I. LINEAR DISTRIBUTED PARAMETER FILTERING: OBSERVATION PROCESSES AND BOUNDARY CONDITIONS FOR ENGINEERING SYSTEMS

II. STATISTICAL ANALYSIS OF AIR POLLUTANT OBSERVATIONS AND MODEL PREDICTIONS

Thesis by

Kenneth Edward Bencala

In Partial Fulfillment of the Requirements for the Degree of

Doctor of Philosophy

California Institute of Technology
Pasadena, California 91125

1979

(Submitted October 23, 1978)

Acknowledgments

I wish to thank my advisor, Professor John H. Seinfeld, for his advice and interest in my thesis work.

The financial support for this research by the National Science Foundation and the Environmental Protection Agency is gratefully acknowledged.

The list of persons at Caltech interested in systems theory and air pollution is far too long to write down here, and I would probably miss somebody. So, without particulars, I will merely thank all of those from whom I have learned while at Caltech.

I also appreciate the assistance of the many members of the Caltech staff that I have dealt with during my stay here. In particular, the unselfish help and friendship of the following people has been very important to me: Edith Huang (Computing Center), Sally Ghan (Placement Office), Kathy Lewis (Chemical Engineering), Katie Foley, Rayma Harrison and Helen Lyons (Engineering Libraries) and, especially, a deeply felt acknowledgment to Lenore Kerner.

For almost four years I shared an office with Tom Peterson and Sudarshan Kumar. I know not what pains and trials this arrangement may have meant for them, but for myself I feel I gained much. I thank them both.

Many other students also contributed to the positive, sensitive and occasionally riotous spirit among those of us in the Chemical and Environmental Engineering programs. Ajit Yoganathan, Gary Whatley, Yanis Yortsos, Freddy Gelbard and Dave Strand have each raised the quality of day-to-day life at Caltech.

The incredibly undaunted faith of my parents, brothers, sisters and in-laws that I would someday graduate has been greatly appreciated.

Before I finish this page, I must play the proud new father and acknowledge the smile of my baby girl, Kristen. She is truly an immense joy.

Finally, my best friend (and wife), Marcia, has been a constant source of support and counsel. I thank her dearly and most of all.

Abstract

Part I

The linear distributed parameter filtering problem is considered for systems of engineering interest. In particular, the boundary conditions and observation processes considered represent those encountered in practice. The intent is to consolidate the theoretical results relevant to engineering applications and to provide a unified structure to facilitate implementation of these results through a complete and consistent set of formal derivations. When necessary, previous results are extended.

The types of observation processes considered are

- continuous time, discrete space
- discrete time, discrete space.

The discrete time case includes both instantaneous and limited time average observations. And the discrete space case includes both pointwise and integral observations. Based on the orthogonal projection lemma, the minimum variance estimate is derived.

To examine the effect of stochastic boundary conditions, second order systems with mixed and Dirichlet boundary conditions are considered.

Part II

Aspects of air quality analysis that must be characterized with statistical methods are considered. The evaluation of air quality models is considered in detail. First, an evaluation framework is developed.

The significance of evaluations of validity, accuracy and efficiency are discussed. Then, the specific aspect of accuracy assessment is addressed. The emphasis is placed on practical and objective methods. An extensive package of specific methods is set forth for use with air quality models. The package has been coded in FORTRAN. A description of the code is included.

In an attempt to increase the understanding of characterizations of long-term data, observed frequency distributions of air pollutant concentration levels are critically analyzed with respect to their statistical description. It is demonstrated that several common distributions can be used to fit observed data, one of which is the popular log-normal distribution. The observation that concentration distributions for all averaging times are approximately log-normal can be explained if the short averaging time data are themselves assumed to be log-normally distributed. The near log-normality of pollutant concentration frequency distributions can be explained on the basis of the near log-normality of wind speed distributions, although this explanation does not establish that wind speed distributions are solely responsible for observed concentration distributions. It is concluded that pollutant concentration frequency distributions are the result of complex phenomena and cannot be predicted exactly, but that the approximate log-normal character of the distributions is useful from a practical point of view and can be understood qualitatively on the basis of the relation between wind speed and concentration.

Table of Contents

Table of Contents (Continued)

Table of Contents (Continued)

PART I. LINEAR DISTRIBUTED PARAMETER FILTERING: OBSERVATION PROCESSES AND BOUNDARY CONDITIONS FOR ENGINEERING SYSTEMS

Chapter 1

Introduction

Research concerned with the linear distributed parameter filtering problem has matured to the stage at which a unifying presentation of the problems, the derivations and the results is beneficial. As our orientation is to eventual applications, we emphasize completely specified results for the problems we consider. We define the linear distributed parameter filtering problem in a consistent manner for a variety of observation processes and boundary conditions. All of the filter derivations are based on the orthogonal projection lemma. In this chapter we introduce the filtering problem and specific problems and aspects to be studied.

1.1 Introduction

In this work we develop linear distributed parameter filters for systems with boundary conditions and observation processes of the types encountered in the analysis of engineering systems. This first section consists of a general introduction to the filtering problem, with particular reference to alternative types of boundary conditions and observation processes. The purpose of such a general introduction is to provide a basis for the summary of previous theoretical work presented in Section 1.2. The particular systems and observation processes we consider are specified in the final section of the chapter.

The linear distributed parameter filtering problem deals with estimating the state of a stochastic system governed by an equation of the type

$$\frac{\partial u}{\partial t} = L_x u + Gw$$

where u is the state vector of the system, w a stochastic input, L_x a linear partial differential operator, and G a matrix. (The system will be more carefully defined later.) We desire to estimate the state based on observations of the system. Due to the fundamentally random nature of physical processes, these observations are corrupted by random errors. The observations are expressed in the form

$$z = Hu + v$$

where z is the observation vector, v a random observation error, and H a bounded linear operator.

There are a number of possible variations to this basic problem. We are concerned, in particular, with the different types of boundary conditions and observation processes. In the derivation of the distributed parameter filters, it is frequently assumed that the boundary conditions are deterministic, i.e.,

$$\beta_x u = h_b$$

where β_x is an appropriate linear operator and h_b is a known deterministic function. A more general form of the boundary conditions includes both deterministic and stochastic inhomogeneities,

$$\beta_x u = h_b + w_b$$

where w_b is a stochastic function. Although there are many systems for which deterministic boundary conditions apply, for the majority of real systems stochastic boundary inputs will be present.

Relative to the observation processes, the most frequently considered process in theoretical developments is, in fact, physically impossible to implement. In this case it is assumed that observations are taken continuously in time and at every point in the spatial domain In practice, observations are generally made at discrete locations, with the sampling being point-wise or integral. For certain state variables it is possible to make observations continuously in time, but frequently it is preferable, or necessary, to sample discretely in time on either an instantaneous or an average basis.

1.2 Distributed Parameter Filters: Summary

The last decade has been a period of active research on the distributed parameter filtering problem. Through both rigorous and formal approaches a reasonably complete theory and set of algorithms have been developed. In the following tables this progress is summarized. Our focus here is on the formal development of filters useful for linear engineering systems. For completeness, references to rigorous theories and to filters for nonlinear systems are provided in Tables 1.6 and 1.7 respectively.

Some background references to the study of distributed parameter systems are given in Table 1.1 A summary of the references to formal approaches for the linear filtering problem is given in Table 1.2. In Tables 1.3 — 1.5 these references are further discussed.

Table 1.1

Distributed Parameter Systems -

Background References

Reference	Comments
Athans 1970	Discussion of tentative approaches to unsolved problems in the practical implementation of the theories of distributed parameter systems.
Robinson 1971	Survey of control--theory and applications.
Curtain 1975a	Review of filtering theory. Focuses on the rigorous theories.
Polis and Goodson 1976	Survey of parameter identification--theory and applications.
Ray 1977a	Survey of filtering--theory and applications.
Ray 1977b	Survey of control, filtering and identification--applications.
Tzafestas 1978	Extensive review of all previous results in distributed parameter filtering.

Table 1.2

Formal Linear Distributed Parameter Filters
Classified by the Type of Observation Process Considered

		OBSERVATIONS IN TIME	
		Continuous	Discrete-Instantaneous
OBSERVATIONS IN SPACE	Continuous	(Table 1.3) Tzafestas and Nightingale 1968a, b Tzafestas 1969, 1972a, b, 1973 Meditch 1971 Atre 1972 Atre and Lamba 1972a, b, c* Kumar and Sage 1972 Shukla and Srinath 1972 Padmanabhan and Colantuoni 1974 Reddy and Rajamani 1975a, b Lee 1976	Padmanabhan and Colantuoni 1974
	Discrete	(Table 1.4) Thau 1969 Meditch 1970**, 1971 Sakawa 1972 Shukla and Srinath 1972 Tzafestas 1972a Padmanabhan and Colantuoni 1974	(Table 1.5) Thau 1969 Padmanabhan and Colantuoni 1974 Aidarous and Ghonaimy 1976

*Observations continuous on the boundary.
**A single integral observation.

Linear Distributed Parameter Filters--

Continuous Time and Continuous Space Observations

Reference	Boundary Conditions	Approach
Tzafestas and Nightingale 1968a	Deterministic	Minimize the error covariance via the orthogonal projection lemma.
Tzafestas and Nightingale 1968b	Deterministic	Find the conditional expectation of the state via characteristic functionals.
Tzafestas 1969	Stochastic	Introduce the boundary kernel matrix to deal with the boundary conditions.
Meditch 1971	Deterministic	Solve a deterministic optimal control problem via the sweep method.
Atre 1972	Stochastic	Introduce extended operators to deal with the boundary conditions.
Atre and Lamba 1972a	Deterministic	Minimize the error covariance via the maximum principle.
Atre and Lamba 1972b	Deterministic	Minimize the error covariance via the innovations approach.
Atre and Lamba 1972c	Stochastic	Minimize the error covariance via the innovations approach. Introduce the boundary Green's kernel to deal with the boundary conditions. (Observations continuous on the boundary.)
Kumar and Sage 1972	Stochastic	Minimize the error covariance via the innovations approach. The boundary conditions are given as part of the problem specification.
Shukla and Srinath 1972	Deterministic	Minimize the error covariance via the orthogonal projection lemma.

Table 1.3 (continued)

Reference	Boundary Conditions	Approach
Tzafestas 1972a	Deterministic	Maximize the marginal density functional
Tzafestas 1972b	Deterministic	Extend previous results to include colored observation noise.
Tzafestas 1973	Deterministic	Minimize the error covariance via an extended Gauss-Markov Theorem and the innovations approach.
Padmanabhan and Colantuoni 1974	Stochastic	Take the limit of the filter for discrete time observations. Introduce the Green's function to deal with the boundary conditions.
Reddy and Rajamani 1975a	Deterministic	Extend previous results to the case of hyperbolic systems with colored observation noise.
Reddy and Rajamani 1975b	Deterministic	Extend previous results to include time delays.
Lee 1976	Deterministic	Minimize the error covariance via the Martingale representation.

Table 1.4

Linear Distributed Parameter Filters--

Continuous Time and Discrete Space Observations

Reference	Boundary Conditions	Approach
Thau 1969	Deterministic	Minimize the error covariance via the calculus of variations.
Meditch 1970	Deterministic	Solve a deterministic optimal control problem via the sweep method. (A single integral observation.)
Meditch 1971	Deterministic	Start with the filter for continuous space observations and specify an assumed weighting of the observations.
Sakawa 1972	Stochastic	Minimize the error covariance via the calculus of variations. Introduce a generalized function to deal with the boundary conditions.
Shukla and Srinath 1972	Deterministic	Start with the filter for continuous space observations and specify an assumed form for the discrete observations.
Tzafestas 1972a	Deterministic	Start with the filter for continuous space observations and specify an assumed form for the observation noise covariance.
Padmanabhan and Colantuoni 1974	Stochastic	Take the limit of the filter for discrete time observations. Introduce the Green's function to deal with the boundary conditions.

Table 1.5

Linear Distributed Parameter Filters--

Discrete-Instantaneous Time and Discrete Space Observations

Reference	Boundary Conditions	Approach
Thau 1969	Deterministic	Minimize the error covariance via direct differentiation.
Padmanabhan and Colantuoni 1974	Stochastic	Find the conditional mean. Introduce the Green's function to deal with the boundary conditions.
Aidarous and Ghonaimy 1976	Deterministic	Find the optimal gain in terms of orthonormal basis functions.

Table 1.6

Linear Distributed Parameter Filters - Rigorous Approaches

Reference	System	Boundary Conditions	Observations
Balakrishnan and Lions, 1967	$\frac{du}{dt}(t) + Au(t) = 0$ A: an infinitesimal generator of a strongly continuous semigroup.		$z(t) = M[u(t)] + v(t)$
Falb, 1967	$du = A(t)udt + G(t)dw$ A: bounded operator.		$dz = M(t)udt + G_v(t)dv$
Kushner, 1970	I. $du(x, t) = \left[L_x u + \int_D L(x', x, t) \cdot u(x', t)dD_{x'}\right]dt + G(x, t)dw$	$u(x, t)$ and $L_x u(x, t) \to 0$ as $x \to \partial D$	$y(t) = \int_0^t \left[\int_D M(x, t')u(x, t')dD_x\right]dt' + \int_0^t G_v(t)dv$
	II. $du(x, t) = [L_x u - h(x, t)]dt + G(x, t)dw$	$\frac{\partial u}{\partial \nu} + B(x, t)u = h_b(x, t) + G_b(x, t)a_b(t)$ $\frac{\partial}{\partial \nu}$: conormal derivative	$dy = \left[\int_{\partial D} M(x, t)u(x, t)d\partial D_x\right]dt + dv$
	L_x: 2nd order linear spatial operator.	$da_b = B_a(t)a_b dt + G_{ab}(t)dw_a$	

Table 1.6 (continued)

Reference	System	Boundary Conditions	Observations
Bensoussan, 1971	$\frac{du}{dt}(t) + A(t)u(t) = h(t) + G(t)w(t)$ $y(0) = y_0 + w_0$ A: linear operator, allowing for a wide class of distributed parameter systems.		
Phillipson, 1971	I. $\frac{\partial u}{\partial t}(x, t) = L_x u(x, t) + h(x, t)$ II. $\frac{\partial^2 u}{\partial t^2}(x, t) = L_x u(x, t) + h(x, t)$ L_x: 2nd order linear spatial operator.	I. $u(x, t) = a_b(x, t)$ II. $\frac{\partial u}{\partial \nu}(x, t) = a_b(x, t)$ III. $\frac{\partial u}{\partial \nu}(x, t) + B(x, t)u(x, t) = a_b(x, t)$ $\frac{\partial}{\partial \nu}$: conormal derivative No apriori information available about a_b.	I. Continuous time/continuous space $z(x, t) = u(x, t) + v(x, t)$ II. Continuous time/discrete space $z(x_i, t) = u(x_i, t) + v(x_i, t)$ III. No state observation $z(x, t) = 0$ Also inexact observations of $u(x, o)$ and a_b are taken.
Balakrishnan, 1974a	$\frac{du}{dt}(t, \omega) = Ax(t; \omega) + Gw(t; \omega)$ A: infinitesimal generator of a strongly continuous semigroup.		$z(t; \omega) = Mu(t; \omega) + G_v w(t; \omega)$ M: linear bounded transformation

Table 1.6 (continued)

Reference	System	Boundary Conditions	Observations
Balakrishnan, 1974b	$\frac{\partial u}{\partial t} = \nabla^2 u$ ∇^2: Laplacian	$u(x, t) = w_b(t)$	$z(t) = M_x u(x, t) + v(t)$ M_x: Bounded linear transformation mapping $L_2(D)$ into another Hilbert Space.
Curtain, 1975b	$du(t, \omega) = A(t)u(t, \omega)dt + G(t)dw(t, \omega)$ $A(t)$: unbounded operator which generates an evolution operator.		$dz(t, \omega) = M(t)u(t, \omega)dt + G_b(t)dv(t, \omega)$
Koda, 1975	$\frac{du}{dt} = Au(t) + G(t)\omega(t)$ A: infinitesimal generator of a semigroup.		$z(t_k) = M(t_k)u(t_k) + v(t_k)$ M: Continuous mapping of a Hilbert Space into the measurement space.
Kumar and Seinfeld, 1978	$\frac{du}{dt} = Au(t) + G(t)\omega(t)$ A: infinitesimal generator of a strongly continuous semigroup.		$z(t) = M(t)u(t) + v(t)$ M: Bound linear operator. Specified to yield spatially-independent integral observations and discrete point observations.

Table 1.7

Nonlinear Distributed Parameter Filters

Reference	System	Boundary Conditions	Observations	Approach	Comments
Tzafestas and Nightingale, 1969	$\frac{\partial u(x,\, t)}{\partial t} = N_x(u,\, x,\, t) + G(x,\, t)w(x,\, t)$ N_x: nonlinear spatial operator	Deterministic $B_{N_x}(u,\, x,\, t) = 0$ B_{N_x}: nonlinear spatial operator.	Continuous time/Continuous space $z(x,\, t) = M_x u(x,\, t) + v(x,\, t) \quad x \epsilon D$	Maximize the likelihood functional via dynamic programming.	Includes smoothing solution.
Seinfeld, Gavalas and Hwang, 1971	$\frac{\partial u(x,\, t)}{\partial t} = N\left(t,\, x,\, u,\, \frac{\partial u}{\partial x},\, \frac{\partial^2 u}{\partial x^2}\right) + w(x,\, t)$	General with stochastic inputs. $B\left(t,\, a_b,\, u,\, \frac{\partial u}{\partial x}\right) = 0$ $\frac{da_b}{dt} = B_a(t,\, a_b) + w_{ab}(t)$	Continuous time/Discrete space $z(t) = M(t,\, u_M) + v(t)$ $u_M(t) = \left[u^T(x,\, t), \ldots,\, u^T(x_p,\, t)\right]^T$ $x_i \epsilon [0,\, 1]$	Discretize the system spatially, apply the lumped parameter filter and take the limit back to the continuous case.	The parameter a_b is also estimated.
Hwang, Seinfeld and Gavalas, 1972	$\frac{\partial u(x,\, t)}{\partial t} = N\left(t,\, x,\, u,\, \frac{\partial u}{\partial x},\, \frac{\partial^2 u}{\partial x^2},\, a\right) + w(x,\, t)$ $\frac{da}{dt} = N_a(t,\, a) + w_a(t)$	General with stochastic inputs. $B\left(t,\, a_b,\, u,\, \frac{\partial u}{\partial x}\right) = w_b(t)$ $\frac{da_b}{dt} = B_a(t,\, a_b) + w_{ab}(t)$	Continuous time/Continuous space $z(x,\, t) = M(t,\, x,\, u) + v(x,\, t)$ $x \epsilon [0,\, 1]$ Discrete space also treated.	Solve a deterministic control problem via invariant imbedding. Discrete space observations treated by the ad hoc specification of the observation weighting matrix.	The parameters a and a_b are also estimated. Includes smoothing solution.
Lamont and Kumar, 1972	$\frac{\partial u(x\, t)}{\partial t} = N_x(u,\, x,\, t) + w(x,\, t)$ N_x: nonlinear second order spatial operator.	Stochastic $B_{N_x}(u,\, x,\, t) = w_b(x,\, t)$ B_{N_x}: nonlinear spatial operator.	Continuous time/Continuous space $z(x,\, t) = M(u,\, x,\, t) + v(x,\, t)$ $x \epsilon \bar{D}$	Solve a deterministic control problem via invariant imbedding.	

Table 1.7 (continued)

Reference	System	Boundary Condition	Observations	Approach	Comments
Sahgal and Webb, 1972	$\frac{\partial u(x, t)}{\partial t} = N_x(u, x, t) + G(x, t)w(x, t)$ N_x: second order spatial operator.	Homogeneous $B_{N_x}(u, x, t) = 0$ B_{N_x}: first order spatial operator.	Continuous time/Continuous space $z(x, t) = M(u, x, t) + v(x, t) \quad x \in D$	Solve a deterministic control problem via invariant imbedding.	Includes smoothing solution.
Tzafestas, 1972a	$\frac{du}{dt}(x, t) = N_x(u, x, t) + G(x, t)w(x, t)$	Deterministic $B_{N_x}(u, x, t) = 0$	Continuous time/Continuous space $z(x, t) = M_x(u, x, t) + v(x, t) \quad x \in D$	Linearize the system around the filter estimate and apply the linear filter.	
Sherry and Shen, 1973	$\frac{\partial u}{\partial t}(x, t) = N_x(u, x, t) + w(x, t)$ N_x: nonlinear spatial operator.	Deterministic $B_{N_x}(u, x, t) = 0$ B_{N_x}: nonlinear spatial operator.	Discrete time/Discrete space $z(x_1, t_j) = M(x_1, t_j)u(x_1, t_j) + v(x_1, t_j) \quad x_1 \in D$	Solve a deterministic control problem via the sweep method.	
Yu, Seinfeld and Ray, 1974	$\frac{\partial u}{\partial t}(x, t) = -N_1(x, t)\frac{\partial u}{\partial x} + N_2(u, x, t) + w(x, t)$ $\frac{\partial a}{\partial t} = N_{a_1}(a, u(x, t), \ldots, u(x_n, t), t) + \int_0^1 N_{a_2}(u, x', t)dx' + w_a(t)$	$u(0, t) = B(x(t))$	Continuous time/Discrete and integral space $z(t) = M_1(a(t), u(x_1, t), \ldots, u(x_p, t)) + \int_0^1 M_2(u(x', t), x', t)dx' + v(t) \quad x_1 \in (0, 1]$	Solve a deterministic control problem employing differential sensitivities.	Includes smoothing solution. The parameter a is also estimated.

Table 1.7 (continued)

Reference	System	Boundary Conditions	Observations	Approach	Comments
Ajinkya, Ray, Yu, and Seinfeld, 1975	$\frac{\partial u}{\partial t}(x, t) = N\left(u, \frac{\partial u}{\partial x}, \frac{\partial^2 u}{\partial x^2}, a, x, t\right) + w(x, t)$ $\frac{\partial a}{\partial t} = N_a\left(a, u(x_1, t), \ldots, u(x_n, t), t\right) + w_a(x, t)$	General with stochastic inputs. $B\left(u, \frac{\partial u}{\partial x}, a, t\right) = w_b(t)$	Continuous time/Discrete space $z(t) = M\left(a(t), u(x_1, t), \ldots, u(x_p, t), t\right) + v(t)$ Discrete time also treated.	Solve a deterministic control problem employing differential sensitivities. Discrete time observations treated by the ad hoc specification of the observation weighting matrix.	The parameter a is also estimated.
Ray and Seinfeld, 1975	$\frac{\partial u_j}{\partial t}(x, t) = N_j\left(u_j, \frac{\partial u_j}{\partial x}, \frac{\partial^2 u_j}{\partial x^2}, x, t\right) + w_j(x, t)$ $j = 1, 2, \ldots, r+1, \; a_{j-1} < x < a_j$	Stochastic $B_0\left(u_1, \frac{\partial u_1}{\partial x}, t\right) + w_{b0} = 0$ $x = a_0 = 0$ $B_{r+1}\left(u_{r+1}, \frac{\partial u_{r+1}}{\partial x}, t\right) + w_{b,\,r+1} = 0$ $x = a_{r+1} = 1$	Continuous time/Discrete space $z(t) = M\left(a, u(x_1, t), \ldots, u(x_p, t)\right) + v(t) \quad x_1 \in [0, 1]$ a and u denote the vectors comprising the moving boundaries and state variables, respectively.	Solve a deterministic control problem employing differential sensitivities.	An extension of the preceding two papers to case of moving boundaries.

Table 1.7 (continued)

Reference	System	Boundary Conditions	Observations	Approach	Comments
Ray and Seinfeld, 1975 (con't)	The locations, a_j, of the r moving boundaries are determined by $\frac{da_j}{dt}(t) = N_a\left(u_j(a_j, t), u_{j+1}(a_j, t), \frac{\partial u_j}{\partial x}(a_j, t), \frac{\partial u_{j+1}}{\partial x}(a_j, t), a_j, t\right) + w_{aj}(t) \quad j = 1, \ldots, r$	$B_j\left(u_j, u_{j+1}, \frac{\partial u_j}{\partial x}, \frac{\partial u_{j+1}}{\partial x}, a_j, t\right) + w_{bj} = 0$ $B'_j\left(u_j, u_{j+1}, \frac{\partial u_j}{\partial x}, \frac{\partial u_{j+1}}{\partial x}, a_j, t\right) + w'_{bj} = 0 \quad j = 1, \ldots, r$			
Tzafestas, 1976	$du(x, t) = N_x(u, a, x, t)dt + G_x(u, a, x, t)dw(x, t)$ $\frac{da}{dt} = N_a(a, t)dt + G_a(a, t)dw_a(t)$	General with stochastic inputs. $B_{N_x}(a_b, x, t) = w_b(t)$ $\frac{da_b}{dt} = B_a(a_b, t)dt + G_{ab}(a_b, t)dw_{ab}(t)$	Continuous time/Continuous space $dz(x, t) = M(u, a, a_b, z, x, t)dt + G(x, a, t)dv(x, t) \quad x \in D$ Discrete space is also treated.	Find the conditional expection of the state via the Fokker-Planck-Kolmogorov equation. Discrete space observations treated by the ad hoc specification of the observation weighting matrix.	The parameters a and a_b are also estimated.

Table 1.7 (continued)

Reference	System	Boundary Conditions	Observations	Approach	Comments
Tzafestas and Nightingale, 1976	$\frac{du}{dt}(x, t) = N_x(u, x, t)dt + G_x(u, x, t)dw(x, t)$	Deterministic $B_x(u, x, t) = 0$	Continuous time/Discrete space Doubly stochastic counting process (the statistics of the counting process are a function of the state, u).	Find the conditional expectation of the state via the Fokker-Planck-Kolmogorov equation.	

1.3 Specification of the System and Observation Processes

As stated in Section 1.1 we are interested in the linear distributed parameter filtering problem for systems with stochastic boundary conditions and for observation processes that are discrete in space. In this work we present a unified discussion of these particular problems. First, in Chapter 2, we treat the problem of discrete spatial observation processes for linear systems with deterministic boundary conditions. Three types of discrete spatial observation processes are considered. They are:

- continuous time
- discrete-instantaneous time
- discrete-average time.

Then, in Chapter 3, we show how these results can be extended in specific cases to systems with stochastic boundary conditions. We have chosen to consider a specific class of linear systems and two types of boundary conditions. This allows us to completely specify the results. In most previous work, consideration of more general cases has necessitated leaving unspecified functions in the final result.

Tables 1.8 and 1.9 summarize the cases considered in this work and list other references, if any, also containing the filters.

Table 1.8

Linear Distributed Parameter Filters Considered in Chapter 2

System: $\frac{\partial u}{\partial t} = L_x u + G(x, t)w(x, t) + h(x, t) \quad x \epsilon D \quad 0 < t$

Boundary Conditions: Deterministic, $\beta_x u = h_b \quad x \epsilon \partial D$

Observation Process - Discrete Space	References Also Containing Filter
Continuous Time $z(t) = \int_D M(x, t)u(x, t)dD_x + v(t)$	As indicated in Section 1.2, this filter is widely available. (For example, see Sakawa, 1972 or Padmanabhan and Colantuoni, 1974.)
Discrete-Instantaneous Time $z(t_k) = \int_D M(x, t_k)u(x, t_k)dD_x + v(t_k)$	Padmanabhan and Colantuoni,1974.
Discrete-Average Time $z(t_k) = \int_{t_{k-1}}^{t_k} \int_D M(x, t)u(x, t)dD_x dt + v(t_k)$	**None**

Table 1.9

Linear Distributed Parameter Filters Considered in Chapter 3

System: $\frac{\partial u}{\partial t} = \alpha_2(x, t) \frac{\partial^2 u}{\partial x^2} + \alpha_1(x, t) \frac{\partial u}{\partial x} + \alpha_0(x, t)u + G(x, t)w(x, t) + h(x, t)$

$0 < x < 1 \quad 0 < t$

Boundary Condition	References Also Containing Filter
Mixed $\left[\phi_1(t)\alpha_2(1, t) \frac{\partial u}{\partial x} + \theta_1(t)u\right]\Big\|_{x = 1}$ $= h_1(t) + w_1(t)$; ϕ_1 nonsingular	None for the specific system considered. Another second order system has been considered by Sakawa (1972).
Dirichlet $\theta_1(t)u \Big\|_{x = 1} = h_1(t) + w_1(t)$; θ_1 nonsingular	None

Observation Process: Discrete Space; Discrete and Continuous Time

We now specify the system and observation processes considered.

The n-dimensional state vector, u(x, t), is governed by the partial differential equation

$$Lu = \frac{\partial u}{\partial t} - L_x u = G(x, t)\, w(x, t) + h(x, t) \qquad (1.1)$$

defined for $t > 0$, $x \varepsilon D$. The domain, D, is a connected subset of ℓ-dimensional Euclidean space R_ℓ with boundary surface ∂D. L_x is a linear, partial differential operator with respect to the spatial coordinate x. w(x, t) is an n-dimensional stochastic input, white in time, h is a known n-dimensional input, and G(x, t) is a known n x n matrix.

The initial condition for (1.1) is

$$u(x, 0) = u_0(x) \qquad (1.2)$$

and the boundary condition* is

$$\beta_x u = h_b(x, t) + w_b(x, t) \qquad x \varepsilon \partial D \qquad (1.3)$$

where β_x is a linear operator. When the stochastic terms are absent, the problem is assumed to be well-posed.

*As indicated in Table 1.9, only for a specific class of these systems is the full stochastic boundary condition problem considered.

The initial condition is unknown, with its first two moments given,

$$E\{u_0(x)\} = \bar{u}_0(x)$$

$$E\left\{\left(u_0(x) - \bar{u}_0(x)\right)\left(u_0(x') - \bar{u}_0(x')\right)^T\right\} = P_0(x, x'). \tag{1.4}$$

The stochastic inputs have the following properties

$$E\{w(x, t)\} = 0$$

$$E\left\{w(x, t)\, w^T(x', t')\right\} = Q(x, x', t)\, \delta(t - t') \tag{1.5}$$

$$E\{w_b(x, t)\} = 0$$

$$E\left\{w_b(x, t)\, w_b^T(x', t')\right\} = Q_b(x, x', t)\, \delta(t - t') \tag{1.6}$$

where $Q(x, x', t)$ and $Q_b(x, x', t)$ are non-negative definite n x n matrix functions. (The matrix $S(x, x', t)$ is said to be non-negative definite, in a generalized sense, if $\int_D \int_D \eta^T(x'')S(x'', x''', t)\eta(x''')dD_{x''}dD_{x'''} \geq 0$ for all $\eta(x)$.)

We are interested in observation processes in which there are p discrete observation locations. $z_i(t)$ denotes the m-dimensional observation vector at the i^{th} observation location. Observations may be made in the domain or on the boundary. We employ the compact notation (Kumar

and Seinfeld, 1978),

$$z(t) = \begin{bmatrix} z_1(t) \\ \vdots \\ z_p(t) \end{bmatrix}$$

to denote the overall pm-dimensional observation vector. The three types of observations* considered are

continuous time

$$z(t) = \int_D M(x', t)\, u(x', t)\, dD_{x'} + v(t) \tag{1.7}$$

discrete-instantaneous time

$$z(t_k) = \int_D M(x', t_k)\, u(x', t_k)\, dD_{x'} + v(t_k) \tag{1.8}$$

discrete-average time**

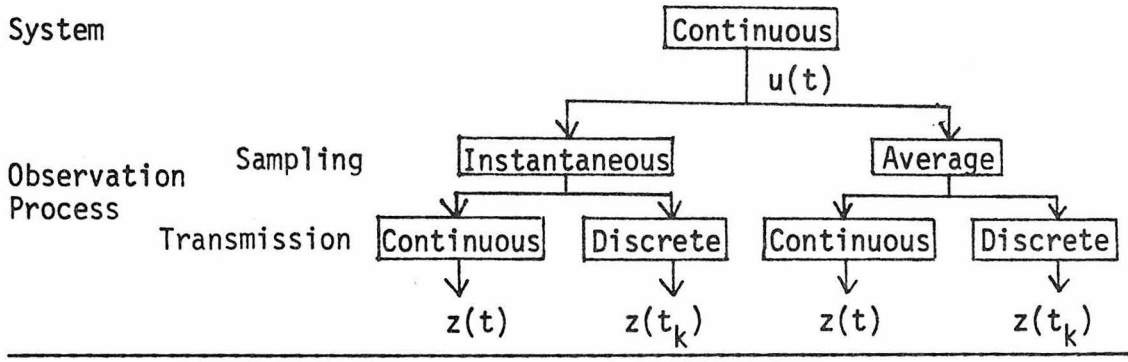
$$z(t_k) = \int_{t_{k-1}}^{t_k} \int_D M(x', t')\, u(x', t')\, dD_{x'} dt' + v(t_k) \tag{1.9}$$

M(x, t) is the known overall pm x n observation matrix, and it represents a matrix consisting of the p observation matrices at the individual locations, i.e.,

$$M(x, t) = \begin{bmatrix} M_1(x, t) \\ \vdots \\ M_p(x, t) \end{bmatrix}.$$

*The following chart explains the temporal observation process nomenclature.

Temporal Observation Process Nomenclature



The recording of observations may be considered as a two-step process; sampling the system and transmitting the observations. For a continuous system, we may sample a system either in an instantaneous or in an average manner and then transmit the result either continuously or discretely. We consider three of these processes. The fourth, continuous-average or moving average as it is commonly referred to, is not considered.

** In the case of discrete-average observations the observation noise includes both sampling and transmission noise. The noise occurring in the sampling process is averaged over the interval t_k to t_{k-1}. Thus, Equation (1.9) may be expressed

$$z(t_k) = \int_{t_{k-1}}^{t_k} \int_D M(x',t')u(x',t')dD_x,dt' + \int_{t_{k-1}}^{t_k} v_S(t')dt' + v_T(t_k) \qquad (1.9a)$$

where

$$v(t_k) = \int_{t_{k-1}}^{t_k} v_S(t')dt' + v_T(t_k),$$

v_S is the sampling noise and v_T is the transmission noise. For v_S and v_T defined as independent white noises with properties

$$E\{v_S(t)v_S^T(t')\} = R_S(t)\delta(t - t')$$

and $$E\{v_T(t_k)v_T^T(t_l)\} = R_T(t)\delta_{kl}$$

Equation (1.12) can be expressed

$$E\{v(t_k)v^T(t_l)\} = R(t_k)\delta_{kl} = \left[\int_{t_{k-1}}^{t_k} R_S(t')dt' + R_T(t_k)\right]\delta_{kl}. \qquad (1.12a)$$

The observation noise, **v**, is a pm-vector, assumed to be zero-mean and white. For continuous time observations the properties of v(t) are

$$E\{v(t)\} = 0$$

$$E\left\{v(t)v^T(t')\right\} = R(t)\delta(t - t') = \begin{bmatrix} R_{1,1}(t)\cdots R_{1,p}(t) \\ \vdots \quad\quad \vdots \\ R_{p,1}(t)\cdots R_{p,p}(t) \end{bmatrix} \delta(t - t') \qquad (1.10)$$

where R is a known positive definite pm x pm matrix and

$$E\left\{v_i(t)v_j^T(t')\right\} = R_{ij}(t)\delta(t - t') \ . \qquad (1.11)$$

Similarly, for discrete time observations the properties of $v(t_k)$ are

$$E\{v(t_k)\} = 0$$

$$E\left\{v(t_k)v^T(t_l)\right\} = R(t_k)\ \delta_{kl} = \begin{bmatrix} R_{1,1}(t_k)\cdots R_{1,p}(t_k) \\ \vdots \quad\quad \vdots \\ R_{p,1}(t_k)\cdots R_{p,p}(t_k) \end{bmatrix} \delta_{kl} \qquad (1.12)$$

where R is a known positive definite pm x pm matrix and

$$E\left\{v_i(t_k)v_j^T(t_l)\right\} = R_{ij}(t_k)\delta_{kl}. \qquad (1.13)$$

Finally, u_0, w, w_b and v are all defined to be independent of each other.

The observation processes have been expressed as integrals over the

domain D. In applications one usually has point-wise observations. By taking the limit to small volumes of integration we can represent such a process. The observation matrices become

$$M_i(x, t) = M_i(t)\delta(x - x_i) \qquad i = 1, \ldots, p$$

and the observations are

$$z_i(t) = M_i(t)u(x_i, t) + v_i(t) \qquad i = 1, \ldots, p.$$

Chapter 2

Distributed Parameter Filters: Deterministic Boundary Conditions

In this chapter we present the linear distributed parameter filters for the three types of discrete space observation processes defined in Chapter 1. We specify the filtering problem, summarize the results, and derive the filters.

2.1 Introduction

The system and observation processes we consider are as given in Section 1.3 with deterministic boundary conditions. Stochastic boundary conditions are considered in Chapter 3. In Section 2.2 we introduce the necessary notation and define the estimate we seek. The filters for the three observation processes are presented in Section 2.3 and derived in Section 2.4. The derivations follow the formal procedure of Tzafestas and Nightingale (1968a) and are based on the orthogonal projection lemma (see Appendix A).

There are three differences in the exact specification of the problems considered in the derivations and in the summary, Table 2.1. These differences arise because of differences in the purposes of the derivations and the summary. In the derivations we wish to present the essentials of the filtering problem uncluttered by slight extensions. However, in the summary we wish to present the filters in forms more useful for applications. The three differences are stated in the following table.

	Derivations	Summary
Spatial Observations	Integral	Point-Wise
Deterministic Inputs	$h = 0$	h
	$h_b = 0$	h_b
Mean of Initial Condition	$\bar{u}_0 = 0$	$\bar{u}_0$

2.2 Specification of the Filters: Notation and Criterion

We begin by defining the following terms:

$\hat{u}(x, t\|t')$		Optimal estimate of $u(x, t)$ based on observations through time t'.
$\hat{u}(x, t)$	$= \hat{u}(x, t\|t)$	
$\tilde{u}(x, t\|t')$	$= u(x, t) - \hat{u}(x, t\|t')$	Error in estimating $u(x, t)$ by the optimal estimate.
$P(x, t\|t'; x', t''\|t''')$	$= E\{\tilde{u}(x, t\|t')\tilde{u}^T(x', t''\|t''')\}$	
$P(x, x', t\|t')$	$= P(x, t\|t'; x', t\|t')$	
$P(x, x', t)$	$= P(x, t\|t; x', t\|t)$	Error covariance matrix.
$\omega(x, t\|t')$		Any admissible estimate of $u(x, t)$.
$\eta(x)$		Arbitrary vector.

where $\hat{u}$, $\tilde{u}$, ω, and η are n-vectors and P is an n x n matrix.

To facilitate working with the optimality criterion, we introduce the following inner product and norm,

$$Y = \langle \eta(x),\, y(x) \rangle = \int_D \eta^T(x')y(x')dD_{x'}$$

$$(Y_1,\, Y_2) = E\{\langle \eta(x),\, y_1(x) \rangle \langle \eta(x),\, y_2(x) \rangle\}$$

$$= \int_D \int_D \eta^T(x')E\{y_1(x')y_2^T(x")\}\eta(x")dD_{x'}dD_{x"}$$

$$\| Y \|^2 = (Y,\, Y) = \int_D \int_D \eta^T(x')E\{y(x')y^T(x")\}\eta(x")dD_{x'}dD_{x"}$$

where y is an n-vector.

We seek an estimate at time t which is unbiased and the optimal linear combination of the observations through time t. The two forms of the estimate for the case $\bar{u}_0 = 0$, $h = 0$, $h_b = 0$ are

$$\omega(x,\, t|t) = \int_0^t B(x,\, t,\, t')z(t')dt' \tag{2.1}$$

for continuous time observations and

$$\omega(x,\, t_k|t_k) = \sum_{j=1}^{k} B(x,\, t_k,\, t_j)z(t_j) \tag{2.2}$$

for discrete time observations. Thus, to find the optimal ω we must find the optimal B. The optimal $B(x, t, t')$ is denoted $A(x, t, t')$, i.e. $\hat{u}(x, t) = \int_0^t A(x, t, t')z(t')dt'$.

The criterion for optimality is that the error covariance matrix be minimized. That is, we seek to minimize the scalar function

$$E\left\{\left[\int_D \eta^T(x')\big(u(x', t) - \omega(x', t|t)\big)\, dD_{x'}\right]^2\right\}$$

for all η. Or, as expressed in the inner product notation; we seek $\hat{U}$, such that

$$\|\tilde{U}\| \leq \|U - \Omega\| \tag{2.3}$$

for all Ω.

2.3 Summary of Results

The filters derived in this chapter are summarized in Table 2.1. For all three of the observation processes, the initial and boundary conditions are

$$\hat{u}(x, 0) = \bar{u}_0(x) \tag{2.4}$$

$$P(x, x', 0) = P_0(x, x')$$

and

$$\beta_x \hat{u}(x) = h_b(x, t) \qquad x \in \partial D \tag{2.5}$$

$$\beta_x P(x, x', t) = 0.$$

Table 2.1

Linear Filters Derived in Chapter 2 - Deterministic Boundary Conditions and Point-Wise Observations

Observation Process	Filter
Continuous Time	$\frac{\partial \hat{u}}{\partial t}(x, t) = L_x \hat{u}(x, t) + \sum_{j=1}^{p} \sum_{i=1}^{p} P(x, x_j, t) M_j^T(t) R_{ji}^{\ddagger}(t) \left[z_i(t) - M_i(t) \hat{u}(x_i, t) \right]$
	$+ h(x, t)$
	$\frac{\partial P}{\partial t}(x, x', t) = L_x P(x, x', t) + P(x, x', t) L_{x'}^T$
	$- \sum_{j=1}^{p} \sum_{i=1}^{p} P(x, x_j, t) M_j^T(t) R_{ji}^{\ddagger}(t) M_i(t) P(x_i, x', t) + G(x, t) Q(x, x' t)$
	$\cdot G^T(x', t)$
Discrete Time	
Prediction Equations $t_{k-1} \le t \le t_k$	$\frac{\partial \hat{u}}{\partial t}(x, t \mid t_{k-1}) = L_x \hat{u}(x, t \mid t_{k-1}) + h(x, t)$
	$\frac{\partial P}{\partial t}(x, x', t \mid t_{k-1}) = L_x P(x, x', t \mid t_{k-1}) + P(x, x', t \mid t_{k-1}) L_{x'}^T$
	$+ G(x, t) Q(x, x', t) G^T(x', t)$
Up-Date Equations	
$t = t_k$	
Instantaneous	$\hat{u}(x, t_k) = \hat{u}(x, t_k \mid t_{k-1}) + \sum_{j=1}^{p} \sum_{i=1}^{p} P(x, x_j, t_k \mid t_{k-1}) M_j^T(t_k) K_{ji}^{\ddagger}(t_k)$
	$\cdot \left[z_i(t_k) - M_i(t_k) \hat{u}(x_i, t_k \mid t_{k-1}) \right]$

Table 2.1 (continued)

Observation Process — Filter

$$P(x, x', t_k) = P(x, x', t_k|t_{k-1}) - \sum_{j=1}^{p} \sum_{i=1}^{p} P(x, x_j, t_k|t_{k-1}) M_j^T(t_k) K_{ji}^{\ddagger}(t_k)$$

$$\cdot M_i(t_k) P(x_i, x', t_k|t_{k-1})$$

$$K_{ji}(t_k) = M_j(t_k) P(x_j, x_i, t_k|t_{k-1}) M_i^T(t_k) + R_{ji}(t_k)$$

Time-Average

$$\hat{u}(x, t_k) = \hat{u}(x, t_k|t_{k-1}) + \int_{t_{k-1}}^{t_k} \sum_{j=1}^{p} \sum_{i=1}^{p} P(x, t_k|t_{k-1}; x_j, t'|t_{k-1})$$

$$\cdot M_j^T(t') K_{ji}^{\ddagger}(t_k) \left[z_i(t_k) - \int_{t_{k-1}}^{t_k} M_i(t'') \hat{u}(x_i, t''|t_{k-1}) dt'' \right] dt'$$

$$P(x, x', t_k) = P(x, x', t_k|t_{k-1}) - \int_{t_{k-1}}^{t_k} \int_{t_{k-1}}^{t_k} \sum_{j=1}^{p} \sum_{i=1}^{p}$$

$$\cdot P(x, t_k|t_{k-1}; x_j, t'|t_{k-1}) M_j^T(t') K_{ji}^{\ddagger}(t_k) M_i(t'')$$

$$\cdot P(x_i, t''|t_{k-1}; x', t_k|t_{k-1}) dt'' dt'$$

Table 2.1 (continued)

Observation Process	Filter
$K_{ji}(t_k) = \int_{t_{k-1}}^{t_k}\int_{t_{k-1}}^{t_k} M_j(t')$	
$\cdot P(x_j, t'\|t_{k-1}; x_i, t''\|t_{k-1})M_i^T(t'')dt'dt'' + R_{ji}(t_k)$	
	$\frac{\partial P}{\partial t}(x, t\|t_{k-1}; x', t'\|t_{k-1}) = \frac{\partial P^T}{\partial t}(x', t'\|t_{k-1}; x, t\|t_{k-1})$
	$= L_x P(x, t\|t_{k-1}; x', t'\|t_{k-1}) \qquad t_{k-1} \le t' < t \le t_k$

The notation $R^{\ddagger}_{ij}$ and $K^{\ddagger}_{ij}$ is used in this chapter. The relationship between the matrices, N_{ik} and $N^{\ddagger}_{kj}$, is defined as follows

$$\sum_{k=1}^{p} N_{ik}N^{\ddagger}_{kj} = I\delta_{ij} .$$

For the discrete-average observations, a boundary condition is also required for $P(x, t|t_{k-1}; x', t'|t_{k-1})$. It is

$$\beta_x P(x, t|t_{k-1}; x', t'|t_{k-1}) = 0 , \quad x\epsilon\partial D. \tag{2.6}$$

2.4 Derivation of the Filters

The approach taken in this chapter to deriving the filters is based on the orthogonal projection lemma. The orthogonal projection lemma characterizes the optimal estimate, $\hat{U}$, of the state, U, out of the set of admissible estimates Ω. The lemma states

$$||U - \hat{U}|| \leq ||U - \Omega|| \qquad \text{if and only if}$$

$$(U - \hat{U}, \Omega) = 0 \tag{2.7}$$

for all Ω, and furthermore, if another estimate $\hat{U}'$ also satisfies (2.7), then

$$||\hat{U} - \hat{U}'|| = 0 \quad . \tag{2.8}$$

2.4.1 Continuous Time Observations

For the continuous time observation process the optimality condition, Equation (2.7), becomes

$$\int_D \int_D \eta^T(x'') \int_0^t E\left\{\tilde{u}(x'', t)z^T(t')\right\} \cdot B^T(x''', t, t')dt' \, \eta(x''')dD_{x''}dD_{x'''} = 0 \quad , \tag{2.9}$$

for all η and B.

The following theorem leads directly to the criterion for A(x, t, t'), the optimal B(x, t, t').

Theorem 2.1

$\hat{u}(x, t)$ satisfies the optimality condition, Equation (2.9), if and only if

$$E\left\{\tilde{u}(x, t)\, z^T(t')\right\} = 0 \qquad , t > t' \tag{2.10}$$

Proof

Sufficiency: Obvious

Necessity: Let $B(x, t, t') = E\left\{\tilde{u}(x, t)z^T(t')\right\}$. Then (2.9) becomes

$$\int_0^t \int_D \int_D \eta^T(x'')B(x'', t, t')B^T(x''', t, t')\eta(x''')dD_{x''}dD_{x'''}\, dt' = 0 \, .$$

Unless $B \equiv 0$, BB^T will be non-negative definite and thus not satisfy this equation. Thus, (2.10) holds. #

We will find it expedient to write Equation (2.10) as

$$E\left\{u(x, t)z^T(t')\right\} - \int_0^t A(x, t, t'')E\left\{z(t'')z^T(t')\right\}dt'' = 0 \, , \; t > t' \tag{2.11}$$

and

$$E\{\tilde{u}(x, t)\hat{u}^T(x', t')\} = 0 \quad , \quad t \geq t' \tag{2.12}$$

follows directly. Equation (2.10) will be referred to as the Wiener-Hopf equation (for continuous time observations).

In a sense, the filtering problem is solved; we have in Equation (2.11) an implicit equation for A. However, the form of Equation (2.11) is impractical in that it is not recursive. We proceed to find the more practical, recursive filter.

Consider the derivative of the Wiener-Hopf equation with respect to time. Using the form in Equation (2.11) we have

$$E\left\{\frac{\partial u(x, t)}{\partial t} z^T(t')\right\} - A(x, t, t)E\left\{z(t)z^T(t')\right\}$$

$$- \int_0^t \frac{\partial A(x, t, t")}{\partial t} E\left\{z(t")z^T(t')\right\} dt" = 0 \quad , \quad t > t'. \tag{2.13}$$

After rearrangement, use of (2.11), (1.7) and (1.1), and attention to the stochastic properties of v and w, Equation (2.13) becomes

$$\int_0^t \left\{L_x A(x, t, t") - A(x, t, t) \int_D M(x', t)A(x', t, t")dD_{x'} - \frac{\partial A(x, t, t")}{\partial t}\right\}$$

$$\cdot E\{z(t")z^T(t')\} dt" = 0 \quad , t > t' . \tag{2.14}$$

We now show that

$$L_x A(x, t, t') - A(x, t, t) \int_D M(x', t)A(x', t, t')dD_{x'}$$

$$- \frac{\partial A(x, t, t')}{\partial t} = 0 \quad , \quad t \geq t' \; . \tag{2.15}$$

Proof of Equation (2.15)

Let $C(x, t, t')$ denote the left hand side of (2.15). Then if $A(x, t, t')$ satisfies the Wiener-Hopf equation, so does $A + C$. Thus both

$$\hat{u}(x, t) = \int_0^t A(x, t, t')z(t')dt' \qquad \text{and}$$

$$\hat{u}'(x, t) = \int_0^t [A(x, t, t') + C(x, t, t')]z(t')dt'$$

are optimal estimates of $u(x, t)$. By the orthogonal projection lemma (Equation (2.8)) we have

$$\int_0^t\int_0^t\int_D\int_D \eta^T(x")C(x", t, t")E\left\{z(t")z^T(t"')\right\}$$

$$\cdot C^T(x"', t, t"')\eta(x"')dD_{x"}dD_{x"'}dt"dt"' = 0.$$

But $E\left\{z(t")z^T(t"')\right\}$ can be shown to be positive-definite. Thus $C \equiv 0$, and (2.15) holds. #

Equation (2.15) is used to find a differential expression for $\hat{u}(x, t)$. Taking the derivative with respect to time of the defining equation for

$\hat{u}(x, t)$, we have

$$\frac{\partial \hat{u}(x, t)}{\partial t} = A(x, t, t)z(t) + \int_0^t \frac{\partial A(x, t, t')}{\partial t} z(t')dt'. \quad (2.16)$$

Substituting (2.15) into (2.16) yields,

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

$$\cdot \left\{ z(t) - \int_D M(x', t)\hat{u}(x', t)dD_{x'} \right\} . \quad (2.17)$$

The gain, $A(x, t, t)$, remains to be determined. We will make a notational change and denote $A(x, t, t)$ as $K(x, t)$. Starting again with the Wiener-Hopf equation and using (2.12) and (1.7) we have

$$\int_D E\left\{\tilde{u}(x, t)u^T(x', t')\right\} M^T(x', t')dD_{x'}$$

$$- \int_D E\left\{\tilde{u}(x, t)\hat{u}^T(x', t')\right\} M^T(x', t')dD_{x'}$$

$$+ E\left\{u(x, t)v^T(t')\right\} - E\left\{\hat{u}(x, t)v^T(t')\right\} = 0 , \; t > t' , \quad (2.18)$$

which reduces to

$$\int_D E\left\{\tilde{u}(x, t)\tilde{u}^T(x', t')\right\} M^T(x', t')dD_{x'} = A(x, t, t')R(t') , \; t > t'. \quad (2.19)$$

Both sides of (2.19) are continuous in t, thus we can take the limit $t' \to t$. Doing so results in

$$K(x, t) = \int_D P(x, x', t)M^T(x', t)dD_{x'}R^{-1}(t) \ . \qquad (2.20)$$

The estimate is specified by Equations (2.17) and (2.20) and the initial and boundary conditions

$$\hat{u}(x, o) = 0 \qquad (2.21)$$

$$\beta_x\hat{u}(x, t) = 0 \ , \quad x\varepsilon\partial D \ . \qquad (2.22)$$

We can now derive the equation for the error covariance matrix. The error, $\tilde{u}(x, t)$, is given by

$$\frac{\partial\tilde{u}}{\partial t}(x, t) = L_x\tilde{u}(x, t) - K(x, t)\int_D M(x', t)\tilde{u}(x', t)dD_{x'}$$

$$+ G(x, t)w(x, t) - K(x, t)v(t) \quad , \quad t > 0 \ , \quad x\varepsilon D \ . \qquad (2.23)$$

With initial and boundary conditions

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

$$\beta\tilde{u}(x, t) = 0 \quad , \ x\varepsilon\partial D \ . \qquad (2.24)$$

The solution to this partial differential equation can be expressed in terms of the Green's function matrix Φ, as follows

$$\tilde{u}(x, t) = \int_D \Phi(x, t; x", 0)u_0(x")dD_{x"} + \int_0^t\int_D \Phi(x, t; x", t")$$

$$\cdot \left\{G(x", t")w(x", t") - K(x", t")v(t")\right\} dD_{x"}dt" \ , \qquad (2.25)$$

where

$$\frac{\partial}{\partial t}\Phi(x, t; x', t') = L_x\Phi(x, t; x', t') - K(x, t)\int_D M(x'', t)\Phi(x'', t; x', t')dD_{x''}$$

$$\lim_{t \to t'^+}\Phi(x, t; x', t') = I\delta(x - x') \tag{2.26}$$

$$\Phi(x, t; x', t') = 0 \quad , \quad x\varepsilon\partial D \ .$$

Directly from the definition of $P(x, x', t)$ and (2.26) we can determine P to be given by

$$\frac{\partial P}{\partial t}(x, x', t) = L_x P(x, x', t) + P(x, x', t)L_{x'}^T - K(x, t)R(t)K^T(x', t) + G(x, t)Q(x, x', t)G^T(x', t) \tag{2.27}$$

$$P(x, x', 0) = P_0(x, x') \tag{2.28}$$

$$P(x, x', t) = 0 \quad , \quad x\varepsilon\partial D \ . \tag{2.29}$$

The complete filter is given by Equations (2.17), (2.21), (2.22), (2.20), (2.27), (2.28), and (2.29). These equations can be rewritten in terms of the individual components of the observation matrix and R as follows (the initial and boundary conditions remain unchanged),

$$\frac{\partial \hat{u}}{\partial t}(x, t) = L_x \hat{u}(x, t) + \int_D P(x, x'', t) \sum_{j=1}^{p} \sum_{i=1}^{p} M_j^T(x'', t) R^{\ddagger}(x_j, x_i, t)$$
$$\cdot \left\{ z_i(t) - \int_D M_i(x''', t) \hat{u}(x''', t) dD_{x'''} \right\} dD_{x''} \qquad (2.30)$$

$$\frac{\partial P}{\partial t}(x, x', t) = L_x P(x, x', t) + P(x, x', t) L_{x'}^T$$
$$+ G(x, t) Q(x, x', t) G^T(x', t) - \int_D \int_D P(x, x'', t) \sum_{j=1}^{p} \sum_{j=1}^{p}$$
$$\cdot M_j^T(x'', t) R^{\ddagger}(x_j, x_i, t) M_i(x''', t) P(x''', x', t) dD_{x''} dD_{x'''}. \qquad (2.31)$$

2.4.2 Discrete-Instantaneous Time Observations*

In considering discrete time observations we divide the problem into two parts, prediction and up-dating. During the interval $t_{k-1} < t < t_k$ no new information is available, thus we predict an estimate of the state based on the past estimate at time t_{k-1}. At time t_k an observation is recorded and we up-date the estimate based on this observation.

The prediction equations were derived by Tzafestas and Nightingale (1968a) and are given in Table 2.1. We derive the up-date equations.

Using difference equations instead of differential equations, the procedure of Section 2.4.1 for continuous time observations is followed here for discrete-instantaneous observations. From the orthogonal pro-

*The orthogonal projection lemma is applied in Appendix B to discrete time observations of lumped parameter systems.

jection lemma, the optimality condition, Equation (2.7), becomes

$$\int_D \int_D \eta^T(x'') \sum_{i=1}^{k} E\left\{\tilde{u}(x'', t_k)\, z^T(t_i)\right\} B^T(x''', t_k, t_i)\eta(x''')\, dD_{x''}\, dD_{x'''} = 0, \quad (2.32)$$

for all η and B.

The theorem for the Wiener-Hopf equation is now stated.

Theorem 2.2

$\hat{u}(x, t_k)$ satisfies the optimality condition, Equation (2.32), if and only if

$$E\left\{\tilde{u}(x, t_k) z^T(t_i)\right\} = 0 \quad , \quad t_i \leq t_k \, . \quad (2.33)$$

Proof

Follows directly from Theorem 2.1. #

Again, Equation (2.33) will be referred to as the Wiener-Hopf equation (for discrete-instantaneous observations), and the following forms will also be useful.

$$E\left\{u(x, t_k) z^T(t_i)\right\} - \sum_{j=1}^{k} A(x, t_k, t_j) E\left\{z(t_j) z^T(t_i)\right\} = 0 \, ,$$

$$1 \leq i \leq k \quad (2.34)$$

$$E\left\{\tilde{u}(x, t_k) \hat{u}^T(x', t_l)\right\} = 0, \quad 1 \leq l \leq k. \quad (2.35)$$

Consider the difference of the Wiener-Hopf equation at times t_k and t_{k-1},

$$E\left\{\tilde{u}(x, t_k)z^T(t_i)\right\} - E\left\{\tilde{u}(x, t_{k-1})z^T(t_i)\right\} = 0 \ , \ 1 \le i \le k-1. \quad (2.36)$$

Using (2.34) and rearranging terms we have

$$E\left\{\left[u(x, t_k) - u(x, t_{k-1})\right]z^T(t_i)\right\} - A(x, t_k, t_k)E\left\{z(t_k)z^T(t_i)\right\}$$
$$- \sum_{j=1}^{k-1}\left[A(x, t_k, t_j) - A(x, t_{k-1}, t_j)\right]E\left\{z(t_j)z^T(t_i)\right\}$$
$$= 0 \ , \ 1 \le i \le k-1 \ . \quad (2.37)$$

The state, $u(x, t_k)$, can be expressed in terms of the Green's function matrix, ψ, as follows

$$u(x, t_k) = \int_D \psi(x, t_k; x', t_{k-1})u(x', t_{k-1})dD_{x'}$$
$$+ \int_{t_{k-1}}^{t_k}\int_D \psi(x, t_k; x', t')G(x', t')w(x', t')dD_{x'}dt' \ , \quad (2.38)$$

where
$$\frac{\partial \psi}{\partial t}(x, t; x', t') = L_x\psi(x, t; x', t')$$

$$\lim_{t \to t'^+}\psi(x, t; x', t') = I\delta(x - x')$$

$$\psi(x, t; x', t') = 0 \ , \ x \in \partial D \ . \quad (2.39)$$

After rearrangement, use of (2.38), (2.34), and (1.8), and attention to the stochastic properties of v and w, Equation (2.37) becomes

$$\sum_{j=1}^{k-1} \Big\{ \Big[\int_D \psi(x, t_k; x', t_{k-1}) A(x', t_{k-1}, t_j) dD_{x'} - A(x, t_{k-1}, t_j) \Big]$$

$$- A(x, t_k, t_k) \int_D \int_D M(x', t_k) \psi(x', t_k; x'', t_{k-1}) A(x'', t_{k-1}, t_j) dD_{x''} dD_{x'}$$

$$- \Big[A(x, t_k, t_j) - A(x, t_{k-1}, t_j) \Big] \Big\} E\left\{ z(t_j) z^T(t_i) \right\} = 0, \quad 1 \le i \le k-1.$$

In a proof analogous to that for Equation (2.15) we can use the orthogonal projection lemma (Equation (2.8)) to prove the following

$$\Big[\int_D \psi(x, t_k; x', t_{k-1}) A(x', t_{k-1}, t_j) dD_{x'} - A(x, t_{k-1}, t_j) \Big]$$

$$- A(x, t_k, t_k) \int_D \int_D M(x', t_k) \psi(x', t_k; x'', t_{k-1}) A(x'', t_{k-1}, t_j) dD_{x''} dD_{x'}$$

$$- \Big[A(x, t_k, t_j) - A(x, t_{k-1}, t_j) \Big] = 0 \ , \quad 1 \le j \le k-1 \ . \qquad (2.40)$$

Taking the difference between $\hat{u}(x, t_k)$ and $\hat{u}(x, t_{k-1})$ and substituting (2.40), we find a recursive equation for $\hat{u}(x, t_k)$,

$$\hat{u}(x, t_k) = \int_D \psi(x, t_k; x', t_{k-1}) \hat{u}(x', t_{k-1}) dD_{x'} + A(x, t_k, t_k)$$

$$\cdot \Big\{ z(t_k) - \int_D M(x', t_k) \int_D \psi(x', t_k; x'', t_{k-1}) \hat{u}(x'', t_{k-1}) dD_{x''} dD_{x'} \Big\}$$

$$= \hat{u}(x, t_k|t_{k-1}) + A(x, t_k, t_k)\left\{z(t_k) - \int_D M(x', t_k)\hat{u}(x', t_k|t_{k-1})dD_{x'}\right\}, \quad (2.41)$$

where $\hat{u}(x, t_k|t_{k-1})$ is the previously defined optimal prediction of $u(x, t_k)$.

$A(x, t_k, t_k)$ is determined by substituting Equations (2.41), (2.35), and (1.2) into the Wiener-Hopf equation for i = k. Renaming $A(x, t_k, t_k)$ to be $K(x, t_k)$ we obtain

$$E\left\{\left[[u(x, t_k) - \hat{u}(x, t_k|t_{k-1})] - K(x, t_k)\int_D M(x', t_k)\right.\right.$$
$$\cdot [u(x', t_k) - \hat{u}(x', t_k|t_{k-1})]dD_{x'} - K(x, t_k)v(t_k)\Big]$$
$$\left.\cdot\left[\int_D M(x'', t_k)[u(x'', t_k) - \hat{u}(x'', t_k|t_{k-1})]dD_{x''} + v(t_k)\right]^T\right\} = 0 \quad (2.42)$$

and thus,

$$K(x, t_k) = \int_D P(x, x''', t_k|t_{k-1})M^T(x''', t_k)dD_{x'''}$$
$$\cdot\left\{\int_D\int_D M(x', t_k)P(x', x'', t_k|t_{k-1})M^T(x'', t_k)\right.$$
$$\left.\cdot dD_{x'}dD_{x''} + R(t_k)\right\}^{-1}. \quad (2.43)$$

The up-dated error covariance matrix is determined by direct substitution. It is

$$P(x, x', t_k) = P(x, x', t_k|t_{k-1}) - \int_D P(x, x'', t_k|t_{k-1}) M^T(x'', t_k) dD_{x''} K^T(x', t_k) \quad . \qquad (2.44)$$

The up-dated filter is given by Equations (2.41), (2.43), and (2.44). These equations can be rewritten in terms of the individual components of the observation matrix and R as follows,

$$\hat{u}(x, t_k) = \hat{u}(x, t_k|t_{k-1}) + \int_D P(x, x''', t_k|t_{k-1}) \sum_{j=1}^{p} \sum_{i=1}^{p} M_j^T(x''', t_k) K_{ji}^{\ddagger}(t_k) \cdot \left\{ z_i(t_k) - \int_D M_i(x'', t_k) \hat{u}(x'', t_k|t_{k-1}) dD_{x''} \right\} dD_{x'''} \qquad (2.45)$$

$$P(x, x', t_k) = P(x, x', t_k|t_{k-1}) - \int_D \int_D P(x, x''', t_k|t_{k-1}) \sum_{j=1}^{p} \sum_{i=1}^{p} M_j^T(x''', t_k) \cdot K_{ji}^{\ddagger}(t_k) M_i(x'', t_k) P(x'', x', t_k|t_{k-1}) dD_{x''} dD_{x'''} \qquad (2.46)$$

$$K_{ji} = \int_D \int_D M_j(x'', t_k) P(x'', x''', t_k|t_{k-1}) M_i^T(x''', t_k) dD_{x''} dD_{x'''} + R_{ji}(t_k) \quad . \qquad (2.47)$$

2.4.3 Discrete-Average Time Observations

For the discrete-average time case the filter derivation proceeds

essentially as in Section 2.4.2. The differences that do exist arise because $z(t_k)$ contains information from the entire interval $t_{k-1} \leq t \leq t_k$ that has been compressed into one datum. The derivation will be sketched here.

The Wiener-Hopf equation for discrete-average time observations is identical to the one for discrete-instantaneous time observations, Equation (2.33). Equations (2.34) and (2.35) also apply. After taking the difference of the Wiener-Hopf equation at times t_k and t_{k-1} we can prove the analog to Equation (2.40), which is,

$$\left[\int_D \psi(x, t_k; x', t_{k-1}) A(x', t_{k-1}, t_j) dD_{x'} - A(x, t_{k-1}, t_j)\right]$$

$$- A(x, t_k, t_k) \int_{t_{k-1}}^{t_k} \int_D \int_D M(x', t') \psi(x', t'; x'', t_{k-1}) A(x'', t_{k-1}, t_j)$$

$$\cdot dD_{x''} dD_{x'} dt' - \left[A(x, t_k, t_j) - A(x, t_{k-1}, t_j)\right] = 0 \ , \ 1 \leq j \leq k-1 \ . \tag{2.48}$$

Taking the difference between $\hat{u}(x, t_k)$ and $\hat{u}(x, t_{k-1})$, after considerable rearrangement and introduction of $\hat{u}(x, t' | t_{k-1})$ we have

$$\hat{u}(x, t_k) = \hat{u}(x, t_k | t_{k-1}) + A(x, t_k, t_k)$$

$$\cdot \left\{ z(t_k) - \int_{t_{k-1}}^{t_k} \int_D M(x', t') \hat{u}(x', t' | t_{k-1}) dD_{x'} dt' \right\} . \tag{2.49}$$

Again, renaming $A(x, t_k, t_k)$ to be $K(x, t_k)$, we employ the Wiener-Hopf equation at $i = k$ and determine $K(x, t_k)$,

$$K(x, t_k) = \int_{t_{k-1}}^{t_k} \int_D P(x, t_k | t_{k-1}; x', t' | t_{k-1}) M^T(x', t') dD_{x'} dt'$$

$$\cdot \left\{ \int_{t_{k-1}}^{t_k} \int_{t_{k-1}}^{t_k} \int_D \int_D M(x', t') P(x', t' | t_{k-1}; x'', t'' | t_{k-1}) \right.$$

$$\left. \cdot M^T(x'', t'') dD_{x''} dD_{x'} dt'' dt' + R(t_k) \right\}^{-1} . \qquad (2.50)$$

The up-dated error covariance matrix is

$$P(x, x', t_k) = P(x, x', t_k | t_{k-1})$$

$$- \int_{t_{k-1}}^{t_k} \int_D P(x, t_k | t_{k-1}; x'', t'' | t_{k-1})$$

$$M^T(x'', t'') dD_{x''} dt'' K^T(x', t_k). \qquad (2.51)$$

In the equations for $K(x, t_k)$ and $P(x, x', t_k)$ we need the value of $P(x, t | t_{k-1}; x', t' | t_{k-1})$ for $t_{k-1} \leq t' < t \leq t_k$. This is straightforward to determine from the prediction equation. $P(x, t | t_{k-1}; x', t' | t_{k-1})$ is given by

$$\frac{\partial P}{\partial t}(x, t | t_{k-1}; x', t' | t_{k-1}) = \frac{\partial P^T}{\partial t}(x', t' | t_{k-1}; x, t | t_{k-1})$$

$$= L_x P(x, t | t_{k-1}; x', t' | t_{k-1}) , \; t_{k-1} \leq t' < t \leq t_k \qquad (2.52)$$

with initial condition $P(x, t'|t_{k-1}; x', t'|t_{k-1})$ given by the prediction covariance equation and boundary condition

$$\beta_x P(x, t|t_{k-1}; x' t'|t_{k-1}) = 0 \quad , \quad x \varepsilon \partial D.$$

The up-dated filter can be rewritten in terms of the individual components of the observation matrix and R as follows,

$$\hat{u}(x, t_k|t_k) = \hat{u}(x, t_k|t_{k-1}) + \int_{t_{k-1}}^{t_k} \int_D P(x, t_k|t_{k-1}; x''', t'''|t_{k-1}) \cdot \sum_{j=1}^{p} \sum_{i=1}^{p} M_j^T(x''', t''') K_{ji}^{\ddagger}(t_k) \cdot \left\{ z_i(t_k) - \int_{t_{k-1}}^{t_k} \int_D M_i(x'', t'') \hat{u}(x'', t''|t_{k-1}) dD_{x''} dt'' \right\} dD_{x'''} dt''' \tag{2.53}$$

$$P(x, x', t_k) = P(x, x', t_k|t_{k-1}) - \int_{t_{k-1}}^{t_k} \int_D \int_{t_{k-1}}^{t_k} \int_D \cdot P(x, t_k|t_{k-1}; x''', t'''|t_{k-1}) \sum_{j=1}^{p} \sum_{i=1}^{p} M_j^T(x''', t''') K_{ji}^{\ddagger}(t_k) \cdot M_i(x'', t'') P(x'', t''|t_{k-1}; x', t_k|t_{k-1}) dD_{x''} dt'' dD_{x'''} dt''' \tag{2.54}$$

$$K_{ji}(t_k) = \int_{t_{k-1}}^{t_k} \int_D \int_{t_{k-1}}^{t_k} \int_D M_j(x'', t'') P(x'', t''|t_{k-1}; x''', t'''|t_{k-1}) \cdot M_i^T(x''', t''') dD_{x'''} dt''' dD_{x''} dt'' + \mathcal{R}_{ji}(t_k) \ . \tag{2.55}$$

CHAPTER 3

Distributed Parameter Filters: Stochastic Boundary Conditions

In this chapter we present the filters for linear distributed parameter systems with stochastic boundary conditions. Such boundary conditions contain both deterministic and stochastic terms and are of the form

$$\beta_x u = h_b + w_b$$

where h_b and w_b are deterministic and stochastic functions, respectively.

3.1 Introduction

A specific class of systems from those given in Section 1.3 is considered in this chapter. In most previous work, consideration of more general cases has necessitated leaving unspecified functions in the final result. Consideration of a specific class of systems enables us to derive the filters completely such that no additional functions need to be determined for implementation. The class of systems we consider is

Domain, D - one dimensional

Spatial Operator, L_x - second order partial differential operator

Boundary Operator, β_x - mixed or Dirichlet.

Many engineering systems are of this class and the extension to different domains, in particular, higher dimensions, involves no new concepts relative to filtering theory.

We shall see that the effect of the stochastic boundary conditions on the filter is independent of the observation process. (There is one interesting exception; the case of Dirichlet

boundary conditions with observations on the boundary. This is discussed in Section 3.3.2.) Although we do not discuss continuous time, continuous space observation processes, such processes are, as noted in Chapter 1, frequently considered in the literature. We note that the results presented in this chapter can be applied directly to such processes.

Section 3.2 contains the specifications of D, L_x, and β_x and the summary of the results. Sections 3.3.1 and 3.3.2 contain the derivations of the filters for mixed boundary conditions and Dirichlet boundary conditions, respectively.

3.2 Summary of Results

As discussed in Section 3.1, the system we consider is as follows

$$Lu = \frac{\partial u}{\partial t} - L_x u = G(x, t)w(x, t) + h(x, t) \quad x\varepsilon D, \quad 0 < t \tag{3.1}$$

where

$$L_x = \alpha_2(x, t) \frac{\partial^2}{\partial x^2} + \alpha_1(x, t) \frac{\partial}{\partial x} + \alpha_0(x, t) \tag{3.2}$$

and D is the normalized interval $0 < x < 1$, with initial condition $u(x, 0) = u_0(x)$, and boundary conditions

$$\beta_0 u = \left[\phi_0(t)\alpha_2(x, t) \frac{\partial u}{\partial x} + \theta_0(t)u\right]_{x = 0} = h_0(t) + w_0(t) \tag{3.3}$$

$$\beta_1 u = \left[\phi_1(t)\alpha_2(x, t) \frac{\partial u}{\partial x} + \theta_1(t)u\right]_{x = 1} = h_1(t) + w_1(t). \tag{3.4}$$

The stochastic boundary inputs, w_0 and w_1, are independent, zero mean, and white with covariance matrices $Q_0(t)$ and $Q_1(t)$, respectively. The coefficients α, ϕ, and θ are nxn matrices. The boundary conditions are considered in three cases,

Case I - mixed (flux and state terms); ϕ_0 and ϕ_1 are nonsingular; observations in the domain, D, and/or on the boundary

Case II - Dirichlet (state term); θ_0 and θ_1 are nonsingular, $\phi_0 = \phi_1 = 0$; observations restricted to the domain

Case III - Dirichlet (state term); θ_0 and θ_1 are nonsingular, $\phi_0 = \phi_1 = 0$; observations on the boundary and in the domain.

It will be clear that the filters for combinations of the above cases (e.g. $\phi_0 = 0$, ϕ_1 nonsingular) follow directly from the three cases considered.

For Case I, the three observation processes discussed in Chapter 2 are all considered. The partial differential equations for $\hat{u}(x, t)$ and $P(x, x', t)$ are not affected by the stochastic boundary conditions for $u(x, t)$. These partial differential equations are those presented in Table 2.1 for each of the respective observation processes. The single set of boundary conditions in Table 3.1 for Case I applies to all of the observation processes considered.

Similarly, for Case II, the partial differential equations for $\hat{u}(x, t)$ and $P(x, x', t)$ are those presented in Table 2.1 for each of the respective observation processes, and the single set of boundary conditions for all three observation processes is given in Table 3.1.

Table 3.1

Linear Filter Boundary Conditions for Stochastic System Boundary Conditions

Case I

$$\beta_0\hat{u}(0, t) = h_0(t) \qquad \beta_0 P(0, x', t) = -Q_0(t)\phi_0^{-1^T}(t)\delta(x') \qquad x = 0$$

$$\beta_1\hat{u}(1, t) = h_1(t) \qquad \beta_1 P(1, x', t) = Q_1(t)\phi_1^{-1^T}(t)\delta(x' - 1) \qquad x = 1$$

Case II

$$\beta_0\hat{u}(0, t) = h_0(t) \qquad \beta_0 P(0, x', t) = -Q_0(t)\left[\alpha_2(0, t)\theta_0^{-1}(t)\right]^T\delta'(x') \qquad x = 0$$

$$\beta_1\hat{u}(1, t) = h_1(t) \qquad \beta_1 P(1, x', t) = Q_1(t)\left[\alpha_2(1, t)\theta_1^{-1}(t)\right]^T\delta'(x' - 1) \qquad x = 1$$

Case III

$$\beta_0\hat{u}(0, t) = h_0(t) + Q_0(t)\theta_0^{-1^T}(t)M_0^{b^T}K_0^{b^{-1}}(t)\left[z_0^b(t) - M_0^b(t)\theta_0^{-1}(t)h_0(t)\right] \qquad x = 0$$

$$\beta_0 P(0, x', t) = -\left[Q_0(t) - Q_0(t)\theta_0^{-1^T}(t)M_0^{b^T}(t)K_0^{b^{-1}}(t)M_0^b(t)\theta_0^{-1}(t)Q_0(t)\right]\theta_0^{-1^T}(t)\alpha_2^T(0, t)\delta'(x')$$

$$\beta_1\hat{u}(1, t) = h_1(t) + Q_1(t)\theta_1^{-1^T}(t)M_1^{b^T}(t)K_1^{b^{-1}}(t)\left[z_1^b(t) - M_1^b(t)\theta_1^{-1}(t)h_1(t)\right] \qquad x = 1$$

$$\beta_1 P(1, x', t) = \left[Q_1(t) - Q_1(t)\theta_1^{-1^T}(t)M_1^{b^T}(t)K_1^{b^{-1}}(t)M_1^b(t)\theta_1^{-1}(t)Q_1(t)\right]\theta_1^{-1^T}(t)\alpha_2^T(1, t)\delta'(x' - 1)$$

$$K_i^b(t) = R_i^b + M_i^b(t)\theta_i^{-1}(t)Q_i(t)\theta_i^{-1^T}(t)M_i^{b^T}(t) \quad , \quad i = 0, 1$$

z_i^b, M_i^b, and R_i^b are defined in Section 3.3.2.

For Case III, the results indicate that if the system boundary conditions are of the Dirichlet type then only for continuous time observations is it useful to take observations at the boundary. Thus, we will only be concerned with this type of observation process. The boundary conditions in Table 3.1 for Case III apply only for continuous time observations. The partial differential equations for $\hat{Q}(x, t)$ and $P(x, x', t)$ are those presented in Table 2.1 for this process. Discrete time boundary observations need not be considered for reasons discussed in Section 3.3.2.

3.3 Derivation of Filters

The central concept in these derivations is the formal use of the Green's function form of the solution for a partial differential equation to show that a boundary inhomogeneity may be represented in the differential equation, (Stakgold, 1967; Sakawa, 1972).

In the derivations, only the $x = 1$ boundary condition will be taken to be stochastic. The $x = 0$ boundary will be taken to be deterministic. The effect of the stochastic term at $x = 0$ will be obvious and is included in Table 3.1. At both boundaries we set the deterministic term equal to 0 to avoid obscuring the derivations.

Before treating Cases I and II separately and in detail, we develop here the common elements of these derivations. Much of the background mathematical development is identical in these derivations and they follow the same logical arguments. (The derivation in Case III also draws on these concepts, but remains substantially different.)

In Cases I and II the derivations proceed as follows.

Step 1 The system boundary conditions are made deterministic by specifying the stochastic inhomogeneity in the partial differential equation.

Step 2 Application of the known filter for the case of deterministic boundary conditions yields an estimate equation with deterministic boundary conditions. The effect of the stochastic inhomogeneity enters in the covariance equation.

Step 3 The effects of the stochastic inhomogeneity is specified in the boundary conditions for the covariance.

As background material, we discuss first the form of the filters previously derived for deterministic boundary conditions and then the Green's function form of the solution to a partial differential equation.

We start by noting that the system and filters presented in Chapters 1 and 2 may be expressed as follows,

- System

$$\frac{\partial u}{\partial t} - L_x u = W(x, t) + H(x, t) \tag{3.5}$$

- Filters

  - Continuous time observations

$$\frac{\partial \hat{u}}{\partial t} - L_x \hat{u} = S_{\hat{u}}(x, t) + H(x, t) \tag{3.6}$$

$$\frac{\partial P}{\partial t} - L_x P - P L_{x'}^T = S_P(x, x', t) + Q_W(x, x', t) \tag{3.7}$$

- Discrete time observations

  - Prediction

$$\frac{\partial \hat{u}}{\partial t} - L_x \hat{u} = H(x, t) \tag{3.8}$$

$$\frac{\partial P}{\partial t} - L_x P - P L_x^T, = Q_W(x, x', t) \tag{3.9}$$

  - Up-Date

  Not a function of W,

with appropriate initial conditions. The boundary conditions are deterministic and in particular those for the covariance equation are homogeneous.

H represents all deterministic inputs and W all stochastic inputs, with $E\left\{ W(x, t)W^T(x', t') \right\} = Q_W(x, x', t)\delta(t - t')$. $S_{\hat{u}}$ and S_P are not functions of W.

Equations (3.6) - (3.9) show the influence of stochastic inputs on the filtering equations. Their effects are the same for all of the observation processes considered. We see that W enters into the covariance equation and that the estimate equation is unaffected by the presence of the stochastic inputs. In the filter derivations we will 'add' stochastic inputs to the system equation, thus it is important that we see clearly what the effects will be.

We now develop the Green's function form of the solution to systems of the forms

$$Lu = \frac{\partial u}{\partial t} - L_x u = T_u(x, t) \tag{3.10}$$

and

$$\frac{\partial P}{\partial t} - L_x P - P L_{x'}^T = T_p(x, x', t) \tag{3.11}$$

where L_x and the domain are as given in Section 3.2.

Introduce the adjoint Green's function matrix, ψ^*, defined by the following system

$$L_x^* \psi^*(x, t; x', t') = \frac{-\partial \psi^*}{\partial t} - L_x^* \psi^* = 0 \tag{3.12}$$

$$L_x^* \psi^* = \frac{\partial^2}{\partial x^2}\left(\alpha_2^T(x, t)\psi^*\right) - \frac{\partial}{\partial x}\left(\alpha_1^T(x, t)\psi^*\right) + \alpha_0^T(x, t)\psi^* \tag{3.13}$$

with end condition

$$\lim_{t \to t'^-} \psi^*(x, t; x', t') = I\delta(x - x') \tag{3.14}$$

and boundary conditions to be specified later. The operators L_x^* and L_x^* are the adjoints of the operators L_x and L_x, respectively. Note that ψ^* is related to the Green's function matrix ψ, introduced in Equation (2.39), by the relationship $\psi^{*T}(x', t'; x, t) = \psi(x, t; x', t')$. In spite of this simple relationship we prefer to use ψ^* in this discussion because it is most natural to do so.

By integration by parts it can be shown that

$$\int_0^t \int_0^1 \left\{ \psi^{*T} L_{x'} u - (L_{x'}^* \psi^*)^T u \right\} dx'dt' = \int_0^1 \psi^{*T} u dx' \Big|_0^t$$

$$+ \int_0^t \left[- \psi^{*T} \alpha_2 \frac{\partial u}{\partial x'} + \left(\frac{\partial}{\partial x'} (\psi^{*T} \alpha_2) - \psi^{*T} \alpha_1 \right) u \right] dt' \Big|_0^1 \tag{3.15}$$

and

$$\int_0^t \int_0^1 \int_0^1 \left\{ \psi^{*T} \left[\frac{\partial}{\partial t'} P - L_{x''} P - P L_{x'''}^T \right] \psi^* \right.$$

$$\left. - \left[\psi^{*T} L_{x''}^{*T} \right] P\psi^* - \psi^{*T} P \left[L_{x'''}^* \psi^* \right] \right\} dx''' dx'' dt' = \int_0^1 \int_0^1 \psi^{*T} P \psi^* dx''' dx'' \Big|_0^t$$

$$+ \int_0^t \int_0^1 \left[-\psi^{*T} \alpha_2 \frac{\partial P}{\partial x'''} + \left(\frac{\partial}{\partial x''} (\psi^{*T} \alpha_2) - \psi^{*T} \alpha_1 \right) P \right] \psi^* dx''' dt' \Big|_0^1$$

$$+ \int_0^t \int_0^1 \psi^{*T} \left[-\frac{\partial P}{\partial x'''} \alpha_2^T \psi^* + P \left(\frac{\partial}{\partial x'''} (\alpha_2^T \psi^*) - \alpha_1^T \psi^* \right) \right] dx'' dt' \Big|_0^1 . \quad (3.16)$$

Using Equations (3.12) and (3.14) we find the 'solution' for u and P to be

$$u(x, t) = \int_0^1 \psi^{*T}(x', 0; x, t) u_0(x') dx'$$

$$+ \int_0^t \int_0^1 \psi^{*T}(x', t'; x, t) L_{x'} u(x', t') dx' dt'$$

$$- \int_0^t \left[-\psi^{*T}(x', t'; x, t) \alpha_2(x', t') \frac{\partial u(x', t')}{\partial x'} \right.$$

$$+ \left(\frac{\partial}{\partial x'} \left(\psi^{*T}(x', t'; x, t) \alpha_2(x', t') \right) \right.$$

$$\left. \left. - \psi^{*T}(x', t'; x, t) \alpha_1(x', t') \right) u(x', t') \right] dt' \Big|_0^1 \quad (3.17)$$

and

$$P(x, x', t) = \int_0^1 \int_0^1 \psi^{*T}(x'', 0; x, t) P_0(x'', x''') \psi^*(x''', 0; x', t) dx''' dx''$$

$$+ \int_0^t \int_0^1 \int_0^1 \psi^{*T}(x'', t'; x, t) \left[\frac{\partial P(x'', x''', t')}{\partial t'} - L_{x''} P(x'', x''', t') \right.$$

$$\left. - P(x'', x''', t') L^T_{x'''} \right] \psi^*(x''', t'; x', t)\, dx'''dx''dt'$$

$$- \int_0^t \int_0^1 \left[-\psi^{*T}(x'', t'; x, t)\alpha_2(x'', t') \frac{\partial P(x'', x''', t')}{\partial x''} \right.$$

$$+ \left(\frac{\partial}{\partial x'''} \left(\psi^{*T}(x'', t'; x, t)\alpha_2(x'', t') \right) \right.$$

$$\left. \left. - \psi^{*T}(x'', t'; x, t)\alpha_1(x'', t') \right) P(x'', x''', t') \right] \psi^*(x''', t'; x', t) dx'''dt' \Big|_0^1$$

$$- \int_0^t \int_0^1 \psi^{*T}(x'', t'; x, t) \left[- \frac{\partial P(x'', x''', t')}{\partial x'''} \alpha_2^T(x''', t') \psi^*(x''', t'; x', t) \right.$$

$$+ P(x'', x''', t') \left(\frac{\partial}{\partial x'''} \left(\alpha_2^T(x''', t') \psi^*(x''', t'; x', t) \right) \right.$$

$$\left. \left. - \alpha_1^T(x''', t') \psi^*(x''', t'; x', t) \right) \right] dx''dt' \Big|_0^1 . \qquad (3.18)$$

Equations (3.17) and (3.18) hold for the boundary conditions of both Case I and Case II. The boundary conditions on ψ^* have yet to be specified. To proceed with our discussion we must now consider the two cases individually.

3.3.1 Mixed Boundary Conditions

For Case I, the adjoint boundary conditions are

$$\beta_0^* \psi^{*T} = \left[\frac{\partial}{\partial x}(\psi^{*T}\alpha_2) - \psi^{*T}(\alpha_1 - \phi_0^{-1}\theta_0)\right]\Big|_{x=0} = 0$$
$$\beta_1^* \psi^{*T} = \left[\frac{\partial}{\partial x}(\psi^{*T}\alpha_2) - \psi^{*T}(\alpha_1 - \phi_1^{-1}\theta_1)\right]\Big|_{x=1} = 0 \quad . \tag{3.19}$$

Recall that we have set only the x = 1 boundary condition on u to be stochastic. Thus the system boundary conditions are

$$\beta_0 u = \left[\phi_0\alpha_2 \frac{\partial u}{\partial x} + \theta_0 u\right]\Big|_{x=0} = 0$$
$$\beta_1 u = \left[\phi_1\alpha_2 \frac{\partial u}{\partial x} + \theta_1 u\right]\Big|_{x=1} = w_1(t) \quad . \tag{3.20}$$

Combining Equations (3.17), (3.19) and (3.20) we have for u,

$$\begin{aligned} u(x, t) = &\int_0^1 \psi^{*T}(x', 0; x, t)u_0(x')dx' \\ &+ \int_0^t \int_0^1 \psi^{*T}(x', t'; x, t)T_u(x', t')dx'dt' \\ &+ \int_0^t \psi^{*T}(1, t'; x, t)\phi_1^{-1}(t')w_1(t')dt' \end{aligned} \tag{3.21}$$

which is equivalent to

$$\begin{aligned} u(x, t) = &\int_0^1 \psi^{*T}(x', 0; x, t)u_0(x')dx' \\ &+ \int_0^t \int_0^1 \psi^{*T}(x', t'; x, t)\left[T_u(x', t')\right. \end{aligned}$$

$$+ \phi_1^{-1}(t')w_1(t')2\delta(x' - 1)\Big] dx'dt' \ . \tag{3.22}$$

Equation (3.22) is the solution to

$$Lu = T_u(x, t) + \phi_1^{-1}(t)w_1(t)2\delta(x - 1) \tag{3.23}$$

with deterministic boundary conditions. Thus we have shown that u can be equivalently described by (3.10) with (3.20) as the boundary condition or by (3.23) with deterministic boundary conditions.

Based on Equations (3.18) and (3.19), we can also show that the covariance, P, described by Equation (3.11) with boundary conditions

$$\begin{aligned} &\beta_0 P(0, x', t) = 0 \\ &\beta_1 P(1, x', t) = Q_1(t)\phi_1^{-1^T}(t)\delta(x' - 1) \ , \end{aligned} \tag{3.24}$$

is equivalent to the specification

$$\begin{aligned} \frac{\partial P}{\partial t} - L_x P - P L_{x'}^T &= T_p(x, x', t) \\ &+ 2\delta(x - 1)\phi_1^{-1}(t)Q_1(t)\phi_1^{-1^T}(t)\delta(x' - 1)2 \end{aligned} \tag{3.25}$$

with homogeneous boundary conditions at both x = 0 and x = 1.

We now have all of the background material and can proceed through the derivation.

Step 1 - The system is described by Equation (3.1) and the stochastic boundary conditions (3.20). An equivalent specification of the system is

$$Lu = \frac{\partial u}{\partial t} - L_x u = G(x, t)w(x, t) + 2\delta(x - 1)\phi_1^{-1}(t)w_1(t) \tag{3.26}$$

with deterministic boundary conditions.

Step 2 - The filter for this system is known from the results of Chapter 2. It is

$$\frac{\partial \hat{u}}{\partial t} - L_x \hat{u} = S_{\hat{u}}(x, t) \tag{3.27}$$

$$\begin{aligned} &\beta_0 \hat{u}(0, t) = 0 \\ &\beta_0 \hat{u}(1, t) = 0 \end{aligned} \tag{3.28}$$

$$\begin{aligned} \frac{\partial P}{\partial t} - L_x P - P L_{x'}^T = {} & S_P(x, x', t) + G(x, t)Q(x, x', t)G^T(x', t) \\ & + 2\delta(x - 1)\phi_1^{-1}(t)Q_1(t)\phi_1^{-1^T}(t)\delta(x' - 1)2 \end{aligned} \tag{3.29}$$

$$\beta_0 P(0, x', t) = 0$$

$$\beta_0 P(1, x', t) = 0 . \tag{3.30}$$

Step 3 - The covariance described by Equations (3.29) and (3.30) is equivalently specified as

$$\frac{\partial P}{\partial t} - L_x P - P L_{x'}^T = S_p(x, x', t) + G(x, t)Q(x, x', t)G^T(x', t) \tag{3.31}$$

$$\beta_0 P(0, x', t) = 0$$

$$\beta_1 P(1, x', t) = Q_1(t)\phi_1^{-1^T}(t)\delta(x' - 1) . \tag{3.32}$$

Equations (3.27), (3.28), (3.31) and (3.32) and the appropriate initial conditions represent the filters for the stochastic boundary conditions of Case I.

3.3.2 Dirichlet Boundary Conditions

In Cases II and III we consider Dirichlet boundary conditions. First we will consider Case II; that is, the case of no observations on the boundary.

The system boundary conditions are

$$\beta_0 u = \theta_0 u \Big|_{x = 0} = 0$$
$$\beta_1 u = \theta_1 u \Big|_{x = 1} = w_1(t). \tag{3.33}$$

As for Case I, Equations (3.17) and (3.18) are the main elements in this derivation and we start by defining the appropriate adjoint boundary conditions. For Case II, the adjoint boundary conditions are

$$\beta_0^* \psi^{*T} = \psi^{*T} \theta_0 \Big|_{x = 0} = 0$$
$$\beta_1^* \psi^{*T} = \psi^{*T} \theta_1 \Big|_{x = 1} = 0 . \tag{3.34}$$

Combine Equations (3.17), (3.33) and (3.34) and note that the following relationship holds by the properties of the derivative of the delta function

$$- \int_0^t \left[\frac{\partial}{\partial x'} \psi^{*T}(x', t'; x, t) \right]_{x' = 1} \alpha_2(1, t') \theta_1^{-1}(t') w_1(t') dt'$$

$$= \int_0^t \int_0^1 \psi^{*T}(x', t'; x, t) \alpha_2(1, t') \theta_1^{-1}(t') w_1(t') 2\delta'(x' - 1) dx' dt' \tag{3.35}$$

It thus follows that u can be described by the equation

$$Lu = T_u(x, t) + \alpha_2(1, t)\theta_1^{-1}(t)w_1(t)2\delta'(x - 1) \tag{3.36}$$

with deterministic boundary conditions.

Based on Equations (3.18) and (3.34), we can also show that the covariance, P, described by Equation (3.11) with boundary conditions

$$\begin{aligned} &\beta_0 P(0, x', t) = 0 \\ &\beta_1 P(1, x', t) = Q_1(t)\left[\alpha_2(1, t)\theta_1^{-1}(t)\right]^T\delta'(x' - 1) , \end{aligned} \tag{3.37}$$

is equivalent to the specification

$$\begin{aligned} \frac{\partial P}{\partial t} - L_x P - P L_{x'}^T = T_p(x, x', t) + 2\delta'(x - 1)\alpha_2(1, t)\theta_1^{-1}(t)Q_1(t) \\ \cdot\left[\alpha_2(1, t)\theta_1^{-1}(t)\right]^T\delta'(x' - 1)2 \end{aligned} \tag{3.38}$$

with homogeneous boundary conditions. The three step derivation previously outlined is now followed and the results are as stated in Table 3.1.

We conclude this chapter on stochastic boundary conditions by considering continuous time observations on the boundary for systems with Dirichlet boundary conditions. As indicated in Table 3.1, we introduce a minor amount of additional notation in treating this type of observation process. Let there be p + 2 observations with the total observation vector denoted Z and defined by

$$Z(t) = \begin{bmatrix} z(t) \\ z_0^b(t) \\ z_1^b(t) \end{bmatrix} = \begin{bmatrix} \int_0^1 M(x', t)u(x', t)dx' \\ M_0^b(t)u(0, t) \\ M_1^b(t)u(1, t) \end{bmatrix} + \begin{bmatrix} v(t) \\ v_0^b(t) \\ v_1^b(t) \end{bmatrix} \tag{3.39}$$

where z represents the p observations in the domain and z_0^b and z_1^b represent the boundary observations at x = 0 and x = 1, respectively. The boundary observation errors, v_0^b and v_1^b, are zero mean and have second moment properties defined by

$$E\left\{ \begin{bmatrix} v(t) \\ v_0^b(t) \\ v_1^b(t) \end{bmatrix} \begin{bmatrix} v^T(t') & v_0^{bT}(t') & v_1^{bT}(t') \end{bmatrix} \right\}$$

$$= \begin{bmatrix} R(t) & 0 & 0 \\ 0 & R_0^b(t) & 0 \\ 0 & 0 & R_1^b(t) \end{bmatrix} \delta(t - t') \tag{3.40}$$

where R(t), $R_0^b(t)$, and $R_1^b(t)$ are positive definite, and we set v_0^b and v_1^b to be independent of each other and of v. The terms z, M, v and R are defined as they were in Chapter 1.

In deriving the filter for this case we will only consider the x = 1 boundary to be stochastic . Thus we will not consider observations at the x = 0 boundary and the terms in Equations (3.39) and (3.40) relating to x = 0 are to be removed. The system boundary conditions are

$$\beta_0 u = \theta_0 u \Big|_{x = 0} = 0$$

$$\beta_1 u = \theta_1 u \Big|_{x = 1} = w_1(t). \tag{3.41}$$

Finally, as in the derivations in Chapter 2, we set $h = 0$ and $\bar{u}_0 = 0$. The system may now be specified with deterministic boundary conditions and system equation

$$\frac{\partial u}{\partial t} - L_x u = G(x, t)w(x, t) + \alpha_2(1, t)\left[\theta_1^{-1}(t)w_1(t)\right]2\delta'(x - 1). \tag{3.42}$$

We seek the estimate of the form

$$\omega(x, t) = \int_0^t \begin{bmatrix} B(x, t, t') & B_1^b(x, t, t') \end{bmatrix} \begin{bmatrix} z(t') \\ z_1^b(t') \end{bmatrix} dt' \tag{3.43}$$

which minimizes the error covariance matrix. Following the conventions in Chapter 2, $\hat{u}$, A, and A_1^b are the optimal ω, B and B_1^b, respectively. Based on the orthogonal projection lemma the optimality condition is

$$\int_0^1\int_0^1 \eta^T(x') \int_0^t E\left\{\tilde{u}(x', t)\begin{bmatrix} z^T(t') & z_1^{bT}(t') \end{bmatrix} \begin{bmatrix} B^T(x'', t, t') \\ B_1^{bT}(x'', t, t') \end{bmatrix}\right\}$$

$$\cdot\, dt'\eta(x'')dx'dx'' = 0 \quad , \tag{3.44}$$

for all η, B and B_1^b. This condition leads directly to a Wiener-Hopf equation which can then be expressed as the following two separate relationships

$$E\{\tilde{u}(x, t)z^T(t')\} = 0 \ , \ t > t' \tag{3.45a}$$

and

$$E\{\tilde{u}(x, t)z_1^{bT}(t')\} = 0 \ , \ t > t' \ . \tag{3.45b}$$

Alternative expressions for these equations are

$$E\{u(x, t)z^T(t')\} - \int_0^t \Big[A(x, t, t")E\{z(t")z^T(t')\} + A_1^b(x, t, t")E\{z_1^b(t")z^T(t')\}\Big] dt" = 0 \ , \ t > t' \tag{3.46a}$$

and

$$E\left\{u(x, t)z_1^{bT}(t')\right\} - \int_0^t \left[A(x, t, t")E\left\{z(t")z_1^{bT}(t')\right\} + A_1^b(x, t, t")E\left\{z_1^b(t")z_1^{bT}(t')\right\}\right] dt" = 0 \ , \ t > t' \ . \tag{3.46b}$$

Differentiating (3.46a) and (3.46b) with respect to t we have

$$\int_0^t C(x, t, t")E\{z(t")z^T(t')\}\, dt" + \int_0^t C_1^b(x, t, t")E\{z_1^b(t")z^T(t')\}dt" = 0 \ , \ t > t' \tag{3.47a}$$

and

$$\int_0^t C(x, t, t")E\{z(t")z_1^{bT}(t')\}\, dt" + \int_0^t C_1^b(x, t, t")E\{z_1^b(t")z_1^{bT}(t')\}\, dt" = 0 \ , \ t > t' \tag{3.47b}$$

where

$$C(x, t, t") = - \frac{\partial A(x, t, t")}{\partial t} + L_x A(x, t, t")$$

$$- A(x, t, t)\int_0^1 M(x', t)A(x', t, t")dx' \qquad (3.48a)$$

and

$$C_1^b(x, t, t") = - \frac{\partial A_1^b(x, t, t")}{\partial t} + L_x A_1^b(x, t, t")$$

$$- A(x, t, t)\int_0^1 M(x', t)A_1^b(x', t, t")dx'. \qquad (3.48b)$$

Equations (3.48a) and 3.48b) are not exactly analogous because, for $t > t'$, $E\left\{z_1^b(t)z_1^{bT}(t')\right\}$ and $E\left\{z_1^b(t)z^T(t')\right\}$ vanish whereas $E\left\{z(t)z^T(t')\right\}$ and $E\left\{z(t)z_1^{bT}(t')\right\}$ do not.

It can be shown that C and C_1^b vanish and that the differential expression for $\hat{u}$ becomes

$$\frac{\partial \hat{u}}{\partial t}(x, t) = L_x \hat{u}(x, t) + A(x, t, t)\left[z(t) - \int_0^1 M(x', t)\hat{u}(x', t)dx'\right]$$

$$+ A_1^b(x, t, t)z_1^b(t) \ , \qquad (3.49)$$

with deterministic boundary conditions. $A(x, t, t)$ and $A_1^b(x, t, t)$ are found to be

$$A(x, t, t) = K(x, t) = \int_0^1 P(x, x', t)M^T(x', t)dx'R^{-1}(t) \qquad (3.50)$$

$$A_1^b(x, t, t) = \alpha_2(1, t)\theta_1^{-1}(t)Q_1(t)\theta_1^{-1^T}(t)M_1^{b^T}(t)K_1^{b^{-1}}(t)2\delta'(x - 1)$$

$$(3.51)$$

where

$$K_1^b(t) = \left[R_1^b(t) + M_1^b(t)\theta_1^{-1}(t)Q_1(t)\theta_1^{-1^T}(t)M_1^{b^T}(t) \right]. \tag{3.52}$$

The error covariance equation is thus,

$$\frac{\partial P}{\partial t}(x, x', t) = L_x P(x, x', t) + P(x, x', t)L_{x'}^T - K(x, t)R(t)K^T(x', t)$$

$$+ G(x, t)Q(x, x', t)G^T(x', t) + 2\delta'(x - 1)\alpha_2(1, t)\theta_1^{-1}(t)$$

$$\cdot\left[Q_1(t) - Q_1(t)\theta_1^{-1^T}(t)M_1^{b^T}(t)K_1^{b^{-1}}(t)M_1^b(t)\theta_1^{-1}(t)Q_1(t) \right]$$

$$\cdot\ \theta_1^{-1^T}(t)\alpha_2^T(1, t)\delta'(x' - 1)2 \tag{3.53}$$

with homogeneous boundary conditions. The filter can alternatively be written with an inhomogeneous boundary condition at x = 1. The partial differential equations for $\hat{u}$ and P are then those in Table 2.1 and the boundary condition is that given in Table 3.1.

The filters for discrete-instantaneous time and discrete-average time boundary observations have also been investigated. Both of these discrete time filters are impractical. In the instantaneous case no information is gained by taking such observations. In the case of averaged observations, the computational burden would be excessive.

In the instantaneous case, the reason for this result is clear. Discrete-instantaneous time boundary observations yield a characterization of an input (the boundary value) at an instant. The influence of an

input at an instant is infinitesimal; thus, no information is gained that can be used to estimate the state.

CHAPTER 4

Summary

In this work, we have considered the linear distributed parameter filtering problem for systems of engineering interest. Formal procedures have been employed in developing filters for a variety of observation processes and boundary conditions. Integral and point-wise spatial observation processes, along with continuous and discrete temporal observation processes, have been considered. For the case of deterministic boundary conditions, the filters have been developed for a general linear system. In order to provide filters which are completely specified, for systems with stochastic boundary conditions a specific linear system has been considered.

References

Aidarous S. E. and Ghonaimy M. A. R. (1976) Optimal filtering for a class of stochastic distributed systems. Int. J. Systems Science, 7, 545 - 556.

Ajinkya M. B., Ray W. H., Yu T. K. and Seinfeld J. H. (1975) The application of an approximate non-linear filter to systems governed by coupled ordinary and partial differential equations. Int. J. Systems Science, 6, 313 - 332.

Athans M. (1970) Toward a practical theory for distributed parameter systems. I.E.E.E. Trans. Auto. Control, AC-15, 245 - 247.

Atre S. R. (1972) A note on the Kalman-Bucy filtering for linear distributed-parameter systems. I.E.E.E. Trans. Auto. Control, AC-17, 712 - 713.

Atre S. R. and Lamba S. S. (1972a) Derivation of an optimal estimator for distributed parameter systems via maximum principle. I.E.E.E. Trans. Auto. Control, AC-17, 388 - 390.

Atre S. R. and Lamba S. S. (1972b) Optimal estimation in distributed processes using the innovations approach. I.E.E.E. Trans. Auto. Control, AC-17, 710 - 712.

Atre S. R. and Lamba S. S. (1972c) Filtering for linear distributed parameter systems via boundary measurements. Proc. I.E.E., 119, 757 - 759.

Balakrishnan A. V. (1974a) Stochastic optimization theory in Hilbert spaces - I. J. Appl. Math. and Opt., 1, 97 - 120.

Balakrishnan A. V. (1974b) Identification and Stochastic Control of a Class of Distributed Systems with Boundary Noise. University of California, Los Angeles, Rept. UCLA-ENG-7461.

Balakrishnan A. V. and Lions J. L. (1967) State estimation for infinite-dimensional systems. J. Computers and Systems Science, 1, 391 - 403.

Bensoussan A. (1971) Filtrage Optimal des Systemes Lineaires, Dunod, Paris.

Curtain R. (1975a) A survey of infinite-dimensional filtering. S.I.A.M. Review, 17, 395 - 411.

Curtain R. (1975b) Infinite-dimensional filtering. S.I.A.M. J. Control, 13, 89 - 104.

Falb P. L. (1967) Infinite dimensional filtering: the Kalman-Bucy filter in Hilbert space. Inf. and Control, 11, 102 - 137.

Fujishige S. (1975) Minimum-variance estimation for a linear continuous-discrete system with noisy state-integral observation. I.E.E.E. Auto. Control, AC-20, 139 - 140.

Hwang M., Seinfeld J. H. and Gavalas G. R. (1972) Optimal least-square filtering and interpolation in distributed parameter systems. J. Math. Anal. Appl., 39, 49 - 74.

Jazwinski A. H. (1970) Stochastic Processes and Filtering Theory. Academic Press, New York.

Kalman R. E. and Bucy R. S. (1961) New results in filtering and prediction theory. J. Basic Eng., Series D, 83, 95 - 108.

Koda M. (1975) Dynamic Estimation of Diffusion Systems-Application to Prediction of Air Pollution, Ph.D. Thesis, University of Tokyo, Tokyo.

Kumar G. R. V. and Sage A. P. (1972) An innovations approach to distributed parameter filtering and smoothing. Inf. Science, 4, 193 - 215.

Kumar S. and Seinfeld J. H. (1978) Optimal location of measurements for distributed parameter estimation. I.E.E.E. Trans. Auto. Control, AC-23, To appear.

Kushner H. J. (1970) Filtering for linear distributed parameter systems. S.I.A.M. J. Control, 8, 346 - 359.

Lamont G. B. and Kumar K. S. P. (1972) State estimation in distributed parameter systems via least squares and invariant imbedding. J. Math. Anal. Appl., 38, 588 - 606.

Lee K. Y. (1976) Optimal estimation of first order distributed systems: multiplicity and martingale approach. Preprints Joint Auto. Control Conf., 634 - 636.

Meditch J. S. (1970) On state estimation for distributed parameter systems. J. Franklin Inst., 290, 49 - 59.

Meditch J. S. (1971) Least-squares filtering and smoothing for linear distributed parameter systems. Automatica, 7, 315 - 322.

Padmanabhan L. and Colantuoni G. (1974) Sequential estimation in distributed systems. Int. J. Systems Science, 5, 973 - 986.

Phillipson G. A. (1971) Identification of Distributed Systems, Elsevier, New York.

Polis M. P. and Goodson R. E. (1976) Parameter identification in distributed systems: a synthesizing overview. Proc. I.E.E.E., 64, 45 - 61.

Ray W. H. (1977a) Some applications of state estimation and control theory to distributed parameter systems. Control Theory of Systems Governed by Partial Differential Equations (Edited by Aziz A. K., Wingate J. W. and Balas M. J.). Academic Press, New York, 231 - 263.

Ray W. H. (1977b) Some recent applications of distributed parameter control theory--a survey. 2nd I.F.A.C. Symposium on Control of Distributed Parameter Systems, Coventry.

Ray W. H. and Seinfeld J. H. (1975) Filtering in distributed parameter systems with moving boundaries. Automatica, 11, 509 - 515.

Reddy K. S. and Rajamani V. S. (1975a) Optimum filtering for a class of distributed-parameter systems with coloured measurement noise. Int. J. Systems Science, 6, 615 - 620.

Reddy K. S. and Rajamani V. S. (1975b) Optimal estimation in linear distributed systems with time delays. Int. J. Systems Science, 6, 859 - 863.

Robinson A. C. (1971) A survey of optimal control of distributed-parameter systems. Automatica, 7, 371 - 388.

Sahgal R. and Webb R. P. (1972) Nonlinear distributed parameter estimation. Proc. 1972 I.E.E.E. Conf. Decision and Control, New Orleans, 89 - 93.

Sakawa Y. (1972) Optimal filtering in linear distributed-parameter systems. Int. J. Control, 16, 115 - 127.

Seinfeld J. H., Gavalas G. R. and Hwang M. (1971) Nonlinear filtering in distributed-parameter systems. J. Dynamic Systems, Measurement, and Control, 93, 157 - 163.

Sherry H. and Shen D. W. C. (1973) Combined state and parameter estimation for distributed-parameter systems using discrete observations. Identification and System Parameter Estimation (Edited by Eykhoff P.). North-Holland, Amsterdam, 671 - 677.

Shukla V. and Srinath M. D. (1972) Optimal filtering in linear distributed-parameter systems with multiple time delays. Int. J. Control, 16, 673 - 687.

Stakgold I. (1967) Boundary Value Problems of Mathematical Physics, vol. 1 and 2. Macmillan, New York.

Thau F. E. (1969) On optimum filtering for a class of linear distributed-parameter systems. J. Basic Eng., Series D, 91, 173 - 178.

Tzafestas S. G. (1969) Boundary and volume filtering of linear distributed-parameter systems. Electronic Letters, 5, 199 - 200.

Tzafestas S. G. (1972a) Bayesian approach to distributed-parameter filtering and smoothing. Int. J. Control, 15, 273 - 295.

Tzafestas S. G. (1972b) On optimum distributed-parameter filtering and fixed-interval smoothing for colored noise. I.E.E.E. Trans. Auto. Control, AC-17, 448 - 458.

Tzafestas S. G. (1973) On the distributed parameter least-squares state estimation theory. Int. J. Systems Science, 4, 833 - 858.

Tzafestas S. G. (1976) Nonlinear distributed-parameter filtering using the Fokker-Planck equation approach. J. Franklin Inst., 301, 429 - 450.

Tzafestas S. G. (1978) Distributed parameter state estimation. Distributed Parameter Systems - Identification, Estimation and Control (Edited by Ray W. H. and Lainiotis D. G.) Marcel Dekker, New York, 135 - 208.

Tzafestas S. G. and Nightingale J. M. (1968a) Optimal filtering, smoothing and prediction in linear distributed-parameter systems. Proc. I.E.E., 115, 1207 - 1212.

Tzafestas S. G. and Nightingale J. M. (1968b) Concerning optimal filtering theory of linear distributed-parameter systems. Proc. I.E.E., 115, 1737 - 1742.

Tzafestas S. G. and Nightingale J. M. (1969) Maximum-likelihood approach to the optimal filtering of distributed-parameter systems. Proc. I.E.E., 116, 1085 - 1093.

Tzafestas S. G. and Nightingale J. M. (1976) Stochastic distributed filtering approach to gamma-ray imaging. Int. J. Systems Science, 7, 1249 - 1274.

Yu T. K., Seinfeld J. H. and Ray W. H. (1974) Filtering in nonlinear time delay systems. I.E.E.E. Trans. Auto. Control, AC-19, 324 - 333.

APPENDIX A

Orthogonal Projection Lemma

In this appendix we provide a proof of the orthogonal projection lemma. The lemma is well known and the proof is essentially that of Kalman and Bucy (1961). Because the proof is available elsewhere, we include it here only as a reference-point for the reader. This being our purpose, we do not use the abstract form of the lemma as it is stated in the introduction to Section 2.4. Rather, we consider it as it specifically applies to distributed parameter estimation.

Lemma

$$\int_D\int_D \eta^T(x')E\{[u(x', t) - \hat{u}(x', t)][u(x'', t) - \hat{u}(x'', t)]^T\}\eta(x'')dD_{x'}dD_{x''}$$

$$\leq \int_D\int_D \eta^T(x')E\{[u(x', t) - \omega(x', t)][u(x'', t) - \omega(x'', t)]^T\}\eta(x'')dD_{x'}dD_{x''} \tag{A.1}$$

(i) if and (ii) only if

$$\int_D\int_D \eta^T(x')E\{[u(x', t) - \hat{u}(x', t)][\omega(x'', t)]^T\}\eta(x'')dD_{x'}dD_{x''} = 0 , \tag{A.2}$$

for all η and ω, and (iii) furthermore, if another estimate $\hat{u}_2$ also satisfies (A.2) then

$$\int_D\int_D \eta^T(x')E\{[\hat{u}(x', t) - \hat{u}_2(x', t)][\hat{u}(x'', t) - \hat{u}_2(x'', t)]^T\}\eta(x'')dD_{x'}dD_{x''}$$

$$= 0 . \tag{A.3}$$

Proof

(i) Consider the identity

$$\int_D\int_D \eta^T(x')E\{[u(x', t) - \omega(x', t)][u(x'', t) - \omega(x'', t)]^T\}\eta(x'')dD_{x'}dD_{x''}$$

$$= \int_D\int_D \eta^T(x')E\{[u(x', t) - \hat{u}(x', t)][u(x'', t) - \hat{u}(x'', t)]^T\}\eta(x'')dD_{x'}dD_{x''}$$

$$+2\int_D\int_D \eta^T(x')E\{[u(x', t) - \hat{u}(x', t)][\hat{u}(x'', t) - \omega(x'', t)]^T\}\eta(x'')dD_{x'}dD_{x''}$$

$$+ \int_D\int_D \eta^T(x')E\{[\hat{u}(x', t) - \omega(x', t)][\hat{u}(x'', t) - \omega(x'', t)]^T\}\eta(x'')dD_{x'}dD_{x''} .$$

Note that $\hat{u} - \omega$ is an admissible estimate of u, thus by the hypothesis, Equation (A.2), the middle term vanishes and (A.1) holds.

(ii) Assume (A.2) does not hold. Then there exists an ω_1 such that

$$\int_D\int_D \eta^T(x')E\{[u(x', t) - \hat{u}(x', t)][\omega_1(x'', t)]^T\}\eta(x'')dD_{x'}dD_{x''} = \rho \neq 0.$$

Consider another estimate $\omega_2 = \hat{u} + \sigma\omega_1$. Then

$$\int_D\int_D \eta^T(x')E\{[u(x', t) - \omega_2(x', t)][u(x'', t) - \omega_2(x'', t)]^T\}\eta(x'')dD_{x'}dD_{x''}$$

$$= \int_D\int_D \eta^T(x')E\{[u(x', t) - \hat{u}(x', t)][u(x'', t) - \hat{u}(x'', t)]^T\}\eta(x'')dD_{x'}dD_{x''}$$

$$- 2\sigma\rho + \sigma^2 \int_D\int_D \eta^T(x')E\left\{\left[\omega_1(x', t)\right]\left[\omega_1(x'', t)\right]^T\right\}\eta(x'')dD_{x'}dD_{x''} .$$

For an appropriate σ, the sum of the last two terms will be negative and thus (A.1) will not hold.

(iii) If both $\hat{u}$ and $\hat{u}_2$ satisfy (A.2), then

$$\int_D\int_D \eta^T(x')E\left\{\left[\hat{u}(x', t) - \hat{u}_2(x', t)\right]\left[\omega(x'', t)\right]^T\right\}\eta(x'')dD_{x'}dD_{x''} = 0 .$$

Note that $\hat{u} - \hat{u}_2$ is an admissible estimate of u, thus (A.3) follows. #

APPENDIX B

Certain Lumped Parameter Filters

In the course of developing the linear distributed parameter filters presented in this work, certain lumped parameter filters were also developed. First, we consider discrete time observation processes with continuous systems. The resulting filters are not new results. However, the derivations are presented here because they represent applications of the orthogonal projection lemma technique and as background material to the distributed parameter filters in Chapter 2. Second, we discuss the influence on a lumped parameter filter of a deterministic input and a non-zero mean initial condition. The results are entirely intuitive and, in fact, in the distributed parameter developments we dismiss them as such. We feel that it is of value to present, at least in a simple case, a proof of the intuitive results.

B.1 Introduction

Consider the system governed by the linear ordinary differential equation

$$\frac{du}{dt}(t) = F(t)u(t) + G(t)w(t) + h(t) \tag{B.1}$$

defined for $t > 0$. The state vector, u, is n-dimensional. $w(t)$ is an n-dimensional stochastic input, white in time, h is a known n-dimensional input, and $F(t)$ and $G(t)$ are known nxn matrices. The initial condition for (B.1) is

$$u(0) = u_0 . \tag{B.2}$$

When the stochastic terms are absent, the problem is assumed to be well-posed.

The initial condition is unknown, with its first two moments given,

$$E\{u_0\} = \bar{u}_0$$

$$E\left\{[u_0 - \bar{u}_0][u_0 - \bar{u}_0]^T\right\} = P_0 \tag{B.3}$$

The stochastic input has the properties

$$E\{w(t)\} = 0$$

$$E\left\{w(t)w^T(t')\right\} = Q(t)\delta(t - t') \tag{B.4}$$

where Q is a non-negative definite nxn matrix.

The three types of observations considered are

continuous

$$z(t) = M(t)u(t) + v(t) \tag{B.5}$$

discrete-instantaneous

$$z(t_k) = M(t_k)u(t_k) + v(t_k) \tag{B.6}$$

discrete-average

$$z(t_k) = \int_{t_{k-1}}^{t_k} M(t')u(t')dt' + v(t_k) \; . \tag{B.7}$$

The observation vector, z, is m-dimensional. M is the known mxn observation matrix and v is the m-dimensional observation noise. For continuous observations the properties of v(t) are

$$E\{v(t)\} = 0$$

$$E\left\{v(t)v^T(t')\right\} = R(t)\delta(t - t') \tag{B.8}$$

where R(t) is a known positive definite mxm matrix. Similarly, for discrete time observations the properties of $v(t_k)$ are

$$E\{v(t_k)\} = 0$$

$$E\left\{v(t_k)v^T(t_l)\right\} = R(t_k)\delta_{kl} \tag{B.9}$$

where $R(t_k)$ is a known positive definite mxm matrix. Finally, u_0, w and v are all defined to be independent of each other.

In working with the filters we will use the following notation

$\hat{u}(t\|t')$		Optimal estimate of u(t) based on observations through time t'.
$\hat{u}(t)$	$= \hat{u}(t\|t)$	
$\tilde{u}(t\|t')$	$= u(t) - \hat{u}(t\|t')$	Error in estimating u(t) by the optimal estimate.
$P(t\|t';\ t"\|t"')$	$= E\left\{\tilde{u}(t\|t')\tilde{u}^T(t"\|t"')\right\}$	
$P(t\|t')$	$= P(t\|t';\ t\|t')$	
$P(t)$	$= P(t\|t;\ t\|t)$	Error covariance matrix.
$\omega(t\|t')$		Any admissible estimate of u(t).

η		Arbitrary vector.
Y	$= \langle \eta y \rangle$	$= \eta^T y$
(Y_1, Y_2)	$= E\{\langle \eta, y_1 \rangle \langle \eta, y_2 \rangle\}$	$= \eta^T E\{y_1 y_2^T\}\eta$
$\|\|Y\|\|^2$	$= (Y, Y)$	$= \eta^T E\{yy^T\}\eta$

where $\hat{u}$, $\tilde{u}$, ω, η and y are n-vectors and P is an nxn matrix.

The form of the estimate we seek will be defined in each of the following sections. The criterion for optimality is that the error covariance matrix be minimized. That is, we seek to minimize the scalar function

$$E\left\{\left[\eta^T[u(t) - \omega(t|t)]\right]^2\right\}$$

for all η. Or, as expressed in the inner product notation, we seek $\hat{U}$, such that

$$||\tilde{U}|| \leq ||U - \Omega|| \qquad (B.10)$$

for all Ω. The optimal estimate, $\hat{U}$, is characterized by the orthogonal projection lemma, which is given in Section 2.4.

In Sections B.2 and B.3 we consider discrete-instantaneous and discrete-average observations, respectively. As in the distributed parameter cases we derive these filters under the conditions $h = 0$ and $\bar{u}_0 = 0$. Then, in Section B.4 we discuss the effects of removing these restrictions in the case of continuous observations.

B.2 Discrete-Instantaneous Observations

This problem has been discussed by others, for example, see Jazwinski (1970). Of course, this problem must be considered in two parts; prediction in between observations, and up-date at the time of an observation. The prediction equations were originally derived by Kalman and Bucy (1961) and are given below. We consider the derivation of the up-date equations. The prediction equations are

$$\frac{d\hat{u}}{dt}(t|t_{k-1}) = F(t)\hat{u}(t|t_{k-1}) \tag{B.11}$$

$$\frac{dP}{dt}(t|t_{k-1}) = F(t)P(t|t_{k-1}) + P(t|t_{k-1})F^T(t) + G(t)Q(t)G^T(t) \ , \quad t_{k-1} \le t \le t_k \ . \tag{B.12}$$

The up-dated estimate which we seek is of the form

$$\omega(t_k|t_k) = \sum_{j=1}^{k} B(t_k, t_j)z(t_j) \ . \tag{B.13}$$

Thus to find the optimal ω we must find the optimal B. The optimal B is denoted A, i.e.

$$\hat{u}(t_k) = \sum_{j=1}^{k} A(t_k, t_j)z(t_j).$$

Using difference equations instead of differential equations, the procedure of Kalman and Bucy (1961) for continuous observations is

followed here for discrete-instantaneous observations. From the orthogonal projection lemma, the optimality condition, Equation (2.7), becomes

$$\eta^T \sum_{i=1}^{k} E\left\{\tilde{u}(t_k)z^T(t_k)\right\}B^T(t_k, t_i)\eta = 0 \ , \tag{B.14}$$

for all η and B. The Wiener-Hopf equation (for discrete observations) follow directly. The estimate must obey the conditions

$$E\left\{\tilde{u}(t_k)z^T(t_i)\right\} = 0 \quad , \quad t_i \leq t_k \ . \tag{B.15}$$

Now consider the difference of the Wiener-Hopf equation at times t_k and t_{k-1}

$$E\left\{\tilde{u}(t_k)z^T(t_i)\right\} - E\left\{\tilde{u}(t_{k-1})z^T(t_i)\right\} = 0 \quad , \quad 1 \leq i \leq k-1. \tag{B.16}$$

The state, $u(t_k)$, can be expressed in terms of the system transition matrix, ψ, as follows

$$u(t_k) = \psi(t_k;t_{k-1})u(t_{k-1}) + \int_{t_{k-1}}^{t_k} \psi(t_k; t')G(t')w(t')dt' \tag{B.17}$$

where

$$\frac{d\psi}{dt}(t; t') = F(t)\psi(t; t')$$

$$\psi(t; t) = I \ . \tag{B.18}$$

Using this expression for $u(t_k)$, Equation (B.16) becomes

$$\sum_{j=1}^{k-1} C(t_k;\, t_j) E\left\{z(t_j) z^T(t_i)\right\} = 0 \quad , \quad 1 \le i \le k-1 \tag{B.19}$$

where

$$\begin{aligned} C(t_k;\, t_j) = &\left[\psi(t_k;\, t_{k-1}) A(t_{k-1},\, t_j) - A(t_{k-1},\, t_j)\right] \\ &- A(t_k,\, t_k) M(t_k) \psi(t_k;\, t_{k-1}) A(t_{k-1},\, t_j) \\ &- \left[A(t_k,\, t_j) - A(t_{k-1},\, t_j)\right] . \end{aligned} \tag{B.20}$$

$C(t_k;\, t_j)$ vanishes and we are then able to find a recursive equation for $\hat{u}(t_k)$,

$$\hat{u}(t_k) = \hat{u}(t_k|t_{k-1}) + A(t_k,\, t_k)\left[z(t_k) - M(t_k)\hat{u}(t_k|t_{k-1})\right] \tag{B.21}$$

where $\hat{u}(t_k|t_{k-1})$ is the previously defined optimal prediction of $u(t_k)$.

$A(t_k,\, t_k)$ is determined by using the Wiener-Hopf equation for $i = k$. The up-dated error covariance matrix follows by direct substitution. Together with (B.21), the following equations comprise the up-dated filter

$$\begin{aligned} A(t_k,\, t_k) &= K(t_k) \\ &= P(t_k|t_{k-1}) M^T(t_k)\left[M(t_k) P(t_k|t_{k-1}) M^T(t_k) + R(t_k)\right]^{-1} \end{aligned} \tag{B.22}$$

$$P(t_k) = P(t_k|t_{k-1}) - P(t_k|t_{k-1})M^T(t_k)K^T(t_k) \ . \qquad (B.23)$$

B.3 Discrete-Average Observations

Using the innovations technique, this filter has also been derived by Fujishige (1975). Because the derivation using the orthogonal projection lemma so closely parallels derivations already presented, we will make only a few comments here.

The appropriate Wiener-Hopf equation is (B.15). From this it follows that the difference equation for A is given by

$$\left[\psi(t_k;t_{k-1})A(t_{k-1}, t_j) - A(t_{k-1}, t_j)\right]$$

$$- A(t_k, t_k)\int_{t_{k-1}}^{t_k} M(t')\psi(t';t_{k-1})dt'A(t_{k-1}, t_j)$$

$$- \left[A(t_k, t_j) - A(t_{k-1}, t_j)\right] = 0 \ , \quad 1 \le j \le k-1 \ . \qquad (B.24)$$

The filter equations are

$$\hat{a}(t_k) = a(t_k|t_{k-1}) + K(t_k)\left[z(t_k) - \int_{t_{k-1}}^{t_k} M(t')\hat{a}(t'|t_{k-1})dt'\right] \qquad (B.25)$$

$$P(t_k) = P(t_k|t_{k-1}) - \int_{t_{k-1}}^{t_k} P(t_k|t_{k-1}; t'|t_{k-1})M^T(t')dt'K^T(t_k) \qquad (B.26)$$

$$K(t_k) = \int_{t_{k-1}}^{t_k} P(t_k|t_{k-1};t'|t_{k-1})M^T(t')dt'$$
$$\cdot \left\{ \int_{t_{k-1}}^{t_k} \int_{t_{k-1}}^{t_k} M(t')P(t'|t_{k-1};t''|t_{k-1})M^T(t'')dt''dt' \right.$$
$$\left. + R(t_k) \right\}^{-1} . \qquad \text{(B.27)}$$

along with

$$\frac{dP(t|t_{k-1};t'|t_{k-1})}{dt} = F(t)P(t|t_{k-1};\ t'|t_{k-1}),$$

$$t_{k-1} \leq t' < t \leq t_k \qquad \text{(B.28)}$$

which is determined from the prediction equations.

B.4 Continuous Observations with a Deterministic Input and Non-Zero Mean Initial Condition

In this section we discuss the influence on a lumped parameter filter of a deterministic input and non-zero mean initial condition. We present a proof of the intuitive results. Of course, the results are that the initial condition for the estimate is the mean of the system initial condition and that the deterministic input appears as an additive term in the differential equation for the estimate.

The essential distinction between the derivations previously presented and this one is the form of the estimate. Instead of seeking an estimate which is only a linear combination of the observations; the

estimate in this case also incorporates our *a priori* knowledge of the influence of h and $\bar{u}_0$ on the state. Thus, the optimal estimate is of the form

$$\hat{u}(t) = \int_0^t A(t, t')z(t')dt' + \int_0^t A_1(t, t')h(t')dt' + A_0(t)\bar{u}_0 . \quad (B.29)$$

In addition to requiring that the estimate minimize the error covariance, we also require that it be unbiased (i.e., $E\{\hat{u}(t)\} = E\{u(t)\}$) . Imposing this requirement constrains A_1 and A_0, and the estimate takes the form

$$\hat{u}(t) = \int_0^t A(t, t')z(t')dt' + \int_0^t \left[\psi(t; t') - \int_{t'}^t A(t, t'')M(t'')\psi(t''; t')dt''\right] h(t')dt' + \left[\psi(t; 0) - \int_0^t A(t, t')M(t')\psi(t'; 0)dt'\right] \bar{u}_0 \quad (B.30)$$

where ψ is the system transition matrix. Note that the estimation error can be expressed as

$$\tilde{u}(t) = u(t) - \hat{u}(t) = \left[\psi(t; 0)(u(0) - \bar{u}_0) + \int_0^t \psi(t; t')G(t')w(t')dt'\right] - \int_0^t A(t; t')\left[M(t')\left[\psi(t'; 0)(u(0) - \bar{u}_0) + \int_0^{t'} \psi(t'; t'')G(t'')w(t'')dt'' + v(t')\right]\right] dt'. \quad (B.31)$$

Now consider the state u_* defined by

$$\frac{du_*}{dt}(t) = F(t)u_*(t) + G(t)w(t) \tag{B.32}$$

where the initial condition has the properties

$$\begin{aligned} E\{u_*(0)\} &= 0 \\ E\{u_*(0)u_*^T(0)\} &= P_0 \end{aligned} \tag{B.33}$$

and F, G, w, and P_0 are as previously defined. Consider also the observation

$$z_*(t) = M(t)u_*(t) + v(t) \tag{B.34}$$

and the estimate

$$\hat{u}_*(t) = \int_0^t A_*(t, t')z_*(t')dt' \ . \tag{B.35}$$

Note that

$$u_*(t) = \psi(t; 0)(u(0) - \bar{u}_0) + \int_0^t \psi(t, t')G(t')w(t')dt' \ , \tag{B.36}$$

and that the estimation error for u_* is the same as that for u, with A replaced by A_*. Thus, $A(t, t') = A_*(t, t')$, and it follows that

$$\hat{u}(t) = \hat{u}_*(t) + \int_0^t \psi(t, t')\, h(t')dt' + \psi(t, 0)\bar{u}_0 \ . \tag{B.37}$$

We know $\hat{u}_*(t)$, thus the filter follows directly,

$$\frac{d\hat{u}}{dt}(t) = F(t)\hat{u}(t) + K(t)\left[z(t) - M(t)\hat{u}(t)\right] + h(t)$$

$$\hat{u}(0) = \bar{u}_0 \tag{B.38}$$

$$\frac{dP}{dt}(t) = F(t)P(t) + P(t)F^T(t) + G(t)Q(t)G^T(t) - K(t)R(t)K^T(t)$$

$$P(0) = P_0 \tag{B.39}$$

$$K(t) = P(t)M^T(t)R^{-1}(t) \ .$$

PART II. STATISTICAL ANALYSIS OF AIR POLLUTANT OBSERVATIONS AND MODEL PREDICTIONS

Chapter 1

Introduction

Efforts to analyze urban air quality are often beset by the seemingly contradictory problems of a simultaneous paucity and abundance of pertinent information. The required meteorological, emissions and ambient concentration data reach virtually unmanageable proportions. But more data, of even finer spatial and temporal resolution, are still needed to answer unsolved questions. To deal effectively with both of these problems, techniques that are essentially statistical must be introduced. Statistical techniques allow large volumes of information to be summarized compactly by identifying basic characteristics of the data. Then, knowledge of these basic characteristics allows the use of relatively small amounts of further information to interpret future or related data sets.

Observations and model predictions serve complementary purposes in engineering. The observations form the ultimate base on which models are built, and the models help in understanding the principles that lead to the observations. In Chapter 2 the role of observations in assessing model accuracy is discussed. Here, great amounts of information have to be summarized by relatively small number of measures of performance. The need is to be able to make a decision as to the adequacy of a given model. In Chapter 3 observed frequency distributions are investigated through the use of highly simplified models. Here, the underlying statistical nature of air quality concentrations is explored.

Chapter 2

Evaluation of Air Quality Models: The Accuracy of Predictions Relative to Observations

1. INTRODUCTION

The development and application of air quality simulation models are areas of fundamental interest in the effort to control air quality in urban airsheds. Air quality modeling has reached the stage of maturity at which a large variety of increasingly sophisticated models are being developed and proposed for applications. Thus there exists a need to concurrently develop a framework for the evaluation of these models. Then, within such a framework, consistent and general schemes need to be developed to perform the various model evaluation tasks.

Model evaluation is a problem of long-standing concern, and nearly every discussion of a given model's development or application contains comments relative to some aspect of that model's performance. In Table 1 representative references to prior air quality model evaluation studies are listed. A clear need exists for a comprehensive discussion of air quality model evaluation along with the development of general evaluation methods.

The research presented here is directed at two aspects of this proble First, in the remainder of this section, we develop a framework for the evaluation of the performance of air quality models. Second, in Section 2 we present quantitative methods for assessing the accuracy of model predictions relative to physical observations. The final result is a package of accuracy assessment methods that when exercised will provide a meaningful evaluation of this aspect of the performance of an air quality model. This package, AQMAAP, has been coded in FORTRAN. A discussion of the code is given in the Appendix.

Table 1

Air Quality Model Evaluation

Reference	Evaluations Discussed
Nappo (1974)	Comparison of nine models. Measures of correlation and bias (based on ratios).
Brier (1975)	Statistically oriented discussion. Measures of bias, spread and correlation. Regression analysis.
Johnson et al. (1976)	Comparison of three models. Measures of correlation and RMS error. Regression analysis.
Liu et al. (1976)	Comparison of three models. Measures of correlation, RMS error, and goodness-of-fit. Plots of scatter and residuals.
Anderson et al. (1977)	Evaluation of a single model. Graphic presentations of time series, frequency distribution, correlation and bias.
Maldonado and Bullin (1977)	Comparison of three models. Measures of bias and spread. Regression analysis.

1.1 Evaluation Framework

The evaluation of an air quality model generally focuses on the accuracy of the model's predictions relative to observations. This is the approach taken in this work, and it is a logical approach because it attempts to answer the 'bottom line' question, "How well does the model predict actual concentrations?" or "Does it work?" However, it is often overlooked that this approach represents only a part of the total model evaluation problem. Two other major issues must be addressed: the first is validity and the second is efficiency.

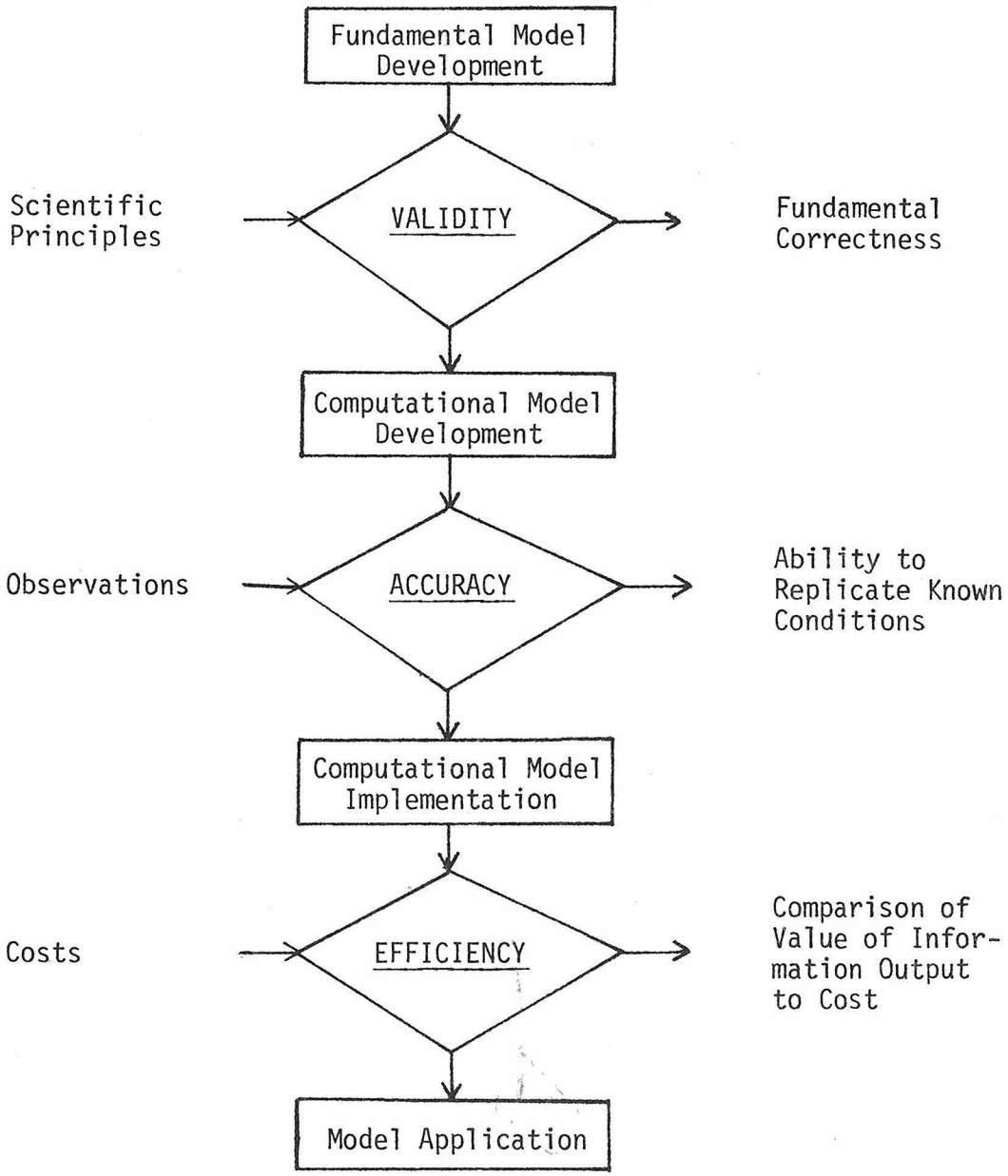
Before focusing on model accuracy in the next section, we will propose an overall evaluation framework, the intent of which is to provide a perspective for accuracy assessment. In Figure 1 we outline the model evaluation problem from the stage of fundamental model development through to application.

Validity refers to the fundamental correctness of the model formulation. The assessment of model validity is based on scientific principles, i.e. the relationships or equations which comprise the model. The first issue is the adequacy of the representation of the physics and the chemistry of the system. When the system is complex, it is difficult to index quantitatively the validity of the entire model. Rather, it may be necessary to assess individually the correctness of the representation of various mechanisms which together influence the system. This separation virtually necessitates that the overall evaluation will be qualitative.

Whereas in assessing model validity the emphasis is placed on the segments that comprise the model, in assessing model accuracy the emphasis

Figure 1

Evaluation of Urban Scale, Photochemical Air Quality Models



shifts to the model as a complete unit, with a shift from evaluation of the beginning of the modeling process, scientific principles, to the end, actual output. Accuracy assessment measures the ability of the model to replicate known conditions. The main limitation to a complete study of accuracy is the availability of data to set the standard against which the model is to be judged. In Section 2 we discuss in detail the quantitative methods for accuracy assessment. However, even here, it is difficult to arrive at a single quantitative index of model accuarcy. We will always be left to form a qualitative summation of the results of the various methods.

As air quality models receive increasing attention in policy decision applications, the interest in their efficiency also increases. Because of limited resources it is usually necessary to compare the value of a model's information output to the costs associated with running the model.* Efficiency may at first appear to be a quantitative measure for model evaluation; the qualitative aspect enters into the 'value of the information output' part. A model does not necessarily have twice the information content if it has four vertical grid layers instead of two. Any measure of efficiency ultimately reflects an opinion in defining the important outputs and the constraining costs.

In concluding this discussion of the evaluation framework, we have

---

*We define cost in a broad sense that includes time and resources in addition to direct dollar outlays. Thus, the obvious cost is in computer expenses, but if three or four months are required to acquire and assemble the input data base, then this is also a significant cost. Of course, we could assign a dollar value to time and resources and then use a single unit of cost. Such an approach becomes artificial, however, when we are constrained by time or computer equipment limitations.

two general comments. The first relates to the process of evaluation and the second to the results of evaluation.

Air quality model evaluation has been presented in this discussion as a serial process, progressing through stages from development to application. As a conceptualization, this allows us to identify and isolate the important aspects of model evaluation. In an ongoing model development effort, the evaluation process will not follow this form explicitly. The various stages will often occur in parallel and with considerable feedback. In practice, all of the stages are interdependent. The framework we have presented must be used as a framework and not a flowsheet.

The result of a model evaluation is not simply a single index of total performance. It is clear that air quality models are too complex for this. In the final analysis, subjective judgements, based on knowledge of the physical problem being studied, are necessary.

2. ACCURACY ASSESSMENT

2.1 Introduction

As defined in Section 1, accuracy refers to the ability of a model to replicate known conditions. In some situations it may be possible, due to characteristics of the process or of the model, to specify a single index of accuracy; this is decidely not the case for air quality models. With air quality models there is typically a very large number of predictions and observations. Thus, a given model is generally not uniformly 'good' or 'bad' in all of its predictions. Rather than insisting on a one-to-one comparison of predictions and observations, it is important to identify the characteristics of a 'good' model and to assess model accuracy on the basis of these characteristics. In Section 2.2 we will identify a variety of methods which, collectively, form the basis for assessing the accuracy of air quality models.

In this work we discuss the direct comparison of model predictions to physical observations. Accuracy assessment need not be solely based on such comparisons. Comparisons can be made relative to input parameters. For example, in a study simulating many days, an index of direct prediction to observation comparison can be analyzed as a function of daily emissions. A model may not necessarily perform as well at one level of emissions as another, and this type of analysis is sensitive to such behavior. Mass fluxes also reflect very basic characteristics of the air quality in an airshed. Observed and predicted mass fluxes can be compared by forming the new variable, mass flux, based on the input parameters, emissions and transport variables.

Direct comparisons of model predictions with physical observations are the basic accuracy assessment methods and thus we focus on them. The methods involving the input parameters are important but are not as general nor as objective. Computer codes for the direct comparison can be written independently of detailed knowledge of the model to be evaluated. Codes involving the input parameters tend to be more model specific because different models, even of the same physical process, will require different details in their respective inputs. Also, the farther removed from the raw data, the more subjective the judgments that may have to be built into the analysis become. If these judgments involve issues similar to those encountered in the model development itself, then the risk exists that the model evaluation is biased by our judgments. An example is in the determination of the wind field from the data at a few observation locations. Different models treat this problem differently and so would different evaluation schemes.

As mentioned earlier, air quality models are sufficiently complex that accuracy must be assessed by a variety of methods. Before presenting specific methods, we discuss the two levels at which predictions can be directly compared to observations. At the second level the evaluations become more meaningful because we are adding to the assessment more implicit information about the nature of the data.

Level 1 assessment does not employ any information concerning the process being modeled and considers the observations and predictions to be two sets of data which should inherently be equivalent. In essence, at this level we measure the degree to which the two sets are in fact equivalent. Level 1 evaluations are generally meaningful to perform on

models of any process and they represent the quality of the model's fit to the observations on an average or overall basis. These two characteristics are, respectively, the strength and the weakness on Level 1 methods.

Level 2 methods are similar to those of Level 1 in that it is assumed that the two sets of data should inherently be equivalent; however, they do account for the particular process being modeled and the use to which the model results will be put. These considerations lead to methods that consider the aspects of the data which indicate whether or not the model will be useful for its intended application. Thus, in contrast to Level 1 methods, these methods are process specific and only measure the model's performance based on specific aspects of the data. In the broadest sense, air quality models may be used either for (1) evaluating current or proposed emission control strategies, or (2) increasing our knowledge in the science of atmospheric pollution. In this discussion we will be concerned with only the first use of air quality models. It is for this use that comparisons to physical observations are most relevant. Various indices of air quality are used to evaluate the effect of control strategies. Thus, a useful model must produce similar values to the observations for the air quality indices. The Level 2 methods of comparison measure this characteristic of the model.

It is extremely important to realize that there are three fundamental difficulties in comparing air quality observations to predictions of the type we are considering. First, on the scale of the model, the observations are spatially pointwise, whereas the predictions represent volume averages. Second, the observations contain instrument errors. Errors in the model input parameters present the final problem. Even if the model is an ideal

formulation of the air pollution process, the predictions will be in error if the inputs are in error. In evaluating the results of an accuracy assessment study these factors must be remembered. The presence of these factors is one reason why many diverse methods are needed for a meaningful assessment.

Next, we discuss methods for Level 1 and Level 2 assessment.

2.2 Assessment Levels 1 and 2

The list of possible methods for assessing model accuracy is indeed lengthy, the fundamental reason for which is simply the sheer number of data values that must be dealt with. (A relatively modest evaluation effort might involve hundreds of values, and in a comprehensive regional model comparison study the number may reach into the millions.) Thus, the number of potentially meaningful ways to organize and reorganize the information for inspection is definitely substantial. The data could be viewed temporally, spatially, spectrally, or relative to frequency distributions. Superposed on these classes of organization are the many possible forms of averaging.

In Table 2 we catalog many of the accuracy assessment methods that have been purposed for comparing air quality model predictions to physical observations. No claim is made that this catalog is necessarily complete; rather, it is representative and includes the important concepts.

From this catalog a general accuracy assessment package, AQMAAP, has been developed. We now discuss the methods that comprise AQMAAP.

Table 2

Accuracy Assessment Methods

Method	Comments	Reference
Level 1		
Analysis of Residuals		
Mean Error	Included in AQMAAP	Anderson et al. (1977) Maldonado and Bullin (1977)
RMS Error Centered at 0	Quadratic Measure of Mean Error	Liu et al. (1976) Johnson et al. (1976) Maldonado and Bullin (1977)
RMS Error Centered at Mean Error	Included in AQMAAP	—
Absolute Error Band	Included in AQMAAP	Maldonado and Bullin (1977)
Percentage Error Band	Included in AQMAAP	—
Residual Histogram	Included in AQMAAP	Liu et al. (1976)
Residual vs. Time Plot	Included in AQMAAP	Liu et al. (1976)
Residual vs. Observed Magnitude Plot	Included in AQMAAP	Liu et al. (1976)

Table 2 (continued)

Method	Comments	Reference
Residual vs. Predicted Magnitude Plot	Yields similar information to the residual vs. observed magnitude plot.	Liu et al. (1976)
Residual vs. Location Plot	Included in AQMAAP	—
Analysis of Trends		
Correlation Coefficient	Included in AQMAAP	Nappo (1974) Liu et al. (1976) Johnson et al. (1976)
Regression Analysis	Included in AQMAAP	Johnson et al. (1976) Maldonado and Bullin (1977)
Scatter Plot	Graphic presentation of information found from correlation coefficient and from regression analyses.	Liu et al. (1976)
Numerical Correlogram	(Same as comment above.)	Anderson et al. (1977)
Other Analyses		
Chi-Square Test	Substantial uncertainty in assuming a frequency distribution for the errors.	Liu et al. (1976)

Table 2 (continued)

Method	Comments	Reference
Mean and Variance of Ratio of Predictions to Observations	Yields similar information to methods for analysis of residuals.	Nappo (1974)
Fractional Mean Deviation from Perfect Correlation	Yields information similar to methods for analysis of residuals.	Anderson et al. (1977)
Spectral Analysis	Requires data over periods long enough to have many cycles. A model accuracy study will typically have data for only a few days.	—
Level 2		
Analysis of Indices of Air Quality		
Magnitude and Time of Concentration Peaks	Included in AQMAAP	—
Location of Concentration Hotspots	Included in AQMAAP	—
Exposure	Included in AQMAAP	—
Geometric Parameters of Highest Values	Included in AQMAAP	—
Frequency Distribution	Included in AQMAAP	—

The basic nomenclature for the concentration values is

$c_i(x_j,d_k,t_l)$: Observed concentration of specie i, at location x_j, on day d_k and at time t_l.

$\hat{c}_i(x_j,d_k,t_l)$: Predicted concentration.

M_i: Number of species.

M_j: Number of locations.

M_k: Number of days.

M_l: Number of hours of interest during each day.

Many of the methods will represent averaged properties of the data. Because there are three independent indices (location, day, hour) there are many ways to treat the data in determining averaged properties. Five such ways appear particularly useful for air quality data. It is important to consider the data in this many ways to first have a general measure of the model performance and then to look at more specific measures. In this manner, systematic or important isolated deficiencies in the model may be uncovered. Most of the Level 1 methods in AQMAAP will treat the data in each of the five ways. To avoid needless repetition we use a notation that allows each of the five to be denoted by one set of equations. Table 3 lists the five averaging processes along with the interpretations of the shorthand notation. The meaning of the notation will become clear once we have discussed the measures of accuracy assessment.

The Level 1 methods are classified into two types: analysis of residual and analysis of trends. In the first, we examine the differences between the observations and predictions. Ideally, the residuals would all be identically zero. In the second, we conceptually order the observations from smallest to largest, then examine the degree to which the

Table 3

Averaging Shorthand Notation

(A) Measure Notation: $\zeta_i(n) = f\left[\frac{1}{M}\sum_{m=1}^{M} g(m,n)\right]$

$\zeta_i(n)$	Measure ζ for species i and set of indices n.
f	Function of the averaged function g.
$\frac{1}{M}\sum_{m=1}^{M}$	Averaging operator.
g	Function of the sets of indices m and n.
m	Set of indices to be averaged over.
n	Set of indices ζ_i is a function of.

Table 3 (continued)

(B) Averaging Notation

	Indices Averaged (Set 'm')	Symbol of Measure	$\frac{1}{M}\sum_{m=1}^{M}$	Set 'n'
Measure over all time and locations.	Location, day, hour	ζ_i^{TL}	$\frac{1}{M_j}\sum_{j=1}^{M_j}\frac{1}{M_k}\sum_{k=1}^{M_k}\frac{1}{M_l}\sum_{l=1}^{M_l}$	–
Measure over all times.	Day, hour	ζ_i^{T}	$\frac{1}{M_k}\sum_{k=1}^{M_k}\frac{1}{M_l}\sum_{l=1}^{M_l}$	$j = 1,\ldots,M_j$
Measure over each day.	Hour	ζ_i^{H}	$\frac{1}{M_l}\sum_{l=1}^{M_l}$	$j = 1,\ldots,M_j; k = 1,\ldots,M_k$
Measure over each hour.	Day	ζ_i^{D}	$\frac{1}{M_k}\sum_{k=1}^{M_k}$	$j = 1,\ldots,M_j; l = 1,\ldots,M_l$
Measure over all locations.	Location	ζ_i^{L}	$\frac{1}{M_j}\sum_{j=1}^{M_j}$	$k = 1,\ldots,M_k; l = 1,\ldots,M_l$

corresponding predictions follow the identical ordering.

(A) Analysis of Residuals

Denote the residual as w_i where

$$w_i(x_j,d_k,t_l) = c_i(x_j,d_k,t_l) - \hat{c}_i(x_j,d_k,t_l).$$

Mean Error

Mean error measures the average bias in the predictions and indicates whether the model predominantly over or under predicts. The general form of the mean error is

$$\mu_i(n) = \frac{1}{M}\sum_{m=1}^{M} w_i(m,n).^{*}$$

Root Mean Square Error Centered at the Mean

This RMS function measures the average spread of the residuals. It is insensitive to the bias. The general form is defined

$$\sigma_i(n) = \left[\frac{1}{M}\sum_{m=1}^{M}\left(w_i(m,n) - \mu_i(n)\right)^2\right]^{\frac{1}{2}}.^{**}$$

---

*As an example of the shorthand notation, g is the residual w_i and there is no function f. The mean error over all times is

$$\mu_i^T(x_j) = \frac{1}{M_k}\sum_{k=1}^{M_k}\frac{1}{M_l}\sum_{l=1}^{M_l} w_i(x_j,d_k,t_l)$$

and the mean error over all locations is

$$\mu_i^L(d_k,t_l) = \frac{1}{M_j}\sum_{j=1}^{M_j} w_i(x_j,d_k,t_l) .$$

**In this equation, f is the square root and g is the squared difference between the residual and the corresponding mean error. Thus,

$$\sigma_i^{TL} = \left[\frac{1}{M_j}\sum_{j=1}^{M_j}\frac{1}{M_k}\sum_{k=1}^{M_k}\frac{1}{M_l}\sum_{l=1}^{M_l}\left(w_i(x_j,d_k,t_l) - \sigma_i^{TL}\right)^2\right]^{\frac{1}{2}}.$$

Error Bands

In contrast to the above averaged measure of the spread, we can evaluate the model relative to a prescribed degree of spread by finding the percentage of the predictions which fall within a given band around the observations. This band may be based on an absolute concentration, $\pm y_i$ ppm, or a percentage, $\pm z_i\%$ of the observed value. The absolute error band measure is $\alpha_i(n)$ = the percentage of residuals from set 'm' within $\pm y_i$ ppm. The percentage error band measure is $\beta_i(n)$ = the percentage of residuals from set 'm' within $\pm z_i\%$ of $c_i(m,n)$.

Residual Plots

Residual plots provide a summary picture of the nature of the bias and the spread for the model. Five types of plots are of specific value for air quality models. Two residual histograms are included in AQMAAP: a histogram based on all of the residuals for a given specie, and individual histograms for each location. By plotting the residuals in other ways we can determine if there are certain conditions under which the model is good, while others under which it is poor. Thus, we also plot the residuals versus time of day, versus location and versus observation magnitude.

(B) Analysis of Trends

Methods for analyzing residuals primarily indicate the absolute fit of the predictions to the observations. Predictions based on fundamentally wrong relationships can have measures of the residuals that appear similar to those for soundly based predictions. By analyzing the

trends of the predictions relative to the observations we gain further information indicating whether or not the predictions at least obey the same relationships as the observations.

Correlation Coefficient

The correlation coefficient measures the degree to which the magnitude of the predictions increases linearly with the magnitude of the observations. The coefficient, however, is insensitive to the extent of the increase. If the predictions increase linearly at only 1/10 the rate of the observations the correlation coefficient will still be one.

The realtionships defining the correlation coefficient are

$$\rho_i(n) = \frac{\frac{1}{M}\sum_{m=1}^{M}\left[\nu_i(m,n)\eta_i(m,n)\right]}{\left[\frac{1}{M}\sum_{m=1}^{M}\left(\nu_i(m,n)\right)^2 \frac{1}{M}\sum_{m=1}^{M}\left(\eta_i(m,n)\right)^2\right]^{\frac{1}{2}}}$$

$$\nu_i(m,n) = \hat{c}_i(m,n) - \bar{\hat{c}}_i(n) \qquad \bar{\hat{c}}_i(n) = \frac{1}{M}\sum_{m=1}^{M}\hat{c}_i(m,n)$$

$$\eta_i(m,n) = c_i(m,n) - \bar{c}_i(n) \qquad \bar{c}_i(n) = \frac{1}{M}\sum_{m=1}^{M}c_i(m,n).$$

Linear Least-Squares Curve Fit

In addition to the correlation coefficient, we can also measure the average increase in the observations as the observations increase. The slope parameter of the linear least-squares curve fit is this measure. The intercept parameter measures the bias if the slope parameter is very nearly one. Otherwise, the slope so strongly influences the intercept

that it has no particular independent interpretation.

We seek to minimize

$$\frac{1}{M} \sum_{m=1}^{M} \left[\xi_i(m,n) - c_i(m,n)\right]^2$$

when ξ_i is defined

$$\xi_i(m,n) = \theta_i(n)\hat{c}_i(m,n) + \phi_i(n).$$

Thus, the slope and intercept are

$$\theta_i(n) = \frac{\frac{1}{M} \sum_{m=1}^{M} [\nu_i(m,n)\eta_i(m,n)]}{\frac{1}{M} \sum_{m=1}^{M} \left(\nu_i(m,n)\right)^2}$$

and

$$\phi_i(n) = \bar{c}_i(n) - \theta_i(n)\bar{\hat{c}}_i(n) ,$$

respectively.

Aside from the plots of residuals the above methods can all be applied in each of the five manners of averaging in Table 3. Of course, we can now further average the quantitative measures themselves. Particularly popular are spatial averages of the temporal averaged measures and the converse. The value of such repeated averaging is, however, dubious because physical interpretations of the measures become tenuous and too much averaging tends to obscure the very deficiencies one is seeking to identify.

The Level 2 methods assess a model relative to its potential use for evaluating control strategies. These methods measure the ability

of the model to produce the same values as the observations for air quality indices.

Analysis of Indices of Air Quality

The most common indices of air quality relate to maximum values. We consider spatial and temporal maximum values.

Concentration Hotspots

On an hourly basis we determine the locations of the observed and predicted hotspots. Then we report γ_i, the percentage of times the model accurately predicted the location of the hotspot of specie i.

Concentration Peaks

At each location we determine both the magnitude and the time concentration peaks. The predictions may be basically accurate near the peak but still fail to reproduce exactly the peak value itself. Instead of basing our judgment on only one value we also characterize the group of highest values for a specie. Because air quality data may span more than one order of magnitude, geometric parameters are appropriate. We compare the geometric mean and standard geometric deviation of the upper ten percent of the data. The peak measures are

magnitude error,

$$\varepsilon_i(x_j,d_k) = \left\{ \max_{t_l} \left[c_i(x_j,d_k,t_l) \right] - \max_{t_l} \left[\hat{c}_i(x_j,d_k,t_l) \right] \right\}$$

time error,

$$\tau_i(x_j,d_k) = \left\{ \left[\text{time of c max} \right] - \left[\text{time of } \hat{c} \text{ max} \right] \right\}$$

geometric mean,

$$\lambda_i(x_j) = \exp\left\{\frac{1}{M_{10}} \sum_{m=1}^{M_{10}} \ln c_i(x_j,m)\right\}$$

$$\hat{\lambda}_i(x_j) = \exp\left\{\frac{1}{M_{10}} \sum_{m=1}^{M_{10}} \ln \hat{c}_i(x_j,m)\right\}$$

and standard geometric deviation,

$$\psi_i(x_j) = \exp\left\{\left[\frac{1}{M_{10}} \sum_{m=1}^{M_{10}} \left(\ln c_i(x_j,m) - \ln\lambda_i(x_j)\right)^2\right]^{\frac{1}{2}}\right\}$$

$$\hat{\psi}_i(x_j) = \exp\left\{\left[\frac{1}{M_{10}} \sum_{m=1}^{M_{10}} \left(\ln \hat{c}_i(x_j,m) - \ln\hat{\lambda}_i(x_j)\right)^2\right]^{\frac{1}{2}}\right\}$$

where $M_{10} = (M_k M_1)/10$

$c_i(x_j,m)$ = m^{th} largest value; $m = 1, \ldots, M_{10}$.

Frequency distribution

Another standard manner of indexing air quality data aside from maximum values is relative to their frequency distributions. It is common, for example, to seek the concentration that occurs at the 99th, 90th or 50th percentile and, conversely, to seek the percentage of time that certain air quality levels are exceeded. All of this information is compactly summarized on a frequency distribution plot. These plots are drawn for each location.

Exposure

The final index of air quality that we will compare is a simple

measure of exposure. We compare the ppm-hour exposure of a receptor at each of the observation locations.

$$\delta_i(x_j) = \frac{1}{M_k} \sum_{k=1}^{M_k} \frac{1}{M_l} \sum_{l=1}^{M_l} c_i(x_j,d_k,t_l)\cdot(t_l - t_{l-1})$$

$$\hat{\delta}_i(x_j) = \frac{1}{M_k} \sum_{k=1}^{M_k} \frac{1}{M_l} \sum_{l=1}^{M_l} \hat{c}_i(x_j,d_k,t_l)\cdot(t_l - t_{l-1}).$$

3. SUMMARY

The final result of this work is the development of the FORTRAN code AQMAAP. AQMAAP represents a general and extensive package for assessing the accuracy of air quality models. The discussion of the total model evaluation framework places AQMAAP in the proper perspective for interpretation of its output.

The methods presented here are both meaningful and practical for studies ranging from small one model validations with a few days data to large multiple model comparisons with possibly a few months data. The package is also quite general and objective. Essentially, it requires only the observed and predicted concentration values for inputs. No information on the model or details of the field study need to be provided. For these reasons, any user can assess any air quality model without extensive alterations and with a minimum of subjective judgments.

REFERENCES

Anderson G. E., Hayes S. R., Hillyer M. J., Killus J. P. and Mundkur P. V. (1977) Air Quality in the Denver Metropolitan Region 1974-2000, Environmental Protection Agency Report EPA 908/1-77-002.

Brier G. W. (1975) Statistical Questions Relating to the Validation of Air Quality Simulation Models, Environmental Protection Agency Report EPA 650/4-75-010.

Johnson W. B., Sklarew R. C. and Turner D. B. (1976) Urban air quality simulation modeling. Air Pollution, Volume I (Edited by Stern A.C.). Academic Press, New York, 503-562.

Liu M. K., Whitney D. C., Seinfeld J. H. and Roth P. M. (1976) Continued Research in Mesoscale Air Pollution Simulation Modeling: Analysis of Prior Evaluation Studies, Environmental Protection Agency Report EPA 600/4-76-016a.

Maldonado C. and Bullin J. A. (1977) Modeling carbon monoxide dispersion from roadways. Environ. Sci. Technol. 11, 1071-1076.

Nappo C. J., Jr. (1974) A method for evaluating the accuracy of air pollution prediction models. Preprints of the Symposium on Atmospheric Diffusion and Pollution, Santa Barbara, CA, American Meteorological Society, 325-329.

APPENDIX. AQMAAP, Air Quality Model Accuracy Assessment Package

A.1 Description

AQMAAP is a FORTRAN coded package of methods for the assessment of the accuracy of air quality models. The scope and attributes of the principles behind AQMAAP have been discussed in the preceding sections of this chapter. This appendix describes the FORTRAN package itself.

AQMAAP consists of 41 subroutines comprising the individual methods, 15 supporting subroutines, and 7 main or driver routines. In Table A.1, the assessment methods and the corresponding subroutines are listed.

The package is running on the IBM 370 at the California Institute of Technology Computing Center. Aside from bookkeeping information (i.e., identification of the data sets and specification of the methods to be exercised), the only inputs are the observation and prediction data sets. A few statements (mainly a single group of DIMENSION statements) in the code are dependent on the size of the data sets. In the following section the necessary information for operating AQMAAP at any installation for any size data set is discussed.

Table A.1

AQMAAP

Assessment Method Subroutines

Assessment Method	Subroutine Name
LEVEL 1	
Mean Error	
Over all times and locations	E11111
Over all times	E11011
Over each day	E12011
Over each hour	E13011
Over all locations	E10111
Root Mean Square Error Centered at the Mean	
Over all times and locations	E11112
Over all times	E11012
Over each day	E12012
Over each hour	E13012
Over all locations	E10112
Absolute Error Bands	
Over all times and locations	E11121
Over all times	E11021
Over each day	E12021
Over each hour	E13021
Over all locations	E10121
Percentage Error Bands	
Over all times and locations	E11122
Over all times	E11022
Over each day	E12022
Over each hour	E13022
Over all locations	E10122
Residual Histogram	
Over all times and locations	E11131
Over all times	E11031

Table A.1 (continued)

Assessment Method	Subroutine Name
Residual Plots	
Residuals versus observation magnitude	E11133
Residuals versus location	E11134
Residuals versus time	E11135
Correlation Coefficient	
Over all times and locations	E11141
Over all times	E11041
Over each day	E12041
Over each hour	E13041
Over all locations	E10141
Linear Least Squares	
Over all times and locations	E11142
Over all times	E11042
Over each day	E12042
Over each hour	E13042
Over all locations	E10142
LEVEL 2	
Magnitude of Daily Concentration Peaks	E22051
Time of Daily Concentration Peaks	E22052
Location of Concentration Hotspots	E20156
Peak Geometric Parameters	E21062
Daily Exposure	E22071
Frequency Distribution Plot	
Over all times	E21061

A.2 Operational Aspects

Tables A.2 - A.5 detail the operation of AQMAAP. Tables A.6 and A.7 provide additional information that may be pertinent for operating AQMAAP at another installation.

The operation of AQMAAP requires that a few parameters be set by the user. Most of these relate to the size of the data set. The remaining parameters give the user control over the error band criteria and some flexibility in the presentation of the plots. In Table A.2 the program parameters are listed.

The three inputs to AQMAAP are, first, the identification of the data sets, then, the data sets themselves and finally, the specification of the methods to be exercised. The proper input sequence and FORTRAN formats are listed in Table A.3. Five averaging processes are used in AQMAAP. As discussed in Table 2.2, a different set of indices is averaged for each averaging process. The corresponding measure of accuracy is then computed as a function of the remaining indices (including the specie index). These indices are specified in Input section (3). Any particular method need only be exercised over a selected portion of the data set. For example, the mean error evaluation over all times might only be computed at a few locations instead of at all locations. Table A.4 elaborates on these input codes. A sample input deck is shown in Table A.5.

Table A.6 indicates which of the 15 supporting subroutines are used in each of the main routines. Eleven additional routines available at the California Institute of Technology Computing Center (CITCC) are used in AQMAAP. These routines are referenced and cataloged in Table A.7.

Table A.2

AQMAAP

Program Parameters

Data set dimensions

(1) In the BLOCK DATA section of the main routines, specify the values and dimension in the statements

INTEGER IPARM/ M_i /, JPARM/ M_j /, KPARM/ M_k /, LPARM/ M_l/

REAL COBS(M_i,M_j,M_k,M_l)/M_a*-1.0/, CPRD(M_i,M_j,M_k,M_l)/M_a*-1.0/

where M_i is the number of species (in the data set), M_j is the number of locations, M_k is the number of days, M_l is the number of hours and $M_a = M_i M_j M_k M_l$.

(2) In the main routines, the method subroutines and support subroutines INCOD3, INDAT1, OTDAT3, RESID and OTRES3 specify the dimensions in the statements

COMMON /DATA/ COBS(M_i,M_j,M_k,M_l), CPRD(M_i,M_j,M_k,M_l)

COMMON /WRKSPC/ WS1(M_b,M_j,M_k,M_l), WS2(M_b,M_j,M_k,M_l)

COMMON /CODE/ NNSPC(M_i), NSSPC(M_i), NSPC(7,M_i), NNLOC(2,M_j), NLOC(7,M_i), NALOC(2,M_j), NDAY(7,M_k)

where M_b is the greater of M_i and 3.

(3) In the subroutines E11131, E11031, E11133, E11134, E11135 and E21061, specify the dimensions in the statements

DIMENSION WS2R(M_l,M_k,M_j,M_i)

DIMENSION $\left\{\begin{matrix}\text{NR}\\ \text{FREQ}\\ \text{RES}\end{matrix}\right\}$ (M_c), X(M_c), Y(M_c)

where $M_c = M_j M_k M_l$ and NR is the variable used in E11133, FREQ in E21061 and RES in the remaining four subroutines.

Table A.2 (continued)

(4) In E11134 (residual versus location plot) and E11135 (residual versus time plot), the length of the abscissas equal the number of locations and days, respectively. These subroutines are currently written to plot on a single sheet of paper. If there are more than 55 locations or 3 days, it will be necessary to make alterations to either the plotting format or the available plotting area. Such changes will require additional familiarity with both the plotter being used and with the coding of these two subroutines.

Error band size

(1) In EB, Exxx21, Exxx22 and SETEB, specify the dimensions in the statement COMMON /ERRBND/ EBA(M_i), EBP(M_i).

(2) In SETEB, specify the values in the statements

EBA(i) = y_i, i=1, ..., M_i

EBP(i) = z_i, i=1, ..., M_i

where y_i and z_i are, for each specie, the user selected sizes of the absolute and percentage error bands, respectively.

Plotting symbol specification

In E11133, E11134 and E11135, specify the values in the statements

MSB(m) = b(m), m = 2,3,4

MSE(m) = e(m), m = 1,2,3

with values between 1 and 99. Four symbols are available to these plotting routines to indicate the frequency of a given residual value. The m^{th} symbol is drawn if b(m) $\leq$ (the frequency of the residual) $\leq$ e(m).

Table A.2 (continued)

Abscissa position

In POSIT, specify the value in the statement

MLOC(j) = v_j

where v_j is the (plotting) position along the abscissa in E11134 of the j^{th} (observation) location. This allows the user the option of grouping the locations in some meaningful manner (e.g. along a typical wind trajectory or ordered in distance from a central point).

Table A.3

AQMAAP

Input Sequence and Format

Input	Format	Number of Cards
(1) Identification of the data sets		
Specie information		
Agency code number , Symbol, Name, Index code number	I2, 2X, 1A4, 2X 7A4, 14X, I4	IPARM
Location information		
Agency code numbers (County and Site), Name, Index code number, Abbreviation	I2, 1X, I5, 2X, 7A4, 14X, I4, 4X, 2A4	JPARM
Day information		
Day, Index code number	10X, 7A4, 14X, I4	KPARM
(2) Data sets		
Observations		
Specie, Location, Day, Hour start, Hour finish, Units , Method, Concentrations (ppm)	5I3, 2I2, 1X, 6E10.3	Dependent on amount of data used.
End of data flag; '0-1'		1
Predictions		
Specie, Location, Day, Hour start, Hour finish, Units, Method , Concentrations (ppm)	5I3, 2I2, 1X, 6E10.3	Dependent on amount of data used.
End of data flag; '0-1'		1

Table A.3 (continued)

Input	Format	Number of Cards
(3) Specification of methods to be exercised		
Method subroutine code	I10	1
For this method:		Dependent on output requested
Number of species for which method will be exercised	10X,I5	
Specie	15X,I5	
For this specie:		
Number of locations for which method will be exercised	20X,I5	
Location	25X,I5	
For this specie/location:		
Number of days for which method will be exercised	30X,I5	
Day	35X,I5	
For this location/day:		
Number of hours for which method will be exercised	40X,I5	
Hour	45X,I5	

Table A.3 (continued)

Notes:

(1) A complete input deck is required for each of the 7 main routines.

(2) In Input section (1) the index code numbers are chosen by the user. They must be unique with $1 \leq i \leq$ IPARM, $1 \leq j \leq$ JPARM and $1 \leq k \leq$ KPARM.

(3) In Input section (1), agency code number, units and method are optional inputs.

(4) Use the appropriate index code numbers for designating specie, location and day in Input sections (2) and (3).

(5) The assumed units for all concentrations are 'ppm'.

(6) To denote a missing concentration value enter '-0.100E+01'. Delete the entire card for a string of missing concentrations.

(7) Complete Input section (3) for each method subroutine in the program in the order given in Table A.4.

(8) For Input section (3), the required inputs vary with the specific method. See Table A.4.

(9) To skip a method, specify 'zero' for the number of species.

Table A.4

AQMAAP

(A) Sequence of method subroutines to be followed for Input section (3)

Main Routine	Subroutine Sequence within Main Routine
MN - Basic Measures of Bias and Spread	E11111 E11112 E11011 E11012 E12011 E12012 E13011 E13012 E10111 E10112
EB - Error Bands	E11121 E11122 E11021 E11022 E12021 E12022 E13021 E13022 E10121 E10122
HT - Residual Histograms	E11131 E11031
RF - Residual Plots	E11133 E11134 E11135
CR - Analysis of Trends	E11141 E11142 E11041 E11042 E12041 E12042 E13041 E13042 E10141 E10142

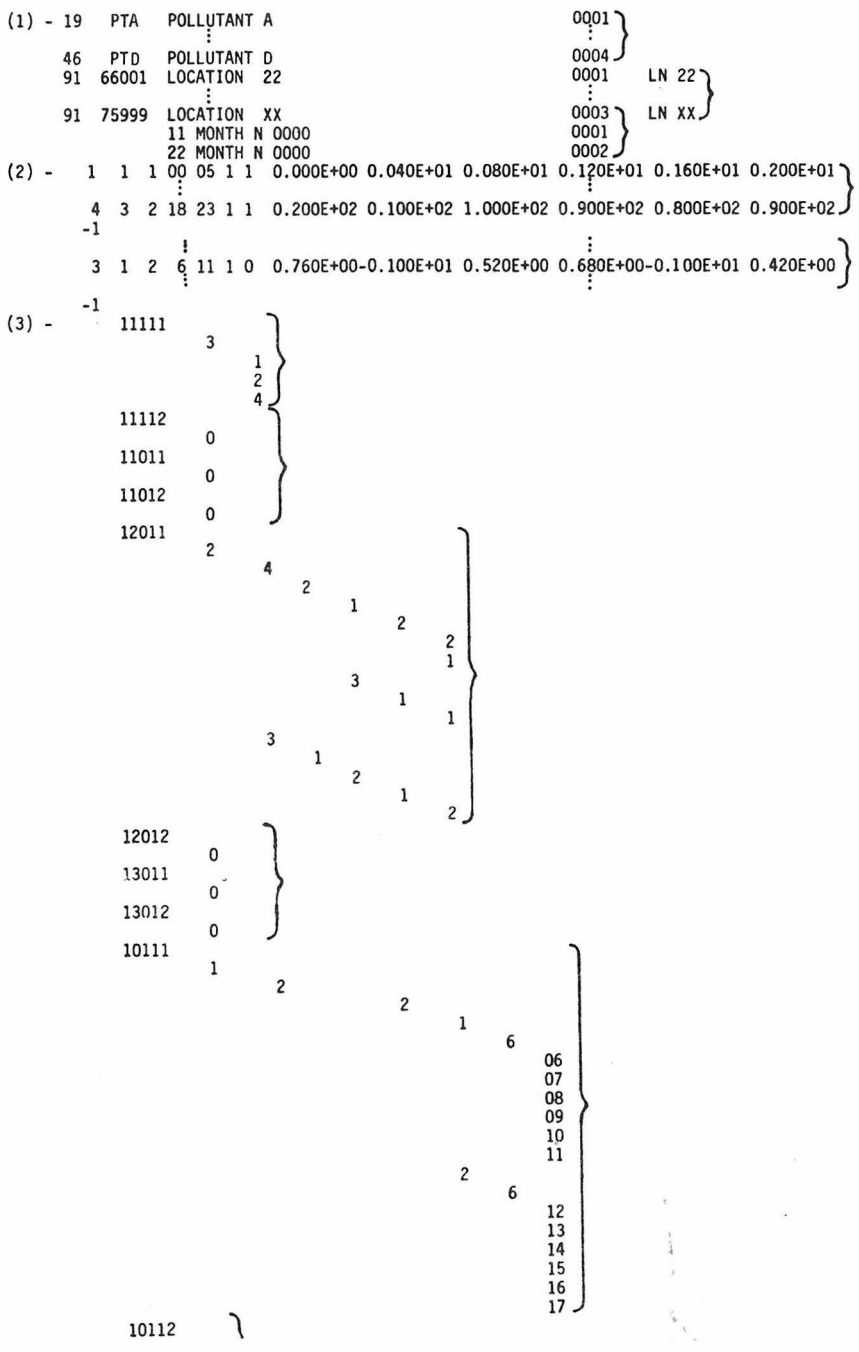
Table A.4 (continued)

Main Routine	Subroutine Sequence within Main Routine
IN - Indices of Air Quality	E22051 E22052 E20156 E22071 E21061
FD - Frequency Distribution Plot	E21061

(B) Sequence of the data set indices to be followed for Input section (3)

Type of Evaluation	Over All Times and Locations	Over All Times	Over Each Day	Over Each Hour	Over all Locations
Method Subroutine Code	x11xx	x10xx	x20xx	x30xx	x01xx
Index: Specie	1	1	1	1	1
Index: Location	-	2	2	2	-
Index: Day	-	-	3	-	2
Index: Hour	-	-	-	3	3

Table A.5
AQMAAP
Sample Input Deck for Program 'MN'



Species, Pollutant A is index #1.

Locations, Location XX is index #3.

Days, The 22th is index #2.

Observations (as available for 4 species, 3 locations, 2 days and 24 hours).
END OF DATA

Predictions (Note: No data for hours 0700 and 1000).
END OF DATA

Exercise method E11111 for the 3 species, 1,2 and 4.

Skip methods E11112, E11011 and E11012.

Exercise method E12011 for:
specie 4, location 1, day 2,
specie 4, location 1, day 1,
specie 4, location 3, day 1
and
specie 3, location 2, day 2.

Skip methods E12012, E13011 and E13012.

Exercise method E10111 for:
specie 2, day 1 and 0600 hours → 1100 hours
and
specie 2, day 2 and 1200 hours → 1700 hours

Skip method E10112.

Table A.6

AQMAAP

SUPPORT-MAIN REFERENCE TABLE

SUPPORT SUBROUTINE	MAIN ROUTINE: MN	EB	HT	RF	CR	IN	FD
INCOD3	X	X	X	X	X	X	X
INDAT1	X	X	X	X	X	X	X
OTDAT3	X	X	X	X	X	X	X
RESID	X	X	X	X			
OTRES3	X	X	X	X			
SETEB		X					
POSIT				X			
AXES			X	X			X
TICKS			X	X			X
SCLS1			X	X			X
SCLS2			X	X			
SCLS4				X			
SCLS5							X
SCLS6				X			
SCLS7				X			

Table A.7

AQMAAP

(A) AQMAAP-CITCC REFERENCE TABLE

AQMAAP ROUTINE	CITCC ROUTINE: RTIME	SORTDE	SORTI2	SORTIR	YNTERP	MAXMIN	OUTCOR	SYSSYM	SYSEND	SYSPLT	XYPLT	-NONE-
E11131						X	X	X			X	
E11031						X	X	X			X	
E11134						X	X	X	X			
E11135						X	X	X	X			
E11133			X	X		X	X	X	X			
E21061		X			X	X	X	X	X		X	
AXES										X		
TICKS										X		
SCLS1							X	X				
SCLS2							X	X				
SCLS4							X	X				
SCLS5							X	X				
SCLS6							X	X				
SCLS7							X	X				
MN	X											
EB	X											
HT	X											
RF	X											
CR	X											
IN	X											
FD	X											
E21062		X										
Exxx11												X
Exxx12												X
Exxx21												X
Exxx22												X
Exxx41												X
Exxx42												X
E22051												X
E22052												X
E20156												X
E22071												X
INCOD3												X
INDAT1												X
OTDAT3												X
RESID												X
OTRES3												X
SETEB												X
POSIT												X

Table A.7 (continued)

(B) Catalog of CITCC Routines

Routine (Name/Number)	Function
RTIME* C 868-249-370	Compute remaining execution time.
SORTDE C1268-279-370	Sort an integer array in decreasing order.
SORTI2 C1268-279-370	Sort and index a real array in increasing order.
SORTIR C1268-279-370	Sort a real array in increasing order.
YNTERP C 766-176-370-10	Evaluation of functions available in tabular form at arbitrary values of the independent variable.
MAXMIN C 267-214-370	Determine the maximum and minimum of an array.
OUTCOR C 169-288-370	Convert data with a FORMAT and place the EBCDIC formated line image in core storage.
SYSSYM C 367-218-370	Plot alphanumeric symbols.
SYSEND C 467-221-370	Terminate a plot.
SYSPLT C 467-222-370	Move plotter pen.
XYPLT C 774-411-370	Plot point data in a 2-D field.

*An optional routine, can be removed without altering basic program function.

Chapter 3

On Frequency Distributions of Air Pollutant Concentrations

Atmospheric Environment Vol. 10, pp. 941–950. Pergamon Press 1976. Printed in Great Britain.

ON FREQUENCY DISTRIBUTIONS OF AIR POLLUTANT CONCENTRATIONS

KENNETH E. BENCALA and JOHN H. SEINFELD*

Department of Chemical Engineering, California Institute of Technology, Pasadena, CA 91125, U.S.A.

(*First received* 21 *January* 1976 *and in final form* 2 *June* 1976)

Abstract—Observed frequency distributions of air pollutant concentration levels are critically analyzed with respect to their statistical description. It is demonstrated that several common distributions can be used to fit observed data, one of which is the popular log-normal distribution. The observation that concentration distributions for all averaging times are approximately log-normal can be explained if the short averaging time data are themselves assumed to be log-normally distributed. The near log-normality of pollutant concentration frequency distributions can be explained on the basis of the near log-normality of wind speed distributions, although this explanation does not establish that wind speed distributions are solely responsible for observed concentration distributions. It is concluded that pollutant concentration frequency distributions are the result of complex phenomena and cannot be predicted exactly, but that the approximate log-normal character of the distributions is useful from a practical point of view and can be understood qualitatively on the basis of the relation between wind speed and concentration.

INTRODUCTION

Most current United States federal air quality standards are stated in terms of the yearly frequency of violation of a specified concentration level for a given averaging time. For example, the 1-h average carbon monoxide concentration may only exceed 35 ppm once during the year. This is equivalent to specifying that the 1-h average CO concentration may exceed 35 ppm only 0.011% of the time. Therefore, if the frequency distribution for hourly average CO levels in a particular urban area could be predicted, then it could be ascertained what degree of emission control would be required to enable meeting of the air quality standard.

There has been interest for some time in the frequency distributions of air pollutant concentrations. Larsen (1971) carried out a comprehensive analysis of the data collected in the Continuous Air Monitoring Program (CAMP) for the years 1962–1968. The CAMP data contain measurements of seven pollutants (CO, NO, NO_2, NO_x, oxidant, SO_2, and hydrocarbons) in eight cities. Readings were recorded every five minutes and then averaged over time periods ranging from ten minutes to one year. Based on his analysis of the data Larsen concluded that regardless of pollutant, city, or averaging time urban concentration frequency distributions could be universally represented as the log-normal. (The log-normal is a two parameter distribution and is represented by a straight line on log-probability paper.) For example, Fig. 1 shows frequency distributions for one hour averaged CO concentration in various cities. Figure 2 shows frequency distributions for CO concentrations in Chicago for various averaging times. Quantitative explanations of why the distributions tend to be log-normal are incomplete, and there is currently no way of predicting how the distributions will shift if emission levels are changed.

The objectives of this work are as follows. First, we wish to consider the question of whether or not

* To whom correspondence should be addressed.

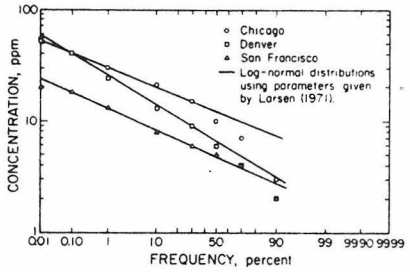


Fig. 1. 1 h average CO concentration distributions; CAMP data (1962–1968).

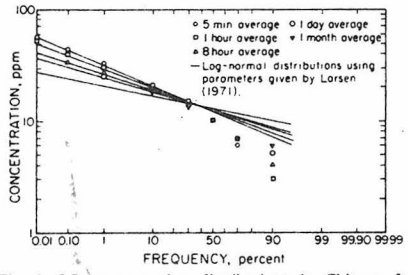


Fig. 2. CO concentration distributions in Chicago for various averaging times; CAMP data (1962–1968).

among common distributions the log-normal does, in fact, provide the best fit for pollutant concentration data. Specifically, we desire to see if other available statistical distributions provide a better fit to a selected sample of data. Second, we wish to analyze the effect of averaging time on the frequency distribution of air pollutant concentrations. We seek specifically to understand the observation that frequency distributions tend to be log-normal regardless of averaging time. Finally, we desire to analyze the possible physical reasons for the near log-normality of air pollutant data.

Our overall aim is to attempt to shed some additional light on the issue of air pollutant concentration frequency distributions. From the outset it must be recognized that these distributions are the result of many complex phenomena, and that we cannot expect to be able to predict them exactly in a given situation. Nevertheless, if the extent of validity of the distributions can be understood at least in a semi-quantitative manner, the use of the distributions can be facilitated.

REPRESENTATION OF CONCENTRATION DATA

Although the log-normal distribution has generally been used to represent air pollution concentration frequency distributions, there are other common distributions which resemble the log-normal and are candidates for representing the data (Lynn, 1974; Mage and Ott, 1975; Pollack, 1975). It is of interest to examine how well different distributions actually fit air quality data. Because chemical reaction behavior may affect the form of concentration distributions, we confine our attention to CO. Our purpose in this section is to examine the ability of a variety of statistical distributions to fit selected air quality data.

The statistics of a set of air quality data may be analyzed in terms of either its probability density function (pdf) or its distribution function. The distribution function is an integral function of the pdf, thus for the purpose of evaluating the fit of mathematical forms, comparison of the data to the pdf provides

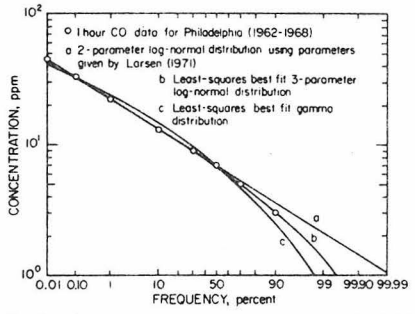


Fig. 3. CO concentration distributions in Philadelphia with lines representing the two-parameter log-normal, the three-parameter log-normal and the gamma distributions.

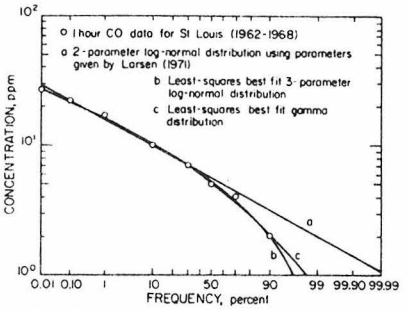


Fig. 4. CO concentration distributions in St. Louis with lines representing the two-parameter log-normal, the three-parameter log-normal and the gamma distributions.

a much more stringent test than comparison to the distribution function. Unfortunately, a plot of the pdf is more difficult to interpret than is a plot of the more commonly used distribution function; therefore, we consider the distribution function.

In attempting to determine which parameter values in a particular mathematical form provide the best fit of that form to the data one has considerable latitude. We selected an unweighted least-squares criterion for determining the parameters in the distributions which best fit the data. (However, an implicit weighting arises in that relatively more data points were available at high concentration values than at low.)

Table 1 presents four potentially applicable pdf's. Figures 3 and 4 show Philadelphia and St. Louis CO data plotted against three of these distributions, the two-parameter and three-parameter log-normal and the gamma. The parameters used for the two-parameter log-normal plots are those given by Larsen, while for the other two the least-squares best fit parameters are used. On log-probability coordinates the Weibull and gamma distributions are concave while the three-parameter log-normal can be either concave or convex depending on the sign of the third parameter, δ. Of the eight data sets investigated, six are to some degree concave while two (Los Angeles and Washington) are slightly convex. Table 2 suggests possible physical interpretations of these deviations from linearity.

Table 3 presents a comparison of the sum of squares error in fitting the distribution functions to the eight data sets. In the least-squares sense the three-parameter log-normal is superior to the two-parameter distributions. The added flexibility afforded by a third parameter accounts for this. In comparing Figs. 3 and 4 with the quantitative measures of goodness of fit given in Table 3, we see that in most cases the two-paramter log-normal distribution provides a useful, if not excellent, approximation to the data, but that in that some cases it is possible to find other two-parameter distributions which provide better fits.

Table 1. Common probability density functions

Name	$p_x(x)$		
2-parameter log-normal	$(\sqrt{2\pi}\ x \ln \beta)^{-1} \exp\left\{-\frac{(\ln x - \ln \alpha)^2}{2 \ln^2 \beta}\right\}$		
		$\delta > 0$	$\delta < 0$
3-parameter log-normal	$\int_{-\delta}^{0} [\sqrt{2\pi}(t+\delta)\ln\beta]^{-1} \exp\left\{-\frac{(\ln(t+\delta) - \ln\alpha)^2}{2\ln^2\beta}\right\} dt$	$x = 0$	
	$[\sqrt{2\pi}(x+\delta)\ln\beta]^{-1} \exp\left\{-\frac{(\ln(x+\delta) - \ln\alpha)^2}{2\ln 2\beta}\right\}$	$x > 0$	$x > -\delta$
Weibull	$\alpha x^{\beta} \exp\left\{-\frac{\alpha x^{\beta+1}}{\beta+1}\right\}$		
Gamma	$\frac{x^{\alpha}}{\Gamma(\alpha+1)\beta^{\alpha+1}} \exp\left(-\frac{x}{\beta}\right)$		

Table 2. Implication of the shape of a distribution function represented on log-probability coordinates relative to a straight line (the log-normal)

	Concave shape	Convex shape
At high concentrations	Fewer days of high concentration. An upper limit in concentration suggested.	More days of higher concentration.
At low concentrations.	More days of low concentration.	Fewer days of low concentration. A lower limit (background) in concentration suggested.

Given the complex dynamic nature of urban air pollution, no one distribution will always be the best or even an adequate representation of the data, however, the two-parameter log-normal distribution is clearly a useful distribution both for describing data and for establishing an understanding of air pollutant statistics.

EFFECT OF AVERAGING TIME ON FREQUENCY DISTRIBUTIONS

Ambient data are generally reported in terms of time averaged values. The time averaged value of concentration C centered at time t and over an averaging period T can be represented as

$$\bar{C}_T(t) = \frac{1}{T}\int_{t-T/2}^{t+T/2} C(\eta)\, d\eta. \quad (1)$$

Typical values of T in air pollution applications range from 5 min to 1 y.

It has been observed that air pollutant concentration distributions approximate the log-normal regardless of averaging time, and that the median concentration is proportional to the averaging time raised to a power (Larsen, 1971). Based on this empirical observation, the standard geometric deviation σ_{gb} for one averaging time T_b can be related to the standard geometric deviation σ_{ga} for another averaging time T_a by

$$\sigma_{gb} = \sigma_{ga}^{V}, \quad (2)$$

where

$$V = \left(\frac{\ln(T/T_b)}{\ln(T/T_a)}\right)^{1/2}. \quad (3)$$

Table 3. Sum of squares error in fitting the distributions in Table 1 to 1-h average CO CAMP data, 1962–1968*

City	Two-parameter log-normal Larsen (1971) values	Best fit	Three-parameter log-normal	Weibull	Gamma
Los Angeles	0.35	0.12	0.03	1.08	0.48
Philadelphia	0.15	0.07	0.01	0.55	0.22
Denver	0.87	0.20	0.03	0.36	0.20
San Francisco	0.76	0.56	0.14	0.30	0.17
Cincinnati	0.64	0.32	0.31	1.14	0.83
St. Louis	1.25	0.44	0.04	0.13	0.04
Washington	0.31	0.08	0.05	0.78	0.57
Chicago	7.24	1.17	0.04	0.08	0.20

* Error based on reduced variate.

where T is the total period over which data are available (usually one year).

It is of interest to examine if this empirical observation can be explained strictly on the basis of the properties of log-normality distributed random variables. Thus, we ask—if the raw data averaged over a period of, say, 5 min are *assumed* to be log-normally distributed, will the data averaged over longer periods continue to be log-normally distributed.

Let $X(t)$ represent the average value of the concentration over the time period from $t - \tau$ to t. Thus, $X(t)$ is considered as the "raw" data, with an inherent averaging time τ, attributable to instrument function. The record of raw data then can be represented as the sequence, $X(t_1), X(t_2), \ldots, X(t_n)$, where $X(t_1)$ is the value in the interval $[t_1, t_1 + \tau]$, $X(t_2)$ is the value in the interval $[t_1 + \tau, t_1 + 2\tau] = [t_2, t_2 + \tau]$, etc. Now, the average concentration over the double interval $t - 2\tau$ to t is written as $Z(t) = \frac{1}{2}[X(t - \tau) + X(t)]$. The series $Z(t_2), \ldots, Z(t_n)$ then represents the sequence of concentrations averaged over periods of length 2τ. The basic problem we wish to consider is—Assuming that $X(t)$ is log-normally distributed, determine the probability density function of $Z(t)$. Thus we seek to relate the statistics of time-averaged concentrations to those of the raw data which are used to construct the averages. We carry out the analysis for an averaging period twice the length of the fundamental averaging period on which the raw data are based. In practice, averaging periods greater than twice the basic average are used. For instance, daily averages are computed from 1-h averages. The salient features of the averaging process will, however, be elucidated with a period twice the length of the basic period.

We assume that the first order density function of $X(t)$, $p_X(x;t)$, is log-normal,

$$p_X(x;t) = \frac{1}{\sqrt{2\pi}\, x \ln \sigma_{gx}} \exp\left\{ - \frac{(\ln x - \ln \mu_{gx})^2}{2 \ln^2 \sigma_{gx}} \right\} \tag{4}$$

and that the second order density function $p_X(x, t - \tau; y, t)$ is the following joint log-normal density,

$$p_X(x, t - \tau; y, t) = [2\pi xy \ln^2 \sigma_{gx} \sqrt{1 - r^2}\,]^{-1} \cdot \exp\left\{ \frac{(\ln x - \ln \mu_{gx})^2 - 2r(\ln x - \ln \mu_{gx})(\ln y - \ln \mu_{gx}) + (\ln y - \ln \mu_{gx})^2}{2(1 - r^2) \ln^2 \sigma_{gx}} \right\}, \tag{5}$$

where μ_{gx} and σ_{gx} are respectively the geometric mean and the standard geometric deviation of $X(t)$, and r is a correlation parameter.

The first order density function of $Z(t)$ can be determined from the relation

$$P_Z(z;t) = 2\int_0^{2z} P_X(x, t - \tau; 2z - x, t)\, dx. \tag{6}$$

The distribution function $F_Z(z;t)$ is related to $P_Z(z;t)$ by

$$F_Z(z;t) = \int_0^z P_Z(\eta;t)\, d\eta. \tag{7}$$

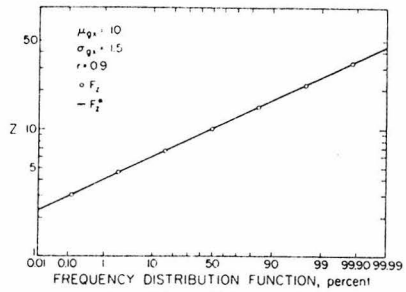


Fig. 5. Comparison of F_Z to F_Z^* for a typically low value of σ_{gx}.

Using (5–7), we obtain F_Z as

$$F_Z(z;t) = \frac{1}{2\sqrt{2\pi} \ln \sigma_{gx}} \int_0^{2z} x^{-1} \times \exp\left(\frac{-(\ln x - \ln \mu_{gx})^2}{2 \ln^2 \sigma_{gx}} \right) \cdot \left\{ 1 + \operatorname{erf}\left(\frac{\ln(2z - x) - r \ln x + (r - 1) \ln \mu_{gx}}{\sqrt{2(1 - r^2)} \ln^2 \sigma_{gx}} \right) \right\} dx. \tag{8}$$

In summary, F_Z is the distribution function of the random variable $Z(t)$, which is the average of two log-normally distributed random variables. We seek to determine how close the distribution of $Z(t)$ approximates a log-normal distribution. Therefore, we compare F_Z from (8) with F_Z^*, a log-normal distribution function with the same mean and variance as F_Z. Comparisons of F_Z and F_Z^* were made in 15 cases: $\sigma_{gx} = 1.18, 1.5, 2, 4, 10$ and $r = 0, 0.5, 0.9$. (The range of σ_{gx} for typical air pollution data is between 1 and 4.) Figures 5 and 6 show comparison of F_Z and F_Z^*. Qualitatively we note that F_Z compares closely with F_Z^*, particularly at large values of z. In the Appendix we discuss in more detail the comparison between F_Z and F_Z^*.

We can conclude from the results shown in Figs. 5 and 6 that the distribution of the average of two correlated log-normal variables approximates a log-

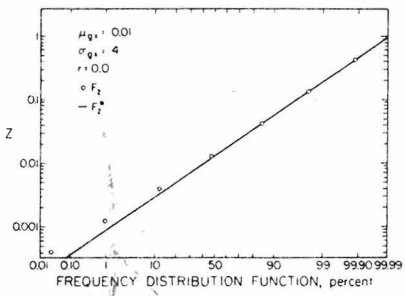


Fig. 6. Comparison of F_Z to F_Z^* for a typically high value of σ_{gx}.

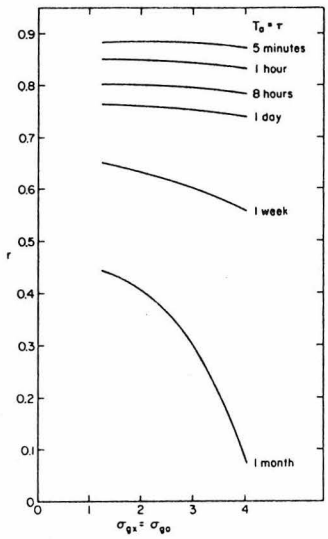


Fig. 7. Value of the correlation parameter r necessary for the standard geometric deviation of F_Z^* to equal σ_{gb} plotted as a function of the standard geometric deviation of the concentration distribution for averaging time T_a.

normal distribution. We now return to the observation (2). If we let $T_a = \tau$ and $T_b = 2\tau$, then we ask—Given $\sigma_{gx} = \sigma_{ga}$, can we, by appropriate choice of the correlation parameter r, have the standard geometric deviation of F_Z^* equal σ_{gb}. Figure 7 shows that for the values of σ_{gx} and τ of interest such an r exists. Although the correlation parameter r is not known for actual data, the results of this section indicate that the observation (2) can be explained simply as a consequence of the near log-normality of the short-averaging time concentrations.

Others have also investigated the effects of averaging time on concentration frequency distributions. There does not appear to be any other work which predicts both the form and parameters of time averaged concentration distributions, however, other relationships for measures of the variance in the distribution versus averaging time have been developed. Saltzman (1970) presented an empirical relationship for the standard geometric deviation, while Shoji and Tsukatani (1973) and Larsen and Peterson (1974) developed relationships based on assumed spectral properties of the concentration time series.

ANALYSIS OF FREQUENCY DISTRIBUTIONS

Up to this point we have shown that several distributions are capable of representing a number of air pollutant concentration frequency distributions. The log-normal distribution, while not, in fact, a perfect representation of the data is a convenient one because the parameters of the distribution can be easily determined from a log-probability plot of the data. Assuming log-normality of the basic data, we then showed that time averages formed from the original data would essentially preserve the log-normality of the data. We now come to the crucial question— why do the concentrations tend to be approximately log-normally distributed in the first place. This section is devoted to an attempt to propose possible explanations for this observed phenomenon.

Because of the approximate universality of pollutant frequency distributions we would expect the principal factors affecting urban concentration frequency distributions also to exhibit universality. That is, the fundamental characteristics of such a factor relative to its influence on urban pollutant concentration must be approximately the same for all pollutants in all cities. Conversely, a factor which is fundamentally different either from city to city or in its effect on different pollutants does not appear to have a major influence on frequency distributions.

Certainly there exists no characteristic similarity over all cities or all pollutants among either the spatial distribution or strength of sources. Similarly, wind direction distributions can be argued not to be a major factor in air pollutant frequency distributions. The concentration of a pollutant near the center of a uniform area source will not be sensitive to wind direction, whereas the concentration at a position near a single point source will be very sensitive to wind direction. Thus the impact of wind direction on concentration can range from negligible to significant and certainly varies from location to location. By such reasoning one is led to the conclusion that the two factors most likely to influence air pollutant frequency distributions are wind speed and mixing height. Increases in both factors will lead to a decrease in concentration.

A number of studies have been carried out investigating the correlation between wind speed, mixing depth, and air pollutant concentrations. Schmidt and Velds (1969) calculated a correlation coefficient between yearly average SO_2 concentrations in Rotterdam and wind speed of -0.97. Marsh and Withers (1969) found significant correlation between 6 h average SO_2 concentrations in Reading and wind speed but not between vertical turbulence and concentration. Similarly, in analyzing hourly ozone data in New York City, Bruntz *et al.* (1974) determined close correlation with wind speed, and were not able to improve the correlation by including mixing height. While certainly not proving that wind speed is the sole factor governing pollutant frequency distributions, these studies do indicate at least that wind speed/concentration correlations are as one might expect. Because of the difficulty in measuring mixing depths relative to wind speeds, fewer studies exist wherein mixing depth/concentration correlations were computed. Undoubtedly, mixing depth does play a role in determining pollutant frequency distributions.

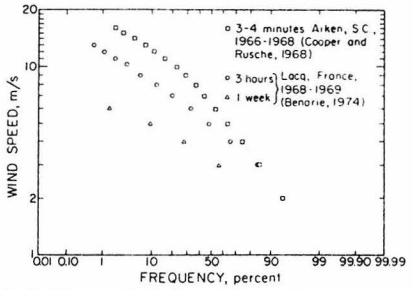


Fig. 8. Wind speed distributions in two cities. Data from Lacq shown for two averaging times.

From the above qualitative arguments we have arrived at the conclusion that wind speed and mixing depth should be the major factors influencing air pollutant frequency distributions. Following this supposition, we need to be more quantitative, that is to determine in what manner wind speed and mixing depth influence frequency distributions. The remainder of this section is devoted to an attempt to explain observed frequency distributions on a fundamental basis. In doing so, we restrict our attention to non-chemically reacting pollutants.

Influence of wind speed on instantaneous concentrations

In this subsection we wish to consider the influence of wind speed on the distribution of instantaneous pollutant concentrations. To begin, we need some notion as to observed frequency distributions of wind speed. Figure 8 shows frequency distributions of wind speed from Lacq, France and Aiken, SC. In both cases we note that the wind speed is approximately log-normally distributed. Why wind speeds seem to be log-normally distributed and even, in fact, if the log-normal is the best distribution to describe wind speeds are questions not central to our purpose here. We seek only to ascertain the consequences of this approximate log-normality in determining pollutant frequency distributions.

We begin in some sense with the most basic problem, that of determining the effect of wind speed variations on the instantaneous concentration of a pollutant released into that wind field. Clearly, attacking this problem on the basis of a three-dimensional urban flow is impossible. Therefore, we need to isolate the key elements of the problem in a much simpler hypothetical situation, which, nevertheless, retains the basic physics. Such a situation is embodied in a one-dimensional flow into which a pollutant is steadily emitted at a plane.

The fundamental equation describing the instantaneous concentration C of an inert atmospheric species is the continuity equation.

$$\frac{\partial C}{\partial t} + \nabla \cdot \mathbf{U}C = \mathscr{D}\nabla^2 C \tag{9}$$

where $\mathbf{U}$ is the instantaneous wind velocity vector, and $\mathscr{D}$ is the molecular diffusivity of the species in air. Molecular diffusion is normally neglected when (9) is applied to atmospheric species, the result being the so-called advection equation.

$$\frac{\partial C}{\partial t} + \nabla \cdot \mathbf{U}C = 0. \tag{10}$$

Because the wind speed $\mathbf{U}$ is a random variable, the instantaneous concentration C is a random variable.

As noted above, let us consider a one-dimensional flow (in the x-direction) into which a pollutant is steadily emitted at the $x = 0$ plane at a rate $S\,\mathrm{g\,cm^{-2}\,s^{-1}}$. The wind velocity in the x-direction is taken to be a function of time only, i.e. $U_x = U(t)$. This case, although highly simplified, exhibits the basic features of the situation in which there are three velocity components.

The instantaneous concentration is described by the one-dimensional form of (10).

$$\frac{\partial C(t,x)}{\partial t} + U(t)\frac{\partial C(t,x)}{\partial x} = 0 \tag{11}$$

subject to

$$C(0,x) = 0 \tag{12}$$

$$C(t,0) = \frac{S}{U(t)} \quad t > 0. \tag{13}$$

The solution of (11–13) is

$$C(t,x) = \left.\frac{S}{U(t')}\right|_{x = \int_0^{t'} U(t'')dt''} \tag{14}$$

Equation (14) relates the concentration at any position x and time t to the source strength S and the wind speed. We recall that $U(t)$ is a random variable, and therefore that $C(t,x)$ is a random variable. Given the pdf of $U(t)$, $p_U(u;t)$, we wish to determine the pdf of C, $p_C(c;t,x)$. It is advantageous to assume that there are only a finite number l of possible wind speeds, with the maximum wind speed denoted by u_l. In addition, we assume that a given wind speed persists for a time Δt. If Δx is the distance a fluid element moves in Δt corresponding to the slowest wind speed u_1, i.e. $u_1 = \Delta x/\Delta t$, and if the wind speeds obey the relations, $u_i = iu_1$, $i = 1,2,\ldots,l$, then the probability of observing concentration c_i at time $t_n = n\Delta t$ and position $x_m = m\Delta x$, $p_C(c_i; t_n, x_m)$, is given by

$$x_m \geq x_1 \text{ and } t_1 \leq t_n < \infty$$

$$p_C(c_i; t_n, x_m) = \sum_{j=1}^{l} p_U(u_j; t_n) p_C(c_i; t_{n-1}, x_{m-j}) + \begin{Bmatrix} p_U(u_i; t_n) & \text{if } x_m \leq i\Delta x \\ 0 & \text{if } x_m > i\Delta x \end{Bmatrix} \quad i = 1,2,3,\ldots,l \tag{15a}$$

$$-\infty < x_m \leq 0 \text{ or } t_n \leq 0$$

$$p_C(c_i; t_n, x_m) = 0 \qquad i = 1,2,\ldots,l. \tag{15b}$$

The first term in (15a) represents the probability of a fluid element with concentration c_i being

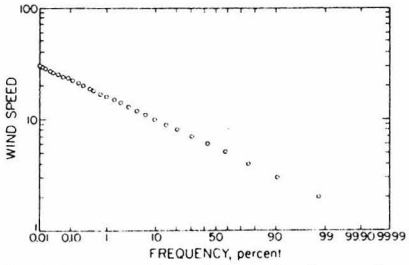


Fig. 9. Assumed wind speed distribution for one-dimensional advection problem.

advected to location x_m while the second term represents the probability of 'fresh' emissions reaching x_m. As l is increased, this formulation yields an approximation to the continuous density $P_C(c;t,x)$. The solution was evaluated for the case $l = 30$, $\Delta t = 1$ and $u_l = 30$. The wind speed was assumed to be log-normally distributed in accordance with observation. Its frequency distribution is shown in Fig. 9. The resulting frequency distribution for the concentration at $x = 30$, $t = 30$, $C(30,30)$, is shown in Fig. 10. The result is a close approximation to a log-normal distribution.

In summary, this situation of a steadily emitting source in a one-dimensional flow in which the wind speed is log-normally distributed leads to an instantaneous concentration that is approximately log-normally distributed. This example, while clearly highly idealized, does, however, demonstrate rigorously that log-normality of wind speed does lead to log-normality of instantaneous concentration. Of course, in an atmospheric flow other phenomena will influence concentration distributions, and so this example does not necessarily establish the cause of approximate log-normality in air quality data.

Influence of wind speed and mixing depth on mean concentrations

In the previous subsection we considered the effect of wind speed distribution on the instantaneous concentration of a species. Monitoring data reflect the instantaneous concentration at a point. This instantaneous concentration is a random quantity because of the turbulent nature of the atmosphere. When air pollutant concentrations are analyzed from a theoretical point of view, only the mean concentration $\langle c \rangle$ can be predicted, where $C = \langle c \rangle + c'$. (The mean concentration $\langle c \rangle$ is a function of location and time but is theoretically the result of an ensemble average.) Virtually all mathematical models of air pollutant behavior are concerned with the prediction of $\langle c \rangle$ (Lamb and Seinfeld, 1973).

There exist a large number of urban air pollution models depending on the dynamic nature (steady vs unsteady), type of source (point, line, area), number of spatial dimensions, meteorological assumptions, boundary conditions, etc. In these models the inputs are usually specified as known, or deterministic, quantities (such as wind speed and mixing depth). However, in attempting to assess the effect of the variability of these inputs on the predicted mean concentration, one can propose to allow the inputs to assume a *distribution* of values and determine the resulting *distribution* of values of the mean concentration. In particular, we are interested, of course, in the effect of the variability of wind speed and mixing depth on the distribution of mean concentrations.*

Box model. It is sometimes assumed that the mean concentration in a local region of an urban area can be represented by the simple box model relationship,

$$\langle c \rangle = \frac{k}{uh}, \tag{16}$$

where u is the mean wind speed, h is the mixing depth, and k is an empirical proportionality constant. If we allow u and/or h to be random variables, then the distribution of $\langle c \rangle$ can be computed directly from (16). Knox and Lange (1974) employed this approach by assuming a frequency distribution for u from hourly average readings and computing the frequency distribution of $\langle c \rangle$. In the cases presented, the approximate log-normality of the observation was reproduced.

We have investigated the effect of assuming different forms for the wind speed distribution on the concentration distribution using (16). In addition to the log-normal, the Weibull, gamma, uniform, and bounded distributions were investigated (Table 4). In all cases we assumed that the mean wind speed was the same and for the two-parameter distributions that the wind speed variance was also the same. The wind speed densities along with the resulting concentration densities are given in Table 4. In most cases the high concentration end of the distribution could be approximated by a straight line (although not the one

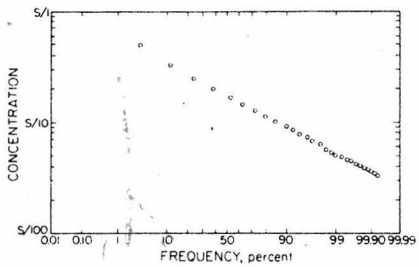


Fig. 10. Calculated concentration distribution corresponding to wind speed distribution in Fig. 9.

* It is reasonable to suppose that the distribution of *mean concentrations* over a time period of the order of several months to one year is not significantly different than the distribution of *instantaneous concentrations* over the same time period. Therefore, this approach should not be wholly inapplicable to the basic issue of analyzing ambient monitoring data.

Table 4. Relationship between probability density functions for wind speed and concentration when $c = k/u$

	Pdf for wind speed $p_u(u)$	Pdf for concentration $p_c(c)$
Log-normal	$(\sqrt{2\pi}\, u \ln \beta)^{-1} \exp\left\{-\frac{(\ln u - \ln \alpha)^2}{2 \ln^2 \beta}\right\}$	$(\sqrt{2\pi}\, c \ln \beta)^{-1} \exp\left\{-\frac{(\ln c - \ln k/\alpha)^2}{2 \ln^2 \beta}\right\}$
Weibull	$\alpha u^{\beta} \exp\left\{-\frac{\alpha u^{\beta+1}}{\beta+1}\right\}$	$\alpha k^{\beta+1} c^{-(\beta+2)} \exp\left\{-\frac{\alpha}{\beta+1}\left(\frac{k}{c}\right)^{\beta+1}\right\}$
Gamma	$\frac{1}{\beta^{\alpha+1}\Gamma(\alpha+1)} u^{\alpha} \exp\left(-\frac{u}{\beta}\right)$	$\frac{k^{\alpha+1} c^{-(\alpha+2)}}{\beta^{\alpha+1}\Gamma(\alpha+1)} \exp\left(-\frac{k}{\beta c}\right)$
Uniform	α^{-1}	$\frac{k}{\alpha c^2}$
Bounded	$-\alpha^{-1} \ln\left(\frac{\alpha - u}{\alpha}\right)$	$\frac{-k}{\alpha c^2} \ln\left(\frac{\alpha - k/c}{\alpha}\right)$

found by assuming the original log-normal distribution for wind speeds). It is reasonable to conclude that with the simple model of (16) a variety of wind speed distributions could actually exist and still the log-normal would be an adequate approximation to the concentration distribution. The explanation, of course, lies in the inverse relationship between concentration and wind speed, as pointed out previously by Benarie (1969, 1971, 1974) and Pollack (1975). Similarly, the inverse relationship between concentration and mixing depth indicates that the same correspondence between mixing depth and concentration is to be expected.

Gaussian plume model. For conditions of (1) a continuous point source located at $(0,0,z_0)$ and (2) wind speed constant and direction aligned with the x-axis, the ground-level mean concentration $\langle c(x,y,0)\rangle$ can be estimated by the familiar Gaussian plume equation,

$$\langle c(x, y, 0)\rangle = \frac{S}{\pi\sigma_y\sigma_z u} \exp\left\{\frac{-y^2}{2\sigma_y^2}\right\} \exp\left\{\frac{-z_0^2}{2\sigma_z^2}\right\}. \quad (17)$$

Let us assume that the wind speed u is distributed according to a log-normal density with parameters μ_u and σ_u. As expected, we then obtain $P_{\langle c\rangle}(\langle c\rangle)$ from (17) as log-normal.

$$p_{\langle c\rangle}(\langle c\rangle) = \frac{1}{\sqrt{2\pi}\langle c\rangle \ln \sigma_c} \times \exp\left\{\frac{-(\ln\langle c\rangle - \ln \mu_c)^2}{2 \ln^2 \sigma_c}\right\}, \quad (18)$$

where $\sigma_c = \sigma_u$ and

$$\mu_c = \frac{S}{\pi\sigma_y\sigma_z\mu_u} \exp\left\{\frac{-y^2}{2\sigma_y^2}\right\} \exp\left\{\frac{-z_0^2}{2\sigma_z^2}\right\}. \quad (19)$$

Eddy diffusion model. Finally, we consider two-dimensional, steady-state diffusion as described by the atmospheric diffusion equation,

$$u(z)\frac{\partial\langle c\rangle}{\partial x} = \frac{\partial}{\partial z}\left(K(z)\frac{\partial\langle c\rangle}{\partial z}\right), \quad (20)$$

where $u(z) = u_1 z^{\alpha}$ and $K(z) = K_1 z^{\beta}$. The solutions of (20) for ground-level crosswind line and area sources are, (Monin and Yaglom, 1971; Lebedeff and Hameed, 1975)

$$\langle c(x, 0)\rangle = \frac{Slp}{u_1\Gamma(q)}\left(\frac{u_1}{K_1 x p^2}\right)^q \quad (21)$$

and

$$\langle c(x, 0)\rangle = \frac{S_a}{K_1(1-\beta)\Gamma(q)}\left(\frac{u_1}{K_1 x p^2}\right)^{\beta-1/p}, \quad (22)$$

where $p = \alpha - \beta + 2 > 0$, $q = (\alpha+1)/(\alpha-\beta+2)$, and $0 \le p < 1$. Equations (21 and 22) can be written in the form $\langle c(x,0)\rangle = Au_1^{\gamma}$. If we now assume u_1 to be log-normally distributed, we find $p_{\langle c\rangle}(\langle c\rangle)$ to be given by (18) with $\sigma_c = \sigma_{u_1}^{|\gamma|}$ and $\mu_c = A\mu_{u_1}^{\gamma}$. Thus, the eddy diffusion model, while more detailed than the Gaussian model, still predicts a log-normal concentration distribution provided that the wind speed is log-normally distributed.

Summary

In this section we have attempted to shed some light on the fundamental question of why urban air pollutant concentration frequency distributions tend to be log-normal. Starting from information available from statistical correlations between wind speed and mixing depth (primarily wind speed), we investigated

how both instantaneous and mean concentrations might be influenced by wind speed and mixing depth (primarily wind speed) variability. In all cases we found that if wind speeds are nearly log-normally distributed then resulting concentrations will be nearly log-normally distributed. In fact, other distributions, namely the gamma and Weibull, are capable of producing nearly log-normal concentrations. This result does not, of course, establish that wind speeds are the primary influence on concentration distributions in the atmosphere, since other effects are most certainly influential. Nevertheless, the results here are convincing of the role of wind speed in pollutant frequency distributions.

It is interesting to note that in the three simple models considered for the mean concentration, the standard geometric deviation is independent of the source strength and the geometric mean varies linearly with it. Thus, if one were using (17, 21, or 22) to predict mean concentrations for source emission changes, only the intercept of the concentration distributions plotted on log-probability paper would change, not the slopes.

CONCLUSIONS

Air pollutant concentration frequency distributions are the result of complex phenomena. The direct prediction of these distributions does not appear to be possible. Observed data are generally represented as log-normal, although other common statistical distributions are capable of representing the data as well as or better than the log-normal. The log-normal is convenient because the mean and variance can be easily determined from a log-probability plot of the data. The fundamental question of why concentration distributions tend to be approximately log-normal cannot be answered unequivocally. The persistence of log-normality for all averaging times can be explained if the raw data are themselves log-normally distributed. The log-normality of the raw data, i.e. the instantaneous concentrations, can be shown to result if wind speed is log-normally distributed. Conventional models for mean concentrations, such as the Gaussian plume and eddy diffusion, can also lead to log-normality for concentration distributions if the wind speeds are log-normal. It is shown how these models may be used to estimate the shift in the distribution resulting from source emission level changes.

Acknowledgement—This work was supported by National Science Foundation Grant ENG71-02486.

REFERENCES

Benarie M. (1969) Calculation of the amount and the nuisance of pollutants emitted from a point source. *Atmospheric Environment* **3**, 467–473.

Benarie M. (1971) About the validity of the log-normal distribution of pollutant concentrations. Paper SU-18D. *Proceedings of the 2nd International Clean Air Congress* (Edited by H. M. Englund and W. T. Berry). Academic Press, New York.

Benarie M. (1974) The use of the relationship between wind velocity and ambient pollutant concentration distributions for the estimation of average concentrations from gross meteorological data. Paper 5, *Proceedings of the Symposium on Statistical Aspects of Air Quality Data*. Environmental Protection Agency Pub. No. 650/4-74-038.

Bruntz S. M., Cleveland W. S., Kleiner G. and Warner J. L. (1974) The dependence of ambient ozone on solar radiation, wind, temperature, and mixing height. *Symposium on Atmospheric Diffusion and Air Pollution*, Santa Barbara, CA. American Meteorological Society, Boston, MA.

Cooper R. E. and Rusche B. C. (1968) The SRL meteorological program and off-site close calculations. E. I. DuPont de Nemours and Co., Savannah River Laboratory, Aiken, SC., Rept. DP-1163.

Knox J. B. and Lange R. (1974) Surface air pollutant concentration frequency distributions: Implications for urban modeling. *J. Air Pollut. Control Ass.* **24**, 48–53.

Lamb R. G. and Seinfeld J. H. (1973) Mathematical modeling of urban air pollution: General theory. *Environ. Sci. Technol* **7**, 253–261.

Larsen R. I. (1971) A mathematical model for relating air quality measurements to air quality standards. Environmental Protection Agency Pub. No. AP-89.

Larsen S. E. and Petersen E. L. (1974) Statistical description of air pollution concentration, averaging time and frequency. *Symposium on Atmospheric Diffusion and Air Pollution*, Santa Barbara, CA. American Meteorological Society, Boston, MA.

Lebedeff S. A. and Hameed S. (1975) Steady state solution of the semi-empirical diffusion equation for area sources. (Preprint).

Lynn D. A. (1974) Fitting curves to urban suspended particulate data. Paper 13, *Proceedings of the Symposium on Statistical Aspects of Air Quality Data*. Environmental Protection Agency Pub. No. 650/4-74-038.

Mage D. T. and Ott W. R. (1975) An improved statistical model for analyzing air pollution concentration data. Paper No. 75-51.4, *68th Meeting of the Air Pollution Control Association*, Boston, MA.

Marsh K. J. and Withers V. R. (1969) An experimental study of the dispersion of the emissions from chimneys in Reading—III. The investigation of dispersion calculations. *Atmospheric Environment* **3**, 281–302.

Mitchell R. L. (1968) Permanence of the log-normal distribution. *J. Opt. Soc. Am.* **58**, 1267–1272.

Monin A. S. and Yaglom A. M. (1971) *Statistical Fluid Mechanics*. MIT Press, Cambridge, MA.

Pollack R. I. (1975) *Studies of pollutant concentration frequency distributions*. Environmental Protection Agency Pub. No. 650/4-75-004.

Saltzman B. E. (1970) Significance of sampling time in air monitoring. *J. Air Pollut. Control Ass.* **20**, 660–665.

Schmidt F. H. and Velds C. A. (1969) On the relation between changing meteorological circumstances and the decrease of sulphur dioxide concentration around Rotterdam. *Atmospheric Environment* **3**, 455–460.

Shoji H. and Tsukatani T. (1973) Statistical model of air pollutant concentration and its application to the air quality standards. *Atmospheric Environment* **7**, 485–501.

APPENDIX. THE DISTRIBUTION OF THE AVERAGE OF TWO CORRELATED LOG-NORMAL VARIATES

Equations (7 and 8) give the distribution function, F_Z, of the average of two correlated log-normal variates. Figures 5 and 6 show typical comparisons of F_Z to F_Z^*, the log-normal distribution function with the same mean

Table A.1. Comparison of the actual difference $F_Z(z)-F_Z^*(z)$ with the first correction term $E(z)$ for the distribution of the average of two log-normal variates

σ_{gx}	μ_{gx}	r	z	$E(z)$	$F_Z(z)-F_Z^*(z)$
1.5	10	0.0			
			4.32	-0.413×10^{-4}	-0.104×10^{-3}
			7.76	-0.816×10^{-3}	-0.950×10^{-3}
			10.4	0.231×10^{-3}	0.791×10^{-3}
			13.9	0.829×10^{-3}	0.696×10^{-3}
			25.0	-0.129×10^{-3}	0.107×10^{-3}
		0.9	3.07	-0.930×10^{-8}	-0.749×10^{-7}
			6.76	-0.300×10^{-6}	-0.549×10^{-6}
			10.0	-0.778×10^{-7}	0.330×10^{-6}
			14.9	0.255×10^{-6}	-0.183×10^{-4}
			32.9	-0.345×10^{-7}	-0.195×10^{-6}
4.0	0.01	0.0	0.397×10^{-3}	-0.222×10^{-6}	-0.111×10^{-2}
			0.411×10^{-2}	-0.232×10^{-4}	-0.416×10^{-1}
			0.132×10^{-1}	-0.573×10^{-4}	-0.249×10^{-1}
			0.425×10^{-1}	-0.501×10^{-4}	0.519×10^{-2}
			0.440	0.519×10^{-5}	-0.494×10^{-5}
		0.9	0.181×10^{-3}	-0.246×10^{-10}	-0.484×10^{-8}
			0.271×10^{-2}	-0.271×10^{-8}	-0.795×10^{-8}
			0.105×10^{-1}	-0.727×10^{-8}	-0.557×10^{-4}
			0.405×10^{-1}	-0.771×10^{-8}	0.378×10^{-5}
			0.606	0.218×10^{-9}	0.330×10^{-7}

and variance as F_Z. Qualitative comparison such as these indicate that F_Z is itself nearly log-normal. Following Mitchell's (1968) analysis of the sum of n independent log-normal variates, we investigate analytically the near log-normality of F_Z in this Appendix.

The pdf, p_Z, of the average of two correlated log-normal variates can be expressed as an orthogonal polynomial expansion in terms of p_Z^*, the pdf of the log-normal density with the same mean and variance as p_Z, i.e.

$$p_Z(z) = b_0 u_0(z) p_Z^*(z) + b_1 u_1(z) p_Z^*(z) + \ldots, \qquad (A.1)$$

where the u_n are the orthogonal polynomials of order n defined by

$$\int u_n(\eta) u_m(\eta) p_Z^*(\eta)\, d\eta = \begin{cases} 1 & m = n \\ 0 & m \neq n \end{cases} \qquad (A.2)$$

and the b_n are constants

$$b_n = \int u_n(\eta) p_Z(\eta)\, d\eta. \qquad (A.3)$$

Because p_Z and p_Z^* have identical means and variances (A.1) becomes

$$p_Z(z) = p_Z^*(z) + b_3 u_3(z) p_Z^*(z) + b_4 u_4(z) p_Z^*(z) + \ldots \qquad (A.4)$$

In this form we see that p_Z is the long-normal density, p_Z^*, plus a "correction" consisting of an infinite series of polynomials multiplying p_Z^*. If we truncate (A.4) after the first correction term and (1) if the truncated expansion is a good approximation to p_Z and (2) if the correction term is small, then we see analytically that p_Z is in fact nearly log-normal. Similarly we can integrate the truncated expansion and form the same conclusions about the distribution function F_Z. The expansion is then

$$F_Z(z) \approx F_Z^*(z) + E(z), \qquad (A.5)$$

where

$$E(z) = \int_0^z b_3 u_3(\eta) p_Z^*(\eta)\, d\eta. \qquad (A.6)$$

The resulting integrated expression for $E(z)$ is

$$E(z) = \left\{\frac{\mu_{3z} - \mu_{3z}^*}{(\alpha_{1z})^3}\right\} \left\{\frac{-\rho^6 N_0(z) + \rho^2(\rho^4 + \rho^2 + 1)N_1(z) - (\rho^4 + \rho^2 + 1)N_2(z) + N_3(z)}{\rho^{12}(\rho^2 - 1)^3(\rho^2 + 1)(\rho^4 + \rho^2 + 1)}\right\} \qquad (A.7)$$

where

$\mu_{3z} \equiv$ third central moment p_Z

$\mu_{3z}^* \equiv$ third central moment of p_Z^*

$\alpha_{1z} \equiv$ first moment about the origin of p_Z

$$\rho^2 \equiv \frac{\text{mean of } p_Z^*}{\text{median of } p_Z^*} = \exp\{\ln^2 \sigma_{gz}^*\} \qquad (A.8)$$

$$N_k(z) = \frac{1}{2}\left\{\text{erf}\left[\frac{1}{\sqrt{2}}\left(\frac{\ln z - \ln \mu_{gz}^*}{\ln \sigma_{gz}^*} - k \ln \sigma_{gz}^*\right)\right] + 1\right\},$$

$\mu_{gz}^* \equiv$ geometric mean of p_Z^*

$\sigma_{gz}^* \equiv$ standard geometric deviation of p_Z^*

or in terms of the parameters of the initial second order density function (equations 3 and 4)

$$\frac{\mu_{3z} - \mu_{3z}^*}{(\alpha_{1z})^3} = \tfrac{1}{8}[\exp(\ln^2 \sigma_{gx}) - \exp(r \ln^2 \sigma_{gx})]^3$$

$$\mu_{gz}^* = \mu_{gx}\left(\frac{\exp(\ln^2 \sigma_{gx})}{\tfrac{1}{2}\exp(\ln^2 \sigma_{gx}) + \tfrac{1}{2}\exp(r \ln^2 \sigma_{gx})}\right)^{1/2} \qquad (A.9)$$

$$\sigma_{gz}^* = \exp\{\ln[1/2 \exp(\ln^2 \sigma_{gx}) + 1/2 \exp(r \ln^2 \sigma_{gx})]\}^{1/2}.$$

For the 15 cases mentioned earlier, E was evlauated and compared to the actual difference between F_Z and F_Z^*. The results of four of these cases are given in Table A.1. These results are typical in that they show that (1) for low values of σ_{gx} the difference between F_Z and F_Z^* is small and E is a good estimate of this difference and (2) for relatively high values of σ_{gx} the difference between F_Z and F_Z^* is still small but E underestimates this difference.