ON THE PROBLEM OF ESTIMATING THE AUTOCORRELATION
FUNCTION OF SPATIO-TEMPORAL VARIABLES

BY

DAG TJØSTHEIM

TECHNICAL REPORT NO. 6

AUGUST 1977

NOTICE
This report was prepared as an account of work sponsored by the United States Government. Neither the United States nor the United States Department of Energy, 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.

STUDY ON STATISTICS AND ENVIRONMENTAL
FACTORS IN HEALTH

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

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

26

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.

ON THE PROBLEM OF ESTIMATING THE AUTOCORRELATION
FUNCTION OF SPATIO-TEMPORAL VARIABLES

Abstract

Motivated by problems in statistical air pollution research, we propose an estimate for the autocorrelation function for spatio-temporal variables. The estimate is meaningful for non-integer lags and for irregularly positioned data points. The statistical properties of the estimate is studied in a special case. Possible applications to air pollution modelling are briefly discussed.

1. INTRODUCTION

As is well known, observations taken in time can often be effectively analyzed using correlation (second order) properties only. This is the subject of time series analysis. Air pollution variables and related weather and health variables are usually observed both in space and in time. In the present paper we propose a technique for examining and estimating the correlation structure of spatio-temporal variables. The technique represents a generalization of a method outlined in Brillinger [4, pp. 166-169] for the time series case.

The problem of estimating the autocorrelation function for a spatio-temporal variable, an air pollutant say, differs from the ideal time series situation in at least three different respects:

- 1) We have multidimensional measurements.
- 2) The corresponding random field is usually not homogeneous-stationary.
- 3) We often have a sparse and irregularly positioned set of measurements, this being particularly true in space. (In the San Francisco Bay Area for example, depending on the particular air pollutant under investigation, only 16-30 measurement stations are available.)

It should be remarked that the last two points are also to a varying degree relevant in empirical time series analysis, and techniques intended to handle such departures from assumptions have

been proposed (see for example Jones [10], Brillinger [3], Akaike [1] and Box and Jenkins [5, Ch. 4]).

The problem of estimating autocorrelation structures for spatial variables has been considered before. For a theoretical treatment we refer to Brillinger [2] which assumes that observations are available everywhere in a bounded region in space. More oriented towards specific practical applications is the work done by the French school in geomathematics, notably Mathéron [11], Delfiner and Delhomme [9] and Delfiner [8]. Working with spatial variables in geology they have suggested "Kriging" (see Mathéron [11, Ch. 3]) as a viable method for spatial interpolation. The method depends crucially on the autocorrelation function, and one of the main obstacles has been to obtain good estimates for it. In this connection we refer especially to the work done by Delfiner [8]. The apparent lack of effective estimation procedures have lead to criticism of the method, Watson [17].

2. DEFINITION OF ESTIMATE

Let $X(t)$ be a second order stationary time series with mean zero and autocorrelation function $K(t)$ given by $K(t) = E[X(t+s)X(s)]$. Assume that $X(t)$ is observed at times $t = 0, 1, \dots, M$. The conventional estimate $\hat{K}(t)$ of $K(t)$ is given by

$$\hat{K}(t) = \frac{1}{M-|t|} \sum_{s=0}^{M-|t|} X(s)X(s+t) \quad (2.1)$$

This estimate can be extended directly to a regular grid in space-time as

$$\hat{K}(x,t) = \frac{1}{M N_{\Pi}} \sum_{y=0}^{N-|x|} \sum_{s=0}^{M-|t|} F(y+x,s+t)F(y,s) \quad (2.2)$$

where $N_{\Pi} = \prod_{i=1}^m N_i$, $x = (x_1, \dots, x_m)$, $y = (y_1, \dots, y_m)$ and

$N-|x| = (N_1 - |x_1|, \dots, N_m - |x_m|)$. The quantity $\hat{K}(x,t)$ estimates the autocorrelation function $K(x,t)$ of a zero mean homogeneous-stationary field $F(x,t) = F(x_1, \dots, x_m, t)$. From our point of view the estimate $\hat{K}(x,t)$ suffers from two weaknesses. It does not make sense if we have irregularly positioned measurements. Furthermore, in interpolation problems we need an estimate of the autocorrelation function $K(x,t) = E[F(x+y,t+s)F(y,s)]$ for arbitrary non-integer values of (x,t) , but this cannot be obtained from (2.2).

In Kriging these two problems are sought avoided by assuming a certain functional form (usually isotropic) for the autocorrelation function in continuous space and the estimation problem is reduced to an estimation of parameters characterizing this function. There are some difficulties [8] both of mathematical and statistical nature associated with this approach. Also it is unfortunate to restrict oneself to a predefined class of correlation functions, especially since there seems to be no clear-cut procedure for actually determining appropriate classes of functions from the data.

In this paper we will obtain an estimate of $K(x,t)$ without assuming anything about its functional form. The estimate is based

on an alternative form of the time series estimate. For time series satisfying the mixing condition $\sum_t K(t) < \infty$ there exists a spectral density $f(\lambda)$ such that

$$K(t) = \int_{-\pi}^{\pi} \exp\{i \lambda t\} f(\lambda) d\lambda \quad (2.3)$$

This suggests [4, Ch. 5.10] the possibility of estimating $K(t)$ via the spectral density as

$$\hat{K}(t) = \int_{-\pi}^{\pi} \exp\{i \lambda t\} \hat{f}(\lambda) d\lambda \quad (2.4)$$

where $\hat{f}(\lambda)$ is a suitable estimate of $f(\lambda)$. As demonstrated in [4, Ch. 3.6] this method of estimation may even be faster than the conventional method represented by (2.1) if Fast Fourier Transforms are used. The estimate (2.4) generalizes directly to the spatio-temporal case as

$$\hat{K}(x, t) = \int \exp\{i([\mu, x] + \lambda t)\} \hat{f}(\mu, \lambda) d\mu d\lambda \quad (2.5)$$

where $\mu = (\mu_1, \dots, \mu_m)$ is a "wave-number" vector and $[\mu, x] = \sum \mu_i x_i$.

Moreover, this estimate remains meaningful for arbitrary values of (x, t) . An estimate $\hat{f}(\mu, \lambda)$ of $f(\mu, \lambda)$ may be based on the quantity

$$d_F(\mu, \lambda) = \sum_{x_i} \sum_{t_i} F(x_i, t_i) \exp\{-i([\mu, x_i] + \lambda t_i)\} \quad (2.6)$$

which makes sense also for irregularly positioned observations

(x_i, t_i) . As might be expected, the theoretical properties of $\hat{K}(x, t)$ are difficult to establish in the general case. We have been able to do so only if the observations are taken on a grid with a certain proportion of the observations missing.

What we are really interested in is the autocorrelation function $K(x,t)$ of the underlying random field in continuous space-time. Assume that $F(x,t)$ as considered in continuous space-time has a spectral density $g(\mu,\lambda)$ such that

$$K(x,t) = \int_{-\infty}^{\infty} \exp\{i([\mu,x] + \lambda t)\} g(\mu,\lambda) d\mu d\lambda \quad (2.7)$$

Suppose that we sample the field at regular grid locations (x_i, t_i) in space and time and let us use the notation $k_j = x_{j,i} - x_{j,i-1}$ and $h = t_i - t_{i-1}$, $j = 1, \dots, m$ for the grid constants in space and time. Consider the spectral estimate

$$I(\mu,\lambda) = \frac{1}{MN_{\Pi}(2\pi)^{m+1}} \left| \sum_{y,s} F(yk,sh) \exp\{-i([\mu,yk] + \lambda sh)\} \right|^2 \quad (2.8)$$

where $yk = (y_1 k_1, \dots, y_m k_m)$ and $[\mu,yk] = \sum_i \mu_i y_i k_i$. Using standard arguments involving $(m+1)$ -dimensional Dirac delta functions we obtain asymptotically (as N and $M \rightarrow \infty$) that

$$E[I(\mu,\lambda)] \rightarrow \sum_j \sum_t g\left(\mu + \frac{2\pi j}{k}, \lambda + \frac{2\pi t}{h}\right) \stackrel{\Delta}{=} f(\mu,\lambda) \quad (2.9)$$

$$\text{for } -\frac{\pi}{k} \leq \mu \leq \frac{\pi}{k} \quad \text{and} \quad -\frac{\pi}{h} \leq \lambda \leq \frac{\pi}{h}$$

where $j = (j_1, \dots, j_m)$ is a vector of integers and where the summation is over $j_i = 0, \pm 1, \pm \dots$ and $t = 0, \pm 1, \pm \dots$, and where (2.9) serves as a defining equation for $f(\mu,\lambda)$. Unless the field $F(x,t)$ is bandlimited in both μ and λ to $-\pi/k \leq \mu \leq \pi/k$ and $-\pi/h \leq \lambda \leq \pi/h$, the estimate $I(\mu,\lambda)$ has an asymptotic bias $g(\mu,\lambda) - f(\mu,\lambda)$ due to aliasing.

From (2.7) and (2.9) we have that for lags corresponding to grid coordinates

$$K(x_k, t_h) = \int_{-\pi/k}^{\pi/k} \int_{-\pi/h}^{\pi/h} f(\mu, \lambda) \exp\{i([\mu, x_k] + \lambda t_h)\} d\mu d\lambda \quad (2.10)$$

for $x = (x_1, \dots, x_m)$ and t integer valued. This relation means that we will be able to obtain good estimates of K from $I(\mu, \lambda)$ at grid coordinate lags even if the corresponding spectral estimates are biased. In fact, it will be demonstrated in Section 3 that using $I(\mu, \lambda)$ one can obtain an asymptotically unbiased and consistent estimate of the quantity

$$\int_{-\pi/k}^{\pi/k} \int_{-\pi/h}^{\pi/h} f(\mu, \lambda) \exp\{i([\mu, x] + \lambda t)\} d\mu d\lambda \quad (2.11)$$

for arbitrary values x and t . Unfortunately, however, for values of x and t other than grid coordinate lags the expression in (2.11) is no longer equal to $K(x, t)$. Actually, it is not difficult to show that (2.11) is identical to

$$\sum_j \sum_s \exp\{-i([\frac{2\pi j}{k}, x] + \frac{2\pi}{h} st)\} \int_{\frac{(2j-1)\pi}{k}}^{\frac{(2j+1)\pi}{k}} \int_{\frac{(2s-1)\pi}{h}}^{\frac{(2s+1)\pi}{h}} g(\mu, \lambda) \exp\{i([\mu, x] + \lambda t)\} d\mu d\lambda \quad (2.12)$$

giving rise to a bias in the estimation of $K(x, t)$.

3. PROPERTIES OF ESTIMATE IN A SPECIAL CASE

In this section we will suppress the time dependence and write $F(x)$ and $K(x)$ instead of $F(x,t)$ and $K(x,t)$. Moreover, we will assume that the grid constants k_i , $i=1, \dots, m$ are all equal to unity. This simplifies notation and does not imply a loss in generality, since for the theory considered here the time coordinate and the space coordinates are treated in exactly the same way, and we can always transform to a coordinate system where $(k_1, \dots, k_m) = (1, \dots, 1)$. We will consider the case where observations are taken on a regular grid, but where observations are allowed to be missing. More precisely, if N_i is the number of grid locations in the x_i -direction with $N_{\Pi} = \prod_i N_i$ and if $S = S(N)$ is the number of observations actually taken, it will be assumed that

$$N_{\Pi} - S = O(N_{\Pi}^{1-\beta}) \quad (3.1)$$

for $0 < \beta \leq 1$ as $N_i \rightarrow \infty$, $i = 1, \dots, m$. Here $A(N) = O(B(N))$ means that the ratio $A(N)/B(N)$ stays finite as $N_i \rightarrow \infty$.

We will study a general r -dimensional homogeneous vector field $\{F_a(x)\}$ $a=1, \dots, r$. This is done partly because it will make it easier to extend the theory to estimation of cross correlations, but also because consistency of the autocorrelation function estimate results as a special case of the more general theory. Following [4, p. 91] we introduce tapering functions $h_a(x)$, $a=1, \dots, r$ which are bounded, vanish outside the unit cube $I = \{x: 0 \leq x_k \leq 1\}$ and satisfy a Lipschitz condition, i.e.,

$$|h_a(x) - h_a(y)| \leq C_a \sum_{i=1}^m |x_i - y_i| \quad (3.2)$$

for a constant C_a . The condition (3.2) may be weakened to a condition requiring bounded variation in some appropriate sense. (See [12, p. 169] and [7] for possible definitions of bounded variation for multivariable functions.) However, this will not be done here since it leads to technical complications and since the Lipschitz condition is fulfilled for the class of functions of interest to us.

We use the notation $h_a^{(N)}(x)$ for $h_a(x/N)$ where $x/N = (x_1/N_1, \dots, x_m/N_m)$. Moreover, $H_{a_1 \dots a_k}^{(N)}(\mu)$ and $H_{a_1 \dots a_k}(\mu)$ are defined by

$$H_{a_1 \dots a_k}^{(N)}(\mu) = \sum_x \left(\prod_{j=1}^k h_{a_j}(x/N) \right) \exp\{-i[\mu, x]\} \quad (3.3)$$

and

$$H_{a_1 \dots a_k}(\mu) = \int_I \left(\prod_{j=1}^k h_{a_j}(x) \right) \exp\{-i[\mu, x]\} dx \quad (3.4)$$

Finally, the functions $b_a(x)$, $a = 1, \dots, r$ will be used to describe the missing observations. We let $b_a(x) = 1$ if an observation of $F_a(x)$ exists at the grid location x , $b_a(x) = 0$ otherwise. The missing observations cannot be included in the functions $h_a(x)$, since using $h_a^{(N)}(x)$ it would imply that the locations of missing observations vary with N . For ease of notation we assume that (3.1) holds for all a with the same β .

The following lemma corresponds to Lemma P.4.1 of [4].

Lemma 3.1: If the functions $h_a(x)$, $a = 1, \dots, r$ satisfy the above

stated conditions, then there exist finite constants L_1 and L_2 such that

$$\begin{aligned} & \left| \sum_x b_{a_k}(x) h_{a_k}^{(N)}(x) \prod_{i=1}^{k-1} (b_{a_i}(x+y_i) h_{a_i}^{(N)}(x+y_i)) \exp\{-i[\mu, x]\} - H_{a_1 \dots a_k}^{(N)}(\mu) \right| \\ & \leq L_1 \sum_{i=1}^{k-1} |y_i| + L_2 \sum_x \left| b_{a_k}(x) \prod_{i=1}^{k-1} b_{a_i}(x+y_i) - 1 \right| \end{aligned} \quad (3.5)$$

where for $y_i = (y_{i,1}, \dots, y_{i,m})$, $|y_i|$ is defined by $|y_i| = \sum_{j=1}^m |y_{i,j}|$.

Proof: Using (3.3) and the boundedness of the functions h_{a_k} it follows that the expression on the left hand side of the inequality sign in (3.5) is

$$\begin{aligned} & \leq L \sum_x \left| b_{a_k}(x) \prod_{i=1}^{k-1} b_{a_i}(x+y_i) h_{a_i}^{(N)}(x+y_i) - \prod_{i=1}^{k-1} h_{a_i}^{(N)}(x) \right| \\ & = L \sum_x \left| b_{a_k}(x) \prod_{i=1}^{k-1} b_{a_i}(x+y_i) h_{a_i}^{(N)}(x+y_i) - \prod_{i=1}^{k-1} h_{a_i}^{(N)}(x+y_i) \right. \\ & \quad + \prod_{i=1}^{k-1} h_{a_i}^{(N)}(x+y_i) - h_{a_1}^{(N)}(x) \prod_{i=2}^{k-1} h_{a_i}^{(N)}(x+y_i) + \dots \\ & \quad \left. + \left(\prod_{i=1}^{k-2} h_{a_i}^{(N)}(x) \right) \cdot h_{a_{k-1}}^{(N)}(x+y_{k-1}) - \prod_{i=1}^{k-1} h_{a_i}^{(N)}(x) \right| \\ & \leq L \sum_x \left| \prod_{i=1}^{k-1} h_{a_i}^{(N)}(x+y_i) \left[b_{a_k}(x) \prod_{i=1}^{k-1} b_{a_i}(x+y_i) - 1 \right] \right| \\ & \quad + L' \sum_{i=1}^{k-1} \sum_x \left| h_{a_i}^{(N)}(x+y_i) - h_{a_i}^{(N)}(x) \right| \end{aligned}$$

and the proof is now completed using the boundedness of

$$\prod_{i=1}^{k-1} h_{a_i}^{(N)}(x+y_i) \text{ and the condition (3.2) for } h_a(x/N) = h_a^{(N)}(x). \quad ||$$

In analogy with (2.6) let us define the tapered version of $d_F(\mu, \lambda)$ as

$$d_{F,a}^{(N)}(\mu) = \sum_x b_a(x) h_a(x/N) F_a(x) \exp\{-i[\mu, x]\} \quad (3.6)$$

and let us use the notation $\text{cum}\{d_{F,a_1}^{(N)}(\mu_1), \dots, d_{F,a_k}^{(N)}(\mu_k)\}$ for the k th

order joint cumulant of these quantities. Let $K_{a_1 \dots a_k}(x_1, \dots, x_{k-1})$ be the k th order cumulant of $F_{a_j}(x)$, $j=1, \dots, k$ and let

$$\begin{aligned} & f_{a_1 \dots a_k}(\mu_1, \dots, \mu_{k-1}) \\ &= \sum_{x_1} \dots \sum_{x_{k-1}} \exp \left\{ -i \sum_{j=1}^{k-1} [\mu_j, x_j] \right\} K_{a_1 \dots a_k}(x_1, \dots, x_{k-1}) \end{aligned} \quad (3.7)$$

be the corresponding joint cumulant spectrum. We assume that

$$\sum_{x_1} \dots \sum_{x_{k-1}} (1 + |x_j|) |K_{a_1 \dots a_k}(x_1, \dots, x_{k-1})| < \infty \quad (3.8)$$

for $j=1, \dots, k-1$. In the following proof the summation in (3.6) will go from $-N_i$ to N_i for $i=1, \dots, m$.

Lemma 3.2: Assume that conditions (3.1) and (3.8) are satisfied.

Then

$$\begin{aligned} & \text{cum}\{d_{F,a_1}^{(N)}(\mu_1), \dots, d_{F,a_k}^{(N)}(\mu_k)\} \\ &= 2\pi^{m(k-1)} H_{a_1 \dots a_k}^{(N)} \left(\sum_{j=1}^k \mu_j \right) f_{a_1 \dots a_k}(\mu_1, \dots, \mu_{k-1}) + o(N_{\Pi}^{1-\beta}). \end{aligned}$$

Proof: Proceeding as in the proof of Lemma P.4.2 of [4] and using Lemma 3.1, the cumulant can be written

$$\begin{aligned} & (2\pi)^{m(k-1)} H_{a_1 \dots a_k}^{(N)} \left(\sum_{j=1}^k \mu_j \right) \sum_{x_1=-(N_1-1)}^{2(N_1-1)} \dots \sum_{x_{k-1}} \exp \left\{ -i \sum_{j=1}^{k-1} [\mu_j, x_j] \right\} \\ & \cdot K_{a_1 \dots a_k}(x_1, \dots, x_{k-1}) + \varepsilon_N^{(1)} + \varepsilon_N^{(2)} \end{aligned}$$

where

$$|\varepsilon_N^{(1)}| \leq L_1 \sum_{x_1=-(2N_1-1)}^{2(N_1-1)} \dots \sum_{x_{k-1}} \left(\sum_{j=1}^{k-1} |x_j| \right) |K_{a_1 \dots a_k}(x_1, \dots, x_{k-1})|$$

and

$$\begin{aligned} |\varepsilon_N^{(2)}| & \leq L_2 \sum_{y_1=-(2N_1-1)}^{2(N_1-1)} \dots \sum_{y_{k-1}} \sum_x |b_{a_k}(x) \prod_{j=1}^{k-1} b_{a_j}(x+y_j) - 1| \\ & \cdot |K_{a_1 \dots a_k}(y_1, \dots, y_{k-1})| \\ & = O(k(N_{\Pi} - S) \sum_y |K_{a_1 \dots a_k}(y_1, \dots, y_{k-1})|) \end{aligned}$$

and the lemma follows using (3.1), (3.7), and (3.8). $\parallel$

Consider now a single series $F(x)$. Corresponding to (2.8) we define $I^{(N)}(\mu)$ as

$$I^{(N)}(\mu) = \frac{1}{(2\pi)^{m_{N_{\Pi}}}} \left| \sum_y F(y) b(y) \exp\{-i[\mu, y]\} \right|^2 \quad (3.9)$$

Corrollary 3.1: Assume that the covariance function K of F is such that $\sum_x |xK(x)| < \infty$ and that (3.1) is fulfilled. Then as $N_{\Pi} \rightarrow \infty$

$$E\{I^{(N)}(\mu)\} = f(\mu) + o(N_{\Pi}^{-\beta}) \quad (3.10)$$

where $f(\mu)$ is the spectral density of $F(x)$.

Proof: Since $F(x)$ is assumed to have zero mean

$$E[I^{(N)}(\mu)] = (2\pi)^{-m} N_{\Pi}^{-1} \text{cum}\{d_F^{(N)}(\mu), d_F^{(N)}(-\mu)\}.$$

For the tapering function $h(x)$ given by

$$h(x) = \begin{cases} 1 & \text{for } 0 \leq x_k \leq 1 \\ 0 & \text{otherwise} \end{cases}$$

we have $H^{(N)}(0) = N_{\Pi}$ and (3.10) follows from Lemma 3.2. ||

We note that the corollary can be generalized to a spatial series with a non-zero mean; see the corresponding Theorem 5.2.2 in [4]. Now let $A_j(\mu)$ $j=1, \dots, L$ be bounded functions satisfying the Lipschitz condition (3.2) (again this can be relaxed to include bounded variation in some appropriate sense) and following [4, p. 166] define $J(A_j)$ and $J^{(N)}(A_j)$ by

$$J(A_j) = \int A_j(\mu) f(\mu) d\mu \quad (3.11)$$

and

$$J^{(N)}(A_j) = \frac{(2\pi)^m}{N_{\Pi}} \sum_x A_j\left(\frac{2\pi x}{N}\right) I^{(N)}\left(\frac{2\pi x}{N}\right) \quad (3.12)$$

We need a strengthened version of assumption (3.8). Denote by $K_k(x_1, \dots, x_{k-1})$ the k th order cumulant of F . It will be assumed

that cumulants of all orders exist for F and that

$$\sum_{x_1} \dots \sum_{x_{k-1}} (1 + |x_j|) |K_k(x_1, \dots, x_{k-1})| < \infty \quad (3.13)$$

for $j=1, \dots, k-1$ and $k=2, 3, \dots$. The following theorem extends Theorem 5.10.1 of [4] to the case of spatial series with observations missing according to (3.1).

Theorem 3.1: Under the assumptions (3.1) and (3.13) we have

$$\begin{aligned} E[J^{(N)}(A_j)] &= \frac{(2\pi)^m}{N_{\Pi}} \sum_x A_j\left(\frac{2\pi x}{N}\right) f\left(\frac{2\pi x}{N}\right) + O(N_{\Pi}^{-\beta}) \\ &= \int A_j(\mu) f(\mu) d\mu + O(N_{\Pi}^{-\beta}) \end{aligned} \quad (3.14)$$

Furthermore

$$\text{cov}(J^{(N)}(A_j), J^{(N)}(A_k)) = O(N_{\Pi}^{-1}) \quad (3.15)$$

and $J^{(N)}(A_j)$, $j=1, \dots, L$ are asymptotically jointly normal.

Proof: This theorem is proved using Lemma 3.2 and the obvious multidimensional generalizations of the arguments in [4, p. 417]. ||

Considering the special function $A(\mu, \lambda) = \exp\{i([\mu, x] + \lambda t)\}$ it is clear how Theorem 3.1 can be used to obtain an estimate of the quantity defined in (2.11). In fact the desired estimate is given by

$$\hat{K}(x, t) = \frac{(\pi)^{m+1}}{h k_{\Pi} M N_{\Pi}} \sum_{y=-N}^N \sum_{s=-M}^M \exp\{i([\frac{\pi y}{kN}, x] + \frac{\pi s}{hM} t)\} I(\frac{\pi y}{kN}, \frac{\pi s}{hM})$$

where $I(\mu, \lambda)$ is defined as in (3.9). From Theorem 3.1 it follows that

$\hat{K}(x,t)$ gives an unbiased and consistent estimate of the integral in (2.11). This integral, as discussed in Section 2, equals the autocorrelation function $K(x,t)$ for lags corresponding to grid intervals.

4. DISCUSSION

We will discuss in this section some of the potential advantages as well as disadvantages of the estimation method outlined in this paper.

The first objection to the method concerns the homogeneity-stationarity assumption. For example, it is well known that this assumption does not hold for air pollution data, but even in such situations homogeneous-stationary models may be useful for pilot studies from which more realistic models can be obtained (compare [5, Ch. 4]).

The observation points in space very often are not located on a grid even if we allow some of the grid locations to be missing. As an illustration, we refer to Figure 1 which shows a map of air pollution monitoring stations in the San Francisco Bay Area. It is obvious that with a spatial arrangement as in Figure 1, only a very crude approximation to the model studied in Section 3 is obtained. So even though the estimation formulas (2.5) and (2.6) remain meaningful, the statistical properties of the corresponding estimates are uncertain. The situation in time domain is usually better. Most of the pollutants in the Bay Area for example are measured at regular time intervals and the percentage of missing observations is small, so the

Fig. 1. Geographical distribution of 24 Bay Area air pollution monitoring stations.

assumption (3.1) may be quite reasonable here. To be able to evaluate the exact properties of a more realistic model in the spatial domain, we need an estimation theory which allows observations to be off grid points. Possible starting points might be [1] or [3].

In practice although the formulas (2.5) and (2.6) are meaningful in the non-grid case, one still has to determine reasonable wavenumber/frequency integration limits in (2.11). From (2.11) it is seen that for a regular grid these are in fact related to the grid constants k and h . Conceivably integration limits could therefore be determined by measuring typical distances between observation points. Alternatively, Capon [6], one could study the spatial measurement point response function given by

$$h(\mu) = \left| \sum_{x_i} \exp\{-i[\mu, x_i]\} \right|^2 \quad (4.1)$$

(x_i are the coordinates of the i th measurement point) and use a typical distance from the main spectral peak at $\mu = (0.0)$ to major secondary peaks.

Another major disadvantage of the proposed estimation method is that it is computationally very slow for a moderately large net of irregularly positioned measurement points. This is because in this case the Fast Fourier Transformation cannot be used for the evaluation of (2.6).

On the other hand, the proposed estimate should have some advantages. Contrary to what is done in Kriging, it does not assume a functional form for the autocorrelation function and is thus avoiding

the element of arbitrariness involved in choosing a proper class of functions. Its main advantage, however, is its generality. In the same way as Theorem 3.1 represents an extension of Theorem 5.10.1 of [4], we could establish a spatio-temporal estimation procedure for cross correlation functions having properties analogous to those given in Theorem 7.6.1 of [4] in the time series case. We consider this fact to be of some importance since it would render feasible a study of linear relations in space and time between several spatio-temporal variables. One possible application could for example be in relating air pollution variables to health.

Our last proposed application has to do with parametric representation of spatio-temporal series, in particular those series $F(x,t)$ which can be approximately represented by autoregressive schemes

$$\sum_{y,s} a(y,s) F(x-y, t-s) = Z(x,t) \quad (4.3)$$

where $Z(x,t)$ is a spatio-temporal white series, i.e., $E[Z(x,t)Z(y,s)] = k\delta(x-y, t-s)$. Such series are discussed from a theoretical point of view in [14], and in [15] we have given several examples of the actual fitting of such series to simulated data. The coefficients $a(y,s)$ describe the correlation structure of the series and as in the one-dimensional case, [13] and [16], should be useful for feature extraction and information compressing purposes. In particular, collection of coefficients $a(y,s)$ may be useful for recognizing and even forecasting extreme air pollution situations. Although the series (4.2) is defined on a regular grid, this is not

critical when it comes to the estimation of coefficients $a(y,s)$. These are computed [14] from Yule Walker type of equations which require the estimated autocorrelation function only on a grid, and this can be achieved from irregularly positioned measurements using the estimation procedure of Section 2. However, the statistical properties of the estimated coefficients remain to be investigated in the non-grid situation.

ACKNOWLEDGMENTS

I am very grateful to Professor Paul Switzer for the opportunity to visit Stanford as a member of the SIMS group. I am grateful to the Norwegian Research Council for providing additional financial support.

REFERENCES

- [1] Akaike, H. Effect of timing-error on the power spectrum of sampled data. Ann. Inst. Statist. Math. 13, (1960), pp. 145-165.
- [2] Brillinger, D.R. The frequency analysis of relations between stationary spatial series. Proc. Twelfth. Bien. Sem. Canadian Math. Congr., Ed. R. Pyke, pp. 39-81. Can. Math. Congr. Montreal, 1970.
- [3] Brillinger, D.R. The spectral analysis of stationary interval functions. In proc. Seventh Berkeley Symp. Prob. Statist. Eds. L. Lecam, J. Neyman, and E. L. Scott, pp. 483-513. Univ. of California Press, Berkeley, 1973.
- [4] Brillinger, D.R. Time Series Data Analysis and Theory. Holt, New York, 1975.
- [5] Box, G.E.P. and G. M. Jenkins. Time Series Analysis Forecasting and Control. Holden-Day, San Francisco, 1970.
- [6] Capon, J. High-resolution frequency-wavenumber spectrum analysis. Proc. IEEE, 57(1969), pp. 1408-1418.
- [7] Clarkson, J.A. and Adams, C.R. On definitions of bounded variation for functions of two variables, Trans. Amer. Math. Soc., 35(1933), pp. 824-854.
- [8] Delfiner, P. Linear estimation of nonstationary spatial phenomena. Proceedings of NATO A.S.I. "Geostat 75," Rome, Italy. Reidel. Publ. Corp. Dordrecht, Netherlands, 1976.
- [9] Delfiner, P. and J. P. Delhomme. Optimum interpolation by Kriging. Proceedings of NATO A.S.I. "Display and analysis of spatial data," Nottingham, England, 1973. Wiley, London, 1975.
- [10] Jones, R.H. Spectral analysis with regularly missed observations. Ann. Math. Statist., 33(1962), pp. 455-461.
- [11] Mathéron, G. The Theory of Regionalized Variables and Its Applications. Ecole National Supérieure des Mines de Paris. Paris, 1971.
- [12] Saks, S. Theory of the Integral. 2nd rev. ed. Dover, New York, 1964.

- [13] Tjøstheim, D. Recognition of waveforms using autoregressive feature extraction. IEEE Trans. Computers, 26(1977), pp. 268-270.
- [14] Tjøstheim, D. Statistical spatial series modelling. Adv. Appl. Prob., to appear.
- [15] Tjøstheim, D. Autoregressive modelling and spectral estimation in the plane. Some simulation results. Stanford Exploration Project, March 1977 Progress Report, Vol. 11.
- [16] Tjøstheim, D. and E. S. Husebye. An improved discriminant for test ban verification using short and long period spectral parameters. Geophysical Research Letters, 3(1976), pp. 499-502.
- [17] Watson, G.S. Trend surface analysis. Journal of Mathematical Geology, 3(1971), pp. 215-226.