* = 0.1, .25, .5, .75, 1, 2, 3$. The misspecified isotropic correlation model, e^{-uh^2} has parameter $u = u^*(1+\alpha)/2$ in accordance with the preceding argument.

The ratio calculations are shown in Table 10. It turns out that interpolation of Z at point $\underline{X}_0$ is robust to a misspecified isotropic correlation replacing a true anisotropic correlation, but the calculation of the estimation variance is sensitive to such misspecifications.

In extreme cases like interpolating at the point $\underline{X}_0 = (0.,0.5)$ under very severe anisotropy ($\alpha = .1$) the suboptimality of an isotropic assumption may be pronounced. For $\alpha = 1/4$, we encounter very mild suboptimality ratios around $R1 = 1.05$. The errors of

$$f = e^{-u(s^2+t^2)} = e^{-uh^2}; \quad \bar{X}_0 = (s, t)$$

Table 10: $f_0 = e^{-u(s^2+t^2)}$ $0 < \alpha < 1$

$$R1 = \sqrt{\frac{\text{Var}_{f_0} Z^*}{\text{Var}_{f_0} Z^*}}; \quad R2 = \sqrt{\frac{\text{Var}_{f_0} Z^*}{\text{Var}_{f_0} Z^*}}$$

Anisotropic	α		Isotropic	$p(1)$		$(0.5, 0.5)$	$(0.1, 0.1)$		$(0.5, 0.)$		$(0., 0.5)$	
	u	n		$R1 \leq 1, R2$	$R1 > 1$		$R2$	$R1$	$R2$	$R1$	$R2$	$R1$
0.1	0.1	0.0550	0.8670	0.9217	1.	1.	0.8693	1.0001	0.5531	1.1446	4.7780	1.
	0.25	0.0625	0.9394	0.9277	0.075	0.1	0.9233	1.0002	0.6279	1.0167	2.4471	1.
	0.50	0.075	0.9277	0.9048	0.1	1.	0.9445	1.0002	0.7522	1.0018	1.4925	1.
	1	0.1	0.9048	0.8715	0.1375	0.815	0.8751	1.0003	0.5581	1.2954	4.1850	1.
0.25	0.1	0.1375	0.8693	0.9234	1.0001	1.	0.8751	1.0003	0.5581	1.2954	4.1850	1.
	0.25	0.1562	0.8553	0.9234	1.0001	1.	0.9277	1.0005	0.6326	1.0351	2.3861	1.
	0.5	0.1875	0.8290	0.9746	1.0001	1.	0.9765	1.0004	0.7560	1.0036	1.4824	1.
	1	0.25	0.7788	0.8715	0.1375	0.815	0.8751	1.0003	0.5581	1.2954	4.1850	1.
0.5	0.1	0.275	0.7596	0.7316	0.3125	0.6065	0.8852	1.0004	0.5674	1.4370	3.7164	1.
	0.25	0.3125	0.7316	0.6873	0.3750	0.6065	0.9352	1.0005	0.6413	1.0510	2.3202	1.
	0.5	0.3750	0.6873	0.6065	0.5	0.6065	0.9798	1.0004	0.7630	1.0046	1.4680	1.
	1	0.5	0.6065	0.5698	0.5625	0.5	0.9832	1.0003	0.7706	1.0040	1.4556	1.
0.75	0.1	0.4125	0.6619	0.6258	0.4687	0.5698	0.8954	1.0003	0.5776	1.4920	3.5248	1.
	0.25	0.4687	0.6258	0.5698	0.5625	0.5698	0.9429	1.0004	0.6509	1.0536	2.2843	1.
	0.5	0.5625	0.5698	0.4723	0.75	0.5698	0.9832	1.0003	0.7706	1.0040	1.4556	1.
	1	0.75	0.4723	0.5769	0.55	0.5769	0.9057	1.0001	0.5853	1.4915	3.4711	1.
1.	0.1	0.55	0.8806	0.9321	1.0007	1.	0.9057	1.0002	0.6612	1.0477	2.664	1.
	0.25	0.625	0.9321	0.9785	1.0007	1.	0.9505	1.0002	0.7786	1.0027	1.4438	1.
	0.5	0.750	0.9785	0.9838	1.0033	1.	0.9963	1.0008	0.8115	1.0003	1.3903	1.
	1	1.1	0.8950	0.9436	1.0053	1.0036	0.9435	1.0004	0.6349	1.2524	3.8737	1.
2.	0.1	1.1	0.8950	0.9436	1.0053	1.0036	0.9435	1.0004	0.6349	1.2524	3.8737	1.
	0.25	1.25	0.9436	0.9770	1.0007	0.9770	0.9963	1.0007	0.7038	1.0074	2.2273	1.
	0.5	1.5	0.9838	1.0033	1.0033	1.0033	0.9963	1.0008	0.8115	1.0003	1.3903	1.
	1	2.	0.9082	1.0081	1.0048	1.0048	0.9711	1.0026	0.6797	1.0311	4.3907	1.
3.	0.1	1.65	0.9082	0.9543	1.0104	1.0048	0.9711	1.0026	0.6797	1.0311	4.3907	1.
	0.25	1.875	0.9543	0.9884	1.0048	1.0048	0.9936	1.0041	0.7444	1.0040	2.1024	1.
	0.5	2.25	0.9884	1.0048	1.0048	1.0048	0.9936	1.0041	0.7444	1.0040	2.1024	1.
	1	3.	0.9082	1.0081	1.0048	1.0048	0.9711	1.0026	0.6797	1.0311	4.3907	1.

variance estimates, though, may be large. Also there seems to be very little sensitivity to the strength of the underlying correlation model. (i.e., to the change in u^*).

Table 11 gives some clues as to how different the weighting coefficients, appearing in the estimator $Z^*(\underline{X}_0)$ in formula (1.1), are. We treated $\underline{X}_0 = (0.5, 0.)$ separately from $\underline{X}_0 = (0., 0.5)$ as under anisotropic correlation models parallel to the main axis these two points are not symmetry points any more, and have a different set of weighting coefficients.

Table 11: Weight Coefficients

$$\alpha = \frac{1}{4} ; \quad u^* = 1 ; \min \rho_u^* = 0.3679 ; \max \rho_u^* = 0.7788$$

$$\rho_u(1) = 0.5353 ; \quad u = 0.625$$

Interpolation point	Observation	Anisotropic weight coefficients	Isotropic weight coefficients
(0.5,0.)	Z(0,0)	0.5347	0.5286
	Z(0,1)	-0.0347	-0.0286
	Z(1,1)	-0.0347	-0.0286
	Z(1,0)	0.5347	0.5286
(0.,0.5)	Z ₁	0.5140	0.5286
	Z ₂	0.5140	0.5286
	Z ₃	-0.0140	-0.0286
	Z ₄	-0.0140	-0.0286
(0.1,0.1)	Z ₁	0.8605	0.8644
	Z ₂	0.0728	0.0730
	Z ₃	-0.0080	-0.0105
	Z ₄	0.0682	0.0730

6. Robustness of a Linear Estimate to Omission of More Distant Data.

One of the problems of linear interpolation is, that for each estimate $Z^*(\underline{X}_0)$ at a location $\underline{X}_0$, we have to invert a matrix of order $k \times k$, where k is the number of observations that "take part" in the estimation procedure. $(Z^*(\underline{X}_0) = \sum_{i=1}^k C_i^* Z(\underline{X}_i)$ where $Z(\underline{X}_i)$ are the observations and C_i^* optimal constrained ($\sum C_i^* = 1$) coefficients.) We have a strong motivation to use a small value of k . On the other hand we will lose accuracy if k is reduced. The question to answer at this point is: "How robust is the interpolation procedure to deletion of the more distant observations? The underlying correlation function (strong or weak correlation) and the data configuration will determine the answer to the above questions.

If we have in mind air pollution stations, each one far away from the adjacent station, the question arises as to which of the stations should we consider, while estimating pollution concentration between the stations, and which of the measurements may be omitted, or, putting it in a different way - do we need to add stations to the ones already operating, and if so - where.

In the case of a regular square grid, the issue of reducing computation by decreasing the number of observations that take part in the estimation procedure, is not as crucial and severe as it is in the irregular sampling case. Once we decide on the number of "active" observations, k , in the regular square grid, we have to calculate the correlation matrix and its inverse only once, for

locations X_0 , where we want to estimate the process. (Only the border locations would have to be calculated with a different "design matrix" of the observations.) But still, even though it is less crucial than in irregular sampling spacing, we would like to know how sensitive the estimation variance is to deletion of data points from the procedure of estimation.

a. Data Points on a Square Grid.

For illustration we have chosen three different interpolation points as shown in Figure 12 with a square data grid. We are interested in composing the variance σ_4^2 of estimates based on the four surrounding data values to the variance σ_{16}^2 of estimates based on sixteen surrounding data values. Once again we use the correlation function $\rho(h) = e^{-uh^2}$ where h is distance and $u > 0$ is a parameter, and variances are calculated according to formula (1.3).

Table 12 shows σ_4^2 , σ_{16}^2 , and the ratio $\sqrt{\sigma_4^2/\sigma_{16}^2}$ as a function of the correlation parameter u and position of the interpolation point.

The results as summarized in Table 12, show that the crucial range, where it would be recommended to use 16 observation points instead of just estimating with four first neighbours, is the range where we expect the correlation $\rho_u(1)$ to be very strong, (i.e., the parameter u to get very low values), using the correlation function e^{-uh^2} , where h is the distance between any two points. When we

Figure 12(a): Coordinate System and Geometrical Setup of Four and Sixteen Observations on a Square Grid.

Figure 12(b): Interpolation Locations

TABLE 12: 1. σ_4^2 2. σ_{16}^2 3. $\sqrt{\frac{\sigma_4^2}{\sigma_{16}^2}}$

$\rho_u(1)$	u	(0.,0.)	(0.0,-0.5)	(-0.4,-0.4)
0.9048	0.1	4.6425·10 ⁻³ 5.1530·10 ⁻⁵ 9.4918*	1.7727·10 ⁻³ 2.3305·10 ⁻⁵ 8.7215*	5.8675·10 ⁻³ 5.6299·10 ⁻⁶ 10.2088*
0.6065	0.5	0.0876 0.0208 2.0538	0.0367 0.0102 1.8992	0.0103 0.0021 2.2059
0.4723	0.75	0.1674 0.0746 1.4977	0.0750 0.0377 1.4114	0.0193 0.0075 1.6037
0.3679	1	0.2547 0.1617 1.2550	0.1222 0.0841 1.2054	0.0291 0.0163 1.3357
0.1353	2	0.5865 0.5706 1.0139	0.3534 0.3444 1.0129	0.0704 0.0648 1.0421
0.0498	3	0.8292 0.8234 1.0035	0.5777 0.5754 1.0019	0.1102 0.1093 1.0042

* Determinant of var-covariance matrix is effectively zero.

assume a very strong correlation between observation points, even as we go far out to large distances, it seems logical to use the far away observations, as they add information, and still "contain" a fair "amount" of correlation with the point to be estimated.

At very low u values - $\rho_u(1) \uparrow 1$ - the determinant of the 16 point correlations is practically zero (the matrix is singular, by having all entries close to 1). The value that we get is just a "strange" value because all variances at this region are very small, and the ratios blow up.

The limit variance for the four points and 16 points in case of $u \uparrow$ is calculated easily by knowing that at $u \rightarrow \infty$, all coefficients tend to get the same values (same weight for all four points or all 16 points). $C_i = C = 1/n$ (n = number of observations). From (1.3):

$$\begin{aligned} \lim_{u \rightarrow \infty} \sigma_n^2 &= 1 + nC^2 = 1.25 & \text{for } n = 4 \\ \lim_{u \rightarrow \infty} \sigma_n^2 &= &= 1.0625 & \text{for } n = 16 \end{aligned}$$

For the symmetric case (0.,0.) the limit ratio

$$\lim_{u \rightarrow \infty} \sqrt{\frac{\sigma_4^2}{\sigma_{16}^2}} = 1.084435$$

is reached already for $u = 20$. (We immediately start with C_i for $i = 1, \dots, 4$ all of them equal to $1/4$ and the C_i 's for $i = 1, \dots, 16$ almost equal. (Two groups of four equivalent

coefficients and one group of eight equal coefficients.)

For other geometrical configuration this limit value is achieved for higher u values, although effectively the observations act as independent observations (resulting in an identity variance-covariance matrix) much earlier, before the limit ratio is achieved, and there is no sense in calculating for higher values of u than those which already give independent observations. Assuming high correlation (u values small) we would use the nearest and next nearest neighbors. From $\rho(1) = .1353$ and down, ($u \geq 2$) it is perfectly satisfying to estimate the value $Z(\underline{X}_0)$ at $\underline{X}_0$ out of the nearest neighbors, deleting the next nearest neighbors.

For the different estimation location, (letting the point $\underline{X}_0$ vary inside the square), we get more or less the same results, having a substantial range of correlations, ($\rho(1) \leq .1353$) where four points in the estimator would be as good as using all 16 points in estimator, for almost all practical purposes. Hence, Figure 13 is a representation of only one configuration $\underline{X}_0 = (0., 0.)$ (the center) as all other configurations result in similar results.

b. Estimation Using Only the Nearest Neighbor in a Triangular Data Configuration.

We shall look at three observations Z_1, Z_2, Z_3 , located on the vertices of an equilateral triangle and an estimation point $\underline{X}_0$ inside the triangle. The estimate of $Z(\underline{X}_0)$ using only the closest observation

Figure 14: Coordinate system, observations and interpolation locations

to $\underline{X}_0$ is compared to the case where we use all three observations in estimation procedure. See Figure 14, for arrangement of observations, estimation points and coordinate system. We shall choose points $\underline{X}_0$ along the line $\overline{Z_1 C}$, where C is the center of the triangle in Figure 14. In this case Z_1 will be the nearest observation to location $\underline{X}_0$.

(1) Taking only the nearest neighbor, Z_1 , (in Figure 14) the best linear unbiased estimator for $Z(\underline{X}_0)$ is $Z^*(\underline{X}_0) = Z_1$ and the estimation variance $V_1 = 2(1 - \rho_{(0,1)})$ (by formula (1.3)).

(2) Taking into account all three observations, we get

$$Z^*(\underline{X}_0) = C_1 Z_1 + C_2 Z_2 + C_3 Z_3 \quad \text{where}$$

$$(6.1) \quad c_i = \frac{1}{1-3\rho^2+2\rho^3} [\rho_{(i,o)}(1+\rho-2\rho^2) + \rho(\rho-1) \sum_{j=1}^3 \rho_{(j,o)}]$$

by formula 1.5 and the estimation variance is - by formula 1.3:

$$(6.2) \quad v_3 = 1 + \sum_{j=1}^3 \sum_{i=1}^3 c_i c_j \rho_{(i,j)} - 2 \sum_{i=1}^3 c_i \rho_{(i,o)}$$

We are interested in the ratio $\sqrt{v_1/v_3}$. For purposes of illustration we again use the correlation model $\rho(h) = e^{-uh^2}$ where $u > 0$ is a parameter.

Five different interpolation points $\underline{X}_0$ are examined, viz. (0.,0.8), (0.,0.7), (0.,0.6), (0.,0.4) and (0., $\sqrt{3}/6$) as shown in Figure 14. The results are summed up in Table 13 and their graphical representation is given in Figure 15. (For (0.,0.8) and (0., $\sqrt{3}/6$).)

(i) Intuitively, we would expect that whenever the estimated point $\underline{X}_0$ is close to the nearest neighbor observation compared to its distances to the other two next nearest neighbors, then the precision of the estimate would not suffer much, by omitting data points. Such a "blanket statement" would be nice to have, but as it appears from the results, is not exactly the case.

Under assumptions of strong correlation it is substantially more accurate to estimate in a "conservative manner" by engaging three surrounding observations rather than only one. For example: interpolate at $\underline{X}_0 = (0.,0.8)$, which is close to the observation Z_1 at (0. $\sqrt{3}/2$) . Assume that the underlying correlation is strong. Let

TABLE 13: $V1$ $R = (V1/V3)^{1/2}$
 $V3$ $\rho(1,0)$
 $\rho(3,0) = \rho(2,0)$

$\rho_u(1)=e^{-u}$		(0.,0.8)		(0.,0.7)		(0.,0.6)		(0.,0.4)		0., $\sqrt{3}/6$	
0.9048	0.1	$V_1=8.7 \cdot 10^{-4}$	$R=2.8471$	$5.50 \cdot 10^{-3}$	3.1543	0.0141	3.5362	0.0429	4.6401	0.0656	5.5534
		$V_3=1.07 \cdot 10^{-4}$	$\rho(1,0)=0.9996$	$5.53 \cdot 10^{-3}$	0.9972	0.0011	0.9929	0.0019	0.9785	0.0021	0.9672
	0.25	$6.11 \cdot 10^{-4}$	$\rho(3,0)=0.9148$		0.9287		0.9408		0.9598		0.9672
0.7788	0.25	$2.17 \cdot 10^{-3}$	1.8878	0.0137	2.0774	0.0351	2.3152	0.1057	3.0091	0.1599	3.5846
			0.9989		0.9931		0.9825		0.9471		0.9200
0.6065	0.5	$6.11 \cdot 10^{-4}$	0.8005	0.0032	0.8311	0.0065	0.8585	0.0117	0.9026	0.0124	0.9200
0.4723	0.5	$4.3546 \cdot 10^{-3}$	1.4397	0.0274	1.5677	0.0695	1.7311	0.2058	2.2161	0.3070	2.6201
		$2.1009 \cdot 10^{-3}$	0.9978		0.9863		0.9652		0.8971		0.8465
	0.75		0.6408	0.0111	0.6907	0.0232	0.7371	0.0419	0.8146	0.0447	0.8465
0.3679	1	$6.5 \cdot 10^{-3}$	1.2627	0.0409	1.3618	0.1034	1.4911	0.3006	1.8820	0.4424	2.2092
			0.9967		0.9795		0.9483		0.8497		0.7788
	1.5	$4.09 \cdot 10^{-3}$	0.5130	0.0221	0.5741	0.0465	0.6329	0.0849	0.7353	0.0906	0.7788
0.2231	1.5	$8.7 \cdot 10^{-3}$	1.1698	0.0544	1.2510	0.1366	1.3592	0.3904	1.6932	0.5669	1.9738
		$6.35 \cdot 10^{-3}$	0.9956		0.9728		0.9317		0.8048		0.7165
	2		0.4106	0.0347	0.4771	0.0740	0.5433	0.1362	0.6636	0.1455	0.7165
0.1353	2	0.0130	1.0800	0.0810	1.1386	0.2014	1.2212	0.5560	1.4863	0.7869	1.7103
			0.9935		0.9595		0.8993		0.7220		0.6065
	3	0.0112	0.2631	0.0625	0.3296	0.1350	0.4005	0.2517	0.5406	0.2690	0.6065
0.0498	3	0.0174	1.0416	0.1073	1.0866	0.2640	1.1538	0.7046	1.3774	0.9732	1.5662
			0.9913		0.9464		0.8680		0.6477		0.5134
	∞	0.0160	0.1686	0.0908	0.2276	0.1983	0.2952	0.3714	0.4404	0.3967	0.5134
0	∞	0.0260	1.0145	0.1587	1.0450	0.3826	1.0958	0.9575	1.2712	1.2642	1.4157
			0.9870		0.9206		0.8087		0.5212		0.3678
	∞	0.0252	0.0692	0.1453	0.1086	0.3186	0.1604	0.5925	0.2923	0.6308	0.3678
0	∞	2	1.2247	2	1.2247	2	1.2247	2	1.2247	2	1.2247
			0		0		0		0		0
	∞	1.3333	0	1.3333	0	1.3333	0	1.3333	0	1.3333	0

Figure 15: Performance Ratio R vs. the parameter u .

$\rho(1) = 0.9048$, ($u = 0.1$) . From Table 13, the ratio obtained by taking one instead of three observations is $\sqrt{V_1/V_3} = 2.8471$! (over 200%). If we weaken the correlation model - for example $\rho(1) = 0.1353$ ($u = 2$), there is no effective difference in precision of Z^* by taking only one observation instead of three. ($\sim 4\%$)

The above mentioned consequence - that for strong underlying correlation models we should use more data points in estimation than for weak models - accounts for the fact that when correlation is strong even the further away observation points carry information relevant to the interpolation point.

(ii) Let the location of the interpolated point X_0 be close to the center of the triangle. In this case we should use all three observations. Large increases in variance are induced by using one - instead of three - observations in the estimate, for both strong and weak correlation models. The limit value for the ratio as $u \rightarrow \infty$ (independent observations) and as we calculate the error of estimation with one observation instead of three - is as big as 22%. This limit is reached at ease for estimation at the center of the triangle, and much slower for X_0 close to an observation point. Therefore, we should hesitate to omit data under an assumption of uncorrelated observations if we estimate at the center. In Figure 15, we see the convergence to the limit of the ratio for interpolation at $(0, \sqrt{3}/6)$ (center of triangle) and at $(0, 0.8)$, (which is close to the data point $(0, \sqrt{3}/2)$).

(iii) It is always true that omission of further observation points, when interpolating close to an observation, is more justified (with respect to the precision of the estimate) than omission of observations in the case of interpolating a point near the center of the data triangle.

References

- [1] Behar, J. V., (1972), "Applications of computer simulation techniques to problems of air pollution". Proceedings of the Sixth Berkeley Symposium, Vol. VI, pp. 29-69.
- [2] Blais, R. A., and Carlier, P. A., (1968), "Applications of geostatistics in ore evaluation". Canadian Inst. of Mining and Metallurgy. Ore reserve estimation and grade control. Special Vol. 9.
- [3] Burnaby, T. P., (1970), "On a method for character weighting a similarity coefficient employing the concept of information". Math. Geology, Vol. 2, #1.
- [4] Huijbregts, C., and Mathéron, G., (1970), "Universal Kriging" (An optimal method of estimating and contouring in trend surface analysis). Ecole de mines de Paris. Dept. Geostatistique, Centre de Morphologie Mathématique Fontainebleau.
- [5] Krumbein, W. C., (1959), "Trend surface analysis of contour type maps with irregular control point spacing". Journal of Geophysical Res., Vol. 64, #7, pp. 823-834.
- [6] Mandelbaum, H., (1963), "Statistical and geological implications of trend mapping with nonorthogonal polynomials". Journal of Geophysical Res., Vol. 68, #2-4.
- [7] Mathéron, G., (1965), "Les variables régionalisées et leur estimation. Une application de la théorie des fonctions aléatoires aux sciences de la nature". Masson et Cie. Editeurs, Paris VI^e.
- [8] Mathéron, G., (1969), "Kriging or polynomial interpolation procedures?" A contribution to polemics in mathematical geology. The Canadian Mining and Metallurgical Bulletin, Vol. 60, #665.
- [9] Mathéron, G., (1969), "Le Krigeage Universel". Les cahiers du centre de morphologie mathématique de Fontainebleau. Fascicule 1. Edité par l'école nationale supérieure des mines de Paris.
- [10] Mathéron, G., (1972), an article in "Traite d'information géologique". Ouvrage Collectif., direction: P. Laffitte, Masson Cie., Paris VI^e.

- [11] Mathéron, G., (1973), "The intrinsic random functions and their applications". Adv. in Appl. Prob. 5, pp. 439-468.
- [12] Norcliffe, G. B., (1969), "On the use and limitations of trend surface models". Canadian Geographer, Vol. 13, #4, pp. 338.
- [13] Olea, R., (1972), "Applications of regionalized variable theory to automatic contouring". The University of Kansas center for research in cooperation with the Kansas geological survey, Lawrence, Kansas.
- [14] Rao, S. V. L. N., and Rao, S. M., (1970), "A study of residual maps in the interpolation of geochemical analysis". Math. Geology, Vol. 2, #1.
- [15] Robinson, J. E., and Merriam, D. F., (1971), "Z-Trend maps for quick recognition of geological patterns". Math. Geology, Vol. 3, #2.
- [16] Switzer, P. C., Mohr, M., and Heitman, R. E., (1964), "Statistical analysis of ocean terrain and contour plotting procedures". Project Trident, Tech. Rep. #1440464, Arthur D. Little Inc., Cambridge, Mass.
- [17] Tinkler, K. J., (1971), "Statistical analysis of tectonic patterns in a real volcanism". Math. Geology, Vol. 3, #4.
- [18] Whitten, T., (1970), "Orthogonal polynomial trend surface for irregularly spaced data". Math. Geology, Vol. 2, #2. Erratum in Vol. 3, #3, 1971.
- [19] Whitten, T., (1974), "Data for spatial variability studies". Math. Geology, Vol. 6, #2.

Bibliography

- Delfiner, P., (1975), "Linear Estimation of Nonstationary Spatial Phenomena". Proceedings of NATO A.S.I. Rome.
- Laffitte, P., (direction) (1972), "Traité d'information géologique". Ouvrage collectif. Masson et Cie. Paris VI^e.
- Matérn, Bertil, (1960), "Spatial Variation", Meddelanden från Statens Skogsforskningsinstitut. Band 49, Nr. 5.
- Mathéron, G., (1965), "Les Variables Régionalisées et leur Estimation". Masson et Cie Editeurs. Paris VI^e.
- Merriam, D. F., (editor) (1970), Geostatistics - a Colloquium. Plenum Press, New York.
- Watson, G. S., (1971), "Trend Surface Analysis", Journal of International Association for Mathematical Geology. Vol. 3, No. 3, p. 215.