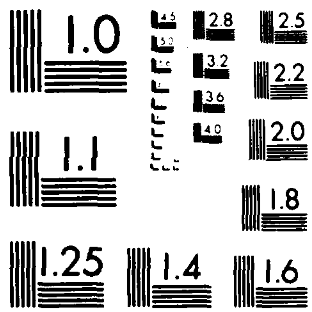
AD-A135 092 ESTIMATION TECHNIQUES FOR TRANSPORT EQUATIONS(U) BROWN UNIV PROVIDENCE RI LEFSCHETZ CENTER FOR DYNAMICAL SYSTEMS H T BANKS ET AL. JUL 83 LCDS-83-23 1/1

UNCLASSIFIED AFOSR-TR-83-0906 AFOSR-81-0198 F/G 12/1 NL

END
DATE
FILMED
12-83
DTIC



MICROCOPY RESOLUTION TEST CHART
NATIONAL BUREAU OF STANDARDS 1963-A

AD-A135092

2

AFOSR-TR- 83-0906

Lefschetz Center for Dynamical Systems

Division of Applied Mathematics Brown University Providence RI 02912

83 11 29 212

2

Estimation Techniques For Transport Equations

by

H. T. Banks, P. L. Daniel, P. Kareiva

July 1983 LCDS #83-23

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR-TR- 83-0906	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) ESTIMATION TECHNIQUES FOR TRANSPORT EQUATIONS		5. TYPE OF REPORT & PERIOD COVERED TECHNICAL
		6. PERFORMING ORG. REPORT NUMBER LCDS #83-23
7. AUTHOR(s) H.T. Banks, P.L. Daniel and P. Kareiva*		8. CONTRACT OR GRANT NUMBER(s) AFOSR-81-0198
9. PERFORMING ORGANIZATION NAME AND ADDRESS Division of Applied Mathematics Brown University Providence RI 02912		10. PROGRAM ELEMENT, PROJECT, TASK AREA & WORK UNIT NUMBERS PE61102F; 2304/A1
11. CONTROLLING OFFICE NAME AND ADDRESS AFOSR/NM BOLLING AFB, DC 20332		12. REPORT DATE JUL 83
		13. NUMBER OF PAGES 12
14. MONITORING AGENCY NAME & ADDRESS(if different from Controlling Office)		15. SECURITY CLASS. (of this report) UNCLASSIFIED
		15a. DECLASSIFICATION/DOWNGRADING SCHEDULE

16. DISTRIBUTION STATEMENT (of this Report)

Approved for public release; distribution unlimited.

17. DISTRIBUTION STATEMENT (of the abstract entered in Block 20, if different from Report)

18. SUPPLEMENTARY NOTES

*P.L. Daniel, Department of Mathematics, Southern Methodist Univ, Dallas TX 75275; P. Kareiva, Division of Biology and Medicine, Brown University, Providence RI 02912.

19. KEY WORDS (Continue on reverse side if necessary and identify by block number)

R sub 2,

20. ABSTRACT (Continue on reverse side if necessary and identify by block number)

The authors present convergence arguments for algorithms developed to estimate spatially and/or time dependent coefficients and boundary parameters in general transport (diffusion, advection, sink/source) models in a bounded domain $\Omega \subset R^2$. A brief summary of numerical results obtained using the algorithms is given.

DD FORM 1 JAN 73 1473 EDITION OF 1 NOV 65 IS OBSOLETE

UNCLASSIFIED
SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

ESTIMATION TECHNIQUES FOR TRANSPORT EQUATIONS

H. T. BANKS
Div. of Applied Math
Brown University
Providence, RI 02912

P. L. DANIEL
Dept. of Mathematics
Southern Methodist Univ.
Dallas, Texas 75275

P. KAREIVA
Div. of Biology and Medicine
Brown University
Providence, RI 02912

Abstract. We present convergence arguments for algorithms developed to estimate spatially and/or time dependent coefficients and boundary parameters in general transport (diffusion, advection, sink/source) models in a bounded domain $\Omega \subset R^2$. A brief summary of numerical results obtained using the algorithms is given.

I. Introduction. In this note we present theoretical results for estimation of function space (i.e., time and spatially varying) parameters in general transport equations. The presentation here is motivated by our own efforts on problems in transport of labeled substances in brain tissue [3], population dispersal (in particular insect movement--see [2], [3], [7]), and bioturbation [6], among other applications in the biological sciences. Due to limitations in space, we shall not discuss here any of those particular efforts. Rather we provide an outline of a general convergence theory for a class of approximation schemes that we have used and are continuing to use successfully in a number of biological applications. In the first two sections we present for the first time general theoretical arguments underlying these approximation schemes; our use of the methods in specific problems is discussed elsewhere ([2], [3]). In a final section we summarize briefly one aspect of the numerical performance of our methods that is pertinent to any convergence theory that one might develop.

As our fundamental state system we consider the scalar equation

$$(1) \quad \frac{\partial u}{\partial t} + \nabla \cdot (\mathfrak{v} u) = \nabla \cdot (\mathcal{D} \otimes \nabla u) + \alpha u + f, \qquad t \in (0,T] ,$$

on the bounded domain $\Omega \subset R^2$ with boundary conditions $u(t,\cdot)|_{\partial\Omega} = 0$ and given initial conditions $u(0,\cdot)|_{\Omega} = u_0(\gamma)$. Here we assume that $\mathfrak{v}$ and $\mathcal{D}$ are functions depending on (t,x,y), $t \geq 0$, $(x,y) \in \Omega$ and $f = f(\beta,t,x,y)$, $\beta = \beta(t,x,y)$, $\alpha = \alpha(t,x,y)$, $\gamma = \gamma(x,y)$. While we treat only trivial Dirichlet boundary conditions in this note, our ideas are sufficiently general to allow nontrivial boundary conditions depending possibly on unknown parameters. We assume that any such boundary conditions have been transformed in the usual manner so that the unknown boundary parameters are included in the vector parameters β and γ in f and u_0 above. We have also simplified our exposition in treating only a scalar equation even though our methods are applicable to (and have been used with) vector systems.

Along with the state equation (1) we assume that we have observations $\hat{u}_i \in H^0(\Omega)$ for $u(t_i,\cdot)$, or $\hat{u}_{ijk} \in R^1$ for $u(t_i,x_j,y_k)$ and that we wish to choose the parameter functions $\mathfrak{D}$, D, α, β, γ so that the corresponding solution of (1) best fits the observations. For our discussions here we shall assume that this problem is posed in terms of an optimization problem for a least squares fit-to-data criterion. Letting $q = (D,\mathfrak{D},\alpha,\beta,\gamma)$ represent the set of unknown parameters and Q represent the class of admissible parameter functions, we denote by $q \to J(q)$ the least squares criterion function. For the observations mentioned above, this function is given by

$$(2) \qquad J(q) = \sum_i |u(t_i,q) - \hat{u}_i|^2$$

in the case of distributed data $\hat{u}_i \in H^0(\Omega)$, and

$$(3) \qquad J(q) = \sum_{i,j,k} |u(t_i,x_j,y_k,q) - \hat{u}_{ijk}|^2$$

in the case of pointwise data $\hat{u}_{ijk} \in R^1$, where in both cases $u(\cdot,q)$ is the solution of (1) for a given $q = (D,\mathfrak{D},\alpha,\beta,\gamma)$. In either case our basic problem consists of minimizing J over Q.

This problem is difficult in part because it is, in general, infinite dimensional in both the state u and the parameters q, each of which lies in a function space. Therefore algorithms for its solution will, in most cases, involve two separate and often unrelated approximation ideas, one for the state space and one for the parameter set. In the next section we consider these approximations and outline convergence arguments that indicate that the schemes can yield useful computational results.

<u>II. Convergence Results</u>. We first rewrite equation (1) in its weak form in the state space $H = H^0(\Omega)$ with the usual inner product. We have, dropping the Kronecker product sign $\otimes$ for ease in notation, that for all $\phi \in H_0^1(\Omega)$, the weak solution u must satisfy

$$(4) \qquad \begin{aligned} &\langle u_t,\phi\rangle + \langle D\nabla u(t),\nabla\phi\rangle - \langle \mathfrak{D}u(t),\nabla\phi\rangle - \langle \alpha u(t) + f,\phi\rangle = 0, \qquad t \in (0,T], \\ &u(0) = u_0(\gamma). \end{aligned}$$

Retaining the notation $q = (D,\mathfrak{D},\alpha,\beta,\gamma)$ we make the standing assumptions on the parameter set Q:

$\boxed{A}$ The set Q is a bounded subset of $L_\infty([0,T] \times \Omega)$.

$\boxed{B}$ There exists a positive constant m such that for every $q = (\mathcal{D},\mathcal{V},\alpha,\beta,\gamma)$ in Q, $\mathcal{D}_i(t,x,y) \geq m$ for $i = 1, 2$, and (t,x,y) in $[0,T] \times \Omega$.

We also assume throughout that the perturbation function f is C^1 in all of its arguments and that $\gamma \to u_0(\gamma)$ is continuous from $H^0(\Omega)$ to $H^0(\Omega)$. This will suffice for most of the results stated below, but the regularity of solutions u assumed later can be guaranteed only under more stringent smoothness assumptions on f along with conditions relating f to the initial data u_0.

Under our standing assumptions on Q and f, we can, further assuming that $u_0 \in H^0(\Omega)$, use rather standard arguments (e.g., see [10, p. 104]) to guarantee existence and uniqueness of solutions u to (4) with $u(t,q) \in H_0^1(\Omega)$.

For the state space approximation of (4), we consider a Galerkin scheme on finite dimensional subspaces H^N of $H^0(\Omega)$. We assume $H^N \subset H_0^1(\Omega)$ for each $N = 1, 2, \ldots$, and define the Galerkin approximation for a given $q \in Q$ as the solution u^N, $u^N(t) \in H^N$, of the equations

$$(5)\qquad \begin{aligned} &\langle u_t^N,\psi\rangle + \langle \mathcal{D}\nabla u^N(t),\nabla\psi\rangle - \langle \mathcal{V}u^N(t),\nabla\psi\rangle - \langle \alpha u^N(t) + f,\psi\rangle = 0, \qquad \psi \in H^N, \\ &u^N(0) = P^N u_0(\gamma), \end{aligned}$$

where P^N is the orthogonal projection of $H^0(\Omega)$ onto H^N.

It is useful to define the bilinear form $\mathcal{L}(q) : H_0^1(\Omega) \times H_0^1(\Omega) \to R^1$ by

$$\begin{aligned} \mathcal{L}(q)(\psi,\phi) &= \langle \mathcal{D}\nabla\psi,\nabla\phi\rangle - \langle \mathcal{V}\psi,\nabla\phi\rangle - \langle \alpha\psi,\phi\rangle \\ &\equiv a(q)(\psi,\phi) - b(q)(\psi,\phi) - \langle \alpha\psi,\phi\rangle . \end{aligned}$$

Then we may rewrite the original equation and its Galerkin approximation in H^N as

$$(6)\qquad \begin{aligned} &\langle u_t,\phi\rangle + \mathcal{L}(q)(u,\phi) = \langle f,\phi\rangle, \qquad \forall\phi \in H_0^1(\Omega), \\ &u(0) = u_0(q) \end{aligned}$$

and

$$(7)\qquad \begin{aligned} &\langle u_t^N,\psi\rangle + \mathcal{L}(q)(u^N,\psi) = \langle f,\psi\rangle, \qquad \forall\psi \in H^N, \\ &u^N(0) = P^N u_0(q). \end{aligned}$$

Defining, for solutions u^N of (7), the approximate fit-to-data function J^N (corresponding to (2)) by

(8) $J^N(q) = \sum_i |u^N(t_i,q) - \hat{u}_i|^2$,

we find that our estimation problems with approximate states (which are still optimization problems over an infinite dimensional function set Q) consist of minimizing J^N over Q. Before turning to a second level of approximation (for the parameter set) we give some convergence results for these approximate problems that will prove useful in discussing the state-and-parameter approximation problems. We make the following standing assumptions on the *approximation properties of* H^N relative to H.

$\boxed{C}$ Let P^N denote the orthogonal projection of $H^0(\Omega)$ onto H^N. Then for any $\psi \in C^2(\Omega) \cap H_0^1(\Omega)$ the following estimates hold: $|P^N\psi - \psi|^2, |(P^N\psi - \psi)_x|^2$, $|(P^N\psi - \psi)_y|^2$ are each dominated by some functional $g(N,\psi)$ satisfying $|g(N,\psi)| \leq \varepsilon(N) \{|\psi_{xx}|^2 + |\psi_{yy}|^2\}$ where $\varepsilon(N) \to 0$ as $N \to \infty$.

We remark that for $\Omega = (0,1) \times (0,1)$, tensor products of the subspaces $L_0(\Delta^N)$ and $S_0(\Delta^N)$ of linear and cubic splines (corresponding to a grid size 1/N) modified to satisfy homogeneous boundary conditions, are readily seen to satisfy the condition C (see [11, Chap. 6] for further discussion and details). The following fundamental convergence statement (e.g., see [2], [3]) is most helpful in establishing the desired approximation theorems.

Theorem 1. Suppose $q^N, q^* \in Q$ with $q^N \to q^*$ in $H^0((0,T) \times \Omega)$. Suppose further that $u(q^*) \in C^2(\overline{(0,T) \times \Omega})$. Then $u^N(t,q^N) \to u(t,q^*)$ in $H^0(\Omega)$ for each $t \in [0,T]$, where u, u^N are solutions of (6), (7), respectively.

Since this theorem is fundamental to our discussions in this note, we shall outline the essential steps of its proof. First, we note that under assumptions A and B above, it is easy to establish the following Gårding inequality (e.g., see also [9, p. 144], [4, p. 34]): There are positive constants c_0, c_1 depending on Ω, the bounds for Q, and m such that $\mathcal{L}(q)(\phi,\phi) \geq c_1|\phi|_1^2 - c_0|\phi|^2$ for all $\phi \in H_0^1(\Omega)$ and all $q \in Q$.

Let u^N, u^* be the solutions of (7), (6) corresponding to q^N, q^*, respectively. We wish to argue that $|u^N(t) - u^*(t)| \to 0$. But from the inequality $|u^N - u^*| \leq |u^N - P^N u^*| + |P^N u^* - u^*|$ (to simplify notation here and throughout we suppress the dependence on t) and the approximation results of condition C, it suffices to argue $|u^N - P^N u^*| \to 0$.

Let f^N denote f at β^N of q^N and note that the convergence hypotheses on q^N and the smoothness of f imply $|f^N - f^*| \to 0$ where f^* corresponds to q^*. We have that u^N, u^* satisfy the equations (see (6), (7))

$$\text{(9)} \qquad \begin{aligned} &\langle u_t^N, \psi\rangle + \mathcal{L}(q^N)(u^N, \psi) = \langle f^N, \psi\rangle \qquad \text{for all} \quad \psi \in H^N, \\ &u^N(0) = P^N u_0(q^N) \end{aligned}$$

and

$$\text{(10)} \qquad \begin{aligned} &\langle u_t^*, \phi\rangle + \mathcal{L}(q^*)(u^*, \phi) = \langle f^*, \phi\rangle \qquad \text{for all} \quad \phi \in H_0^1(\Omega)\ , \\ &u^*(0) = u_0(q^*). \end{aligned}$$

Since $H^N \subset H_0^1(\Omega)$, letting D_t denote $\frac{\partial}{\partial t}$, we have for all $\psi \in H^N$

$$\begin{aligned} &\langle D_t(u^N - P^N u^*), \psi\rangle + \mathcal{L}(q^N)(u^N - P^N u^*, \psi) \\ &\quad = - \langle D_t P^N u^*, \psi\rangle - \mathcal{L}(q^N)(P^N u^*, \psi) + \langle f^N, \psi\rangle \\ &\quad = \langle D_t(u^* - P^N u^*), \psi\rangle + \mathcal{L}(q^*)(u^*, \psi) - \mathcal{L}(q^N)(P^N u^*, \psi) + \langle f^N - f^*, \psi\rangle \end{aligned}$$

where we have used (9) and (10). Choosing $\psi = z^N = u^N - P^N u^*$ in H^N we thus find using the above identity

$$\begin{aligned} \frac{1}{2} D_t |z^N|^2 + \mathcal{L}(q^N)(z^N, z^N) &= \langle D_t(u^* - P^N u^*), z^N\rangle \\ &\quad + \mathcal{L}(q^*)(u^*, z^N) - \mathcal{L}(q^N)(P^N u^*, z^N) + \langle f^N - f^*, z^N\rangle\ . \end{aligned}$$

Use of the Gårding inequality then yields

$$\begin{aligned} \frac{1}{2} D_t |z^N|^2 + c_1 |z^N|_1^2 - c_0 |z^N|^2 &\le \langle D_t(u^* - P^N u^*), z^N\rangle \\ &\quad + \mathcal{L}(q^*)(u^*, z^N) - \mathcal{L}(q^N)(P^N u^*, z^N) + \langle f^N - f^*, z^N\rangle \\ \text{(11)} \qquad &\le \frac{1}{2} |u_t^* - P^N u_t^*|^2 + \frac{1}{2} |z^N|^2 + a(q^*)(u^*, z^N) - a(q^N)(P^N u^*, z^N) \\ &\quad + b(q^N)(P^N u^*, z^N) - b(q^*)(u^*, z^N) \\ &\quad + \langle \alpha^N P^N u^* - \alpha^* u^*, z^N\rangle + \langle f^N - f^*, z^N\rangle\ . \end{aligned}$$

But

$$a(q^*)(u^*,z^N) - a(q^N)(P^N u^*,z^N) = \langle \mathcal{D}^* \nabla u^* - \mathcal{D}^N \nabla P^N u^*, \nabla z^N \rangle$$

$$\le \frac{1}{2}\frac{2}{c_1}|\mathcal{D}^*\nabla u^* - \mathcal{D}^N \nabla P^N u^*|^2 + \frac{1}{2}\frac{c_1}{2}|\nabla z^N|^2$$

while

$$b(q^N)(P^N u^*,z^N) - b(q^*)(u^*,z^N) = \langle \mathcal{V}^N P^N u^* - \mathcal{V}^* u^*, \nabla z^N \rangle$$

$$\le \frac{1}{2}\ \frac{2}{c_1}|\mathcal{V}^N P^N u^* - \mathcal{V}^* u^*|^2 + \frac{1}{2}\ \frac{c_1}{2}|\nabla z^N|^2 .$$

Thus, using these estimates in (11) we find

$$\frac{1}{2}D_t|z^N|^2 + c_1|z^N|_1^2 - c_0|z^N|^2$$

$$\le \frac{1}{2}|u_t^* - P^N u_t^*|^2 + \frac{1}{2}|z^N|^2 + \frac{1}{c_1}|\mathcal{D}^*\nabla u^* - \mathcal{D}^N\nabla P^N u^*|^2$$

$$+ \frac{1}{c_1}|\mathcal{V}^N P^N u^* - \mathcal{V}^* u^*|^2 + \frac{c_1}{2}|\nabla z^N|^2$$

$$+ \frac{1}{2}|\alpha^N P^N u^* - \alpha^* u^*|^2 + \frac{1}{2}|z^N|^2 + \frac{1}{2}|f^N - f^*|^2 + \frac{1}{2}|z^N|^2 ,$$

or, since $|z^N|_1^2 \ge |\nabla z^N|^2$,

$$\text{(12)}\quad \frac{1}{2}D_t|z^N|^2 + (-\frac{3}{2} - c_0)\,|z^N|^2 \le h^N$$

where

$$h^N \equiv \frac{1}{2}|u_t^* - P^N u_t^*|^2 + \frac{1}{c_1}|\mathcal{D}^*\nabla u^* - \mathcal{D}^N\nabla P^N u^*|^2$$

$$+ \frac{1}{c_1}|\mathcal{V}^N P^N u^* - \mathcal{V}^* u^*|^2 + \frac{1}{2}|\alpha^N P^N u^* - \alpha^* u^*|^2 + \frac{1}{2}|f^N - f^*|^2 .$$

Using the Gronwall inequality and defining $\varepsilon^N = |u^N(0) - P^N u^*(0)|$, we may obtain from (12) the estimate

$$|u^N(t) - P^N u^*(t)| \le \varepsilon^N + e^{2\delta t}\int_0^t e^{-2\delta s}\, 2h^N(s)ds$$

(13)

$$\le \varepsilon^N + 2e^{2\delta t}\,(T\int_0^T |h^N(s)|^2 ds)^{\frac{1}{2}}$$

where $\delta = c_0 + \frac{3}{2}$ is independent of N. Since $\epsilon^N \to 0$ follows from the continuity assumption on u_0, for the desired convergence it suffices to argue that $h^N \to 0$ in $H^0((0,T) \times \Omega)$. However, using the convergence $q^N \to q^*$, the assumption that u^* is in $C^2(\overline{(0,T) \times \Omega})$, the estimates of condition C, and the bounds from condition A, this convergence is readily established. Thus is the statement of Theorem 1 proved.

We remark that the regularity required of $u(q^*)$ in Theorem 1 can be guaranteed by rather standard smoothness theorems (e.g., see [4, p. 141]). It is a straightforward exercise to verify that such theorems require sufficient smoothness of the coefficients as well as of the perturbing function f (this also involves the initial data u_0). An alternate approach, which permits relaxation of the smoothness of the coefficients (and requiring this smoothness only on $(\epsilon,T] \times \Omega$) could be taken (see [1], [5]) at the expense of some technical tedium. However, we shall be approximating the parameter functions q below on $[0,T]$, not $(\epsilon,T]$, and hence the stronger smoothness assumptions are more appropriate here.

To use the statements in Theorem 1 to obtain a convergence theory for approximate parameters, we shall need a continuous dependence result for solutions u of (4). To state this result, we define $C_B^1(\Omega)$ as the set of C^1 functions with bounded derivatives on Ω.

<u>Theorem 2</u>. For any solution u of (4) such that $u(t,q) \in C_B^1(\Omega)$, we have that $q \to u(t,q)$ is continuous on Q in the $H^0((0,T) \times \Omega)$ topology.

The arguments for this theorem, which involve estimates for $|u(t,q) - u(t,\tilde{q})|$, make use of conditions A and B above. They are very similar to the arguments outlined r Theorem 1 and so we shall not give them here. Instead we explain how these results are used to obtain a parameter convergence theorem.

<u>Theorem 3</u>. Suppose Q is compact in the $H^0((0,T) \times \Omega)$ topology. Then a solution $\bar{q}^N$ to the problem of minimizing J^N over Q exists, $N = 1,2,\ldots$. Let $\{\bar{q}^{N_k}\}$ be any convergent subsequence, $\bar{q}^{N_k} \to q^*$ in $H^0((0,T) \times \Omega)$. If $u(q) \in C^2(\overline{(0,T) \times \Omega})$ for each $q \in Q$, then q^* is a solution to the problem of minimizing J over Q.

Since the arguments are similar to those we have presented elsewhere (see [1, pp. 28-29]), we only sketch them here. The existence statement follows once one establishes continuity of $q \to J^N(q)$ on the compact set Q (the continuity arguments are similar to those behind Theorem 2). From Theorem 1, (2), and (8) we have

$$J(q^*) = \lim_{N_k\to\infty} J^{N_k}(\bar{q}^{N_k}) \leq \lim_{N_k\to\infty} J^{N_k}(q)$$

for any $q \in Q$. But Theorem 1 (with the constant sequence $\{q\}$) also guarantees $J^{N_k}(q) \to J(q)$. Thus $J(q^*) \leq J(q)$ for any $q \in Q$.

As we have indicated previously, the state approximation results of Theorem 3 are only first level results that are not satisfactory from a computational point of view since the approximate problems still involve minimization over the infinite dimensional set Q. We turn next to a second level of approximation where we combine the state approximation ideas outlined above with ideas for approximation of the parameter set Q.

We suppose that Q^M, $M = 1,2,\ldots$, are subsets of $H^0((0,T) \times \Omega)$ defined by $Q^M = i_M(Q)$ where i_M is a mapping from $Q \subset H^0((0,T) \times \Omega)$ into $H^0((0,T) \times \Omega)$. The approximation properties for the Q^M are given in terms of the mappings i_M. Specifically we assume

$\boxed{D}$ (a) The mapping $i_M : Q \to H^0$ is continuous;

(b) For each $q \in Q$, $i_M(q) \to q$ as $M \to \infty$ and the convergence is moreover uniform in $q \in Q$.

We note that we do <u>not</u> require that $Q^M \subset Q$. Furthermore, in the event $\Omega = (0,1) \times (0,1)$, there are several useful special cases of approximation sets that satisfy the assumptions of condition D. Under sufficient regularity assumptions on Q, we may choose $i_M = I^M$ = the linear spline (or cubic spline) interpolatory map--for precise definitions and details, see [1], [11]. As a second example, we (again for sufficient regularity on Q) may verify that condition D is satisfied when we choose $i_M = P^M$ = the orthogonal projection mapping (in H^0) onto the subspace $L(\Delta^M)$ of linear B-splines (or the subspace $S(\Delta^M)$ of cubic B-splines)--see [11].

To see that this approximation idea does indeed fulfill our needs theoretically, we first observe that if Q is compact in H^0, then part (a) of condition D guarantees that $Q^M = i_M(Q)$ is compact. Hence the problem of minimizing J^N of (8) over Q^M has a solution $\bar{q}_M^N$. From the compactness of Q^M, we have a convergent subsequence $\bar{q}_M^{N_k} \to \bar{q}_M$ in Q^M. It follows (under sufficient regularity) from Theorem 3 that $\bar{q}_M$ is a solution of the problem of minimizing J over Q^M. Let $\hat{q}_M \in Q$ be chosen such that $\bar{q}_M = i_M(\hat{q}_M)$. Then $\{\hat{q}_M\} \subset Q$ and the compactness of Q guarantees existence of a subsequential limit $q^* = \lim \hat{q}_{M_j}$. From (b) of condition D and the fact that $\bar{q}_{M_j} = i_{M_j}(\hat{q}_{M_j})$, it is easily seen that $\bar{q}_{M_j}$ also converges to q^*. We finally observe that $J(\bar{q}_{M_j}) \leq J(q)$ for all $q \in Q^{M_j}$. But since $Q^{M_j} = i_{M_j}(Q)$, we actually have $J(\bar{q}_{M_j}) \leq J(i_{M_j}(q))$ for all $q \in Q$. Taking the limit as $M_j \to \infty$ in this inequality and using the continuity of J and part (b) of condition D, we find $J(q^*) \leq J(q)$ for all $q \in Q$.

For details of the above double limit results in the case of cubic spline state approximations and linear or cubic interpolatary splines for the parameter approximations, the reader may consult [1]. We further observe that a careful consideration of the detailed arguments in [1] and those sketched above will reveal that the order of the limits in the double limit procedure is immaterial. Summarizing we have

Theorem 4. Let Q^M be as given above where Q is compact and let $\bar{q}_M^N$ be a solution of the problem of minimizing J^N over Q^M. Then for any convergent subsequence $\bar{q}_{M_j}^{N_k} \to q^*$, the limit q^* is a solution of minimizing J over Q.

III. Numerical Findings. We have carried out extensive numerical tests of the methods described in this note on problems involving estimation of both constant and time and/or spatially varying coefficients in parabolic equations. Computations (involving linear and cubic splines) for both test examples and inverse problems using experimental data have been performed. See, for example [1], [2], [3], where descriptions of the algorithms and software packages employed can also be found. We report here briefly on one aspect of our numerical findings which concerns the possible difference in performance of the algorithms when one employs a pointwise (in Ω--see (3)) fit-to-data criterion as opposed to an integral criterion (e.g., as in (2), where the $H^0(\Omega)$ norm is used).

We first observe that the theory presented in this note (for $H^0(\Omega)$ convergence of the approximating states) promises adequate performance of the methods when a distributed criterion such as (2) (and the analogous form of J^N--see (8)) is used in computations. One can, in some cases (e.g., see [8]), at the expense of technical tedium, establish pointwise (in Ω) convergence of states so that a convergence theory for pointwise criterion such as (3) can be developed. This raises the natural question as to which, if either, criterion is preferable from a computational viewpoint. Our preliminary numerical investigations suggest that whenever one is estimating constant or temporally varying parameters (the ones of interest in the insect dispersal studies of [2], [3]), it doesn't matter whether the problem is posed using (2) or (3).

To be more specific, consider the Example 4.4 of [1], which is a test example where both the true parameters and true solution are known. The equation involved has the form (1) with $\alpha = 0$, f known, $\Omega = (0,1)$, $\nu = 100(4-t)(.5-x)$ and $D = 20$. Tests to estimate the temporal part of ν (i.e., the term $4-t$) or D were performed (see [1] for details). The methods yield essentially the same parameter values and functions (to 3 decimal places) regardless of whether the integral form (i.e., (2)) or the pointwise form (see (3)) of the criterion function J (and the corresponding J^N) are used. In general the computer time to carry out the estimation with distributed criterion was equal to (or less than) that needed for the same example using a pointwise criterion.

We also raised this "criterion" question in some of our extensive uses of the methods with field data from insect dispersal experiments (see [2], [3] for a full report). Again we compared performance of the algorithm using a pointwise criterion with that employing a distributed (in Ω) criterion. Results very similar to those reported above for the test examples were obtained (i.e., same parameters with comparable computational efficiency). For a more complete discussion see [2].

Finally we note that we have successfully used the methods discussed in this paper for test examples and experimental data (again for insect dispersal) in the case of two-dimensional spatial domains Ω. These results will be reported in a manuscript currently in preparation.

Acknowledgements. This research was supported in part by the National Science Foundation under NSF grants MCS-8205335 (to H.T.B), MCS-8200883 (to P.L.D), and DEB-8207117 (to P.K.), and by the U.S. Air Force Office of Scientific Research under contract AFOSR 81-0198. Part of the research was carried out while the first two authors were visitors at the Institute for Computer Applications in Science and Engineering (ICASE), NASA Langley Research Center, Hampton, VA. which is operated under NASA contracts NAS1-15810 and NAS1-16394. Part of the research was carried out while the first author was a Visiting Professor in the Department of Mathematics, Southern Methodist University.

References

[1] H.T. Banks and P.L. Daniel, Estimation of variable coefficients in parabolic distributed systems, LCDS #82-22, Brown University, Sept. 1982; IEEE Trans. Auto. Control, to appear.

[2] H.T. Banks, P.L. Daniel and P. Kareiva, Estimation of temporally and spatially varying coefficients in models for insect dispersal, LCDS #83-14, Brown University, June, 1983; J. Math. Biology, to be submitted.

[3] H.T. Banks and P. Kareiva, Parameter estimation techniques for transport equations with application to population dispersal and tissue bulk flow models, LCDS #82-13, Brown University, July, 1982; J. Math. Biology, 17 (1983), 253-273.

[4] A. Friedman, Partial Differential Equations, Holt, Rinehart and Winston, 1969.

[5] A. Friedman, Partial Differential Equations of Parabolic Type, Prentice-Hall, 1964.

[6] N.L. Guinasso and D.R. Schink, Quantitative estimates of biological mixing rates in abyssal sediments, J. Geophys. Res. 80 (1975), 3032-3043.

[7] P. Kareiva, Experimental and mathematical analyses of herbivore movement: quantifying the influence of plant spacing and quality of foraging discrimination, Ecological Monographs, 52(3), (1982), 261-282.

[8] K. Kunisch and L. White, The parameter estimation problem for parabolic equations in multidimensional domains in the presence of point evaluations, preprint, 1983.

[9] O.A. Ladyzhenskaya and N.N. Ural'tseva, Linear and Quasilinear Elliptic Equations, Academic Press, 1968.

[10] J.L. Lions, Optimal Control of Systems Governed by Partial Differential Equations, Springer-Verlag, 1971.

[11] M.H. Schultz, Spline Analysis, Prentice-Hall, 1973.

ATE
LME