

Estimation of Atmospheric Species Concentrations from Remote Sensing Data

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

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.

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 episodes [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 de-

rive 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 Wiener-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 theory.

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]¹:

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

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 $u(t_0, x_1, x_2, x_3) = u_0(x_1, x_2, x_3)$, and typical boundary conditions are

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

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

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

¹In 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].

Manuscript received May 15, 1981; revised October 30, 1981. This work was supported by NASA under Grant NAG1-71.

S. Omatu was with the Department of Chemical Engineering, California Institute of Technology, Pasadena, CA 91125. He is now with 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.

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:

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

Let $S(t, \xi)$ be

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

Thus (1) can be represented as

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

and (2) can be written as

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

We assume that the initial condition $u_0(x)$ can be represented as a Gaussian process with statistics

$$\begin{aligned} E[u_0(x)] &= \bar{u}_0(x) \\ E[(u_0(x) - \bar{u}_0(x))(u_0(y) - \bar{u}_0(y))] &= P_0(x, y) \end{aligned} \quad (5)$$

and the random disturbance $w(t, x)$ is stochastically independent of $u_0(x)$ and is a white Gaussian process with statistics

$$\begin{aligned} E[w(t, x)] &= 0 \\ E[w(t, x)w(s, y)] &= Q(t, x, y)\delta(t - s). \end{aligned} \quad (6)$$

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 $(x_1^{m(k)}, x_2^{m(k)}, 0)$, $m = 1, 2, \dots, M$.

We are interested in considering the vertically integrated measurement given by

$$\begin{aligned} Z_{m(k)}(t_k, n) &= \int_0^{h_n} \tilde{J}_{m(k)}(x_3) u(t_k, x_1^{m(k)}, x_2^{m(k)}, x_3) dx_3 \\ &+ v(t_k, x_1^{m(k)}, x_2^{m(k)}, h_n), \\ m &= 1, 2, \dots, M; \quad n = 1, 2, \dots, N \\ h_1 &< h_2 < \dots < h_N; \quad k = 1, 2, \dots \end{aligned} \quad (7)$$

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

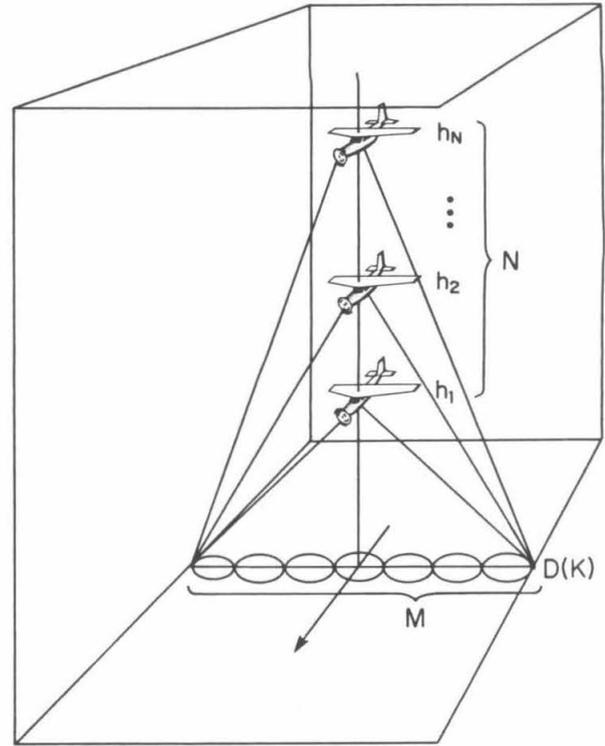


Fig. 1. Remote sensing measurement configuration considered in this work.

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:

$$\begin{aligned} x(m(k)) &= (x_1^{m(k)}, x_2^{m(k)}) \\ J_{m(k)}^n &= \begin{cases} \tilde{J}_{m(k)}^n(x_3), & x_3 \leq h_n \\ 0, & x_3 > h_n \end{cases} \\ J^n(t_k, x_3) &= \begin{bmatrix} J_{1(k)}^n(x_3) & \dots & 0 \\ \vdots & \ddots & \vdots \\ 0 & \dots & J_{M(k)}^n(x_3) \end{bmatrix} \end{aligned}$$

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

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

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

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

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

and

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

Then (7) can be represented compactly as

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

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 $E[v(t_k)v'(t_l)] = R(t_k) \delta_{kl}$, where $'$ denotes the transpose operator and $R(t_k)$ is an $MN \times MN$ positive-definite matrix.

The problem considered here is to estimate $u(t, x)$ over D on the basis of the measurement $Z(t_\sigma)$, $\sigma = 0, 1, \dots, k$. 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 $Z(t_\sigma)$, $\sigma = 0, 1, \dots, k$ by $\hat{u}(t_\tau, x/t_k)$ which is given by the following linear transformation of $Z(t_\sigma)$, $\sigma = 0, 1, \dots, k$:

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

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

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

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 ($t > t_k$)

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

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

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

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

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

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 $\zeta = 0, 1, \dots, k$ and $x \in \bar{D} = D \cup \partial D$:

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

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

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

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 $\tilde{F}(t_\tau, x, t_\sigma) + \tilde{F}_\Delta(t_\tau, x, t_\sigma)$ be also the optimal kernel function satisfying the Wiener-Hopf equation

(14). Then it follows that $\tilde{F}_\Delta(t_\tau, x, t_\sigma) \equiv 0$, $\sigma = 0, 1, \dots$, 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 ($t > t_k$)

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

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

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

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

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

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

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

for $\zeta = 0, 1, \dots, k$ 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

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

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

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

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

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

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

$$\Gamma_\xi P(t, \xi, y/t_k) = 0, \quad \xi \in \partial D. \quad (23)$$

Proof: From (3) and (19) we have

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

and from (4) and (20)

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

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

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

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

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

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

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

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

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

$$\begin{aligned} P(t, x, y/t_k) &= \int_D \int_D G(t, t_k, x, \alpha) P(t_k, \alpha, \beta/t_k) \\ &\cdot G(t, t_k, y, \beta) d\alpha d\beta + \int_{t_k}^t \int_D \int_D G(t, \sigma, x, \alpha) \\ &\cdot Q(\sigma, \alpha, \beta) G(t, \sigma, y, \beta) d\alpha d\beta d\sigma. \end{aligned} \quad (28)$$

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

$$\begin{aligned} u(t_{k+1}, x) &= \int_D G(t_{k+1}, t_k, x, \alpha) u(t_k, \alpha) d\alpha \\ &\quad + \int_{t_k}^{t_{k+1}} \int_D G(t_{k+1}, \eta, x, \alpha) d\alpha d\eta \quad (29) \\ u_{t_{k+1}}(x_3) &= \int_D G_M(t_{k+1}, t_k, x_3, \alpha) u(t_k, \alpha) d\alpha \\ &\quad + \int_{t_k}^{t_{k+1}} \int_D G_M(t_{k+1}, \eta, x_3, \alpha) w(\eta, \alpha) d\alpha d\eta \quad (30) \end{aligned}$$

where

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

From (17) we have

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

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

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

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

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

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

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

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

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

$$\begin{aligned} E[Z(t_{k+1})Z'(t_\zeta)] &= \int_0^h J(t_{k+1}, x_3) \\ &\quad \cdot \int_D G_M(t_{k+1}, t_k, x_3, \alpha) E[u(t_k, \alpha)Z'(t_\zeta)] d\alpha dx_3. \end{aligned}$$

Again we use the Wiener-Hopf equation (17) in the above equation and

$$\begin{aligned} E[Z(t_{k+1})Z'(t_\zeta)] &= \int_0^h J(t_{k+1}, x_3) \int_D G_M(t_{k+1}, t_k, x_3, \alpha) \\ &\quad \cdot \sum_{\sigma=0}^k F(t_k, \alpha, t_\sigma) E[Z(t_\sigma)Z'(t_\zeta)] d\alpha dx_3. \quad (34) \end{aligned}$$

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

$$\begin{aligned} \tilde{F}_\Delta(t_k, x, t_\sigma) &= F(t_{k+1}, x, t_{k+1}) \int_0^h J(t_{k+1}, x_3) \\ &\quad \cdot \int_D G_M(t_{k+1}, t_k, x_3, \alpha) F(t_k, \alpha, t_\sigma) d\alpha dx_3 \\ &\quad + F(t_{k+1}, x, t_\sigma) - \int_D G(t_{k+1}, t_k, x, \alpha) F(t_k, \alpha, t_\sigma) d\alpha. \end{aligned}$$

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

$$\begin{aligned} F(t_{k+1}, x, t_\sigma) &= \int_D G(t_{k+1}, t_k, x, \alpha) F(t_k, \alpha, t_\sigma) d\alpha \\ &\quad - F(t_{k+1}, x, t_{k+1}) \int_0^h J(t_{k+1}, x_3) \\ &\quad \cdot \int_D G_M(t_{k+1}, t_k, x_3, \alpha) F(t_k, \alpha, t_\sigma) d\alpha dx_3. \quad (35) \end{aligned}$$

Theorem 4: The optimal filtering estimate $\hat{u}(t_{k+1}, x/t_{k+1})$ is given by

$$\hat{u}(t_{k+1}, x/t_{k+1}) = \hat{u}(t_{k+1}, x/t_k) + F(t_{k+1}, x, t_{k+1})v(t_{k+1}) \quad (36)$$

$$\nu(t_{k+1}) \triangleq Z(t_{k+1}) - \int_0^h J(t_{k+1}, x_3) \cdot \hat{u}_{t_{k+1}}(x_3/t_k) dx_3 \quad (37)$$

$$\hat{u}(t_0, x/t_0) = \bar{u}_0(x) \quad (38)$$

$$\Gamma_{\xi} \hat{u}(t_{k+1}, \xi/t_{k+1}) = S(t_{k+1}, \xi), \quad \xi \in \partial D \quad (39)$$

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

$$\begin{aligned} \hat{u}(t_{k+1}, x/t_{k+1}) &= F(t_{k+1}, x, t_{k+1})Z(t_{k+1}) \\ &+ \int_D G(t_{k+1}, t_k, x, \alpha) \sum_{\sigma=0}^k F(t_k, \alpha, t_{\sigma})Z(t_{\sigma}) d\alpha \\ &- F(t_{k+1}, x, t_{k+1}) \int_0^h J(t_{k+1}, x_3) \\ &\cdot \int_D G_M(t_{k+1}, t_k, x_3, \alpha) \\ &\cdot \sum_{\sigma=0}^k F(t_k, \alpha, t_{\sigma})Z(t_{\sigma})Z(t_{\sigma}) d\alpha dx_3. \end{aligned}$$

Then from (12) and (27) we have

$$\begin{aligned} \hat{u}(t_{k+1}, x/t_{k+1}) &= \int_D G(t_{k+1}, t_k, x, \alpha) \hat{u}(t_k, \alpha/t_k) d\alpha \\ &+ F(t_{k+1}, x, t_{k+1}) \left(Z(t_{k+1}) - \int_0^h J(t_{k+1}, x_3) \right. \\ &\cdot \left. \int_D G_M(t_{k+1}, t_k, x_3, \alpha) \hat{u}(t_k, \alpha/t_k) d\alpha dx_3 \right) \\ &= \hat{u}(t_{k+1}, x/t_k) + F(t_{k+1}, x, t_{k+1})\nu(t_{k+1}). \end{aligned}$$

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:

$$\begin{aligned} P_M(t_{\tau}, x, y_3/t_k) &= [P(t_{\tau}, x, y_1^{1(k)}/t_k), \dots, \\ &P(t_{\tau}, x, y_1^{M(k)}/t_k)] \end{aligned} \quad (40)$$

and

$$\begin{aligned} P_{MM}(t_{\tau}, x_3, y_3/t_k) &= \begin{bmatrix} P(t_{\tau}, x_1^{1(k)}, y_3/t_k) \\ \vdots \\ P(t_{\tau}, x_1^{M(k)}, y_3/t_k) \end{bmatrix} \\ &= \begin{bmatrix} P(t_{\tau}, x_1^{1(k)}, y_1^{1(k)}/t_k), \dots, P(t_{\tau}, x_1^{1(k)}, y_1^{M(k)}/t_k) \\ \vdots \\ P(t_{\tau}, x_1^{M(k)}, y_1^{1(k)}/t_k), \dots, P(t_{\tau}, x_1^{M(k)}, y_1^{M(k)}/t_k) \end{bmatrix} \end{aligned} \quad (41)$$

where $x^{m(k)} = (x_1^{m(k)}, x_2^{m(k)}, x_3)$ and $y^{m(k)} = (y_1^{m(k)}, y_2^{m(k)}, y_3)$, $m = 1, 2, \dots, M$. 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

$$P_M(t_{\tau}, x, y_3/t_k) = E[\tilde{u}(t_{\tau}, x/t_k)\tilde{u}_{t_{\tau}}(y_3/t_k)] \quad (42)$$

and

$$P_{MM}(t_{\tau}, x_3, y_3/t_k) = E[\tilde{u}_{t_{\tau}}(x_3/t_k)\tilde{u}_{t_{\tau}}(y_3/t_k)] \quad (43)$$

where

$$\tilde{u}_{t_{\tau}}(x_3/t_k) = u_{t_{\tau}}(x_3) - \hat{u}_{t_{\tau}}(x_3/t_k) \quad (44)$$

and

$$\hat{u}_{t_{\tau}}(x_3/t_k) = \begin{bmatrix} \hat{u}(t_{\tau}, x_1^{1(k)}/t_k) \\ \vdots \\ \hat{u}(t_{\tau}, x_1^{M(k)}/t_k) \end{bmatrix}. \quad (45)$$

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

$$\begin{aligned} \Gamma(t_{k+1}/t_k) &= \int_0^h \int_0^h J(t_{k+1}, x_3) P_{MM}(t_{k+1}, x_3, y_3/t_k) \\ &\cdot J'(t_{k+1}, y_3) dx_3 dy_3 + R(t_{k+1}). \end{aligned} \quad (46)$$

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

$$\begin{aligned} F(t_{k+1}, x, t_{k+1}) &= \int_0^h P_M(t_{k+1}, x, y_3/t_k) J'(t_{k+1}, y_3) \\ &\cdot dy_3 \Gamma^{-1}(t_{k+1}/t_k). \end{aligned} \quad (47)$$

Proof: From the Wiener-Hopf equation (17) we have

$$\begin{aligned} F(t_{k+1}, x, t_{k+1}) E[Z(t_{k+1})Z'(t_{k+1})] \\ + \sum_{\sigma=0}^k F(t_{k+1}, x, t_{\sigma}) E[Z(t_{\sigma})Z'(t_{k+1})] \\ = E[u(t_{k+1}, x)Z'(t_{k+1})]. \end{aligned}$$

Substituting (35) into the above equation yields

$$\begin{aligned} F(t_{k+1}, x, t_{k+1}) E[(Z(t_{k+1}) - \int_0^h J(t_{k+1}, x_3) \hat{u}_{t_{k+1}} \\ \cdot (x_3/t_k) dx_3) Z'(t_{k+1})] = E[(u(t_{k+1}, x) \\ - \hat{u}(t_{k+1}, x/t_k)) Z'(t_{k+1})]. \end{aligned}$$

Using (8) and the orthogonality condition of Corollary 1 yields

$$\begin{aligned} E[\tilde{u}(t_{k+1}, x/t_k) Z'(t_{k+1})] &= \int_0^h E[\tilde{u}(t_{k+1}, x/t_k) \\ &\cdot u'_{t_{k+1}}(x_3)] J'(t_{k+1}, x_3) dx_3 \\ &= \int_0^h P_M(t_{k+1}, x, x_3/t_k) J'(t_{k+1}, x_3) dx_3 \end{aligned}$$

and

$$\begin{aligned} E[v(t_{k+1})Z'(t_{k+1})] &= \int_0^h \int_0^h J(t_{k+1}, x_3) \\ &\cdot P_{MM}(t_{k+1}, x_3, y_3/t_k) J'(t_{k+1}, y_3) dx_3 dy_3 \\ &+ R(t_{k+1}) = \Gamma(t_{k+1}/t_k). \end{aligned} \quad (48)$$

Then we have

$$\begin{aligned} F(t_{k+1}, x, t_{k+1})\Gamma(t_{k+1}/t_k) \\ = \int_0^h P_M(t_{k+1}, x, x_3/t_k) J'(t_{k+1}, x_3) dx_3 \end{aligned}$$

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

$$\begin{aligned} P(t_{k+1}, x, y/t_{k+1}) &= P(t_{k+1}, x, y/t_k) \\ &- \int_0^h \int_0^h P_M(t_{k+1}, x, x_3/t_k) J'(t_{k+1}/x_3) \Gamma^{-1}(t_{k+1}/t_k) \\ &\cdot J(t_{k+1}, y_3) P'_M(t_{k+1}, y, y_3/t_k) dx_3 dy_3 \end{aligned} \quad (49)$$

$$P(t_0, x, y/t_0) = P_0(x, y) \quad (50)$$

$$\Gamma_\xi P(t_{k+1}, \xi, y/t_k) = 0, \quad \xi \in \partial D. \quad (51)$$

Proof: From (3) and (36) we have

$$\tilde{u}(t_{k+1}, x/t_{k+1}) = \tilde{u}(t_{k+1}, x/t_k) - F(t_{k+1}, x, t_{k+1})v(t_{k+1}) \quad (52)$$

and from (4) and (39)

$$\Gamma_\xi \tilde{u}(t_{k+1}, \xi/t_{k+1}) = 0, \quad \xi \in \partial D. \quad (53)$$

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

$$\begin{aligned} P(t_{k+1}, x, y/t_{k+1}) &= E[\tilde{u}(t_{k+1}, x/t_{k+1})\tilde{u}(t_{k+1}, y/t_{k+1})] \\ &= P(t_{k+1}, x, y/t_k) + F(t_{k+1}, x, t_{k+1})E[v(t_{k+1})v'(t_{k+1})] \\ &\cdot F'(t_{k+1}, y, t_{k+1}) - F(t_{k+1}, x, t_{k+1}) \\ &\cdot \int_0^h J(t_{k+1}, x_3) E[\tilde{u}_{t_{k+1}}(x_3/t_k)\tilde{u}(t_{k+1}, y/t_k)] dx_3 \\ &- \int_0^h E[\tilde{u}(t_{k+1}, x/t_k)\tilde{u}_{t_{k+1}}(y_3/t_k)] J'(t_{k+1}, y_3) \\ &\cdot dy_3 F'(t_{k+1}, y, t_{k+1}). \end{aligned}$$

Using (40) and (47) yields

$$\begin{aligned} P(t_{k+1}, x, y/t_{k+1}) &= P(t_{k+1}, x, y/t_k) \\ &- \int_0^h \int_0^h P_M(t_{k+1}, x, x_3/t_k) J'(t_{k+1}, x_3) \Gamma^{-1}(t_{k+1}/t_k) \\ &\cdot J(t_{k+1}, y_3) P'_M(t_{k+1}, y, y_3/t_k) dx_3 dy_3. \end{aligned}$$

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

$$\begin{aligned} B(t_\tau, t_{k+1}, x, t_\sigma) &= B(t_\tau, t_k, x, t_\sigma) - B(t_\tau, t_{k+1}, x, t_{k+1}) \\ &\cdot \int_0^h J(t_{k+1}, x_3) \int_D G_M(t_{k+1}, t_k, x_3, \alpha) \\ &\cdot F(t_k, \alpha, t_\sigma) d\alpha dx_3. \end{aligned} \quad (54)$$

Proof: From the Wiener-Hopf equation (18) for the smoothing problem we have

$$\begin{aligned} \sum_{\sigma=0}^{k+1} B(t_\tau, t_{k+1}, x, t_\sigma) E[Z(t_\sigma)Z'(t_\xi)] &= E[u(t_\tau, x)Z'(t_\xi)], \\ \xi &= 0, 1, \dots, k+1 \end{aligned} \quad (55)$$

and

$$\begin{aligned} \sum_{\sigma=0}^k B(t_\tau, t_k, x, t_\sigma) E[Z(t_\sigma)Z'(t_\xi)] &= E[u(t_\tau, x)Z'(t_\xi)], \\ \xi &= 0, 1, \dots, k. \end{aligned} \quad (56)$$

Subtracting (56) from (55) yields

$$\begin{aligned} B(t_\tau, t_{k+1}, x, t_{k+1}) E[Z(t_{k+1})Z'(t_\xi)] \\ + \sum_{\sigma=0}^k (B(t_\tau, t_{k+1}, x, t_\sigma) - B(t_\tau, t_k, x, t_\sigma)) E[Z(t_\sigma) \\ \cdot Z'(t_\xi)] = 0. \end{aligned}$$

From (8) and (17) we have

$$\begin{aligned} E[Z(t_{k+1})Z'(t_\xi)] &= \int_0^h J(t_{k+1}, x_3) \int_D G_M(t_{k+1}, t_k, x_3, \alpha) \\ &\cdot E[u(t_k, \alpha)Z'(t_\xi)] d\alpha dx_3 = \int_0^h J(t_{k+1}, x_3) \\ &\cdot \int_D G_M(t_{k+1}, t_k, x_3, \alpha) \sum_{\sigma=0}^k F(t_k, \alpha, t_\sigma) \\ &\cdot E[Z(t_\sigma)Z'(t_\xi)] d\alpha dx_3. \end{aligned}$$

Then it follows that

$$\sum_{\sigma=0}^k \tilde{F}_\Delta(t_\tau, t_k, x, t_\sigma) E[Z(t_\sigma)Z'(t_\xi)] = 0$$

where

$$\begin{aligned} \tilde{F}_\Delta(t_\tau, t_k, x, t_\sigma) &= B(t_\tau, t_{k+1}, x, t_\sigma) - B(t_\tau, t_k, x, t_\sigma) \\ &+ B(t_\tau, t_{k+1}, x, t_{k+1}) \int_0^h J(t_{k+1}, x_3) \\ &\cdot \int_D G_M(t_{k+1}, t_k, x_3, \alpha) F(t_k, \alpha, t_\sigma) d\alpha dx_3. \end{aligned}$$

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 $\tilde{F}_\Delta(t_\tau, t_k, x, t_\sigma) \equiv 0, \sigma = 0, 1, \dots, k$. 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

$$\hat{u}(t_\tau, x/t_{k+1}) = \hat{u}(t_\tau, x/t_k) + B(t_\tau, t_{k+1}, x, t_{k+1})\nu(t_{k+1}) \quad (57)$$

$$\Gamma_\xi \hat{u}(t_\tau, \xi/t_{k+1}) = S(\tau, \xi), \quad \xi \in \partial D. \quad (58)$$

Proof: From (13) it follows that

$$\begin{aligned} \hat{u}(t_\tau, x/t_{k+1}) &= B(t_\tau, t_{k+1}, x, t_{k+1})Z(t_{k+1}) \\ &+ \sum_{\sigma=0}^k B(t_\tau, t_{k+1}, x, t_\sigma)Z(t_\sigma). \end{aligned}$$

Substituting (54) into the above equation yields

$$\begin{aligned} \hat{u}(t_\tau, x/t_{k+1}) &= B(t_\tau, t_{k+1}, x, t_{k+1})\nu(t_{k+1}) \\ &+ \sum_{\sigma=0}^k B(t_\tau, t_k, x, t_\sigma)Z(t_\sigma) \end{aligned}$$

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

$$\begin{aligned} B(t_\tau, t_{k+1}, x, t_{k+1}) &= \int_0^h N(t_\tau, x, x_3/t_{k+1})J'(t_{k+1}, x_3) \\ &\cdot dx_3 \Gamma^{-1}(t_{k+1}/t_k) \end{aligned} \quad (59)$$

where

$$N(t_\tau, x, x_3/t_{k+1}) = \int_D M(t_\tau, x, y/t_k)G'_M(t_{k+1}, t_k, x_3, y) dy \quad (60)$$

$$M(t_\tau, x, y/t_k) = E[\tilde{u}(t_\tau, x/t_k)\tilde{u}(t_k, y/t_k)]. \quad (61)$$

Proof: From the Wiener-Hopf equation (18) we have

$$\begin{aligned} B(t_\tau, t_{k+1}, x, t_{k+1})E[Z(t_{k+1})Z'(t_{k+1})] \\ + \sum_{\sigma=0}^k B(t_\tau, t_{k+1}, x, t_\sigma)E[Z(t_\sigma)Z'(t_{k+1})] \\ = E[u(t_\tau, x)Z'(t_{k+1})]. \end{aligned}$$

Substituting (54) into the above equation yields

$$B(t_\tau, t_{k+1}, x, t_{k+1})E[\nu(t_{k+1})] = E[\tilde{u}(t_\tau, x/t_k)Z'(t_{k+1})]. \quad (62)$$

On the other hand, from (27) and (29)

$$\begin{aligned} \tilde{u}(t_{k+1}, x/t_k) &= \int_D G(t_{k+1}, t_k, x, y)\tilde{u}(t_k, y/t_k) dy \\ &+ \int_{t_k}^{t_{k+1}} \int_D G(t_{k+1}, \eta, x, y)w(\eta, y) dy d\eta. \end{aligned}$$

Then we have

$$\begin{aligned} E[\tilde{u}(t_\tau, x/t_k)Z'(t_{k+1})] &= \int_0^h \int_D M(t_\tau, x, y/t_k)G'_M \\ &\cdot (t_{k+1}, t_k, x_3, y) dy J'(t_{k+1}, x_3) dx_3 \\ &= \int_0^h N(t_\tau, x, x_3/t_{k+1})J'(t_{k+1}, x_3) dx_3. \end{aligned}$$

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

$$M(t_\tau, x, y/t_{k+1}) = E[u(t_\tau, x)\tilde{u}(t_{k+1}, y/t_{k+1})]. \quad (63)$$

Substituting (52) into the above equation yields

$$\begin{aligned} M(t_\tau, x, y/t_{k+1}) &= \int_D G(t_{k+1}, t_k, y, \alpha)M(t_\tau, x, \alpha/t_k) d\alpha \\ &- \int_0^h N(t_\tau, x, x_3/t_{k+1})J'(t_{k+1}, x_3) \\ &\cdot dx_3 F'(t_{k+1}, y, t_{k+1}). \end{aligned}$$

From (4) and (58) we have

$$\Gamma_\xi \tilde{u}(t_\tau, \xi/t_{k+1}) = 0, \quad \xi \in \partial D. \quad (64)$$

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

$$\begin{aligned} M(t_\tau, x, y/t_{k+1}) &= \int_D G(t_{k+1}, t_k, y, \alpha)M(t_\tau, x, \alpha/t_k) d\alpha \\ &- \int_0^h N(t_\tau, x, x_3/t_{k+1})J'(t_{k+1}, x_3) \\ &\cdot dx_3 F'(t_{k+1}, y, t_{k+1}) \end{aligned} \quad (65)$$

$$M(t_\tau, x, y/t_\tau) = P(t_\tau, x, y/t_\tau) \quad (66)$$

$$\Gamma_\xi M(t_\tau, \xi, y/t_{k+1}) = 0, \quad \xi \in \partial D. \quad (67)$$

TABLE I
OPTIMAL FILTER AND SMOOTHER FOR LINE SOURCE APPLICATION

Filter	Smoother
(i) $t_k \leq t < t_{k+1}$ $\frac{\partial \hat{u}}{\partial t} = \frac{\partial^2 \hat{u}}{\partial x^2}$ $\frac{\partial \hat{u}}{\partial x} = -\phi \delta(t), \quad x = 0$ $\frac{\partial \hat{u}}{\partial x} = 0, \quad x = h$ $\hat{u}(t_k, x) = \hat{u}(t_k, x/t_k)$ $\frac{\partial P}{\partial t} = \frac{\partial^2 P}{\partial x^2} + \frac{\partial^2 P}{\partial y^2} + Q$ $\frac{\partial P}{\partial x} = 0, \quad x = 0, h$ $P(t_k, x, y) = P(t_k, x, y/t_k)$	$t = t_{k+1}, \quad t_\tau < t_{k+1}$ $\frac{\partial G}{\partial t} = \frac{\partial^2 G}{\partial x^2}$ $\frac{\partial G}{\partial x} = -\phi \delta(t), \quad x = 0$ $\frac{\partial G}{\partial x} = 0, \quad x = h$ $G(\sigma, \sigma, x, y) = \delta(x-y)$ $\hat{u}(t_\tau, x/t_{k+1}) = \hat{u}(t_\tau, x/t_k) + B(t_\tau, t_{k+1}, x, t_{k+1})v(t_{k+1})$ $P(t_\tau, x, y/t_{k+1}) = P(t_\tau, x, y/t_k)$ $- \int_0^h \int_0^h N(t_\tau, x, \alpha/t_{k+1}) \Gamma^{-1}(t_{k+1}/t_k) \cdot$ $N(t_\tau, y, \beta/t_{k+1}) \, d\alpha \, d\beta$
(ii) $t = t_k$ $\hat{u}(t_{k+1}, x/t_{k+1}) = \hat{u}(t_{k+1}, x/t_k) + F(t_{k+1}, x, t_{k+1})$ $v(t_{k+1})$ $v(t_{k+1}) = Z(t_{k+1}) - \int_0^h J(t_{k+1}, x) \hat{u}(t_{k+1}, x/t_k) \, dx$ $F(t_{k+1}, x, t_{k+1}) = \int_0^h J(t_{k+1}, x) P(t_{k+1}, x, y/t_k) \, dy \Gamma^{-1}(t_{k+1}/t_k)$ $\Gamma(t_{k+1}/t_k) = \int_0^h \int_0^h J(t_{k+1}, x) P(t_{k+1}, x, y/t_k) J(t_{k+1}, y) \cdot$ $\, dx \, dy + R(t_{k+1})$ $P(t_{k+1}, x, y/t_{k+1}) = P(t_{k+1}, x, y/t_k)$ $- \int_0^h \int_0^h J(t_{k+1}, \alpha) P(t_{k+1}, x, \alpha/t_k) \cdot$ $\Gamma^{-1}(t_{k+1}/t_k) P(t_{k+1}, \beta, y/t_k) \, d\alpha \, d\beta$	$B(t_\tau, t_{k+1}, x, t_{k+1}) = \int_0^h J(t_{k+1}, \alpha) N(t_\tau, x, \alpha/t_k) \, d\alpha$ $\Gamma^{-1}(t_{k+1}/t_k)$ $N(t_\tau, x, \alpha/t_{k+1}) = \int_0^h M(t_\tau, x, \beta/t_k) G(t_{k+1}, t_k, \alpha, \beta) \, d\beta$ $M(t_\tau, x, y/t_k) = \int_0^h G(t_k, t_{k-1}, y, \alpha) M(t_\tau, x, \alpha/t_{k-1}) \, d\alpha$ $- \int_0^h J(t_\tau, \alpha) N(t_\tau, x, \alpha/t_k) \, d\alpha F^{-1}(t_k, y, t_k)$ $M(t_\tau, x, y/t_\tau) = P(t_\tau, x, y/t_\tau)$

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

$$\tilde{u}(t_\tau, x/t_{k+1}) = \tilde{u}(t_\tau, x/t_k) - B(t_\tau, t_{k+1}, x, t_{k+1})v(t_{k+1}). \quad (68)$$

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

$$\begin{aligned} P(t_\tau, x, y/t_{k+1}) &= P(t_\tau, x, y/t_k) \\ &- \int_0^h \int_0^h N(t_\tau, x, \alpha_3/t_{k+1}) J'(t_{k+1}, \alpha_3) \\ &\cdot \Gamma^{-1}(t_{k+1}/t_k) J(t_{k+1}, \alpha_3) \\ &\cdot N(t_\tau, y, \beta_3/t_{k+1}) \, dx_3 \, dy_3 \end{aligned} \quad (69)$$

and

$$\Gamma_\xi P(t_\tau, \xi, y/t_{k+1}) = 0, \quad \xi \in \partial D. \quad (70)$$

Proof: From (68) we have

$$\begin{aligned} P(t_\tau, x, y/t_{k+1}) &= E[\tilde{u}(t_\tau, x/t_{k+1})\tilde{u}(t_\tau, y/t_{k+1})] \\ &= P(t_\tau, x, y/t_k) + B(t_\tau, t_{k+1}, x, t_{k+1}) \\ &\cdot \Gamma(t_{k+1}/t_k) B'(t_\tau, t_{k+1}, y, t_{k+1}) \\ &- B(t_\tau, t_{k+1}, x, t_{k+1}) E[v(t_{k+1})\tilde{u}(t_\tau, y/t_k)] \\ &- E[\tilde{u}(t_\tau, x/t_k)v'(t_{k+1})] B'(t_{k+1}) \\ &\cdot B'(t_\tau, t_{k+1}, y, t_{k+1}). \end{aligned} \quad (71)$$

(69) But we have

$$E[\tilde{u}(t_\tau, x/t_k) v'(t_{k+1})] = \int_0^h \int_D G'_L(t_{k+1}, t_k, x_3, \alpha) \cdot M(t_\tau, x, \alpha/t_k) J'(t_{k+1}, x_3) d\alpha dx_3$$

and

$$E[v(t_{k+1}) \tilde{u}(t_\tau, x/t_k)] = \int_0^h \int_D J(t_{k+1}, x_3) \cdot G_L(t_{k+1}, t_k, x_3, \alpha) M(t_\tau, x, \alpha) d\alpha dx_3.$$

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

$$\hat{u}(t_\tau, x/t_k) = \hat{u}(t_\tau, x/t_\tau) + \sum_{l=\tau+1}^k B(t_\tau, t_l, x, t_l) v(t_l) \quad (72)$$

and the optimal smoothing error covariance function $P(t_\tau, x, y/t_k)$ is given by

$$P(t_\tau, x, y/t_k) = P(t_\tau, x, y/t_\tau) - \sum_{l=\tau+1}^k \int_0^h \int_0^h N(t_\tau, x, x_3/t_l) J'(t_l, x_3) \Gamma^{-1}(t_l/t_{l-1}) \cdot J(t_l, y_3) N(t_\tau, y, y_3/t_l) dx_3 dy_3. \quad (73)$$

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 [10]-[12]. 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 [13], 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]:

$$V_1 \frac{\partial u}{\partial x_1} = \frac{\partial}{\partial x_3} \left(K_v(x_3) \frac{\partial u}{\partial x_3} \right) + w(x_1, x_3) \quad (74)$$

$$u(0, x_3) = u_0(x_3) \quad (75)$$

$$-K_v(0) \frac{\partial u}{\partial x_3} = \phi \delta(x_1), \quad x_3 = 0 \quad (76)$$

$$\frac{\partial u}{\partial x_3} = 0, \quad x_3 = h \quad (77)$$

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

TABLE II
PARAMETER VALUES USED IN LINE SOURCE ESTIMATION EXAMPLE

Truncation number $N = 5$

Measurement time $t_{k+1} = t_k + 0.001$, $k = 1, 2, 3, \dots$
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

$$E[u_0(x)] = u_0(x) = \sum_{i=1}^N u_i^0 \phi_i(x)$$

$$\text{Cov}[u_0(x), u_0(y)] = P_0(x, y) = \sum_{i=1}^N p_{ii}^0 \phi_i(x) \phi_i(y)$$

$$\text{Cov}[w(t, x), w(s, y)] = Q(t, x, y) \delta(t-s), \quad \text{Cov}[v(t_k), v(t_n)] = R(t_k) \delta_{kn}$$

$$Q(t, x, y) = \sum_{i=1}^N q_{ii} \phi_i(x) \phi_i(y), \quad R(t_k) = \text{diag}[r_1, r_2, r_3, r_4]$$

$$\phi_i(x) = \begin{cases} 1 & i = 1 \\ \sqrt{2} \cos(i-1)\pi x & i \geq 2 \end{cases}$$

$$\lambda_i = - (i-1)^2 \pi^2 \quad i \geq 1$$

i	1	2	3	4	5
u_i^0	3.0	1.0	0.03	0.003	0.0003
p_{ii}^0	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	

$$\frac{\partial u}{\partial t} = \frac{\partial^2 u}{\partial x^2} + w(t, x) \quad (78)$$

$$u(0, x) = u_0(x) \quad (79)$$

$$\frac{\partial u}{\partial x} = -\phi \delta(t), \quad x = 0 \quad (80)$$

$$\frac{\partial u}{\partial x} = 0, \quad x = h. \quad (81)$$

In this case the measurements $Z(t_k)$ are related to the concentration $u(t_k, x)$ by (8)

$$Z(t_k) = \int_0^h J^n(t_k, x) u(t_k, x) dx + v(t_k) \quad (82)$$

where the instrument kernel function will be taken to have the form

$$J^n(t_k, x) = \begin{cases} 1, & x \leq h_n, \\ 0, & x > h_n, \end{cases} \quad n = 1, 2, \dots, N. \quad (83)$$

The theory developed in the prior sections can be applied directly to this problem, and the optimal filter and smoother are given in Table I. 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

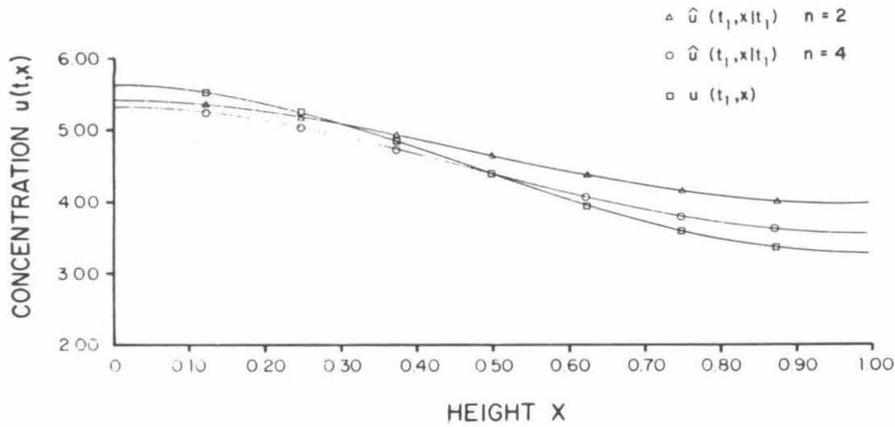


Fig. 2. Comparison of true concentration $\hat{u}(t_1, x)$ and the filter estimates $u(t_1, x/t_1)$ based on 2 and 4 measurement elevations. $t_1 = 0.0002$.

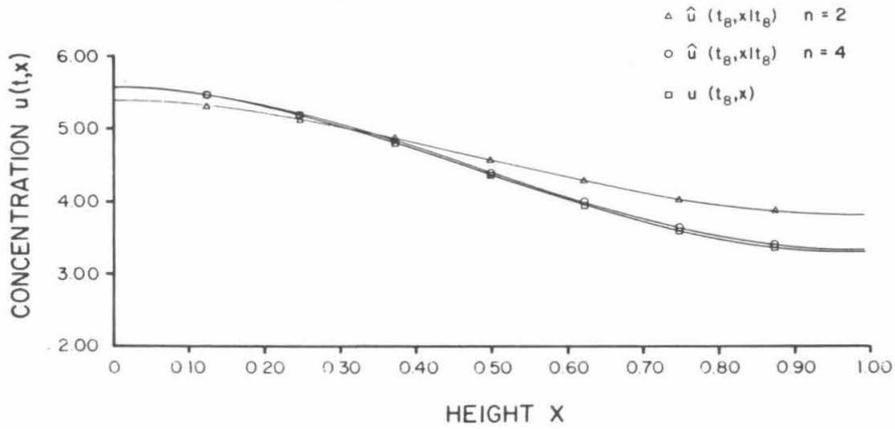


Fig. 3. Comparison of true concentration $u(t_8, x)$ and the filter estimates $\hat{u}(t_8, x/t_8)$ based on 2 and 4 measurement elevations. $t_8 = 0.0082$.

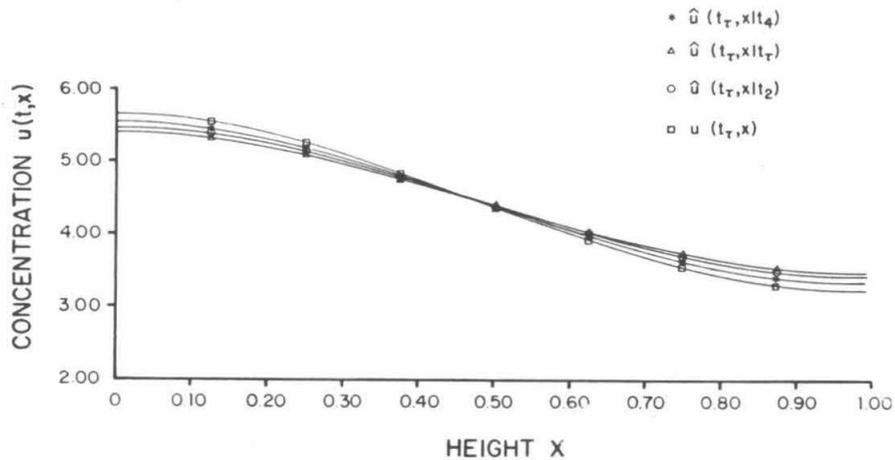


Fig. 4. Comparison of true concentration $u(t_T, x)$, the filter estimate $\hat{u}(t_T, x/t_T)$, and the fixed point smoothing estimates $\hat{u}(t_T, x/t_2)$ and $\hat{u}(t_T, x/t_4)$. $t_T = 0.0002$, $t_2 = 0.0012$, $t_4 = 0.0032$.

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, \dots 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 num-

TABLE III
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

ber of elevations at which data are simultaneously taken. Values of all parameters used in the calculation are given in Table II.

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 III 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 Rep. 159107, Langley Research Center, VA, Aug. 1979.

- [2] W. A. Lyons, J. C. Dooley, Jr., and K. T. Whitby, "Satellite detection of long-range pollution transport and sulfate aerosol hazes," *Atmos. Environ.*, 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 Rep. CR-145097, Langley Research Center, VA, Jan. 1976.
- [4] J. F. Kibler and J. T. Suttles, "Air pollution model parameter estimation using simulated LIDAR data," *AIAA J.*, vol. 15, pp. 1381-1384, 1977.
- [5] J. H. Seinfeld, *Air Pollution: Physical and Chemical Fundamentals*. New York: McGraw-Hill, 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, Springer-Verlag, 1978.
- [7] K. E. Bencala and J. H. Seinfeld, "Distributed parameter filtering: Boundary noise and discrete observations," *Int. J. Syst. Sci.*, vol. 10, no. 5, pp. 493-512, 1979.
- [8] Y. Sakawa, "Observability and related problems for partial differential equations of parabolic type," *SIAM J. Contr.*, vol. 13, no. 1, pp. 14-27, 1975.
- [9] M. Koda and J. H. Seinfeld, "Reconstruction of atmospheric pollutant concentrations from remote sensing data—An application of distributed parameter observer theory," *IEEE Trans. Automat. Contr.*, vol. AC-27, pp. 74-80, Feb. 1982.
- [10] A. J. Alkezweeny, N. E. Davis, and R. C. Easter, "Comparison of ozone in polluted and clean air masses over Lake Michigan," IAMAP Symp. Atmospheric Ozone, Boulder, CO, Aug. 4-9, 1980.
- [11] S. Sandroni, P. Bacci, and D. Anfossi, "Aircraft observations of plumes emitted from elevated sources," *Atmos. Environ.*, vol. 15, no. 1, pp. 95-100, 1981.
- [12] C. J. Nappo, Jr., "Atmospheric turbulence and diffusion estimates derived from observations of a smoke plume," *Atmos. Environ.*, vol. 15, no. 4, pp. 541-548, 1981.
- [13] M. S. Shumate, R. T. Menzies, W. B. Grant, and D. S. McDougal, "The laser absorption spectrometer: Remote measurement of tropospheric ozone," *Appl. Opt.*, vol. 20, pp. 545-553, 1981.

*



Sigeru Omatu (M'76) received the Ph.D. degree in electronics engineering from the University of Osaka Prefecture, Osaka, Japan, in 1974.

During 1980-1981 he was Visiting Associate at the California Institute of Technology. He is currently Professor in the Department of Information Science and Systems Engineering, University of Tokushima, Tokushima 770, Japan. He is co-author of *Modeling, Estimation, and their Applications for Distributed Parameter Systems*. His research interests are in distributed parameter estimation theory and remote sensing.



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

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

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 1980 NASA Public Service Medal.