Filtering in Distributed Parameter Systems

A Survey Introduction

Curtain, Ruth

1973

Document Version:
Publisher's PDF, also known as Version of record

Link to publication

Citation for published version (APA):
Curtain, R. (1973). *Filtering in Distributed Parameter Systems: A Survey Introduction*. (Research Reports TFRT-3050). Department of Automatic Control, Lund Institute of Technology (LTH).

Total number of authors:
1

LUND UNIVERSITY

PO Box 117
221 00 Lund
+46 46-222 00 00

3050

FILTERING IN DISTRIBUTED PARAMETER SYSTEMS - - A SURVEY INTRODUCTION

RUTH CURTAIN

Report 7301 January 1973
Lund Institute of Technology
Division of Automatic Control

FILTERING IN DISTRIBUTED PARAMETER SYSTEMS - - A SURVEY INTRODUCTION.

Ruth Curtain

TABLE OF CONTENTS

1. INTRODUCTION.

In the early days of Kalman-type filtering, there appeared in the journals much literature about the problem of estimating a state vector $u(T)$ from an observation process $z(t)$ on $0 \leq t \leq T$, assuming a linear dynamical system for $u(t)$ with a "white noise" forcing term, i.e.

$$(1.1)\quad \begin{cases} \dot{u}(t) = A(t)u(t) + B(t)\xi(t) \\ u(0) = u_0 \\ z(t) = C(t)u(t) + \eta(t) \end{cases}$$

where A, B, C are continuous matrix functions and $\xi(t)$, $\eta(t)$ are independent vector Gaussian white noises, i.e.

$$\begin{aligned} E\{\xi(t)\xi(s)^1\} &= \delta(t-s)Q(t) \\ E\{\eta(t)\eta(s)^1\} &= \delta(t-s)R(t) \end{aligned} \qquad (1.2)$$

where Q,R are continuous in t, symmetric and positive matrices and R(t) is invertible.

Using this formal approach, filtering equations for a best least squares estimate can be derived as for the discrete time case. However, the definition of the disturbance is not probabilistically well founded, because $\xi(t)$ is equivalent to the derivative of a Wiener process which simply does not exist. In other words, the state vector $u(t)$ and observation $z(t)$ are not stochastic processes in the probabilistic sense.

All of these problems are solved, if one uses the Itô integral and writes

$$\int_0^t B(t)dw(t)$$

for the "white noise" and defines the stochastic dynamical system in integral form, i.e.

$$(1.3)\quad \begin{cases} du(t) = A(t)u(t)dt + B(t)dw(t) & E\{w(t)w(s)^1\} = \int_0^{\min(t,s)} Q(\tau)d\tau \\ u(0) = u_0 & \\ dz(t) = C(t)u(t)dt + dv(t) & E\{v(t)v(s)^1\} = \int_0^{\min(t,s)} R(\tau)d\tau \end{cases}$$

where w, v are independent vector Wiener processes. The solutions u(t) and z(t) are now well defined vector stochastic processes, and the best least squares estimate of u(t) based on the observations z(s), $0 \leq s \leq t$ is $\hat{u}(t)$ and is given by the familiar Kalman-Bucy equations:

$$(1.4)\quad \begin{cases} d\hat{u}(t) = A(t)\hat{u}(t)dt + P(t)C(t)^{*}R^{-1}(t)\left(dz(t) - C(t)\hat{u}(t)dt\right) \\ \hat{u}(0) = 0 \end{cases}$$

$$(1.5)\quad \begin{cases} \frac{dP}{dt} - A(t)P(t) - P(t)A(t)^{*} + P(t)C(t)^{*}R^{-1}(t)C(t)P(t) = \\ \qquad = B(t)Q(t)B(t)^{*} \\ P(0) = P_0 = \mathrm{Cov}(x_0x_0^1) \end{cases}$$

(P(t) of course is the covariance of the error $u(t) - \hat{u}(t)$.)

In many physical examples, the dynamical system one wishes to estimate is described some partial-integro differential equation; to mention a few of the often quoted ones from the literature:

1. a heat exchanger

2. electrical, optical or acoustic waves which are randomly excited

3. random variations on an electrical transmission line.

It seems to me that these motivations are only secondary, the primary one being a generalization of the Kalman-Bucy filtering theory to distributed parameter systems. As one would expect, the class of distributed parameter system which has been considered is small, namely parabolic partial differential equations with initial data and boundary conditions, so that the dynamical system always has a unique solution which depends continuously on the data. (There are some exceptions to this which will be mentioned later.) As for the finite dimensional case, the tricky mathematical problem is to suitably define the noise disturbances and as for the finite dimensional case, most authors have ignored the existence of this problem and have proceeded formally. So this survey will fall into two parts, namely the formal approaches and the mathematically rigorous ones.

2. FORMAL SOLUTIONS TO THE FILTERING PROBLEM.

2.1. Basic Linear Filtering Model.

Many authors [2], [14], [15], [16], [20], [21], [22], take as their basic system equations, the following model

$$(2.1)\quad \begin{cases} \dfrac{\partial u(x,t)}{\partial t} = A(x,t)u(x,t) + B(x,t)\xi(x,t) \quad \text{on } D\subset R^n \\ \beta_x u(x,t) = 0 \quad \text{on } \partial D \\ u(x,0) = u_0(x) \\ z(x,t) = M(x,t)u(x,t) + \eta(x,t) \end{cases}$$

where the deterministic system with $B \equiv 0$ is a well-found problem of Hadamard, $A(x,t)$ is some spatial, linear, partial differential operator, β_x boundary conditions and $u_0(x)$ the initial data. $B(x,t)$ and $M(x,t)$ are matrix-valued with suitably smooth coefficients and $\xi(x,t)$, $\eta(x,t)$ are independent Gaussian distributed "stochastic processes" with zero expectation and covariances

$$(2.2)\quad \begin{aligned} E\{\xi(x,t)\xi(y,s)^1\} &= Q(x,y,t)\delta(t-s) \\ E\{\eta(x,t)\eta(y,s)^1\} &= R(x,y,t)\delta(t-s) \\ E\{\xi(x,t)\eta(y,s)^1\} &= S(x,y,t)\delta(t-s) \end{aligned}$$

$u_0(x)$ is a distributed Gaussian random variable with zero expectation and covariance matrix $P_0(x,y)$ and is independent of ξ and η. $R(x,y,t)$ is assumed to have an inverse.

Although several authors give less general examples, their ideas and techniques can be generalized to this case.

2.2. Best Least Squares Estimate.

A major contribution was made by Tzafestas and Nightingale [21] - [24] around 1968, 1969. In [21] he assumes that his best estimate $\hat{u}(x,t|t)$ of $u(x,t)$ based on $z(s,x)$ $0 \leqslant s \leqslant t$ has the form

$$\hat{u}(x,t|t) = \int_0^t \int_D H(x,t;y,s)z(y,s)dDds \tag{2.3}$$

and minimizes the error in the sense of least squares. Then he is able to follow the finite dimensional proof to get analogous filtering equations to (1.4) and (1.5),

$$\left.\begin{aligned} &\frac{\partial \hat{u}}{\partial t} = A(x,t)\hat{u} + \int_D H(x,t;y,t)[z(y,t) - M(y,t)\hat{u}(y,t|t)]dD \\ &\text{and} \\ &H(x,t;y,t) = P(x,y,t)M^*(y,t)R^{-1}(x,y,t) \end{aligned}\right\} \quad (2.4)$$

where

$$\begin{aligned} \frac{\partial P}{\partial t}(x,y,t) = {} & A(x,t)P(x,y,t) - P(x,y,t)A^*(y,t) + \\ & + B(x,t)Q(x,y,t)B^*(y,t) - \\ & - \int_D P(y,\alpha,t)M^*(\alpha,t)R^{-1}(y,\alpha,t)M(\alpha,t)P(\alpha,y,t)dD \end{aligned}$$

As these are partial differential equations, one needs boundary and initial conditions

$$(2.5) \quad \begin{cases} \hat{u}(x,0|0) = 0 \\ \beta_x \hat{u}(x,t|t) = 0 & x \in \partial D \end{cases} \quad \begin{cases} P(x,y,0) = P_0(x,y) \\ \beta_x P(x,y,t|t) = 0 = \beta_y P(x,y,t|t) & x,y \in \partial D \end{cases}$$

In [22] he gets the same result by introducing generalized distributed characteristic functions and in [23] he shows that this filter is also the best estimate in the maximum likelihood sense.

Thau [20] in 1969 considered a special case of (2.1) for the random heat equation, assuming his noise disturbances were independent of x and taking a single point observation $z(t,x_1)$. He used the same least squares criterion as Tzafestas and Nightingale and derived similar filter equations. If nothing else, this paper demonstrates the information gap between work in England and in America.

2.3. Minimization of Functional Approach.

Meditch in [14] 1971 considers (2.1) with $B \equiv 0$ and avoids all specification of the random nature of $u_0(x)$ and $\eta(x,t)$; they are simply "errors". He defines his best estimate $\hat{u}(x,T|T)$ to be that which minimizes the following functional

$$(2.6) \qquad J = \frac{1}{2} \int_D \big(u(x,0) - u_0(x)\big)^1 A_{x,y} \big(u(y,0) - u_0(y)\big) dD +$$

$$+ \int_0^T \int_D \big(z(x,t) - M(x,t)u(x,t)\big)^1 R^{-1}(x,t) \cdot$$

$$\cdot \big(z(x,t) - M(x,t)u(x,t)\big) dDdt$$

where

$$A_{x,y}(\cdot) = \Big(\int_D P_0(x,y)(\cdot) dD\Big)^{-1}$$

and P_0, R are suitable weighting matrices, symmetric and positive.

This recasts the filtering problem as the deterministic optimal control problem of minimizing (2.6), which he solves by classical techniques, obtaining a two point boundary value problem (TPBVP) and then decouples to get the usual filtering equations for u and P. He notes that the case $B \neq 0$ can be included by adding an appropriate weighting term in J and he draws the analogy between P_0, R and the covariance matrices of the errors. Although in finite dimensions one can show that minimizing a quadratic functional is equivalent to finding the best least squares estimate, in infinite dimensions, this must be regarded as an "ad hoc" hypothesis. Both Phillipson and Mitter [15] 1967 and Balakrishnan and Lions [2] 1967 use this technique to solve the problem of estimating the control data $u_0(x)$ from noisy observations, nearly avoiding a probabilistic treatment. In fact, Bensoussan [4] does provide a probabilistic interpretation for minimizing a functional to find $\hat{u}$ (see §3).

2.4. Discrete Observations.

Meditch also suggests that in practice discrete observations are more desirable and provides a formal way of including these.

$$(2.7) \qquad z(t,x) = \sum_{i=1}^{k} M(t,s_i)\delta(s_i - x) + \sum_{i=1}^{k} \eta(t,s_i)\delta(t-s_i)$$

Other authors have since included these in their work.

2.5. Innovations Approach.

Atre and Lamba [1] 1971 borrow the innovations approach from Kailath and use it to formally reproduce the filtering equations. (This approach may be fruitful in a rigorous treatment of linear distributed parameter systems.)

2.6. Noise on Boundary.

Sakawa [16] 1972 includes the case of noisy boundary data by assuming a boundary condition of the form

$$F(y)u(t,y) + \frac{\partial u(t,y)}{\partial n_A} = S(t,y)v(t,y) \qquad y \in \partial\Omega,$$

where F, S are smooth matrices, $\partial/\partial n_A$ is the normal spatial derivative corresponding to A and $v(t,y)$ is Gaussian white noise independent of ξ and η. He also assumes discrete observations. However, he readily reduces the problem to one equivalent to (2.1) by adding the term $\delta(\psi(x)) \cdot S(t,x)v(t,x)$ to the state equation, where $\delta(\psi(x))$ is a generalized function concentrated on ∂D. Solutions of (2.1) are taken in the generalized sense and the usual formal analysis yields the generalized Kalman-Bucy filter equations.

2.7. Colored Noise.

Tzafestas in a recent paper [24], 1972, has generalized his filtering results to include coloured noise in the observation, i.e.

$$
(2.8)\quad \begin{cases}
z(x,t) = M(x,t)u(x,t) + N(x,t)\Omega(x,t) \\
\text{with} \\
\dfrac{\partial\Omega(x,t)}{\partial t} = B_x\Omega(x,t) + D(x,t)\eta(x,t) \\
\Omega(x,0) = \Omega_0(x) \\
\gamma_x\Omega(x,t) = 0 \quad \text{on } \partial D
\end{cases}
$$

where η is a distributed Gaussian white noise as before and the coloured noise Ω is the output of the randomly forced linear partial differential equation. In essence his solution amounts to transforming this system into one equivalent to (2.1).

2.8. Solution of the Filtering Equations.

Several of the authors included examples of solutions of (2.4), (2.5) in their papers ([14], [20], [21] - [24]) and all were obtained by looking for a separation of variables solution using eigenfunction expansions for $\hat{u}$ and P, where the eigenfunctions are those defined by the deterministic version of (2.1). Sakawa [16] does a general eigenfunction expansion, obtaining Kalman-Bucy equations for an infinite estimation vector $\left(\hat{u}_1(t), \hat{u}_2(t), \ldots\right)$ and an infinite covariance matrix $\left(P_{ij}(t)\right)$, $i = 1, \ldots \infty$, $j = 1, \ldots \infty$. Naturally these look just like the finite dimensional Kalman-Bucy equations

However, these examples are purely illustrative and do not represent any general theory. For a theory based on approximation on finite dimensional subspaces, see Bensoussan [5] and a thesis to appear by J.C. Nedelec (I.R.I.A).

2.9. Other Extensions.

Some authors have included partial-integro differential operators in A(x,t), for example [22]. In fact all the theories can be easily generalized to include integral operators explicitly, as they are just bounded operators. Using a formal approach it is possible to generalize most results from lumped parameter to distributed parameter systems. For example, Shukla and Srinath [19] give results for optimal filtering in a linear, distributed parameter system with time delays.

2.10. Smoothing and Estimation.

Although this report concerns the filtering problem, I should note that similar techniques to the above ones are applicable to the smoothing and estimation problem and several authors have exploited these (see [1], [2], [14], [15], [21], [23]. [24]).

2.11. Nonlinear Filtering.

The most recent papers in this field consider the filtering problem when the state equations are nonlinear. Tzafestas and Nightingale in [23] 1969 used a maximum likelihood approach to find a best filter for (2.1) where $A(x,t)$, β_x, $M(x,t)$ are allowed to be nonlinear differential or integro-differential operators. He uses a generalization of a likelihood functional to infinite dimensional Gaussian random variables, and shows that the best maximum likelihood estimate is one which minimizes a functional

$$J = \frac{1}{2} \int_{D_1} \int_{D_2} \left(u(x,0) - u_0\right)^1 P_0(x,y) \cdot \tag{2.9}$$

$$\cdot \left(u(y,0) - u_0(y)\right) dD_1 dD_2 +$$

$$+ \frac{1}{2} \int_0^t \int_{D_1} \int_{D_2} \left(z(t,x) - M(x,t)u(x,t)\right)^1 R^{-1}(x,y,t) \cdot$$

$$\cdot \left(z(t,y) - M(y,t)u(y,t)\right) dD_1 dD_2 dt +$$

$$+ \frac{1}{2} \int_0^t \int_{D_1} \int_{D_2} \left[\frac{\partial u(x,t)}{\partial t} - A(x,t,u)u\right]^1$$

$$Q^{-1}(x,y,t) \left[\frac{\partial u(y,t)}{\partial t} - A(y,t,u)u(t,y)\right] dD_1 dD_2 dt$$

subject to $\beta_x(u,x,t) = 0$ on ∂D and Q is actually the covariance matrix of $B(x,t)\xi(x,t)$

So this provides some justification for the functional approach. Seinfeld [17], [18] 1969, 1971, Hwang, Seinfeld and Gavalas [11] 1972 and Lamont and Kumar [12] 1972 all treat various nonlinear dynamical systems and add in noise as desired and then write down some massive functional which includes a term for every noise component in the system. They define their best estimate to be the solution of this deterministic minimization problem and then proceed to solve it as a classical calculus of variations problem. This yields a TPBVP which they embed in a wider class of problems (i.e. "invariant embedding") to yield a recursive Kalman-Bucy type filter. The mathematics is quite formal and there is even some difference in the choice of cost functionals by various authors. Cf. [12] and [10].

3. MATHEMATICALLY RIGOROUS APPROACHES.

An early paper which posed an infinite dimensional filtering problem was that by Falb [9] in 1967. The idea was to generalize the Itô integral to a Hilbert space and so to define a stochastic evolution equation, the analogue of (1.3) in a Hilbert space context, and then to proceed as for the finite dimensional case. The main difficulty here is that for partial differential equations, the operator A(t) is a differential operator and hence unbounded. In [9], Falb only considered A(t) bounded which would cover integral equations, but in [7], Curtain and Falb gave an existence theorem for a stochastic evolution equation with unbounded A(t), which could be applied fruitfully to the filtering problem.

Kushner [11] 1970 considers the filtering problem for two classes of distributed systems. The first is a second order parabolic partial differential equation with an integral operator in the state equation and Dirichelet-type boundary conditions. The observation is a weighted average of the state vector over space and time and the disturbance processes are Itô differentials of Wiener processes in time only, i.e. he considers the degenerate case where the noise is uniformly distributed in space. Taking his best estimate to be the conditional expectation of the state, given the observation process, he derives filtering equations for this estimate and for the covariance of the error, rigorously establishing existence, uniqueness of the solutions. The second problem is a second order parabolic partial differential equation with mixed boundary conditions. Again the disturbance processes are only random in time, but now he allows noise on the boundary as well as in the state equation. Taking his observations averaged over the boundary this time, he does a similar analysis to the first problem.

Duncan [8] 1972 generalizes filtering results to Banach valued stochastic processes, but as the operators are bounded it does not apply to partial differential equations.

The rest of this survey will be concerned with the work of Bensoussan and for a complete exposition, the reader is referred to his recent book [4] 1971. He has solved the problem of filtering for a very wide class of linear distributed parameter systems, namely

$$(3.1)\quad \begin{cases} \dot{u}(t) + A(t)u = f(t) + B(t)\xi(t) & 0 \leq t \leq T \\ u(0) = u_0 + \rho & \end{cases}$$
$$z(t) = C(t)u(t) + \eta(t)$$

where the state $u(t)$ has values in a Hilbert space H, the observation $z(t)$ in another Hilbert space F, the operators $B(t)$, $C(t)$ are bounded and linear, but $A(t)$ is linear and closed, i.e. a differential operator. If we consider (3.1) as a deterministic system, then Lions [13] has developed a very nice theory for the solution of control problems of this form, and Bensoussan uses this format. The random components are $\xi(t)$, ρ and $\eta(t)$ and have values in some Hilbert space, but instead of treating them as random variables or stochastic processes in a Hilbert space, he introduces a more general concept of a linear random functional or equivalently a cylindrical probability. Really he treats them as "generalized random variables" in a manner analogous to the use of generalized functions in partial differential equations.

The class of second order linear random functionals (LRF) which have uniquely defined expectations and covariance operators are closed under affine continuous maps. (3.1) defines $u(t)$ and $z(\cdot)$ as affine, continuous maps from

$(\zeta,\xi(\cdot),\ \eta(\cdot))$ using appropriate spaces and so if we assign a LRF associated with $(\zeta,\ \xi(\cdot),\ \eta(\cdot))$, this induces LRF's associated with $u(T)$ and $z(\cdot)$. The estimation problem is then defined as the best linear estimate of the LRF associated with $u(T)$ by the LRF associated with $z(\cdot)$. Using results from estimation theory in a Hilbert space, Bensoussan reduces the problem to the deterministic problem of minimizing the functional*

$$(3.2)\quad J(\zeta,\ \xi(\cdot)) = \langle P_0^{-1}\zeta,\zeta\rangle_H + \int_0^T \langle Q(t)^{-1}\xi(t),\ \xi(t)\rangle_E dt$$

$$+\int_0^T \langle R^{-1}(t)z_d(t)-C(t)u(t;\zeta,\xi)\ ,\ z_d(t)-$$

$$- C(t)\ u(t;\ \zeta,\xi)\ \rangle_F dt$$

with respect to ζ, $\xi(\cdot)$, where P_0, $Q(t)$, $R(t)$ are the covariance operators for the LRF corresponding to ζ, $\xi(t)$ and $\eta(t)$ respectively. P_0, $Q(t)$, $R(t)$ are self adjoint and positive and invertible.

The best estimate is $\hat{\omega}_T = u(T;\hat{\zeta},\hat{\xi}(\cdot))$, where $(\hat{\zeta},\ \hat{\xi}(\cdot))$ is the minimizing pair.*

He also derives the Kalman - Bucy equations for $\hat{\omega}_r$.

$$(3.3)\quad \frac{dr}{dt} + A(t)\ r(t)+P(t)C^*(t)R^{-1}(t)\ C(t)\ r(t) =$$

$$g(t) + P(t)\ C^*(t)R^{-1}(t)z_d(t)$$

$$r(0) = u_0$$

---

* This gives some insight into the justification for defining the best estimate a priori in terms of minimizing a functional. Cf. 2.3.

$$(3.4)\quad \begin{cases} \dfrac{dP(t)}{dt} + P(t)A^{*}(t) + A(t)P(t) + \\ \qquad + P(t)C(t)^{*}R^{-1}(t)C(t)P(t) = \\ \quad = B(t)\phi(t)B^{*}(t)\phi(t) \\ P(0) = P_0 \end{cases}$$

(actually only a weaker version of the P(t) equation holds.)

and $\hat{u}_T = r(T)$.

Finally he shows that (3.3), (3.4) define the best LRF estimate for the general case where P_0 and $Q(t)$ are not invertible. Applying this result to the special equation (2.1), one finds agreement between both filtering equations. If one applies these results to the finite dimensional case, then (3.3), (3.4) correspond to the filtering equations one obtains using the formal approach of (1.1). So this LRF approach gives new insight into the finite dimensional filtering problem as well.

Note that (3.1) is not a stochastic differential equation in the strict sense and that u(t) is not a well-defined stochastic process, but has some linear random functional identified with it. For a stochastic evolution equation approach, Bensoussan defines a Hilbert space-valued Wiener process and stochastic integration in a Hilbert space along the lines of Falb [9]. Then he considers the following system model

$$
(3.5)\quad \begin{cases} u(t,\omega) + \int_0^t A(\tau)u(\tau,\omega)d\tau = \\ \qquad = u_0(\omega) + \int_0^t f(\tau)d\tau + \int_0^t B(t)d\xi(t) \\ z(t,\omega) = \int_0^t C(\tau)u(\tau,\omega)d\tau + \eta(t,\omega) \end{cases}
$$

where $\xi(t)$, $\eta(t)$ are Hilbert space valued independent Wiener processes with covariance operators $Q(t)$, $R(t)$ respectively and $u_0(\omega)$ is a Gaussian stochastic process independent of $\xi(t)$ and $\eta(t)$ with zero expectation and covariance P_0.

This theory requires that P_0, $Q(t)$, $R(t)$ be nuclear operators (as well as positive and self adjoint). For filtering, one requires $R(t)$ to be invertible and so F must be finite dimensional (i.e. a finite dimensional observation space).

Under these conditions, Bensoussan proves that the best global estimate for the random variable $u(T,\omega)$ based on the stochastic process $z(t,\omega)$; $0 \leq t \leq T$ is given by the solution of the stochastic differential equation

$$
(3.6)\quad \hat{u}(T,\omega) + \int_0^T A(t)\hat{u}(t)dt + \int_0^T P(t)C^*(t)R^{-1}(t)C(t)\hat{u}(t)dt =
$$

$$
= \int_0^T f(t)dt + \int_0^T P(t)C^*(t)R^{-1}(t)dz(t) + u_0(\omega)
$$

where $P(t)$ is given by (3.4).

This is the required generalization of the Kalman-Bucy filter equations (1.4), (1.5) to infinite dimensions.

Bensoussan has also entended these results to systems with second order in t to obtain similar results. In [6] he has looked at the problem of optimization of location of sensors when you have discrete observations and he has proved a separation principle for distributed parameter systems in [5]. The reader is referred to [3] for his treatment of the estimation problem.

REFERENCES.

[1] Atre, S.R., and Lamba, S.S.: Optimal Estimation in Distributed Processes Using Innovations Approach. Proc. IFAC Banf, Canada.

[2] Balakrishnan and Lions: State Estimation for Infinite Dimensional Systems. J. Comp. Syst. Sci. 1967 p. 391 - 403.

[3] Bensoussan, A.: L'Identification et le Filtrage. I.R.I.A 1969 (Thesis).

[4] Bensoussan, A.: Filtrage Optimal des Systèmes Linéaires (Dunod 1971).

[5] Bensoussan, A., Bossavit, A., Nedelec, J.C.: Approximation des problèmes de Contrôle. Cahiers de l'Iria, No. 2.

[6] Bensoussan, A.: Optimization of Sensors' Location in a Distributed Filtering Problem (Proceedings on Stability of Stochastic Dynamical Systems 1972. Lecture Notes in Mathematics No. 294 Springer).

[7] Curtain and Falb: Stochastic Differential Equations in Hilbert Space, J. Diff. Eqns. Vol. 10, No. 3.

[8] Duncan, T.L.: Some Banach valued processes with applications (Proc. on Stability of Stochastic Dynamical Systems 1972. Lecture Notes in Mathematics No. 294 Springer).

[9] Falb, P.: Infinite Dimensional Filtering, Inf. & Control, 1967, 102 - 137.

[10] Hwang, M., Seinfeld, J.H., and Gavalas, G.R.: Optimal Least Squares Filtering & Interpolation in Distributed Parameter Systems. J. Math. An. & Appl. 39, 49 - 74.

[11] Kushner, H.J.: Filtering for Linear Distributed Parameter Systems, SIAM. J. Control, (1970) 346 - 359.

[12] Lamont and Kumar: State Estimation in distributed parameter systems via least squares and invariant embedding. J. Math. Anal. & Appl. 38, 588 - 606.

[13] Lions, J.L.: Contrôle Optimal de Systèmes gouverné par des équations aux dérivées partielles (Dunod 1968).

[14] Meditch, J.S.: Least Squares Filtering and Smoothing for linear distributed parameter systems. Automatica, 1971 Vol. 7, p. 315 - 322.

[15] Phillipson and Mitter, S.K.: State Identification of a class of linear distributed Systems. PROC. 4th IFAC Congress (Warsaw) 1969.

[16] Sakawa, Y.: Optimal Filtering in Linear Distributed-Parameter Systems. Int. J. Control 1972 p 115 - - 127.

[17] Seinfeld, J.H.: Nonlinear Estimation for partial differential equations. Chem. Eng. Sci. (1969) 75 - 83.

[18] Seinfeld, J.H.: J. Dynamical Systems, Measurement & Control G93 1971 157 - 163.

[19] Shukla, V., and Srinath, M.D.: Optimal Filtering in Linear Distributed Parameter Systems with Multiple Time Delays. Int. J. Control, 1972, 16, 4 p 673 - - 687.

[20] Thau, F.E.: On Optimal Filtering for a class of linear distributed parameter systems. J. Basic Eng. 1969 p 173 - 178.

[21] Tzafestas and Nightingale: Optimal Filtering, smoothing and prediction in linear distributed parameter systems, PROC. IEE 1968, 1207 - 1212.

[22] Tzafestas and Nightingale: Concerning Optimal filtering of linear distributed parameter systems. PROC. IEE 1968, 1737 - 1742.

[23] Tzafestas, S.G. and Nightingale: Maximum Likelihood Approach to the Optimal Filtering of distributed parameter systems. PROC. IEE, 1969, p 1085 - 1093.

[24] Tzafestas, S.G.: On Optimum Distributed Parameter Filtering and Fixed Interval Smoothing for Colored Noise, IEE Trans. AC-17, No. 4 p 448 - 458.