E83-10062
CR-169574

FINAL TECHNICAL REPORT

NASA GRANT NAG-1-71

DEVELOPMENT OF MATHEMATICAL TECHNIQUES FOR THE ASSIMILATION OF REMOTE SENSING DATA INTO ATMOSPHERIC MODELS

John H. Seinfeld, Principal Investigator
California Institute of Technology
Pasadena, California 91125

ABSTRACT

The object of this project was to define the problem of the assimilation of remote sensing data into mathematical models of atmospheric pollutant species. An object of remote sensing of the atmosphere is to enable reconstruction of the concentration distribution of trace species over a region based on the data available from the instrument. The data assimilation problem is posed in terms of the matching of spatially integrated species burden measurements to the predicted three-dimensional concentration fields from atmospheric diffusion models. General conditions have been derived for the reconstructability of atmospheric concentration distributions from data typical of remote sensing applications, and a computational algorithm (filter) for the processing of remote sensing data has been developed.

(E83-10062) DEVELOPMENT OF MATHEMATICAL TECHNIQUES FOR THE ASSIMILATION OF REMOTE SENSING DATA INTO ATMOSPHERIC MODELS Final Technical Report (California Inst. of Tech.) 106 p HC A06/MF A01 CSCL 13B G3/43

N83-14571

Unclas
00062

DEVELOPMENT OF MATHEMATICAL TECHNIQUES FOR THE ASSIMILATION OF REMOTE SENSING DATA INTO ATMOSPHERIC MODELS

Table of Contents

to appear in February 1982 issue of IEEE Transactions on Automatic Control

RECONSTRUCTION OF ATMOSPHERIC POLLUTANT CONCENTRATIONS FROM REMOTE SENSING DATA - AN APPLICATION OF DISTRIBUTED PARAMETER OBSERVER THEORY†

Masato Koda* and John H. Seinfeld
Department of Chemical Engineering
California Institute of Technology
Pasadena, California 91125

ABSTRACT

The reconstruction of a concentration distribution from spatially-averaged and noise-corrupted data is a central problem in processing atmospheric remote sensing data. Distributed parameter observer theory is used to develop reconstructibility conditions for distributed parameter systems having measurements typical of those in remote sensing. The relation of the reconstructibility condition to the stability of the distributed parameter observer is demonstrated. The theory is applied to a variety of remote sensing situations, and it is found that those in which concentrations are measured as a function of altitude satisfy the conditions of distributed state reconstructibility.

*Permanent address: Department of Aeronautics, University of Tokyo, Hongo, Bunkyo-ku, Tokyo 113, Japan

†This work was supported by NASA Research Grant NAG-1-71

I. INTRODUCTION

In the remote sensing of tropospheric species, a ground-, aircraft-, or satellite-based platform performs an instantaneous scan of a region of the atmosphere and measures the species burden within the field of view. With aircraft or satellite remote sensing the platform is in motion and the field of view is constantly changing. An object of remote sensing of the atmosphere is to enable reconstruction of the concentration distribution of trace species over an entire region based on the data available from the instrument.

The reconstruction of a concentration distribution from spatially-averaged and possibly noise-corrupted data is a central problem in processing remote sensing data. In the absence of a mathematical model describing the spatial and temporal concentration distributions, the reconstruction can be carried out by standard data interpolation methods. However, when a mathematical model exists, the problem becomes one of matching the remote sensing data to the model solution in such a way that the incomplete data can be used in conjunction with the model to produce an estimate of the region-wide concentration distribution. This problem of the matching or assimilation of remote sensing data into mathematical models for atmospheric constituents is the subject of this paper.

There exist a few recent studies that assess the capabilities of remote sensing for monitoring regional air pollution episodes. For example, Barnes et al. [1] conducted a comparative analysis of satellite visible channel imagery in ground-based aerosol measurements. For three cases, each of which represented a significant pollution episode based on low surface visibility and high sulfate levels, the results show that the extent and transport of the haze pattern can be monitored from satellite data. The study demonstrated the potential of the satellite to monitor both magnitude and aerial extent of pollution episodes. In a related study, Lyons, et al. [2] reported on a demonstration project showing that currently available synchronous satellite data can detect the aerial extent of large scale hazy air masses associated with sulfate and ozone episodes.

A study related to that of the present work was reported by Diamonte, et al. [3] in which they considered the comparison of remote and in situ data on pollutant concentrations from point sources. They considered typical remote sensing geometries to provide insight on estimation of plume properties from these measurements. In a study also related to the present, Kibbler and Suttles [4] considered the estimation of unknown parameters in a pollutant dispersion model by comparing model predictions with remotely sensed air quality data. A ground-based sensor provided relative pollutant concentration measurements as a function of space and time. The measured data were compared with the dispersion model output through a numerical estimation procedure to yield parameter estimates that best fit the data.

The object of this paper is to define the problem of the assimilation of atmospheric remote sensing data into mathematical models of pollutant behavior. Since the atmosphere is a three-dimensional system, models of pollutant behavior are of the distributed parameter type [5]. Remote sensing data represent spatial averages of concentrations, so that the assimilation problem is, in essence, one of distributed parameter state estimation.

First, the concept of distributed state reconstructibility is developed for the class of problems of interest. That is, the first question to be faced is - can the desired spatial-temporal concentration distribution information be recovered from the measurements in the absence of noise. The derivation of general conditions that allow one to answer this question is the subject of Section II. In Section III a variety of common remote sensing measurement configurations and atmospheric models are tested for reconstructibility. We conclude in Section IV with general observations concerning the inherent potential of remote sensing data in analyzing regional air pollution.

II. RECONSTRUCTIBILITY AND OBSERVERS FOR DISTRIBUTED PARAMETER SYSTEMS

Atmospheric pollutant models consist of partial differential equations, linear in the case in which the species does not react chemically or in which it is produced or destroyed by a first-order reaction of the form $A \rightarrow$. This case represents a wide class of important situations and is the one to which we direct our attention here. Nonlinear distributed models must be handled by linearization and therefore also fall within the present framework.

Our interest in this section is to derive distributed parameter observers for systems described by linear partial differential equations with inhomogeneous boundary conditions characteristic of atmospheric models. An observer is an algorithm that processes measurements of the state of a system to yield an estimate of the entire system state. An observer is most frequently employed when not all of the states of a system are accessible for measurement. In the present application, we will be generally interested in only a single state variable, the measurements of which have limited spatial resolution. The observer is stable if its estimated state converges to the true state after a sufficiently long time. The concept of state reconstructibility is useful as a condition for the stability of the observer. Thus, if a measurement strategy satisfies the condition of state reconstructibility, then the corresponding observer is stable, and, the state (i.e. the concentrations) can, in principle, be estimated from the measurements. The condition that allows the reconstruction of the system state on the entire field is called distributed state reconstructibility. Associated with distributed state reconstructibility, the concept of uniform n-mode reconstructibility can be developed. Both conditions, n-mode and distributed state reconstructibility, will be applied, in Section III, to typical remote sensing measurement configurations.

There exists some previous work on observer theory for distributed parameter systems [6-8]. Kitamura et al. [6] formally extended the lumped parameter observer to the distributed parameter case. Gressang and Lamont [7] developed a more complete theory of the distributed parameter observer, including reduced order observers. An application of distributed parameter observer theory has been presented by Köhne [9]. The most complete treatment of observer theory is that of Dolecki and Russell [8]. In the current work, distributed parameter observers are derived in a form appropriate for application to the class of systems representing atmospheric species behavior. In addition, a result of the present work is an explicit relation between distributed parameter reconstructibility and the stability of the observer. Observer stability is demonstrated using a technique of Hale [10] in which Lyapunov stability theory s extended to function spaces.

We consider the linear distributed parameter system,

$$\frac{\partial u(x,t)}{\partial t} = L_x\, u(x,t) + B(x,t)f(x,t) \tag{1}$$

defined for $t > 0$, $x \in D$. The domain D is a connected subset of a d-dimensional Euclidean space E^d with boundary surface ∂D. The d-dimensional spatial coordinate vector is denoted by x. The state $u(x,t)$ is a scalar function and L_x is a linear partial differential operator with respect to x. It is assumed that the operator L_x is well-posed. The input $f(x,t)$ is a known scalar function and $B(x,t)$ is a known coefficient.

The boundary condition on (1) is

$$\beta_x u(x,t) = h(x,t) \qquad x \in \partial D \tag{2}$$

where β_x is a linear, spatial differential operator of suitable order over ∂D and $h(x,t)$ is a known function. The initial condition is assumed to be unknown or incompletely known.

We are interested in considering three types of measurements:

Case 1: Spatially-Independent Integral Measurements

The measurement takes the form

$$w(t) = \int_D H(x,t)u(x,t)dx \tag{3}$$

where $H(x,t)$ is a spatial weighting function.

Case 2: Spatially-Continuous Measurements

$$w(x,t) = C(x,t)u(x,t) \tag{4}$$

where $C(x,t)$ is a square-integrable function, i.e., $C \in L_2$.

Case 3: Spatially-Discrete Measurements

$$w_i(t) = H_i(t)u(x_i,t) \qquad i = 1,2,\dots \ell \tag{5}$$

where $w_i(t)$ denotes a measurement at the ith measurement location x_i. By taking the limit to small volumes of integration in (3), we can represent a system such as (5) by choosing $H(x,t) = H_i(t)\delta(x-x_i)$, $i = 1,2,\dots, \ell$.

For the moment let us restrict the problem to one spatial dimension, i.e., $0 \leq x \leq 1$. Accordingly, boundary condition (2) can be expressed as

$$\begin{aligned} \beta_0 u(x,t) &= h_0(t) \qquad & x = 0 \\ \beta_1 u(x,t) &= h_1(t) \qquad & x = 1 \end{aligned} \tag{6}$$

Then the solution of (1) and (6) with initial condition $u(x,0) = u_0(x)$ can be expressed in the form*

$$u(x,t) = \int_0^1 \Phi^* (r,0;x,t)u_0(r)dr + \int_0^t\int_0^1 \Phi^* (r,\tau;x,t)B(r,\tau)f(r,\tau)drd\tau$$

$$+ \int_0^t\int_0^1 \Phi^* (r,\tau;x,t)g(r,\tau)drd\tau \tag{7}$$

where

$$g(x,t) = 2h_1(t)\delta(x-1) - 2h_0(t)\delta(x) \; . \tag{8}$$

The adjoint Green's function $\Phi^*(x,t;y,\tau)$ is governed by

$$\frac{\partial\Phi^*(x,t;y,\tau)}{\partial t} + L_x^*\Phi^*(x,t;y,\tau) = 0 \tag{9}$$

with the terminal condition

$$\Phi^*(x,t;y,t) = \delta(x-y) \tag{10}$$

*The explicit form of operators L_x, β_0, and β_1 are assumed as follows:

$$L_x(\cdot) = \alpha_2(x,t)\frac{\partial^2(\cdot)}{\partial x^2} + \alpha_1(x,t)\frac{\partial(\cdot)}{\partial x} + \alpha_0(x,t)(\cdot)$$
$$\beta_0(\cdot) = \alpha_2(0,t)\frac{\partial(\cdot)}{\partial x} + \theta_0(t)(\cdot)$$
$$\beta_1(\cdot) = \alpha_2(1,t)\frac{\partial(\cdot)}{\partial x} + \theta_1(t)(\cdot)$$

and boundary conditions

$$\beta_0^* \Phi^* = 0$$
$$\beta_1^* \Phi^* = 0 \tag{11}$$

The operators L_x^*, β_0^*, and β_1^* are the adjoints of the operators L_x, β_0, and β_1, respectively.

The extension of the adjoint Green's functions to higher spatial dimensions is straightforward. In higher dimensions, (9) and (10) remain the same with the general boundary conditions

$$\beta_x^* \Phi^* (x,t;y,\tau) = 0 \qquad x \in \partial D \tag{12}$$

where β_x^* is the adjoint of the operator β_x. In general, we note that Φ^* is related to the Green's function Φ associated with the system (1) with homogeneous boundary conditions by the relationship $\Phi(x,t;y,\tau) = \Phi^*(y,\tau;x,t)$.

The adjoint Green's function for well-posed distributed parameter systems can be constructed in a variety of ways. Expansion in spatial eigenfunctions and construction of the adjoint Green's function from eigenvalues and eigenfunctions is a powerful method for linear systems. Let us assume that L_x^* has an infinite series of discrete eigenvalues $\{\lambda_i\}$, $i = 1,2,\ldots.$ Using standard methods, the adjoint Green's function that satisfies (9)-(12) is found to be [11]

$$\Phi^*(x,t;y,\tau) = \sum_{n=1}^{\infty} \phi_n(x)\phi_n(y)e^{-\lambda_n(t-\tau)} \tag{13}$$

where the eigenfunctions $\{\phi_i\}$, $i = 1,2,\ldots$, are the solution of the equation, $L_x^*\phi_i = \lambda_i\phi_i$, satisfying the boundary conditions (11) or (12).

II.1 Reconstructibility Conditions

The objective of an observer is to reconstruct the system state when the measurements are incomplete. To be able to reconstruct the state the observer must be asymptotically stable.

An identity or non-reduced observer for the system (1) with measurements (4) takes the form

$$\frac{\partial \hat{u}(x,t)}{\partial t} = L_x \hat{u}(x,t) + B(x,t)f(x,t) + G[w(x,t) - C(x,t)\hat{u}(x,t)] \tag{14}$$

where $\hat{u}(x,t)$ is the observer output and G is a suitably chosen integral operator with the kernel $G(x,y,t)$.

Before presenting a derivation of the observer, we will establish the conditions under which the system (1) and (4) is reconstructible. We define the reconstructibility kernel function by

$$Q(x,y,t) = \int_0^t \int_D \Phi^*(x,t;r,\tau)C^2(r,\tau)\Phi^*(y,t;r,\tau)drd\tau \tag{15}$$

It will be shown later that the observer (14) is stable if $Q(x,y,t)$ has a so-called generalized inverse, i.e., if there exists $P(x,y,t)$ such that

$$\int_D P(x,r,t)Q(r,y,t)dr = \delta(x-y) \tag{16}$$

By formal differentiation of (15) with respect to time and use of the properties of the adjoint Green's function (9) - (12), it is found that $Q(x,y,t)$ satisfies the following Lyapunov equation,

$$\frac{\partial Q(x,y,t)}{\partial t} = - L_x^* Q(x,y,t) - Q(x,y,t)L_y^* + c^2(x,t)\delta(x-y) \tag{17}$$

with the initial condition

$$Q(x,y,0) = 0 \tag{18}$$

and boundary conditions

$$\beta_x^* Q = 0, \qquad Q\beta_y^* = 0 \tag{19}$$

where $L_y^* Q = QL_y^*$. Although $Q(x,y,t)$ is formally defined by (15), it is important to note that $Q(x,y,t)$ may be computed from (17)-(19) without using the adjoint Green's function.

By using the identity

$$\frac{\partial P(x,y,t)}{\partial t} = - \int_D \int_D P(x,r,t) \frac{\partial Q(r,s,t)}{\partial t} P(s,y,t)dsdr \tag{20}$$

$P(x,y,t)$ can be shown to obey the following Riccati equation,*

$$\frac{\partial P(x,y,t)}{\partial t} = L_x P(x,y,t) + P(x,y,t)L_y - \int_D P(x,r,t)\, c^2(r,t)P(r,y,t)dr \tag{21}$$

with boundary conditions

$$\beta_x P = 0, \qquad P\beta_y = 0 \tag{22}$$

$P(x,y,t)$ may be considered as the kernel of the integral operator P defined as

*The impact of observation error on the design of an observer can be assessed from (21) by comparing P to that from the corresponding distributed parameter filter.

$$Pf(x) = \int_D P(x,y,t)f(y)dy \tag{23}$$

for $f \in L_2$.

A linear distributed parameter system (1) and (2) with measurement (4) is said to be *distributed state reconstructible* if and only if $Q(x,y,t)$ defined by (15) has a bounded generalized inverse $P(x,y,t)$ for $t > 0$. It may be shown that $Q(x,y,t)$ has a bounded generalized inverse when $Q(x,y,t)$ is bounded and positive-definite for $t > 0$ [11]*. The system (1), (2), and (4) will be defined to be *uniformly n-mode reconstructible* if there exists positive constants M_1, M_2, and σ such that

$$M_1 \leq \int_D \int_D \phi_n(x) Q^\sigma(x,y,t) \phi_n(y) \; dxdy \leq M_2 \tag{24}$$

for all $t > 0$, where $\phi_n(x)$ is the eigenfunction of L_x^* and the modified reconstructibility kernel $Q^\sigma(x,y,t)$ is defined by

$$Q^\sigma(x,y,t) = \int_{t-\sigma}^{t} \int_D \Phi^*(x,t;r,\tau) C^2(r,\tau) \Phi^*(y,t;r,\tau) \; drd\tau \tag{25}$$

The system is distributed state reconstructible if (24) is satisfied for each of the eigenfunctions. The uniform n-mode reconstructibility test (24) is useful when $P(x,y,t)$ cannot be found directly from $Q(x,y,t)$. Since it is straightforward to extend the concept of distributed state reconstructibility to measurement Cases 1 and 3, detailed discussion is omitted here.

*Positive-definiteness of the kernel Q implies that

$$\int_D \int_D f(x) Q(x,y,t) f(y) \; dxdy > 0$$

for all $t > 0$ and $f \in L_2$.

II.2 Minimum Variance Observers

II.2.1 Observer for Case 1

For the system described by (1) and (3), we define the reconstructibility kernel function by

$$Q(x,y,t) = \int_0^t \int_D \Phi^*(x,t;r,\tau)H(r,\tau)dr \int_D H(s,\tau)\Phi^*(y,t;s,\tau)dsd\tau \qquad (26)$$

where $Q(x,y,t)$ obeys

$$\frac{\partial Q(x,y,t)}{\partial t} = - L_x^* Q(x,y,t) - Q(x,y,t)L_y^* + H(x,t)H(y,t) \qquad (27)$$

with initial and boundary conditions given by (18) and (19). Assuming that the system is distributed state reconstructible, the existence of the generalized inverse $P(x,y,t)$ of $Q(x,y,t)$, that satisfies

$$\frac{\partial P(x,y,t)}{\partial t} = L_x P(x,y,t) + P(x,y,t)L_y$$
$$- \int_D P(x,r,t)H(r,t)dr \int_D H(s,t)P(s,y,t)ds \qquad (28)$$

will establish the observer for the system (1), (2) and (3).

Following Meditch [12], we define the cost functional associated with the observer as

$$J_0 = \frac{1}{2}\int_D [u(x,0) - u_0(x)]' \; A_x \; [u(y,0) - u_0(y)]dx$$

$$+ \frac{1}{2}\int_0^{t_f} \left\{ w(t) - \int_D H(x,t)u(x,t)dx \right\}^2 dt \tag{29}$$

where t_f is an arbitrary final time, $u_0(x)$ is an initial estimate of $u(x,0)$, and

$$A_x \, (\cdot) = \left[\int_D P_0(x,y)\{\cdot\}dy \right]^{-1} \tag{30}$$

$P_0(x,y)$ is a bounded, symmetric, and positive-definite weighting function. The observer is found by selecting $u(x,t)$ so as to minimize (29) subject to (1) and (2). By minimizing the augmented functional,

$$J = J_0 + \int_0^{t_f}\int_D \lambda(x,t)\left[\frac{\partial u(x,t)}{\partial t} - L_x u(x,t) - B(x,t)f(x,t) \right] dxdt \tag{31}$$

the result is the Euler-Lagrange equation,

$$\frac{\partial \lambda(x,t)}{\partial t} = - L_x^* \lambda(x,t) - H(x,t)\left[w(t) - \int_D H(y,t)\hat{u}(y,t)dy \right] \tag{32}$$

with the transversality conditions,

$$\begin{aligned} \lambda(x,0) &= A_x \; [\hat{u}(y,0) - u_0(y)] \\ \lambda(x,t_f) &= 0 \end{aligned} \tag{33}$$

Equations (32) and (33) constitute a two-point boundary value problem that may be solved by the sweep method. We assume the following Riccati transformation for $\hat{u}(x,t)$,

$$\hat{u}(x,t) = \int_D P(x,y,t)\lambda(y,t)dy + p(x,t) \tag{34}$$

where the kernel $P(x,y,t)$ and $p(x,t)$ have to be determined.
Substitution of (34) into (1), (32) and (33) yields

$$\frac{\partial p(x,t)}{\partial t} = L_x p(x,t) + B(x,t)f(x,t)$$

$$+ \int_D P(x,r,t)H\ (r,t)dr \left[w(t) - \int_D H(s,t)p(s,t)ds\right] \tag{35}$$

$$p(x,0) = u_0(x) \tag{36}$$

$$\beta_x p(x,t) = h(x,t) \qquad x \in \partial D \tag{37}$$

$$\frac{\partial P(x,y,t)}{\partial t} = L_x P(x,y,t) + P(x,y,t)L_y$$

$$- \int_D P(x,r,t)H\ (r,t)dr \int_D H(s,t)P(s,y,t)ds \tag{38}$$

$$P(x,y,0) = P_0(x,y) \tag{39}$$

$$\beta_x P(x,y,t) = 0, \qquad P(x,y,t)\beta_y = 0 \qquad x,y \in \partial D \tag{40}$$

Equations (33) and (34) imply that $p(x,t_f) = \hat{u}(x,t_f)$ is the state estimate at an arbitrary final time t_f. It is important to note that (38) is identical to (28). Thus we may conclude that the symmetric, positive-definite kernel $P(x,y,t)$ completely characterizes the minimum variance observer.

Equation (35) can be rewritten as

$$\frac{\partial \hat{u}(x,t)}{\partial t} = L_x \hat{u}(x,t) + B(x,t)f(x,t)$$

$$+ K(x,t)\left[w(t) - \int_D H(y,t)\hat{u}(y,t)dy\right] \tag{41}$$

where a time-varying observer gain $K(x,t)$ is defined by

$$K(x,t) = \int_D P(x,y,t)\, H(y,t)dy \tag{42}$$

The structure of the observer is identical to that of the distributed parameter filter [13].

We introduce the reconstruction or observer error $e(x,t) = \hat{u}(x,t) - u(x,t)$. Then we obtain the following equation for $e(x,t)$,

$$\frac{\partial e(x,t)}{\partial t} = L_x e(x,t) - K(x,t) \int_D H(y,t)e(y,t)dy \tag{43}$$

with initial and boundary conditions, $e(x,0) = u_0(x) - u(x,0)$, and $\beta_x e(x,t) = 0$. If the initial state is known exactly and the observer is initialized such that $\hat{u}(x,0) = u(x,0)$, then the observer will reconstruct the state exactly. It is not reasonable, however, to expect that the initial state will be known exactly. It is, therefore, important to insure that if errors are present in the initial conditions applied to the observer that the estimate will converge to the true value of the state, i.e., the reconstruction error $e(x,t)$ must have the property $\lim_{t\to\infty} \|e(x,t)\| = 0$, for all $e(x,0)$.

Asymptotic stability of the observer can be demonstrated by using (16), (26), (27), and (43). We will consider a Lyapunov function defined by

$$V(e,t) = \int_D\int_D e(x,t)Q(x,y,t)e(y,t)dydx \ . \tag{44}$$

It is first necessary to note that $Q(x,y,t)$ is positive-definite and bounded from below. Then the time derivative of the Lyapunov function is calculated using (27) and (43). The result is

$$\frac{d}{dt} V(e,t) = - \int_D \int_D e(x,t)\; H(x,t)H(y,t)e(y,t)dydx \tag{45}$$

which is a negative-semidefinite quadratic form. This is sufficient to show that (43) is stable in the sense of Lyapunov [10].

II.2.2 Observer for Case 2

In a similar manner to that of Case 1, we can obtain the minimum variance observer for Case 2, i.e., for spatially-continuous measurements (4). The observer dynamics are described by

$$\frac{\partial \hat{u}(x,t)}{\partial t} = L_x \hat{u}(x,t) + B(x,t)f(x,t) + \int_D G(x,y,t)[w(y,t) - C(y,t)\hat{u}(y,t)]dy \tag{46}$$

with initial and boundary conditions

$$\hat{u}(x,t) = u_0(x) \tag{47}$$

$$\beta_x \hat{u}(x,t) = h(x,t), \qquad x \in \partial D \tag{48}$$

where the optimal gain kernel G(x,y,t) is defined by

$$G(x,y,t) = P(x,y,t)\; C(y,t). \tag{49}$$

The Riccati equation for P(x,y,t) in (49) is identical to (21) with boundary conditions given by (22). The reconstruction error $e(x,t) = \hat{u}(x,t) - u(x,t)$ satisfies

$$\frac{\partial e(x,t)}{\partial t} = F\; e(x,t) \tag{50}$$

where the integro-differential operator F is defined by

$$F\, e(x,t) = L_x\, e(x,t) - \int_D G(x,y,t)C(y,t)e(y,t)dy \tag{51}$$

We can demonstrate the stability of the observer by using the reconstructibility kernel Q(x,y,t) defined by (15) and the Lyapunov function (44). Under the reconstructibility assumption, the derivative of the Lyapunov function becomes

$$\frac{d}{dt} V(e,t) = - \int_D e(x,t)\, C^2(x,t)e(x,t)dx \tag{52}$$

which is a negative-semidefinite quadratic form.

II.2.3 Observer for Case 3

For the spatially-discrete measurements (5), i.e., Case 3, the observer is given by the following system:

$$\begin{aligned}\frac{\partial \hat{u}(x,t)}{\partial t} &= L_x\hat{u}(x,t) + B(x,t)f(x,t) \\ &\quad + \sum_{i=1}^{\ell} G_i(x,t)[w_i(t) - H_i(t)\hat{u}(x_i,t)]\end{aligned} \tag{53}$$

where

$$G_i(x,t) = P(x,x_i,t)H_i(t) \tag{54}$$

and

$$\begin{aligned}\frac{\partial P(x,y,t)}{\partial t} &= L_x P(x,y,t) + P(x,y,t)L_y \\ &\quad - \sum_{i=1}^{\ell} P(x,x_i,t)H_i(t)H_i(t)P(x_i,y,t)\end{aligned} \tag{55}$$

with initial and boundary conditions given by (47), (48), (39), and (40). Under the distributed state reconstructibility assumption, the stability of the observer can be demonstrated.

II.3 Comments

The relationship has been established between distributed state reconstructibility and the existence of an observer. Distributed state reconstructibility is defined through the existence of the generalized inverse to the reconstructibility kernel. The kernel associated with the observer gain satisfies the same Riccati equation as does the generalized inverse of the reconstructibility kernel.

III. REMOTE SENSING MEASUREMENTS AND ATMOSPHERIC MODELS

In this section we will test both the n-mode and state reconstructibility of common remote sensing measurements with models of atmospheric pollutant behavior. By far the predominant mode of remote sensing is to measure the integrated quantity (burden) of material between the ground and some known altitude. Thus, both cases we consider here involve vertically integrated data. Various assumptions concerning the horizontal characteristics of the measurements will be tested. Three-dimensional models of pollutant behavior are generally based on the atmospheric diffusion equation [5] that describes the flow and diffusion of species. The object of this section is to ascertain if the customary remote sensing measurements allow one, in principle, to reconstruct the detailed concentration distribution. The distributed parameter reconstructibility condition derived in Section II will therefore be tested in each case.

III.1 Measurements in a Layer with Horizontal Homogeneity

The vertical concentration distribution of a pollutant in a layer with horizontal homogeneity can be described by the one-dimensional diffusion equation,

$$\frac{\partial u}{\partial t} = K \frac{\partial^2 u}{\partial z^2} \tag{56}$$

subject to

$$K \frac{\partial u}{\partial z} = h_o(t) \qquad z = 0 \tag{57}$$

$$K \frac{\partial u}{\partial z} = 0 \qquad z = 1 \tag{58}$$

where h_o is a given flux at the ground (z=0) and K is the turbulent diffusion coefficient.

The adjoint Green's function for the system (56)-(58) is

$$\Phi^*(z,t;z',\tau) = 1 + 2\sum_{n=1}^{\infty} \cos(n\pi z)\cos(n\pi z')e^{(n\pi)^2 K(t-\tau)} \tag{59}$$

State reconstructibility is then to be assessed by condition (24) using the modified reconstructibility kernel (25).

We consider each of the measurement types (3), (4) and (5). The condition for uniform n-mode reconstructibility is (24), which is written for $\phi_n = \cos(n\pi z)$, $n = 0, 1,2,...,$ as

$$0 < M_1 \leq \int_0^1 \int_0^1 \cos(n\pi z)Q^\sigma(z,z',t)\cos(n\pi z')dzdz' \leq M_2 < \infty \tag{60}$$

For each of the three types of measurement, the integral in (60) is:

Case 1: Spatially-Independent Integral Measurements

$$\int_{t-\sigma}^{t} e^{2(n\pi)^2K(t-\tau)}\left\{\int_0^1 H(r,\tau)\cos(n\pi r)dr\right\}^2 d\tau \tag{61}$$

Case 2: Spatially-Continuous Measurements

$$\int_{t-\sigma}^{t} e^{2(n\pi)^2K(t-\tau)}\int_0^1 \left\{C(r,\tau)\cos(n\pi r)\right\}^2 drd\tau \tag{62}$$

Case 3: Spatially-Discrete Measurements

$$\int_{t-\sigma}^{t} e^{2(n\pi)^2K(t-\tau)}\left\{\sum_{i=1}^{\ell} H_i(\tau)\cos(n\pi z_i)\right\}^2 d\tau \tag{63}$$

for $n = 0, 1, 2, \ldots$.

From (61)-(63), we see that uniform n-mode reconstructibility is completely dependent on the form of the measurement weighting functions, $H(z,t)$, $C(z,t)$ and $H_i(t)$ and on the eigenfunction, $\cos(n\pi z)$. The condition (60) implies that $\int_0^1 H(z,t)\cos(n\pi z)dz \neq 0$. We may note that this inequality is essentially equivalent to the observability condition derived by McGlothin [14]. Similarly, (63) implies that the system state is reconstructible by point sensors if the sensors are not located at the zeros of any of the eigenfunctions.

In the remote sensing problem, the measurement weighting functions are often taken as $H(z,t) = 1$ or $C(z,t) = 1$. When $H(z,t) = 1$, the condition (60) holds only for $n = 0$ implying that the spatially-independent integral measurements do not allow reconstruction of the system state on entire fields.* This can be directly checked by computing the reconstructibility kernel $Q(z,z',t)$. The system with integral measurements cannot be distributed state reconstructible since the generalized inverse of $Q(z,z',t) = t$ does not exist, since Q is not an explicit function of z and z'.

When the measurements are spatially-continuous and $C(z,t) = 1$, the system is distributed state reconstructible. From the definitions of $Q(x,y,t)$ and $P(x,y,t)$ in (15) and (16), we have

$$Q(z,z',t) = t + \frac{1}{\pi^2 K} \sum_{n=1}^{\infty} \frac{1}{n^2} \cos(n\pi z)\cos(n\pi z')\left\{e^{2(n\pi)^2 Kt} - 1\right\} \tag{64}$$

and

$$P(z,z',t) = \frac{1}{t} + 4\pi^2 K \sum_{n=1}^{\infty} n^2 \cos(n\pi z)\cos(n\pi z')\left\{e^{2(n\pi)^2 Kt} - 1\right\}^{-1} \tag{65}$$

We may note that the integral equation (16) is satisfied when it is recognized that

*A mode associated with the eigenfunction $\phi_0 = 1$ $(n = 0)$ can be reconstructible and the appropriate observer is

$$\frac{\partial \hat{u}}{\partial t} = K \frac{\partial^2 \hat{u}}{\partial z^2} + \frac{1}{t}\left[w(t) - \int_0^1 \hat{u}(z',t)\, dz'\right]$$

Stability of the observer can be demonstrated by constructing the Lyapunov function

$$V(e,t) = \int_0^1\int_0^1 e(z,t)te(z',t)\, dzdz'$$

$$\delta(z-z') = 1 + 2\sum_{n=1}^{\infty} \cos(n\pi z)\cos(n\pi z') \tag{66}$$

$P(z,z',t)$ is bounded and positive-definite, and for $t > 0$, the series (65) is uniformly convergent.

III.2 Measurements of a Steady State, Point Source Plume

The concentration distribution in a plume from a continuously emitting, elevated point source can be described by

$$\frac{\partial u}{\partial t} = K_y \frac{\partial^2 u}{\partial y^2} + K_z \frac{\partial^2 u}{\partial z^2} \tag{67}$$

where t is the time an element of fluid spends in the plume from emission, equal to downwind distance x divided by the wind speed. The source is of strength q located at $t = 0$, $y = 1/2$, $z = z_H (0 \leq z_H \leq 1)$. The boundary conditions on (67) are

$$u(0,y,z) = q\delta(y - 1/2)\delta(z-z_H) \tag{68}$$

$$\frac{\partial u}{\partial y} = 0 \qquad y = 0,1 \tag{69}$$

$$K_z \frac{\partial u}{\partial z} = h_o \qquad z = 0 \tag{70}$$

$$\frac{\partial u}{\partial z} = 0 \qquad z = 1 \tag{71}$$

The adjoint Green's function for this system is

$$\Phi^*(y,z,t;\, y',z',\tau) = \left\{1 + 2\sum_{n=1}^{\infty} \cos(n\pi y)\cos(n\pi y')e^{(n\pi)^2 K_y(t-\tau)}\right\} \times \left\{1 + 2\sum_{m=1}^{\infty} \cos(m\pi z)\cos(m\pi z')e^{(m\pi)^2 K_z(t-\tau)}\right\} \tag{72}$$

Consider first a scanning measurement performed at a horizontal position $y = y_*$,

$$w(t) = \int_0^1 J(z)u(t,y_*,z)dz \tag{73}$$

where the scanning data $w(t)$ are taken on a coordinate that moves along the t-axis. $J(z)$ in (73) is the altitude-dependent weighting function for the measurements. When $J = 1$, the reconstructibility kernel function becomes

$$\begin{aligned} Q\ (t,y,z;y',z') &= t \\ &+ \frac{2}{K_H}\sum_{n=1}^{\infty}\frac{\cos(n\pi y_*)}{(n\pi)^2}\left\{\cos(n\pi y) + \cos(n\pi y')\right\}\left\{e^{(n\pi)^2 K_H t} - 1\right\} \\ &+ \frac{4}{K_H}\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\frac{\cos(n\pi y_*)\cos(m\pi y_*)}{(n^2+m^2)\pi^2}\cos(n\pi y)\cos(m\pi y')\left\{e^{(n^2+m^2)\pi^2 K_H t} - 1\right\} \end{aligned} \tag{74}$$

The system is not distributed state reconstructible since the generalized inverse of $Q(t,y,z;y',z')$ does not exist. Therefore, we con[illegible] that the scanning measurement (73) cannot, in principle, allow re[illegible]tion of the system state.

The same results can be obtained for the following measurement systems:

$$w(t) = \int_0^1\int_0^1 u(t,y,z)dydz \tag{75}$$

$$w(t,y) = \int_0^1 u(t,y,z)dz \tag{76}$$

$$w(t,z) = \int_0^1 u(t,y,z)dy \tag{77}$$

In these cases, the reconstructibility kernel function $Q(t,y,z;y',z')$ cannot be written explicitly in terms of all the spatial variables y,z, and y',z'. Thus, the generalized inverse $P(t,y,z;y',z')$ does not exist, which allows reconstruction of the system state on the whole domain. As a rule, if $Q(t,y,z;y',z')$ is expressible as an explicit function of all the spatial variables and if it satisfies the uniform n-mode reconstructibility test, then the system is distributed state reconstructible.

Indeed, we can show that the system state is distributed state reconstructible for the measurement

$$w(t,y,z) = u(t,y,z) \tag{78}$$

In this case, we have

$$\begin{aligned} Q\,(t,y,z;y',z') = t &+ \frac{1}{K_H}\sum_{n=1}^{\infty}\frac{1}{(n\pi)^2}\cos(n\pi y)\cos(n\pi y')\left\{e^{2(n\pi)^2K_Ht} - 1\right\} \\ &+ \frac{1}{K_V}\sum_{n=1}^{\infty}\frac{1}{(n\pi)^2}\cos(n\pi z)\cos(n\pi z')\left\{e^{2(n\pi)^2K_Vt} - 1\right\} \\ &+ 2\sum_{n=1}^{\infty}\sum_{m=1}^{\infty}\frac{\cos(n\pi y)\cos(m\pi z)\cos(n\pi y')\cos(m\pi z')}{(n\pi)^2K_H+(m\pi)^2K_V}\left[e^{2\{(n\pi)^2K_H+(m\pi)^2K_V\}t} - 1\right] \end{aligned} \tag{79}$$

The generalized inverse of (79) is given by

$$P(t,y,z;y',z') = \frac{1}{t} + 4K_H \sum_{n=1}^{\infty} (n\pi)^2 \cos(n\pi y)\cos(n\pi y')\left\{e^{2(n\pi)^2K_Ht} - 1\right\}^{-1}$$

$$+ 4K_V \sum_{n=1}^{\infty} (n\pi)^2 \cos(n\pi z)\cos(n\pi z')\left\{e^{2(n\pi)^2K_Vt} - 1\right\}^{-1}$$

$$+ 8 \sum_{n=1}^{\infty} \sum_{m=1}^{\infty} \left\{(n\pi)^2K_H + (m\pi)^2K_V\right\} \cos(n\pi y)\cos(m\pi z)\cos(n\pi y')\cos(m\pi z')$$

$$\times\left[e^{2\{(n\pi)^2K_H+(m\pi)^2K_V\}t} - 1\right]^{-1} \tag{80}$$

where (80) satisfies the Riccati equation associated with the measurement (78).

IV. CONCLUSION

This paper has examined the possibility of estimating atmospheric species concentration distributions from remote sensing data. Atmospheric concentrations can be modeled by partial differential equations of the diffusion type. Remote sensing data generally represent spatial averages of the concentrations, frequently in the vertical direction. The essential problem, therefore, is to assess the possibility of estimating the state of a distributed parameter system on the basis of spatially-averaged measurements. The theoretical basis of the assessment is a condition for state reconstructibility of distributed parameter systems. (The connection between state reconstructibility and the stability of the distributed parameter observer has also been developed.)

A variety of remote sensing measurement configurations were tested for reconstructibility. It was found, not unexpectedly, that those measurements based on integration of the vertical concentration distribution over the entire layer do not lead to distributed state reconstructibility, i.e., there does not exist a generalized inverse of the reconstructibility matrix kernel and therefore do not afford the possibility of estimating the concentration distribution over entire field. Those measurement configurations that, on the other hand, enable sampling of the concentration at vertical positions lead to distributed state reconstructibility.

REFERENCES

[1] J. C. Barnes, C. J. Bowley and H. K. Burke, "Evaluation of the Capabilities of Satellite Imagery for Monitoring Regional Air Pollution Episodes," NASA Contractor Report 159107, Langley Research Center, Va., August 1979

[2] W. A. Lyons, J. C. Dooley, Jr. and K. T. Whitby, "Satellite Detection of Long-Range Pollution Transport and Sulfate Aerosol Hazes," Atmospheric Environment, vol. 12, pp. 621-631, 1978.

[3] J. M. Diamante, T. S. Englar, Jr. and A. H. Jazwinski, "Final Report Phase I - Urban Air Quality Estimation Study," NASA Report CR- 145097, Langley Research Center, Va., January 1976.

[4] J. F. Kibler and J. T. Suttles, "Air Pollution Model Parameter Estimation Using Simulated LIDAR Data," AIAA Journal, vol. 15, pp. 1381-1384 1977.

[5] J. H. Seinfeld, Air Pollution: Physical and Chemical Fundamentals, McGraw-Hill, New York, 1975.

[6] S. Kitamura, S. Sakairi and M. Nishimura, "Observers for Distributed Parameter Systems of Diffusion Type," Trans. IEE Japan, vol. 92-C, pp. 407 - , 1972.

[7] R. V. Gressang and G. B. Lamont, "Observers for Systems Characterized by Semigroups," IEEE Trans. Auto. Control, vol. AC-20, pp. 523-528, 1975.

[8] S. Dolecki and D. L. Russell, "A General Theory of Observations and Control," SIAM J. Control and Optimization, vol. 15, pp. 185-220, 1977.

[9] M. Köhne, "The Control of Vibrating Elastic Systems," in Distributed Parameter Systems, W. H. Ray and D. G. Lainiotis, editors, Marcel Dekker, New York, pp. 387-456, 1977.

[10] J. K. Hale, "Dynamical Systems and Stability," J. Math. Anal. Appl., vol. 26, pp. 39-59, 1969

[11] K. Yosida, Functional Analysis, Springer-Verlag, New York, 1966.

[12] J. S. Meditch, "Least Squares Filtering and Smoothing for Linear Distributed-Parameter Systems," Automatica, vol. 7, pp. 315-322, 1971.

[13] S. G. Tzafestas, "Distributed Parameter State Estimation," in Distributed Parameter Systems, W. H. Ray and D. G. Lainiotis, editors, Marcel Dekker, New York, pp. 135-208, 1977.

[14] G. E. McGlothin, "Controllability, Observability, and Quality in a Distributed Parameter System with Continuous and Point Spectrum, IEEE Trans. Auto. Control, vol. AC-23, pp. 687-690, 1978.

to appear in IEEE Transactions on Systems, Man, and Cybernetics

Filtering and Smoothing for Linear Discrete-Time Distributed Parameter Systems Based on Wiener–Hopf Theory with Application to Estimation of Air Pollution

SIGERU OMATU, MEMBER, IEEE, AND JOHN H. SEINFELD

IEEE: SYST., MAN, CYBERN.
Vol. ______ Issue ______
Au: Omatu
Galley No. 1

Abstract—Optimal filtering and smoothing algorithms for linear discrete-time distributed parameter systems are derived by a unified approach based on the Wiener–Hopf theory. The Wiener–Hopf equation for the estimation problems is derived using the least-squares estimation error criterion. Using the basic equation, three types of the optimal smoothing estimators are derived, namely, fixed-point, fixed-interval, and fixed-lag smoothers. Finally, the results obtained are applied to estimation of atmospheric sulfur dioxide concentrations in the Tokushima prefecture of Japan.

I. Introduction

A NUMBER of important physical phenomena may be modeled as discrete-time distributed parameter systems. When estimation problems are encountered in such systems, the measurements are also frequently discrete in time. A great deal of work has been carried out on estimation problems for continuous-time distributed parameter systems [1]–[4]. Tzafestas [5], [6] and Nagamine *et al.* [7] have derived optimal estimators for discrete-time distributed parameter systems. Tzafestas employed a Bayesian approach, where Nagamine *et al.* considered only the filtering problem based on the Wiener–Hopf theory. Recently, Bencala and Seinfeld [3] have derived the optimal filter for continuous-time distributed parameter systems with discrete-time observations by the Wiener–Hopf approach.

The object of this paper is twofold. First, we seek to derive optimal filtering and smoothing algorithms for discrete-time distributed parameter systems by a unified Wiener–Hopf approach. Fixed-point, fixed-interval, and fixed lag smoothers are considered. Second, we wish to apply the results to the estimation of atmospheric sulfur dioxide concentrations in the Tokushima prefecture of Japan.

II. Description of the Distributed Parameter System

Let D be a bounded open domain of an r-dimensional Euclidean space with smooth boundary ∂D. The spatial coordinate vector will be denoted by $x = (x_1, \cdots, x_r) \in D$. Consider a linear distributed parameter system described by

$$u(k+1, x) = \mathcal{L}_x u(k, x) + G(k, x) w(k, x), \qquad x \in D \tag{1}$$

where $u(k+1, x)$ is an n-dimensional vector function of the system, $w(k, x)$ is a vector-valued Gaussian process, $\mathcal{L}_x$ is a linear spatial matrix differential operator, and $G(k, x)$ is a known matrix function.

The initial and boundary conditions are given by

$$u(0, x) = u_0(x) \tag{2}$$

$$\Gamma_\xi u(k+1, \xi) = S(k+1, \xi), \qquad \xi \in \partial D \tag{3}$$

$$\Gamma_{\underline{\xi}}[\cdot] = \alpha(\xi)[\cdot] + (1 - \alpha(\xi))\partial[\cdot]/\partial n \tag{4}$$

$\hat{\xi}/(\ )$

where n is an exterior normal vector to the boundary ∂D at a point $\xi \in \partial D$ and $\alpha(\xi)$ is a function of class c^2 on ∂D satisfying $0 \leq \alpha(\xi) \leq 1$. $S(k+1, \xi)$ denotes a source function at the boundary and is assumed to be known.

ORIGINAL PAGE IS
OF POOR QUALITY

IEEE: SYST., MAN, CYBERN.
Vol. ______ Issue ______
Au: Omatu
Galley No. 2

Assume that $u_0(x)$ is a Gaussian random function the mean and covariance functions of which are given by

$$E[u_0(x)] = 0 \tag{5}$$

$$E[u_0(x)u_0'(y)] = P_0(x, y) \tag{6}$$

where $E[\cdot]$ and the prime symbol denote the expectation and transpose operators, respectively.

Let the observed data be taken at m points, $x^1, \cdots, x^m \in \bar{D} = D \cup \partial D$ and let an mn-dimensional column vector $u_m(k)$ be defined by

$$u_m(k) = \mathrm{Col}[u(k, x^1), \cdots, u(k, x^m)]. \tag{7}$$

Let the observations be related to the states by

$$z(k) = H(k)u_m(k) + v(k). \tag{8}$$

where $z(k)$ is a p-dimensional observations vector at the m observation points, $x^1, \cdots, x^m \in \bar{D}$, $H(k)$ is a known $pxmn$ matrix, and $v(k)$ is a p-dimensional vector-valued white Gaussian process. Assume that the white Gaussian process $w(k, x)$ in (1) and $v(k)$ in (8) are statistically independent of each other and also independent of the initial condition $u_0(x)$. Their mean and covariance functions are given by

$$E[w(k, x)] = 0, \qquad E[v(k)] = 0 \tag{9}$$

$$E[w(k, x)w'(s, y)] = Q(k, x, y)\delta_{ks}, \qquad x, y \in D \tag{10}$$

$$E[v(k)v'(s)] = R(k)\delta_{ks}, \tag{11}$$

where δ_{ks} is the Kronecker delta function, and $Q(k, x, y)$ and $R(k)$ are symmetric positive-semidefinite and positive-definite matrices, respectively.

III. Description of the Estimation Problems

The general problem considered here is to find an estimate $\hat{u}(\tau, x, k)$ of the state $u(\tau, x)$ at time τ based on the measurement data z_0^k, denoting a family of $z(\sigma)$ from $\sigma = 0$ up to the present time k. Specifically, for $\tau > k$ we have the prediction problem, for $\tau = k$ the filtering problem, and for $\tau < k$ the smoothing problem. As in the Kalman–Bucy approach, an estimate $\hat{u}(\tau, x/k)$ of $u(\tau, x)$ is sought through a linear operation on the past and present observation values z_0^k as follows:

$$\hat{u}(\tau, x/k) = \sum_{\sigma=0}^{k} \tilde{F}(\tau, x, \sigma)z(\sigma) \tag{12}$$

where $\tilde{F}(\tau, x, \sigma)$ is an nxp matrix kernel function.

To differentiate between the prediction, filtering, and smoothing problems, we replace (12) with different notation for each problem:

1) Prediction ($\tau > k$)

$$\hat{u}(\tau, x/k) = \sum_{\sigma=0}^{k} A(\tau, x, \sigma)z(\sigma). \tag{13}$$

2) Filtering ($\tau = k$)

$$\hat{u}(k, x/k) = \sum_{\sigma=0}^{k} F(\tau, x, \sigma)z(\sigma). \tag{14}$$

3) Smoothing ($\tau < k$)

$$\hat{u}(\tau, x/k) = \sum_{\sigma=0}^{k} B(\tau, k, x, \sigma)z(\sigma). \tag{15}$$

The estimation error is denoted by $\tilde{u}(\tau, x/k)$,

$$\tilde{u}(\tau, x/k) = u(\tau, x) - \hat{u}(\tau, x/y). \tag{16}$$

The estimate $\hat{u}(\tau, x/k)$ that minimizes

$$J(\hat{u}) = E[\|\tilde{u}(\tau, x/k)\|^2] \tag{17}$$

is said to be optimal, where $\|\cdot\|$ denotes the Euclidian norm.

IEEE: SYST., MAN, CYBERN.
Vol. ______ Issue ______
Au: Omatu
Galley No. 3

Theorem 1 (Wiener–Hopf theory): A necessary and sufficient condition for the estimate $\hat{u}(\tau, x/k)$ to be optimal is that the following Wiener–Hopf equation holds for $\alpha = 0, 1, \cdots, k$ and $x \in \bar{D}$.

$$\sum_{\sigma=0}^{k} \hat{F}(\tau, x, \sigma) E[z(\sigma) z'(\alpha)] = E[u(\tau, x) z'(\alpha)]. \tag{18}$$

Furthermore, (18) is equivalent to

$$E[\tilde{u}(\tau, x/k) z'(\alpha)] = 0 \tag{19}$$

for $\alpha = 0, 1, \cdots, k$ and $x \in \bar{D}$.

Proof: Let $F_\Delta(\tau, x, \sigma)$ be an $n \times p$ matrix function and let ϵ be a scalar-valued parameter. The trace of the covariance of the estimate,

$$\hat{u}_\epsilon(\tau, x/k) = \sum_{\sigma=0}^{k} \left(\hat{F}(\tau, x, \sigma) + \epsilon F_\Delta(\tau, x, \sigma)\right) z(\sigma)$$

is given by

$$J(\hat{u}_\epsilon) = E\left[\left\| u(\tau, x) - \hat{u}(\tau, x/k) - \epsilon \sum_{\sigma=0}^{k} F_\Delta(\tau, x, \sigma) z(\sigma) \right\|^2\right]$$

$$= E\left[\|\tilde{u}(\tau, x/k)\|^2\right] - 2\epsilon E\left[\tilde{u}'(\tau, x/k) \sum_{\sigma=0}^{k} F_\Delta(\tau, x, \sigma) z(\sigma)\right]$$

$$+ \epsilon^2 E\left[\left\| \sum_{\sigma=0}^{k} F_\Delta(\tau, x, \sigma) z(\sigma) \right\|^2\right].$$

A necessary and sufficient condition for $\hat{u}(\tau, x/k)$ to be optimal is that

$$\left.\frac{\partial J(\hat{u}_\epsilon)}{\partial \epsilon}\right|_{\epsilon=0} = 0,$$

that is,

$$E\left[\tilde{u}'(\tau, x/k) \sum_{\sigma=0}^{k} F_\Delta(\tau, x, \sigma) z(\sigma)\right] = 0$$

for any $n \times p$ matrix $F_\Delta(\tau, x, \sigma)$. Using the relation between the trace and inner product yields

$$E\left[\tilde{u}'(\tau, x/k) \sum_{\sigma=0}^{k} F_\Delta(\tau, x, \sigma) z(\sigma)\right]$$

$$= \operatorname{tr}\left[E\left[\tilde{u}(\tau, x/k) \sum_{\sigma=0}^{k} z'(\sigma) F'_\Delta(\tau, x, \sigma)\right]\right]$$

$$= \sum_{\sigma=0}^{k} \operatorname{tr}\left[E[\tilde{u}(\tau, x/k) z'(\sigma)] F'_\Delta(\tau, x, \sigma)\right] = 0.$$

Setting $F_\Delta(\tau, x, k) = E[\tilde{u}(\tau, x/k) z'(\sigma)]$ in the above equation, it follows that (19) is a necessary condition for $\hat{u}(\tau, x/k)$ to be optimal. Sufficiency of (19) also follows from the above equation. Q.E.D.

Corollary 1: (Orthogonal projection lemma). The following orthogonality condition holds,

$$E[\tilde{u}(\tau, x/k) \hat{u}'(\zeta, y/k)] = 0, \qquad x, y \in \bar{D} \tag{20}$$

where ζ is any time instant, for example, $\zeta < k$, $\zeta = k$ or $\zeta > k$.

IEEE: SYST., MAN, CYBERN.
Vol. ______ Issue ______
Au: Omatu
Galley No. 4

Proof: Multiplying each side of (19) by $\hat{F}(\zeta, y, \alpha)$ and summing from $\sigma = 0$ to $\sigma = k$ yields

$$E\left[\tilde{u}(\tau, x/k)\sum_{\alpha=0}^{k} z'(\alpha)\hat{F}'(\zeta, y, \alpha)\right] = 0.$$

Substituting (12) into the above equation yields (20).

Q.E.D.

Then the following lemma can be proved.

Lemma 1: (Uniqueness of the optimal kernel). Let $\hat{F}(\tau, x, \sigma)$ and $\hat{F}(\tau, x, \sigma) + N(\tau, x, \sigma)$ be optimal matrix kernel functions satisfying the Wiener–Hopf equation (18). Then it follows that

$$N(\tau, x, \sigma) \equiv 0, \qquad \sigma = 0, 1, \cdots, k \text{ and } x \in \bar{D}. \tag{21}$$

and the optimal matrix kernel function $\hat{F}(\tau, x, \sigma)$ is unique.

Proof: From (18) we have

$$\sum_{\sigma=0}^{k} \hat{F}(\tau, x, \sigma)E[z(\sigma)z'(\alpha)] = E[u(\tau, x)z'(\alpha)] = \sum_{\sigma=0}^{k} (\hat{F}(\tau, x, \sigma) + N(\tau, x, \sigma))E[(z(\sigma)z'(\alpha)].$$

Thus,

$$\sum_{\sigma=0}^{k} N(\tau, x, \sigma)E[z(\sigma)z'(\alpha)] = 0.$$

Multiplying each side of the above equation by $N'(\tau, x, \alpha)$ and summing from $\alpha = 0$ to $\alpha = k$ yields

$$\sum_{\sigma=0}^{k}\sum_{\alpha=0}^{k} N(\tau, x, \sigma)E[z(\sigma)z'(\alpha)]N'(\tau, x, \alpha) = 0.$$

On the other hand, from (8) and (11) we have

$$E[z(\sigma)z'(\alpha)] = H(\sigma)E[u_m(\sigma)u_m'(\alpha)]H'(\alpha) + R(\sigma)\delta_{\sigma\alpha}.$$

Then it follows that

$$\sum_{\sigma=0}^{k}\sum_{\alpha=0}^{k} N(\tau, x, \sigma)H(\sigma)E[u_m(\sigma)u_m'(\alpha)]H'(\alpha)N'(\tau, x, \alpha) + \sum_{\sigma=0}^{k} N(\tau, x, \sigma)H(\sigma)R(\sigma)H'(\sigma)N'(\tau, x, \sigma) = 0.$$

Since both terms on the right side of the above equation are positive-semidefinite because of the positive-definiteness of $R(\sigma)$, a necessary and sufficient condition for the above equation to hold is $N(\tau, x, \sigma) \equiv 0$, $\sigma = 0, 1, \cdots, k$ and $x \in \bar{D}$. Thus, the proof of the lemma is complete. Q.E.D.

In order to facilitate the derivation of the optimal estimators, we rewrite (18) in terms of the following corollary.

Corollary 2: The Wiener–Hopf equation (18) is rewritten for the prediction, filtering, and smoothing problems as follows.

1) Prediction ($\tau > k$)

$$\sum_{\sigma=0}^{k} A(\tau, x, \sigma)E[z(\sigma)z'(\sigma)] = E[u(\tau, x)z'(\alpha)]. \tag{22}$$

for $\alpha = 0, 1, \cdots, k$ and $x \in \bar{D}$.

2) Filtering ($\tau = k$)

$$\sum_{\sigma=0}^{k} F(k, x, \sigma)E[z(\sigma)z'(\alpha)] = E[u(k, x)z'(\alpha)] \tag{23}$$

for $\alpha = 0, 1, \cdots, k$ and $x \in \bar{D}$.

3) Smoothing ($\tau < k$)

$$\sum_{\sigma=0}^{k} B(\tau, k, x, \sigma)E[z(\sigma)z'(\alpha)] = E[u(\tau, x)z'(\alpha)] \tag{24}$$

for $\alpha = 0, 1, \cdots, k$ and $x \in \bar{D}$.

In what follows, let us denote the estimation error covariance matrix function by $P(\tau, x, y/k)$,

$$P(\tau, x, y/k) = E[\tilde{u}(\tau, x/k)\tilde{u}'(\tau, y/k)]. \tag{25}$$

III. Derivation of the Optimal Predictor

In this section, we derive the optimal prediction estimator by using the Wiener–Hopf theory in the previous section.

Theorem 2: The optimal prediction estimator is given by

$$\hat{u}(k+1, x/k) = \mathcal{L}_x \hat{u}(k, x/k) \tag{26}$$

$$\Gamma_\xi \hat{u}(k+1, \xi/k) = S(k+1, \xi), \qquad \xi \in \partial D. \tag{27}$$

Proof: From (22) and (1) we have

$$\sum_{\sigma=0}^{k} A(k+1, x, \sigma) E[z(\sigma) z'(\alpha)] = \mathcal{L}_x E[u(k, x) z'(\alpha)]$$

since $w(k, x)$ is independent of $z(\alpha)$, $\alpha = 0, 1, \cdots, k$. From the Wiener–Hopf equation (23) for the optimal filtering problem we have

$$\sum_{\sigma=0}^{k} \{A(k+1, x, \sigma) - \mathcal{L}_x F(k, x, \sigma)\} E[z(\sigma) z'(\alpha)] = 0.$$

Defining $\bar{N}(k, x, \sigma)$ by

$$\bar{N}(k+1, x, \sigma) = A(k+1, x, \sigma) - \mathcal{L}_x F(k, x, \sigma),$$

it is clear that $A(k+1, x, \sigma) + \bar{N}(k+1, x, \sigma)$ also satisfies the Wiener–Hopf equation (22). From the uniqueness of $A(k+1, x, \sigma)$ by Lemma 1 it follows that $\bar{N}(k+1, x, \sigma) \equiv 0$, that is,

$$A(k+1, x, \sigma) = \mathcal{L}_x F(k, x, k). \tag{28}$$

Thus, from (13) and (14) we have

$$\hat{u}(k+1, x/k) = \mathcal{L}_x \sum_{\sigma=0}^{k} F(\sigma, x, \sigma) z(\sigma) = \mathcal{L}_x \hat{u}(k, x/k).$$

Since the forms of Γ_ξ and $S(k+1, \xi)$ are known, the predicted estimate $\hat{u}(k+1, \xi/k)$ also satisfies the same boundary condition as (3), $\Gamma_\xi \hat{u}(k+1, \xi/k) = S(k+1, \xi)$, $\xi \in \partial D$. Thus, the proof of the theorem is complete. Q.E.D.

Theorem 3: The optimal prediction error covariance matrix function $P(k+1, x, y/k)$ is given by

$$P(k+1, x, y/k) = \mathcal{L}_x P(k, x, y/k) \mathcal{L}_y' + \bar{Q}(k, x, y), \tag{29}$$

$$\Gamma_\xi P(k+1, \xi, y/k) = 0, \qquad \xi \in \partial D \tag{30}$$

where

$$\bar{Q}(k, x, y) = G(k, x) Q(k, x, y) G'(k, y). \tag{31}$$

Proof: From (1), (16), and (26) it follows that

$$\tilde{u}(k+1, x/k) = \mathcal{L}_x \tilde{u}(k, x/k) + G(k, x) w(k, x) \tag{32}$$

and from (3), (16), and (27),

$$\Gamma_\xi \tilde{u}(k+1, \xi/k) = 0, \qquad \xi \in \partial D. \tag{33}$$

Then we have from (31) $P(k+1, x, y/k) = E[\tilde{u}(k+1, x/k) \tilde{u}'(k+1, y/k)] = \mathcal{L}_x P(k, x, y/k) \mathcal{L}_y' + \bar{Q}(k, x, y)$ and from (33), $E[\Gamma_\xi \tilde{u}(k+1, \xi/k) \tilde{u}'(k+1, y/k)] = \Gamma_\xi P(k+1, \xi, y/k) = 0$. Thus, the proof of the theorem is complete. Q.E.D.

IV. Derivation of the Optimal Filter

Let us derive the optimal filter by using the Wiener–Hopf theorem for the filtering problem. From (23) it follows that

$$\begin{aligned} &F(k+1, x, k+1) E[z(k+1) z'(\alpha)] \\ &\quad + \sum_{\sigma=0}^{k} F(k+1, x, \sigma) E[z(\sigma) z'(\alpha)] \\ &= E[u(k+1, x) z'(\alpha)] \end{aligned} \tag{34}$$

for $\alpha = 0, 1, \cdots, k+1$.

From (1) and the independence of z_0^{k+1} and $w(k+1, x)$, it follows that $E[u(k+1, x)z'(\alpha)] = \mathcal{L}_x E[u(k, x)z'(\alpha)]$. Applying the Wiener–Hopf equation (23) to the right side of the above equation yields

$$E\big[u(k+1, x)z'(\alpha)\big] = \mathcal{L}_x \sum_{\sigma=0}^{k} F(k, x, \sigma)E\big[z(\sigma)z'(\alpha)\big]. \tag{35}$$

Furthermore, from (8) and the whiteness of $v(k+1)$ we have

$$E\big[z(k+1)z'(\alpha)\big] = H(k+1)E\big[u_m(k+1)z'(\alpha)\big].$$

Let us introduce $\mathcal{L}_*[\cdot]$ and $[\cdot]\mathcal{L}_*'$ as follows,

$$\mathcal{L}_*[\cdot] = \begin{bmatrix} \mathcal{L}_{x^1}[\cdot] & & 0 \\ & \ddots & \\ 0 & & \mathcal{L}_{x^m}[\cdot] \end{bmatrix} \tag{36}$$

and

$$[\cdot]\mathcal{L}_*' = (\mathcal{L}_*[\cdot])'. \tag{37}$$

Then from (1) and (7) it follows that

$$u_m(k+1) = \mathcal{L}_* u_m(k) + w_m(k) \tag{38}$$

$$w_m(k) = \mathrm{Col}\big[G(k, x^1)w(k, x^1), \cdots, \cancel{\delta} G(k, x^m)w(k, x^m)\big]. \tag{39}$$ e

Then we have for $\alpha < k+1$, $E[z(k+1)z'(\alpha)] = H(k+1)\mathcal{L}_* E[u_m(k)z'(\alpha)]$. Applying the Wiener–Hopf equation (23) to the right side of the above equation yields

$$E\big[z(k+1)z'(\alpha)\big] = H(k+1)\mathcal{L}_* \sum_{\sigma=0}^{k} F_m(k, \sigma)E\big[z(\sigma)z'(\alpha)\big] \tag{40}$$

where

$$F_m(k, \sigma) = \begin{bmatrix} F(k, x^1, \sigma) \\ F(k, x^m, \sigma) \end{bmatrix}. \tag{41}$$

Substituting (35) and (40) into (34) yields

$$\sum_{\sigma=0}^{k} N_3(k, x, \sigma)E\big[z(\sigma)z'(\alpha)\big] = 0, \qquad \alpha = 0, 1, \cdots, k$$

where

$$N_3(k, x, \sigma) = F(\cancel{k}+1, x, k+1)H(k+1)\mathcal{L}_* F_m(k, \sigma) - \mathcal{L}_x F(k, x, \sigma) + F(k+1, x, \sigma).$$ l.c.k

Since it is clear that $F(k, x, \sigma) + N_3(k, x, \sigma)$ also satisfies the Wiener–Hopf equation (23), it follows from Lemma 1 that $N_3(k, x, \sigma) \equiv 0$. Thus, we have the following lemma.

Lemma 2: The optimal matrix kernel function $F(k, x, \sigma)$ of the filter is given by

$$F(k+1, x, \sigma) = \mathcal{L}_x F(k, x, \sigma) - F(k+1, x, k+1)H(k+1)\mathcal{L}_* F_m(k, \sigma), \quad \sigma = 0, 1, \cdots, k. \tag{42}$$

Theorem 4: The optimal filtering estimate $\hat{u}(k, x/k)$ is given by

$$\hat{u}(k+1, x/k+1) = \mathcal{L}_x \hat{u}(k, x/k) + F(k+1, x, k+1)\nu(k+1) \tag{43}$$

$$\nu(k+1) = z(k+1) - H(k+1)\mathcal{L}_* \hat{u}_m(k/k) \tag{44}$$

$$\hat{u}(0, x/0) = 0 \tag{45}$$

$$\Gamma_\xi \hat{u}(k+1, \xi/k+1) = S(k+1, \xi), \qquad \xi \in \partial D \tag{46}$$

where

$$\hat{u}_m(k/k) = \mathrm{Col}\big[\hat{u}(k, x^1/k), \cdots, \hat{u}(k, x^m/k)\big]. \tag{47}$$

Proof: Using (14) and (42) yields

$$\hat{u}(k+1, x/k+1) = F(k+1, x, k+1)z(k+1) + \mathcal{L}_x \sum_{\sigma=0}^{k} F(k, x, \sigma)z(\sigma) - F(k+1, x, k+1)H(k+1)\mathcal{L}_x \cdot \sum_{\sigma=0}^{k} F_m(k, \sigma)z(\sigma).$$

Again from (14) we have

$$\hat{u}(k+1, x/k+1) = \mathcal{L}_x \hat{u}(k, x/k) + F(k+1, x, k+1)\nu(k+1).$$

Since we have no information at the initial time, it is suitable to assume an initial value of $\hat{u}(k+1, x/k+1)$ as $\hat{u}(0, x/0) = E[u_0(x)] = 0$. Furthermore, since we know the exact forms of Γ_ξ and $S(k+1, \xi)$, the boundary value $\hat{u}(k+1, \xi/k+1)$ also satisfies the same boundary condition as $u(k+1, \xi)$. Thus, we have $\Gamma_\xi \hat{u}(k+1, \xi/k+1) = S(k+1, \xi)$, $\xi \epsilon \partial D$, and the proof the of the theorem is complete. Q.E.D.

ϵ

Note that $\nu(k+1)$ defined by (44) is rewritten by using the prediction value of (26) as follows.

$$\nu(k+1) = z(k+1) - H(k+1)\hat{u}_m(k+1/k) \quad (48)$$

or

$$\nu(k+1) = H(k+1)\tilde{u}_m(k+1/k) + v(k+1) \quad (49)$$

where

$$\hat{u}_m(k+1/k) = \mathrm{Col}\left[\hat{u}(k+1, x^1/k), \cdots, \hat{u}(k+1, x^m/k)\right] \quad (50)$$

and

$$\tilde{u}_m(k+1/k) = u_m(k+1) - \hat{u}_m(k+1/k). \quad (51)$$

$\nu(k+1)$ is termed the innovation process [8], [9].

In order to find the optimal matrix kernel function $F(k+1, x, k+1)$ for the filtering problem, we introduce the following notation.

$$p_m(\tau, x/k) = \left[p(\tau, x, x^1/k), \cdots, p(\tau, x, x^m/k)\right] \quad (52)$$

and

$$p_{mm}(\tau/k) = \begin{bmatrix} p(\tau, x^1/k) \\ \vdots \\ p(\tau, x^m/k) \end{bmatrix} \cdot \begin{bmatrix} p(\tau, x^1, x^1/k), \cdots, p(\tau, x^1, x^m/k) \\ \vdots \quad \ddots \quad \vdots \\ p(\tau, x^m, x^1/k), \cdots, p(\tau, x^m, x^m/k) \end{bmatrix}. \quad (53)$$

Note from the definitions of $p_m(\tau, x/k)$ and $p_{mm}(\tau/k)$ that

$$p_m(\tau, x/k) = E\left[\tilde{u}(\tau, x/k)\tilde{u}_m'(\tau/k)\right] \quad (54)$$

and

$$p_{mm}(\tau/k) = E\left[\tilde{u}_m(\tau/k)\tilde{u}_m'(\tau/k)\right]. \quad (55)$$

Furthermore, we define the covariance matrix of the innovation process $\nu(k+1)$ by $\Gamma(k+1/k)$.

$$\Gamma(k+1/k) = E\left[\nu(k+1)\nu'(k+1)\right]. \quad (56)$$

Then from (49) it follows that

$$\Gamma(k+1/k) = H(k+1)p_{mm}(k+1/k) \cdot H'(k+1) + R(k+1). \quad (57)$$

Then the following theorem holds.

ORIGINAL PAGE IS OF POOR QUALITY

IEEE: SYST., MAN, CYBERN.
Vol. ______ Issue ______
Au: Curta fu
Galley No. 7

Theorem 5: The optimal filtering gain matrix function $F(k+1, x, k+1)$ is given by

$$F(k+1, x, k+1) = p_m(k+1, x/k) \cdot H'(k+1)\Gamma^{-1}(k+1/k) \quad (58)$$

or

$$F(k+1, x, k+1) = p_m(k+1, x/k) \cdot \psi(k+1/k)H'(k+1)R^{-1}(k+1) \quad (59)$$

where

$$\psi(k+1/k) = \left(I + \bar{R}(k+1)p_{mm}(k+1/k)\right)^{-1} \quad (60)$$

and

$$\bar{R}(k+1) = H'(k+1)R^{-1}(k+1)H(k+1). \quad (61)$$

Proof: From the Wiener–Hopf equation (23) it follows that

$$F(k+1, x, k+1)E\left[z(k+1)z'(k+1)\right] + \sum_{\sigma=0}^{k} F(k+1, x, \sigma)E\left[z(\sigma)z'(k+1)\right] = E\left[u(k+1, x)z'(k+1)\right].$$

Substituting (42) into the above equation yields

$$F(k+1, x, k+1)\left(E\left[\left\{z(k+1) - H(k+1)\mathcal{E}_* \cdot \sum_{\sigma=0}^{k} F_m(k, \sigma)z(\sigma)\right\}z'(k+1)\right]\right) = E\left[\left\{u(k+1, x) - \mathcal{E}_x \sum_{\sigma=0}^{k} F(k, x, \sigma)z(\sigma)\right\}z'(k+1)\right].$$

Substituting (14) into the right side of the above equation and using (26) and the orthogonality condition of (20) yields

$$\begin{aligned} E\left[\left\{u(k+1, x) - \mathcal{E}_x \sum_{\sigma=0}^{k} F(k, x, \sigma)z(\sigma)\right\}z'(k+1)\right] &= E\left[\tilde{u}(k+1, x/k)z'(k+1)\right] \\ &= E\left[\tilde{u}(k+1, x/k)\tilde{u}'_m(k+1)\right]H'(k+1) \\ &= p_m(k+1, x/k)H'(k+1). \end{aligned}$$

Using the orthogonality condition of (20) gives

$$\begin{aligned} E\left[\nu(k+1)z'(k+1)\right] &= H(k+1)E\left[\tilde{u}_m(k+1/k)\tilde{u}'_m(k+1)\right]H'(k+1) + R(k+1) \\ &= H(k+1)p_{mm}(k+1/k) \cdot H'(k+1) + R(k+1) \\ &= \Gamma(k+1/k). \end{aligned} \quad (62)$$

Then we have

$$F(k+1, x, k+1)\Gamma(k+1/k) = p_m(k+1, x/k)H'(k+1). \quad (63)$$

Thus (58) is derived. In order to show the equivalence between (58) and (59), we use the following matrix inversion lemma,

$$PH'(HPH' + R)^{-1} = P(I + H'R^{-1}HP)^{-1}H'R^{-1}. \quad (64)$$

From (58) and (64) we have

$$F(k+1, x, k+1) = p_m(k+1, x/k) \cdot \psi(k+1/k)H'(k+1)R^{-1}(k+1).$$

Then (59) is derived, and the proof of the theorem is complete. Q.E.D.

The equation for the optimal filtering error covariance matrix function $p(k+1, x, y/k+1)$ now must be derived.

ORIGINAL PAGE IS
OF POOR QUALITY

IEEE: SYST., MAN, CYBERN.
Vol. ______ Issue ______
Au: Omatu
Galley No. 8

ORIGINAL PAGE IS
OF POOR QUALITY

IEEE: SYST., MAN, CYBERN.
Vol. ______ Issue ______
Au: Omatu
Galley No. 709

Theorem 6 The optimal filtering error covariance matrix function $p(k+1, x, y/k+1)$ is given by

$$p(k+1,x,y/k+1) = p(k+1,x,y/k) - p_m(k+1,x/k)H'(k+1) \cdot \Gamma^{-1}(k+1/k)H(k+1) \cdot p_m'(k+1,y/k) \quad (65)$$

or

$$p(k+1,x,y/k+1) = p(k+1,x,y/k) - p_m(k+1,x/k)\psi(k+1/k) \cdot \bar{R}(k+1)p_m'(k+1,y/k) \quad (66)$$

where

$$p(0,x,y/0) = p_0(x,y) \quad (67)$$

and

$$\Gamma_\xi p(k+1,\xi,y/k+1) = 0, \qquad \xi \in \partial D. \quad (68)$$

Proof From (1) and (43) we have

$$\tilde{u}(k+1,x/k+1) = \tilde{u}(k+1,x/k) - F(k+1,x,k+1)\nu(k+1) \quad (69)$$

and from (3) and (46)

$$\Gamma_\xi \tilde{u}(k+1,\xi/k+1) = 0, \qquad \xi \in \partial D. \quad (70)$$

Using the independence property between $\nu(k+1)$ and $\tilde{u}(k+1,x/k)$ or $\tilde{u}(k+1,y/k)$ yields from (69),

$$\begin{aligned} p(k+1,x,y/k+1) &= E\big[\tilde{u}(k+1,x/k+1) \cdot \tilde{u}'(k+1,y/k+1)\big] \\ &= p(k+1,x,y/k) \\ &\quad + F(k+1,x,k+1)E\big[\nu(k+1) \cdot \nu'(k+1)\big]F'(k+1,y,k+1) \\ &\quad - F(k+1,x,k+1)H(k+1) \cdot E\big[\tilde{u}_m(k+1/k)\tilde{u}'(k+1,y/k)\big] \\ &\quad - E\big[\tilde{u}(k+1,x/k)\tilde{u}_m'(k+1/k)\big] \cdot H'(k+1)F'(k+1,y,k+1). \end{aligned}$$

Using (58) and (63) it follows that

$$\begin{aligned} p(k+1,x,y/k+1) &= p(k+1,x,y/k) - p_m(k+1,x/k)H'(k+1) \cdot F'(k+1,y,k+1) \\ &= p(k+1,x,y/k) - p_m(k+1,x/k)H'(k+1)\Gamma^{-1} \cdot (k+1/k)H(k+1)p_m'(k+1,y/k). \end{aligned}$$

Thus (65) is derived. The equivalence between (65) and (66) is easily shown by using (64). Since the initial value $\tilde{u}(0,x/0)$ of $\tilde{u}(k+1,x/k+1)$ is zero from (45), it is clear that $\tilde{p}(0,x,y/0) = E[\tilde{u}(0,y/0)] = \tilde{p}_0(x,y)$. Multiplying each side of (70) by $\tilde{u}'(k+1,y/k+1)$ and taking the expectation yields $\Gamma_\xi \tilde{p}(k+1,\xi,y/k+1) = 0$, $\xi \in \partial D$. Thus, the proof of the theorem is complete. Q.E.D.

l.c. p

l.c. p

Corollary 3: $\hat{u}_m(k+1/k+1)$ and $p_m(k+1,x/k+1)$ satisfy the following relations,

$$\hat{u}_m(k+1/k+1) = \hat{u}_m(k+1/k) + F_m(k+1,k+1)\nu(k+1) \quad (71)$$

$$F_m(k+1,k+1) = p_{mm}(k+1/k)\psi(k+1/k)H'(k+1)R^{-1}(k+1) \quad (72)$$

or

$$F_m(k+1,k+1) = p_{mm}(k+1/k+1) \cdot H'(k+1)R^{-1}(k+1) \quad (73)$$

$$p_{mm}(k+1/k+1) = p_{mm}(k+1/k) - p_{mm}(k+1/k) \cdot \psi(k+1/k)\bar{R}(k+1)p_{mm}(k+1/k) \quad (74)$$

or

$$p_{mm}(k+1/k+1) = p_{mm}(k+1/k)\psi(k+1/k). \tag{75}$$

Proof: From the definitions (41) and (50) of $F_m(k+1, k+1)$ and $\hat{u}_m(k+1/k)$, it is clear that (71), (72), and (74) hold. From (60) and (74) it follows that

$$\begin{aligned} p_{mm}(k+1/k+1) &= p_{mm}(k+1/k)\psi(k+1/k) \\ &\quad \cdot \left\{\psi^{-1}(k+1/k) - \hat{R}(k+1)\right. \\ &\quad \left. \cdot p_{mm}(k+1/k)\right\} \\ &= p_{mm}(k+1/k)\psi(k+1/k) \\ &\quad \cdot \left\{I + \hat{R}(k+1)p_{mm}(k+1/k)\right. \\ &\quad \left. - \hat{R}(k+1)p_{mm}(k+1/k)\right\} \\ &= p_{mm}(k+1/k)\psi(k+1/k). \end{aligned}$$

Thus, (75) is derived and (73) is clear from (72) and (75). Q.E.D.

The present result corresponds to that of Santis *et al.* [17] which is an abstract form of the filter.

V. Derivation of the Equations for the Optimal Smoothing Estimator

In this section, we derive the basic equations for the optimal smoothing estimator by using the Wiener–Hopf theory.

Lemma 3: The optimal matrix kernel function $B(\tau, k+1, x, \sigma)$ of the smoothing estimator is given by

$$\begin{aligned} B(\tau, k+1, x, \sigma) &= B(\tau, k, x, \sigma) \\ &\quad - B(\tau, k+1, x, k+1)H(k+1)\mathcal{E}_{\bullet}F_m(k,\sigma), \\ \sigma &= 0, 1, \cdots, k. \end{aligned} \tag{76}$$

Proof: From the Wiener–Hopf equation (24) we have

$$\sum_{\sigma=0}^{k+1} B(\tau, k+1, x, \sigma)E\left[z(\sigma)z'(\alpha)\right] = E\left[u(\tau, x)z'(\alpha)\right], \quad \alpha = 0, \cdots, k+1$$

and

$$\sum_{\sigma=0}^{k} B(\tau, k, x, \sigma)E\left[z(\sigma)z'(\alpha)\right] = E\left[u(\tau, x)z'(\alpha)\right], \quad \alpha = 0, \cdots, k.$$

Subtracting the latter equation from the former yields

$$\begin{aligned} B(\tau, k+1, x, k+1)E\left[z(k+1)z'(\alpha)\right] &+ \sum_{\sigma=0}^{k}\left(B(\tau, k+1, x, \sigma)\right. \\ &\left. - B(\tau, k, x, \sigma)\right)E\left[z(\sigma)z'(\alpha)\right] = 0. \end{aligned}$$

From (8) and (23) we have

$$\begin{aligned} E\left[z(k+1)z'(\alpha)\right] &= H(k+1)\mathcal{E}_{\bullet}E\left[u_m(k)z'(\alpha)\right] \\ &= H(k+1)\mathcal{E}_{\bullet}\sum_{\sigma=0}^{k} F_m(k,\sigma) \\ &\quad \cdot E\left[z(\sigma)z'(\alpha)\right]. \end{aligned}$$

Then it follows that

$$\sum_{\sigma=0}^{k} \tilde{N}(\tau, k, x, \sigma)E\left[z(\sigma)z'(\alpha)\right] = 0$$

where

$$\begin{aligned} \tilde{N}(\tau, k, x, \sigma) &= B(\tau, k+1, x, \sigma) - B(\tau, k, x, \sigma) \\ &\quad + B(\tau, k+1, x, k+1)H(k+1)\mathcal{E}_{\bullet}F_m(k,\sigma). \end{aligned}$$

Since it is easily seen that $B(\tau, k, x, \sigma) + \tilde{N}(\tau, k, x, \sigma)$ also satisfies the Wiener–Hopf equation (24), from Lemma 1 we have $\tilde{N}(\tau, k, x, \sigma) \equiv 0$, and the proof of the lemma is complete. Q.E.D.

Theorem 7: The optimal smoothing estimate $\hat{u}(\tau, x/k+1)$ is given by

$$\begin{aligned} \hat{u}(\tau, x/k+1) &= \hat{u}(\tau, x/k) + B(\tau, k+1, x, k+1) \\ &\quad \cdot \nu(k+1) \end{aligned} \tag{77}$$

$$\Gamma_\xi \hat{u}(\tau, \xi/k+1) = S(\tau, \xi), \ \xi \in \partial D \tag{78}$$

$$k = \tau, \tau+1, \cdots.$$

ORIGINAL PAGE IS
OF POOR QUALITY

IEEE: SYST., MAN, CYBERN.
Vol. ______ Issue ______
Au: Omatu
Galley No. 10

Proof: From (15) it follows that

$$\hat{u}(\tau, x/k+1) = B(\tau, k+1, x, k+1) z(k+1) + \sum_{\sigma=0}^{k} B(\tau, k+1, x, \sigma) z(\sigma).$$

Substituting (76) into the above equation yields

$$\hat{u}(\tau, x/k+1) = B(\tau, k+1, x, k+1) \cdot \left(z(k+1) - H(k+1) \mathfrak{L}_* \sum_{\sigma=0}^{k} F_m(k, \sigma) z(\sigma) \right) + \sum_{\sigma=0}^{k} B(\tau, k, x, \sigma) z(\sigma).$$

Substituting (14) and (15) into the above equation yields

$$\hat{u}(\tau, x/k+1) = \hat{u}(\tau, x/k) + B(\tau, k+1, x, k+1) \nu(k+1).$$

Since we have no additional information about the boundary value of $u(\tau, x)$, except for $S(\tau, \xi)$ and the exact form Γ_ξ, we have $\Gamma_\xi \hat{u}(\tau, \xi/k+1) = S(\tau, \xi)$, $\xi \in \partial D$, and the proof of the theorem is complete. Q.E.D.

Theorem 8: The optimal smoothing gain matrix function $B(\tau, k+1, x, k+1)$ is given by

$$B(\tau, k+1, x, k+1) = L_m(\tau, x/k) \mathfrak{L}_*' H' \cdot (k+1) \Gamma^{-1}(k+1/k) \tag{79}$$

or

$$B(\tau, k+1, x, k+1) = J(\tau, x/k+1) \cdot H'(k+1) R^{-1}(k+1) \tag{80}$$

where

$$J(\tau, x/k+1) = L_m(\tau, x/k) \mathfrak{L}_*' \cdot \left(I + \hat{R}(k+1) p_{mm}(k+1/k)\right)^{-1} \tag{81}$$

$$L_m(\tau, x/k) = \left[L(\tau, x, x^1/k), \cdots, L(\tau, x, x^m/k) \right] \tag{82}$$

and

$$L(\tau, x, y/k) = E\left[\tilde{u}(\tau, x/k) \tilde{u}'(k, y/k) \right]. \tag{83}$$

Proof: From the Wiener–Hopf equation (24) it follows that

$$B(\tau, k+1, x, k+1) E\left[z(k+1) z'(k+1) \right] + \sum_{\sigma=0}^{k} B(\tau, k+1, x, \sigma) E\left[z(\sigma) z'(k+1) \right] = E\left[u(\tau, x) z'(k+1) \right].$$

Substituting (76) into the above equation yields

$$B(\tau, k+1, x, k+1) E\left[\nu(k+1) z'(k+1) \right] = E\left[\tilde{u}(\tau, x/k) z'(k+1) \right].$$

On the other hand, from (48) and (49) we have

$$E\left[\nu(k+1) z'(k+1) \right] = E\left[\nu(k+1) (\nu(k+1) + H(k+1) \hat{u}_m(k+1/k))' \right] = E\left[\nu(k+1) \nu'(k+1) \right] = \Gamma(k+1/k). \tag{84}$$

From (8) and the independence of $v(k+1)$ and $\tilde{u}(\tau, x/k)$, we have

$$E\left[\tilde{u}(\tau, x/k) z'(k+1) \right] = E\left[\tilde{u}(\tau, x/k) \tilde{u}_m'(k+1/k) \right] H'(k+1).$$

But from (26) and (38) it follows that

$$\tilde{u}_m(k+1/k) = \mathfrak{L}_* \tilde{u}_m(k/k) + \tilde{w}_m(k). \tag{85}$$

Then we have

$$B(\tau, k+1, x, k+1) \Gamma(k+1/k) = L_m(\tau, x/k) \mathfrak{L}_*' H'(k+1).$$

ORIGINAL PAGE IS OF POOR QUALITY

IEEE: SYST., MAN, CYBERN.
Vol. ______ Issue ______
Au: Omatu
Galley No. 42/11

using the matrix inversion lemma (64). Thus the proof of the theorem is complete. Q.E.D.

Let us now derive the equation for $L(\tau, x, y/k+1)$. Using the orthogonality condition (20) yields

$$L(\tau, x, y/k+1) = E\big[u(\tau, x)\tilde{u}'(k+1, y/k+1)\big].$$

Substituting (69) into the above equation yields

$$L(\tau, x, y/k+1) = L(\tau, x, y/k)\mathfrak{L}'_{\xi}$$
$$-L_m(\tau, x/k)\mathfrak{L}'_{\xi}H'(k+1)F'(k+1, y, k+1). \quad (86)$$

From (3) and (78) it follows that $\Gamma_\xi \tilde{u}(\tau, \xi/k+1) = 0, \xi \in \partial D$. Multiplying each side by $\tilde{u}'(k+1, y/k+1)$ and taking the expectation yields

$$\Gamma_\xi L(\tau, \xi, y/k+1) = 0, \qquad \xi \in \partial D. \quad (87)$$

Then the following theorem holds.

Theorem 9: $J(\tau, x/k+1)$ in (80) is given by

$$J(\tau, x/k+1) = J(\tau, x/k)\mathfrak{L}'_{\xi}\psi(k+1/k) \quad (88)$$
$$J(\tau, x/\tau) = p_m(\tau, x/\tau) \quad (89)$$
$$\Gamma_\xi J(\tau, \xi/k+1) = 0, \qquad \xi \in \partial D. \quad (90)$$

Proof: From (86) and (59) it follows that

$$L_m(\tau, x/k+1) = L_m(\tau, x)\mathfrak{L}'_{\xi}$$
$$\cdot\big\{I - \bar{R}(k+1)\psi'(k+1/k)p_{mm}(k+1/k)\big\}.$$

But we have

$$\begin{aligned} I - \bar{R}(I + P\bar{R})^{-1}P &= \bar{R}(I + P\bar{R})^{-1} \\ &\quad \cdot \big((I + P\bar{R})\bar{R}^{-1} - P\big) \\ &= \big((I + P\bar{R})R^{-1}\big)^{-1}\bar{R}^{-1} \\ &= \big(\bar{R}(\bar{R}^{-1} + P)\big)^{-1} \\ &= (I + \bar{R}P)^{-1}. \end{aligned}$$

Thus,

$$L_m(\tau, x/k+1) = L_m(\tau, x)\mathfrak{L}'_{\xi}$$
$$\cdot\big(I + \bar{R}(k+1)p_{mm}(k+1/k)\big)^{-1}.$$

Therefore, from (81) it follows that

$$J(\tau, x/k+1) = L_m(\tau, x/k+1) \quad (91)$$

and from (81) we have

$$J(\tau, x/k+1) = J(\tau, x/k)\mathfrak{L}'_{\xi}$$
$$\cdot\big(I + \bar{R}(k+1)p_{mm}(k+1/k)\big)^{-1}.$$

Then it follows that

$$J(\tau, x/\tau) = L_m(\tau, x/\tau) = p_m(\tau, x/\tau).$$

Since (90) is clear from (87) and (91), the proof of the theorem is complete. Q.E.D.

Let us now derive the equation for the optimal smoothing error covariance matrix function $p(\tau, x, y/k)$ defined by

$$p(\tau, x, y/k) = E\big[\tilde{u}(\tau, x/k)\tilde{u}(\tau, y/k)\big]. \quad (92)$$

From (77) and (78) it follows that

$$\tilde{u}(\tau, x/k+1) = \tilde{u}(\tau, x/k)$$
$$-B(\tau, k+1, x, k+1)\nu(k+1) \quad (93)$$

$$\Gamma_\xi \tilde{u}(\tau, \xi/k+1) = 0, \qquad \xi \in \partial D. \quad (94)$$

Then the following theorem holds.

Theorem 10: The optimal smoothing error covariance matrix function $p(\tau, x, y/k+1)$ is given by

$$p(\tau, x, y/k+1) = p(\tau, x, y/k)$$
$$-L_m(\tau, x/k)\mathfrak{L}'_{\xi}H'(k+1)\Gamma^{-1}$$
$$\cdot(k+1/k)H(k+1)\mathfrak{L}'_{\xi}L'_m(\tau, y/k) \quad (95)$$

or

$$p(\tau, x, y/k+1) = p(\tau, x, y/k)$$
$$-J(\tau, x/k+1)\psi^{-1}$$
$$\cdot(k+1/k)\bar{R}(k+1)J'(\tau, y/k+1) \quad (96)$$

$$\Gamma_\xi p(\tau, \xi, y/k+1) = 0, \qquad \xi \in \partial D. \quad (97)$$

Proof: From (93) it follows that

$$p(\tau, x, y/k+1) = p(\tau, x, y/k) + B(\tau, k+1, x, k+1)E[\nu(k+1)\nu'(k+1)]B'(\tau, k+1, y, k+1) - B(\tau, k+1, x, k+1)E[\nu(k+1)\tilde{u}'(\tau, y/k)] - E[\tilde{u}(\tau, x/k)\nu'(k+1)] \cdot B'(\tau, k+1, y, k+1).$$

$$E[\tilde{u}(\tau, x/k)\nu'(k+1)] = E[\tilde{u}(\tau, x/k)\tilde{u}_m'(k/k)] \cdot \mathfrak{L}_*' H'(k+1) = L_m(\tau, x/k)\mathfrak{L}_*' H'(k+1)$$

and

$$E[\nu(k+1)\tilde{u}'(\tau, y/k)] = H(k+1)\mathfrak{L}_* L_m'(\tau, y/k).$$

Thus, we have

$$p(\tau, x, y/k+1) = p(\tau, x, y/k) + B(\tau, k+1, x, k+1) \cdot \Gamma(k+1/k)B'(\tau, k+1, y, k+1) - B(\tau, k+1, x, k+1)H(k+1) \cdot \mathfrak{L}_* L_m'(\tau, y/k) - L_m(\tau, x/k) \cdot \mathfrak{L}_*' H'(k+1)B'(\tau, k+1, y, k+1).$$

Substituting (79) into the above equation yields

$$p(\tau, x, y/k+1) = p(\tau, x, y/k) - L_m(\tau, x/k)\mathfrak{L}_*' H' (k+1)\Gamma^{-1}(k+1/k)H(k+1)\mathfrak{L}_* L_m'(\tau, y/k).$$

In order to derive (96), note that from (81),

$$L_m(\tau, x/k)\mathfrak{L}_*' = J(\tau, x/k+1)\psi^{-1}(k+1/k)$$

and from the matrix inversion lemma (64),

$$H'(HPH' + R)^{-1}H = (I + H'R^{-1}HP)^{-1}H'R^{-1}H.$$

Then we have

$$H'(k+1)\Gamma^{-1}(k+1/k)H(k+1) = \psi(k+1/k)\tilde{R}(k+1)$$

and

$$p(\tau, x, y/k+1) = p(\tau, x, y/k) - J(\tau, x/k+1) \cdot \psi^{-1}(k+1/k)\tilde{R}(k+1)J'(\tau, y/k+1).$$

Multiplying each side of (94) by $\tilde{u}(\tau, y/k+1)$ and taking the expectation yields $\Gamma_\xi' p(\tau, \xi, y/k+1) = 0$, $\xi \in \partial D$. Thus, the proof of the theorem is complete. Q.E.D.

Corollary 4: $J(\tau, x/k)$ satisfies the following relations.

$$J(\tau, x/k+1) = A(\tau, x)J_m(\tau+1/k+1) \quad (98)$$

and

$$J(\tau+1, x/k) = D(\tau, x)J_m(\tau/k) \quad (99)$$

where

$$J_m(\tau/k) = \begin{bmatrix} J(\tau, x^1/k) \\ \vdots \\ J(\tau, x^m/k) \end{bmatrix} \quad (100)$$

$$A(\tau, x) = p_m(\tau, x/\tau)\mathfrak{L}_*' p_{mm}^{-1}(\tau+1/\tau) \quad (101)$$

$$D(\tau, x) = p_m(\tau+1, x/\tau)\left(p_{mm}(\tau/\tau)\mathfrak{L}_*'\right)^{-1}. \quad (102)$$

Proof: Letting $\Phi(k+1)$ be given by $\Phi(k+1) = \mathfrak{L}_*'(I + \tilde{R}(k+1)p_{mm}(k+1/k))^{-1}$, from (88) and (89) it follows that $J(\tau, x/k+1) = p_m(\tau, x/\tau)\Phi(\tau+1)\Phi(\tau+2)\cdots\Phi(k+1)$ and $J_m(\tau+1/k+1) = p_{mm}(\tau+1/\tau+1)\Phi(\tau+2)\cdots\Phi(k+1)$. From the above equations and (75) we have

$$\begin{aligned} J(\tau, x/k+1) &= p_m(\tau, x/\tau)\Phi(\tau+1) \cdot p_{mm}^{-1}(\tau+1/\tau+1)J_m(\tau+1/k+1) \\ &= p_m(\tau, x/\tau)\mathfrak{L}_*'\psi(\tau+1/\tau)\psi^{-1}(\tau+1/\tau) \cdot p_{mm}^{-1}(\tau+1/\tau)J_m(\tau+1/k+1) \\ &= A(\tau, x)J_m(\tau+1/k+1). \end{aligned}$$

From (88) and (89) it follows that $J(\tau+1, x/k) = p_m(\tau+1, x/\tau+1)\Phi(\tau+2)\cdots\Phi(k)$ and $J_m(\tau/k) = p_{mm}(\tau/\tau)\Phi(\tau+1)\Phi(\tau+2)\cdots\Phi(k)$. Thus, we have from the above equations

ORIGINAL PAGE IS OF POOR QUALITY

IEEE: SYST., MAN, CYBERN.
Vol. ______ Issue ______
Au: Omatu
Galley No. 13

IEEE: SYST., MAN, CYBERN.
Vol. ______ Issue ______
Au: Omatu
Galley No. ~~15~~ 14

$$J(\tau+1, x/k) = p_m(\tau+1, x/\tau+1)\big(I + \tilde{R}(\tau+1) \cdot p_{mm}(\tau+1/\tau)\big)\big(p_{mm}(\tau/\tau)\mathfrak{L}^*_{\bullet}\big)^{-1} J_m(\tau/k)$$
$$= p_m(\tau+1, x/\tau)\big(p_{mm}(\tau/\tau)\mathfrak{L}^*_{\bullet}\big)^{-1} J_m(\tau/k)$$

where the following equality derived from (66) has been used,

$$p_m(\tau+1, x/\tau+1) = p_m(\tau+1, x/\tau) \cdot \big(I + \tilde{R}(\tau+1)p_{mm}(\tau+1/\tau)\big)^{-1}. \quad (103)$$

Thus, the proof of the corollary is complete. Q.E.D.

Theorem 11: The optimal smoothing estimator is given by

$$\hat{u}(\tau, x/k) = \hat{u}(\tau, x/\tau) + \sum_{l=\tau+1}^{k} J(\tau, x/l)\bar{\nu}(l) \quad (104)$$

$$\Gamma_\xi \hat{u}(\tau, \xi/k) = S(\tau, \xi), \qquad \xi \in \partial D \quad (105)$$

where

$$\bar{\nu}(l) = H'(l)R^{-1}(l)\nu(l). \quad (106)$$

Furthermore, the optimal smoothing error covariance matrix function $p(\tau, x, y/k)$ is given by

$$p(\tau, x, y/k) = p(\tau, x, y/\tau) - \sum_{l=\tau+1}^{k} J(\tau, x/l)\psi^{-1}(l/l-1)\tilde{R}(l)J'(\tau, y/l) \quad (107)$$

$$\Gamma_\xi p(\tau, \xi, y/k) = 0, \qquad \xi \in \partial D. \quad (108)$$

Proof: From (77) and (80), (104) can be directly obtained and from (96), (107) is clear. Thus, the proof of the theorem is complete. Q.E.D.

VI. Summary of the Optimal Smoothing Estimators

A. Fixed-Point Smoother (τ = fixed, $k = \tau+1, \tau+2, \cdots$)

Theorem 12: The optimal fixed-point smoothing estimator is given by

$$\hat{u}(\tau, x/k+1) = \hat{u}(\tau, x/k) + J(\tau, x/k+1)\bar{\nu}(k+1) \quad (109)$$

$$J(\tau, x/k+1) = J(\tau, x/k)\mathfrak{L}^*_{\bullet}\psi(k+1/k) \quad (110)$$

$$\psi(k+1/k) = \big(I + \tilde{R}(k+1)p_{mm}(k+1)\big)^{-1} \quad (111)$$

$$J(\tau, x/\tau) = p_m(\tau, x/\tau) \quad (112)$$

$$\Gamma_\xi \hat{u}(\tau, \xi/k+1) = S(\tau, \xi), \qquad \xi \in \partial D \quad (113)$$

$$\Gamma_\xi J(\xi, x/k+1) = 0, \qquad \xi \in \partial D. \quad (114)$$

Furthermore, the optimal fixed-point smoothing error covariance matrix function $p(\tau, x, y/k+1)$ is given by

$$p(\tau, x, y/k+1) = p(\tau, x, y/k) - J(\tau, x/k+1)\psi^{-1}(k+1/k)\tilde{R}(k+1)J'(\tau, y/k+1) \quad (115)$$

$$\Gamma_\xi p(\tau, \xi, y/k+1) = 0, \qquad \xi \in \partial D. \quad (116)$$

B. Fixed-Interval Smoothing Estimator (k = fixed, $\tau = k-1, k-2, \cdots$)

From Theorem 11 it follows that

$$\hat{u}(\tau+1, x/k) = \hat{u}(\tau+1, x/\tau+1) + \sum_{l=\tau+2}^{k} J(\tau+1, x/l)\bar{\nu}(l) \quad (117)$$

and

$$p(\tau+1, x, y/k) = p(\tau+1, x, y/\tau+1) - \sum_{l=\tau+2}^{k} J(\tau+1, x/l)\psi^{-1}(l/l-1) \cdot \tilde{R}(l)J'(\tau+1, y/l). \quad (118)$$

Then the following theorem holds.

IEEE: SYST., MAN, CYBERN.
Vol. ______ Issue ______
Au: ______
Galley No. ______

Theorem 13: The optimal fixed-interval smoothing estimator is given by

$$\hat{u}(\tau+1, x/k) = \hat{u}(\tau+1, x/\tau+1)$$
$$+A(\tau+1, x)\left[\hat{u}_m(\tau+2/k) - \hat{u}_m(\tau+2/\tau+1)\right] \quad (119)$$
$$\Gamma_\xi \hat{u}(\tau+1, \xi/k) = S(\tau+1, \xi), \quad \xi \in \partial D. \quad (120)$$

$\in /$

Furthermore, the optimal fixed-interval smoothing error covariance matrix function is given by

$$p(\tau+1, x, y/k) = p(\tau+1, x, y/\tau+1)$$
$$-A(\tau+1, x)\left(p_{mm}(\tau+1/k) - p_{mm}(\tau+1/\tau)\right)A'(\tau+1, y) \quad (121)$$
$$\Gamma_\xi p(\tau+1, \xi, y/k) = 0, \quad \xi \in \partial D. \quad (122)$$

Proof. From (98) and (117) we have

$$\hat{u}(\tau+1, x/k) = \hat{u}(\tau+1, x/\tau+1) + A(\tau+1, x) \sum_{l=\tau+2}^{k} J_m(\tau+1/l)\tilde{\nu}(l).$$

But from Theorem 11,

$$\hat{u}(\tau+2, x/k) = \hat{u}(\tau+2, x/\tau+2) + \sum_{l=\tau+3}^{k} J(\tau+2, x/l)\tilde{\nu}(l)$$

and from (43) and (59),

$$\hat{u}(\tau+2, x/\tau+2) = \hat{u}(\tau+2, x/\tau+1) + F(\tau+2, x, \tau+2)\nu(\tau+2)$$
$$= \hat{u}(\tau+2, x/\tau+1) + J(\tau+2, x/\tau+2)\tilde{\nu}(\tau+2).$$

Thus, we have

$$\hat{u}_m(\tau+2/k) - \hat{u}_m(\tau+2/\tau+1) = \sum_{l=\tau+2}^{k} J_m(\tau+2/l)\tilde{\nu}(l).$$

Then we have

$$\hat{u}(\tau+1, x/k) = \hat{u}(\tau+1, x/\tau+1) + A(\tau+1, x) \cdot\left[\hat{u}_m(\tau+2/k) - \hat{u}_m(\tau+2/\tau+1)\right].$$

From (98) and (118),

$$p(\tau+1, x, y/k) = p(\tau+1, x, y/\tau+1) - A(\tau+1, x) \sum_{l=\tau+2}^{k} J_m(\tau+2/l)\psi^{-1}(l/l-1)\tilde{R}(l)J_m'(\tau+2/l) \cdot A'(\tau+1, y).$$

From Theorem 11,

$$p(\tau+2, x, y/k) = p(\tau+2, x, y/\tau+2) - \sum_{l=\tau+3}^{k} J(\tau+2, x/l)\psi^{-1}(l/l-1)\tilde{R}(l)J'(\tau+2, y/l)$$

and from (66).

$$p(\tau + 2, x, y/\tau + 2) = p(\tau + 2, x, y/\tau + 1) - p_{m}(\tau + 2, x/\tau + 1)\psi(\tau + 2/\tau + 1) \cdot \bar{R}(\tau + 2)p_{m}'(\tau + 2, y/\tau + 1).$$

Taking into consideration that, from(103), $J(\tau + 2, x/\tau + 2) = p_{m}(\tau + 2, x/\tau + 1)\psi(\tau + 2/\tau + 1)$ and

$$p_{mm}(\tau + 2/k) - p_{mm}(\tau + 2/\tau + 1) = -\sum_{l=\tau+2}^{k} J_m(\tau + 2/l)\psi^{-1}(l/l - 1)\bar{R}(l)J_m'(\tau + 2/l),$$

we have $p(\tau + 1, x, y/k) = p(\tau + 1, x, y/\tau + 1) - A(\tau + 1, x)[p_{mm}(\tau + 2/k) - p_{mm}(\tau + 2/\tau + 1)]A'(\tau + 1, y)$. Since the boundary conditions (120) and (122) are clear from (105) and (108), respectively, the proof of the theorem is complete. Q.E.D.

C. Fixed-lag Smoothing Estimator ($\tau = k + 1, k = k + 1 + \Delta, \Delta = fixed$)

From Theorem 11 we have

$$\hat{u}(k + 1, x/k + 1 + \Delta) = \hat{u}(k + 1, x/k + 1) + \sum_{l=k+2}^{k+1+\Delta} J(k + 1, x/l)\bar{r}(l) \quad (123)$$

$$p(k + 1, x, y/k + 1 + \Delta) = p(k + 1, x, y/k + 1) - \sum_{l=k+2}^{k+1+\Delta} J(k + 1, x/l)\psi^{-1}(l/l - 1)\bar{R}(l)J'(k + 1, y/l). \quad (124)$$

ORIGINAL PAGE IS
OF POOR QUALITY

IEEE: SYST., MAN, CYBERN.
Vol. ______ Issue ______
Au: Omath
Galley No. 17/6

ORIGINAL PAGE IS OF POOR QUALITY

IEEE: SYST., MAN, CYBERN.
Vol. ______ Issue ______
Au: Chen Fu
Galley No. #17

Then the following theorem holds.

Theorem 14. The optimal fixed-lag smoothing estimator is given by

$$\begin{aligned}&\hat{u}(k+1, x/k+1+\Delta)\\ &= \mathfrak{L}_x \hat{u}(k, x/k+\Delta) + C(x, k+1, \Delta)\\ &\cdot F_m(k+1+\Delta/k+1+\Delta)\nu(k+1+\Delta)\\ &+ \hat{Q}_m(k, x)\left(p_{mm}(k/k)\mathfrak{L}_*'\right)^{-1}\\ &\cdot (\hat{u}_m(k/k+\Delta) - \hat{u}_m(k/k)) \qquad (125)\end{aligned}$$

$$\Gamma_\xi \hat{u}(k+1, \xi/k+1+\Delta) = S(k+1, \xi), \qquad \xi \in \partial D \qquad (126)$$

where

$$C(x, k+1, \Delta) = A(k+1, x)A_m(k+1), \cdots, A_m(k+\Delta) \qquad (127)$$

and

$$A_m(k) = \begin{bmatrix} A(k, x^1) \\ \vdots \\ A(\underline{l}, x^m) \end{bmatrix}.$$ k/

Furthermore, the optimal fixed-lag smoothing error covariance matrix function $p(k+1, x, y/k+1+\Delta)$ is given by

$$\begin{aligned}&p(k+1, x, y/k+1+\Delta)\\ &= p(k+1, x, y/k) - C(x, k+1, \Delta)\\ &\cdot F_m(k+1+\Delta/k+1+\Delta)H(k+1+\Delta)\\ &\cdot p_{mm}(k+1+\Delta/k+\Delta)C'(y, k+1, \Delta) - D(k, x)\\ &\cdot \left[p_{mm}(k/k) - p_{mm}(k/k+\Delta)\right]D'(k, y) \qquad (128)\end{aligned}$$

$$\Gamma_\xi p(k+1, \xi, y/k+1+\Delta) = 0, \qquad \xi \in \partial D. \qquad (129)$$

Proof: From (43) and (59) we have

$$\hat{u}(k+1, x/k+1) = \mathfrak{L}_x \hat{u}(k, x/k) + J(k+1, x/k+1)\tilde{\nu}(k+1).$$

From (123) and the above equation it follows that

$$\begin{aligned}\hat{u}(k+1, x/k+1+\Delta) &= \mathfrak{L}_x \hat{u}(k, x/k)\\ &+ \sum_{l=k+1}^{k+\Delta} J(k+1, x/k)\tilde{\nu}(l)\\ &\cdot J(k+1, x/\underline{l}+1+\Delta)\tilde{\nu}(k+1+\Delta).\end{aligned}$$ k/

From (88) it follows that

$$\begin{aligned}J(k+1, x/k+1+\Delta) &= p_m(k+1, x/k+1)\\ &\cdot \mathfrak{L}_*'\psi(k+2/k+1)\\ &\cdot \mathfrak{L}_*'\psi(k+3/k+2)\cdots\\ &\cdot \mathfrak{L}_*'\psi(k+2+\Delta/k+1+\Delta).\end{aligned}$$

IEEE SYST., MAN, CYBERN.
Vol. Issue
Au: Cimatv
Galley No. 77 18

Substituting (75) into the right side yields

$$J(k+1, x/k+1+\Delta)$$
$$= p_m(k+1, x/k+1)\mathcal{E}_*' p_{mm}^{-1}(k+2/k+1)$$
$$\cdot p_{mm}(k+2/k+2)\mathcal{E}_*'\psi(k+3/k+2)\cdots$$
$$\cdot \mathcal{E}_*'\psi(k+2+\Delta/k+1+\Delta).$$

Repeating the same procedure and using (101) yields

$$J(k+1, x/k+1+\Delta)$$
$$= A(k+1, x)A_m(k+2)\cdots A_m(k+\Delta)$$
$$\cdot p_{mm}(k+1+\Delta/k+1+\Delta).$$

Thus we have

$$J(k+1, x/k+1+\Delta)$$
$$\bar{r}(k+1+\Delta) = C(x, k+1, \Delta)p_{mm}(k+1+\Delta/k+1+\Delta). \tag{130}$$

From (99) it follows that

$$\sum_{l=k+1}^{k+\Delta} J(k+1, x/l)\bar{r}(l) = \sum_{l=k+1}^{k+\Delta} p_m(k+1, x/k)$$
$$(p_{mm}(k/k)\mathcal{E}_*')^{-1}J_m(k/l)\underline{\varepsilon}(l).$$

9/

But from (29) we have

$$p_m(k+1, x/k) = \mathcal{E}_x p_m(k, x/k)\mathcal{E}_*' + Q_m(k, x).$$

From (98) and (99) we have

$$J(k, x/l)$$
$$= A(k, x)J_m(k+1/l) = p_m(\tau, x/\tau)\mathcal{E}_*'(p_{mm}(k/k)\mathcal{E}_*')^{-1}J_m(k/l).$$

Then it follows that

$$\sum_{l=k+1}^{k+\Delta} J(k+1, x/l)\bar{r}(l) = \mathcal{E}_x \sum_{l=k+1}^{k+\Delta} J(k, x/l)\bar{r}(l)$$
$$+ \bar{Q}_m(k, x)\sum_{l=k+1}^{k+\Delta}(p_{mm}(k/k)\mathcal{E}_*')^{-1}J_m(k/l)\bar{r}(l)$$

and

$$
\begin{aligned}
&\hat{u}(k+1, x/k+1+\Delta) \\
&= \hat{\mathcal{L}}_x \hat{u}(k, x/k+\Delta) + C(x, k+1, \Delta) \\
&\quad \cdot F_m(k+1+\Delta/k+1+\Delta)\nu(k+1+\Delta) \\
&\quad + \hat{Q}_m(k, x)\left(p_{mm}(k/k)\mathcal{L}_x^*\right)^{-1} \sum_{l=l+1}^{k+\Delta} J_m(k/l)\bar{\nu}(l).
\end{aligned}
$$

But from Theorem 11 we have

$$
\hat{u}_m(k/k+\Delta) - \hat{u}_m(k/k) = \sum_{l=k-1}^{k+\Delta} J_m(k/l)\bar{\nu}(l).
$$

Thus we have (125). From (65) and (124) it follows that

$$
p(k+1, x, y/k+1+\Delta) = p(k+1, x, y/k) - J_1 - J_2
$$

1+ /

where

$$
\begin{aligned}
J_1 &= \sum_{l=k-1}^{k-\Delta} J(k+1, x/l)\psi^{-1}(l/l-1)\tilde{R}(l) \\
&\quad \cdot J'(k+1, y/l) \\
J_2 &= J(k+1, x/k+1+\Delta)\psi^{-1}(k+1+\Delta/k+\Delta) \\
&\quad \cdot \tilde{R}(k+1+\Delta)J'(k+1, y/k+1+\Delta).
\end{aligned}
$$

From (75) and (130) we have

$$
\begin{aligned}
J_2 &= C(x, k+1, \Delta)p_{mm}(k+1+\Delta/k+1+\Delta) \\
&\quad \cdot \tilde{R}(k+1+\Delta)p_{mm}(k+1+\Delta/k+\Delta) \\
&\quad \cdot C'(y, k+1, \Delta) \\
&= C(x, k+1, \Delta)F_m(k+1+\Delta/k+1+\Delta) \\
&\quad \cdot H(k+1+\Delta)p_{mm}(k+1+\Delta/k+\Delta) \\
&\quad \cdot C'(y, k+1, \Delta).
\end{aligned}
$$

Substituting (99) into J_1 yields

$$
\begin{aligned}
J_1 &= D(k, x) \sum_{l=k-1}^{k-\Delta} J_m(k/l)\psi^{-1}(l/l-1) \\
&\quad \cdot \tilde{R}(l)J_m'(k/l)D'(k, y).
\end{aligned}
$$

But from Theorem 11 we have

$$
\begin{aligned}
&p(k, x, y/k+\Delta) - p(k, x, y/k) \\
&= -\sum_{l=k-1}^{k-\Delta} J(k, x/l)\psi^{-1}(l/l-1)\tilde{R}(l)J'(k, y/l)
\end{aligned}
$$

and

$$
\begin{aligned}
&p_{mm}(k/k+\Delta) - p_{mm}(k/k) \\
&= -\sum_{l=k-1}^{k-\Delta} J_m(k/l)\psi^{-1}(l/l-1)\tilde{R}(l)J_m'(k/l).
\end{aligned}
$$

Then we have

$$
\begin{aligned}
&p(k+1, x, y/k+1+\Delta) \\
&= p(k+1, x, y/k) - C(x, k+1, \Delta) \\
&\quad \cdot F_m(k+1+\Delta/k+1+\Delta)H(k+1+\Delta) \\
&\quad \cdot p_{mm}(k+1+\Delta/k+\Delta)C'(y, k+1, \Delta) - D(k, x) \\
&\quad \cdot \left[p_{mm}(k/k) - p_{mm}(k/k+\Delta)\right]D'(k, y).
\end{aligned}
$$

Since the boundary conditions (126) and (129) are clear from (105) and (108), respectively, the proof of the theorem is complete. Q.E.D.

Kelly and Anderson [18] proved that the fixed-lag smoothing algorithm of Theorem 14 may be unstable, but Chirarattananon and Anderson [19] derived a stable version of the algorithm. It is possible to derive a comparable version here, although stability problems should not arise in our use of the algorithm of Theorem 14 as long as it is used over a finite time interval.

IEEE: SYST., MAN, CYBERN.
Vol. ______ Issue ______
Au: Omati
Galley No. 50 / 14

IX. Application to Estimation of Air Pollution

ORIGINAL PAGE IS OF POOR QUALITY

IEEE: SYST., MAN, CYBERN.
Vol. ______ Issue ______
Au: Omatu
Galley No. 20

Distributed parameter estimation theory has recently been applied to simulated air pollution data to demonstrate the capability of estimating atmospheric concentration levels from routine monitoring data [10], [11]. A problem identified in these early studies was how to specify the statistical properties of the assumed system and observation noise. In this section we expand upon the prior studies in two respects. First, we consider actual monitoring data for sulfur dioxide (SO_2), in particular those measured each hour during the period December 1–31, 1975 at four locations in Tokoshima Prefecture, Japan (see Fig. 1). Second, we apply the method of Sage and Husa [12] to estimate the unknown noise covariances in the system equation and measurements.

e/

Fig. 1

Hourly sulfur dioxide data are available at the four locations shown in Fig. 1 for the period December 1–31, 1975. The data for day k at location i may be denoted by $e_k(x^i, t)$. It is useful to average the data for December 1–30 to produce

$$\langle e(x^i, t)\rangle = \frac{1}{30}\sum_{k=1}^{30} e_k(x^i, t) \tag{131}$$

where we will consider December 31 as a day to test the algorithms.

If it can be assumed that the wind flows are such that there are no north-south variations of concentration and that vertical mixing is rapid enough to eliminate variations of concentration with altitude, then the region can be considered to be one-dimensional along the east-west coordinate. The SO_2 concentration at any particular time can be assumed to be described by the atmospheric diffusion equation [13].

$$\frac{\partial c}{\partial t} + \zeta\frac{\partial c}{\partial x} = a\frac{\partial^2 c}{\partial x^2} + S(x, t) \tag{132}$$

∂

where ζ is the wind velocity, a is a diffusion coefficient, and S is the rate of emission of SO_2 as a function of location and time.

tr/

Equation (132) holds at any instant of time, but we desire an equation governing the monthly mean concentration $\langle c\rangle$. Although no such equation exists, we can formally average (132) over the 30 realizations (days) to produce

$$\frac{\partial\langle c\rangle}{\partial t} + \left\langle \zeta\frac{\partial c}{\partial x}\right\rangle = \left\langle a\frac{\partial^2 c}{\partial x^2}\right\rangle + S. \tag{133}$$

One object will be to estimate the diffusion parameter a. This parameter will in general vary with location and time of day, although for simplicity we seek a constant value for the month. Thus, the first term on the right side of (133) becomes a $\partial^2\langle u\rangle/\partial x^2$. We can form the residuals, $u = c - \langle c\rangle$ and $z = e - \langle e\rangle$. By subtracting (133) from (132) we obtain

$$\frac{\partial u}{\partial t} + \zeta\frac{\partial c}{\partial x} - \left\langle \zeta\frac{\partial c}{\partial x}\right\rangle = a\frac{\partial^2 u}{\partial x^2}. \tag{134}$$

Since wind data are not available with which to evaluate the second and third terms on the left side of (134) let us rewrite (134) as

$$\frac{\partial u}{\partial t} = a\frac{\partial^2 u}{\partial x^2} + w(x, t) \tag{135}$$

where $w(x, t)$ includes those unknown features associated with the velocity terms.

The boundary conditions on (132) are

$$\frac{\partial c}{\partial x} = 0, \quad x = 0, 1 \tag{136}$$

expressing the assumption that there is no diffusive flux of SO_2 into or out of the region at the boundaries. After averaging and forming the residual, (136) becomes

$$\frac{\partial u}{\partial x} = 0, \quad x = 0, 1. \tag{137}$$

IEEE: SYST., MAN, CYBERN.
Vol. ______ Issue ______
Au: Amato
Galley No. 21

The problem is now to estimate $u(x, t)$ based on the data,

$$z(x^i, t) = u(x^i, t) + v_i(t), \qquad i = 1,2,3,4. \quad (138)$$

Since hourly data are available, (135) can be cast into the discrete-time form (1),

$$u(k+1, x) = \mathfrak{L}_x u(k, x) + w(x, t) \qquad (139)$$

with $\mathfrak{L}_x = 1 + a\partial^2/\partial x^2$. Observation error is estimated from the mean square error of predicted values and observed data,

$$PA(i) = \frac{1}{24} \sum_{k=1}^{24} \left(z_i(k) - \hat{z}_i(k/k-1)\right)^2, \qquad i = 1,2,3,4 \qquad (140)$$

An index of overall estimation error is

$$J = \sum_{i=1}^{4} PA(i). \qquad (141)$$

To apply discrete-time distributed parameter estimation theory to predict air pollution levels, we must consider three problems. The first problem is how to simulate the distributed parameter system. The second is how to determine the covariances of system and observation noise. The last is how to determine the diffusion coefficient a. For the first problem we use the Fourier expansion method and approximate the original distributed parameter system by a finite-dimensional system. For the second problem, we apply the algorithm of Sage and Husa [12] that necessitates the simultaneous application of the optimal filtering and smoothing algorithms. For the third problem we apply the maximum likelihood approach in the smoothing form [14]. We now consider these problems in more detail.

ital.

Fourier Expansion Method: It is well-known that the state $u(k, x)$ of the distributed parameter system (139) with boundary condition (137) can be represented by using the eigenfunctions $\phi_i(x)$ as follows,

$$u(k, x) = \sum_{i=1}^{\infty} u_i(k)\phi_i(x) \qquad (142)$$

where

$$\mathfrak{L}_x \phi_i(x) = \lambda_i \phi_i(x), \qquad x \in (0, 1)$$

$$\frac{\partial \phi_i(\xi)}{\partial \xi} = 0, \qquad \xi = 0, 1$$

and (143)

Au/Ed: placement of eq. no?

$$\int_0^1 \phi_i(x)\phi_j(x)\, dx = \delta_{ij}.$$

λ_i is the eigenvalue of $\mathfrak{L}_x$ corresponding to $\phi_i(x)$. In this case, it is easily seen that the eigenfunction $\phi_i(x)$ and the eigenvalue λ_i are given by

$$\phi_1(x) = 1, \qquad \phi_i(x) = \sqrt{2}\cos \pi i x, \qquad i = 2, \cdots$$

and

$$\lambda_i = 1 - a\pi^2(i-1)^2, \qquad i = 1, 2, \cdots, \quad (144)$$

Then $\hat{u}(\tau, x/k)$, $p(\tau, x, y/k)$, and $A(\tau, x)$ can be represented as follows:

$$\hat{u}(\tau, x/k) = \sum_{i=1}^{\infty} \hat{u}_{ij}(\tau/k)\phi_i(x)$$

$$p(\tau, x, y/k) = \sum_{i,j=1}^{\infty} \hat{p}_{ij}(\tau/k)\phi_i(x)\phi_j'(y)$$

$$A(\tau, x) = \sum_{i=1}^{\infty} a_i(\tau)\phi_i(x). \qquad (145)$$

Let us approximate these infinite expansions by the first N terms and define the following matrices and vectors,

$$\hat{u}(\tau/k) = \mathrm{Col}\left[\hat{u}_1(\tau/k),\cdots,\hat{u}_N(\tau/k)\right],$$
$$A(\tau) = \mathrm{Col}\left[a_1(\tau),\cdots,a_N(\tau)\right],$$
$$\Lambda = \mathrm{diag}\left[\lambda_1,\cdots,\lambda_N\right],$$
$$P(\tau/k) = \begin{bmatrix} \tilde{p}_{11}(\tau/k),\cdots,\tilde{p}_{1N}(\tau/k) \\ \vdots \quad \vdots \\ \tilde{p}_{N1}(\tau/k),\cdots,\tilde{p}_{NN}(\tau/k) \end{bmatrix}$$
$$Q(k) = \begin{bmatrix} q_{11}(k),\cdots,q_{1N}(k) \\ \vdots \quad \vdots \\ q_{N1}(k),\cdots,q_{NN}(k) \end{bmatrix}$$

and

$$\Phi = \begin{bmatrix} \phi_1(x^1),\cdots,\phi_N(x^1) \\ \vdots \quad \vdots \\ \phi_1(x^m),\cdots,\phi_N(x^m) \end{bmatrix}$$

where $q_{ij}(k)$ denotes the (i, j)th Fourier coefficient of $\bar{Q}(k, x, y)$.

Then, from Theorems 3–5 we have

$$\hat{u}(k+1/k+1) = \Lambda\hat{u}(k/k) + F(k+1)\nu(k+1)$$
$$F(k+1) = P(k+1/k)\Phi'H'(k+1)$$
$$\left[H(k+1)\Phi P(k+1/k)\Phi' H'\right.$$
$$\left.\cdot(k+1) + R(k+1)\right]^{-1},$$
$$P(k+1/k) = \Lambda P(k/k)\Lambda' + Q(k),$$
$$P(k+1/k+1) = (I - F(k+1)H(k+1)\Phi)$$
$$\cdot P(k+1/k). \tag{146}$$

Furthermore, from Theorem 12 we have 13

$$\hat{u}(\tau+1/k) = \hat{u}(\tau+1/k-1) + A(\tau+1)\Phi(\hat{u}(\tau+2/k)$$ τ /
$$-\hat{u}(\tau+2/\tau+1)),$$
$$A(\tau+1) = P(\tau+1/\tau+1)\Lambda P^{-1}(\tau+1/\tau)\Phi^{-1},$$
$$P(\tau+1/k) = P(\tau+1/\tau+1)$$
$$-A(\tau+1)\Phi(P(\tau+1/k)$$
$$-P(\tau+1/\tau))\Phi' A'(\tau+1). \tag{147}$$

Note that the fixed-interval smoothing estimator does not depend on the matrix Φ which reflects the effect of sensor location.

Determination of the Noise Covariances: In order to determine the unknown covariance matrices of the system and observation noises, we adopt Sage and Husa's algorithm [12] given by

$$\hat{Q}(k) = \frac{1}{k}\sum_{j=1}^{k}(\hat{u}(j/k) - \Lambda\hat{u}(j-1/k))$$
$$\cdot(\hat{u}(j/k) - \Lambda\hat{u}(j-1/k))' \tag{148}$$

and

$$\hat{R}(k) = \frac{1}{k}\sum_{j=1}^{k}(z(j) - H(j)\Phi\hat{u}(j/k))$$
$$\cdot(z(j) - H(j)\Phi\hat{u}(j/k))' \tag{149}$$

where $\hat{Q}(k)$ and $\hat{R}(k)$ denotes the estimated values of $Q(k)$ and $R(k)$, respectively. Note that in these identification algorithm the fixed-interval smoothing estimate $\hat{u}(j/k)$ is used. e this

Identification of the unknown Parameter a: To determine the unknown parameter a we use the maximum likelihood approach in smoothing form [14]. The log-likelihood function $\gamma(k; a)$ is given from [14] by cap U

$$\gamma(k;a) = \frac{1}{2}(\gamma_{\mathrm{bias}} + \gamma_{\mathrm{obs}}) \tag{150}$$

ORIGINAL PAGE IS
OF POOR QUALITY

IEEE: SYST., MAN, CYBERN.
Vol. ______ Issue ______
Au: Omati
Galley No. ______

ORIGINAL PAGE IS OF POOR QUALITY

IEEE: SYST., MAN, CYBERN.
Vol. ______ Issue ______
Au: Omatu
Galley No. 22

where

$$\gamma_{\text{bias}} = -kp \ln(2\pi) - \sum_{j=1}^{k} \ln \det \Sigma_r(j/j-1; a)$$

$$\gamma_{\text{obs}} = -\sum_{j=1}^{k} \{ \nu'(j; a) R^{-1} \nu(j; a) + (\hat{u}(j/k, a)$$

$$-\hat{u}(j-1/k, a))' Q^{-1}$$

$$\cdot (\hat{u}(j/k, a) - \hat{u}(j-1/k, a)) \}$$

$$\nu(j; a) = z(j) - H(j) \hat{u}(j/j-1, a)$$ Φ

$$\Sigma_r(j/j-1, a) = E[\nu(j; a) \nu'(j; a)].$$

where p is the dimension of $z(k)$, and $\hat{u}(j/k-1, a)$ denotes $\hat{u}(j/k-1)$ under the condition that the unknown parameter is assumed to be a.

To maximize $\gamma(k, a)$ we use the following gradient method.

$$a_{i+1} = a_i + G(i) \gamma_g(k; a_i)$$

$$\gamma_g(k, a_i) = \frac{\partial \gamma(k; a)}{\partial a}\bigg|_{a=a_i} \tag{151}$$

where $G(i)$ is a suitable matrix. Therefore, we adopt the following recursive algorithm to identify the unknown parameters Q, R, and a:

1) Make an initial guess a_0 of a.
2) Compute $\hat{Q}(a_i)$ and $\hat{R}(a_i)$ by using (148) and (149).
3) Compute $\hat{a}$ by using (150).
4) compute $\hat{Q}(\hat{a})$ and $\hat{R}(\hat{a})$ by using (148) and (149).
5) Return three by changing i to $i-1$ and repeat until these values do not change. tr

Numerical Results: We use the observed data from December 1–30 to identify the unknown parameter a and noise covariances Q and R. After four iterations the algorithm for determining a converged to the value $\hat{a} = 0.001$. The Fourier expansion has been truncated at $N = 4$. The estimated diagonal elements of noise covariance matrices are

$$Q_{11} = 6.44 \qquad R_{11} = 0.29$$
$$Q_{22} = 1.40 \qquad R_{22} = 0.61$$
$$Q_{33} = 5.75 \qquad R_{33} = 1.96$$
$$Q_{44} = 3.68 \qquad R_{44} = 1.34.$$

To consider the effect of the number and location of monitoring stations, we assume that we have data at only one monitoring station. In this case from the previous results of Kumar and Seinfeld [15] and Omatu *et al.* [16] we expect that the optimal sensor location is closest to the boundary. Thus, either x^1 or x^4 is the optimal single sensor location among the four monitoring stations, x^1, x^2, x^3, x^4. In Table I we show the values of $PA(i)$ and J for several monitoring stations. We see that Aizumi or Matsushige is optimal for the one-point sensor location case. Similar conclusions hold for two or three monitoring stations. Finally, we illustrate the actual observation data and one-hour ahead predicted values for December 31 in Figs. 2–5 for Aizumi, Kitajima, Kawauchi, and Matsushige, respectively. b/

Table I, Figs. 2–5 ε/

Comparison with Other Approaches: It is of interest to compare results of the present filtering and smoothing approaches with others available for air pollution estimation. We consider, therefore, the same SO_2 estimation problem by the following methods: 1) AR-model, 2) persistence, and 3) weighted ensemble.

The AR-model method is based on the following $AR(p)$ model

$$u_k^{(i)} = a_1 u_{k-1}^{(i)} + a_2 u_{k-2}^{(i)} + \cdots + a_p u_{k-p}^{(i)} + e_k^{(i)},$$
$$i = 1, 2, 3, 4 \tag{152}$$

where the $u_k^{(i)}$'s are the concentration levels at time k and at monitoring station x^i, $a_1, a_2, \cdots, a_p$ are the corresponding AR-parameters, and the $e_k^{(i)}$'s are residuals. We used the Levinson algorithm to determine the AR-parameters, while the optimal order p of the AR-process is determined by using the minimum Akaike's information criterion (AIC) [20]. Then the one-hour ahead predicted concentration is given by

$$\hat{u}_{k}^{(i)}{}_{k-1} = a_1 u_{k-1}^{(i)} + \cdots + a_p u_{k-p}^{(i)} \qquad (153)$$

and the prediction error variance is

$$J = \sum_{i=1}^{4} \left\{ \frac{1}{24} \sum_{k=1}^{24} \left(\hat{u}_k^{(i)} - u_k^{(i)}{}_{k-1} \right)^2 \right\} \qquad (154)$$

Table II shows the AR-parameters and minimum AIC value at each monitoring station. Table II

The persistence method consists merely of using the observation data $u_{k-1}^{(i)}$ as the one-hour ahead prediction value $\hat{u}_{k}^{(i)}{}_{k-1}$.

The weighted ensemble method uses the mean of the past observation data at each time k weighted by a linear function of the source strength as the prediction value at time k. Based on the number of emission sources, the weighting functions are assumed here to be 0.15, 0.41, 0.26, and 0.18 at x^1, x^2, x^3, x^4, respectively. Table III shows the performance criteria of the four methods. From Table III we can see that the present method possesses almost the same accuracy as the AR-model method. By multiplying each eigenfunction coefficient by the corresponding eigenfunction and summing them, however, the present method enables us to estimate concentrations over the entire region. Therefore, the present method is more powerful than the AR-model method. Table III

IX. Conclusions

Optimal estimators for discrete-time distributed parameter systems have been derived based on Wiener–Hopf theory. A notable point of the present work is that the smoothing estimators have been derived by the same approach as the filter, thus providing a unified approach for this class of distributed parameter estimation problems. The estimation algorithms have been applied to the problem of predicting atmospheric sulfur dioxide levels in the Tokushima prefecture of Japan.

References

[1] R. F. Curtain, "A survey of infinite dimensional filtering," *SIAM Rev.*, vol. 17, no. 3, pp. 395–411, 1975.

[2] Y. Sawaragi, T. Soeda, and S. Omatu, "Modeling, estimation, and Their Applications for Distributed Parameter Systems," in *Lecture Notes in Control and Information Sciences*, vol. 11. New York: Springer, 1978.

[3] K. E. Bencala and J. H. Seinfeld, "Distributed parameter filtering: boundary noise and discrete observations," *Int. J. Systems Sci.*, vol. 10, no. 5, pp. 493–512, 1979.

[4] J. S. Meditch, "A survey of data smoothing for linear and nonlinear dynamic systems," *Automatica*, vol. 9, no. 2, pp. 151–162, 1973.

[5] S. G. Tzafestas, "Bayesian approach to distributed-parameter filtering and smoothing," *Int. J. Control*, vol. 15, no. 2, pp. 273–295, 1972.

[6] ——, "On optimal distributed-parameter filtering and fixed-interval smoothing for colored noise," *IEEE Trans. Automat. Contr.*, vol. AC-17, no. 4, pp. 448–458, 1972.

[7] H. Nagamine, S. Omatu, and T. Soeda, "The optimal filtering problem for a discrete-time distributed parameter system," *Int. J. Systems Sci.*, vol. 10, no. 7, pp. 735–749, 1979.

[8] T. Kailath, "An innovation approach to least-squares estimation, Part I: Linear filtering in additive white noise," *IEEE Trans. Automat. Contr.*, vol. AC-13, no. 6, pp. 646–655, 1968.

[9] T. Kailath and P. Frost, "An innovation approach to least-squares estimation, Part II: Linear smoothing in additive white noise," *IEEE Trans. Automat. Contr.*, vol. AC-13, no. 6, pp. 655–660, 1968.

[10] A. A. Desalu, L. A. Gould, and F. C. Schweppe, "Dynamic Estimation of Air Pollution," *IEEE Trans. Automat. Contr.*, vol. AC-19, no. 6, pp. 904–910, 1974.

[11] M. Koda and J. H. Seinfeld, "Estimation of urban air pollution," *Automatica*, vol. 14, no. 6, pp. 583–595, 1978. Au: vol., yr and pgs?

IEEE: SYST., MAN, CYBERN.
Vol. ____ Issue 7
Au: ____
Galley No. 21

[12] A. P. Sage and G. W. Husa, "Adaptive filtering with known prior statistics," JACC, pp. 760–769, 1969.
[13] J. H. Seinfeld, Air Pollution: Physical and Chemical Fundamentals. New York: McGraw-Hill, 1975.
[14] F. C. Schweppe, Uncertain Dynamic Systems. Englewood Cliffs, NJ: Prentice-Hall, 1973.
[15] S. Kumar and J. H. Seinfeld, "Optimal location of measurements for distributed parameter estimation," IEEE Trans. Automat. Contr., vol. AC-23, no. 4, pp. 690–698, 1978.
[16] S. Omatu, S. Koide, and T. Soeda, "Optimal sensor location problem for a linear distributed parameter system," IEEE Trans. Automat. Contr., vol. AC-23, no. 4, pp. 665–673, 1978.
[17] R. M. D. Santis, R. Saeks, and L. J. Tung, "Basic optimal estimation and control problems in Hilbert space," Math. Syst. Theory, 12, pp. 175–203, 1978.
[18] C. N. Kelly and B. D. O. Anderson, "On the stability of fixed-lag smoothing algorithms," J. Franklin Inst., vol. 291, no. 4, pp. 271–281, 1971.
[19] S. Chirarattananon and B. D. O. Anderson, "The fixed-lag smoother as a stable finite-dimensional linear system," Automatica, vol. 7, pp. 657–669, 1971.
[20] R. K. Mehra and D. G. Lainiotis, System Identification: Advances and Case Studies. New York: Academic, 1976, pp. 27–96.

Number of sensor locations	Sensor location	J
1	Aizumi (x^1)	39.9
	Kitajima (x^2)	83.1
	Kawauchi (x^3)	46.3
	Matsushige (x^4)	39.1
2	x^1, x^2	26.9
	x^1, x^3	25.8
	x^1, x^4	23.9
	x^2, x^3	41.2
	x^2, x^4	37.[illegible]
	x^3, x^4	29.2
3	x^1, x^2, x^3	22.[illegible]
	x^1, x^2, x^4	22.[illegible]
	x^1, x^3, x^4	23.4
	x^2, x^3, x^4	29.[illegible]
4	x^1, x^2, x^3, x^4	1[illegible]

MAIC (optimal p)	x^1 3.24 ($p = 5$)	x^2 6.8[illegible] ($p = 1$)	x^3 9.8[illegible] ($p = 10$)	x^4 7.52 ($p = 6$)
a_1	−0.8[illegible]	−0.8[illegible]	−0.7[illegible]	−0.8[illegible]
a_2	0.05[illegible]		−0.0[illegible]	−0.0[illegible]
a_3	0.0[illegible]		−0.0[illegible]	−0.04[illegible]
a_4	−0.152		0.0[illegible]	0.01[illegible]
a_5	−0.05[illegible]		0.0[illegible]	−0.079
a_6			0.01[illegible]	0.133
a_7			−0.029	
a_8			0.04[illegible]	
a_9			0.014	
a_{10}			0.069	

Method	x^1	x^2	x^3	x^4	Total
Current	1.02	2.32	4.21	3.08	10.63
AR	0.99	2.36	4.31	3.11	10.77
Persistence	0.92	2.87	6.08	2.71	12.58
Weighted ensemble	2.10	6.55	10.46	5.48	24.59

Fig. 1. Map of Tokushima Prefecture, Japan. The four air pollution monitoring stations shown are located as follows: x^1: Aizumi, x^2: Kitajima, x^3: Kawauchi, x^4: Matsushige. Sources of sulfur dioxide have been lumped according to the three sources sizes indicated by the open circles: ◯ 30–50 m³/h, ○ 10–30 m³/h, • < 10 m³/h.

Fig. 2. Measured and estimated sulfur dioxide concentrations on December 31, 1975 at Aizumi monitoring station (x^1).

Fig. 3. Measured and estimated sulfur dioxide concentrations on December 31, 1975 at Kitajima monitoring station (x^2).

Fig. 4. Measured and estimated sulfur dioxide concentrations on December 31, 1975 at Kawauchi monitoring station (x^3).

Fig. 5. Measured and estimated sulfur dioxide concentrations on December 31, 1975 at Matsushige monitoring station (x^4).

ORIGINAL PAGE IS
OF POOR QUALITY

TABLE I
EFFECT OF THE NUMBER OF OBSERVATION LOCATIONS ON THE OVERALL ERROR J

TABLE II
AR-PARAMETERS AND MINIMUM AIC ($MAIC$)

TABLE III
COMPARISON OF THE FOUR METHODS AT THE FOUR MONITORING SITES FOR ONE-HOUR AHEAD PREDICTED VALUES-PREDICTION ERROR SQUARED

IEEE: SYST., MAN, CYBERN.
Vol. ______ Issue ______
Au: ______
Galley No. ______

Sigeru Omatu was born in Ehime, Japan, on December 16, 1946. He received the B.E. degree in electrical engineering from University of Ehime, Matsuyama, Japan, in 1969 and the M.E. and Ph.D. degrees in electronics engineering from the University of Osaka Prefecture, Osaka, Japan, in 1971 and 1974, respectively.

From 1974 to 1975 he was a Research Associate, from 1975 to 1980 he was a Lecturer, and since 1980 he has been an Associate Professor of Information Science and Systems Engineering at the University of Tokushima, Tokushima, 770, Japan. From November 1980 to February 1981 he was Visiting Associate in Chemical Engineering at the California Institute of Technology, Pasadena. His current interests are in the estimation theory of distributed parameter systems and image data processing, especially remote sensing data analysis.

Dr. Omatu is a member of the Society of Instrument and Control Engineers in Japan, the Institute of Electrical Engineers of Japan, and the Society for Industrial and Applied Mathematics. He received the award of the Society of Instrument and Control Engineers of Japan in 1973. He has co-authored five books in Japanese and two books in English.

John H. Seinfeld was born in Elmira, NY, on August 3, 1942. He received the B.S. degree in chemical engineering from the University of Rochester, Rochester, NY, in 1964 and the Ph.D. degree in chemical engineering from Princeton University, Princeton, NJ, in 1967.

He joined the faculty of the California Institute of Technology, Pasadena, in 1967, where he is currently Louis E. Nohl Professor and Executive Officer for Chemical Engineering. His research interests are in distributed parameter control and estimation and mathematical modeling of air pollution.

Dr. Seinfeld received the 1970 Donald P. Eckman Award of the American Automatic Control Council, the 1976 Allan P. Colburn Award of the American Institute of Chemical Engineers, the 1976 Curtis W. McGraw Research Award of the American Society for Engineering Education, and a 198[illegible] NASA Public Service Medal.

Ed query no photos?

Manuscript received February 3, 1981; revised August 3, 1981. This work was supported by NASA Research Grant NAG1-71.

S. Omatu was with the Department of Chemical Engineering, California Institute of Technology, Pasadena, CA, on leave from the Department of Information Science and Systems Engineering, University of Tokushima, Tokushima, Japan.

J. H. Seinfeld is with the Department of Chemical Engineering, California Institute of Technology, Pasadena, CA 91125.

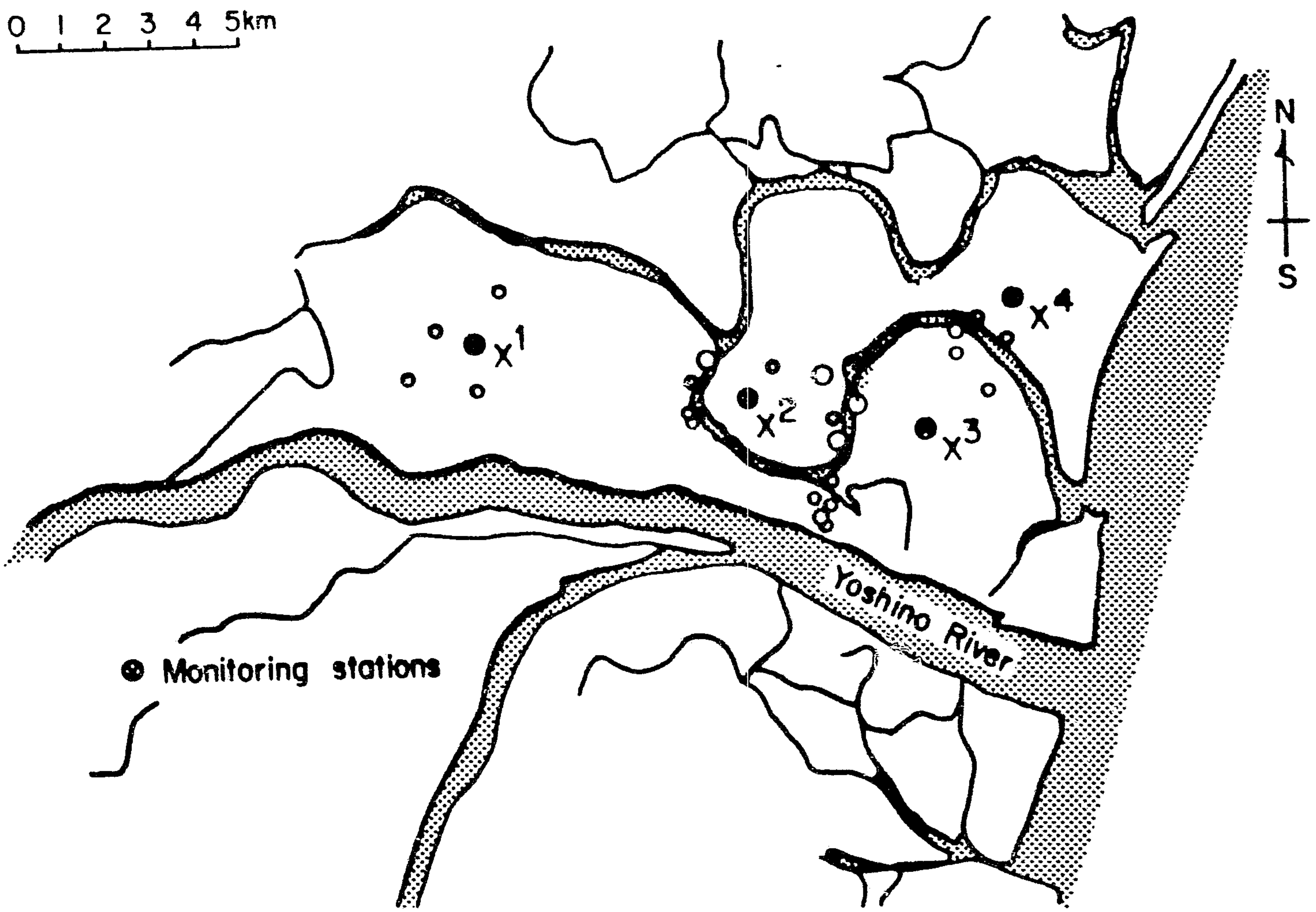


Figure 1

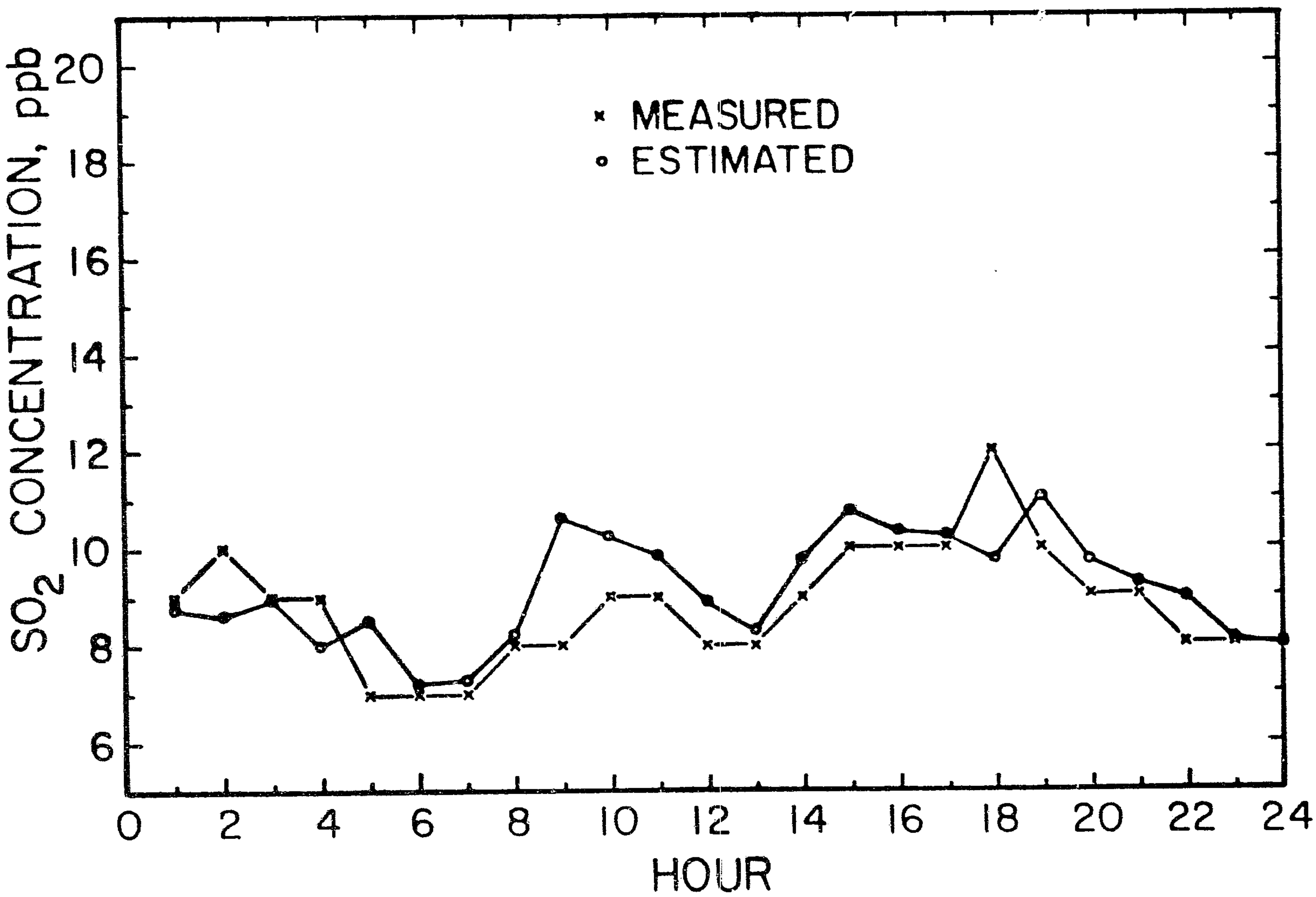

MEASURED
ESTIMATED
SO2 CONCENTRATION, ppb
20
18
16
14
12
10
8
6
0
2
4
6
8
10
12
14
16
18
20
22
24
HOUR

Figure 2

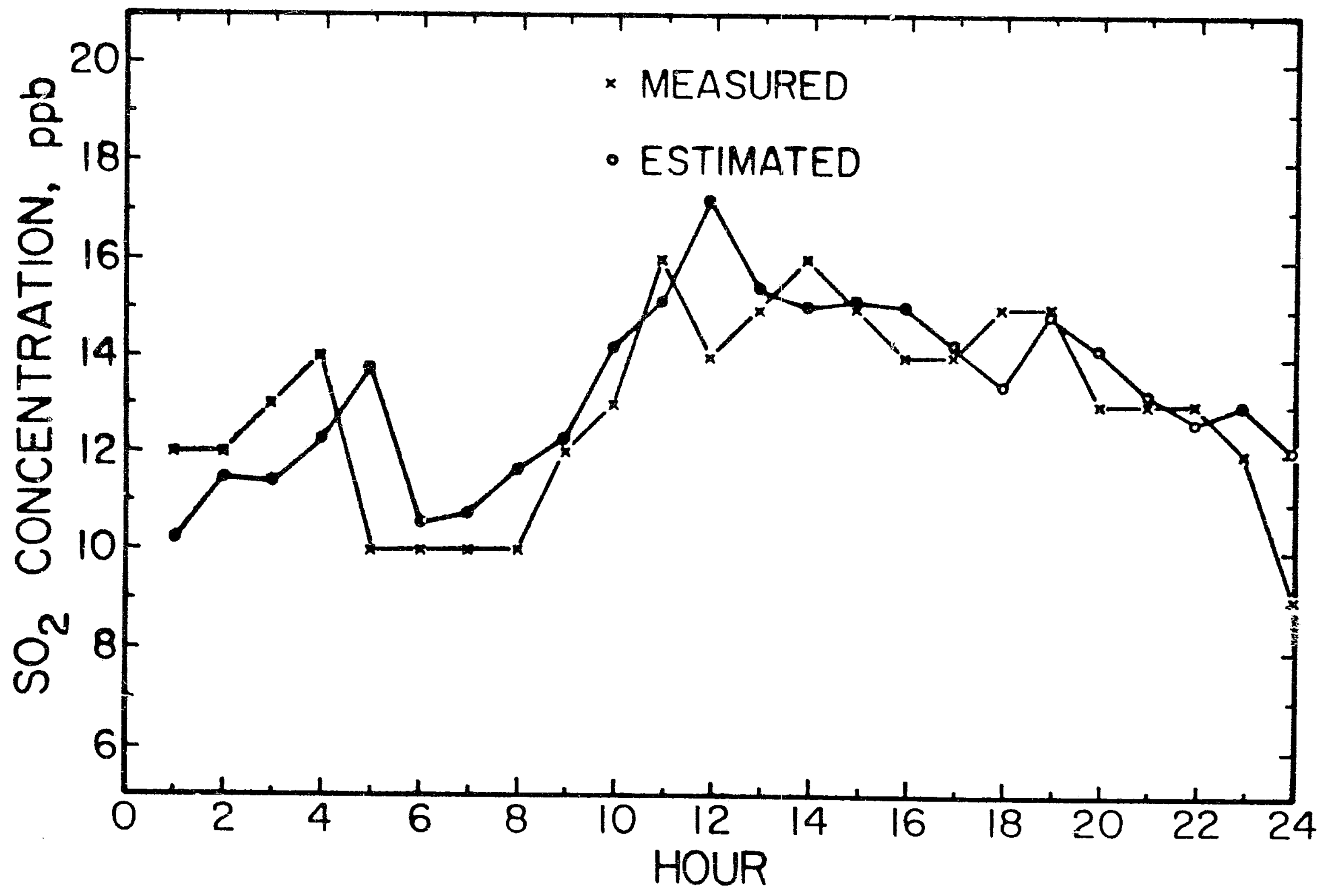

× MEASURED
∘ ESTIMATED
SO_2 CONCENTRATION, ppb
20
18
16
14
12
10
8
6
0
2
4
6
8
10
12
14
16
18
20
22
24
HOUR

Figure 3

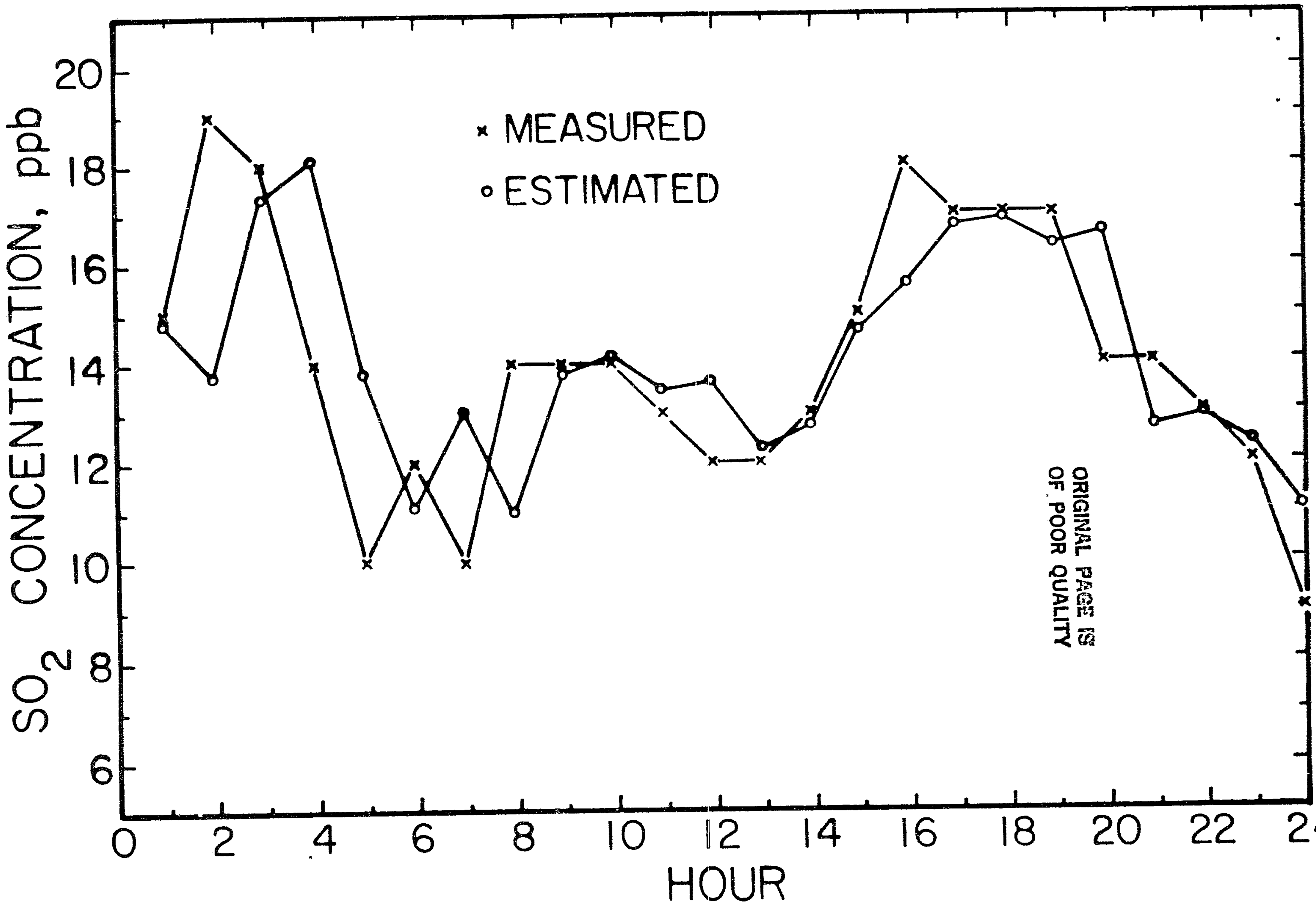

× MEASURED
○ ESTIMATED
SO2 CONCENTRATION, ppb
20
18
16
14
12
10
8
6
0
2
4
6
8
10
12
14
16
18
20
22
24
HOUR

Figure 4

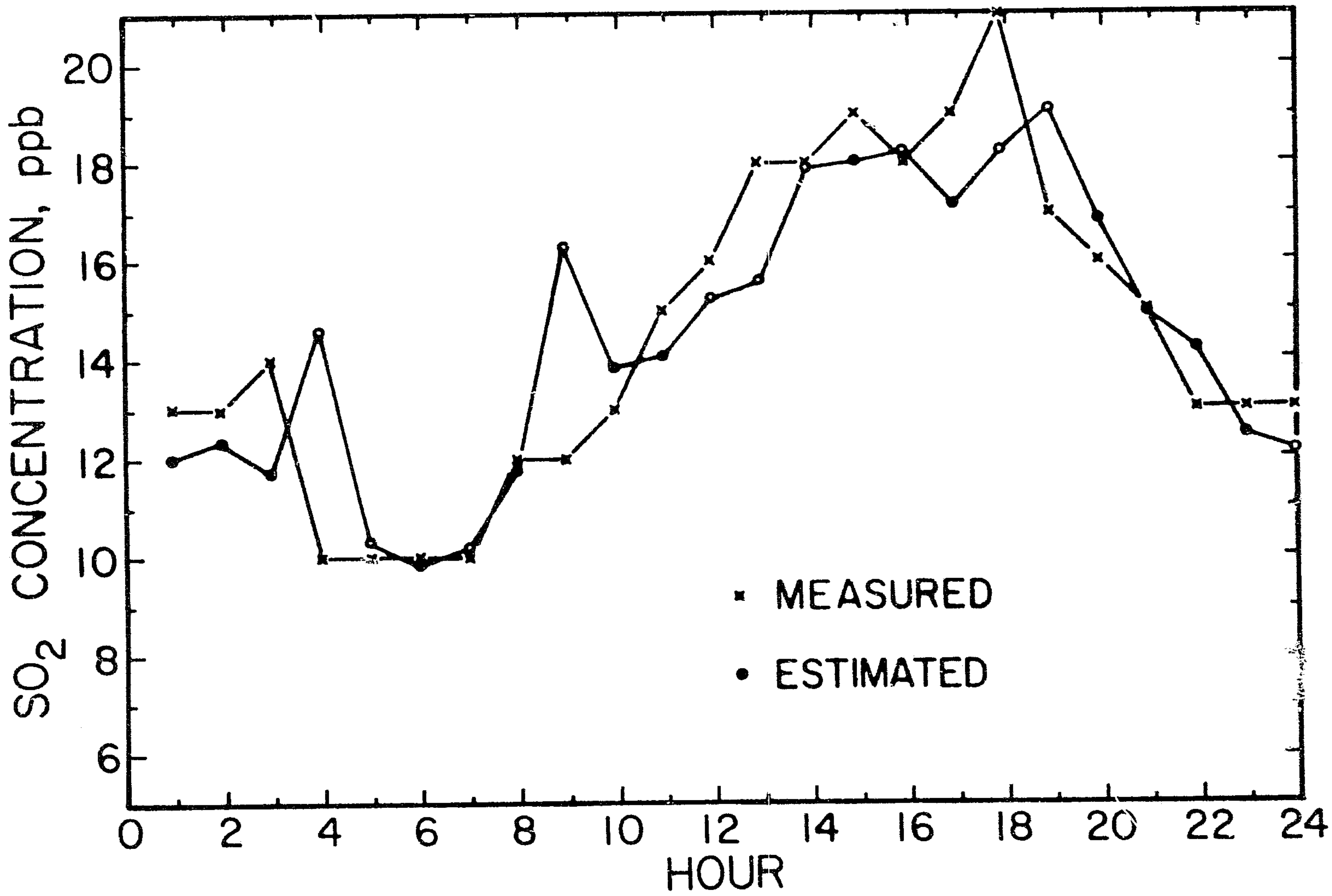

SO_2 CONCENTRATION, ppb
20
18
16
14
12
10
8
6
0
2
4
6
8
10
12
14
16
18
20
22
24
HOUR
MEASURED
ESTIMATED

Figure 5

, to appear in IEEE Transactions on Geoscience and Remote Sensing

ESTIMATION OF ATMOSPHERIC SPECIES CONCENTRATIONS FROM REMOTE SENSING DATA[+]

Sigeru Omatu[*] and John H. Seinfeld
Department of Chemical Engineering
California Institute of Technology
Pasadena, California 91125

ABSTRACT

A basic problem in the interpretation of atmospheric remote sensing data is to estimate species concentration distributions. Typical remote sensing data involve a field of view that moves across the region and represent integrated species burdens from the ground to the altitude of the instrument. The estimation problem arising from this special measurement configuration is solved based on the partial differential equation for atmospheric diffusion and Wiener-Hopf theory. The estimation of the concentration distribution downwind of a hypothetical continuous, ground-level source of pollutants is studied numerically.

[+]Research supported by NASA research grant NAG1-71.

[*]Permanent address: Department of Information Science and Systems Engineering
University of Tokushima
Tokushima, Japan

I. Introduction

In the remote sensing of atmospheric species, a ground-, aircraft-, or satellite-based platform scans a region of the atmosphere and measures the species burden within the field of view. An object of atmospheric remote sensing is to reconstruct species concentration distributions over a region based on the data available from the instrument.

There exist two recent studies that assess the capabilities of remote sensing for monitoring regional air pollution episdoes [1,2]. Diamonte et al. [3] developed theoretical results for the estimation of point source plume dispersion parameters from remote sensing data. In a similar vein, Kibbler and Suttles [4] studied the estimation of unknown parameters in a pollutant dispersion model by comparing model predictions with remotely sensed data. No results have yet been reported in which actual remote sensing data have been used to estimate species concentration distributions.

The present paper deals with the theoretical foundation of estimating atmospheric concentration distributions from remote sensing data. Since the atmosphere is a three-dimensional system, mathematical models of pollutant behavior are of the distributed parameter type [5]. Remote sensing data usually represent spatial averages of concentrations, so that the estimation problem concerns a distributed parameter system with spatially integrated, scanning data. Although distributed parameter state estimation has been considered extensively (see, for example, [6] and [7]), such problems with scanning and spatially integrated measurements have not been considered previously. The purpose of the present paper is to derive the required optimal estimators for the scanning and spatially integrated measurement case by a unified method based on the Wiener-Hopf theory.

In Section II, we define the remote sensing data analysis problem mathematically. Sections III-VI are devoted to derivation of the optimal prediction, filtering and smoothing algorithms for the problem by Weiner-Hopf theory. Finally, in Section VII we present a detailed numerical example of estimating the concentration distribution downwind of a continuous, ground-level line source to illustrate the application of the thoery.

II. Problem Statement

We consider a single atmospheric species (nonreactive), the mean concentration $u(t,x_1,x_2,x_3)$ of which over a certain region is described by the following form of the atmospheric diffusion equation [5],*

$$\frac{\partial u}{\partial t} + V_1 \frac{\partial u}{\partial x_1} + V_2 \frac{\partial u}{\partial x_2} = \frac{\partial}{\partial x_3}\left(K_v(x_3)\frac{\partial u}{\partial x_3}\right) + w(t,x_1,x_2,x_3) \tag{1}$$

where V_1 and V_2 are the mean velocities in the x_1- and x_2-directions, respectively, $K_v(x_3)$ is the vertical turbulent eddy diffusivity, and $w(t,x_1,x_2,x_3)$ is a random disturbance accounting for inaccuracies inherent in the basic model. The initial condition for (1) is $u(t_o,x_1,x_2,x_3) = u_o(x_1,x_2,x_3)$, and typical boundary conditions are

$$\begin{aligned} -K_v(x_3)\frac{\partial u}{\partial x_3} &= \tilde{S}(t,x_1,x_2), \qquad & x_3 &= 0 \\ \frac{\partial u}{\partial x_3} &= 0, & x_3 &= h \end{aligned} \tag{2}$$

where $\tilde{S}(t,x_1,x_2)$ is the ground-level species source emission rate, presumably a known function, and h denotes the upper vertical boundary of the pollutant-containing region, for example, the base of an inversion (stable) layer. For convenience, we denote the coordinate vector by x and let

$$L_x[\cdot] = -V_1\frac{\partial[\cdot]}{\partial x_1} - V_2\frac{\partial[\cdot]}{\partial x_2} + \frac{\partial}{\partial x_3}\left(K_v(x_3)\frac{\partial[\cdot]}{\partial x_3}\right)$$

*In this form of the atmospheric diffusion equation, turbulent diffusion in the horizontal direction is neglected relative to transport by the mean flow, a common assumption in treating atmospheric diffusion problems [5].

Assume that the concentration of a species over a fixed spatial domain D with its boundary ∂D is of interest. Let us define the operator Γ_ξ, $\xi \in \partial D$ as follows,

$$\Gamma_\xi = \begin{cases} -K_v(x_3)\dfrac{\partial[\cdot]}{\partial x_3}, & x_3 = 0 \\ \dfrac{\partial[\cdot]}{\partial x_3}, & x_3 = h. \end{cases}$$

Let $S(t,\xi)$ be

$$S(t,\xi) = \begin{cases} \tilde{S}(t,x_1,x_2), & x_3 = 0 \\ 0 \qquad\qquad\quad , & x_3 = h. \end{cases}$$

Thus, (1) can be represented as

$$\frac{\partial u(t,x)}{\partial x} = L_x u(t,x) + w(t,x) \tag{3}$$

and (2) can be written as

$$\Gamma_\xi u(t,\xi) = S(t,\xi), \qquad \xi \in \partial D. \tag{4}$$

We assume that the initial condition $u_o(x)$ can be represented as a Gaussian process with statistics,

$$E[u_o(x)] = \bar{u}_o(x)$$

$$E[\left(u_o(x) - \bar{u}_o(x)\right)\left(u_o(y) - \bar{u}_o(y)\right)] = P_o(x,y) \tag{5}$$

and the random disturbance $w(t,x)$ is stochastically independent of $u_o(x)$ and is a white Gaussian process with statistics,

$$E[w(t,x)] = 0$$
$$E[w(t,x)w(s,y)] = Q(t,x,y)\delta(t-s). \tag{6}$$

We assume that the remote sensing measurements are taken at time t_k over a view volume $D(k)$ consisting of M pixels, as shown in Fig. 1. Since the sensing platform may be in motion, the field of view, in general, moves with time across the entire spatial domain D. We assume that the shape and extent of the field of view $D(k)$ remain fixed and only the location of the centroid of each pixel changes with time. The ground-level location of the centroid of each pixel of $D(k)$ is denoted as $(x_1^{m(k)}, x_2^{m(k)}, 0)$, $m = 1,2,\ldots, M$

We are interested in considering the vertically integrated measurement given by

$$Z_{m(k)}(t_k,n) = \int_0^{h_n} \tilde{J}_{m(k)}(x_3)u(t_k,x_1^{m(k)},x_2^{m(k)},x_3)\, dx_3 + v(t_k,x_1^{m(k)},x_2^{m(k)},h_n) \tag{7}$$

$$m = 1,2,\ldots, M, \quad n = 1,2,\ldots, N, \; h_1 < h_2 < \ldots < h_N, \quad k = 1,2,\ldots$$

where $\tilde{J}_{m(k)}(x_3)$ is an altitude-dependent instrument weighting function, and h_n is the vertical position of the scanning sensor. Physically, $Z_{m(k)}(t_k,n)$ represents the vertically-integrated species concentrations within each of the M pixels, indicated by $m(k)$, at each time, t_k, from an altitude of h_n. $v(t_k,x_1^{m(k)},x_2^{m(k)}, h_n)$ represents measurement errors.

Some comments concerning the measurement configuration shown in Fig. 1 are in order. Ordinarily remote sensing from an airborne platform would be carried out at a single altitude. In such a case, it is not possible to estimate the concentration distribution between the platform and the ground based only on the integral of the concentration. Sakawa [8] and Koda and Seinfeld [9] have shown that in problems of this nature it is impossible to estimate the state uniquely based on integrated measurements from only a single sensor position since the required distributed parameter observability condition does not hold. Therefore, the estimation of species concentration distributions necessitates traverses over the region at different altitudes. From a practical point of view this requirement restricts this type of monitoring to aircraft platforms, which, for purposes of measuring air pollution, are the most useful. Considering that atmospheric concentration distributions change gradually and that airplane speeds are fast, the configuration sketched in Fig. 1 implies that repeated measurements at several altitudes are possible using only one airborne platform.

In order to represent (7) more compactly we introduce the following notation:

ORIGINAL PAGE IS OF POOR QUALITY

$$x(m(k)) = (x_1^{m(k)}, x_2^{m(k)})$$

$$J^n_{m(k)} = \begin{cases} \tilde{J}^n_{m(k)}(x_3), & x_3 \leq h_n \\ 0 \quad , & x_3 > h_n \end{cases}$$

$$J^n(t_k,x_3) = \begin{bmatrix} J^n_{1(k)}(x_3) & \cdots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \cdots & J^n_{M(k)}(x_3) \end{bmatrix}$$

$$u_{t_k}(x_3) = \begin{bmatrix} u(t_k,\ x(1(k)),\ x_3) \\ u(t_k,\ x(2(k)),\ x_3) \\ \vdots \\ u(t_k,\ x(M(k)), x_3) \end{bmatrix}$$

$$J(t_k,x_3) = \begin{bmatrix} J^1(t_k,x_3) \\ \vdots \\ J^N(t_k,x_3) \end{bmatrix}$$

$$Z(t_k,n) = \begin{bmatrix} Z_{1(k)}(t_k,n) \\ \vdots \\ Z_{M(k)}(t_k,n) \end{bmatrix}$$

$$Z(t_k) = \begin{bmatrix} Z(t_k,1) \\ \vdots \\ Z(t_k,N) \end{bmatrix}$$

$$v(t_k,n) = \begin{bmatrix} v(t_k,\ x(1(k)),\ h_n) \\ \vdots \\ v(t_k,\ x(M(k)),\ h_n) \end{bmatrix}$$

and

$$v(t_k) = \begin{bmatrix} v(t_k,1) \\ \vdots \\ v(t_k,N) \end{bmatrix}.$$

Then (7) can be represented compactly as

$$Z(t_k) = \int_0^h J(t_k,x_3)u_{t_k}(x_3)dx_3 + v(t_k). \tag{8}$$

We assume that $v(t_k)$ is independent of $w(t,x)$ and $u_0(x)$ and is a white Gaussian process with statistics, $E[v(t_k)] = 0$ and $E[v(t_k)v'(t_\ell)] = R(t_k)\delta_{k\ell}$, where $'$ denotes the transpose operator and $R(t_k)$ is an MNxMN positive-definite matrix.

The problem considered here is to estimate $u(t,x)$ over D on the basis of the measurement $Z(t_\sigma)$, $\sigma = 0, 1, \ldots, k$. The novel aspect of this problem from the point of view of distributed parameter estimation arises because of the scanning and vertically integrated nature of the measurements. In what follows, we use k instead of t_k as long as there is no ambiguity.

III. Estimation Problems and Wiener-Hopf Theory

Let us denote the estimate of $u(t_\tau,x)$ based on the observation data $Z(t_\sigma)$, $\sigma = 0,1,\ldots, k$ by $\hat{u}(t_\tau,x/t_k)$ which is given by the following linear transformation of $Z(t_\sigma)$, $\sigma = 0,1,\ldots, k$,

$$\hat{u}(t_\tau,x/t_k) = \sum_{\sigma=0}^{k} \tilde{F}(t_\tau,x,t_\sigma)Z(t_\sigma) \tag{9}$$

where $\tilde{F}(t_\tau,x,t_\sigma)$ is an unknown MN-dimensional row vector called the estimation kernel function. When there is no ambiguity, we write (9) compactly as

$$\hat{u}(\tau,x/k) = \sum_{\sigma=0}^{k} \tilde{F}(t_\tau,x,t_\sigma)Z(t_\sigma). \tag{10}$$

Furthermore, we denote the estimation error and error covariance functions by $\tilde{u}(t_\tau,x/t_k)$ and $P(t_\tau,x,y/t_k)$, respectively, where $\tilde{u}(t_\tau,x/t_k) = u(t_\tau,x) - \hat{u}(t_\tau,x/t_k)$ and $P(t_\tau,x,y/t_k) = E[\tilde{u}(t_\tau x/t_k)\tilde{u}(t_\tau,y/t_k)]$. The estimate $\hat{u}(t_\tau,x/t_k)$ that minimizes $J(\hat{u}) = E[\tilde{u}(t_\tau,x/t_k)^2]$ is said to be optimal. Note that by using $P(t_\tau,x,y/t_k)$, $J(\hat{u})$ can be rewritten as $J(\hat{u}) = P(t_\tau,x,x/t_k)$.

To clarify the differences between the prediction, filtering, and smoothing problems, we express $\tilde{F}(t_\tau,x,t_\sigma)$ differently for each problem as follows:

(i) Prediction ($t > t_k$)

$$\hat{u}(t,x/t_k) = \sum_{\sigma=0}^{k} A(t,x,t_\sigma)Z(t_\sigma). \tag{11}$$

(ii) Filtering ($t_\tau = t_k$)

$$\hat{u}(t_k,x/t_k) = \sum_{\sigma=0}^{k} F(t_k,x,t_\sigma)Z(t_\sigma). \tag{12}$$

(iii) Smoothing ($t_\tau < t_k$)

$$\hat{u}(t_\tau,x/t_k) = \sum_{\sigma=0}^{k} B(t_\tau,t_k,x,t_\sigma)Z(t_\sigma). \tag{13}$$

Here we use three temporal arguments t_τ, t_k and t_σ for the smoothing kernel $B(t_\tau,t_k,x,t_\sigma)$ since these parameters should be changed according to the measurement data acquisition time. Then the following theorem can be proved similarly to that of [6] for the continuous-time observation case.

[Theorem 1] (Wiener-Hopf Theorem)
A necessary and sufficient condition for the estimate $\hat{u}(t_\tau,x/t_k)$ to be optimal is that the following Wiener-Hopf equation holds for $\zeta = 0,1,\ldots, k$ and $x \in \bar{D} = D \cup \partial D$,

$$\sum_{\sigma=0}^{k} \bar{F}(t_\tau,x,t_\sigma)E[Z(t_\sigma)Z'(t_\zeta)] = E[u(t_\tau,x)Z'(t_\zeta)], \tag{14}$$

or equivalently, for $\zeta = 0,1,\ldots, k$ and $x \in \bar{D}$,

$$E[\tilde{u}(t_\tau,x/t_k)Z'(t_\zeta)] = 0. \tag{15}$$

[Corollary 1] (Orthogonal projection lemma)
The orthogonality condition, $E[\tilde{u}(t_\tau,x/t_k)\hat{u}(t_\eta,y/t_k)] = 0$, $x,y \in \bar{D}$, holds where t_η is any time instant such as $t_\eta < t_k$, $t_\eta = t_k$, or $t_\eta > t_k$.

[Proof] Multiplying each side of (15) by $\tilde{F}'(t_\eta,y,t_\zeta)$ and summing from $\zeta = 0$ to $\zeta = k$ yields

$$E[\tilde{u}(t_\tau,x/t_k)\sum_{\zeta=0}^{k} Z'(t_\zeta)\tilde{F}'(t_\eta,y,t_\zeta)] = 0.$$

Using (9) in the above equation yields the desired relation completing the proof of the corollary. Q.E.D.

[Lemma 1] (Uniqueness of the optimal kernel)

Let $\tilde{F}(t_\tau,x,t_\sigma)$ be the optimal kernel function satisfying the Wiener-Hopf equation (14) and let $\tilde{F}(t_\tau,x,t_c) + \tilde{F}_\Delta(t_\tau,x,t_\sigma)$ be also the optimal kernel function satisfying the Wiener-Hopf equation (14). Then it follows that $\tilde{F}_\Delta(t_\tau,x,t_\sigma) \equiv 0$, $\sigma = 0,1,\ldots, k$ and $x \in \bar{D}$, i.e. the optimal kernel function is unique.

In order to consider the prediction, filtering, and smoothing problems, separately, we rewrite (14) using the notation of (11) - (13).

[Corollary 2] The Wiener-Hopf equation (14) is rewritten for the prediction, filtering, and smoothing problems as follows:

(i) Prediction $(t > t_k)$

$$\sum_{\sigma=0}^{k} A(t,x,t_\sigma)E[Z(t_\sigma)Z'(t_\zeta)] = E[u(t,x)Z'(t_\zeta)] \tag{16}$$

for $\zeta = 0,1,\ldots, k$ and $x \in \bar{D}$.

(ii) Filtering $(t_\tau = t_k)$

$$\sum_{\sigma=0}^{k} F(t_k,x,t_\sigma)E[Z(t_\sigma)Z'(t_\zeta)] = E[u(t_k,x)Z'(t_\zeta)] \tag{17}$$

for $\zeta = 0,1,\ldots, k$ and $x \in \bar{D}$.

(iii) Smoothing ($t_\tau < t_k$)

$$\sum_{\sigma=0}^{k} B(t_\tau,t_k,x,t_\sigma)E[Z(t_\sigma)Z'(t_\zeta)] = E[u(t_\tau,x)Z'(t_\zeta)] \tag{18}$$

for $\zeta = 0,1,\ldots, k$ and $x \in \bar{D}$.

IV. Derivation of the Optimal Prediction Estimator

In this section we derive the optimal prediction estimator by using the Wiener-Hopf theory in the previous section.

[Theorem 2] The optimal prediction estimator is given by

$$\frac{\partial \hat{u}(t,x/t_k)}{\partial t} = L_x \hat{u}(t,x/t_k), \qquad t > t_k \tag{19}$$

$$\Gamma_\xi \hat{u}(t,\xi/t_k) = S(t,\xi), \qquad \xi \in \partial D. \tag{20}$$

[Proof] Differentiating (16) with respect to t and substituting (3) yields

$$\sum_{\sigma=0}^{k} \frac{\partial A(t,x,t_\sigma)}{\partial t} E[Z(t_\sigma)Z'(t_\zeta)] = L_x E[u(t,x)Z'(t_\zeta)]$$

where the independence of $w(t,x)$ and $Z(t_\zeta)$ is used. Substituting (16) into the above equation yields

$$\sum_{\sigma=0}^{k} \tilde{F}_\Delta(t,x,t_\sigma)E[Z(t_\sigma)Z'(t_\zeta)] = 0$$

where

$$\tilde{F}_\Delta(t,x,t_\sigma) = \frac{\partial A(t,x,t_\sigma)}{\partial t} - L_x A(t,x,t_\sigma).$$

From Lemma 1 we have

$$\frac{\partial A(t,x,t_\sigma)}{\partial t} = L_x A(t,x,t_\sigma). \tag{21}$$

Differentiating (11) with respect to t and substituting (21) yields (19). Since the forms of Γ_ξ and $S(t,\xi)$ are known, the predicted estimate $\hat{u}(t,x/t_k)$ also satisfies the same boundary condition (4). Q.E.D.

[Theorem 3] The optimal prediction error covariance function $P(t,x,y/t_k)$ is governed by

$$\frac{\partial P(t,x,y/t_k)}{\partial t} = (L_x + L_y)P(t,x,y/t_k) + Q(t,x,y), \tag{22}$$

$$\Gamma_\xi P(t,\xi,y/t_k) = 0, \qquad \xi \in \partial D. \tag{23}$$

[Proof] From (3), and (19) we have

$$\frac{\partial \tilde{u}(t,x/t_k)}{\partial t} = L_x \tilde{u}(t,x/t_k) + w(t,x) \tag{24}$$

and from (4), and (20)

$$\Gamma_\xi \tilde{u}(t,\xi/t_k) = 0, \qquad \xi \in \partial D. \tag{25}$$

Differentiating the definition of P with respect to t and using (24) yields

$$\frac{\partial P(t,x,y/t_k)}{\partial t} = (L_x + L_y)P(t,x,y/t_k) + \Sigma(t,x,y)$$

where

$$\Sigma(t,x,y) = E[w(t,x)\tilde{u}(t,y/t_k)] + E[\tilde{u}(t,x/t_k)w(t,y)].$$

Let the fundamental solution of L_x be $G(t,\sigma,x,y)$, where

$$\frac{\partial G(t,\sigma,x,y)}{\partial t} = L_x G(t,\sigma,x,y),$$

$$\Gamma_\xi G(t,\sigma,\xi,y) = S(t,\xi), \qquad \xi \in \partial D$$

$$G(\sigma,\sigma,x,y) = \delta(x-y).$$

Then $\tilde{u}(t,x/t_k)$ of (24) can be represented in terms of $G(t,\sigma,x,y)$ as follows,

$$\tilde{u}(t,x/t_k) = \int_D G(t,t_k,x,\alpha)\,\tilde{u}(t_k,\alpha/t_k)d\alpha + \int_{t_k}^{t}\int_D G(t,\sigma,x,\alpha)w(\sigma,\alpha)d\alpha d\sigma. \qquad (26)$$

Substituting (26) into $\Sigma(t,x,y)$ and using (6) yields $\Sigma(t,x,y) = Q(t,x,y)$. Multiplying each side of (25) by $\tilde{u}(t,y/t_k)$ and taking the expectation yields (23). Q.E.D.

[Corollary 3] The optimal prediction estimate $\hat{u}(t,x/t_k)$ and prediction error covariance function $P(t,x,y/t_k)$ can be represented as

$$\hat{u}(t,x/t_k) = \int_D G(t,t_k,x,\alpha)\hat{u}(t_k,\alpha/t_k)d\alpha \qquad (27)$$

and

$$P(t,x,y/t_k) = \int_D\int_D G(t,t_k,x,\alpha)P(t_k,\alpha,\beta/t_k)G(t,t_k,y,\beta)\; d\alpha d\beta$$

$$+ \int_{t_k}^{t}\int_D\int_D G(t,\sigma,x,\alpha)Q(\sigma,\alpha,\beta)G(t,\sigma,y,\beta)\; d\alpha d\beta d\sigma. \qquad (28)$$

[Proof] It is clear that (19) and (22) possess unique solutions. Differentiating (27) and (28) with respect to t yields (19) and (22), respectively. Since (19) and (22) have unique solutions, (27) and (28) are those solutions. Q.E.D.

V. Derivation of the Optimal Filter

In order to derive the optimal filter by using the Wiener-Hopf theorem for the filtering problem, we represent the solution of (3) in terms of the fundamental solution $G(t,\sigma,x,y)$ as

$$u(t_{k+1},x) = \int_D G(t_{k+1},t_k,x,\alpha)u(t_k,\alpha)\, d\alpha + \int_{t_k}^{t_{k+1}}\int_D G(t_{k+1},\eta,x,\alpha)w(\eta,\alpha)\, d\alpha d\eta \tag{29}$$

and

$$u_{t_{k+1}}(x_3) = \int_D G_M(t_{k+1},t_k,x_3,\alpha)u(t_k,\alpha)\, d\alpha + \int_{t_k}^{t_{k+1}}\int_D G_M(t_{k+1},\eta,x_3\alpha)w(\eta,\alpha)\, d\alpha d\eta \tag{30}$$

where

$$G_M(t_{k+1},\eta,x_3,\alpha) = \begin{bmatrix} G(t_{k+1},\eta,\ x_1^{1(k+1)},x_2^{1(k+1)},x_3,\alpha) \\ \vdots \\ G(t_{k+1},\eta,\ x_1^{M(k+1)},x_2^{M(k+1)},x_3,\alpha) \end{bmatrix}. \tag{31}$$

From (17) we have

$$F(t_{k+1},x,t_{k+1})E[Z(t_{k+1})Z'(t_\zeta)] + \sum_{\sigma=0}^{k} F(t_{k+1},x,t_\sigma)E[Z(t_\sigma)Z'(t_\zeta)] = E[u(t_{k+1},x)Z'(t_\zeta)] \tag{32}$$

for $\zeta = 0,1,\ldots, k+1$.

From (29) and the independence of $Z(t_\zeta)$, $\zeta = 0,1,\ldots, k$ and $w(\eta,x)$, $t_k < \eta \leq t_{k+1}$ it follows that

$$E[u(t_{k+1},x)Z'(t_\zeta)] = \int_D G(t_{k+1},t_k,x,\alpha)E[u(t_k,\alpha)Z'(t_\zeta)]\, d\alpha.$$

Using the Wiener-Hopf equation (17), we have

$$E[u(t_{k+1},x)Z'(t_\zeta)] = \int_D G(t_{k+1},t_k,x,\alpha) \sum_{\sigma=0}^{k} F(t_k,\alpha,t_\sigma)E[Z(t_\sigma)Z'(t_\zeta)]. \tag{33}$$

On the other hand, from (8) and the whiteness of $v(t_{k+1})$, we have, for $t_\zeta \leq t_k$,

$$E[Z(t_{k+1})Z'(t_\zeta)] = \int_0^h J(t_{k+1},x_3)E[u_{t_{k+1}}(x_3)Z'(t_\zeta)]\, dx_3.$$

Substituting (30) into the above equation and using the independence of $Z(t_\zeta)$, $t_\zeta \leq t_k$ and $w(\eta,\alpha)$, $t_k < \eta \leq t_{k+1}$ yields

$$E[Z(t_{k+1})Z'(t_\zeta)] = \int_0^h J(t_{k+1},x_3) \int_D G_M(t_{k+1},t_k,x_3,\alpha)E[u(t_k,\alpha)Z'(t_\zeta)]\, d\alpha dx_3.$$

Again, we use the Wiener-Hopf equation (17) in the above equation and

$$E[Z(t_{k+1})Z'(t_\zeta)] = \int_0^h J(t_{k+1},x_3) \int_D G_M(t_{k+1},t_k,x_3,\alpha) \sum_{\sigma=0}^{k} F(t_k,\alpha,t_\sigma)$$
$$E[Z(t_\sigma)Z'(t_\zeta)]\, d\alpha dx_3. \tag{34}$$

Substituting (33) and (34) in (32) yields

$$\sum_{\sigma=0}^{k} \tilde{F}_{\Delta}(t_k,x,t_\sigma)E\ Z(t_\sigma)Z'(t_\zeta) = 0$$

where

$$\tilde{F}_{\Delta}(t_k,x,t_\sigma) = F(t_{k+1},x,t_{k+1}) \int_0^h J(t_{k+1},x_3)\int_D G_M(t_{k+1},t_k,x_3,\alpha)F(t_k,\alpha,t_\sigma)\ d\alpha dx_3$$

$$+ F(t_{k+1},x,t_\sigma) - \int_D G(t_{k+1},t_k,x,\alpha)F(t_k,\alpha,t_\sigma)\ d\alpha.$$

Then from Lemma 1 we have $\tilde{F}_{\Delta}(t_k,x,t_\sigma) \equiv 0$, and we have the following lemma.

[Lemma 2] The optimal kernel function $F(t_{k+1},x,t_\sigma)$ of the filter is given by

$$F(t_{k+1},x,t_\sigma) = \int_D G(t_{k+1},t_k,x,\alpha)F(t_k,\alpha,t_\sigma)\ d\alpha$$

$$- F(t_{k+1},x,t_{k+1}) \int_0^h J(t_{k+1},x_3) \int_D G_M(t_{k+1},t_k,x_3,\alpha)F(t_k,\alpha,t_\sigma)\ d\alpha dx_3. \quad (35)$$

[Theorem 4] The optimal filtering estimate $\hat{u}(t_{k+1},x/t_{k+1})$ is given by

$$\hat{u}(t_{k+1},x/t_{k+1}) = \hat{u}(t_{k+1},x/t_k) + F(t_{k+1},x,t_{k+1})\nu(t_{k+1}), \quad (36)$$

$$\nu(t_{k+1}) \triangleq Z(t_{k+1}) - \int_0^h J(t_{k+1},x_3)\hat{u}_{t_{k+1}}(x_3/t_k)\ dx_3, \quad (37)$$

$$\hat{u}(t_\sigma,x/t_o) = \bar{u}_o(x), \quad (38)$$

$$\Gamma_\xi \hat{u}(t_{k+1},\xi/t_{k+1}) = S(t_{k+1},\xi), \quad \xi \in \partial D \quad (39)$$

where

$$\hat{u}_{t_{k+1}}(x_3/t_k) = \begin{bmatrix} \hat{u}(t_{k+1},x_1^{1(k+1)},x_2^{1(k+1)},x_3/t_k) \\ \vdots \\ \hat{u}(t_{k+1},x_1^{M(k+1)},x_2^{M(k+1)},x_3/t_k) \end{bmatrix}.$$

[Proof] Using (12) and (35) yields

$$\hat{u}(t_{k+1},x/t_{k+1}) = F(t_{k+1},x,t_{k+1})Z(t_{k+1})$$

$$+\int_D G(t_{k+1},t_k,x,\alpha) \sum_{\sigma=0}^{k} F(t_k,\alpha,t_\sigma)Z(t_\sigma)\, d\alpha$$

$$- F(t_{k+1},x,t_{k+1}) \int_0^h J(t_{k+1},x_3) \int_D G_M(t_{k+1},t_k,x_3,\alpha) \sum_{\sigma=0}^{k} F(t_k,\alpha,t_\sigma)Z(t_\sigma)\, d\alpha dx_3.$$

Then from (12) and (27) we have

$$\hat{u}(t_{k+1},x/t_{k+1}) = \int_D G(t_{k+1},t_k,x,\alpha)\hat{u}(t_k,\alpha/t_k)\, d\alpha$$

$$+ F(t_{k+1},x,t_{k+1})(Z(t_{k+1}) - \int_0^h J(t_{k+1},x_3) \int_D G_M(t_{k+1},t_k,x_3,\alpha)\hat{u}(t_k,\alpha/t_k)\, d\alpha dx_3$$

$$= \hat{u}(t_{k+1},x/t_k) + F(t_{k+1},x,t_{k+1})\nu(t_{k+1}).$$

Since the initial and boundary conditions are clear, the proof of the theorem is complete. Q.E.D.

To determine the optimal kernel function $F(t_{k+1},x,t_{k+1})$, we introduce the following notation,

$$P_M(t_\tau,x,y_3/t_k) = [P(t_\tau,x,y^{1(k)}/t_k),\ldots,\ P(t_\tau,x,y^{M(k)}/t_k)] \tag{40}$$

and

$$P_{MM}(t_\tau,x_3,y_3/t_k) = \begin{bmatrix} P(t_\tau,x^{1(k)},y_3/t_k) \\ \vdots \\ P(t_\tau,x^{M(k)},y_3/t_k) \end{bmatrix}$$

$$= \begin{bmatrix} P(t_\tau,x^{1(k)},y^{1(k)}/t_k), & \ldots, & P(t_\tau,x^{1(k)},y^{M(k)}/t_k) \\ \vdots & & \vdots \\ P(t_\tau,x^{M(k)},y^{1(k)}/t_k), & \ldots, & P(t_\tau,x^{M(k)},y^{M(k)}/t_k) \end{bmatrix} \tag{41}$$

where $x^{m(k)} = (x_1^{m(k)}, x_2^{m(k)}, x_3)$ and $y^{m(k)} = (y_1^{m(k)}, y_2^{m(k)}, y_3)$, $m = 1,2,\ldots, M$. From the definitions of $P_M(t_\tau,x,y_3/t_k)$ and $P_{MM}(t_\tau,x_3,y_3/t_k)$ it follows that

$$P_M(t_\tau,x,y_3/t_k) = E[\tilde{u}(t_\tau,x/t_k)\tilde{u}_{t_\tau}(y_3/t_k)] \tag{42}$$

and

$$P_{MM}(t_\tau,x_3,y_3/t_k) = E[\tilde{u}_{t_\tau}(x_3/t_k)\tilde{u}'_{t_\tau}(y_3/t_k)] \tag{43}$$

where

$$\tilde{u}_{t_\tau}(x_3/t_k) = u_{t_\tau}(x_3) - \hat{u}_{t_\tau}(x_3/t_k) \tag{44}$$

and

$$\hat{u}_{t_\tau}(x_3/t_k) = \begin{bmatrix} \hat{u}(t_\tau,x^{1(k)}/t_k) \\ \vdots \\ \hat{u}(t_\tau,x^{M(k)}/t_k) \end{bmatrix}. \tag{45}$$

Furthermore, we define the covariance matrix of the innovation process $\nu(t_{k+1})$ by $\Gamma(t_{k+1}/t_k) = E[\nu(t_{k+1})\nu'(t_{k+1})]$. Then from (37) we have

$$\Gamma(t_{k+1}/t_k) = \int_0^h \int_0^h J(t_{k+1},x_3)P_{MM}(t_{k+1},x_3,y_3/t_k)J'(t_{k+1},y_3)\, dx_3 dy_3$$
$$+ R(t_{k+1}). \tag{46}$$

[Theorem 5] The optimal filtering gain function $F(t_{k+1},x,t_{k+1})$ is given by

$$F(t_{k+1},x,t_{k+1}) = \int_0^h P_M(t_{k+1},x,y_3/t_k)J'(t_{k+1},y_3)dy_3\Gamma^{-1}(t_{k+1}/t_k). \quad (47)$$

[Proof] From the Wiener-Hopf equation (17) we have

$$F(t_{k+1},x,t_{k+1})E[Z(t_{k+1})Z'(t_{k+1})]$$

$$+ \sum_{\sigma=0}^{k} F(t_{k+1},x,t_\sigma)E[Z(t_\sigma)Z'(t_{k+1})] = E[u(t_{k+1},x)Z'(t_{k+1})].$$

Substituting (35) into the above equation yields

$$F(t_{k+1},x,t_{k+1})E[(Z(t_{k+1}) - \int_0^h J(t_{k+1},x_3)\hat{u}_{t_{k+1}}(x_3/t_k)dx_3)Z'(t_{k+1})]$$

$$= E[(u(t_{k+1},x) - \hat{u}(t_{k+1},x/t_k))Z'(t_{k+1})].$$

Using (8) and the orthogonality condition of Corollary 1 yields

$$E[\tilde{u}(t_{k+1},x/t_k)Z'(t_{k+1})] = \int_0^h E[\tilde{u}(t_{k+1},x/t_k)u'_{t_{k+1}}(x_3)]J'(t_{k+1},x_3)\,dx_3$$

$$= \int_0^h P_M(t_{k+1},x,x_3/t_k)J'(t_{k+1},x_3)\,dx_3$$

and

$$E[v(t_{k+1})Z'(t_{k+1})] = \int_0^h\int_0^h J(t_{k+1},x_3)P_{MM}(t_{k+1},x_3,y_3/t_k)J'(t_{k+1},y_3)\,dx_3dy_3$$

$$+ R(t_{k+1}) = \Gamma(t_{k+1}/t_k). \quad (48)$$

Then we have

$$F(t_{k+1},x,t_{k+1})\Gamma(t_{k+1}/t_k) = \int_0^h P_M(t_{k+1},x,x_3/t_k)J'(t_{k+1},x_3)\,dx_3$$

and the proof of the theorem is complete. Q.E.D.

[Theorem 6] The optimal filtering error covariance function $P(t_{k+1},x,y/t_{k+1})$ is given by

$$P(t_{k+1},x,y/t_{k+1}) = P(t_{k+1},x,y/t_k)$$

$$-\int_0^h\int_0^h P_M(t_{k+1},x,x_3/t_k)J'(t_{k+1},x_3)\Gamma^{-1}(t_{k+1}/t_k)J(t_{k+1},y_3)P_M'(t_{k+1},y,y_3/t_k)\,dx_3dy_3 \tag{49}$$

$$P(t_o,x,y/t_o) = P_o(x,y) \tag{50}$$

$$\Gamma_\xi P(t_{k+1},\xi,y/t_k) = 0, \qquad \xi \in \partial D. \tag{51}$$

[Proof] From (3) and (36) we have

$$\tilde{u}(t_{k+1},x/t_{k+1}) = \tilde{u}(t_{k+1},x/t_k) - F(t_{k+1},x,t_{k+1})\nu(t_{k+1}) \tag{52}$$

and from (4) and (39),

$$\Gamma_\xi\tilde{u}(t_{k+1},\xi/t_{k+1}) = 0, \qquad \xi \in \partial D. \tag{53}$$

Using the independence of $v(t_{k+1})$ and $\tilde{u}(t_{k+1},x/t_k)$ or $\tilde{u}(t_{k+1},y/t_k)$ yields

$$P(t_{k+1},x,y/t_{k+1}) = E[\tilde{u}(t_{k+1},x/t_{k+1})\tilde{u}(t_{k+1},y/t_{k+1})]$$

$$= P(t_{k+1},x,y/t_k) + F(t_{k+1},x,t_{k+1})E[\nu(t_{k+1})\nu'(t_{k+1})]F'(t_{k+1},y,t_{k+1})$$

$$- F(t_{k+1},x,t_{k+1}) \int_0^h J(t_{k+1},x_3)E[\tilde{u}_{t_{k+1}}(x_3/t_k)\tilde{u}(t_{k+1},y/t_k)]\, dx_3$$

$$- \int_0^h E[\tilde{u}(t_{k+1},x/t_k)\tilde{u}'_{t_{k+1}}(y_3/t_k)]J'(t_{k+1},y_3)dy_3F'(t_{k+1},y,t_{k+1}).$$

Using (40) and (47) yields

$$P(t_{k+1},x,y/t_{k+1}) = P(t_{k+1},x,y/t_k)$$

$$- \int_0^h \int_0^h P_M(t_{k+1},x,x_3/t_k)J'(t_{k+1},x_3)\Gamma^{-1}(t_{k+1}/t_k)$$

$$J(t_{k+1},y_3)P_M'(t_{k+1},y,y_3/t_k)\, dx_3dy_3.$$

Since the initial value $\hat{u}(t_0,x/t_0)$ is equal to $\bar{u}_0(x)$, it is clear that $P(t_0,x,y/t_0) = P_0(x,y)$. Multiplying each side of (53) by $\tilde{u}(t_{k+1},y/t_{k+1})$ and taking the expectation yields (51). Q.E.D.

VI. Derivation of the Optimal Smoothing Estimator

In this section we derive the optimal smoothing estimator by using the Wiener-Hopf theory.

[Lemma 3] The optimal kernel function $B(t_\tau,t_{k+1},x,t_\sigma)$ of the smoothing estimator is given by

$$B(t_\tau,t_{k+1},x,t_\sigma) = B(t_\tau,t_k,x,t_\sigma)$$

$$- B(t_\tau,t_{k+1},x,t_{k+1}) \int_0^h J(t_{k+1},x_3) \int_D G_M(t_{k+1},t_k,x_3,\alpha)F(t_k,\alpha,t_\sigma)\, d\alpha dx_3. \qquad (54)$$

[Proof] From the Wiener-Hopf equation (18) for the smoothing problem we have

$$\sum_{\sigma=0}^{k+1} B(t_\tau,t_{k+1},x,t_\sigma)E[Z(t_\sigma)Z'(t_\zeta)] = E[u(t_\tau,x)Z'(t_\zeta)], \qquad (55)$$

$$\zeta = 0,1,\ldots, k+1$$

and

$$\sum_{\sigma=0}^{k} B(t_\tau,t_k,x,t_\sigma)E[Z(t_\sigma)Z'(t_\zeta)] = E[u(t_\tau,x)Z'(t_\zeta)], \qquad (56)$$

$$\zeta = 0,1,\ldots, k.$$

Subtracting (56) from (55) yields

$$B(t_\tau,t_{k+1},x,t_{k+1})E[Z(t_{k+1})Z'(t_\zeta)]$$

$$+ \sum_{\sigma=0}^{k}(B(t_\tau,t_{k+1},x,t_\sigma) - B(t_\tau,t_k,x,t_\sigma))E[Z(t_\sigma)Z'(t_\zeta)] = 0.$$

From (8) and (17) we have

$$E[Z(t_{k+1})Z'(t_\zeta)] = \int_0^h J(t_{k+1},x_3)\int_D G_M(t_{k+1},t_k,x_3\alpha)E[u(t_k,\alpha)Z'(t_\zeta)]\, d\alpha dx_3$$

$$= \int_0^h J(t_{k+1},x_3)\int_D G_M(t_{k+1},t_k,x_3,\alpha) \sum_{\sigma=0}^{k} F(t_k,\alpha,t_\sigma)E[Z(t_\sigma)Z'(t_\zeta)]\, d\alpha dx_3.$$

Then it follows that

$$\sum_{\sigma=0}^{k} \tilde{F}_\Delta(t_\tau,t_k,x,t_\sigma)E[Z(t_\sigma)Z'(t_\zeta)] = 0$$

where

$$\tilde{F}_\Delta(t_\tau,t_k,x,t_\sigma) = B(t_\tau,t_{k+1},x,t_\sigma) - B(t_\tau,t_k,x,t_\sigma)$$

$$+ B(t_\tau,t_{k+1},x,t_{k+1}) \int_0^h J(t_{k+1},x_3) \int_D G_M(t_{k+1},t_k,x_3,\alpha)F(t_k,\alpha,t_\sigma).$$

Since it is clear that $B(t_\tau,t_k,x,t_\sigma) + \tilde{F}_\Delta(t_\tau,t_k,x,t_\sigma)$ also satisfies the Wiener-Hopf equation (18), from Lemma 1 $\tilde{F}_\Delta(t_\tau,t_k,x,t_\sigma) \equiv 0$, $\sigma = 0,1,\ldots, k$. Thus, the proof of the lemma is complete. Q.E.D.

[Theorem 7] The optimal smoothing estimate $\hat{u}(t_\tau,x/t_{k+1})$ is given by

$$\hat{u}(t_\tau,x/t_{k+1}) = \hat{u}(t_\tau,x/t_k) + B(t_\tau,t_{k+1},x,t_{k+1})\nu(t_{k+1}) \tag{57}$$

$$\Gamma_\xi\hat{u}(t_\tau,\xi/t_{k+1}) = S(\tau,\xi), \qquad \xi \in \partial D. \tag{58}$$

[Proof] From (13) it follows that

$$\hat{u}(t_\tau,x/t_{k+1}) = B(t_\tau,t_{k+1},x,t_{k+1})Z(t_{k+1})$$
$$+ \sum_{\sigma=0}^{k} B(t_\tau,t_{k+1},x,t_\sigma)Z(t_\sigma).$$

Substituting (54) into the above equation yields

$$\hat{u}(t_\tau,x/t_{k+1}) = B(t_\tau,t_{k+1},x,t_{k+1})\nu(t_{k+1}) + \sum_{\sigma=0}^{k} B(t_\tau,t_k,x,t_\sigma)Z(t_\sigma)$$

and substituting (13) into the above equation yields (57). Since we have no additional information about the boundary value of $u(t_\tau,x)$ except for $S(t_\tau,\xi)$, we have (58). Thus, the proof of the theorem is complete. Q.E.D.

[Theorem 8] The optimal smoothing gain function $B(t_\tau,t_{k+1},x,t_{k+1})$ is given by

$$B(t_\tau,t_{k+1},x,t_{k+1}) = \int_0^h N(t_\tau,x,x_3/t_{k+1})J'(t_{k+1},x_3)dx_3\Gamma^{-1}(t_{k+1}/t_k) \quad (59)$$

where

$$N(t_\tau,x,x_3/t_{k+1}) = \int_D M(t_\tau,x,y/t_k)G_M'(t_{k+1},t_k,x_3,y)\,dy \quad (60)$$

and

$$M(t_\tau,x,y/t_k) = E[\tilde{u}(t_\tau,x/t_k)\tilde{u}(t_k,y/t_k)]. \quad (61)$$

[Proof] From the Wiener-Hopf equation (18) we have

$$B(t_\tau,t_{k+1},x,t_{k+1})E[Z(t_{k+1})Z'(t_{k+1})] + \sum_{\sigma=0}^{k} B(t_\tau,t_{k+1},x,t_\sigma)E[Z(t_\sigma)Z'(t_{k+1})] = E[u(t_\tau,x)Z'(t_{k+1})].$$

Substituting (54) into the above equation yields

$$B(t_\tau,t_{k+1},x,t_{k+1})E[\nu(t_{k+1})Z'(t_{k+1})] = E[\tilde{u}(t_\tau,x/t_k)Z'(t_{k+1})]. \tag{62}$$

On the other hand, from (27) and (29)

$$\tilde{u}(t_{k+1},x/t_k) = \int_D G(t_{k+1},t_k,x,y)\tilde{u}(t_k,y/t_k)\,dy$$

$$+ \int_{t_k}^{t_{k+1}}\int_D G(t_{k+1},\eta,x,y)w(\eta,y)\,dyd\eta.$$

Then we have

$$E[\tilde{u}(t_\tau,x/t_k)Z'(t_{k+1})] = \int_0^h\int_D M(t_\tau,x,y/t_k)G_M'(t_{k+1},t_k,x_3,y)\,dy\,J'(t_{k+1},x_3)\,dx_3$$

$$= \int_0^h N(t_\tau,x,x_3/t_{k+1})J'(t_{k+1},x_3)\,dx_3.$$

Substituting (48) and the above equation into (62) yields (59). Thus, the proof of the theorem is complete. Q.E.D.

Let us now derive the equation for $M(t_\tau,x,y/t_{k+1})$. Using the orthogonality condition of Corollary 1 yields

$$M(t_\tau,x,y/t_{k+1}) = E[u(t_\tau,x)\tilde{u}(t_{k+1},y/t_{k+1})]. \tag{63}$$

Substituting (52) into the above equation yields

$$M(t_\tau,x,y/t_{k+1}) = \int_D G(t_{k+1},t_k,y,\alpha)M(t_\tau,x,\alpha/t_k)\,d\alpha$$

$$- \int_0^h N(t_\tau,x,x_3/t_{k+1})J'(t_{k+1},x_3)\,dx_3F'(t_{k+1},y,t_{k+1}).$$

From (4) and (58) we have

$$\Gamma_\xi \tilde{u}(t_\tau,\xi/t_{k+1}) = 0, \qquad \xi \in \partial D. \tag{64}$$

Multiplying each side of the above equation by $\tilde{u}(t_{k+1},y/t_{k+1})$ and taking the expectation yields $\Gamma_\xi M(t_\tau,\xi,y/t_{k+1}) = 0$, $\xi \in \partial D$. Thus, the following theorem holds.

[Theorem 9] $M(t_\tau,x,y/t_{k+1})$ is given by

$$M(t_\tau,x,y/t_{k+1}) = \int_D G(t_{k+1},t_k,y,\alpha)M(t_\tau,x,\alpha/t_k)\, d\alpha$$

$$- \int_0^h N(t_\tau,x,x_3/t_{k+1})J'(t_{k+1},x_3)\, dx_3\ F'(t_{k+1},y,t_{k+1}), \tag{65}$$

$$M(t_\tau,x,y/t_\tau) = P(t_\tau,x,y/t_\tau), \tag{66}$$

$$\Gamma_\xi M(t_\tau,\xi,y/t_{k+1}) = 0, \qquad \xi \in \partial D. \tag{67}$$

It remains to derive the equation for the optimal smoothing error covariance function $P(t_\tau,x,y/t_{k+1})$. From (57) we have

$$\tilde{u}(t_\tau,x/t_{k+1}) = \tilde{u}(t_\tau,x/t_k) - B(t_\tau,t_{k+1},x,t_{k+1})\nu(t_{k+1}). \tag{68}$$

[Theorem 10] The optimal smoothing error covariance function $P(t_\tau,x,y/t_{k+1})$ is given by

$$P(t_\tau,x,y/t_{k+1}) = P(t_\tau,x,y/t_k)$$

$$- \int_0^h \int_0^h N(t_\tau,x,x_3/t_{k+1})J'(t_{k+1},x_3)\Gamma^{-1}(t_{k+1}/t_k)J(t_{k+1},y_3)$$

$$N(t_\tau,y,y_3/t_{k+1})\, dx_3 dy_3\ , \tag{69}$$

$$\Gamma_\xi P(t_\tau,\xi,y/t_{k+1}) = 0, \qquad \xi \in \partial D. \tag{70}$$

[Proof] From (68) we have

$$P(t_\tau,x,y/t_{k+1}) = E[\tilde{u}(t_\tau,x/t_{k+1})\tilde{u}(t_\tau,y/t_{k+1})]$$

$$= P(t_\tau,x,y/t_k) + B(t_\tau,t_{k+1},x,t_{k+1})\Gamma(t_{k+1}/t_k)B'(t_\tau,t_{k+1},y,t_{k+1})$$

$$- B(t_\tau,t_{k+1},x,t_{k+1})E[\nu(t_{k+1})\tilde{u}(t_\tau,y/t_k)]$$

$$- E[\tilde{u}(t_\tau,x/t_k)\nu'(t_{k+1})]B'(t_\tau,t_{k+1},y,t_{k+1}). \tag{71}$$

But we have

$$E[\tilde{u}(t_\tau,x/t_k)\nu'(t_{k+1})] = \int_0^h\int_D G_L'(t_{k+1},t_k,x_3,\alpha)M(t_\tau,x,\alpha/t_k)$$

$$J'(t_{k+1},x_3)\, d\alpha dx_3$$

and

$$E[\nu(t_{k+1})\tilde{u}(t_\tau,x/t_k)] = \int_0^h\int_D J(t_{k+1},x_3)G_L(t_{k+1},t_k,x_3,\alpha)M(t_\tau,x,\alpha)\, d\alpha dx_3.$$

Substituting the above equations and (47) into (71) yields (69). Multiplying each side of (64) by $\tilde{u}(t_\tau,y/t_{k+1})$ and taking the expectation yields (70). Q.E.D.

[Theorem 11] The optimal smoothing estimator is given by

$$\hat{u}(t_\tau,x/t_k) = \hat{u}(t_\tau,x/t_\tau) + \sum_{\ell=\tau+1}^{k} B(t_\tau,t_\ell,x,t_\ell)\nu(t_\ell) \tag{72}$$

and the optimal smoothing error covariance function $P(t_\tau,x,y/t_k)$ is given by

$$P(t_\tau,x,y/t_k) = P(t_\tau,x,y/t_\tau)$$

$$- \sum_{\ell=\tau+1}^{k} \int_0^h \int_0^h N(t_\tau,x,x_3/t_\ell)J'(t_\ell,x_3)\Gamma^{-1}(t_\ell/t_{\ell-1})$$

$$J(t_\ell,y_3)N(t_\tau,y,y_3/t_\ell)\ dx_3dy_3. \tag{73}$$

VII. Estimation of the Concentration Distribution Downwind of a Continuous, Ground-Level Line Source

There has been much recent interest in the airborne measurement of pollutant concentrations downwind of sources [9] - [11]. Here we wish to consider a hypothetical, but realistic, situation in which an aircraft with a downward-looking instrument, such as for example the JPL Laser Absorption Spectrometer [12], is flown at different altitudes downwind of the source, and total species burdens are measured at a series of downwind distances.

The steady-state concentration of a species downwind of a continuously emitted ground-level line source (e.g. a highway) situated normal to the direction of the wind flow is governed by the following form of the atmospheric diffusion equation [5],

$$V_1 \frac{\partial u}{\partial x_1} = \frac{\partial}{\partial x_3}\left(K_V(x_3) \frac{\partial u}{\partial x_3}\right) + w(x_1,x_3) \tag{74}$$

$$u(0,x_3) = u_0(x_3) \tag{75}$$

$$- K_V(0) \frac{\partial u}{\partial x_3} = \phi\delta(x_1), \qquad x_3 = 0 \tag{76}$$

$$\frac{\partial u}{\partial x_3} = 0, \qquad x_3 = h \tag{77}$$

where ϕ is the constant rate of release. For convenience we will take $K_V = 1$, since vertical variations of this constant are not essential to the estimation problem we will consider. If we let $t = x_1/V_1$ and $x = x_3$, (74)-(77) become

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + w(t,x) \tag{78}$$

$$u(0,x) = u_0(x) \tag{79}$$

$$\frac{\partial u}{\partial x} = -\ \phi\delta(t), \qquad x = 0 \tag{80}$$

$$\frac{\partial u}{\partial x} = 0, \qquad x = h. \tag{81}$$

In this case the measurements $Z(t_k)$ are related to the concentration $u(t_k,x)$ by (8),

$$Z(t_k) = \int_0^h J^n(t_k,x)u(t_k,x)\ dx + v(t_k) \tag{82}$$

where the instrument kernel function will be taken to have the form,

$$J^n(t_k,x) = \begin{cases} 1 & x \le h_n \\ 0 & x > h_n \end{cases} \qquad n = 1,2,\ldots,\ N \tag{83}$$

The theory developed in the prior sections can be applied directly to this problem, and the optimal filter and smoother are given in Table 1. The prediction, filtering and smoothing algorithms were applied to hypothetical data generated by solving (74)-(77) and forming $Z(t_k)$ from (82), using noise processes $w(t,x)$ and $v(t_k)$ with prescribed properties. The algorithms were applied to estimate the concentration distribution $u(t_k,x)$ as a function of height x at several downwind distances, t_1, t_2,... based on measurements taken at one to four elevations. It is of interest to study the behavior of the estimates as a function of downwind distance and of the number of elevations at which data are simultaneously taken. Values of all parameters used in the calculation are given in Table 2.

Figs. 2-4 show selected results of the application of the filtering and smoothing algorithms to the synthetic data of this example. Fig. 2 shows a comparison of the true concentration distribution $u(t_1,x)$ and the filter estimates, $\hat{u}(t_1,x/t_1)$ based on two and four measurement elevations ($t_1 = 0.0002$). As expected, the profile estimated on the basis of four measurement elevations is superior to that based only on two altitudes. Fig. 3 shows similar results at $t_8 = 0.0082$. The filter estimate based on $n = 4$ virtually coincides with the actual concentration distribution. The performance of the smoothing algorithm is illustrated in Fig. 4, in which the true concentration $u(t_\tau,x)$ is compared with the filter estimate, $\hat{u}(t_\tau,x/t_\tau)$, and the smoothed estimates, $\hat{u}(t_\tau,x/t_2)$, and $\hat{u}(t_\tau,x/t_4)$, with $t_\tau = 0.0002$, $t_2 = 0.0012$, and $t_4 = 0.0032$. Table 3 gives the trace of the filtering error covariance matrix, $P(t,x,x/t)$, for the four measurement configurations at three downwind distances t. As expected, the trace decreases as the number of measurement elevations is increased from 1 to 4.

VIII. Conclusions

Filtering and smoothing algorithms for the processing of remote sensing data on atmospheric species concentrations have been derived using Wiener-Hopf theory. The algorithms were applied successfully to estimate concentration distributions from a hypothetical ground-level line source of material (e.g. a highway) based on remote sensing data taken from several elevations at a number of points downwind from the source. Although there has been increasing interest in the remote sensing of airborne concentrations, a data set sufficient for application of the theory developed in this paper does not yet appear to exist. Nevertheless, it is hoped that the availability of the algorithms developed here will facilitate processing of remote sensing data in conjunction with mathematical models of air pollutant behavior.

References

[1] J. C. Barnes, C. J. Bowley, and H. K. Burke, "Evaluation of the Capabilities of Satellite Imagery for Monitoring Regional Air Pollution Episodes," NASA Contractor Report 159107, Langley Research Center, VA., August 1979.

[2] W. A. Lyons, J. C. Dooley, Jr. and K. T. Whitby, "Sattelite Detection of Long-Range Pollution Transport and Sulfate Aerosol Hazes," Atmospheric Environment, Vol. 12, pp. 621-631, 1978.

[3] J. M. Diamante, T. S. Englar, Jr. and A. H. Jazwinski, "Final Report Phase I - Urban Air Quality Estimation Study," NASA Report CR-145097, Langley Research Center, Va., January 1976.

[4] J. F. Kibler and J. T. Suttles, "Air Pollution Model Parameter Estimation Using Simulated LIDAR Data," AIAA Journal, Vol. 15, pp. 1381-1384 1977.

[5] J. H. Seinfeld, Air Pollution: Physical and Chemical Fundamentals, McGraw-Hill, New York, 1975.

[6] Y. Sawaragi, T. Soeda, and S. Omatu, "Modeling, Estimation, and Their Applications for Distributed Parameter Systems," Lecture Notes in Control and Information Sciences, Vol. 11, 1978.

[7] K. E. Bencala and J. H. Seinfeld, "Distributed Parameter Filtering: Boundary Noise and Discrete Observations," Int. J. Systems Science, Vol. 10, No. 5, pp. 493-512, 1979.

[8] Y. Sakawa, "Observability and Related Problems for Partial Differential Equations of Parabolic Type," SIAM J. Control, Vol. 13, No. 1, pp. 14-27, 1975.

[9] A. J. Alkezweeny, N. E. Davis and R. C. Easter, "Comparison of Ozone in Polluted and Clean Air Masses Over Lake Michigan," IAMAP Symposium on Atmospheric Ozone, Boulder, CO, August 4-9, 1980.

[10] S. Sandroni, P. Bacci and D. Anfossi, "Aircraft Observations of Plumes Emitted from Elevated Sources," Atmospheric Environment, Vol. 15, No. 1, pp. 95-100, 1981.

[11] C. J. Nappo, Jr., "Atmospheric Turbulence and Diffusion Estimates Derived from Observations of a Smoke Plume," Atmospheric Environment, Vol. 15, No. 4, pp. 541-548, 1981.

[12] M. S. Shumate, R. T. Menzies, W. B. Grant and D. S. McDougal, "The Laser Absorption Spectrometer: Remote Measurement of Tropospheric Ozone," Applied Optics, Vol. 20, pp. 545-553, 1981.

Table 1. Optimal Filter and Smoother for Line Source Application

Filter	Smoother
(i) $t_k \leq t < t_{k+1}$	$t = t_{k+1}$, $t_\tau < t_{k+1}$
$\frac{\partial \hat{u}}{\partial t} = \frac{\partial^2 u}{\partial x^2}$	$\frac{\partial G}{\partial t} = \frac{\partial^2 G}{\partial x^2}$
$\frac{\partial \hat{u}}{\partial x} = -\phi\delta(t), \quad x = 0$	$\frac{\partial G}{\partial x} = -\phi\delta(t), \quad x = 0$
$\frac{\partial \hat{u}}{\partial x} = 0, \quad x = h$	$\frac{\partial G}{\partial x} = 0, \quad x = h$
$\hat{u}(t_k,x) = \hat{u}(t_k,x/t_k)$	$G(\sigma,\sigma,x,y) = \delta(x-y)$
$\frac{\partial P}{\partial t} = \frac{\partial^2 P}{\partial x^2} + \frac{\partial^2 P}{\partial y^2} + Q$	$\hat{u}(t_\tau,x/t_{k+1}) = \hat{u}(t_\tau,x/t_k) + B(t_\tau,t_{k+1},x,t_{k+1})\nu(t_{k+1})$
$\frac{\partial P}{\partial x} = 0 \quad x = 0,h$	$P(t_\tau,x,y/t_{k+1}) = P(t_\tau,x,y/t_k) - \int_0^h \int_0^h N(t_\tau,x,\alpha/t_{k+1})\Gamma^{-1}(t_{k+1}/t_k) \cdot N(t_\tau,y,\beta/t_{k+1})\, d\alpha d\beta$
$P(t_k,x,y) = P(t_k,x,y/t_k)$	
ii) $t = t_k$	
$\hat{u}(t_{k+1},x/t_{k+1}) = \hat{u}(t_{k+1},x/t_k) + F(t_{k+1},x,t_{k+1}) \nu(t_{k+1})$.	$B(t_\tau,t_{k+1},x,t_{k+1}) = \int_0^h J(t_{k+1},\alpha)N(t_\tau,x,\alpha/t_k)\, d\alpha \; \Gamma^{-1}(t_{k+1}/t_k)$

Table 1. Optimal Filter and Smoother for Line Source Application (Continued)

Filter	Smoother
ii) (Continued)	
$\nu(t_{k+1}) = Z(t_{k+1}) - \int_0^h J(t_{k+1},x)\hat{u}(t_{k+1},x/t_k)dx$	$N(t_\tau,x,\alpha/t_{k+1}) = \int_0^h M(t_\tau,x,\beta/t_k)G(t_{k+1},t_k,\alpha,\beta)\, d\beta$
$F(t_{k+1},x,t_{k+1}) = \int_0^h J(t_{k+1},x)P(t_{k+1},x,y/t_k)dy\Gamma^{-1}(t_{k+1}/t_k)$	$M(t_\tau,x,y/t_k) = \int_0^h G(t_k,t_{k-1},y,\alpha)M(t_\tau,x,\alpha/t_{k-1})d\alpha - \int_0^h J(t_\tau,\alpha)N(t_\tau,x,\alpha/t_k)d\alpha F'(t_k,y,t_k)$
$\Gamma(t_{k+1}/t_k) = \int_0^h\int_0^h J(t_{k+1},x)P(t_{k+1},x,y/t_k)J(t_{k+1},y)\cdot dxdy + R(t_{k+1})$	$M(t_\tau,x,y/t_\tau) = P(t_\tau,x,y/t_\tau)$
$P(t_{k+1},x,y/t_{k+1}) = P(t_{k+1},x,y/t_k) - \int_0^h\int_0^h J(t_{k+1},\alpha)P(t_{k+1},x,\alpha/t_k)\Gamma^{-1}(t_{k+1}/t_k)P(t_{k+1},\beta,y/t_k)\, d\alpha d\beta$	

Table 2. Parameter Values Used in Line Source Estimation Example

Truncation number $N = 5$

Measurement time $t_{k+1} = t_k + 0.001$, $k = 1,2,3,\cdots$
where $t_1 = 0.0002$

Fixed-point time for smoothing $t_\tau = 0.0002$

Constant rate of release $\phi = 0.3$

Measurement points $h_n = n/4$, $n = 1,2,3,4$

Initial values and noise covariances

$$E[u_0(x)] = u_0(x) = \sum_{i=1}^{N} u_i^o \phi_i(x)$$

$$\mathrm{Cov}[u_0(x),u_0(y)] = P_0(x,y) = \sum_{i=1}^{N} p_{ii}^o \phi_i(x)\phi_i(y)$$

$$\mathrm{Cov}[w(t,x),w(s,y)] = Q(t,x,y)\delta(t-s), \quad \mathrm{Cov}[v(t_k),v(t_n)] = R(t_k)\delta_{kn}$$

$$Q(t,x,y) = \sum_{i=1}^{N} q_{ii}\phi_i(x)\phi_i(y), \qquad R(t_k) = \mathrm{diag}[r_1,r_2,r_3,r_4]$$

$$\phi_i(x) = \begin{cases} 1 & i = 1 \\ \sqrt{2}\cos(i-1)\pi x & i \geq 2 \end{cases}$$

$$\lambda_i = -(i-1)^2\pi^2 \qquad i \geq 1$$

i	1	2	3	4	5
u_i^o	3.0	1.0	0.03	0.003	0.0003
p_{ii}^o	1	0.1^2	0.01^2	0.001^2	0.0001^2
q_{ii}	1	0.5^2	0.25^2	0.125^2	0.0625^2
r_i	0.1^2	0.07^2	0.05^2	0.03^2	

Table 3. Trace of the Filtering Error Covariance Matrix $P(t,x,x/t)$

Measurements	$t = 0.0002$	$t = 0.0032$	$t = 0.0062$
4 point (h_1,h_2,h_3,h_4)	0.2405×10^{-1}	0.1189×10^{-1}	0.9616×10^{-2}
3 point (h_1,h_2,h_3)	0.3441×10^{-1}	0.1577×10^{-1}	0.1335×10^{-1}
2 point (h_1,h_2)	0.1678	0.1195	0.1137
1 point (h_1)	0.6700	0.2914	0.2338

Figure Captions

Fig. 1. Remote Sensing Measurement Configuration Considered in This Work.

Fig. 2. Comparison of True Concentration $\hat{u}(t_1,x)$ and the Filter Estimates $u(t_1,x/t_1)$ based on 2 and 4 Measurement Elevations. $t_1 = 0.0002$.

Fig. 3. Comparison of True Concentration $u(t_8,x)$ and the Filter Estimates $\hat{u}(t_8,x/t_8)$ based on 2 and 4 Measurement Elevations. $t_8 = 0.0082$.

Fig. 4. Comparison of True Concentration $u(t_\tau,x)$, the Filter Estimate $\hat{u}(t_\tau,x/t_\tau)$ and the Fixed Point Smoothing Estimates $\hat{u}(t_\tau,x/t_2)$ and $\hat{u}(t_\tau,x/t_4)$. $t_\tau = 0.0002$, $t_2 = 0.0012$, $t_4 = 0.0032$.

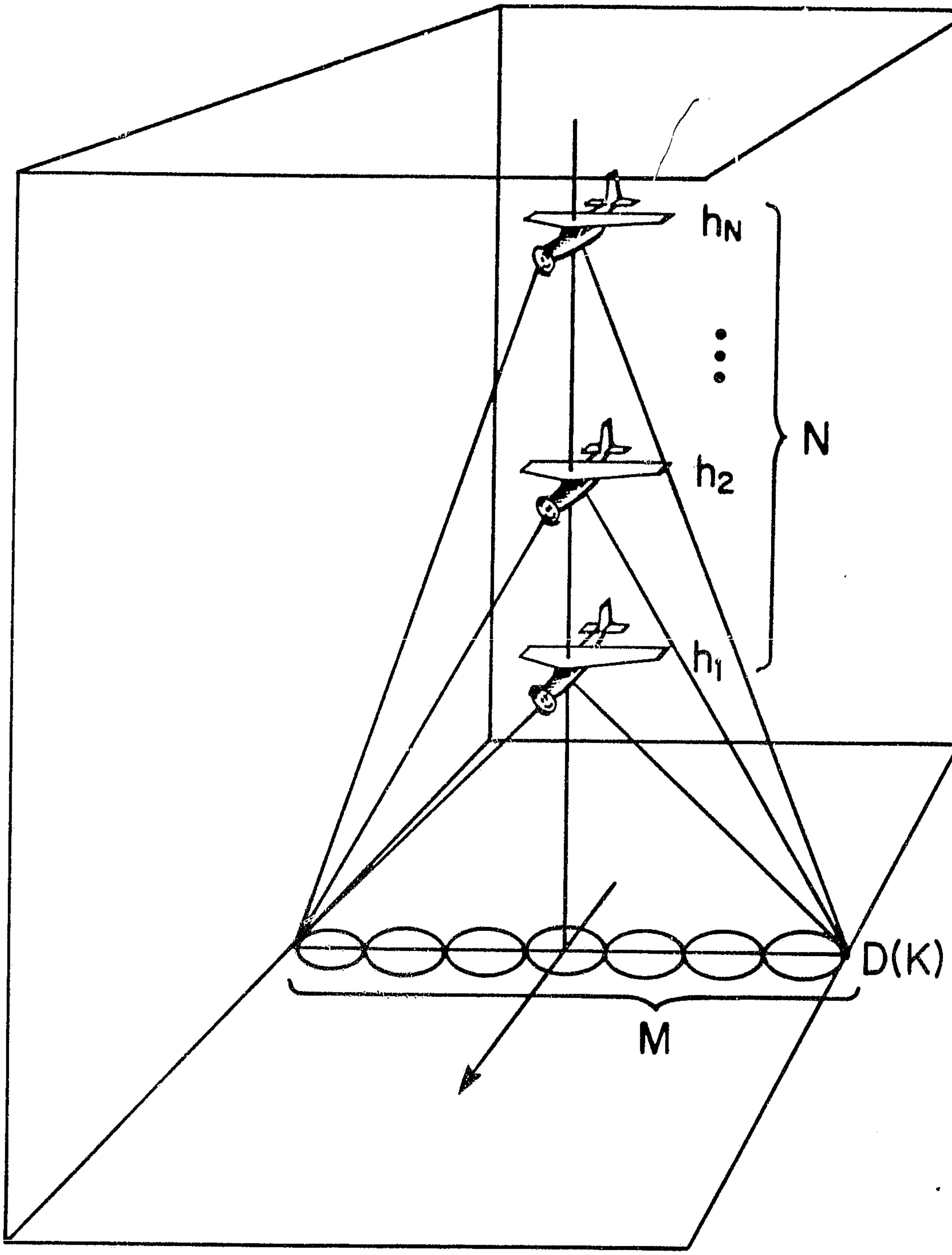


Figure 1

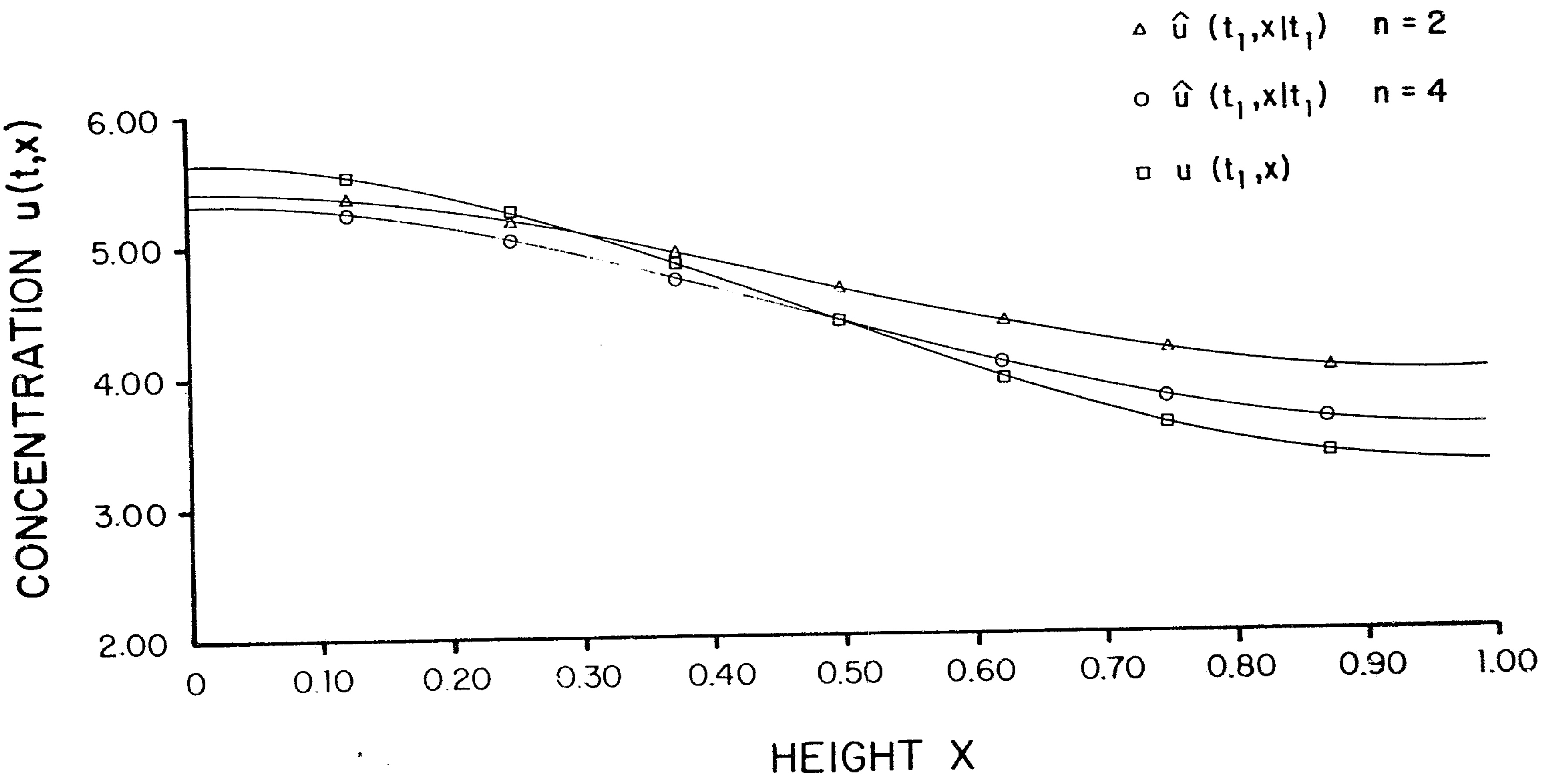
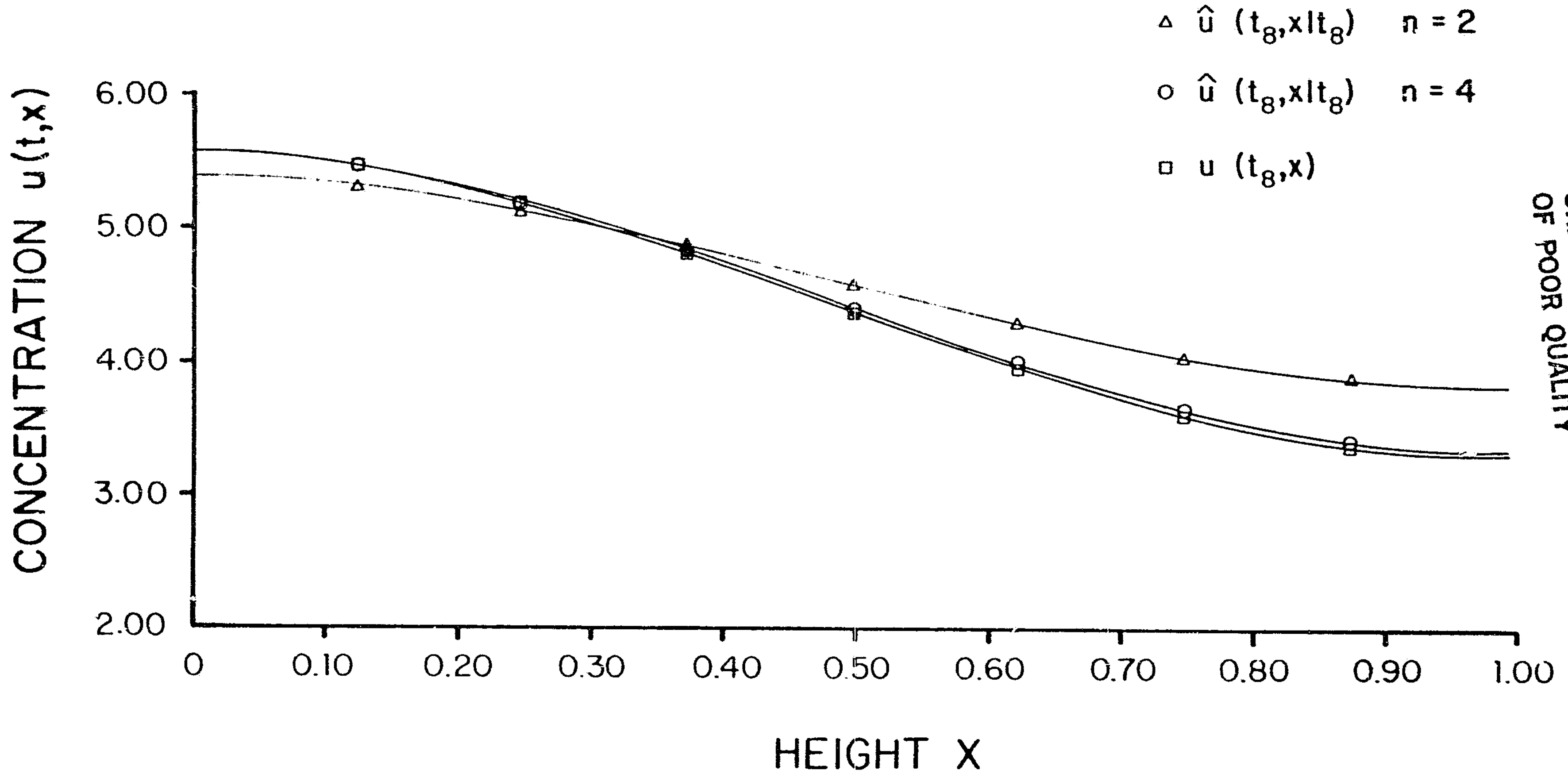

△ $\hat{u}(t_1,x|t_1)$ n = 2
○ $\hat{u}(t_1,x|t_1)$ n = 4
□ $u(t_1,x)$
CONCENTRATION u(t,x)
6.00
5.00
4.00
3.00
2.00
0
0.10
0.20
0.30
0.40
0.50
0.60
0.70
0.80
0.90
1.00
HEIGHT X

Figure 2



△ $\hat{u}$ (t_8,x|t_8) n = 2
○ $\hat{u}$ (t_8,x|t_8) n = 4
□ u (t_8,x)
CONCENTRATION u(t,x)
6.00
5.00
4.00
3.00
2.00
0
0.10
0.20
0.30
0.40
0.50
0.60
0.70
0.80
0.90
1.00
HEIGHT X

Figure 3

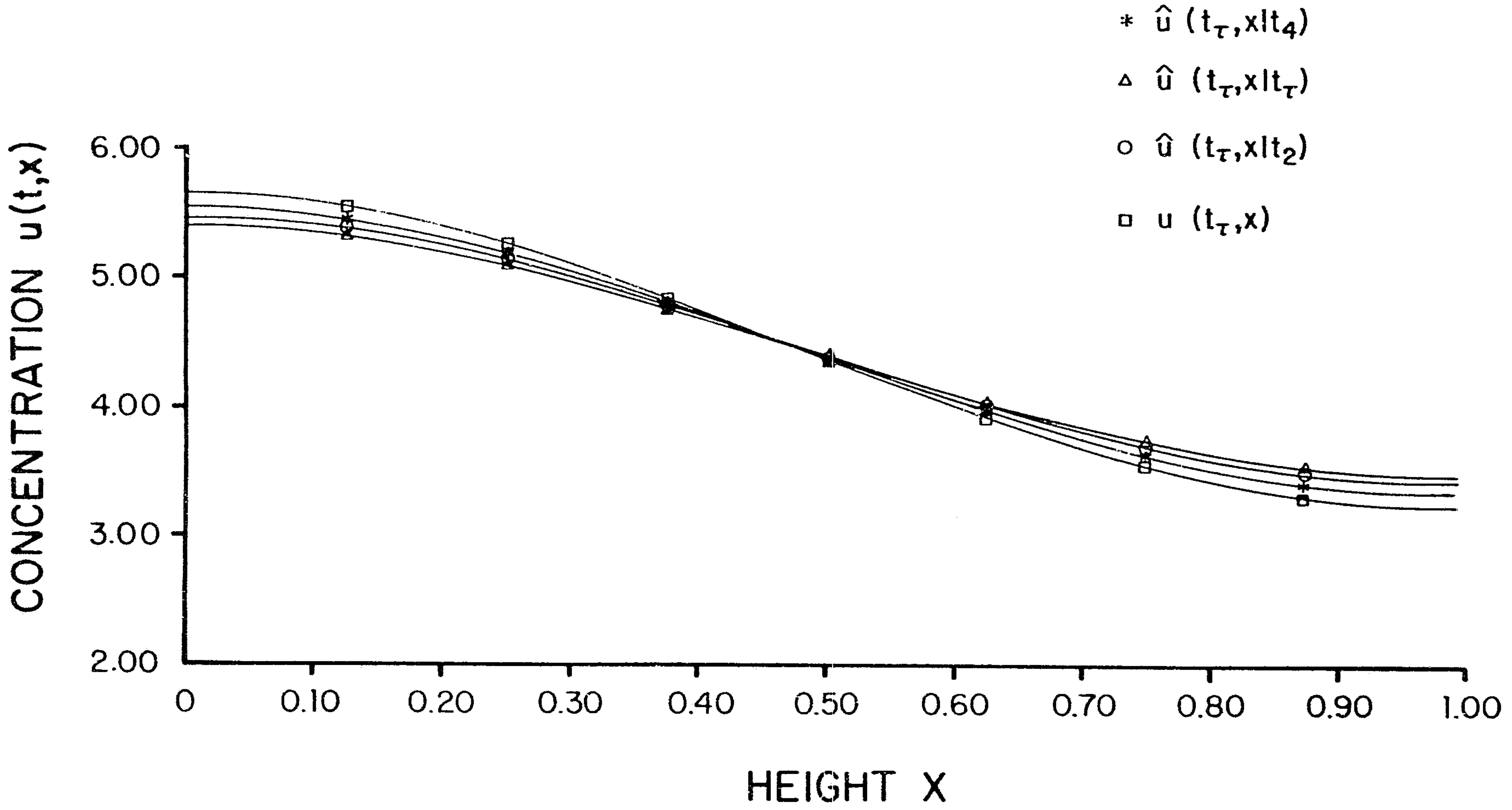

* $\hat{u}\ (t_\tau, x | t_4)$
△ $\hat{u}\ (t_\tau, x | t_\tau)$
○ $\hat{u}\ (t_\tau, x | t_2)$
□ $u\ (t_\tau, x)$
CONCENTRATION u(t,x)
6.00
5.00
4.00
3.00
2.00
0
0.10
0.20
0.30
0.40
0.50
0.60
0.70
0.80
0.90
1.00
HEIGHT X

Figure 4

SUMMARY AND CONCLUSIONS

The object of this research grant was to initiate an evaluation of the analysis of remote sensing data on pollutant concentrations in the troposphere. Remote sensing measurements of pollutant concentrations are becoming increasingly important in understanding the transport and transformation of pollutants over moderate to long distances in the atmosphere. Traditionally such data have not been analyzed beyond the point of constructing mass fluxes and total budgets over a region. The question studied in this research grant was that of the further analysis of such data, particularly when one has a mathematical model available. The specific problem then is to see how typical remote sensing data can be used in conjunction with a mathematical model to extract additional information about the pollutant behavior in the region being studied.

The essential problem is one of estimation, that is, of using the typical remote sensing data to determine full concentration distributions. Once full concentration distributions are available, one can then assess the mechanisms of the process through the mathematical model. The first step in the research was to look theoretically at the question of the minimum amount of data needed to reconstruct a concentration distribution from finite data typical of those collected in remote sensing. Chapter I of this report presents a development and derivation of a condition of reconstructability, namely rigorous conditions that can be applied to a data sampling program to determine whether it will be possible to estimate a species concentration distribution from such measurements. Chapters II and III of this report are then devoted to the development of a numerical algorithm that will process the data to produce concentration distribution estimates in the cases when the data are a priori reconstructable.

Perhaps the most important result of this study is the indication of the types of measurement strategies one should follow in remote sensing programs.

In particular, it appears that the best measurement strategy is to attempt to obtain pollutant burdens at a certain location at a number of elevations at times as close as possible. This strategy is recommended because the vertical distribution of pollutant concentrations in the first 1,000 meters of the atmosphere is a crucial element of a mathematical model of such species. The theory and numerical techniques developed in this study will tell one when devising a measurement program and monitoring strategy the number of vertical levels at which one should make measurements to be able to estimate relatively accurately the complete vertical concentration profile of the species of interest. It is anticipated that these results will be of value to those contemplating remote sensing measurement programs of tropospheric species that involve measurements at several vertical levels.