MASTER

MISSPECIFICATION OF TREND IN SPATIAL RANDOM-FUNCTION INTERPOLATION
WITH APPLICATION TO OXIDANT MAPPING

RAY FAITH and RUTH SHESHINSKI

TECHNICAL REPORT NO. 28

SEPTEMBER 1979

PREPARED UNDER THE AUSPICES OF
SIAM INSTITUTE FOR MATHEMATICS AND SOCIETY

SIMS

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

Prepared For
THE U.S. DEPARTMENT OF ENERGY
UNDER CONTRACT NO. EY-76-S-02-2874

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.

MASTER

DISCLAIMER
This book 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.

**MISSPECIFICATION OF TREND IN SPATIAL RANDOM-FUNCTION INTERPOLATION
WITH APPLICATION TO OXIDANT MAPPING**

**Ray Faith
Stanford University**

and

**Ruth Sheshinski
The Hebrew University
of Jerusalem**

TECHNICAL REPORT NO. 28

SEPTEMBER 1979

**STUDY ON STATISTICS AND ENVIRONMENTAL
FACTORS IN HEALTH (SIMS)**

PREPARED UNDER SUPPORT TO SIMS FROM

DEPARTMENT OF ENERGY (DOE)

ROCKEFELLER FOUNDATION

SLOAN FOUNDATION

ENVIRONMENTAL PROTECTION AGENCY (EPA)

NATIONAL SCIENCE FOUNDATION (NSF)

**DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA**

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

MISSPECIFICATION OF TREND IN SPATIAL RANDOM-FUNCTION INTERPOLATION
WITH APPLICATION TO OXIDANT MAPPING

Ray Faith
Stanford University

and

Ruth Sheshinski
The Hebrew University
of Jerusalem

Abstract

In epidemiological studies of the effects of air pollution it is desirable to obtain estimates of the exposure to various pollutants of persons living throughout a geographical region under study. On the other hand, it is typically the case that pollution levels are measured at only a few locations. The problems of interpolating these values and of determining the accuracy of interpolation are amenable to analysis by way of spatial stochastic models. Such models can lead to an optimal interpolation technique, in a manner which is reviewed here.

In Section 1 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 section 2 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 section 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 these sections 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.

In Section 4 we use the Kriging method for interpolating (mapping) values of oxidant levels from 1973 at 16 stations located in the San Francisco Bay Area. We use for this purpose constant and linear trends. The method requires for its application knowledge of the spatial covariance structure of the pollutant to which it is applied. Some difficulties encountered in estimating this for oxidant levels in the San Francisco Bay Area are mentioned.

Key Words: Kriging, spatial correlation, temporal correlation, isopleths, oxidant levels, trend misspecification, isotropy, anisotropy, spatial and temporal correlation estimates, spatial linear interpolation.

0. Introduction

The Kriging Method. The method discussed here has been developed extensively by other authors, notably Mathéron (1963, 1973) and Delfiner (1976), and takes its name from Krige (1951). The following is an elementary special case of the more general theory.

Let $Z(\underline{X})$, $\underline{X} \in R$, be a real valued stochastic process, where R is a region in the plane. Let $\underline{X}_1, \underline{X}_2, \dots, \underline{X}_n$ be n points in R at which $Z(\cdot)$ is known. Let us denote $S = (\underline{X}_1, \underline{X}_2, \dots, \underline{X}_n)'$ and $Z_S = (Z(\underline{X}_1), Z(\underline{X}_2), \dots, Z(\underline{X}_n))'$. In the following, we shall consider only linear estimators of $Z(\underline{X})$, of the form

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

The error of estimation at a point $\underline{X}$ is $Z^*(\underline{X}) - Z(\underline{X})$.

We introduce the following assumptions concerning the stochastic model.

- (a) The process $Z(\underline{X})$ has a finite variance.
- (b) $E Z(\underline{X}) = M$ a constant.
- (c) $\frac{1}{2} E[Z(\underline{X}+\underline{h}) - Z(\underline{X})]^2 = c(\underline{0}) - c(\underline{h})$ depends only on $\underline{h}$, where

$$c(\underline{h}) = \text{cov}(Z(\underline{X}+\underline{h}), Z(\underline{X})) .$$

Assuming, without loss of generality, that the stationary variance $E(Z(\underline{X}) - M)^2$ equals one, we deal with the correlation function $\rho(\underline{h})$ and $\rho(\underline{0}) = 1$, hence

$$(c') \quad \frac{1}{2} E[Z(\underline{X}+\underline{h}) - Z(\underline{X})]^2 = 1 - \rho(\underline{h}) .$$

Assumption (c), however, requires Z to be a weakly stationary stochastic process. (In case Z is a process with drift, the more general representation of Mathéron has to be used.)

It is immediate from assumption (b) that a linear estimator $Z^*(\underline{X})$ is unbiased in the sense that the expected error is zero, $(E(Z^*(\underline{X}) - Z(\underline{X})) = 0)$ if and (except for the special case $M=0$ which is worked out in the next section) only if $\sum_{i=1}^n C_i \equiv 1$. An unbiased linear estimator is, therefore, a weighted average of the data values.

The Kriging method consists of taking $Z^*(\underline{X})$ to be that unbiased linear estimator of $Z(\underline{X})$ which minimizes the mean squared error $MSE = E[Z^*(\underline{X}) - Z(\underline{X})]^2$. An application of Lagrange multipliers shows MSE to be minimized by

$$(0.1) \quad Z^*(\underline{X}) = \hat{M} + \Sigma'_{SX} \Sigma^{-1}_{SS} (Z_S - \hat{M}\underline{1})$$

where

$$\Sigma_{SX} = (\rho(\underline{X}_1 - \underline{X}), \dots, \rho(\underline{X}_n - \underline{X}))'; \quad n \times 1$$

$$\Sigma_{SS} = ((\rho(\underline{X}_i - \underline{X}_j))); \quad n \times n$$

$$\underline{1} = (1, \dots, 1)'; \quad (n \times 1)$$

and

$$\hat{M} = \frac{\underline{1}' \Sigma^{-1}_{SS} Z_S}{\underline{1}' \Sigma^{-1}_{SS} \underline{1}}.$$

$\hat{M}$ is the usual minimum variance unbiased estimator of M and $Z^*(\underline{X})$ the usual regression estimator of $Z(\underline{X})$ given Z_S , with $\hat{M}$ substituted for M . For proof see Goldberger (1962).

In addition to producing the optimal estimator $Z^*(\underline{X})$, the above framework tells us its mean squared error, namely

$$(0.2) \quad MSE = E[Z^*(\underline{X}) - Z(\underline{X})]^2 = \rho(0) - \Sigma'_{SX} \Sigma^{-1}_{SS} \Sigma_{SX} + \frac{(\underline{1}' \Sigma^{-1}_{SS} \Sigma_{SX} - 1)^2}{\underline{1}' \Sigma^{-1}_{SS} \underline{1}}.$$

Examination of (0.1) and (0.2) shows that, at each data location $\underline{X} = \underline{X}_i$ the estimator $Z^*(\underline{X})$ equals the data value $Z(\underline{X}_i)$ and the error variance is zero. If $\rho(\underline{h})$ tends to zero for large $\underline{h}$ then the estimate tends to $\hat{M}$ at points far from the data and the error variance tends to

$$\rho(0) + \frac{1}{\underline{1}' \Sigma_{SS}^{-1} \underline{1}} = \text{Var } Z(\underline{X}) + \text{Var } \hat{M} .$$

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

Suppose we have observed pollutant concentrations $Z(\underline{x}_i)$, $i=1, \dots, n$ at n monitoring stations $\underline{x}_1, \dots, \underline{x}_n$, in the plane region R^2 . We regard these as a single realization $z(\underline{x})$ out of the (weak stationary) stochastic process $Z(\underline{x})$ on R^2 with mean function $M(\underline{x})$ and spatial correlation function $\rho(\underline{x})$.

We define a linear predictor $Z^*(\cdot)$ at the point $\underline{x}_0$ as

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

We calculate "optimal" coefficients $C_0, \dots, C_n$, to fulfill the criteria

$$(i) \quad E[Z^*(\underline{x}_0)] = M(\underline{x}_0) ,$$

$$(ii) \quad \text{minimize } E[Z^*(\underline{x}_0) - Z(\underline{x}_0)]^2 .$$

Condition (i) implies

$$(1.2) \quad C_0 = M(1 - \sum_{i=1}^n C_i) , \text{ for a specified constant mean } M .$$

The resulting MSE will then not depend on M , and hence also the optimal coefficients $C_1, \dots, C_n$ will be independent of M .

$$(1.3) \quad \text{MSE} = \sigma_{\text{opt}}^2 = E[Z^*(\underline{x}_0) - Z(\underline{x}_0)]^2 = E[M(1 - \sum_{i=1}^n C_i) + \sum_{i=1}^n C_i Z(\underline{x}_i) - Z(\underline{x}_0)]^2 \\ = 1 + \sum_{i=1}^n \sum_{j=1}^n C_i C_j \rho_{ij} - 2 \sum_{i=1}^n C_i \rho_{i0} ,$$

where $\rho_{ij} = \rho(|\underline{x}_i - \underline{x}_j|)$. Here $E(Z-M)^2$ is taken to equal one. Minimizing σ_{opt}^2 w.r.t. the C_i 's results in n optimal unconstrained coefficients.

If, however, M is incorrectly specified, then the prediction $Z^*(\underline{X}_0)$ will be neither optimal, nor will $Z^*(\underline{X}_0)$ be unbiased.

Let M' be the incorrectly specified mean. The mean squared estimation error $\sigma_{M'}^2$ is:

$$(1.4) \quad \sigma_{M'}^2 = E[Z^*(\underline{X}_0) - Z(\underline{X}_0)]^2 = E\left\{[M(1 - \sum_{i=1}^n C_i) + \sum_{i=1}^n C_i Z(\underline{X}_i) - Z(\underline{X}_0)] + (M-M') \left(\sum_{i=1}^n C_i - 1\right)\right\}^2 = \sigma_{opt}^2 + (M-M')^2 \left(\sum_{i=1}^n C_i - 1\right)^2.$$

On the other hand if we do not specify M , and wish to use an estimator that is unbiased for all M , then we require - according to criterion (i) - that $\sum_{i=1}^n C_i \equiv 1$ and hence, equation (1.2) results in $C_0 \equiv 0$. This furnishes a constrained solution of $C_1, \dots, C_n$, denoted $\{C_i^*\}$. The "optimal" unbiased estimator, $Z^*(\underline{X}_0)$ defined as:

$$(1.5) \quad Z^*(\underline{X}_0) = \sum_{i=1}^n C_i^* Z(\underline{X}_i) \quad (C_0 \equiv 0)$$

uses the "optimal" constrained coefficients, which are achieved by minimizing the MSE denoted σ_C^2 , and calculated as in (1.3).

$$(1.6) \quad \sigma_C^2 = 1 + \sum_{i=1}^n \sum_{j=1}^n C_i^* C_j^* \rho_{ij} - 2 \sum_{i=1}^n C_i^* \rho_{i0}$$

under the linear constraint $\sum_{i=1}^n C_i^* = 1$.

We define a relative measure of performance, denoted r , for a specified constant trend M versus an unspecified constant trend, as follows:

$$(1.7) \quad r = [\sigma_C^2 / \sigma_{opt}^2]^{1/2}.$$

We have simulated prediction outcomes, for the situation of four observation points, interpolating as well as extrapolating. (In Section 4 air pollution data with the constant trend assumption are plotted.)

Simulation of Constant Trend Models. 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 1.1 for the geometrical representation of the observations and estimation locations.

Figure 1.1. Geometrical set-up for four observations, and the coordinate system.

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

$$(1.8) \quad \rho(h) = e^{-uh^2}; \quad u > 0.$$

Ratio (1.7) is easily calculated for $n=4$ uncorrelated observations. As $u \rightarrow \infty$, 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 the ratio r in (1.7) gives:

$$(1.9) \quad \lim_{u \rightarrow \infty} (\sigma_C^2 / \sigma_{\text{opt}}^2)^{1/2} = 1.118 .$$

For data with moderately correlated observations, this relative performance ratio can be very close to unity. See Table 1.1 for examples of the above mentioned ratio (1.7). Also tabulated in Table 1.1 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 (1.7) is much closer to the value 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 $\rho(1) = 0.05$. The performance ratio is similarly not a monotonic function of the strength of the correlation model, but the worst relative performance 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 (1.7)

Table 1.1. 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	r	Uncon- strained C_i	σ_C^2	r	Constrained Coefficient		Unconstrained Coefficient	
		σ_{opt}^2			σ_{opt}^2		$C_1^*=C_2^*$	$C_3^*=C_4^*$	$C_1=C_2$	$C_3=C_4$
0.9048	.1	0.0046 0.0025	1.3639	0.2622	0.0130 0.0111	1.0844	-.2256	0.7256	-.2372	0.7410
.6065	.50	0.0876 0.0600	1.2087	0.3017	0.1965 0.1769	1.0541	-.1394	0.6394	-.1830	0.5958
.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
.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

starts to increase monotonically with weaker correlation assumptions.

The unconstrained coefficients are zero in the uncorrelated case, which means that the linear estimator $Z^* = M$ 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 $\sum_{i=1}^n C_i < 1$. The ratio r in (1.7) appears to decrease with decreasing correlation, as long as the sum of absolute values of uncon-

strained coefficients > 1 and increases until reaching the limit in

(1.9) 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 (1.7) is not a monotonic

function of the correlation parameter. It achieves the worst performance

rate 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 (1.5) which is universally unbiased. In (1.4) 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 $\sigma_{M'}^2$ is still less than the corresponding value of σ_C^2 (1.6), for the constrained estimator (1.5) under the same underlying correlation assumptions.

Using (1.4) the estimation variance σ_C^2 is less than σ_M^2 , whenever

$$(1.10) \quad (\sigma_C^2 - \sigma_{opt}^2)^{1/2} / \left| \left(\sum_{i=1}^n C_i - 1 \right) \right| < |M - M'|.$$

In Table 1.2 are shown values of the LHS of (1.10) 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, i.e. $|M - M'| = 1$, then, applying (1.10), the constrained universally unbiased estimator is preferable to the biased unconstrained estimator. However, for uncorrelated models, a discrepancy between M and M' of ± 0.5 standard deviations would make the unconstrained biased estimator preferable.

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 minimizing (1.3) and

$$(1.11) \quad C_0 = M(\underline{X}_0) = \sum_{i=1}^n C_i M(\underline{X}_i).$$

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.

Table 1.2. C_i^* = constrained coefficient; $\sum C_i^* = 1$

C_i = unconstrained coefficient; $S = \sum_{i=1}^4 C_i$

σ_C^2 = constrained estimation variance

σ_{opt}^2 = unconstrained estimation variance

$$|M-M'| = (\sigma_C^2 - \sigma_{opt}^2)^{1/2} / |(S-1)|$$

$\rho_u(1)$ u	$\underline{x}_0 = (0.5, 0.5)$			$\underline{x}_0 = (1.5, 0.5)$			$\underline{x}_0 = (0.1, 0.1)$		
	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	0.00059 0.00032	0.9553
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
0.3679 1	1.2964	0.2547 0.2135	0.6848	0.7362	0.4611 0.4285	0.6844	1.1004	0.02913 0.02441	0.6843
0.2725 1.3	1.2892	0.3607 0.3268	0.6366	0.6926	0.5917 0.5535	0.6358	1.1022	0.4142 0.3719	0.6363
0.0498 3	0.8100	0.8293 0.8193	0.5263	0.4058	1.0022 0.9049	0.5249	1.0166	0.1102 0.1101	0.6024
0 ∞	0	1.25 1	0.50	0	1.25 1	0.5	0	1.25 1	0.5

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 (1973) where the trend is given as an analytic expression

$$(1.12) \quad M(\underline{X}) = \sum_{i=0}^k a_i f_i(\underline{X})$$

where $\{f_i(\underline{X})\}$ are specified functions and $\{a_i\}$ are unspecified constants. For example, (1.12) can be used to represent a polynomial trend in the mean function. Then the conditions on the weights $\{C_j\}$ in (1.1), so that $Z^*(\underline{X}_0)$ is unbiased for any $\{a_i\}$ are:

$$(1.13) \quad \begin{aligned} E[Z^*(\underline{X}_0)] &= \sum_{j=1}^n C_j E[Z(\underline{X}_j)] = \sum_{j=1}^n C_j \sum_{i=0}^k a_i f_i(\underline{X}_j) \\ &= \sum_{i=0}^k \sum_{j=1}^n a_i C_j f_i(\underline{X}_j) . \end{aligned}$$

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

$$(1.14) \quad \sum_j C_j f_i(\underline{X}_j) = f_i(\underline{X}_0), \quad i=0, 1, 2, \dots, k .$$

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

The term:

$$(1.15) \quad \sigma_{ij} = E[(Z(\underline{X}_i) - M(\underline{X}_i))(Z(\underline{X}_j) - M(\underline{X}_j))]$$

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.

1. 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 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

$$(2.3) \quad \left\{ \begin{array}{l} \sum_{i=1}^n C_i^* = 1 \\ \sum_{i=1}^n C_i^* U_i = U_0 \\ \sum_{i=1}^n C_i^* V_i = V_0 \end{array} \right.$$

where (U_0, V_0) are the coordinates of $\underline{X}_0$. These constraints follow from (1.14). The estimation variance for the optimal constrained estimator, which is unbiased for all α, β, γ , was denoted 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 set-up of four observations located at the corners of a unit square. See Figure 1.1

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 (1.6) subject to the constraints (2.3), using Lagrange multipliers. For illustration, autocorrelations were taken from the model $\rho(h) = e^{-uh^2}$ where h denotes distance and $u > 0$ is a parameter.

Let

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

See Table 2.1 - for the points of interpolation except $\underline{X}_0 = (0.5, 0.5)$ and Table 2.2 - for extrapolation locations, using underlying autocorrelation models ranging from $\rho(1) = 0.05$ to $\rho(1) = 0.9$.

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 (1.7). For other locations of interpolation, as given in Table 2.1 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 2.1.

The universal unbiasedness constraints have their largest effect on estimation variances when the correlation model is the strongest for

Table 2.1: Interpolation points.

(defined in 2.1) $r = (\sigma_C^2 / \sigma_{\text{opt}}^2)^{1/2}$; constant mean

(defined in 3.3) $r_1 = (\sigma_C^2 / \sigma_{\text{opt}}^2)^{1/2}$; linear trend with α, β, γ

(defined in 3.4) $r_3 = (\sigma_b^2 / \sigma_C^2)^{1/2}$; only one constraint in σ_b^2 , $\left(\sum_{i=1}^4 C_i^* = 1 \right)$,
versus a 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
0.6065	.5		1.0982	1.1210	0.9797	0.9797	0.9797	0.9797
0.3679	1		1.0390	1.0565	0.9834	0.9834	0.9834	0.9834
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
0.6065	.5		1.2058	1.2343	0.9777	0.9815	0.9952	2.1352
0.3679	1		1.0916	1.1273	0.9691	0.9730	0.9868	2.1296
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
0.6065	.5		1.0982	1.1210	0.9810	0.9878	1.0116	2.7047
0.3679	1		1.0390	1.0565	0.9842	0.9882	1.0026	2.1849
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 2.2: Extrapolation points.

(defined in 2.1) $r = (\sigma_C^2 / \sigma_{\text{opt}}^2)^{1/2}$; constant mean.

(defined in 3.3) $r_1 = (\sigma_C^2 / \sigma_{\text{opt}}^2)^{1/2}$; linear trend with α, β, γ

(defined in 3.4) $r_3 = (\sigma_b^2 / \sigma_C^2)^{1/2}$; only one constraint in σ_b^2 , $\left(\sum_{i=1}^4 C_i^* = 1 \right)$,
versus a 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	(1.5, 0.5)	1.0844	1.1022	0.9875	1.0060	1.0698	4.3138
0.6065	.5		1.0541	1.1340	0.9341	0.9580	1.0388	4.7302
0.3679	1		1.0373	0.1783	0.8862	0.9165	1.0174	5.1767
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
0.6065	.5		1.1992	1.5578	0.7781	0.8204	0.9565	5.7290
0.3679	1		1.1274	1.5946	0.7152	0.7564	0.8883	5.4239
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
0.6065	.5		1.0541	1.1340	0.9295	0.9295	0.9295	0.9295
0.3679	1		1.0373	1.1783	0.8803	0.8803	0.8803	0.8803
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

the interpolation location given in Table 1 (as $r_1 = r$, for $\underline{x}_0 = (0.5, 0.5)$) and in Table 2.1. For example, when

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

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:

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

The optimal constrained estimation variance is given by

$$(2.6) \quad \sigma_C^2 = \rho_{00} - (C^*)' \cdot \rho^*$$

where

$$\begin{aligned}C^* &= (c_1^*, \dots, c_4^*, \mu_0, \mu_1, \mu_2)' \\ \rho^* &= (\rho_{01}, \rho_{02}, \rho_{03}, \rho_{04}, 1, U_0, V_0)'\end{aligned}$$

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

$$(2.7) \quad C^* = (R^*)^{-1} \cdot \rho^*,$$

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.

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

$$(2.8) \quad \rho_{ij} = \begin{cases} 0 & \text{for } i \neq j \\ 1 & \text{for } i = j \end{cases}$$

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

$$(2.9) \quad \sigma_C^2 = 1 + \sum_{i=1}^4 C_i^{*2}.$$

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 constraints 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.

For an application of mapping air pollution data with a linear trend model see Section 4.

3. 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 (1.5) (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) into the MSE equation results in:

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

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 1.6) 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 (2.6) 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 2.1 and 2.2 for the ratio r_3 , with the above four mentioned β parameters using the autocorrelation model $\rho(h) = e^{-uh^2}$, where h denotes the distance and $u > 0$ is a parameter, in the range $\rho(1) = 0.9$ 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 (1.7) at any interpolation location, and therefore has not been tabulated.

In Table 2.1 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 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 (1.14). However, for extreme trend angles ($\beta = 10, \theta = 85^\circ$), the additional constraints, which assure an unbiased estimate, will also provide an estimate with much smaller MSE.

In Table 2.2 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 (1.5). However, for very steep trends the reverse is true at the extrapolation points $(1.5, .5)$ and $(1.5, -.5)$.

4. Interpolation of Oxidant Levels.

a) Extension of the Model: The discussion so far has dealt solely with the interpolation of data on a two dimensional surface. The situation with air pollution data is somewhat different since the data are collected not only at fixed locations throughout a region but in addition are taken repeatedly in time. Thus it is natural to attempt to estimate, for example, the average pollution at some specified location over a specified period of time, using the values collected at all monitoring stations in the region over an interval of time containing the period of interest.

In fact the addition of the temporal dimension need not increase the theoretical difficulty of the problem. Precisely the same method of interpolation as described above still applies to the present situation, provided we are willing to make the assumption that the pollution level $Z(\underline{X},t)$, now indexed by location and time, is a stationary random process with mean function $M(\underline{X},t)$, variance σ_Z^2 , and correlation function $\rho(\underline{X},t)$. If $M(\underline{X},t)$ is known a priori or if it is assumed to be a linear combination of known functions of $(\underline{X},t)$ with unknown coefficients, then the interpolation methods described above apply equally here. This is the approach we have taken, together with simplifying assumptions to reduce the computational difficulty of estimation and interpolation.

b) The Data: The raw data consist of average oxidant levels at each of 16 stations in the San Francisco Bay Area air basin, collected daily for 112 days in mid 1973. We transformed these to a log scale (after adding a small constant) to eliminate the potential harmful effects of their long tailed distribution on our linear models.

Overall there are $n \cdot T = 16 \cdot 112 = 1792$ observations $Z(\underline{X}_i, t)$,
with overall mean

$$\bar{Z} = 3.0828$$

and variance

$$s^2 = \sum_i \sum_t (Z(\underline{X}_i, t) - \bar{Z})^2 / nT = .1574 .$$

The means for different stations are substantially different from each other, with between station variance

$$s_{\bar{Z}}^2 = \sum_i (\bar{Z}(\underline{X}_i, \cdot) - \bar{Z})^2 / n = .0550 ,$$

which accounts for 35 percent of the total variance. As can be observed in Figure (4.1), the station mean values do not follow a simple spatial pattern although there may be a tendency to higher values at inland locations.

The observed difference in station means can be accounted for in either of two essentially equivalent ways: either they are the result of differences in expected value at these locations, to be reflected in $M(\underline{X}, t)$, or they are caused by random perturbations in the process $Z(\underline{X}, t)$ which appear fixed for the duration of the observations, but which might be fluctuating on a longer time scale, say over millions of years. In the latter case, the mean function could be a constant throughout the region, but the correlation function $\rho(\underline{X}, \tau)$ would tend to a nonzero asymptote for large τ . Although these two models are different, their practical consequences are virtually the same, so we elect the one which is simpler to use, namely the latter one. To be precise we begin by modelling the process as having a constant mean, which we estimate by

$\bar{Z}$, then use deviations from this value to estimate the variance and correlation function. We estimate the variance σ_Z^2 by the sample variance S^2 . The correlation estimator is given in the following section.

c) Further Simplifications. In order to simplify both the problem of estimating the model parameters and the subsequent problem of interpolating, we make the following important assumption

$$(4.1) \quad \rho(\underline{X}, t) = g(\underline{X}) \cdot h(t) , \quad g(\underline{0}) = h(0) = 1 .$$

In terms of the estimation problem, this assumption permits the estimation problem to be broken into two parts, that of estimating $g(\underline{X})$ and that of estimating $h(t)$. Elementary estimators of these are provided by simply setting the parameters t and $\underline{X}$ to zero, respectively. In particular, for each pair of stations i and j we have the following estimator of $g(\underline{X}_i - \underline{X}_j) = \rho(\underline{X}_i - \underline{X}_j, 0)$:

$$(4.2) \quad g_{i,j}^* = \sum_t [Z(\underline{X}_i, t) - \bar{Z}] [Z(\underline{X}_j, t) - \bar{Z}] / (TS^2) .$$

Similarly, for each time lag s we have the following estimator of $h(s) = \rho(\underline{0}, s)$:

$$(4.3) \quad h_s^* = \sum_i \sum_{t=1}^{T-s} [Z(\underline{X}_i, t+s) - \bar{Z}] [Z(\underline{X}_i, t) - \bar{Z}] / [n(T-s)S^2] .$$

Except for the use of the constant mean in these expressions, $g_{i,j}^*$ is just the sample lag zero cross correlation between stations i and j , and h_s^* is the average sample lag s autocorrelation.

We calculated and smoothed these estimates by plotting them and then fitting them to reasonable parametric families. In the case of the temporal correlation $h(t)$, the asymptotic correlation between

$Z(\underline{X}_1, t)$ and $Z(\underline{X}_1, t+s)$, as s grows large, should equal the variance of the "constant" part of $Z(\underline{X}_1, t)$. This variance is estimated by the between station sample variance $S_{\bar{Z}}^2$. We used the ratio of this to S^2 as the asymptote of $\hat{h}(s)$ (see Figure (4.2)). Otherwise the fitting consisted of choosing parameters to minimize the sum of absolute deviations between the observed correlations and the parametric values.

This procedure yielded the following smoothed correlation functions:

$$(4.4) \quad \hat{h}(s) = .349 + (1 - .349) \exp(-.515s)$$

and, letting $\underline{X} = (X_1, X_2)$,

$$(4.5) \quad \hat{g}(\underline{X}) = \begin{cases} 1 & \text{if } \underline{X} = \underline{0} \\ \hat{\alpha} + \hat{\beta} \exp(-\hat{\gamma} X_1^2 - \hat{\delta} X_2^2) & \text{otherwise} \end{cases}$$

In (4.5) two sets of best fitting parameter values were obtained:

Type I. Isotropic: $\hat{\alpha} = .3$, $\hat{\beta} = .56$, $\hat{\gamma} = \hat{\delta} = .005$,

Type II. Anisotropic: $\hat{\alpha} = .2$, $\hat{\beta} = .48$, $\hat{\gamma} = .006$, $\hat{\delta} = .001$.

Note: The above anisotropic parameter estimators for $g(\underline{X})$ are not quite the best fitting ones found, in fact. The best fitting value of δ among multiples of .001 is the value zero. However this choice leads to isopleth maps which are relatively uninteresting (essentially a series of vertical lines) and so the above nearly optimal value was used here instead.

d) Optimal Interpolation. The splitting of the correlation function of (4.1) simplifies not only the problem of estimating this function, but more importantly for our purpose, it greatly reduces the computational burden of interpolation, by allowing the estimation to proceed in a

sequence of manageable steps. We illustrate this by the following proposition, stated without proof.

Assumptions: Suppose $Z(\underline{X}, t)$ is stationary with mean function

$$M(\underline{X}) = \sum_{j=1}^P \alpha_j f_j(\underline{X}) \quad f_j(\cdot) \text{ given, } \underline{\alpha} \text{ unknown}$$

(note: $M(\cdot)$ depends only on $\underline{X}$),

and known correlation function $\rho(\underline{X}, t) = g(\underline{X}) h(t)$.

Notation: Let F be the $n \times p$ array of values of the p mean functions evaluated at the n stations:

$$F_{n \times p} = ((F_{ij})) = ((f_j(\underline{X}_i))) .$$

Let G be the $n \times n$ array of spatial correlations between the n stations:

$$G_{n \times n} = ((g(\underline{X}_i - \underline{X}_j))) .$$

Let H be the $T \times T$ array of temporal autocorrelations:

$$H_{T \times T} = ((h(t-s))) .$$

Conclusions:

a) The best linear unbiased estimator of $\underline{\alpha}$ is

$$(4.6) \quad \hat{\underline{\alpha}} = (F' G^{-1} F)^{-1} F' G^{-1} \underline{M}^*$$

where $\underline{M}^* = (M_1^*, \dots, M_n^*)'$ and

$$(4.7) \quad M_i^* = (\underline{1}' H^{-1} \underline{1})^{-1} \underline{1}' H^{-1} \begin{pmatrix} z(\underline{X}_i, 1) \\ \vdots \\ z(\underline{X}_i, T) \end{pmatrix}, \quad \forall i, i=1, \dots, n .$$

b) The best linear unbiased estimator of $M(\underline{X})$ at any point $\underline{X}$ is

$$(4.8) \quad \hat{M}(\underline{X}) = \sum_{j=1}^p \hat{\alpha}_j f_j(\underline{X}) .$$

c) The best linear unbiased estimator of $Z(\underline{X}, t)$ is

$$(4.9) \quad Z^*(\underline{X}, t) = \hat{M}(\underline{X}) + (\dots, g(\underline{X} - \underline{X}_1), \dots) G^{-1} \begin{bmatrix} Z(\underline{X}_1, t) \\ \vdots \\ \hat{M}(\underline{X}_1) \\ \vdots \end{bmatrix} .$$

d) The best linear unbiased estimator of any given linear combination of $Z(\underline{X}, t)$, e.g.

$$\sum_{j=1}^q w_j Z(\underline{X}, t_j) ,$$

is the corresponding linear combination of $Z(\underline{X}, t)$, viz.

$$(4.10) \quad \sum_{j=1}^q w_j Z^*(\underline{X}, t_j) .$$

In each of these conclusions, by "best" is meant that function of all $n \times T$ observations which has smallest mean squared error.

From the computational perspective, each of the estimation steps described above is quite manageable with the exception of (4.7), which requires the inversion of a $T \times T = 112 \times 112$ matrix. This could be circumvented by simply using for M_i^* the station mean $\bar{Z}(\underline{X}_i, \cdot)$, to which it will be very nearly equal (M_i^* is a weighted average over t of $Z(\underline{X}_i, t)$ with nearly constant weights). We circumvented the problem by estimating M_i^* by a weighted average of 16 seven-day means, using (4.7) with a modified 16×16 H matrix of autocorrelations for seven-day means calculated from the daily autocorrelations of (4.4).

e) Results. Once the $\hat{h}(t)$ and $\hat{g}(X)$ models were estimated, from the four month mid-1973 data, we applied them to plot daily or one week or one month data over all 16 stations. For the trend, though, we encountered two different approaches: 1) Used the trend that was estimated by the four month data (constant or linear). 2) Used an estimated trend function that resulted from the partial data (day, week, or month) which is being plotted.

The justification to use the daily interpolated BLUE $Z^*(X,t)$ as given in (4.9) for averaged values (weekly or four-month) is given in the reasoning of formula (4.10).

The trend estimates from the four month data are different for each of the various correlation models. (Type I or II, given in (4.5).)

Linear and constant trend estimates based on four month data and estimates based on different data bases are described in Table 4.1.

Figure 4.3 a and b show four month data, plotted according to spatial correlation model Type I, with (a) a linear and (b) a constant trend estimate. The isopleths are only moderately altered by the different underlying trend models. Figure 4.4 a, b, and c are maps based on data taken for one week averages (for all 16 stations). The week was chosen to be the sixth week of the four month data base, at random. The Kriging method uses $\hat{h}(t)$ and $\hat{g}(X)$ of the Type I model. Map (a) uses a linear trend achieved from the four month data, (b) uses the estimated linear trend based on the one week's data, (c) as in b - only for a constant trend model. For results of the trend estimates consult Table 4.1.

Table 4.1. Summary of estimated spatial correlation parameters and trend parameters for Type I and II correlation models.

$$\text{Type I: } g(\underline{X}) = 0.3 + 0.56 \exp\{-0.005(X_1^2 + 1 \cdot X_2^2)\}$$

Figure	Data Plotted	Trend Model	Trend Estimate
4.3a	4 month	Linear	$2.8678 + 0.01X_1 + 0.0009X_2$
b	"	Constant	3.1091
4.4a	6th week	Linear-4 month	Same as above
b	"	Linear	$2.7107 + 0.0129X_1 + 0.0027X_2$
c	"	Constant	3.0931
4.5a	32nd day	Linear-4 month	Same as above
b	"	Linear	$3.0105 + 0.0022X_1 + 0.0020X_2$

$$\text{Type II: } g(\underline{X}) = 0.2 + 0.48 \exp\{-0.006(X_1^2 + 0.1667X_2^2)\}$$

Figure	Data Plotted	Trend Model	Trend Estimate
4.6a	4 month	Linear	$2.9557 + 0.0072X_1 + 0.0004X_2$
b	"	Constant	3.1192
4.7a	6th week	Linear-4 month	Same as above
b	"	Linear	$2.7738 + 0.0105X_1 + 0.0027X_2$
c	"	Constant	3.1031
4.8a	32nd day	Linear-4 month	Same as above
b	"	Linear	$3.0885 - 0.0019X_1 + 0.0030X_2$

The temporal correlation used with both types of spatial correlations:

$$\hat{h}(t) = 0.3493 + 0.6507 \exp(-0.51|t|). \quad \rho(\underline{X}, t) = g(\underline{X}) \cdot h(t).$$

It seems evident - from checking high and low values on the map - that different trend models or trend estimates from different data bases have almost no effect on the general pattern of the maps.

Figure 4.5 describes the one day (32nd day chosen at random) Ozone levels, Type I correlation model with (a) a linear estimated trend achieved from the full four month data, and (b) estimated linear trend based only on 32nd day's data.

The maps seem to indicate robustness to trend variations and mis-specifications.

We describe the maps resulting from an anisotropic Type II correlation function ($\hat{\delta} \neq 0$, $\hat{\delta} \neq 1$) in Figures 4.6, 4.7, and 4.8.

Figure 4.6 - Four Month Ozone Data

- a) with a linear trend function
- b) with a constant trend function

Figure 4.7 - One Week Ozone Data

- a) with a linear trend based on the Ozone data of the four month (Type II correlation function)
- b) with a linear trend estimated from the one week data
- c) constant trend estimated from the one week data. See Table 4.1.

Figure 4.8 - One Day Ozone Data

- a) with a linear trend based on the Ozone data of the four month (Type II correlation model)
- b) with an estimated linear trend based on the one day's data.

For correlation parameters and trends, see Table 4.1.

The isopleth maps in each data base (four month, a week, a day) are insensitive to the trend changes. The peaks and lows stay in about the same places.

The isopleth lines for four month and one week have a tendency to run from north to south.

It is of some interest to compare Type I and Type II maps. The patterns in 4.3a and 4.6a show some similarities. The same holds for 4.3b and 4.6b. The anisotropy factor straightened out the lines to run somewhat more in the north-south direction, where elsewhere in the isotropic Type I model the isopleth lines run in a slight north-west direction.

Conclusions.

Our personal experience with attempting to find a reasonably smooth function for $\rho(\underline{x},t)$ using daily mean oxidant observations from 1971 to 1973 at 16 stations located in the San Francisco Bay Area (which is meteorologically quite heterogeneous for a small area) has demonstrated the difficulty of this task. It occurs that there are stations which are located near each other which tend to differ greatly in oxidant levels, whereas there are other stations located far apart and which show relatively strong agreement. This phenomenon remains true after a yearly seasonal component is subtracted for each station's time series and even when a weekly seasonal effect is removed. It is possible that the high correlation observed over long distances is an artifact of the influence of weather on oxidant levels, the data being suggestive of this, but the issue has not been sufficiently studied to confirm this. Anisotropic spatial correlation assumption alters the pattern of the maps when compared to the patterns achieved by isotropic correlation models.

The state of the art is at this point very crude and it is impossible to say what approach is going to prove most effective in the determination of $\rho(\underline{h},s)$ from the type of data available in air pollution

work. In particular, it remains to be examined how best to incorporate seasonal and meteorological information into the estimation of $\rho(\underline{h}, s)$ and/or into the reformulation of the spatial model itself. The maps, though, as well as the simulations show strong evidence that the interpolated and extrapolated values are robust under changes in the assumed trend (linear or constant), within different frames of correlation models.

Figure 4.1. Sixteen Monitoring Stations in the San Francisco Bay Area.
112 Days Station Means in Parentheses (log-Oxidant Measurements).

Overall mean: 3.0828

Variance: 0.1574

Interstation Mean Variance:
0.0550

Asymptotic Correlation:
0.3493

Stations

1. Petaluma
2. Napa
3. San Rafael
4. Vallejo
5. Fairfield
6. Richmond
7. San Francisco
8. Oakland
9. Pleasant Hill
10. Pittsburg
11. Burlingame
12. San Leandro
13. Redwood City
14. San Jose
15. Fremont
16. Livermore

Figure 4.2. 1973 Log-Oxidant, Month 6-9 Autocorrelations, Subtracting Constant Mean.

Figure 4.3.a. Four Month Ozone Data $\rho(\underline{X},t) = g(\underline{X}) \cdot h(t)$
 $g(\underline{X})$ Type I. Linear Trend.

Figure 4.3.b. Four Month Ozone Data $\rho(\underline{X}, t) = g(\underline{X}) \cdot h(t)$
 $g(\underline{X})$ Type I. Constant Trend.

Figure 4.4.a. One Week Ozone Data $\rho(\underline{X}, t) = g(\underline{X}) \cdot h(t)$
 $g(\underline{X})$ Type I. Four Month Linear Trend.

Figure 4.4.b. One Week Ozone Data $\rho(\underline{X}, t) = g(\underline{X}) \cdot h(t)$
 $g(\underline{X})$ Type I. Estimated Linear Trend Based on One Week's Data.

Figure 4.4.c. One Week Ozone Data $\rho(\underline{X}, t) = g(\underline{X}) \cdot h(t)$
 $g(\underline{X})$ Type I. Estimated Constant Trend Based on One
 Week's Data.

Figure 4.5.a. One Day Ozone Data $\rho(\underline{X}, t) = g(\underline{X}) \cdot h(t)$.
 $g(\underline{X})$ Type I. Four Month Linear Trend.

Figure 4.5.b. One Day Ozone Data $\rho(\underline{X}, t) = g(\underline{X}) \cdot h(t)$. $g(\underline{X})$ Type I. Estimated Linear Trend Based on One Day's Data.

Figure 4.6.a. Four Month Ozone Data $g(\underline{X})$ Type II. Anisotropic Linear Trend.

Figure 4.6.b. Four Month Ozone Data $g(X)$ Type II.
Anisotropic Constant Trend.

Figure 4.7.a. One Week Ozone Data $g(X)$ Type II.
Anisotropic Linear Trend Based on Four Month Data.

Figure 4.7.b. One Week Ozone Data $g(X)$ Type II. Anisotropic Estimated Linear Trend Based on One Week's Data.

Figure 4.7.c. One Week Ozone Data $g(X)$ Type II. Anisotropic Estimated Constant Trend Based on One Week's Data.

Figure 4.8.a. One Day Ozone Data $z(X)$ Type II.
Anisotropic Linear Trend Based on Four Month Data.

Figure 4.8.b. One Day Ozone Data $g(X)$ Type II.
Anisotropic Estimated Linear Trend Based on One
Day's Data.

Acknowledgements

This work was supported by a grant from SIAM Institute for Mathematics and Society (SIMS) to the Stanford University Department of Statistics funded by the Rockefeller Foundation, the Sloane Foundation, the Department of Energy (US), the Environmental Protection Agency (US), and the National Science Foundation (US).

The authors are indebted to Paul Switzer, Stanford University, for his guidance and help.

References

- [1] P. Delfiner (1976), Linear Estimation of Non-Stationary Spatial Phenomena, Proceedings of NATO A.S.I. "Geostat 75", Rome, Italy. Reidel Publ. Corp., Dordrecht, Netherlands.
- [2] A. S. Goldberger (1962), Best Linear Unbiased Prediction in the Generalized Linear Regression Model, JASA, 57, 369-375.
- [3] D. G. Krige (1951), A Statistical Approach to Some Basic Mine Value Problems in Witwatersrand, Journal of Chemical, Metallurgical and Mining Society of South Africa, 119-139.
- [4] G. Mathéron (1965), Les Variables Regionalisées et Leur Estimation, Paris, Masson et Cie.
- [5] G. Mathéron (1973), The Intrinsic Random Functions and Their Applications, Adv. in Appl. Prob., 5, 439-468.