PROBABILITY MODELING AND ESTIMATION FOR HOURLY VARIATION
OF AIR POLLUTION CONCENTRATIONS

CHEN-HUI FOO

TECHNICAL REPORT NO. 31

DECEMBER 1979

PREPARED UNDER THE AUSPICES OF
SIAM INSTITUTE FOR MATHEMATICS AND SOCIETY

DEPARTMENT OF STATISTICS
STANFORD UNIVERSITY
STANFORD, CALIFORNIA

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

DISTRIBUTION OF THIS DOCUMENT IS UNLIMITED

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.

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.

PROBABILITY MODELING AND ESTIMATION FOR HOURLY VARIATION
OF AIR POLLUTION CONCENTRATIONS

by

CHEN-HUI FOO

TECHNICAL REPORT NO. 31

DECEMBER 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

ACKNOWLEDGMENTS

I would like to express my gratitude to Professor Paul Switzer, my dissertation advisor, for his valuable suggestions, encouragement, guidance and assistance, and to Professor Ray Faith for his assistance and patience on the model simulation and time series analysis. I also thank Professors Thomas Sager, Theodore Anderson and Ingram Olkin for their helpful suggestions especially on multivariate analysis and time series analysis. Special thanks to Ms. Judi Davis for her arrangement in the typing of this dissertation and to Ms. Rosemarie Stampfel for her excellent typing.

I also would like to express my appreciation to my father and my husband for their full support and encouragement.

TABLE OF CONTENTS

CHAPTER		PAGE
I	INTRODUCTION AND SUMMARY	1
	1.1. A Model for Hourly Oxidant Concentrations	1
	1.2. Summary	4
II	DERIVATION OF THE MAXIMUM LIKELIHOOD ESTIMATORS OF THE MODEL PARAMETERS	6
	2.1. Identifiability of the Model	6
	2.2. Derivation of the Maximum Likelihood Estimators	7
III	ASYMPTOTIC BEHAVIOR OF THE MAXIMUM LIKELIHOOD ESTIMATORS OF THE MODEL PARAMETERS.	17
	Theorem 3.1.....	17
	Theorem 3.2.....	34
IV	THRESHOLD EXCEEDANCES OF MID-DAY OXIDANT CONCENTRATIONS	41
	4.1. Estimation of the Probability of Thres- hold Exceedances from San Jose Data and from Model Simulations	41
	4.2. Model Sensitivity Analysis	44
	REFERENCES	53

CHAPTER I

INTRODUCTION AND SUMMARY

1.1. A Model for Hourly Oxidant Concentrations

Model building in air pollution research is often complicated by various kinds of correlations. However, the period which most concerns us is the time when the concentration is high. The probability model presented here is not intended to fit the entire time series but rather to describe the behavior of high pollutant concentrations.

The model (1.1) below describes the variation of hourly average concentration, $X_{k,t}$, at the t -th hour on the k -th day during the peak concentration period. After a suitable transformation function f , the fluctuation of $f(X_{k,t})$ is modelled to be the superposition of a daily random effect Z_k and an hourly random effect $u_{k,t}$.

The model is

$$(1.1) \quad Y_{k,t} = f(X_{k,t}) = \mu_t + Z_k + u_{k,t}, \quad 1 \leq k \leq K, 1 \leq t \leq n$$

where f mentioned above is some appropriate function such that $f(X_{k,t})$ has a normal distribution. μ_t is the overall hourly mean at the t -th hour. K is the total number of days observed and n is the number of hours during the peak concentration period. The daily variation, Z_k , is distributed as $N(0, \sigma_D^2)$, while the correlated hourly variations on the k -th day, $u_k = (u_{k,1}, \dots, u_{k,n})'$, are

modelled to have the joint normal distribution with mean vector zero and covariance $\sigma_H^2 R$ where R is the Markov matrix with $R_{ij} = \rho^{|i-j|}$, $1 \leq i, j \leq n$.

Three types of independence are needed in the above model. At first, we assume that Z_k 's are independent variables by neglecting the day-to-day correlation. Secondly, we assume the independence between Z_k and $u_{k',t}$ for all k, k', t . Thirdly, for $k \neq k'$, we assume that $u_{k,t}$ and $u_{k',t'}$ are independent for all t, t' .

If $\sigma_D^2 = 0$, then for fixed k , $\{Y_{k,t} - \mu_t\}$ is an autoregressive process of order 1 with correlation ρ [Box and Jenkins (1970)].

If $\rho = 0$, the model (1.1) becomes a random effects model for a balanced one-way classification [Searle (1971)].

To check whether the model (1.1) can fit the observed phenomenon, we choose the oxidant data at the San Jose Piedmont station between May 1, 1974 and October 31, 1974. It has been observed that high concentrations of oxidant occur almost invariably only in the mid-day period, 11:00 a.m. - 5:00 p.m. (refer to Figure A on the next page). Therefore we have taken $n = 6$. Furthermore, from Larsen's model (1971) the oxidant concentration is lognormally distributed. Hence we take $f(x) = \ln x$ here.

Those days without complete six hourly averages are deleted from our data set. It turned out that there are $K = 172$ days left for the analysis.

There are three reasons to select this special time category. At first, during the summer time, the daytime meteorological conditions are fairly constant, and since pollutant emission sources are also

Figure A. Daily Variation Pattern at San Jose-Piedmont
Data Source: May - October 1974

constant, our day-to-day stationarity assumption may be reasonable. Secondly, since the meteorological conditions make the oxidant concentrations relatively high throughout the summer mid-day period, we can avoid the problems of dealing with very low concentrations derived from truncation error and the undefined value for $\ln 0$.

1.2. Summary

In Chapter II, the maximum likelihood estimates (m.l.e.) were obtained for the parameters μ , ρ , σ_H^2 , η ($= \sigma_D^2/\sigma_H^2$) in the model. There is no closed form solution for $\hat{\rho}$. However, for fixed ρ , we find $\hat{\eta}(\rho)$ and $\hat{\sigma}_H^2(\rho)$ which are the m.l.e.'s in terms of ρ and sample covariances. By using numerical iteration, we can find that $\hat{\rho}$ which maximizes the likelihood function. Just as in the variance components problem [Herbach (1959)], it is possible to get a negative estimate of η . In this case, the constrained m.l.e. is used. In the San Jose oxidant data mentioned above, we encountered the estimate $\hat{\eta} = 0.34 \times 10^{-4}$ which is very close to 0.

In Chapter III, the asymptotic theory of the m.l.e. for this model is discussed. The information matrix $I(\underline{\theta})$ with $\underline{\theta} = (\eta, \sigma_H^2, \rho)$ is obtained. Tables 1 and 2 show the asymptotic standard deviations and the asymptotic correlations of $\hat{\eta}$, $\hat{\sigma}_H^2$, $\hat{\rho}$ with true parameters η , σ_H^2 , ρ in the following range.

$\eta = 0.1, 1, 2$; $\rho = 0.2, 0.5, 0.8, 0.93$; $\sigma_H^2 = 0.22$. However when $\eta \approx 0$, the asymptotic theory is not appropriate. P. A. P. Moran (1971) derived the modified asymptotic theory for the constrained m.l.e. when one of the true parameters is on the boundary of the parameter space.

By using Moran's theory, we find the joint distribution of the m.l.e., conditional on $\hat{\eta} > 0$. The asymptotic means and variances of the m.l.e., conditional on $\hat{\eta} > 0$, when $\eta = 0$, are tabulated in Table 5. Table 5 shows that when ρ is close to 1, the asymptotic variances of $\hat{\eta}$, $\hat{\sigma}_H^2$, increase in the order of $[(1-\rho)^2]^{-1}$, but the asymptotic variance of $\hat{\rho}$ tends to a constant. This phenomenon was shown in the model simulation.

The comparison of the asymptotic means and variances of model parameter estimates using simulation, the modified asymptotic theory and the usual asymptotic theory is shown in Table 6.

Section 4.1 discusses the estimation of the distribution of the number of hourly exceedances of a specified threshold level for given mid-day oxidant concentrations for the San Jose data. The observed distribution is well-predicted by the model (1.1). Three different threshold levels used here are 8.5 pphm, 10.5 pphm and 12.5 pphm.

In Section 4.2, the sensitivity study of the estimated probability function of exactly i hourly exceedances of the threshold level c during the mid-day is done by changing the values of the model parameters in the simulation. Six different patterns are shown in Figure 1.a-3.b. For large hourly correlation ρ , there is a high probability of obtaining zero or six exceedances in the mid-day six hourly readings. For small ρ , there is high probability of having two or three exceedances, but a very small chance of having six exceedances.

The observed cumulative distribution function and probability function from the San Jose data also are shown in Figure 1.a-3.b. The closeness of the observed values to the curve from the model simulation with $\rho = 0.93$, $\hat{\sigma}_H^2 = 0.22$, $\eta = 0$ demonstrates the goodness-of-fit.

CHAPTER II

DERIVATION OF THE MAXIMUM LIKELIHOOD ESTIMATORS OF THE MODEL PARAMETERS

2.1. Identifiability of the Model

The model (1.1) for mid-day hourly oxidant concentrations can be generalized in various ways. For example, it may be desirable to lessen the independence condition among the Z_k 's. However, model (1.1) is not identifiable without restrictions imposed on the correlation matrix (ρ_{ij}) of the u_k . The model is called not identifiable if, for different sets of parameters $\{\sigma_D^2, \sigma_H^2, \rho_{ij}, i, j = 1, 2, \dots, n\}$, it is possible to produce the same covariance matrix for $\underline{Y}_k = (Y_{k,1}, Y_{k,2}, \dots, Y_{k,n})'$ which has the form $((\sigma_D^2 + \sigma_H^2 \rho_{ij}))$. For example, we can choose $1 \geq c > 0$ such that $\min_{i,j} \rho_{ij} \geq 1 - 1/c \geq 0$. Define a new set of $\{\sigma_D'^2, \sigma_H'^2, \rho_{ij}'\}$ as follows,

$$\begin{aligned}\rho_{ij}' &= (1-c) + c\rho_{ij} \\ \sigma_H'^2 &= \frac{1}{c} \sigma_H^2 \\ \sigma_D'^2 &= \sigma_D^2 + (1 - \frac{1}{c})\sigma_H^2\end{aligned}$$

Then

$$\sigma_D^2 + \sigma_H^2 \rho_{ij} = \sigma_D'^2 + \sigma_H'^2 \rho_{ij}'$$

Therefore, $\{\sigma_D^2, \sigma_H^2, \rho_{ij}'\}$ and $\{\sigma_D'^2, \sigma_H'^2, \rho_{ij}'\}$ produce the same covariance matrix of $\underline{Y}_k$.

In the model (1.1), no meteorological parameters such as temperature, wind velocity are explicitly mentioned. We can further modify the model by assuming that the hourly mean μ_t is a function of those meteorological variables, but this will not be undertaken here.

2.2. Derivation of the Maximum Likelihood Estimators

In the model (1.1), $Y_{k,t}$, the average concentration on the k -th day at the t -th hour after normalization, is modelled as

$$Y_{k,t} = \mu_t + Z_k + u_{k,t}, \quad 1 \leq t \leq n, \quad 1 \leq k \leq K$$

where for $k = 1, 2, \dots, K$, Z_k and u_k are independently and identically distributed (i.i.d.) as $N(0, \sigma_D^2)$ and $N(0, \sigma_H^2 R)$ respectively. R has been defined in Section 1.1. It follows that the daily concentrations

$\tilde{Y}_k = (Y_{k,1}, \dots, Y_{k,n})$ are i.i.d. multivariate random variables.

Furthermore, the independence between Z_k and u_k implies that each

$\tilde{Y}_k$ has a multivariate normal distribution with mean $\mu = (\mu_1, \mu_2, \dots, \mu_K)'$ and covariance

$$\Sigma = ((\sigma_D^2 + \sigma_H^2 \rho^{|i-j|})) = \sigma_H^2 R + \sigma_D^2 \mathbf{1}\mathbf{1}',$$

where $\mathbf{1} = (1, 1, \dots, 1)'$, an $n \times 1$ column vector.

The log likelihood function of $\tilde{Y}_1, \tilde{Y}_2, \dots, \tilde{Y}_K$ is

$$(2.1) \quad \mathcal{L}(\mu, \theta', \tilde{Y}_1, \dots, \tilde{Y}_K) \\ = -\frac{nK}{2} \ln 2\pi - \frac{K}{2} \ln |\Sigma| - \frac{1}{2} \sum_{k=1}^K (\tilde{Y}_k - \mu)' \Sigma^{-1} (\tilde{Y}_k - \mu)$$

with $\theta' = (\sigma_D^2, \sigma_H^2, \rho)$. (See Anderson (1958).) The determinant and inverse of Σ can be obtained by using the following two results in Rao (1967), pp. 29.

Result (a). If $B = (b_{ij})$ is any $n \times n$ matrix and $\text{adj } B = (B_{ij})$ where B_{ij} is a cofactor of b_{ij} , then $|B + c\mathbf{1}\mathbf{1}'| = |B| + c\mathbf{1}'(\text{adj } B)\mathbf{1}$ where c is any constant.

Result (b). If B is a nonsingular matrix, and U, V are two column vectors, then

$$(B + UV')^{-1} = B^{-1} - \frac{(B^{-1}U)(V'B^{-1})}{1 + V'B^{-1}U}$$

It was shown in Szatrowksi (1976) that $\det R = (1-\rho^2)^{n-1}$ and $R^{-1} = (1/1-\rho^2)\tilde{R}$ where

$$\tilde{R} = \begin{pmatrix} 1 & -\rho & & & & \\ -\rho & 1+\rho^2 & -\rho & & & \\ & -\rho & 1+\rho^2 & & & \\ & & & \ddots & & \\ 0 & & & & 1+\rho^2 & -\rho \\ & & & & -\rho & 1 \end{pmatrix}$$

Hence using result (a) with $B = \sigma_H^2 R$, we get

$$(2.2) \quad |\Sigma| = (\sigma_H^2)^n (1-\rho^2)^{n-1} \left\{ 1 + \frac{\sigma_D^2}{\sigma_H^2(1-\rho^2)} [2(1-\rho) + (n-2)(1-\rho)^2] \right\}$$

Replacing $\underline{B}$, $\underline{U}$ and $\underline{Y}$ in result (b) by $\sigma_H^2 \underline{R}$, $\sigma_D^2 \underline{1}$ and $\sigma_D^2 \underline{1}$ respectively, we get

$$(2.3) \quad \underline{\Sigma}^{-1} = \frac{1}{\sigma_H^2} \left\{ \frac{\tilde{\underline{R}}}{1 - \rho^2} - \frac{(\sigma_D^2/\sigma_H^2) \cdot \tilde{\underline{R}}}{(1+\rho)[(1+\rho) + (\sigma_D^2/\sigma_H^2)(n-(n-2)\rho)]} \right\}$$

where

$$\tilde{\underline{R}} = \begin{pmatrix} 1 & 1-\rho & 1-\rho & \dots & 1-\rho & 1 \\ 1-\rho & (1-\rho)^2 & (1-\rho)^2 & \dots & (1-\rho)^2 & 1-\rho \\ \vdots & \vdots & & & \vdots & \vdots \\ 1-\rho & (1-\rho)^2 & & \dots & (1-\rho)^2 & 1-\rho \\ 1 & (1-\rho) & & \dots & 1-\rho & 1 \end{pmatrix}$$

substituting (2.2), (2.3) into (2.1) and $\text{tr } \underline{B} = \sum_{i=1}^n b_{ii}$, we get

$$(2.4) \quad \mathcal{L}(\underline{\mu}, \underline{\theta}', \underline{y}_1, \dots, \underline{y}_K) \\ = -\frac{nK}{2} \ln 2\pi - \frac{K}{2} \left[n \ln \sigma_H^2 + (n-1) \ln(1-\rho^2) + \ln \left(1 + \frac{\sigma_D^2}{\sigma_H^2} \frac{n-(n-2)\rho}{1+\rho} \right) \right] \\ - \frac{1}{2\sigma_H^2} \text{tr} \left\{ \left[\sum_{k=1}^K (\underline{y}_k - \underline{\mu})(\underline{y}_k - \underline{\mu})' \right] \left[\frac{\tilde{\underline{R}}}{1-\rho^2} - \frac{(\sigma_D^2/\sigma_H^2) \tilde{\underline{R}}}{(1+\rho)[(1+\rho) + (\sigma_D^2/\sigma_H^2)(n-(n-2)\rho)]} \right] \right\}$$

Let $\underline{A} = (1/K) \sum_{k=1}^K (\underline{y}_k - \underline{\mu})(\underline{y}_k - \underline{\mu})' = (a_{t,t'})$ where

$a_{t,t'} = (1/K) \sum_{k=1}^K (y_{k,t} - \mu_t)(y_{k,t'} - \mu_{t'})$. Since σ_D^2 always appears together with σ_H^2 in the form of σ_D^2/σ_H^2 in (2.4), we reparametrize the parameter

space $\Omega' = \{\theta' | \sigma_D^2 \geq 0, \sigma_H^2 \geq 0, -1 \leq \rho \leq 1\}$ into

$$\Omega = \{\theta = (\eta, \sigma_H^2, \rho) | \eta \geq 0, \sigma_H^2 \geq 0, -1 \leq \rho \leq 1\} \quad \text{with } \eta = \frac{\sigma_D^2}{\sigma_H^2}$$

in order to make the following calculations simpler.

Using the new parameter $\eta = \sigma_D^2 / \sigma_H^2$ and the fact that

$\text{tr } \underline{BC} = \text{tr } \underline{CB}$, we have

$$(2.5) \quad \mathcal{L}(\underline{\mu}, \underline{\theta}; \underline{Y}_1, \dots, \underline{Y}_K) \\ = - \frac{nK}{2} \ln 2\pi - \frac{K}{2} \left[n \ln \sigma_H^2 + (n-1) \ln(1-\rho^2) + \ln \left(1 + \eta \frac{n-(n-2)\rho}{1+\rho} \right) \right] \\ - \frac{K}{2\sigma_H^2} \text{tr} \left[\frac{\tilde{\underline{R}}}{1-\rho^2} \cdot \underline{\tilde{A}} - \frac{\eta \tilde{\underline{R}}}{(1+\rho)[(1+\rho) + \eta(n-(n-2)\rho)]} \cdot \underline{\tilde{A}} \right]$$

Expression (2.5) can be further expressed in the following form

$$(2.6) \quad \mathcal{L}(\underline{\mu}, \underline{\theta}; \underline{Y}_1, \dots, \underline{Y}_K) \\ = - \frac{nK}{2} \ln 2\pi - \frac{K}{2} \left[n \ln \sigma_H^2 + (n-1) \ln(1-\rho^2) + \ln \left(1 + \eta \frac{n-(n-2)\rho}{1+\rho} \right) \right] \\ - \frac{K}{2\sigma_H^2} \left\{ \frac{1}{1-\rho^2} \left[\sum_{t=1}^n a_{t,t} - 2\rho \sum_{t=1}^{n-1} a_{t,t+1} + \rho^2 \sum_{t=2}^{n-1} a_{t,t} \right] \right. \\ \left. - \frac{\eta \left[a_{1,1} + 2a_{1,n} + a_{n,n} + 2(1-\rho) \left(\sum_{t=2}^{n-1} (a_{1,t} + a_{n,t}) \right) + (1-\rho)^2 \sum_{2 \leq t, t' \leq n-1} a_{t,t'} \right]}{(1+\rho)[(1+\rho) + \eta(n-(n-2)\rho)]} \right\}$$

Differentiating $\mathcal{L}(\underline{\mu}, \underline{\theta}; \underline{Y}_1, \underline{Y}_2, \dots, \underline{Y}_K)$ in (2.4) with respect to $\underline{\mu}'$ and setting it to zero, we have the maximum likelihood estimate (m.l.e.) $\hat{\underline{\mu}}$ of $\underline{\mu}$, $\hat{\underline{\mu}} = \bar{\underline{Y}} = (\bar{Y}_{.,1} + \bar{Y}_{.,2} + \dots + \bar{Y}_{.,n})'$ where $\bar{Y}_{.,t} = (1/K) \sum_{k=1}^K Y_{k,t}$.

To find the m.l.e.'s of σ_H^2 and η , we differentiate $(\underline{\mu}, \underline{\theta}; \underline{Y}_1, \dots, \underline{Y}_K)$ in (2.6) with respect to η and σ_H^2 respectively and set their derivatives to zero. We obtain

$$(2.7) \quad \hat{\sigma}_H^2(\rho) = \frac{1}{n} \left\{ \frac{1}{1-\rho^2} \left[\sum_{t=1}^n c_{t,t} - 2\rho \sum_{t=1}^{n-1} c_{t,t+1} + \rho^2 \sum_{t=2}^{n-1} c_{t,t} \right] \right. \\ \left. - \frac{\hat{\eta}(\rho) [c_{1,1} + 2c_{1,n} + c_{n,n} + 2(1-\rho) \left(\sum_{t=2}^{n-1} (c_{1,t} + c_{n,t}) \right) + (1-\rho)^2 \sum_{2 \leq t, t' \leq n-1} c_{t,t'}]}{(1+\rho)[(1+\rho) + \hat{\eta}(\rho)(n-(n-2)\rho)]} \right\}$$

$$(2.8) \quad \hat{\sigma}_H^2(\rho) \equiv \frac{[c_{1,1} + 2c_{1,n} + c_{n,n} + 2(1-\rho) \left(\sum_{t=2}^{n-1} (c_{1,t} + c_{n,t}) \right) + (1-\rho)^2 \sum_{2 \leq t, t' \leq n-1} c_{t,t'}]}{[n-(n-2)\rho][(1+\rho) + \hat{\eta}(\rho)(n-(n-2)\rho)]}$$

where $c_{t,t'} = (1/K) \sum_{k=1}^K (Y_{k,t} - \bar{Y}_{.,t})(Y_{k,t'} - \bar{Y}_{.,t'})$ which is the same as $a_{t,t'}$ with μ_t being replaced by $\hat{\mu}_t = \bar{Y}_{.,t}$.

For simplicity, we define

$$(2.9) \quad G_1(\rho) = \sum_{t=1}^n c_{t,t} - 2\rho \sum_{t=1}^{n-1} c_{t,t+1} + \rho^2 \sum_{t=2}^{n-1} c_{t,t}$$

$$(2.10) \quad G_2(\rho) = c_{1,1} + 2c_{1,n} + c_{n,n} + 2(1-\rho) \left[\sum_{t=2}^{n-1} (c_{1,t} + c_{n,t}) \right] \\ + (1-\rho)^2 \sum_{2 \leq t, t' \leq n-1} c_{t,t'}$$

Then we simplify (3.7) and (3.8) into

$$(2.11) \quad \hat{\sigma}_H^2(\rho) = \frac{1}{n} \left\{ \frac{G_1(\rho)}{1-\rho^2} - \frac{\hat{\eta}(\rho) G_2(\rho)}{(1+\rho)[(1+\rho) + \hat{\eta}(\rho)(n-(n-2)\rho)]} \right\}$$

$$(2.12) \quad \hat{\sigma}_H^2(\rho) = \frac{G_2(\rho)}{[n-(n-2)\rho][(1+\rho) + \hat{\eta}(\rho)(n-(n-2)\rho)]}$$

Equating (2.11) and (2.12), we have

$$(2.13) \quad \hat{\eta}(\rho) = \frac{nG_2(\rho)(1-\rho^2) - G_1(\rho)(1+\rho)(n-(n-2)\rho)}{(n-(n-2)\rho)[(n-(n-2)\rho)G_1(\rho) - (1-\rho)G_2(\rho)]}$$

$$(2.14) \quad \hat{\sigma}_D^2(\rho) = \hat{\sigma}_H^2(\rho) \hat{\eta}(\rho)$$

Both $\hat{\eta}(\rho)$ and $\hat{\sigma}_H^2(\rho)$ are functions of ρ and $c_{t,t'}$'s. The derivative with respect to ρ is very complicated so that there is no closed form solution for $\hat{\rho}$. However numerical iteration can be used to find that $\hat{\rho}$ which maximizes $\mathcal{L}(\hat{\eta}(\rho), \hat{\sigma}_H^2(\rho), \hat{\mu}', \underline{Y}_1, \dots, \underline{Y}_K)$. Then the m.l.e. of σ_H^2 , η are as follows.

$$\hat{\sigma}_H^2 = \hat{\sigma}_H^2(\hat{\rho}), \quad \hat{\eta} = \hat{\eta}(\hat{\rho}).$$

For $\rho = 1$ expression (2.13) gives $\hat{\eta}(1) = -1$. Since $\hat{\eta}(\rho)$ is a rational function of ρ , it is a continuous function of ρ . We can find a neighborhood of $\rho = 1$, i.e., $(1-\epsilon, 1)$ for some $\epsilon > 0$, such that for all $\rho \in (1-\epsilon, 1)$, $\hat{\eta}(\rho) < 0$. Therefore, if the true hourly correlation ρ is close to 1, we will tend to get a negative estimate $\hat{\eta}$ which is outside the parameter space

$$\Omega = \{(\eta, \sigma_H^2, \rho, \mu') | \eta \geq 0, \sigma_H^2 \geq 0, -1 \leq \rho \leq 1\}.$$

The necessary and sufficient condition such that $\hat{\eta} > 0$ is

$$\frac{1 - \hat{\rho}}{n - (n-2)\hat{\rho}} G_2(\hat{\rho}) < G_1(\hat{\rho}) < \frac{n(1-\hat{\rho})}{n - (n-2)\hat{\rho}} G_2(\hat{\rho}).$$

Since $\hat{\eta} > 0$ does not always hold, we use the modified estimate $\hat{\eta}^+ = \max(\hat{\eta}, 0)$.

By using the following lemma, we can show that $\hat{\eta} \geq 0 \implies \hat{\sigma}_H^2 \geq 0$.

LEMMA (2.1). $G_2(\rho) \geq 0 \forall \rho \in (-1, 1)$.

Proof. From (2.5), (2.6), (2.10), $G_2(\rho)$ can be expressed as

$$G_2(\rho) = \text{tr}(\tilde{R} \cdot \tilde{A}) = \frac{1}{K} \text{tr}[(1+\rho)^2 (R^{-1} \tilde{1})(\tilde{1}' R^{-1}) (\sum_{k=1}^K \tilde{z}_k \tilde{z}_k')]]$$

where $\tilde{z}_k = Y_k - \bar{Y}$. Then

$$G_2(\rho) = \frac{1}{K} \sum_{k=1}^K (1+\rho)^2 [(\tilde{z}_k' R^{-1} \tilde{1})' (\tilde{z}_k R^{-1} \tilde{1}')] \geq 0 \text{ for all } \rho \in (-1, 1).$$

From (2.12) and Lemma (2.1), $\hat{\eta} \geq 0$ implies $\hat{\sigma}_H^2 \geq 0$. Q.E.D.

When $\rho = 0$, the model becomes a component of variance problem for the balanced one-way classification. The m.l.e. of σ_H^2 and η can be expressed as follows.

$$\hat{\eta} = \frac{\sum_{t=1}^n \sum_{t'=1}^n c_{t,t'} - \sum_{t=1}^n c_{t,t}}{n \sum_{t=1}^n c_{t,t} - \sum_{t=1}^n \sum_{t'=1}^n c_{t,t'}}$$

$$\hat{\sigma}_H^2 = \frac{1}{n} \left(\sum_{t=1}^n c_{t,t} - \frac{\hat{\eta} \sum_{t=1}^n \sum_{t'=1}^n c_{t,t'}}{1 + n\hat{\eta}} \right) - \frac{1}{n(n-1)} \left(n \sum_{t=1}^n c_{t,t} - \sum_{t=1}^n \sum_{t'=1}^n c_{t,t'} \right)$$

$$\hat{\sigma}_D^2 = \hat{\sigma}_H^2 \hat{\eta} = \frac{1}{n(n-1)} \left(\sum_{t=1}^n \sum_{t'=1}^n c_{t,t'} - \sum_{t=1}^n c_{t,t} \right)$$

From the definition of $c_{t,t'}$, we express $\hat{\sigma}_D^2, \hat{\sigma}_H^2$ in terms of the $Y_{k,t}$'s and the $\bar{Y}_{.,t}$'s

$$\hat{\sigma}_H^2 = \frac{1}{K(n-1)} \sum_{k=1}^K \sum_{t=1}^n (Y_{k,t} - \bar{Y}_{.,t} - \bar{Y}_{k,.} + \bar{Y}_{..})^2$$

which is equal to $SS(E)/K(n-1)$ in variance components problem [Herbach (1959)].

$$\hat{\sigma}_D^2 = \hat{\sigma}_H^2 \hat{\eta} = \frac{1}{n} \left[\frac{\sum_{k=1}^K (\bar{Y}_{k..} - \bar{Y}_{...})^2}{K} - \frac{\sum_{k=1}^K \sum_{t=1}^n (Y_{k,t} - \bar{Y}_{k..} - \bar{Y}_{t..} + \bar{Y}_{...})^2}{K(n-1)} \right]$$

which is equal to

$$\frac{1}{n} \left[\frac{SS(A)}{K} - \frac{SS(E)}{K(n-1)} \right]$$

in variance components problem [Herbach (1959)].

When $\eta = 0$, model (1.1) describes the daily pollutant concentration $Y_{k,t}$ as autoregressive process of order 1. Therefore (2.6) becomes

$$(2.15) \quad \mathcal{L}(\mu, \theta; Y_1, Y_2, \dots, Y_K)$$

$$\begin{aligned} &= -\frac{nK}{2} \ln 2\pi - \frac{K}{2} [n \ln \sigma_H^2(1-\rho^2) - \ln(1-\rho^2)] \\ &\quad - \frac{K}{2\sigma_H^2(1-\rho^2)} \left[\sum_{t=1}^n a_{t,t} - 2\rho \sum_{t=1}^{n-1} a_{t,t+1} + \rho^2 \sum_{t=2}^{n-1} a_{t,t} \right] \end{aligned}$$

In Box and Jenkins (1970) Appendix A7.5, the approximate maximum likelihood estimates of ρ and σ_H^2 can be obtained by first reparametrizing $\sigma_H^2(1-\rho^2)$ as one parameter σ^2 and then using

$$D_{t,t'} = \frac{c_{t,t'} + c_{t+1,t'+1} + \dots + c_{n+1-t',n+1-t}}{n - (t-1) - (t'-1)}$$

as an estimate for $\text{cov}(Y_{k,t}, Y_{k,t'}) \forall k, t, t'$. Therefore

$$\hat{\rho} = \frac{D_{1,2}}{D_{2,2}} = \frac{n-2}{n-1} \frac{\sum_{t=1}^{n-1} c_{t,t+1}}{\sum_{t=2}^{n-2} c_{t,t}},$$

$$\hat{\sigma}^2 = \frac{1}{n} G_1(\hat{\rho})$$

and

$$\hat{\sigma}_H^2 = \frac{G_1(\hat{\rho})}{n(1-\hat{\rho}^2)}.$$

$\hat{\rho}$ is different from the least squares estimate of ρ which is $\frac{\sum_{t=1}^{n-1} c_{t,t+1}}{\sum_{t=2}^{n-2} c_{t,t}}$ by a constant $(n-2)/(n-1)$.

Anderson (1978) discussed the repeated measurements on normal autoregressive processes similar to the model (1.1) with $\eta = 0$, i.e. the homogeneous autoregressive process of order 1. However, Anderson developed maximum likelihood estimates only for a non-stationary case where the marginal normal distribution of the first-hour concentration does not depend on σ_H^2 and ρ . Anderson (1978, pp. 373) mentioned that for the stationary case, the maximum likelihood estimators of ρ and σ_H^2 are much more complicated. This is the case we have developed and explains why we used numerical iteration and Box-Jenkins used approximation to get the maximum likelihood estimators.

CHAPTER III

ASYMPTOTIC BEHAVIOR OF THE MAXIMUM LIKELIHOOD ESTIMATORS OF THE MODEL PARAMETERS

The maximum likelihood estimators $\hat{\eta}$, $\hat{\rho}$, $\hat{\sigma}_H^2$ derived in Chapter 3 are functions of the sample covariances $c_{t,t'}$'s with

$$c_{t,t'} = \frac{1}{K} \sum_{k=1}^K (Y_{k,t} - \bar{Y}_{\cdot,t})(Y_{k,t'} - \bar{Y}_{\cdot,t'})$$

and the m.l.e. $\hat{\mu} = \bar{Y}_{\cdot}$.

The usual asymptotic theory of maximum likelihood estimates of $\beta = (\mu', \theta')$ is as follows.

Theorem 3.1 [Cox and Hinkley (1974)]

Let β be a vector of parameters in m -dimensional space and $\hat{\beta}$ be the maximum likelihood estimate of β . β_0 is the vector of true parameter. K is the sample size. If the following regularity conditions hold,

- (i) the parameter space Ω of β has finite dimension m , is closed and compact, and the true β_0 is an interior point in Ω .
- (ii) The probability density function $f_Y(\beta; Y)$ is one-to-one, i.e. $f_Y(\beta; Y) \neq f_Y(\beta'; Y)$ if $\beta \neq \beta'$ where Y is one observation.

(iii) $\exists N(\beta_0)$: a neighborhood of β_0 in Ω such that

$$\forall \beta' \in N(\beta_0)$$

$$\left. \frac{\partial \mathcal{L}(\beta; Y)}{\partial \beta_i} \right|_{\beta=\beta'}, \quad \left. \frac{\partial^2 \mathcal{L}(\beta; Y)}{\partial \beta_i \partial \beta_j} \right|_{\beta=\beta'}, \quad \text{and} \quad \left. \frac{\partial^3 \mathcal{L}(\beta; Y)}{\partial \beta_i \partial \beta_j \partial \beta_k} \right|_{\beta=\beta'}$$

exist a.s. for all $i, j, k \in \{1, 2, \dots, m\}$. Also

$$\frac{1}{K} \left| \left. \frac{\partial^3 \mathcal{L}(\beta; Y)}{\partial \beta_i \partial \beta_j \partial \beta_k} \right|_{\beta=\beta'} \right| \leq \text{some integrable function } z \text{ for all } i, j, k.$$

(iv) The equality

$$E \left\{ - \frac{\partial^2 \mathcal{L}(\beta; Y)}{\partial \beta_i \partial \beta_j} \right\} = E \left(\frac{\partial \mathcal{L}(\beta; Y)}{\partial \beta_i} \right) \left(\frac{\partial \mathcal{L}(\beta; Y)}{\partial \beta_j} \right)$$

holds for all i, j .

(v) The information matrix $I(\rho)$ defined as the $m \times m$ matrix with (i, j) -element $E\{-(\partial^2 \mathcal{L}(\beta; Y)/(\partial \beta_i \partial \beta_j))\}$ is finite and positive definite, $\forall \beta \in N(\beta_0)$.

Then the limiting distribution of $K^{1/2} (\hat{\beta} - \beta_0)$ is $N(0, I^{-1}(\beta_0))$.

In the model (1.1), the distribution of Y_k has a multivariate normal distribution with mean μ and covariance Σ . (μ', θ) is a 1×9 vector, so the information matrix $I(\mu', \theta)$ is a 9×9 matrix.

From (2.1) with $K = 1$, we have

$$\mathcal{L}(\mu, \theta; Y_1) = -\frac{n}{2} \ln 2\pi - \frac{1}{2} \ln |\Sigma| - \frac{1}{2} (Y_1 - \mu)' \Sigma^{-1} (Y_1 - \mu).$$

Since

$$-\frac{\partial^2 \mathcal{L}(\mu, \theta, \underline{y}_1)}{\partial \eta \partial \mu'} = -\left(\frac{\partial}{\partial \eta} \Sigma^{-1}\right)(\underline{y}_1 - \mu),$$

then

$$E \left[-\frac{\partial^2 \mathcal{L}(\mu, \theta, \underline{y}_1)}{\partial \eta \partial \mu'} \right] = \underline{0},$$

where $\mu' = (\mu_1, \mu_2, \dots, \mu_6)$. Similarly

$$E \left[-\frac{\partial^2 \mathcal{L}(\mu, \theta, \underline{y}_1)}{\partial \sigma_H^2 \partial \mu'} \right] = E \left[\frac{\partial^2 \mathcal{L}(\mu, \theta, \underline{y}_1)}{\partial \rho \partial \mu'} \right] = \underline{0}.$$

Hence the information matrix

$$(3.1) \quad I(\mu', \theta) = \begin{bmatrix} \overbrace{D_1(\theta)}^3 & \overbrace{\underline{0}}^6 \\ \underline{0} & D_2(\mu') \end{bmatrix} \begin{matrix} \} 3 \\ \} 6 \end{matrix} \quad \text{with } D_1(\theta), D_2(\mu'),$$

information matrices of $\theta = (\eta, \sigma_H^2, \rho)$ and $\mu' = (\mu_1, \dots, \mu_6)$, respectively. Furthermore,

$$\underline{D}_2 = E \left(\left[\frac{\partial \mathcal{L}(\mu, \theta, \underline{y}_1)}{\partial \mu'} \right] \left[\frac{\partial \mathcal{L}(\mu, \theta, \underline{y}_1)}{\partial \mu'} \right] \right) = \Sigma.$$

The inverse of $I(\mu', \theta)$ is

$$\begin{bmatrix} D_1^{-1}(\theta) & 0 \\ 0 & \Sigma^{-1} \end{bmatrix}.$$

To derive the information matrix $D_1(\theta)$ of $\theta = (\eta, \sigma_H^2, \rho)$, we substitute $K = 1$ into (2.6). We have

$$(3.2) \quad \mathcal{L}(\mu, \theta; \mathbf{y}_1)$$

$$= -\frac{n}{2} \ln 2\pi - \frac{1}{2} \left[n \ln \sigma_H^2 + (n-1) \ln(1-\rho^2) + \ln \left(1 + \eta \frac{n-(n-2)\rho}{(1+\rho)} \right) \right]$$

$$- \frac{1}{2\sigma_H^2} \left\{ \frac{1}{1-\rho^2} \left[\sum_{t=1}^n a_{t,t}^1 - 2\rho \sum_{t=1}^{n-1} a_{t,t+1}^1 + \rho^2 \sum_{t=2}^{n-1} a_{t,t}^1 \right] \right.$$

$$\left. \eta \left[a_{1,1}^1 + 2a_{1,n}^1 + a_{n,n}^1 + 2(1-\rho) \left(\sum_{t=2}^{n-1} (a_{1,t}^1 + a_{n,t}^1) \right) + (1-\rho)^2 \sum_{2 \leq t, t' \leq n-1} a_{t,t'}^1 \right] \right\}$$

$$- \frac{1}{(1+\rho) [(1+\rho) + \eta(n - (n-2)\rho)]}$$

where $a_{t,t'}^1 = (y_{1,t} - \mu_t)(y_{1,t'} - \mu_{t'})$ and $Ea_{t,t'}^1 = \sigma_D^2 + \rho |t-t'| \sigma_H^2$.

In order to simplify the derivation of the second partial derivatives of $\mathcal{L}(\mu, \theta; \mathbf{y}_1)$ we define

$$G_1^1 = \sum_{t=1}^n a_{t,t}^1 - 2\rho \sum_{t=1}^{n-1} a_{t,t+1}^1 + \rho^2 \sum_{t=2}^{n-1} a_{t,t}^1$$

and

$$G_2^1 = a_{1,1}^1 + 2a_{1,n}^1 + a_{n,n}^1 + 2(1-\rho) \left(\sum_{t=2}^{n-1} (a_{1,t}^1 + a_{n,t}^1) \right)$$

$$+ (1-\rho)^2 \sum_{2 \leq t, t' \leq n-1} a_{t,t'}^1.$$

Therefore the six second partial derivatives and their expectations are as follows.

$$\frac{\partial^2 \mathcal{L}(\mu, \theta; \underline{y}_1)}{\partial \eta^2} = \frac{1}{2} \frac{[n - (n-2)\rho]^2}{[(1+\rho) + \eta(n-(n-2)\rho)]^2} - \frac{G_2^1[n - (n-2)\rho]}{\sigma_H^2[(1+\rho) + \eta(n-(n-2)\rho)]^3}$$

$$E \left[- \frac{\partial^2 \mathcal{L}(\mu, \theta; \underline{y}_1)}{\partial \eta^2} \right] = \frac{[n - (n-2)\rho]^2}{2[(1+\rho) + \eta(n-(n-2)\rho)]^2}$$

$$\frac{\partial^2 \mathcal{L}(\mu, \theta; \underline{y}_1)}{\partial (\sigma_H^2)^2} = \frac{n}{2(\sigma_H^2)^2} - \frac{1}{(\sigma_H^2)^3} \left\{ \frac{G_1^1}{1-\rho^2} - \frac{\eta G_2^1}{(1+\rho)[(1+\rho) + \eta(n-(n-2)\rho)]} \right\}$$

$$E \left[- \frac{\partial^2 \mathcal{L}(\mu, \theta; \underline{y}_1)}{\partial (\sigma_H^2)^2} \right] = - \frac{n}{2(\sigma_H^2)^2} + \frac{1}{(\sigma_H^2)^2} \left\{ \frac{n(1-\rho^2) + \eta(1-\rho)(n-(n-2)\rho)}{1-\rho^2} - \frac{\eta(n-(n-2)\rho)}{1+\rho} \right\}$$

$$= \frac{n}{2(\sigma_H^2)^2}$$

$$\frac{\partial^2 \mathcal{L}(\mu, \theta; \underline{y}_1)}{(\partial \sigma_H^2)(\partial \rho)} = \frac{1}{2(\sigma_H^2)^2} \left\{ \frac{2\rho G_1^1}{(1-\rho^2)^2} + \frac{-2 \sum_{t=1}^{n-1} c_{t,t+1}^1 + 2\rho \sum_{t=2}^{n-1} c_{t,t}^1}{1-\rho^2} \right.$$

$$- \frac{\eta[-2 \sum_{t=1}^{n-1} (c_{1,t}^1 + c_{n,t}^1) - 2(1-\rho) \sum_{2 \leq t, t' \leq n-1} c_{t,t'}^1]}{(1+\rho)[(1+\rho) + \eta(n-(n-2)\rho)]}$$

$$+ \left. \frac{2\eta G_2^1[(1+\rho) + \eta(1 - (n-2)\rho)]}{(1+\rho)^2 [(1+\rho) + \eta(n-(n-2)\rho)]^2} \right\}$$

$$E \left[- \frac{\partial^2 \mathcal{L}(\mu, \theta; \underline{y}_1)}{(\partial \sigma_H^2)(\partial \rho)} \right] = - \frac{(n-1)}{\sigma_H^2} \left\{ \frac{\eta(1-\rho) + \rho[(1+\rho) + \eta(n-(n-2)\rho)]}{(1-\rho)^2 [(1+\rho) + \eta(n-(n-2)\rho)]} \right\}$$

$$\frac{\partial^2 \mathcal{L}(\mu, \theta; \underline{Y}_1)}{\partial \eta \partial \rho}$$

$$= -\frac{1}{2} \frac{[-(n-2)][(1+\rho) + (n-(n-2)\rho)\eta] - (n-(n-2)\rho)(1-(n-2)\eta)}{[(1+\rho) + \eta(n-(n-2)\rho)]^2}$$

$$+ \frac{1}{2\sigma_H^2} \frac{[-2(\sum_{t=2}^{n-1}(c_{1,t}^1 + c_{n,t}^1)) - 2(1-\rho) \sum_{2 \leq t, t' \leq n-1} c_{t,t'}^1]}{[(1+\rho) + \eta(n-(n-2)\rho)]^2}$$

$$E \left[-\frac{\partial^2 \mathcal{L}(\mu, \theta; \underline{Y}_1)}{\partial \eta \partial \rho} \right] = \frac{n-1}{[(1+\rho) + \eta(n-(n-2)\rho)]^2}$$

$$\frac{\partial^2 \mathcal{L}(\mu, \theta; \underline{Y}_1)}{\partial \sigma_H^2 \partial \eta} = \frac{1}{2(\sigma_H^2)^2} \left\{ \frac{-G_2^1}{[(1+\rho) + \eta(n-(n-2)\rho)]^2} \right\}$$

$$E \left[-\frac{\partial^2 \mathcal{L}(\mu, \theta; \underline{Y}_1)}{\partial \sigma_H^2 \partial \eta} \right] = \frac{n - (n-2)\rho}{2(\sigma_H^2) [(1+\rho) + \eta(n-(n-2)\rho)]}$$

$$\frac{\partial^2 \mathcal{L}(\mu, \theta; \underline{Y}_1)}{\partial \rho^2} = \frac{(n-1)(1+\rho^2)}{(1-\rho^2)^2} + \frac{[1 - (n-2)\eta]^2}{2[(1+\rho) + \eta(n-(n-2)\rho)]^2} - \frac{1}{2(1+\rho)^2}$$

$$- \frac{1}{\sigma_H^2} \left[\frac{(1+3\rho^2)G_1^1}{(1-\rho^2)^3} + \frac{\sum_{t=2}^{n-1} c_{t,t}^1}{1-\rho^2} + \frac{4\rho(-\sum_{t=1}^{n-1} c_{t,t+1}^1 + \rho \sum_{t=2}^{n-1} c_{t,t}^1)}{(1-\rho^2)^2} \right]$$

$$+ \frac{\eta}{\sigma_H^2} \left\{ \frac{\sum_{2 \leq t, t' \leq n-1} c_{t,t'}^1}{(1+\rho)[(1+\rho)+\eta(n-(n-2)\rho)]} + \frac{4[(1+\rho) + \eta(1-(n-2)\rho)]}{(1+\rho)^2[(1+\rho)+\eta(n-(n-2)\rho)]^2} \right\}$$

$$\begin{aligned}
& \times \left[\sum_{t=2}^{n-1} (c_{1,t}^1 + c_{n,t}^1) + (1-\rho) \sum_{2 \leq t, t' \leq n-1} c_{t,t'}^1 \right] \\
& - \frac{G_2^1 [1 - (n-2)\eta]}{(1+\rho)^2 [(1+\rho) + \eta(n-(n-2)\rho)]^2} + \frac{4G_1^2 [(1+\rho) + \eta(1-(n-2)\rho)]^2}{(1+\rho)^3 [(1+\rho) + \eta(n-(n-2)\rho)]^3} \Big\} \\
& E \left[- \frac{\partial^2 \mathcal{L}(\underline{\mu}, \underline{\theta}; \underline{Y}_1)}{\partial \rho^2} \right] = - \frac{(n-1)(1+\rho^2)}{(1-\rho^2)^2} - \frac{1}{2} \frac{[1 - (n-2)\eta]^2}{[(1+\rho) + \eta(n-(n-2)\rho)]^2} + \frac{1}{2(1+\rho)^2} \Big\} \\
& + \left\{ \frac{1 + 3\rho^2}{(1-\rho^2)^3} [n(1-\rho^2) + \eta(1-\rho)(n-(n-2)\rho)] + \frac{(n-2)(1+\eta)}{1-\rho^2} + \frac{4\rho[(-\rho+\eta)-\eta(n-(n-2)\rho)]}{(1-\rho^2)^2} \right\} \\
& - \eta \left\{ \frac{(n-2)^2\eta + (n-2) + [2(n-3)\rho]/(1-\rho) - 2(\rho^2-\rho^{n-1})/(1-\rho)^2}{(1+\rho) [(1+\rho) + \eta(n-(n-2)\rho)]} \right. \\
& + \frac{4(n-2)[(1+\rho) + \eta(1-(n-2)\rho)]}{(1+\rho)^2 [(1+\rho) + \eta(n-(n-2)\rho)]} - \frac{[1 - (n-2)\eta][n - (n-2)\rho]}{(1+\rho)^2 [(1+\rho) + \eta(n-(n-2)\rho)]} \\
& \left. + \frac{4[n-(n-2)\rho] [(1+\rho) + \eta(1-(n-2)\rho)]^2}{(1+\rho)^3 [(1+\rho) + \eta(n-(n-2)\rho)]^2} \right\} \\
& = \{2(n-1)^2\eta^2(1-\rho)^2 + 2\eta[(1+\rho)(1-\rho^{n-1}) - 2(n-1)(1-\rho)][(1+\rho) + \eta(n-(n-2)\rho)] \\
& + (n-1)(1+\rho^2)[(1+\rho) + \eta(n-(n-2)\rho)]^2\} / (1-\rho^2)^2 [(1+\rho) + \eta(n-(n-2)\rho)]^2
\end{aligned}$$

Therefore, we obtain the information matrix $D_1(\underline{\theta})$ of $\underline{\theta} = (\eta, \sigma_H^2, \rho)$

which is

$$\begin{bmatrix} \frac{1}{2} \frac{[n-(n-2)\rho]^2}{F^2} & \frac{n-(n-2)\rho}{2(\sigma_H^2)F} & \frac{n-1}{F^2} \\ \frac{n-(n-2)\rho}{2(\sigma_H^2)F} & \frac{n}{2(\sigma_H^2)^2} & -\frac{(n-1)[\eta(1-\rho) + \rho F]}{\sigma_H^2(1-\rho^2)F} \\ \frac{n-1}{F^2} & -\frac{(n-1)[\eta(1-\rho)+\rho F]}{\sigma_H^2(1-\rho^2)F} & d_{33} \end{bmatrix}$$

where $F = (1 + \rho) + \eta[n-(n-2)\rho]$ and

$$d_{33} = \frac{2(n-1)^2 \eta^2 (1-\rho)^2 + 2\eta[(1+\rho)(1-\rho^{n-1}) - 2(n-1)(1-\rho)] F + (n-1)(1+\rho^2)F^2}{(1-\rho^2)^2 \cdot F^2}$$

Since for all k , the distribution of $\tilde{Y}_k$ has a multivariate normal distribution $N_n(\mu, \Sigma)$ and the parameter space Ω has dimension 9, is closed and compact for any bounded subset, hence (i) and (ii) are satisfied. The log likelihood function $\mathcal{L}(\mu, \theta; \tilde{Y}_k)$ in (3.2) has continuous first, second and third derivatives w.r.t. μ and θ at any interior point of Ω . Since the dimension of Ω is finite, we can find an integrable function z dominating all third partial derivatives of $\mathcal{L}(\mu, \theta; \tilde{Y})$. Therefore (iii) and (iv) are satisfied. The information matrix $I(\mu', \theta)$ with $D_1(\theta)$ given above and $D_2(\mu')$ mentioned before is finite and positive definite for any interior point (μ', θ) in the parameter space. Hence the regularity conditions (i)-(v) in Theorem 3.1 are satisfied.

The asymptotic covariance matrix of $\hat{\theta} = (\hat{\eta}, \hat{\sigma}_H^2, \hat{\rho})$ is $(1/K)D_1^{-1}(\theta)$. The limiting distribution of $\sqrt{K}(\hat{\theta} - \theta)$ is $N(0, D_1^{-1}(\theta))$. $D_1^{-1}(\theta)$

can be computed for any interior point $\underline{\theta}$ in the parameter space $\Omega = \{\underline{\theta} | \eta \geq 0, \sigma_H^2 \geq 0, -1 \leq \rho \leq 1\}$. Let d_1^{ij} indicate the (i,j) -element in $D_1^{-1}(\underline{\theta})$. Then $(1/K)d_1^{11}$, $(1/K)d_1^{22}$, $(1/K)d_1^{33}$ are the asymptotic variances of $\hat{\eta}$, $\hat{\sigma}_H^2$, $\hat{\rho}$, respectively. The off-diagonal elements in D_1^{-1} are K times the asymptotic covariances of $\hat{\underline{\theta}} = (\hat{\eta}, \hat{\sigma}_H^2, \hat{\rho})$.

Tables 1 and 2 show the asymptotic standard deviations (Asym.S.D.) and asymptotic correlations (Asym. corr.) of $\hat{\eta}$, $\hat{\sigma}_H^2$, $\hat{\rho}$ respectively as η and ρ vary in the interior of Ω with $\sigma_H^2 = 0.22$, $n = 6$, $K = 200$ fixed.

From Table 1, the asymptotic standard deviations of $\hat{\eta}$ and $\hat{\sigma}_H^2$ have larger row-wise variations than columnwise variations. Table 2 shows that as $\rho \rightarrow 1$, the absolute values of asymptotic correlations of $\hat{\eta}$, $\hat{\rho}$, $\hat{\sigma}_H^2$ are close to 1. Hence model (1.1) becomes unidentifiable.

The data chosen to fit model (1.1) were mentioned in Section 1.1. 172 days with complete 6 hourly averages of oxidant concentrations from 11:00 a.m. to 5:00 p.m. were selected from the Piedmont station in San Jose during the summer 1974. The m.l.e.'s of η , σ_H^2 , ρ and $\underline{\mu}$ for this oxidant data are as follows.

$$\begin{aligned} \hat{\eta} &= 0.34 \times 10^{-4}, & \hat{\sigma}_H^2 &= 0.22, & \hat{\rho} &= 0.93 \\ \hat{\underline{\mu}} &= (1.82, 1.97, 2.06, 2.04, 1.93, 1.70)' \end{aligned}$$

Since P. A. P. Moran (1971) proved that the constrained m.l.e. $(\hat{\underline{\mu}}, \hat{\underline{\theta}})$ is consistent for $(\underline{\mu}, \underline{\theta})$, the above $(\hat{\underline{\mu}}, \hat{\underline{\theta}})$ is a consistent

TABLE 1

THE ASYMPTOTIC STANDARD DEVIATIONS OF $\hat{\eta}$, $\hat{\sigma}_H^2$, $\hat{\rho}$ IN THE FORM

$$\begin{pmatrix} \text{asym.S.D. of } \hat{\eta} \\ \text{asym.S.D. of } \hat{\sigma}_H^2 \\ \text{asym.S.D. of } \hat{\rho} \end{pmatrix} \quad \text{with } \hat{\sigma}_H^2 = 0.22, n = 6, K = 200 \text{ fixed.}$$

	$\rho = 0.2$	$\rho = 0.5$	$\rho = 0.8$	$\rho = 0.93$
$\eta = 0.1$	$\begin{pmatrix} 0.04 \\ 0.01 \\ 0.04 \end{pmatrix}$	$\begin{pmatrix} 0.08 \\ 0.02 \\ 0.04 \end{pmatrix}$	$\begin{pmatrix} 0.23 \\ 0.05 \\ 0.04 \end{pmatrix}$	$\begin{pmatrix} 0.67 \\ 0.14 \\ 0.04 \end{pmatrix}$
$\eta = 1$	$\begin{pmatrix} 0.14 \\ 0.01 \\ 0.04 \end{pmatrix}$	$\begin{pmatrix} 0.20 \\ 0.02 \\ 0.04 \end{pmatrix}$	$\begin{pmatrix} 0.45 \\ 0.05 \\ 0.04 \end{pmatrix}$	$\begin{pmatrix} 1.26 \\ 0.14 \\ 0.04 \end{pmatrix}$
$\eta = 2$	$\begin{pmatrix} 0.26 \\ 0.01 \\ 0.04 \end{pmatrix}$	$\begin{pmatrix} 0.33 \\ 0.02 \\ 0.04 \end{pmatrix}$	$\begin{pmatrix} 0.71 \\ 0.05 \\ 0.04 \end{pmatrix}$	$\begin{pmatrix} 1.91 \\ 0.14 \\ 0.04 \end{pmatrix}$

TABLE 2

THE ASYMPTOTIC CORRELATION OF $\hat{\eta}$, $\hat{\sigma}_H^2$, $\hat{\rho}$ IN THE FORM

$$\begin{pmatrix} \text{asym. corr. } (\hat{\eta}, \hat{\sigma}_H^2) \\ \text{asym. corr. } (\hat{\eta}, \hat{\rho}) \\ \text{asym. corr. } (\hat{\rho}, \hat{\sigma}_H^2) \end{pmatrix} \text{ with } \sigma_H^2 = 0.22, n = 6, K = 200 \text{ fixed.}$$

	$\rho = 0.2$	$\rho = 0.5$	$\rho = 0.8$	$\rho = 0.93$
$\eta = 0.1$	$\begin{pmatrix} -0.60 \\ -0.57 \\ 0.56 \end{pmatrix}$	$\begin{pmatrix} -0.78 \\ -0.76 \\ 0.85 \end{pmatrix}$	$\begin{pmatrix} -0.94 \\ -0.93 \\ 0.98 \end{pmatrix}$	$\begin{pmatrix} -0.99 \\ -0.99 \\ 0.997 \end{pmatrix}$
$\eta = 1$	$\begin{pmatrix} -0.52 \\ -0.37 \\ 0.57 \end{pmatrix}$	$\begin{pmatrix} -0.72 \\ -0.65 \\ 0.86 \end{pmatrix}$	$\begin{pmatrix} -0.93 \\ -0.92 \\ 0.98 \end{pmatrix}$	$\begin{pmatrix} -0.99 \\ -0.99 \\ 0.997 \end{pmatrix}$
$\eta = 2$	$\begin{pmatrix} -0.50 \\ -0.33 \\ 0.57 \end{pmatrix}$	$\begin{pmatrix} -0.70 \\ -0.62 \\ 0.86 \end{pmatrix}$	$\begin{pmatrix} -0.93 \\ -0.91 \\ 0.98 \end{pmatrix}$	$\begin{pmatrix} -0.99 \\ -0.99 \\ 0.997 \end{pmatrix}$

estimator for (μ, θ) . "Constrained" here is in the sense that $(\hat{\mu}, \hat{\theta})$ is not allowed to be outside the boundary of the parameter space. Therefore when the sample size K is large enough, $\hat{\theta}$ is in the small neighborhood of θ . In the case when $\hat{\eta} \approx 0$, either the true parameter η is zero, or the model (1.1) has to be modified.

In order to judge whether the model (1.1) is reasonable and the asymptotic theory approximates well the sampling distribution of the m.l.e.'s of the model parameters, we simulated 200 lognormally distributed samples of pseudo oxidant concentrations using a random generator. Each sample consists of 200 days in which 6 hourly averages were taken. The pseudo daily observation $\tilde{Y}_k = (Y_{k,1}, Y_{k,2}, \dots, Y_{k,6})'$ follows model (1.1),

$$Y_{k,t} = \mu_t + Z_k + u_{k,t}, \quad k = 1, 2, \dots, 200, \quad t = 1, 2, \dots, 6$$

using model parameter values equal to the m.l.e.'s derived from the actual data.

For each simulated sample, we obtain $\hat{\mu}$ and the Pearson correlation matrix $\hat{\Sigma}$ of the log transformed values for that sample. From the 200 simulated samples, we obtain the approximate sampling distribution of $\hat{\mu}$ and $\hat{\Sigma}$. Table 3 shows the m.l.e. of the hourly mean values μ from the observed data in San Jose as well as the quantiles from the above approximate sampling distribution.

To compare the empirical estimate of the correlation matrix $\mathbf{r} = (r_{ij})$ from the actual data with the m.l.e. of that matrix from the model (1.1), i.e., $\hat{r}_{ij} = (\hat{\eta} + \hat{\rho}^{|i-j|}) / (1 + \hat{\eta})$, the numerical values

TABLE 3

THE m.l.e. $\hat{\mu}$ FROM THE OBSERVED DATA AND THE
 QUANTILES OF THE SAMPLING DISTRIBUTION OF $\hat{\mu}$

Quantiles of the Sampling Distribution of $\hat{\mu}$						
	observed	0%	25%	50%	75%	100%
$\hat{\mu}_1$	1.82	1.72	1.79	1.82	1.84	1.89
$\hat{\mu}_2$	1.97	1.86	1.95	1.97	1.99	2.05
$\hat{\mu}_3$	2.06	1.95	2.04	2.06	2.08	2.13
$\hat{\mu}_4$	2.04	1.91	2.02	2.04	2.06	2.11
$\hat{\mu}_5$	1.93	1.83	1.90	1.93	1.95	1.99
$\hat{\mu}_6$	1.70	1.60	1.68	1.70	1.73	1.77

of r_{ij} 's with $i < j$ are shown in the columns 1 and 2 of Table 4. The means and standard deviations of r_{ij} 's from the model simulations mentioned in the previous page with $\hat{\eta} = 0.34 \times 10^{-4}$, $\sigma_H^2 = 0.22$, $\hat{\rho} = 0.93$ and $\hat{\mu} = (1.82, 1.97, 2.06, 2.04, 1.93, 1.70)'$ are also shown in the columns 3 and 4 of Table 4.

Therefore Tables 3 and 4 indicate the consistency between the observed data and the simulated data with the model parameters replaced by the m.l.e. from the observed data. Hence model (1.1) is appropriate to describe the behaviour of the oxidant concentrations up to second moments even though one of the parameter estimates, i.e. $\hat{\eta}$ in our case is close to the boundary of the parameter space.

Since $\hat{\eta} \approx 0$, the Box-Jenkins approximate m.l.e.'s of $\hat{\rho}$ and σ_H^2 for the autoregressive model of order 1 [see the last paragraph in Section 2.2] were also calculated giving estimates of 0.94 and 0.22 respectively. The naive least-squared estimate of the hour-to-hour correlation parameter, ρ , would be 1.17.

The asymptotic behavior of the Box-Jenkins approximate m.l.e.'s mentioned above was discussed in Box and Jenkins (1970) Appendix A7.5. It is proved that for large n , the off-diagonal elements of the information matrix are 0 and it will follow that the asymptotic variance of the Box-Jenkins estimate of ρ in our case is approximately $(1/K) \cdot (1-\hat{\rho})^2/n$ which is consistent with the result proved in Anderson (1971), Theorem 5.5.6. The asymptotic variance of the Box-Jenkins estimate of σ_H^2 is approximately $(1/K) \cdot [2(\hat{\sigma}_H^2)^2]/n$. Table 5 given below shows the calculated m.l.e. and Box-Jenkins approximate estimates of ρ and σ_H^2 as

TABLE 4

THE COMPARISON OF THE ESTIMATES OF HOUR-TO-HOUR CORRELATIONS

	Empirical estimate from actual data	m.l.e. from actual data using model (1.1) $r_{ij} = \frac{\hat{\eta} + \hat{\rho}^{ i-j }}{1 + \hat{\eta}}$	200 samples from model simulations	
			mean	standard deviation
r_{12}	0.93	0.93	0.93	0.01
r_{13}	0.83	0.86	0.86	0.02
r_{14}	0.77	0.80	0.80	0.02
r_{15}	0.71	0.75	0.74	0.03
r_{16}	0.69	0.69	0.69	0.04
r_{23}	0.92	0.93	0.93	0.01
r_{24}	0.86	0.86	0.86	0.02
r_{25}	0.77	0.80	0.80	0.03
r_{26}	0.72	0.75	0.74	0.03
r_{34}	0.94	0.93	0.93	0.01
r_{35}	0.86	0.86	0.86	0.02
r_{36}	0.78	0.80	0.80	0.03
r_{45}	0.94	0.93	0.93	0.01
r_{46}	0.83	0.86	0.86	0.02
r_{56}	0.92	0.93	0.93	0.01

TABLE 5

COMPARISON BETWEEN THE m.l.e. AND THE BOX-JENKINS

APPROXIMATE ESTIMATES OF ρ AND σ_H^2

	$\hat{\rho}$	$\hat{\sigma}_H^2$
m.l.e. calculated from data	0.93	0.22
mean of simulated m.l.e. distribution	0.91	0.19
Box-Jenkins estimate calculated from data	0.94	0.24
Asymptotic S.D. of the m.l.e. by using Theorem 3.1	0.04	0.13
S.D. of simulated m.l.e. distribution	0.025	0.04
Asymptotic S.D. of Box-Jenkins estimate	0.01	0.01

well as their asymptotic standard deviations (S.D.) for comparison. Moreover, the simulated m.l.e. distributions of $\hat{\rho}$ and $\hat{\sigma}_H^2$ were obtained from the simulated 200 samples generated before by calculating $\hat{\rho}$, $\hat{\sigma}_H^2$ in each sample. The means and standard deviations of those simulated m.l.e. distributions also are shown in Table 5.

There is a noticeable discrepancy between the asymptotic S.D. of the m.l.e. calculated using Theorem 3.1 and the S.D. of the simulated m.l.e. distribution of $\hat{\rho}$ and $\hat{\sigma}_H^2$ in Table 5. In the 200 simulated samples, the estimate $\hat{\eta}$ is positive in 99 samples and negative in 101 samples, i.e. about 50% of the time we get a negative estimate of η . For the samples with $\hat{\eta} < 0$, we set $\hat{\eta} = 0$ and adjust $\hat{\sigma}_H^2$ and $\hat{\rho}$ accordingly. [See (2.12) and the derivation of $\hat{\rho}$.] However the asymptotic theory stated in Theorem 3.1 does not make this adjustment and still asserts that the marginal limiting distribution of $\sqrt{K} \hat{\eta}$ is normally distributed with mean 0 and variance

$$\frac{2(1+\rho)^2}{(n-1)(n-2)(1-\rho)^2}$$

which is $d(1,1;0)$ in (3.4). Hence this results in larger standard deviations in the theoretical calculation than those in the model simulation.

P. A. P. Moran (1971) was aware of this problem and derived the following asymptotic theory for the constrained maximum likelihood estimates when one of the parameters is on the boundary.

Theorem 3.2.

Suppose that the likelihood function satisfies the assumptions 1-8 [Moran (1971)]. If the true $\theta_1 = 0$, $\theta_i > 0$ for $i = 2, 3, \dots, p$, the distribution function $P[\sqrt{K}(\hat{\theta}_K - \underline{\theta}) < \underline{t}; \underline{\theta}]$, where $\hat{\theta}_K$ is the constrained m.l.e. of $\underline{\theta} = (\theta_1, \theta_2, \dots, \theta_p)$ in the sample of size K , converges, as $K \rightarrow \infty$, uniformly in $\underline{t}$ and $\underline{\theta}$, to the mixture of the two distributions, i.e.

$$P[\sqrt{K}(\hat{\theta}_K - \underline{\theta}) < \underline{t}; \underline{\theta}] = \frac{1}{2} F_1(\underline{t}; \underline{\theta}) + \frac{1}{2} F_2(\underline{t}; \underline{\theta})$$

F_1 is the limiting distribution function of $\sqrt{K}(\hat{\theta}_K - \underline{\theta})$, conditional on $\hat{\theta}_1 > 0$ and is a p -dimensional multivariate distribution defined on the region

$$t_1 > 0, \quad -\infty < t_i < \infty, \quad i = 2, 3, \dots, p$$

and having in this region a probability density equal to twice the density of a multivariate normal distribution with mean zero and covariance matrix equal to $I^{-1}(\underline{\theta}^0)$ with $\underline{\theta}^0 = (0, \theta_2, \dots, \theta_p)$, the usual asymptotic covariance matrix in the unconstrained case. F_2 , the other part of the mixture, is the limiting distribution of $\sqrt{K}(\hat{\theta}_K - \underline{\theta})$ conditional on $\hat{\theta}_1 = 0$ whose expression is in Moran (1971).

Therefore in our case when $\theta_1 = \eta = 0$, and $\hat{\underline{\theta}} = (\hat{\eta}, \hat{\sigma}_H^2, \hat{\rho})$ is the constrained m.l.e. of $\underline{\theta}$ in Ω , the joint distribution of $\sqrt{K}(\hat{\underline{\theta}} - \underline{\theta})$ conditional on $\hat{\eta} > 0$, converges, uniformly in $\underline{\theta}$ in any bounded subset of Ω , to a tri-variate distribution function whose

density is twice the multivariate normal density with mean zero and covariance $(D_1(\varrho^0))^{-1}$ mentioned before with $\varrho^0 = (0, \sigma_H^2, \rho)$.

Since $D_1(\varrho)$ is continuous at $\varrho = \hat{\varrho}$ and if K is large enough, by consistency $D_1(\hat{\varrho})$ can be used to approximate $D_1(\varrho)$ when the true ϱ is unknown. [See Cox and Hinkley (1974), pp. 302.] Furthermore, $D_1(\hat{\varrho}) = D_1((0.34 \times 10^{-4}, 0.22, 0.93))$ can be approximated by $D_1(\varrho^0)$ with $\varrho^0 = (0, 0.22, 0.93)$ using the continuity property of $D_1(\varrho)$. The calculated $D_1(\varrho^0)$ has the following form (3.3).

$$(3.3) \quad D_1(\varrho^0) = ((d_{ij}^0))$$

$$= \begin{bmatrix} \frac{1}{2} \cdot \frac{[n-(n-2)\rho]^2}{(1+\rho)^2} & \frac{n-(n-2)\rho}{2(\sigma_H^2)(1+\rho)} & \frac{n-1}{(1+\rho)^2} \\ \frac{n-(n-2)\rho}{2(\sigma_H^2)(1+\rho)} & \frac{n}{2(\sigma_H^2)^2} & -\frac{(n-1)\rho}{(\sigma_H^2)(1-\rho^2)} \\ \frac{n-1}{(1+\rho)^2} & -\frac{(n-1)\rho}{(\sigma_H^2)(1-\rho^2)} & \frac{(n-1)(1+\rho^2)}{(1-\rho^2)^2} \end{bmatrix}$$

The determinant $|D_1(\varrho^0)|$ is

$$\frac{(n-1)^2}{(\sigma_H^2)^2 (1+\rho)^4} \cdot \frac{(n-1)[n-(n-2)\rho^2]}{4}$$

which is positive. Hence $(D_1(\varrho^0))^{-1}$ exists and its (i,j) element can be denoted by $d(i,j;0)$. The elements of $(D_1(\varrho^0))^{-1}$ can be explicitly expressed as follows.

$$\begin{aligned}
(3.4) \quad d(1,1;0) &= \frac{2(1+\rho)^2}{(n-1)(n-2)(1-\rho)^2} \\
d(1,2;0) &= - \frac{2(\sigma_H^2)(1+\rho)^2}{(n-2)(n-1)(1-\rho)^2} = d(2,1;0) \\
d(1,3;0) &= - \frac{2(1+\rho)^2}{(n-1)(n-2)(1-\rho)} = d(3,1;0) \\
d(2,2;0) &= \frac{4(\sigma_H^2)^2 \left\{ \frac{[n(n-2)\rho]^2(1+\rho^2)}{2(1-\rho)^2} - n + 1 \right\}}{(n-1)(n-2)[n-(n-2)\rho^2]} \\
d(2,3;0) &= \frac{2[n-(n-2)\rho](1+\rho)(\sigma_H^2)[1 + (\rho/(1-\rho))(n-(n-2)\rho)]}{(n-1)(n-2)[n-(n-2)\rho^2]} \\
&= d(3,2;0) \\
d(3,3;0) &= \frac{(1+\rho)^2 [n-(n-2)\rho]^2}{(n-1)(n-2)[n-(n-2)\rho^2]}
\end{aligned}$$

The asymptotic means and asymptotic variances of $\hat{\eta}|\hat{\eta} > 0$, $\hat{\sigma}_H^2|\hat{\eta} > 0$ and $\hat{\rho}|\hat{\eta} > 0$ can be calculated as follows by using Theorem 3.2.

$$E[\hat{\eta}|\hat{\eta} > 0] = \sqrt{\frac{2}{\pi}} \sigma_1 \quad \text{and} \quad \text{Var}[\hat{\eta}|\hat{\eta} > 0] = \sigma_1^2 \left(1 - \frac{2}{\pi}\right)$$

with $\sigma_1^2 = d(1,1;0)$. Also

$$\begin{pmatrix} \hat{\sigma}_H^2 \\ \hat{\eta} \end{pmatrix} = \begin{pmatrix} \sigma_H^2 \\ 0 \end{pmatrix}$$

is asymptotically distributed as $N_2(0, \Sigma_1)$ with $\Sigma_1 = (b_{ij})$ and $b_{ij} = \sigma_1 \sigma_2 \rho_{12}^{|i-j|}$, $1 \leq i, j \leq 2$ in which $\sigma_2^2 = d(2, 2, 0)$ and $\rho_{12} = d(1, 2, 0) / [d(1, 1, 0) \cdot d(2, 2, 0)]^{1/2}$. Then the conditional mean and variance of $\hat{\sigma}_H^2$ conditioned on $\hat{\eta}$ is

$$E[\hat{\sigma}_H^2 | \hat{\eta}] = \sigma_H^2 + (\sigma_2 / \sigma_1) \rho_{12} \hat{\eta}$$

and

$$E[(\hat{\sigma}_H^2 - E[\hat{\sigma}_H^2 | \hat{\eta}])^2 | \hat{\eta}] = \sigma_2^2 (1 - \rho_{12}^2)$$

[see Anderson (1958)]. Then

$$E[\hat{\sigma}_H^2 | \hat{\eta} > 0] = \int_0^\infty E[\hat{\sigma}_H^2 | \hat{\eta}] f_{\hat{\eta} | \hat{\eta} > 0}(\hat{\eta}) d\hat{\eta} = \sigma_H^2 + \sigma_2 \rho_{12} \sqrt{\frac{2}{\pi}}$$

Moreover, the conditional mean squared error is

$$\begin{aligned} E[(\hat{\sigma}_H^2 - \sigma_H^2)^2 | \hat{\eta} > 0] &= \int_0^\infty E[(\hat{\sigma}_H^2 - \sigma_H^2)^2 | \hat{\eta}] f_{\hat{\eta} | \hat{\eta} > 0}(\hat{\eta}) d\hat{\eta} \\ &= \sigma_2^2 (1 - \rho_{12}^2) + (\sigma_2^2 / \sigma_1^2) \rho_{12}^2 E[\hat{\eta}^2 | \hat{\eta} > 0] \\ &= \sigma_2^2 \end{aligned}$$

Therefore,

$$\begin{aligned} \text{Var}[\hat{\sigma}_H^2 | \hat{\eta} > 0] &= E[(\hat{\sigma}_H^2 - \sigma_H^2)^2 | \hat{\eta} > 0] - (E[\hat{\sigma}_H^2 - \sigma_H^2 | \hat{\eta} > 0])^2 \\ &= \sigma_2^2 (1 - \rho_{12}^2 \cdot \frac{2}{\pi}) \end{aligned}$$

Similarly, we can calculate $E[\hat{\rho}|\hat{\eta} > 0]$ and $\text{Var}[\hat{\rho}|\hat{\eta} > 0]$ by using the above method with σ_2^2 and ρ_{12} replaced by $\sigma_3^2 = d(3,3;0)$ and $\rho_{13} = d(1,3;0)/[d(1,1;0) \cdot d(3,3;0)]^{1/2}$ respectively. We have $E[\hat{\rho}|\hat{\eta} > 0] = \rho + \sigma_3 \rho_{13} (\sqrt{2/\pi})$ and $\text{var}[\hat{\rho}|\hat{\eta} > 0] = \sigma_3^2 (1 - \rho_{13}^2 \cdot 2/\pi)$. Those results are tabulated in Table 6.

Therefore, from Table 6, we note that when ρ is close to 1, the asymptotic variances of $\hat{\eta}$ and $\hat{\sigma}_H^2$ increase in the order of $(1/(1-\rho)^2)$, while the asymptotic variance of $\hat{\rho}$ tends to a constant as $\rho \rightarrow 1$.

In order to see how well the asymptotic means and variances in Table 6 work in the finite sample situation, we use the 200 simulated samples generated before in which each sample consists of 200 daily vectors of six hourly observations with the probability structure as in the model (1.1). The actual m.l.e.'s $\hat{\theta} = (0.34 \times 10^{-4}, 0.22, 0.93)$ and $\mu = \hat{\mu}$ were used as the true parameter values in the model simulations. For each sample, we obtain an m.l.e. $\hat{\theta} = (\hat{\eta}, \hat{\sigma}_{II}^2, \hat{\rho})$. Therefore we have the approximate sampling distribution of $\hat{\theta}$ from 200 simulated samples. As mentioned before, 99 out of the 200 samples, $\hat{\eta}$ is positive. The mean and standard deviation (S.D.) of $\hat{\theta}$ in this subsample are calculated and tabulated in Table 7. The numerical calculation of the means and standard deviations of the m.l.e. $\hat{\theta}$ conditional on $\hat{\eta} > 0$ in Table 6 by using Moran's asymptotic theory is also in Table 7.

TABLE 6

THE ASYMPTOTIC MEANS AND VARIANCES OF THE m.l.e., CONDITIONAL ON $\hat{\eta} > 0$ WHEN THE TRUE $\eta = 0$

m.l.e.	$\hat{\eta}$	$\hat{\sigma}_H^2$	$\hat{\rho}$
mean	$\sqrt{\frac{2}{\pi}} \frac{\sqrt{2} (1+\rho)}{\sqrt{E} (1-\rho)}$	$\sigma_H^2 \left[1 - \sqrt{\frac{2}{\pi}} \frac{\sqrt{2} (1+\rho)}{\sqrt{E} (1-\rho)} \right]$	$\rho - \frac{\sqrt{2} (1+\rho)^2}{\sqrt{E}} \sqrt{\frac{2}{\pi}}$
variance	$(1 - \frac{2}{\pi}) \frac{2(1+\rho)^2}{E(1-\rho)^2}$	$\frac{4(\sigma_H^2)^2 G}{E[n-(n-2)\rho^2]}$	$\frac{(1+\rho)^2}{E} \left\{ \frac{(n-(n-2)\rho)^2}{[n-(n-2)\rho^2]} - \frac{4}{\pi} (1+\rho)^2 \right\}$

where

$$E = K(n-1)(n-2)$$

and

$$G = \left\{ \frac{[n-(n-2)\rho]^2 (1+\rho^2)}{2(1-\rho)^2} - n + 1 - \frac{(1+\rho)^2 [n-(n-2)\rho^2]}{(1-\rho)^2 \pi} \right\}$$

TABLE 7

COMPARISON BETWEEN THE SIMULATION AND
MORAN'S ASYMPTOTIC THEORY

Estimator	$\hat{\eta}$		$\hat{\sigma}_H^2$		$\hat{\rho}$	
	mean	S.D.	mean	S.D.	mean	S.D.
Simulations with $\hat{\eta} > 0$	0.46	0.39	0.16	0.04	0.90	0.03
Moran's asymptotic theory as in Table 6	0.48	0.36	0.11	0.08	0.89	0.03

CHAPTER IV

THRESHOLD EXCEEDANCES OF MID-DAY OXIDANT CONCENTRATIONS

4.1. Estimation of the Probability of Threshold Exceedances from San Jose Data and from Model Simulations

The Federal standard for oxidant concentration stipulates that the daily maximum of the 24 hourly average concentrations may not exceed a threshold of 12 parts per hundred million (pphm) more than one day in a year. Until recently the standard was more stringent--8pphm. However, this standard is not concerned with the number of hours in any one day during which the threshold is exceeded. For measuring human health effects, it may be that longer exposures to high levels are more damaging [National Academy of Sciences (1977)]. Therefore, we further investigate the frequency distribution of the number of hourly exceedances of a threshold level c during the six-hour mid-day period when the oxidant concentrations are elevated.

As before, $X_{j,t}$ denotes the concentration on day j for hour t . The number of hourly exceedances between 11:00 a.m. and 5:00 p.m. of the threshold level c on the j th day is denoted by N_c^j and can be expressed as $\sum_{t=1}^6 I_{\{X_{j,t} > c\}}$ where I_A is the indicator function of A . If the actual oxidant concentrations behave as model (1.1) and satisfy the assumptions mentioned in Section 1.1, then $\tilde{Y}_j = (\ln X_{j,1}, \dots, \ln X_{j,6})'$, $j = 1, 2, \dots, K$ are K independently identically normally distributed

random vectors. Therefore $N_c^1, N_c^2, \dots, N_c^K$ are i.i.d. discrete random variables with probability function $\pi(i; c) = P[N_c^j = i], 0 \leq i \leq 6$ for $j = 1, 2, \dots, K$. The direct calculation of $\pi(i; c)$ involves complicated integration of a six-variate normal density.

Since the probability function $\pi(i; c)$ depends on the values of the model parameter (μ, θ) , with $\theta = (\eta, \sigma_H^2, \rho)$, we use the more explicit notation $\pi(i; c, \mu, \theta)$, $0 \leq i \leq 6$. Then the m.l.e. of $\pi(i; c, \mu, \theta)$ is $\pi(i; c, \hat{\mu}, \hat{\theta})$ where $\hat{\mu}$ and $\hat{\theta}$ are the m.l.e.'s of μ and θ respectively. However, $\pi(i; c, \hat{\mu}, \hat{\theta})$ is difficult to calculate directly from $\hat{\theta}$ and $\hat{\mu}$. Instead we simulate 40,000 pseudo daily observations Y_k with six hourly averages for each day by using model (1.1) with θ, μ replaced by $\theta = (0.34 \times 10^{-4}, 0.22, 0.93)$ and $\mu = (1.82, 1.97, 2.06, 2.04, 1.93, 1.70)$ estimated from the actual oxidant data. Then $\pi(i; c, \hat{\mu}, \hat{\theta})$ is estimated by the proportion of simulated days with exactly i hours exceeding the threshold level c , for $0 \leq i \leq 6$. Then the simulated $\pi(i; c, \hat{\mu}, \hat{\theta})$ has standard deviation

$$\sqrt{\frac{\pi(i; c, \hat{\mu}, \hat{\theta}) [1 - \pi(i; c, \hat{\mu}, \hat{\theta})]}{40,000}} \leq 0.0025$$

for $0 \leq i \leq 6$.

The three different threshold levels used for the simulations are 8.5 pphm, 10.5 pphm and 12.5 pphm, since the actual oxidant data reported by the Bay Area Air Quality Management District have been truncated to the nearest integer. Table 8 displays the discrepancy which would be observed without this continuity correction using the San Jose data mentioned in Section 1.1. The values of $\pi(i; 8.5, \hat{\mu}, \hat{\theta})$

TABLE 8

THE COMPARISON OF

- (a) DISTRIBUTION OF OBSERVED HOURLY EXCEEDANCES,
 (b) SIMULATED DISTRIBUTION USING MODEL(1.1) WITH m.l.e. AND
 (c) SIMULATED DISTRIBUTION WITHOUT CONTINUITY CORRECTION

No. of mid-day hours exceeding threshold c	c = 8.5 pphm			c = 10.5 pphm		c = 12.5 pphm	
	(a)	(b)	(c)	(a)	(b)	(a)	(b)
i = 0	49%	47.6%	45%	68%	65.9%	78%	78.4%
i = 1	9%	8%	7.9%	8%	7.1%	5%	5.7%
i = 2	8%	7.3%	7.6%	5%	6%	6%	4.4%
i = 3	7%	7.7%	8.2%	5%	5.8%	3%	3.7%
i = 4	8%	8.7%	9.4%	5%	5.5%	3%	3.3%
i = 5	7%	9.8%	10.7%	5%	5.2%	4%	2.8%
i = 6	11%	10.9%	13.8%	5%	4.5%	1%	1.7%

$0 \leq i \leq b$ also are shown in Table 8, to explain why we chose 8.5 pphm, 10.5 pphm and 12.5 pphm instead of 8 pphm, 10 pphm and 12 pphm. The estimated probability $\pi(i; c, \hat{\mu}, \hat{\theta})$, mentioned in the previous paragraph, and the observed $\pi(i; c)$ from San Jose data of exactly i hourly exceedances of the threshold level $c = 8.5$ pphm, 10.5 pphm, 12.5 pphm are shown in Table 8.

4.2. Model Sensitivity Analysis

In the previous section, we estimate $\pi(i; c, \mu, \theta)$, $0 \leq i \leq 6$, i.e. the probability function of the number of the hourly exceedances during mid-day, according to the model (1.1) with the m.l.e. $\hat{\theta}$, $\hat{\mu}$ estimated from the data. However, for different sets of the data, $\hat{\theta}$ may be different. In this section we further look into how sensitive the probability function of the number of hourly exceedances $\pi(i; c, \mu, \theta)$ will be to changes in the model parameters η , σ_H^2 , ρ .

Six different combinations of parameters were selected from the set $\Theta = \{(\eta, \sigma_H^2, \rho) | \eta \in \{0, 1\}, \sigma_H^2 = 0.22, \rho \in \{0.93, 0.5, 0\}\}$. The probabilities $\pi(i; c, \mu, \theta)$, $0 \leq i \leq 6$, were estimated from 40,000 simulated days from the model (1.1) with $(\eta, \sigma_H^2, \rho) \in \Theta$ and $\mu = (1.82, 1.97, 2.06, 2.04, 1.93, 1.71)'$.

Figure 1a shows six different patterns of the simulated probability function $\pi(i; 8.5, \mu, \theta)$, $0 \leq i \leq 6$. For large hour-to-hour correlation ρ , the pattern shows that high probabilities appear at two extremes. It indicates that most days will have either no hours or the entire six-hour period exceeding 8.5 pphm. However, if ρ were

small, the pattern shows that most days would have one to three hours exceeding 8.5 pphm at current mean levels.

Figure 1a also shows that the difference of the model simulated probabilities $\pi(i; 8.5, \mu, \theta)$, $0 \leq i \leq 6$, between $\eta = 0$ (the daily variance σ_D^2 is zero) and $\eta = 1$ (the daily variance σ_D^2 is equal to the hourly variance σ_H^2), increases as ρ increases. This means that for large hour-to-hour correlation ρ , the calculated $\pi(i; 8.5, \mu, \theta)$, $0 \leq i \leq 6$, is more sensitive to the changes in η , the ratio of the daily variance to the hourly variance. The observed relative frequencies in the original data, are shown by the circled points. The closeness of the circled points to the solid line with $\eta = 0$, $\rho = 0.93$, $\sigma_H^2 = 0.22$ indicates the goodness of fit.

Figure 1b shows the corresponding cumulative distribution functions (C.D.F.) for the number of hourly exceedances of the 8.5 pphm threshold. By comparing Figure 1a and Figure 1b, we expect to have smooth and flat curves for large ρ with large values for zero hour frequencies, but sharply increasing curves for small ρ with smaller values for zero hour frequencies.

For the threshold level 10.5 pphm and 12.5 pphm, similar patterns of the observed and simulated model probability functions for number of hourly exceedances are shown in Figure 2a and 3a, respectively. However, the model for $\rho = 0$ has been shifted from two hours in Figure 1a to one hour in Figure 2a and further to zero hour in Figure 3a. The corresponding patterns of the calculated C.D.F. for level 10.5 pphm and 12.5 pphm are shown in Figure 2b and 3b, respectively.

The observed relative frequencies of the number of the hourly exceedances at threshold levels 10.5 pphm and 12.5 pphm are shown in the circled points in Figure 2a-3b. By looking at Table 8, it shows that the probability distributions calculated according to the model (1.1) with the m.l.e. $\hat{\theta}$, $\hat{\mu}$ as the model parameter are very close to the observed one for the cases of threshold values $c = 8.5$ pphm, 10.5 pphm and 12.5 pphm. Figure 1a-3b also demonstrate how well the model (1.1) explains the behavior of the original data.

Figure 1a. The estimated probability function of hourly oxidant exceedances at 8.5 pphm during the midday period from 11 a.m. to 5 p.m. using model simulations with $\sigma_H^2 = 0.22$ fixed. The observed probability function for San Jose, Summer '74, is also shown.

Figure 1b. The estimated cumulative distribution function (C.D.F.) of hourly oxidant exceedances at 8.5 ppm during the midday period from 11 a.m. to 5 p.m. using model simulations with $\sigma_H^2 = 0.22$ fixed. The observed C.D.F. for San Jose, Summer '74, is also shown.

Figure 2a. The estimated probability mass function of hourly oxidant exceedances at 10.5 ppm during the midday period from 11 a.m. to 5 p.m. using model simulations with $\sigma_H^2 = 0.22$ fixed. The observed mass function for San Jose, Summer '74, is also shown.

Figure 2b. The estimated cumulative distribution function (C.D.F.) of hourly exceedances at 10.5 ppm during the midday period from 11 a.m. to 5 p.m. using model simulations with $\sigma_H^2 = 0.22$ fixed. The observed C.D.F. for San Jose, Summer '74, is also shown.

Figure 3a. The estimated probability mass function of hourly oxidant exceedances at 12.5 ppm during the midday period from 11 a.m. to 5 p.m. using model simulations with $\sigma_H^2 = 0.22$ fixed. The observed mass function for San Jose, Summer '74, is also shown.

Figure 3b. The estimated cumulative distribution function (C.D.F.) of hourly exceedances at 12.5 pphm during the midday period from 11 a.m. to 5 p.m. using model simulations with $\sigma_H^2 = 0.22$ fixed. The observed C.D.F. for San Jose, Summer '74, is also shown.

REFERENCES

- Anderson, T. W. (1958). An Introduction to Multivariate Analysis, Wiley, New York.
- Anderson, T. W. (1971). The Statistical Analysis of Time Series, Wiley, New York.
- Anderson, T. W. (1978). Repeated Measurements on Autoregressive Processes, JASA 73, No. 362 (June).
- Box, G. E. P. and Jenkins, G. M. (1970). Time Series Analysis, Forecasting and Control, Holden-Day, San Francisco.
- Cox, D. R. and Hinkley, D. V. (1974). Theoretical Statistics, Chapman and Hall.
- Herbach, L. H. (1959). Properties of Model II Type Analysis of Variance Tests, Ann. Math. Stat. 30, 939-959.
- Larsen, R. I. (1971). A Mathematical Model for Relating Air Quality Measurements to Air Quality Standards, EPA.
- Moran, P. A. P. (1971). The Uniform Consistency of Maximum-Likelihood Estimators. Proc. Cambridge Philos. Soc. 70, pp. 435-439.
- Moran, P. A. P. (1971). Maximum Likelihood Estimation in Non-standard Conditions. Proc. Cambridge Philos. Soc. 70, pp. 441-450.
- National Academy of Sciences (1977), Committee on Medical and Biologic Effects of Environmental Pollutants. Ozone and other Photochemical Oxidants, Chapter 9.
- Rao, C. Radhakrishna (1967). Linear Statistical Inference and its Application, Wiley, New York.
- Searle, S. R. (1971). Linear Model, Wiley, New York.
- Szatrowski, Ted Howard (1976). Estimation and Testing for Block Compound Symmetry and Other Patterned Covariance Matrices with Linear and Nonlinear Structure. Technical Report No. 107, Department of Statistics, Stanford University.