ROBUSTNESS OF LINEAR ESTIMATORS TO
MISSPECIFICATION OF TREND

Ruth Sheshinski

TECHNICAL REPORT NO. 4
August 16, 1976

PREPARED UNDER THE AUSPICES OF
SIAM INSTITUTE FOR MATHEMATICS AND SOCIETY

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

MASTER

There is no objection from the patent
point of view to the publication or
dissemination of the document(s)
listed in this letter.

BROOKHAVEN PATENT GROUP

12/8 1976 By *chc*

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 OF LINEAR ESTIMATORS TO MISSPECIFICATION OF TREND

by

RUTH SHESHINSKI

Hebrew University, Jerusalem
and Stanford University

TECHNICAL REPORT NO. 4

AUGUST 16, 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

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

leg

ACKNOWLEDGEMENTS

This work is a second paper to be submitted as part of the Ph.D. requirements 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

The following work is a continuation of an earlier report, Sheshinski (1976). In this paper we continue our examination of the problem of interpolation of air pollution concentrations over an air basin, using a model of a spatially autocorrelated random process.

In Chapter I we compare estimators based on a correctly and incorrectly specified constant mean with estimators based on an unspecified constant mean. The comparison is done by calculating the corresponding values for the MSE. The ratios of the mean squared errors decreases with decreasing correlation between observations, and are, in general, very close to one.

In Chapter II we compare estimation done with a specified linear trend function to estimation with an unspecified linear trend function. For interpolation, in general, the MSE's tend to be close except for high underlying correlation models. For extrapolation - the MSE values can be very different.

The third chapter deals with a biased partially constrained estimator versus a fully constrained universally unbiased estimator in the presence of a linear trend. In general, the biased and unbiased estimators have similar performance, except for very steep linear slopes of the trend. The comparisons of this paper are made in the context of rather special and perhaps unrealistic model assumptions, but they should provide useful insights into the importance or unimportance of the need for correct parameter specifications.

I. Interpolation Comparisons with Specified Mean Levels and Unspecified Mean Levels.

We have observed pollutant concentrations $Z(\underline{X}_i)$, $i=1, 2, \dots, n$ at n monitoring stations $\underline{X}_1, \underline{X}_2, \dots, \underline{X}_n$ in a plane region. Define a linear estimator at an interpolation point $\underline{X}_0$ as

$$(1.1) \quad Z^*(\underline{X}_0) = C_0 + \sum_{i=1}^n C_i Z(\underline{X}_i) .$$

Suppose we take $Z(\underline{X})$ to be a realization of a stationary process on R^2 with mean function $m(\underline{X}) = m$ and an isotropic autocorrelation function $\rho(h)$, where h is distance. Then a simple calculation gives the coefficients $C_0, C_1, \dots, C_n$ in (1.1) which (i) determine an unbiased estimator, i.e., $E Z^*(\underline{X}_0) = m$, and (ii) minimizes $E [Z^*(\underline{X}_0) - Z(\underline{X}_0)]^2$. The unbiasedness of $Z^*(\cdot)$ requires

$$(1.2) \quad C_0 = m(1 - \sum_{i=1}^n C_i)$$

for a specified constant mean m . The resulting mean-squared-error

will then not depend on m and, therefore, neither will the optimal coefficients $C_1, C_2, \dots, C_n$. We write:

$$\begin{aligned}
 \sigma_{\text{opt}}^2 &= E[Z^*(\underline{X}_0) - Z(\underline{X}_0)]^2 = E\left[m\left(1 - \sum_{i=1}^n C_i\right) + \sum_{i=1}^n C_i Z(\underline{X}_i) - Z(\underline{X}_0)\right]^2 \\
 (1.3) \qquad &= 1 + \sum_{i=1}^n \sum_{j=1}^n C_i C_j \rho_{ij} - 2 \sum_{i=1}^n C_i \rho_{i0}
 \end{aligned}$$

where $\rho_{ij} = \rho(|\underline{X}_i - \underline{X}_j|)$ and, to simplify notation, the stationary variance $E[Z-m]^2$ is taken to be equal to one.

Differentiating (1.3) with respect to $C_1, C_2, \dots, C_n$ we get the "usual" normal equations for the optimal weight coefficients, viz.

$$(1.4) \qquad C = R^{-1} \rho$$

where in matrix notation

$$\begin{aligned}
 C &= (C_1, C_2, \dots, C_n)^T, \\
 \rho &= (\rho_{10}, \rho_{20}, \dots, \rho_{n0})^T,
 \end{aligned}$$

matrix R (of order $n \times n$) is $((\rho_{ij}))$, and T denotes transpose.

The interpolation $Z^*(\underline{X}_0)$ will be neither optimal nor will $Z^*(\underline{X}_0)$ be unbiased if the mean m is incorrectly specified. Let m' be the incorrectly specified mean. Then the mean-squared-estimation-error is:

$$\begin{aligned}
(1.5) \quad \sigma_m^2 &= E[Z^*(X_0) - Z(X_0)]^2 = \\
&= E \left[\left(m(1 - \sum_{i=1}^n C_i) + \sum_{i=1}^n C_i Z(X_i) - Z(X_0) \right) + (m-m') \left(\sum_{i=1}^n C_i - 1 \right) \right]^2 \\
&= \sigma_{opt}^2 + (m-m')^2 \left(\sum_{i=1}^n C_i - 1 \right)^2.
\end{aligned}$$

On the other hand, if we want an estimate which is unbiased for all m and we do not wish to specify m , then we require $\sum_{i=1}^n C_i = 1$ which makes $C_0 \equiv 0$. Denote the coefficients $C_1, C_2, \dots, C_n$ constrained to add to one by $\{C_i^*\}$. We now seek the optimal constrained estimator

$$(1.6) \quad Z^*(X_0) = \sum_{i=1}^n C_i^* Z(X_i) \quad (\text{as } C_0 \equiv 0)$$

which minimizes $E[Z^*(X_0) - Z(X_0)]^2$. We have

$$(1.7) \quad \sigma_c^2 = E[Z^*(X_0) - Z(X_0)]^2 = 1 + \sum_{i=1}^n \sum_{j=1}^n C_i^* C_j^* \rho_{ij} - 2 \sum_{i=1}^n C_i^* \rho_{i0}.$$

This estimation variance has to be minimized. (1.7) may be written in terms of $C_1^*, \dots, C_{n-1}^*$. (Since $C_n^* = 1 - \sum_{i=1}^{n-1} C_i^*$) viz.

$$\begin{aligned}
(1.8) \quad \sigma_c^2 &= 1 + \sum_{i=1}^{n-1} \sum_{j=1}^{n-1} C_i^* C_j^* \rho_{ij} - 2 \sum_{i=1}^{n-1} C_i^* \rho_{i0} + \\
&\quad + \left(1 - \sum_{i=1}^{n-1} C_i^* \right)^2 + 2 \left(1 - \sum_{i=1}^{n-1} C_i^* \right) \sum_{i=1}^{n-1} C_i^* \rho_{in} - 2 \left(1 - \sum_{i=1}^{n-1} C_i^* \right) \rho_{n0}.
\end{aligned}$$

Differentiating (1.8) w.r.t. $C_1^*, \dots, C_{n-1}^*$ we obtain $n-1$ normal equations for $C_1^*, \dots, C_{n-1}^*$:

$$(1.9) \quad \sum_{j=1}^{n-1} C_j^* [\rho_{kj} + 1 - \rho_{jn} - \rho_{kn}] = \rho_{ko} + 1 - \rho_{kn} - \rho_{no}$$

for all $k = 1, 2, \dots, (n-1)$.

Rewriting the $(n-1)$ equations in matrix notation, denoting

$$\begin{aligned} C^* &= (C_1^*, C_2^*, \dots, C_{n-1}^*)^T \\ \rho_{ij}^* &= 1 + \rho_{ij} - \rho_{in} - \rho_{jn} \\ \rho^* &= (\rho_{10}^*, \rho_{20}^*, \dots, \rho_{(n-1)0}^*)^T \end{aligned}$$

and R of order $(n-1) \times (n-1)$ is $((\rho_{ij}^*))$, results in

$$(1.10) \quad \begin{cases} C = (R^*)^{-1} \rho^* \\ \text{and } C_n^* = 1 - \sum_{i=1}^{n-1} C_i^* . \end{cases}$$

These results have been given by Mathéron (1965).

The relative measure of performance for a specified m versus an unspecified m is the ratio of the squared error using the constrained universally unbiased estimator (1.8) and the optimal squared error (1.3). We define

$$(1.11) \quad r = (\sigma_C^2 / \sigma_{opt}^2)^{1/2} .$$

For a simple illustration, let us take four observations located at the corners of a unit square, interpolating and extrapolating inside and outside the square. See Figure 1 for the geometrical representation of the observations and estimation locations.

Figure 1: Geometrical setup for four observations, and the coordinate system.

As an illustration, we shall choose the following isotropic autocorrelation model:

$$(1.12) \quad \rho(h) = e^{-uh^2}; \quad u > 0$$

Ratio (1.11) is easily calculated for uncorrelated observations.

As $u \uparrow \infty$, the coefficients $\{C_i^*\}$ for the constrained set of equations

equal $1/4$, $\forall i$, at any interpolation point $\underline{X}_0$. The unconstrained estimator has diminishing coefficients, i.e., $C_i \rightarrow 0$, $\forall i$, as correlations become weaker. Hence substituting into (1.11) gives:

$$(1.13) \quad \lim_{u \uparrow \infty} (\sigma_C^2 / \sigma_{\text{opt}}^2)^{1/2} = 1.118.$$

For data with moderately correlated observations, this relative performance ratio can be very close to unity. See Table 1 for examples of the above mentioned ratio (1.11). Tabulated as well in Table 1 are the corresponding estimation errors σ_{opt}^2 and σ_C^2 , and the unconstrained and constrained coefficients $\{C_i\}$ and $\{C_i^*\}$, for correlations in the range 0.05 to 0.90. With strong underlying correlation models, the ratio (1.11) is much closer to one at the extrapolation point $\underline{X}_0 = (1.5, .5)$ than at the interpolation point $\underline{X}_0 = (.5, .5)$.

Let us look at the behavior of the weighting coefficients and the performance ratios exhibited in Table 1. For interpolation at $\underline{X}_0 = (0.5, 0.5)$ the constrained coefficients are always equal to $1/4$, under all correlation models. The unconstrained coefficients at this location are also all equal, but in general are not equal to $1/4$. For very high correlation models the coefficients are slightly larger than $1/4$. They increase as correlation decreases to a maximum near $1/3$ and then decrease with decreasing correlation reaching $1/5$ for $\rho(1) \approx 0.05$. The performance ratio is similarly not a monotonic function of the strength of the correlation model, but the worst relative performance

Table 1: Interpolation at Center of Unit Square $\underline{X}_0 = (0.5, 0.5)$.
Extrapolation at $\underline{X}_0 = (1.5, 0.5)$

1. Variance of constrained estimate $\equiv \sigma_C^2$
2. Variance of unconstrained estimate $\equiv \sigma_{opt}^2$
3. $r = (\sigma_C^2 / \sigma_{opt}^2)^{1/2}$
4. Coefficients resulting from unconstrained $\{C_i\}$
and constrained $\{C_i^*\}$ minimization problem
5. $\rho(h) = e^{-uh^2}$; $u > 0$

$\rho_u(1)$	u	Interpolation			Extrapolation					
		σ_C^2 σ_{opt}^2	r	Unconst. Coef. C_i	σ_C^2 σ_{opt}^2	r	Const. Coef.		Unconst. Coef.	
							$C_1^* = C_2^*$	$C_3^* = C_4^*$	$C_1 = C_2$	$C_3 = C_4$
0.9048	.1	0.0046 0.0025	1.3639	0.2622	0.0130 0.0111	1.0844	-.2256	0.7256	-.2372	0.7140
.7788	.25	0.0260 0.0155	1.2976	0.2789	0.0668 0.0583	1.0707	-.1912	0.6912	-.2172	0.6653
.6065	.50	0.0876 0.0600	1.2087	0.3017	0.1965 0.1769	1.0541	-.1394	0.6394	-.1830	0.5958
.4724	.75	0.1674 0.1284	1.1417	0.3170	0.3343 0.3070	1.0435	-.0936	0.5936	-.1498	0.5375
.3679	1	0.2547 0.2135	1.0921	0.3241	0.4611 0.4285	1.0373	-.0533	0.5533	-.1192	0.4873
.2725	1.3	0.3607 0.3268	1.0506	0.3223	0.5917 0.5535	1.0340	-.0110	0.5110	-.0878	0.4342
.1353	2	0.5865 0.5800	1.0055	0.2854	0.8148 0.7583	1.0366	0.0661	0.4339	-.0386	0.3292
.0498	3	0.8292 0.8193	1.0061	0.2025	1.0022 0.9049	1.0523	0.1384	0.3616	-.0101	0.2130
0	∞	1.25 1	1.118	0	1.25 1	1.118	0.25	0.25	0	0

from the table ($r = 1.36$) occurs in the presence of the strongest correlation. It appears that at the very point where the set of optimal unconstrained coefficients fall below $1/4$, the performance ratio (1.11) starts to increase monotonically with weaker correlation assumptions. The unconstrained coefficients are zero in the uncorrelated case, which means that the linear estimator is $Z^* = m$ and the estimation variance is 1. For extrapolation the unconstrained estimator has large positive weight for the near data points and negative weight for the distant data points, with $\sum_{i=1}^n C_i < 1$. Ratio r (1.11) appears to decrease with decreasing correlation, as long as the sum of absolute values of unconstrained coefficients > 1 and increases till reaching the limit in (1.13) for $\sum_{i=1}^n |C_i| < 1$. For the constrained linear estimator ($\sum_{i=1}^n C_i^* = 1$) the weights on the distant data points become positive for weak correlation models. Once again, the relative performance ratio (1.11) is not a monotonic function of the correlation parameter with the worst performance for very weak and very strong correlation models.

However, neither in interpolation nor in extrapolation does the use of the universally unbiased constrained linear estimator, as opposed to the unconstrained optimal linear estimator have a severe influence on the precision.

An interesting question is whether to use the unconstrained optimal linear estimator (1.1), when the mean m may be misspecified. The estimator with a misspecified value of m (say m') will no longer be unbiased and its mean squared error may be larger than that for the

constrained estimator (1.6) which is universally unbiased. In (1.5) we calculated the experienced mean squared error for the unconstrained estimator (1.1) with a misspecified mean m' . It would be better to use the unconstrained estimator as long as this MSE - denoted σ_m^2 , is still less than the corresponding value of σ_C^2 (1.7), for the constrained estimator (1.6) under the same underlying correlation assumptions.

Using (1.5) the estimation variance σ_C^2 is less than σ_m^2 , whenever

$$(1.14) \quad (\sigma_C^2 - \sigma_{opt}^2)^{1/2} / \left| \left(\sum_{i=1}^n C_i - 1 \right) \right| < |m - m'|.$$

In Table 2 are shown values of the LHS of (1.14) for two different interpolation locations $\underline{X}_0 = (0.5, 0.5)$, $\underline{X}_0 = (0.1, 0.1)$ and an extrapolation location $\underline{X}_0 = (1.5, 0.5)$.

(i) It is a striking result that the actual interpolation point $\underline{X}_0$ has no influence on the possible difference between true and misspecified mean values. No matter where estimation takes place - the allowed discrepancy $|m - m'|$ is almost identical.

(ii) For strong underlying correlation models, if we can specify the value of m to within about one standard deviation, then the biased unconstrained estimator is preferable to the universally unbiased constrained estimator. However, for uncorrelated models, a discrepancy between m and m' of ± 0.5 standard deviations would make the constrained unbiased estimator preferable.

Table 2 C_i^* = constrained coefficient; $\sum C_i^* = 1$
 C_i = unconstrained coefficient
 σ_C^2 = constrained estimation variance
 σ_{opt}^2 = unconstrained estimation variance
 $S = \sum_{i=1}^4 C_i$
 $|m-m'| = |(\sigma_C^2 - \sigma_{opt}^2)^{1/2} / (S - 1)|$

$\rho_u(1)$ u	$X_0 = (0.5, 0.5)$			$X_0 = (1.5, 0.5)$			$X_0 = (0.1, 0.1)$		
	S	σ_C^2 σ_{opt}^2	$ m-m' $	S	σ_C^2 σ_{opt}^2	$ m-m' $	S	σ_C^2 σ_{opt}^2	$ m-m' $
0.9048 0.1	1.0488	0.0046 0.0025	0.9390	0.9536	0.0130 0.0111	0.9394	1.0172	0.0005867 0.0003169	0.9524
0.7788 .25	1.1156	0.0260 0.0155	0.8864	0.8962	0.0668 0.0583	0.8882	1.0402	0.003187 0.001910	0.8894
0.6065 .50	1.2068	0.0876 0.0600	0.8033	0.8256	0.1965 0.1769	0.8027	1.0702	0.01032 0.00714	0.8033
0.4724 .75	1.2680	0.1674 0.1284	0.7369	0.7754	0.3343 0.3070	0.7356	1.0901	0.01929 0.01490	0.7362
0.3679 1	1.2964	0.2547 0.2135	0.6848	0.7362	0.4611 0.4285	0.6844	1.1004	0.02913 0.02441	0.6839
0.2725 1.3	1.2892	0.3607 0.3268	0.6366	0.6926	0.5917 0.5535	0.6358	1.1022	0.04142 0.03719	0.6363
0.1353 2	1.1416	0.5865 0.5800	0.5694	0.5812	0.8148 0.7583	0.5676	1.0767	0.07043 0.06853	0.5677
0.0498 3	0.8100	0.8292529 0.8192934	0.5237	0.4058	1.0022 0.9049	0.5249	1.0166	0.11017 0.11009	0.5249
0 ∞	0	1.25 1	0.5	0	1.25 1	0.5	0	1.25 1	0.5

The form of the optimal linear estimator (1.1) for an arbitrary mean function $m(\underline{X})$, not necessarily constant over the air basin, still has $C_1, \dots, C_n$ given by (1.4) and

$$(1.15) \quad C_0 = m(\underline{X}_0) - \sum_{i=1}^n C_i m(\underline{X}_i) .$$

The squared error formula (1.3) is also unchanged. This may be seen by writing $Z(\underline{X}) = m(\underline{X}) + Y(\underline{X})$ where $Y(\underline{X})$ is a stationary random function with constant mean $\equiv 0$. Then $Z(\underline{X})$ and $Y(\underline{X})$ have the same autocorrelation function and interpolation of $Z(\underline{X})$ is the same as interpolation of $Z(\underline{X}) - m(\underline{X})$, when $m(\underline{X})$ is correctly specified.

However, if we misspecify the mean function $m(\underline{X})$ by $m'(\underline{X})$, we get, as already seen in the case of constant mean, a suboptimal and biased estimation. Similar calculations may be executed, as have been done for the constant mean function, to determine the discrepancy between different functional forms of $m(\underline{X})$ and $m'(\underline{X})$ at which unconstrained (presumably optimal) solution is less accurate compared to a constrained solution which is required to be unbiased over some class of mean functions $m(\underline{X})$.

Constrained estimators are described in Mathéron (1972, 1973) where the trend is given as an analytic expression

$$(1.16) \quad m(\underline{X}) = \sum_{i=0}^k a_i f_i(\underline{X})$$

where $\{f_i(\underline{X})\}$ are specified functions and $\{a_i\}$ are unspecified constants. For example, (1.16) can be used to represent a polynomial trend in the mean function. Then the conditions on the weights $\{C_i\}$ in (1.1), so that $Z^*(\underline{X}_0)$ is unbiased for any $\{a_i\}$ is:

$$(1.17) \quad E[Z^*(\underline{X}_0)] = \sum_{j=1}^n C_j E[Z(\underline{X}_j)] = \sum_{j=1}^n C_j \sum_{i=0}^k a_i f_i(\underline{X}_j) \\ = \sum_{i=0}^k \sum_{j=1}^n a_i C_j f_i(\underline{X}_j) .$$

As shown in the above reference, the necessary and sufficient condition for unbiasedness of $Z^*(\underline{X}_0)$ is:

$$(1.18) \quad \sum_j C_j f_i(\underline{X}_j) = f_i(\underline{X}_0) , \quad i = 0, 1, 2, \dots, k .$$

When assuming a misspecified trend other than $m(\underline{X}) = m$ (constant) we actually miscalculate the estimates for the autocorrelation functions. The term:

$$(1.19) \quad \sigma_{ij} = E[(Z(\underline{X}_i) - m(\underline{X}_i))(Z(\underline{X}_j) - m(\underline{X}_j))]$$

that enters in ρ_{ij} will be different for misspecified $m(\underline{X}_i)$.

These errors in estimating autocorrelation functions imply wrong coefficients in the estimator. For an arbitrary trend (nonconstant) the location of the estimation point will certainly be a factor in determining the severity of the discrepancy.

II. Interpolation Comparisons of Specified Linear Trends and Unspecified Linear Trends.

We now assume a linear trend, represented by the mean function

$$(2.1) \quad E[Z(\underline{X})] = m(\underline{X}) = \alpha + \beta U + \gamma V ,$$

where (U,V) are the coordinates of the point $\underline{X}$ in the plane, and α, β, γ are fixed constants. Let $Z^*(\underline{X}_0)$ be of the form (1.1) i.e., a linear combination of the observations used to estimate $Z(\underline{X}_0)$ at an interpolation point $\underline{X}_0$, $(Z^*(\underline{X}_0) = \sum_{i=1}^n C_i Z(\underline{X}_i))$.

If the trend constants α, β, γ are not specified and we want

$$(2.2) \quad E[Z^*(\underline{X}_0)] = m(\underline{X}_0)$$

for all α, β, γ , then the constraints on the weighting coefficients are

$$(2.3) \quad \begin{cases} \sum_{i=1}^n C_i^* = 1 \\ \sum_{i=1}^n C_i^* U_i = U_0 \\ \sum_{i=1}^n C_i^* V_i = V_0 \end{cases}$$

where (U_0, V_0) are the coordinates of $\underline{X}_0$. These constraints follow from (1.18). The estimation variance for the optimal constrained

estimator, which is unbiased for all α, β, γ , we denote by σ_C^2 . Of course it is larger than the variance σ_{opt}^2 for the optimal unconstrained linear estimator, where α, β, γ are correctly specified.

We shall compare σ_C^2 with σ_{opt}^2 in a geometrical setup of four observations located at the corners of a unit square. See Figure 1 for the coordinate system and estimation locations for interpolation and extrapolation. We selected four interpolation locations:

$\underline{X}_0 = (0.5, 0.5), (0.5, 0.), (0.75, 0.25), (1.0, 0.5)$ and three extrapolation locations: $\underline{X}_0 = (1.5, 0.5), (0.5, -0.5), (1.5, -0.5)$.

Values of σ_{opt}^2 may be calculated according to (1.3), in which the specified mean function plays no role. Values of σ_C^2 may be calculated by minimizing (1.7) subject to the constraints (2.3), using Lagrange multipliers. For illustration, autocorrelations were taken from the model $\rho(h) = e^{-uh^2}$ where h is distance and $u > 0$ is a parameter.

Let

$$(2.4) \quad r_1 = (\sigma_C^2 / \sigma_{opt}^2)^{1/2}.$$

See Table 3 - for the points of interpolation except $\underline{X}_0 = (0.5, 0.5)$ and Table 4 - for extrapolation locations, using underlying autocorrelation models ranging from $\rho(1) = 0.05$ to $\rho(1) = 0.9$.

Table 3: Interpolation points. $r = (\sigma_C^2 / \sigma_{opt}^2)^{1/2}$ constant mean

$r_1 = (\sigma_C^2 / \sigma_{opt}^2)^{1/2}$ linear trend with α, β, γ

$r_3 = (\sigma_b^2 / \sigma_C^2)^{1/2}$ $\sum_{i=1}^4 C_i^* = 1$ only and linear trend with α, β, γ

$\rho_u(1)$	u	$\underline{x}_0$	r	r_1	r_3			
					$\beta=0.2$	$\beta=0.5$	$\beta=1.$	$\beta=10$
0.9048	.1	(0.5, 0.)	1.1915	1.2002	0.9927	0.9927	0.9927	0.9927
0.7788	.25		1.1499	1.1672	0.9852	0.9852	0.9852	0.9852
0.6065	.5		1.0982	1.1210	0.9797	0.9797	0.9797	0.9797
0.4724	.75		1.0629	1.0846	0.9800	0.9800	0.9800	0.9800
0.3679	1		1.0390	1.0565	0.9834	0.9834	0.9834	0.9834
0.1353	2		1.0021	1.0038	0.9984	0.9984	0.9984	0.9984
0.0498	3		1.0024	1.0046	0.9978	0.9978	0.9978	0.9978
0	∞		1.1180	1.2247	0.9129	0.9129	0.9129	0.9129
0.9048	.1	(0.75, 0.25)	1.3626	1.3707	0.9947	0.9982	1.0103	2.0598
0.7788	.25		1.2952	1.3130	0.9871	0.9908	1.0036	2.0960
0.6065	.5		1.2058	1.2343	0.9777	0.9815	0.9952	2.1352
0.4724	.75		1.1397	1.1736	0.9719	0.9758	0.9898	2.1463
0.3679	1		1.0916	1.1273	0.9691	0.9730	0.9868	2.1296
0.1353	2		1.0084	1.0340	0.9758	0.9784	0.9879	1.8549
0.0498	3		1.0012	1.0131	0.9885	0.9897	0.9942	1.4647
0	∞		1.1180	1.1726	0.9544	1.0063	1.0247	2.4496
0.9048	.1	(1., 0.5)	1.1915	1.2002	0.9943	1.0027	1.0323	3.0023
0.7788	.25		1.1499	1.1672	0.9867	0.9946	1.0224	2.9062
0.6065	.5		1.0982	1.1210	0.9810	0.9878	1.0116	2.7047
0.4724	.75		1.0629	1.0846	0.9811	0.9865	1.0057	2.4597
0.3679	1		1.0390	1.0565	0.9842	0.9882	1.0026	2.1849
0.1353	2		1.0021	1.0038	0.9984	0.9987	1.0000	1.1520
0.0498	3		1.0024	1.0046	0.9979	0.9938	1.0000	1.1947
0	∞		1.1180	1.2247	0.9166	0.9354	0.9129	4.1835

Table 4: Extrapolation points. $r = (\sigma_c^2 / \sigma_{opt}^2)^{1/2}$ constant mean

$$r_1 = (\sigma_c^2 / \sigma_{opt}^2)^{1/2} \text{ linear trend with } \alpha, \beta, \gamma$$

$$r_3 = (\sigma_b^2 / \sigma_c^2)^{1/2} \sum_{i=1}^4 C_i^* = 1 \text{ only and linear trend with } \alpha, \beta, \gamma$$

$\rho_u(1)$	u	X_0	r	r_1	r_3			
					$\beta=0.2$	$\beta=0.5$	$\beta=1.$	$\beta=10.$
0.9048	.1	(1.5, 0.5)	1.0844	1.1022	0.9875	1.0060	1.0698	4.3138
0.7788	.25		1.0707	1.1133	0.9656	0.9862	1.0563	4.4753
0.6065	.5		1.0541	1.1340	0.9341	0.9580	1.0388	4.7302
0.4724	.75		1.0435	1.1560	0.9079	0.9351	1.0262	4.9645
0.3679	1		1.0373	1.1783	0.8862	0.9165	1.0174	5.1767
0.1353	2		1.0366	1.2617	0.8296	0.8705	1.0030	5.8120
0.0498	3		1.0523	1.3314	0.7998	0.8478	1.0005	6.1844
0	∞		1.1180	1.5	0.7572	0.8165	1.	6.7082
0.9048	.1	(1.5, -0.5)	1.3473	1.4669	0.9278	0.9754	1.1292	6.6319
0.7788	.25		1.2749	1.5156	0.8500	0.8947	1.0388	6.1537
0.6065	.5		1.1992	1.5578	0.7781	0.8204	0.9565	5.7290
0.4724	.75		1.1548	1.5799	0.7391	0.7805	0.9132	5.5237
0.3679	1		1.1274	1.5946	0.7152	0.7564	0.8883	5.4239
0.1353	2		1.0888	1.6407	0.6722	0.7153	0.8517	5.3797
0.0498	3		1.0881	1.6854	0.6546	0.6992	0.8421	5.4450
0	∞		1.1180	1.8028	0.6300	0.6793	0.8320	5.5815
0.9048	.1	(0.5, -0.5)	1.0844	1.1022	0.9839	0.9839	0.9839	0.9839
0.7788	.25		1.0707	1.1133	0.9617	0.9617	0.9617	0.9617
0.6065	.5		1.0541	1.1340	0.9295	0.9295	0.9295	0.9295
0.4724	.75		1.0435	1.1560	0.9027	0.9027	0.9027	0.9027
0.3679	1		1.0373	1.1783	0.8803	0.8803	0.8803	0.8803
0.1353	2		1.0366	1.2617	0.8216	0.8216	0.8216	0.8216
0.0498	3		1.0523	1.3314	0.7904	0.7904	0.7904	0.7904
0	∞		1.1180	1.5	0.7453	0.7453	0.7453	0.7453

At the center of the square the weighting coefficients $\{C_i\}$ are always all equal to $1/4$ by symmetry considerations, regardless of the correlation model, which makes r_1 the same as the ratio r of (1.11). For other locations of interpolation, as given in Table 3, the ratio r_1 is larger than r , (where the mean function was assumed constant), reflecting the fact that we not only have the constraint $\sum_{i=1}^n C_i^* = 1$, but also $\sum_{i=1}^n C_i^* U_i = U_0$ and $\sum_{i=1}^n C_i^* V_i = V_0$. The difference between r_1 and r is small throughout most of the range of Table 3.

The universal unbiasedness constraints have their largest effect on estimation variances when the correlation model is the strongest for the interpolation location given in Table 1 (as $r_1 = r$, for $\underline{X}_0 = (0.5, 0.5)$) and in Table 3. For example, when

$$\begin{aligned} \rho(1) = 0.9, \quad r_1 &\doteq 1.36 \quad \text{for } \underline{X}_0 = (0.5, 0.5) \\ r_1 &\doteq 1.20 \quad \text{for } \underline{X}_0 = (0.5, 0) \\ r_1 &\doteq 1.37 \quad \text{for } \underline{X}_0 = (0.75, 0.25) \\ r_1 &\doteq 1.20 \quad \text{for } \underline{X}_0 = (1., 0.5), \end{aligned}$$

whereas when $\rho(1) = .05$, the corresponding values for r_1 are 1.006, 1.002, 1.001, 1.002. We minimize the mean square error, under the constraints given in (2.3), i.e., minimize the auxiliary function, with three Lagrange multipliers:

$$\begin{aligned} (2.5) \quad A = 1 + \sum_{i=1}^n \sum_{j=1}^n C_i^* C_j^* \rho_{ij} - 2 \sum_{i=1}^n C_i^* \rho_{i0} + 2 \mu_0 \left(\sum_{i=1}^4 C_i^* - 1 \right) + \\ + 2 \mu_1 \left(\sum_{i=1}^4 C_i^* U_i - U_0 \right) + 2 \mu_2 \left(\sum_{i=1}^4 C_i^* V_i - V_0 \right). \end{aligned}$$

For uncorrelated observations (i.e., $u \rightarrow \infty$), we substitute

$$(2.6) \quad \rho_{ij} = \begin{cases} 0 & \text{for } i \neq j \\ 1 & \text{for } i = j \end{cases}$$

and $\rho_{i0} = 0$ unless $i = 0$, in which case $\rho_{00} = 1$. The optimal constrained estimation variance is given by

$$(2.7) \quad \sigma_C^2 = \rho_{00} - (C^*)^T \cdot \rho^*$$

which for the correlogram given in (2.6) results in

$$(2.8) \quad \sigma_C^2 = 1 + \sum_{i=1}^4 C_i^{*2},$$

where

$$C^* = (C_1^*, \dots, C_4^*, \mu_0, \mu_1, \mu_2)^T$$

$$\rho^* = (0, 0, 0, 0, 1, U_0, V_0)^T$$

The optimal constrained coefficients are given by differentiating the auxiliary function w.r.t. C^* and equating to zero,

$$(2.9) \quad C^* = (R^*)^{-1} \cdot \rho^*,$$

where R^* is a 7×7 matrix, having the 4×4 correlation matrix in the upper left side and the other vectors are due to the constraints.

In Table 5 we list the optimal constrained coefficients and the minimal mean squared error for the uncorrelated case for some estimation locations. See the calculated limit value (uncorrelated observations) for r_1 in the last entries ($u \rightarrow \infty$) of Tables 3 and 4.

For extrapolation locations, though, the situation is quite different: The values of r_1 are generally higher than for interpolation and increase with decreasing correlation. The universal unbiasedness constraint have their largest effect at the extrapolation location $\underline{x}_0 = (1.5, -0.5)$ where the ratio r_1 reaches 1.80 for uncorrelated models.

Table 5: The limit values of $\{C_i^*\}$, $i = 1, \dots, 4$ and σ_C^2 for the uncorrelated observations ($u \rightarrow \infty$), assuming a nonspecified linear trend with parameters α, β, γ .

Estimation Location	C_1^*	C_2^*	C_3^*	C_4^*	$\lim_{u \rightarrow \infty} \sigma_C^2$	$\lim_{u \rightarrow \infty} r_1$
(0.5,0.5)	1/4	1/4	1/4	1/4	1.25	1.1180
(0.5,0)	1/2	0	0	1/2	1.5	1.2247
(0.75,0.25)	1/4	0	1/4	1/2	1.375	1.1726
(1.,0.5)	0	0	1/2	1/2	1.5	1.2247
(1.5,0.5)	-1/4	-1/4	3/4	3/4	2.25	1.50
(1.5,-0.5)	-3/4	1/4	1.25	-1/4	3.25	1.8028
(0.5,-0.5)	3/4	-1/4	-1/4	3/4	2.25	1.5

III. Assuming a Constant Mean in the Presence of a Linear Trend.

Now, instead of introducing all three constraints (2.3) on the weight coefficients for unbiased interpolation in the presence of a general linear trend, we simplify the procedure by using only the constraint $\sum_{i=1}^n C_i^* = 1$. The resulting estimator will be the same as that derived in (1.6) (for a constant mean function), however, it will no longer be universally unbiased for all linear trends. The mean squared estimation error may be found by writing

$$(3.1) \quad Z(\underline{X}) = m(\underline{X}) + Y(\underline{X})$$

where $Y(\underline{X})$ is a zero mean random function with the same autocorrelation as $Z(\underline{X})$. Hence, inserting (3.1)

$$\begin{aligned} (3.2) \quad E[Z^*(\underline{X}_0) - Z(\underline{X}_0)]^2 &= E\left[\sum_{i=1}^n C_i^* Z(\underline{X}_i) - Z(\underline{X}_0)\right]^2 = \\ &= E\left[\sum_{i=1}^n C_i^* (Y(\underline{X}_i) + m(\underline{X}_i)) - (Y(\underline{X}_0) + m(\underline{X}_0))\right]^2 \\ &= 1 + \sum_{i=1}^n \sum_{j=1}^n C_i^* C_j^* \rho_{ij} - 2 \sum_{i=1}^n C_i^* \rho_{i0} + \\ &\quad + \left[\sum_{i=1}^n C_i^* m(\underline{X}_i) - m(\underline{X}_0)\right]^2. \end{aligned}$$

We shall denote this MSE of the biased estimator by σ_b^2 . It is equal to the sum of the estimation variance of Table 1 (see 1.7) and

the squared bias $[\sum C_i^* m(\underline{X}_i) - m(\underline{X}_0)]^2$ which depends on $m(\underline{X})$ and the interpolation point $\underline{X}_0$. [By incorrectly assuming a constant mean we would also tend to overestimate the autocorrelations. However, we ignore this difficulty here and use the same correlation function for both the biased and unbiased calculations.]

We shall define two ratios

$$(3.3) \quad r_2 = (\sigma_b^2 / \sigma_{opt}^2)^{1/2}$$

and

$$(3.4) \quad r_3 = (\sigma_b^2 / \sigma_c^2)^{1/2},$$

where σ_c^2 is the optimal constrained MSE, given in (2.7) and σ_{opt}^2 is the optimal unconstrained MSE, with specified known trend, given in (1.3). Each of the above defined ratios will depend on the parameters α, β, γ of the linear trend function

$$m(\underline{X}) = \alpha + \beta U + \gamma V \quad \text{where } \underline{X} = (U, V).$$

To simplify comparisons we take $\alpha = \gamma = 0$ throughout and values of $\beta = (0.2, 0.5, 1.0, 10.0)$, which correspond to an angle θ between the U axis and the linear trend of about $12^\circ, 30^\circ, 45^\circ$ and 85° .

See Tables 3 and 4 for the ratio r_3 , with the above four mentioned β parameters using the autocorrelation model $\rho(h) = e^{-uh^2}$, where h is distance and $u > 0$ is a parameter, in the range $\rho(1) \doteq 0.9$ till $\rho(1) \equiv 0$.

For interpolation at the center of the unit square, the value of the ratios r_2, r_3 are the same, for any β . Also $r = r_1 = r_2$ for this interpolation point. In general, the ratio r_2 behaves very similarly to the ratio r (1.11) at any interpolation location, and therefore has not been tabulated.

In Table 3 the ratio r_3 is shown for three different interpolation points $\underline{X}_0$. For moderate trend angles (up to 45°) this ratio is remarkably very close to one, regardless of the strength of correlations. [At the interpolation point $\underline{X}_0 = (0.5, 0)$, r_3 is the same for all values of β .] Hence, there is nothing lost by using only the constraint $\sum_{i=1}^n C_i^* = 1$ in the estimate rather than the full set of constraints given in (1.18). However, for extreme trend angles ($\beta = 10, \theta \doteq 85^\circ$), the additional constraints, which assure an unbiased estimate, will also provide an estimate with much smaller MSE.

In Table 4 three different extrapolation locations are shown. [At the extrapolation point $\underline{X}_0 = (0.5, -0.5)$ the ratio r_3 can be shown to be the same for any value of the trend parameter β .]

We see from the table that when the angle of the trend is less than about 45° , then the biased estimator using the single constraint $\sum_{i=1}^n C_i^* = 1$ will perform as well as, and sometimes noticeably better, than the fully constrained unbiased estimator (1.6). However, for very steep trends the reverse is true at the extrapolation points $(1.5, .5)$ and $(1.5, -.5)$.

References

- [1] Mathéron, G., (1965), "Les Variables Regionalisées et Leur Estimation". Paris, Masson et Cie.
- [2] Mathéron, G., (1972), "Geostatistique", an article in Ouvrage Collectif., "Traite d'information Geologique", direction: P. Laffitte, Masson et Cie., p. 357.
- [3] Mathéron, G., (1973), "The intrinsic random functions and their applications". Adv. in Appl. Prob. 5, pp. 439-468.
- [4] Sheshinski, R., (1976), "Robustness to Misspecification of Correlation Models". Stanford Technical Report No. 2, Prepared under support of SIMS from ERDA Rockefeller Foundation.