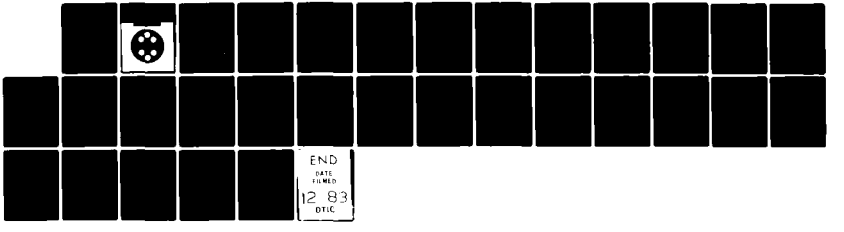
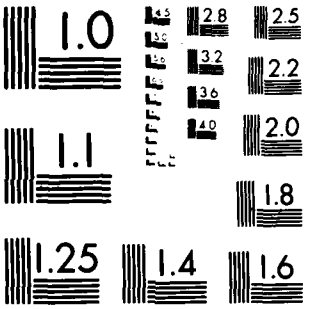
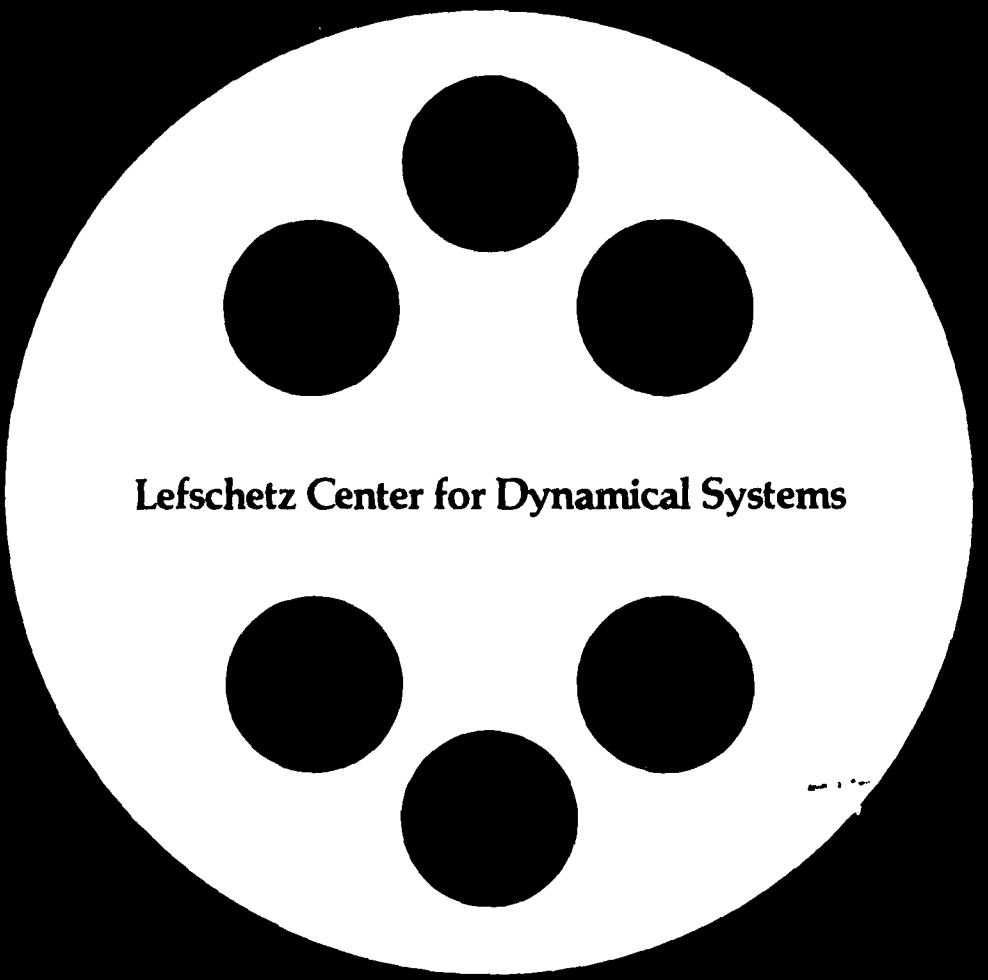
AD-A135 093 ESTIMATION OF TEMPORALLY AND SPATIALLY VARYING 1/1
COEFFICIENTS IN MODELS FOR..(U) BROWN UNIV PROVIDENCE
RI LEFSCHETZ CENTER FOR DYNAMICAL SYSTE..
UNCLASSIFIED H T BANKS ET AL. JUN 83 LCDS-83-14 F/G 12/1 NL



END
DATE FILMED
12 83
DTIC



AFOSR-TR- 83-0927



Lefschetz Center for Dynamical Systems

Division of Applied Mathematics Brown University Providence RI 02912

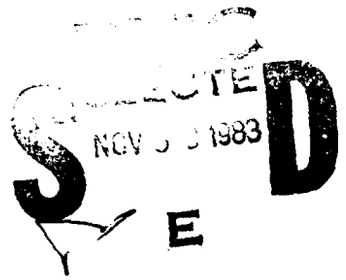
2

ESTIMATION OF TEMPORALLY AND SPATIALLY VARYING COEFFICIENTS IN MODELS FOR INSECT DISPERSAL

by

H. T. Banks, P. K. Daniel Lamm, P. M. Kareiva

June 1983 LCDS #83-14



UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

REPORT DOCUMENTATION PAGE		READ INSTRUCTIONS BEFORE COMPLETING FORM
1. REPORT NUMBER AFOSR-TR- 83-0927	2. GOVT ACCESSION NO.	3. RECIPIENT'S CATALOG NUMBER
4. TITLE (and Subtitle) ESTIMATION OF TEMPORALLY AND SPATIALLY VARYING COEFFICIENTS IN MODELS FOR INSECT DISPERSAL		5. TYPE OF REPORT & PERIOD COVERED TECHNICAL 6. PERFORMING ORG. REPORT NUMBER LCDS #83-14
7. AUTHOR(s) H.T. Banks, P.K. Daniel Lamm, P.M. Kareiva*		8. CONTRACT OR GRANT NUMBER(s) AFOSR-81-0198
9. PERFORMING ORGANIZATION NAME AND ADDRESS Lefschetz Center for Dynamical Systems, 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 JUN 83 13. NUMBER OF PAGES 29
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.K. Daniel Lamm, Department of Mathematics, Southern Methodist University, Dallas TX 75275; P.M. 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)

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

The authors describe techniques for estimating temporally and spatially dependent parameters (including coefficients) that appear in general transport models. Convergence properties of the resulting algorithms are given and sample computational findings with test examples are presented. The authors conclude with a summary of their use of the methods analyzing experiments on the movements of marked flea beetles in cultivated arrays of the cole crop, collards (Brassica oleraceae).

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

UNCLASSIFIED

SECURITY CLASSIFICATION OF THIS PAGE (When Data Entered)

ESTIMATION OF TEMPORALLY AND SPATIALLY VARYING COEFFICIENTS IN MODELS FOR INSECT DISPERSAL

H. T. Banks*
Lefschetz Center for Dynamical Systems
Division of Applied Mathematics
Brown University
Providence, Rhode Island 02912

P. K. Daniel Lamm†
Department of Mathematics
Southern Methodist University
Dallas, Texas 75275

P. M. Kareiva#
Division of Biology and Medicine
Brown University
Providence, Rhode Island 02912

June 1983

* This research was supported in part by the National Science Foundation under NSF #MCS-8205335 and by the U. S. Air Force Office of Scientific Research under contract AFOSR 81-0198. Part of the research was carried out while this author was a Visiting Professor in the Department of Mathematics at Southern Methodist University.

† This research was supported in part by the National Science Foundation under NSF #MCS-8200883. 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.

This research was supported in part by the National Science Foundation under NSF #DEB-8207117.

ESTIMATION OF TEMPORALLY AND SPATIALLY VARYING COEFFICIENTS IN MODELS FOR INSECT DISPERSAL

H. T. Banks, P. L. Daniel Lamm and P. M. Kareiva

ABSTRACT

We describe techniques for estimating temporally and spatially dependent parameters (including coefficients) that appear in general transport models. Convergence properties of the resulting algorithms are given and sample computational findings with test examples are presented. We conclude with a summary of our use of the methods analyzing experiments on the movements of marked flea beetles in cultivated arrays of the cole crop, collards (*Brassica oleraceae*).

I. INTRODUCTION

Transport equations appropriately model numerous biological systems and have played an important role in mathematical biology (cf Rubinow [14], Okubo [12]). Recently such transport equations have been increasingly used with experimental data, ranging from developmental biology (cf Kauffman et al. [10]) to population biology (cf Levin [11]). This connection between model and data can be facilitated by parameter identification schemes (i.e., algorithms for generating the parameters in a model that provide the best match between model predictions and observed data). For the past several years we have been developing and testing spline-based algorithms for identifying parameters in transport equations (cf Banks, Crowley and Kunisch [1]). Our most recent efforts have been directed towards models of insect dispersal where,working with insect mark-recapture data, we have had some success in identifying constant coefficients and spatially varying coefficients in transport equations of insect movement (Banks and Kareiva [4]). However, our successful efforts involved only our more abbreviated data sets (those spanning one or two days) and were counterbalanced by failures at describing shifts in insect distributions over the course of three days. We hypothesized that our models failed in some instances because they lacked temporal variation in the parameters reflecting rates of insect movement and migration (Banks and Kareiva [4]).

The idea of temporally varying dispersal rates is not new to biologists. In general, the complication of temporal variation in rate constants or model parameters is ubiquitous in biology. For example, many biological processes vary with time because the environment changes

seasonally. Biological processes may also vary temporally without any driving environmental force; a familiar example is the inexorable senescence of organs in aging individuals. Motivated by the apparent temporal variation in the mobility of the insects in our experiments, and the recognition that temporal variation is a general feature of biological processes, we have made efforts to extend our parameter estimation algorithms to enable us to treat transport equations that contain time-varying parameters.

In this paper we describe our methods for identifying both spatially and temporally varying coefficients in transport equations. Although the particular examples with which we illustrate our methods here represent scalar equations, our techniques are readily used with vector systems. Beginning with a basic transport model, we sketch the ideas behind our spline based techniques and summarize the main convergence results. We then report on tests of our methods on numerical examples involving coefficients that simultaneously vary in space and time. Because we view our methods as tools, we discuss practical computing aspects of implementing these methods. We then summarize results obtained when we applied the methods to experimental data that describes the dispersal of flea beetles. It was this data that initially prompted us to consider models with time-varying coefficients. Since this paper represents the culmination of several mathematical and biological investigations, we have sacrificed many details in order to present an overview. Readers interested mainly in the biological issues of flea beetle dispersal should consult Kareiva [7], [8], [9], and Banks and Kareiva [4] ; for more details concerning our general approximation techniques, including convergence

proofs, readers may consult Banks and Daniel [2] and Banks, Daniel and Kareiva [3] .

II. SPLINE BASED ESTIMATION TECHNIQUES FOR TRANSPORT EQUATIONS

We continue here our investigations (Banks and Kareiva [4], Sec. 2) of general transport equations which describe advection, Fickian diffusion, and sink/source mechanisms containing variable and possibly unknown rate parameters. Specifically we consider the problem of estimating the unknown functional parameters V, $\mathcal{D}$, α, β, and γ appearing in the system

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x}(Vu) = \frac{\partial}{\partial x}\left(\mathcal{D}\frac{\partial u}{\partial x}\right) + \alpha u + f, \quad t\in(0,T],\ x\in(0,1)$$

$$u(t,0) = u(t,1) = 0 \tag{1}$$

$$u(0,\cdot) = u_0(\gamma) .$$

It is assumed that $\gamma = \gamma(x)$ and that V, $\mathcal{D}$, α, and β are functions of $(t,x)\in(0,T]\times(0,1)$. We note that although (1) is formulated in terms of homogeneous boundary conditions, our ideas are sufficiently general to include systems with nontrivial boundary conditions, possibly even dependent upon unknown parameters. A standard transformation (see Banks and Kareiva [4]) may be used to reformulate such systems as special cases of (1) and we shall therefore restrict our considerations in this paper to estimation problems for (1).

The parameter estimation problem associated with the set $q = (\mathcal{D},V,\alpha,\beta,\gamma)$ of unknown parameters appearing in (1) is the following: Given spatially distributed observations $\hat{u}_i \in L_2(0,1)$ corresponding to times t_i, $i=1,\ldots,L$, find $q \in Q$ (for Q a given class of admissible parameter functions) that minimizes a given fit-to-data criterion. We employ here a least squares

criterion (although we are not limited only to functionals of this type) of the form

$$(2) \qquad J(q) = \sum_{i=1}^{L} |u(t_i;q) - \hat{u}_i|^2$$

where u is the solution to (1) corresponding to $q \in Q$; the parameter set Q is assumed to possess boundedness, compactness, and regularity properties (these assumptions are made precise in Banks, Daniel and Kareiva [3]) .

Due to the infinite dimensional nature of both the state u and the parameters q (and the computational difficulties associated with optimazation over infinite dimensional spaces), one of our objectives has been to develop numerical estimation algorithms based on approximation schemes for both the state and the parameters. We describe such an algorithm here. To this end we define finite dimensional state spaces H^N and approximate parameter sets Q^M and reformulate the parameter estimation problem in a computationally more tractable finite-dimensional setting.

To consider state approximations, we first rewrite (1) in weak form in the state space $L_2(0,1)$ with the usual inner product $\langle \cdot,\cdot \rangle$. For a fixed $q = (\mathcal{D},\mathcal{V},\alpha,\beta,\gamma) \in Q$, this weak form is given by

$$(3) \qquad \begin{aligned} &\langle u_t,\phi \rangle - \langle \mathcal{V}u,D\phi \rangle + \langle \mathcal{D}Du,D\phi \rangle - \langle \alpha u,\phi \rangle = \langle f,\phi \rangle, \ t \in (0,T] \\ &u(0) = u_0(\gamma) \end{aligned}$$

which solutions u of (1) must satisfy for all sufficiently smooth ϕ with $\phi(0) = \phi(1) = 0$. Here D denotes the spatial differentiation operator. We next use Galerkin ideas with (3) to approximate solutions u of (1): For each N, we construct finite dimensional subspaces H^N of $L_2(0,1)$, $H^N = \text{span}\{B_1^N,\ldots,B_{N+1}^N\}$, where the basis elements B_j^N are linear com-

binations of standard cubic B-spline elements (see Schultz [15], for example) defined on a uniform mesh of size 1/N. The linear combinations are chosen so that the resulting basis elements B_j^N satisfy the homogeneous boundary conditions in (1). For q and N given, we then seek the solution u^N, $u^N(t) = \sum_{j=1}^{N+1} w_j^N(t) B_j^N$, which satisfies the Galerkin equations

$$
\begin{aligned}
&\langle u_t^N, \psi\rangle - \langle \mathcal{V} u^N, D\psi\rangle + \langle \mathcal{D} D u^N, D\psi\rangle - \langle \alpha u^N, \psi\rangle = \langle f, \psi\rangle \quad t \in (0,T], \\
&\langle u^N(0), \psi\rangle = \langle u_0(\gamma), \psi\rangle \ ,
\end{aligned} \tag{4}
$$

for all $\psi \in H^N$. Since H^N is finite dimensional, it is easy to see that (4) is an ordinary differential equation in the "Fourier coefficients" $w_j^N(t)$; we defer further discussions on the nature of the approximating equations to a later section where we provide details on the computer implementation of the overall algorithm.

Our approach to approximation of the functional parameters $q = (\mathcal{D}, \mathcal{V}, \alpha, \beta, \gamma) \in Q$ is similar to that of the state approximations in that we construct finite dimensional sets Q^M that "approximate" (in an appropriate sense) the original infinite dimensional parameter set Q. Our theory allows for a wide variety of classes of approximation sets Q^M (we need not have $Q^M \subseteq Q$) which in general are unrelated to the N^{th} order state spaces H^N. In the next section we indicate how we have successfully used cubic and linear spline-based approximations for the parameters; a more precise statement of these general ideas may be found in Banks and Daniel [2] and Banks, Daniel, and Kareiva [3] .

In practice we combine both state and parameter approximation

ideas to formulate an approximate estimation problem: For given levels of approximation N and M, the problem consists of determining $\bar{q}_M^N \in Q^M$ that minimizes

$$(5) \qquad J^N(q) = \sum_{i=1}^{L} |u^N(t_i;q) - \hat{u}_i|^2$$

over $q \in Q^M$. As before, u^N is the solution to (4) corresponding to q. Each approximate estimation problem, being finite dimensional, enjoys significant computational advantages over our original estimation problem; in addition we have the following convergence results (valid under sufficient regularity assumptions on f, u_0 and Q -- see Banks, Daniel and Kareiva [3]) for solutions of the approximate problems:

Theorem: For N=1,2,..., and M=1,2,..., there exists a solution $\bar{q}_M^N$ to the problem of minimizing J^N over Q^M. In addition, there exists a convergent subsequence $\bar{q}_{M_j}^{N_k} \to q^*$ such that $q^* \in Q$ is a solution to the original problem of minimizing J over Q. Finally, we obtain state variable convergence $u^{N_k}(t;\bar{q}_{M_j}^{N_k}) \to u(t;q^*)$, where u^{N_k}, u are solutions to (4), (3) corresponding to $\bar{q}_{M_j}^{N_k}$, q^* respectively.

The proof of this theorem, which relies heavily on variational arguments and spline estimates (e.g., see Schultz [15]), may be found for a general transport equation in a two-dimensional domain in Banks, Daniel and Kareiva [3] and thus will not be presented here. Instead we turn to a discussion of computer implementation and numerical testing of our algorithms, and to our findings upon use of these techniques with flea beetle dispersal data.

III. NUMERICAL IMPLEMENTATION

In this section we first outline some essential features of computer implementation of the spline-based approximation ideas introduced above. Then, before examining the more difficult problems associated with estimating parameters for models using experimental data, we shall present our findings for a representative test example.

For given values of M and N, we wish to consider the approximate parameter estimation problem, namely that of minimizing J^N of (5) over Q^M. In developing a computational package to effect this minimization, one can often take into account special characteristics of the state space H^N and the parameter set Q^M. Turning first to the state space approximations, we examine equation (4) for any given value of parameter $q = (\mathcal{D},V,\alpha,\beta,\gamma)$ in Q^M. We recall that $u^N(t) \in H^N$ may be written $u^N(t) = \sum_{j=1}^{N+1} w_j^N(t) B_j^N$ (where B_j^N are the cubic spline basis elements defined in sec. 2); substituting this expression into (4) and letting $\psi = B_i^N$, $i=1,\ldots,N+1$, we obtain the system of ordinary differential equations (ODE) in $w^N(t) = \operatorname{col}_j(w_j^N(t))$,

$$
\begin{aligned}
&Q^N \dot{w}^N(t) = (K^N(t) + G^N(t)) w^N(t) + F^N(t) \\
&Q^N w^N(0) = w_0^N.
\end{aligned}
\tag{6}
$$

Here Q^N, $K^N(t)$ and $G^N(t)$ are $(N+1) \times (N+1)$ matrices, where for $i,j=1,\ldots,N+1$,

$$Q_{i,j}^N = \langle B_j^N, B_i^N \rangle ,$$

$$K_{i,j}^N = \langle \mathcal{D}(t,\cdot)\, DB_j^N + V(t,\cdot) B_j^N, DB_i^N \rangle$$

$$G_{i,j}^N = \langle \alpha B_j^N, B_i^N \rangle .$$

In general, these matrices are sparse, banded, and (in the case of Q^N, G^N) symmetric. In addition, the (N+1)-vectors $F^N(t)$ and w_0^N are given by

$$F^N(t) = \operatorname*{col}_i(\langle f(\beta,t,\cdot),B_i^N\rangle)$$

$$w_0^N = \operatorname*{col}_i(\langle u_0,B_i^N\rangle).$$

We remark here that, with the exception of the time-varying nature of the matrices above, the ODE (6) in w^N is essentially the same as the corresponding system (equation (12)) in Banks and Kareiva [4]. The fundamental difference between the two, however, involves the way in which variable parameters are treated in each. In Banks and Kareiva [4], a fixed parametrization is chosen that imposes a specific a priori shape for the coefficients (e.g., in that reference, the choice of parametrizations for V, namely $V(x) = c(x-.5)$, constrains V to a class of affine functions). Here we allow the iterative estimation scheme itself to determine the shape of the coefficients by searching over functions in a class Q^M which contains a large number of different functional shapes. (For the special case described below, we use linear and cubic splines to generate the elements of Q^M.) We shall briefly indicate some of the computational aspects of this approach and refer the reader to Banks and Daniel [2] for a more thorough treatment of these ideas (we note that the underlying model equation in Banks and Daniel [2] differs from the one considered here, but that the approximation ideas for variable parameters presented there are directly applicable here).

To explain the implementation of approximation techniques for parameters, we shall simplify our presentation by focusing on the parameter α only and note that the parameters D, V, β and γ are treated in a similar

manner. To illustrate our ideas, we assume, in the examples to follow, a separable form for the parameters (i.e., $\alpha(t,x) = \alpha_1(t)\alpha_2(x)$) and that the approximation from Q^M for α takes the form $\alpha^M = \alpha_1^M\alpha_2^M$ where

$$(7) \qquad \alpha_1^M(t) = \sum_{k=1}^{M} a_k^M C_k^M(t)$$

$$(8) \qquad \alpha_2^M(x) = \sum_{k=1}^{M} \hat{a}_k^M \hat{C}_k^M(x).$$

Here C_k^M, k=1,...,M, are linear spline basis elements based on a mesh size T/(M-1) and $\hat{C}_k^M$, k=1,...,M are standard cubic spline (B-splines -- see Schultz [15]) elements with mesh size 1/(M-3). (We note that the orders of the approximation for α_1 and α_2 need not be related; the form assumed here is only for ease in expostion.) For fixed M, N, the matrix $G^N(t)$ of equation (6) has entries

$$(9) \qquad G_{ij}^n(t) = \Big(\sum_{k=1}^{M} a_k^M C_k^M(t)\Big)\Big\{\sum_{k=1}^{M} \hat{a}_k^M \ \langle \hat{C}_k^M B_j^N, B_i^N\rangle\Big\},$$

where a_k^M, $\hat{a}_k^M$, k=1,...,M, are to be estimated. We note that the inner products $\langle \hat{C}_k^M B_j^N, B_i^N\rangle$ may be computed and stored prior to beginning the optimization process, so that as the a_k^M, $\hat{a}_k^M$ are updated, we need only recombine the terms in (9) to obtain the current value of $G^N(t)$. It is easy to see how to extend these ideas to the matrices involving approximations for $\mathcal{D}$, V, β and γ.

To minimize J^N over Q^M using any one of a number of iterative optimization procedures, it is necessary to select start-up values for the parameters (i.e., for the coefficients a_k^M, $\hat{a}_k^M$ in expansions such as (7), (8)). For the examples reported here we used the minimization package LMDIF1, which is MINPACK's version of the Levenberg - Marquardt modified Newton algorithm. We also used IMSL's DGEAR (stiff ODE solver) to compute the solution to (6) for each parameter iterate, and the IMSL numerical quadrature package DCADRE to compute J^N (the L_2 norm

difference between u^N and sample data at sample times t_i). All the computations reported on here were executed on the CDC 6600 at Southern Methodist University and the IBM 370/158 at Brown University.

We now consider a test example where (synthetic) data generated from a known solution (see Banks and Daniel [2]) is used with the estimation algorithm; by comparing the parameters identified by our estimation process with known "true" parameter values, we are able to evaluate the performance of our algorithm. In this test example, "true" parameters are given by $\mathcal{D}$=20, $V(t,x) = -100(x-.5)(4-t)$, and $\alpha = \beta = \gamma = 0$, while the "true" state is given by $u(t,x) = -400x(x-1)\cos(t/2)$, $0 \leq t \leq 2$, $0 \leq x \leq 1$. Starting from initial guesses of the parameters, we apply the optimization procedure to J^N, where the distributed sample data needed to compute J^N is given by $\hat{u}_i = u(t_i,\cdot)$, $t_i = .25i$, i=1,...,8. For each example presented below we let N=3. (Further details on the implementation of test examples may be found in Banks and Daniel [2].)

Example 1.

(a) We estimate $\mathcal{D}$ only, holding V fixed at its true value. We allow time varying estimates for $\mathcal{D}$, letting $\mathcal{D}^M(t) = \sum_{k=1}^{M} d_k^M C_k^M(t)$ where M=4 and the C_k^M are defined as in (7). A time-varying initial guess of $\mathcal{D}$=40-15t yields the following results:

	d_1^M	d_2^M	d_3^M	d_4^M
initial guess	40.	30.	20.	10.
estimated values (N=3,M=4)	20.0002	20.0002	20.0001	20.0000
"true" values	20.	20.	20.	20.

Execution time was 110 seconds with a final value of $J^N(\bar{\mathcal{D}}_M^N) = 6.7 \times 10^{-7}$.

(b) We estimate the time-varying part of V only, holding other parameters fixed at true values; i.e., we let $V(t,x) = V_1(t)V_2(x)$ and estimate $V_1(t) = 4-t$ only. As before, N=3, M=4, with $V_1^M(t) = \sum_{k=1}^{M} v_k^M C_k^M(t)$. Starting from an initial guess of $V_1 \equiv 1$ we obtain the following estimates:

	v_1^M	v_2^M	v_3^M	v_4^M
initial guess	1.0	1.0	1.0	1.0
estimated values (N=3,M=4)	3.99994	3.33329	2.66664	1.99999
"true" values	4.0	3.33333	2.66667	2.0

The execution time for this example was approximately 62.55 seconds and the converged value of J^N was given by $J^N(\bar{V}_M^N) = 1.2 \times 10^{-6}$.

Example 2.

We repeat the above examples (same start-up values, N=3, M=4) but consider the case where sample data is available only at discrete spatial points x_j in (0,1). The least squares fit-to-data criterion in this case is given by

$$\tilde{J}(q) = \sum_{i,j} |u(t_i,x_j;q) - \hat{u}_{ij}|^2 , \tag{10}$$

with the corresponding approximate criterion

$$\tilde{J}^N(q) = \sum_{i,j} |u^N(t_i,x_j;q) - \hat{u}_{ij}|^2; \tag{11}$$

here u, u^N are solutions to (3), (4) respectively and $\hat{u}_{ij} \in R$ are observed data points. Repeating example 1(a) (where we estimate $\mathcal{D}$), we obtain the following converged values.

	d_1^M	d_2^M	d_3^M	d_4^M
estimated values (N=3,M=4)	19.9998	19.9996	19.9997	19.9997

The CP time here was 475.33 sec. while $\tilde{J}^N(\bar{\mathcal{D}}_M^N) = 1.44 \times 10^{-3}$. Repeating

example 1(b) resulted in the following converged values of v_i^M, with a CPU time of 143.99 sec. and the least squares criterion value $\tilde{J}^N(\bar{V}_M^N) = 4.57 \times 10^{-7}$.

	v_1^M	v_2^M	v_3^M	v_4^M
estimated values (N=3, M=4)	3.99994	3.33329	2.66664	1.99999

Example 3.

Here we estimate both spatially and temporally varying parts of $V(t,x) = [4-t][-100(x-.5)] \equiv V_1(t)V_2(x)$ using the "pointwise" criterion $\tilde{J}^N$. As in the previous examples we approximate V_1 by V_1^M, and also approximate V_2 by $V_2^M(x) = \sum_{k=1}^{M} \hat{v}_k^M \hat{C}_k^M(x)$ where $\hat{C}_k^M$ is defined as in (8) and M is fixed at M=4. Since V^M is the product of V_1^M and V_2^M, at least one of their basis coefficients must be fixed or there will be an infinite number of possible combinations of V_1^M and V_2^M that still yield $V_1^M V_2^M = \bar{V}^M$. For the case considered here we fixed $\hat{v}_1^M$ and $\hat{v}_2^M$ at their "true" values throughout the entire optimization process. Initial guesses for the remaining coefficients, as well as their converged values (for N=3), are given below.

	$\hat{v}_3^M$	$\hat{v}_4^M$	v_1^M	v_2^M	v_3^M	v_4^M
initial guess	-5.0	-15.0	1.0	1.0	1.0	1.0
estimated values (N=3,M=4)	-8.3333	-24.9883	4.0005	3.3338	2.6670	2.0003
"true" values	-8.3333	-25.0	4.0	3.3333	2.6667	2.0

Additional numerical findings for this example, as well as numerous other examples (with both temporally and spatially varying parameters), may be found in Banks and Daniel [2]; Section 4]

We consider in the next section the estimation of parameters using models for insect dispersal with experimental data from field studies.

The numerical results reported below were generated using a state approximation index of N=32 and parameter approximation index of M=4. In addition, for convenience, the "pointwise" least squares criterion $\tilde{J}^N$ was used in the minimization scheme instead of J^N since sample data $\hat{u}_{ij}$ was readily available in discrete form. As in the test examples above, we determined that the use of $\tilde{J}^N$ vs. J^N made little difference in the outcome when one is estimating time-varying parameters. (We repeated several of the computations with J^N as the criterion, where the distributed data $\hat{u}_i$ was constructed by interpolating linearly between data points $\hat{u}_{i1},\ldots,\hat{u}_{i\rho}$; here ρ is the number of observed spatial locations.)

IV. TEMPORALLY VARYING INSECT DISPERSAL

Since dispersal is important to the population dynamics of insects (cf Stinner et al. [17], Joyce [6]), mathematical models have been used extensively to investigate the consequences of differing patterns of insect movement (Kareiva [7], [8], [9]; Okubo [12]). The development of accurate models may be especially valuable in the design of pest management schemes, where certain cropping arrangements appear to promote outbreaks by altering pest movements (Risch et al. [13]). As part of our long-term study of insect dispersal, we have been examining population models that are special cases of equation (1) (cf Kareiva [7]; Banks and Kareiva [4]). An extensive review of field mark-recapture data has indicated that, although passive diffusion models are a good beginning, models of insect dispersal must also allow for rates that vary in space and time (Kareiva [8]). In our first attempt at treating spatially varying dispersal, we analyzed the spread of marked populations of flea beetles and found that transport models with spatially varying convection (or oriented movement) performed significantly better than passive diffusion models, but still

not as well as we would like (Banks and Kareiva [4]). We hypothesized that in these flea beetle mark-release experiments, coefficients representing certain mechanisms were time dependent. Here we use our new techniques to test the hypothesis that flea beetle dispersal and migration varies temporally in these mark-recapture experiments. Beetle movement was observed in 1 m × 80 m cultivated strips of collard patches. Since there were no food plants for flea beetles outside of these linear arrays, their local movement was confined to short hops up and down each linear array. This justifies the following model as a reasonable hypothesis:

$$(12) \qquad \frac{\partial u}{\partial t} = \mathcal{D}(t)\,\frac{\partial^2 u}{\partial x^2} - \alpha(t)u, \qquad t \in (0,T],\ x \in (0,1),$$

where u is the density of marked beetles, $\mathcal{D}$ is the diffusion coefficient, α represents population "decay" rate due to death and emigration, and x is the rescaled spatial scale (where the interval [0,1] corresponds to an actual interval of length 100m). On this rescaled interval the center of each experimental array corresponds to x=.5, and the data are observations of u at evenly spaced sample stations between .2 and .8. Initial data is given by $u(0,x)=0$ for all $x \neq .5$ and $u(0,.5)$ is the number of marked beetles that were released at the mid-point of each array to begin the experiment. Since less than 1% of the marked beetles reached the ends of the cultivated arrays ($x = .1$ and $x = .9$), and since there was no food to support beetles outside the arrays, we used $u(t,0) = u(t,1) = 0$ as our boundary conditions. More details concerning the marking and recapture procedure and flea beetle biology can be found in Kareiva [7]. Our subsequent analyses of experimental data and equation (12) are reported in terms of days after release. In fact, the data were collected in terms of hours, with 8 hours (9 a.m. - 5 p.m.) representing a beetle's "activity day." Outside of this activity period

beetles are inactive, neither feeding nor moving (Kareiva, personal observation). To explore the hypothesis of time-varying dynamics, we systematically examine the ability of equation (12) to describe experimental data, using different combinations of $\mathcal{D}$ constant, $\mathcal{D}$ varying in time, α constant and α varying in time.

The spline algorithm we have developed estimates the coefficients in equation (12) that yield the "best fit" between model dynamics and experimental data; in particular, the algorithm minimizes the sums of the squared differences between observed beetle densities $\hat{u}_{ij}$ and predicted beetle densities u at all points (t_i,x_j) for which there are data. As a standard against which to compare the fit of the model, we consider the null hypothesis that beetle density is simply a constant normal random variable. Using this null hypothesis, we have evaluated the success of differing forms of equation (12) by an ad hoc modification of multiple regression analyses and significance tests based on the F-distribution. The basic idea in this approach is the "extra sums of squares" principle (Draper and Smith [5]), which involves examining the reduction in the residual sum of squares that is achieved by adding parameters to the model. Although for such non-linear models use of the F-test cannot be rigorously justified, it at least provides a guide to the significance of our findings.

Following the protocol of traditional regression analysis, we define a total sums of squares (TSSQ) for any data set $\{\hat{u}_{ij}\}$ as

$$\sum_{j=1}^{m} \sum_{i=1}^{n} (\hat{u}_{ij}-\bar{u})^2$$

where nm is the number of observations of beetle density (one for each (t_i,x_j)) and $\bar{u}$ is the observed mean beetle density. Unlike regression

analysis, however, our model does not hypothesize that beetle density is a simple function of selected independent variables. Instead, our model is a partial differential equation, the solution of which depends on initial data and boundary conditions (known parameters) and on (possibly time-varying) coefficients. These time-varying coefficients are the unknown parameters we are seeking to identify. In our statistical analyses, we calculated the degrees of freedom (df) in each model identification run as simply the number of parameters in equation (12) that were estimated by our algorithm. For example, if we were testing a model with $\alpha = 0$ and $\mathcal{D}$ an unknown constant, the model df is 1. If we sought to estimate a constant $\mathcal{D}$ and a time-varying α that was to be represented by 4 linear splines, then our model df is 1 (for $\mathcal{D}$) plus 4 (for the a_k^M, k=1,...,4 in (7)) or 5 in total. Given a total sums of squares (hereafter TSSQ) and error sums of squares ($\tilde{J}^N$ of Section 3; hereafter referred to as ERRSQ), we calculated explained sums of squares (EXSSQ) as TSSQ - ERRSQ. In traditional regression analyses EXSSQ must always be nonnegative. When testing dynamic models such as equation (12), it is possible, however, to obtain ERRSQ > TSSQ and consequently EXSSQ < 0. This possibility arises because a dynamical model may be so inappropriate that treating the data as simple normally distributed (white) noise provides a better description than "foolish" dynamics. Thus the examination of particular models such as (12) through parameter estimation techniques is not an exercise in simple curve fitting with guaranteed success whenever enough parameters are included. In our situation, no amount of refining of parameter fitting can resurrect a poorly chosen dynamic model.

Three separate dispersal experiments were performed with the flea beetle Phyllotreta striolata: one in a linear array with 3 m between collard patches, one with 6 m between collard patches, and one with 9 m between patches. At all three interpatch spacings, our EXSSQ exceeded 75% of the TSSQ when both $\mathcal{D}$ and α were allowed to vary in time (Table 1). However, a good fit between predicted and observed data could in fact, be consistently obtained with a much simpler model in which we treated $\mathcal{D}$ as a constant but allowed α to vary temporally. In this case we always explained a significant portion of the TSSQ, a portion which was only negligibly less than the portion explained when $\mathcal{D}$ also varied temporally (Table 2). Without temporal variation in either $\mathcal{D}$ or α, solutions of equation (12) could provide a good fit with data for only one day (Figure 1).

An interesting methodological point concerning the search for time-varying parameters is the importance of initial guesses. In Table 3 we see that using the same experimental data and the same model but differing initial guesses, we obtained dramatically different parameter values and percentages of TSSQ explained. Our "good" first guesses (those eventually leading to large EXSSQ) were the constant α's or $\mathcal{D}$'s identified by our spline-based algorithm under the assumption of no temporal variation in parameter values. These estimates of constant α and $\mathcal{D}$ then could be effectively used as starting points for our search of temporally varying parameters.

Not only was α consistently identified as significantly varying in time, the quantitative nature of that temporal variation was remarkably identical among different experiments (Figure 2). We interpret Figure 2 to represent the following general biological phenomenon: Early after their release (i.e., in the first day), there is a peak in beetle emigration (high α) due to the disturbance of being marked and handled;

after one day of natural activity in the field this disturbance effect disappears and α falls to its "natural" level. Disturbance following marking and release is so common in insect dispersal experiments that ecological methodology books warn against it as an experimental artifact that needs to be controlled for or taken into account when analyzing data (cf. Southwood[16]; p.91). Because our method allows us to explicitly identify the disturbance effect, in studying specific models we have a quantitative technique for estimating intrinsic natural diffusivity and emigration of flea beetles after disturbance has been factored out. We are now applying our methods to insect dispersal systems where we suspect that day to day variation in temperature produces marked variation in movement rates. Our identification algorithms provide a tool for describing the relationship between D and temperature under field circumstances that do not permit the experimental control or manipulation of temperature.

TABLE 1. Least Squares Analysis of the Ability of Differing Transport Models to Describe Beetle Dispersal Data.

DESCRIPTION OF DATA

3 m spacing: 9 points in space sampled 1 and 3 days after beetles were released, TSSQ = 76.3

6 m spacing: 9 points in space sampled 1 and 3 days after beetles were released, TSSQ = 53.5

9 m spacing: 7 points in space sampled 1 and 3 days after beetles were released, TSSQ = 109.1

PERFORMANCE OF VARIOUS MODELS

Model	Data Set	% of TSSQ explained by model	F-statistic and significance level
$\frac{\partial u}{\partial t} = \mathcal{D} \frac{\partial^2 u}{\partial x^2} - \alpha u$	3 m	32.7%	$F_{2,15} = 3.64$, $p > .05$
(with constant	6 m	NONE	(model ERSSQ > TSSQ)
$\mathcal{D}$ and α)	9 m	73.4%	$F_{2,11} = 15.97$, $p < .005$
$\frac{\partial u}{\partial t} = \mathcal{D} \frac{\partial^2 u}{\partial x^2} - \alpha(t)u$	3 m	83.9%	$F_{5,12} = 12.48$, $p < .005$
(with constant	6 m	70.4%	$F_{5,12} = 5.71$, $p < .01$
$\mathcal{D}$, variable α)	9 m	97.6%	$F_{4,9} = 68.7$, $p < .005$
$\frac{\partial u}{\partial t} = \mathcal{D}(t) \frac{\partial^2 u}{\partial x^2} - \alpha(t)u$	3 m	92.5%	$F_{8,9} = 13.8$, $p < .005$
(with variable	6 m	76.7%	$F_{8,9} = 3.71$, $p < .05$
$\mathcal{D}$, variable α)	9 m	97.6%	$F_{8,15} = 19.2$, $p < .005$

TABLE 2. Choosing the "Best" Model By Examining How the Percent of TSSQ Explained Can Be Increased as the Models Gain Complexity (i.e., Number of Parameters)

A. Moving from constant α and $\mathcal{D}$ (2 parameters) to a constant $\mathcal{D}$ but variable α model (5 parameters)		
3 m spacing:	% TSSQ explained goes from 32.7 to 83.9	$F_{3,12} = 12.7$, $p < .005$ significant improvement
6 m spacing:	% TSSQ explained goes from 0 to 70.4	$F_{3,12} = 9.51$, $p < .005$ significant improvement
9 m spacing:	% TSSQ explained goes from 73.4 to 97.6	$F_{3,8} = 18.82$, $p < .005$ significant improvement
B. Moving from constant $\mathcal{D}$ and variable α model (5 parameters) to variable $\mathcal{D}$ and α model (8 parameters)		
3 m spacing:	% TSSQ explained goes from 83.9 to 92.5	$F_{3,9} = 3.43$, $p > .05$ improvement is not significant
6 m spacing:	% TSSQ explained goes from 70.4 to 76.7	$F_{3,9} = 8.11$, $p > .5$ improvement is not significant
9 m spacing:	there is no change in % TSSQ explained	no improvement

TABLE 3. Dependence of Performance of Algorithm on Initial Guess

Data from 6 m spacing experiments (see Table 1)
TSSQ = 53.5

MODEL	$\frac{\partial u}{\partial t} = \mathcal{D} \frac{\partial^2 u}{\partial x^2} - \alpha(t)u$	% of TSSQ explained by model
EX1:	$\mathcal{D}$ = .03101, INITIAL $\alpha(t) \equiv .20$ SEARCH FOR VARIABLE α	NONE (model ERSSQ > TSSQ)
EX2:	$\mathcal{D}$ = .00046, INITIAL $\alpha(t) \equiv .2567$ SEARCH FOR VARIABLE α	70.4%
MODEL	$\frac{\partial u}{\partial t} = \mathcal{D}(t) \frac{\partial^2 u}{\partial x^2} - \alpha(t)u$	
EX3:*	INITIAL $\mathcal{D}(t)$: .0296 .0302 .00008 .0000 INITIAL $\alpha(t) \equiv 0$ SEARCH FOR VARIABLE $\mathcal{D}$ AND α	12%
EX4:*	INITIAL $\mathcal{D}(t)$: .0296 .0302 .00008 .0000 INITIAL $\alpha(t)$: .0097 .0263 .4834 .0058 SEARCH FOR VARIABLE $\mathcal{D}$ AND α	NONE (model ERSSQ > TSSQ)
EX5:*	INITIAL $\mathcal{D}(t) \equiv .00046$ INITIAL $\alpha(t)$: .2559 .2486 .0056 .00014 SEARCH FOR VARIABLE $\mathcal{D}$ AND α	76.7%

*FUNCTIONAL INITIAL GUESSES EXPRESSED IN TERMS OF SPLINE REPRESENTATION COEFFICIENTS - SEE SEC. 3

FIGURE LEGENDS

Figure 1. Fitting models to data on the dispersal of marked flea beetles. Observed average densities of recaptured beetles in linear array containing collard patches at 9 m spacing are represented by solid dots; the dashed line represents best-fit equation (12) with constant D and constant α; the solid line represents best-fit equation (12) with constant D and variable α.

Figure 2. Temporal variation in beetle disappearance rate (α of equation 12).

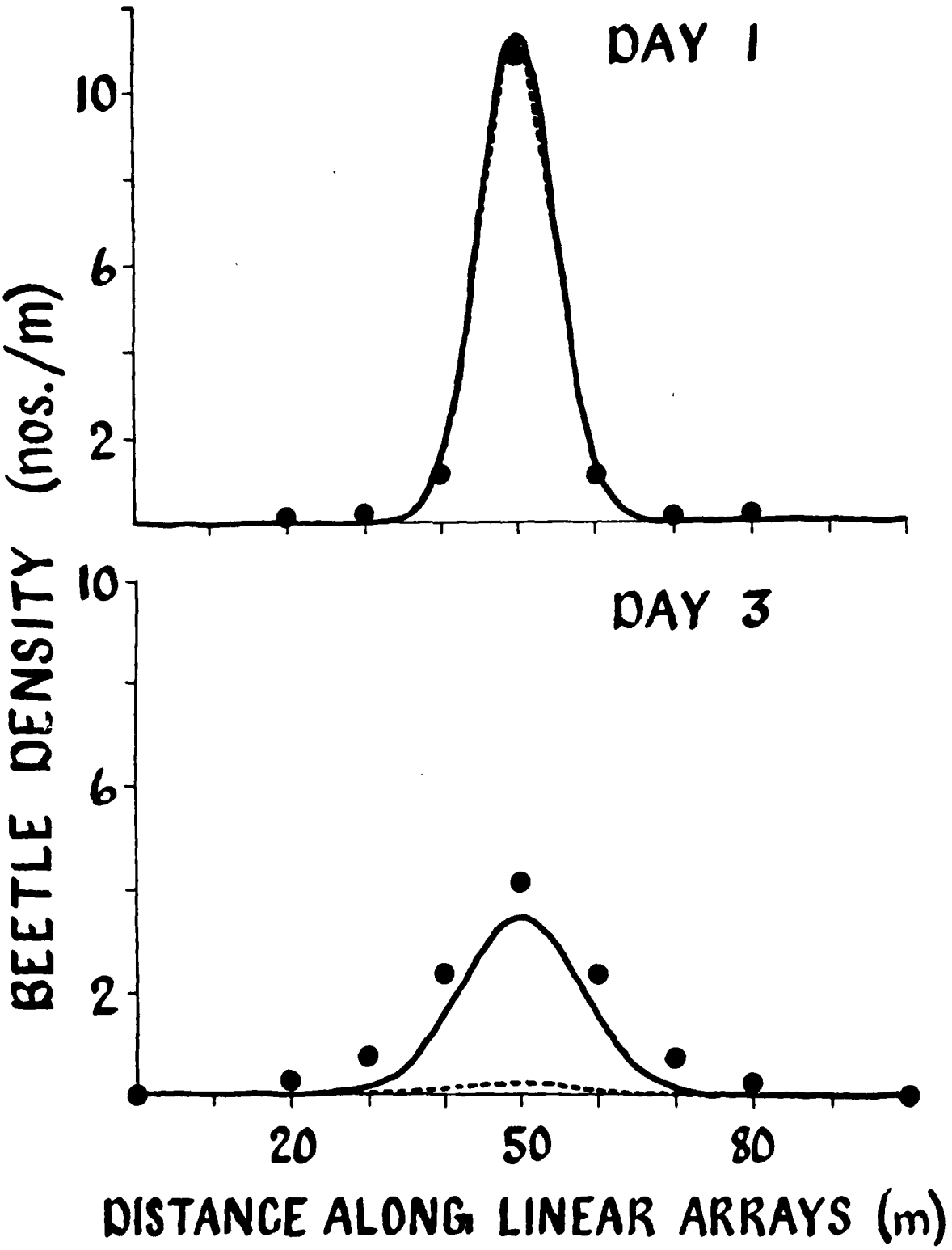

DAY 1
DAY 3
BEETLE DENSITY (nos./m)
10
6
2
10
6
2
20
50
80
DISTANCE ALONG LINEAR ARRAYS (m)

FIGURE 1

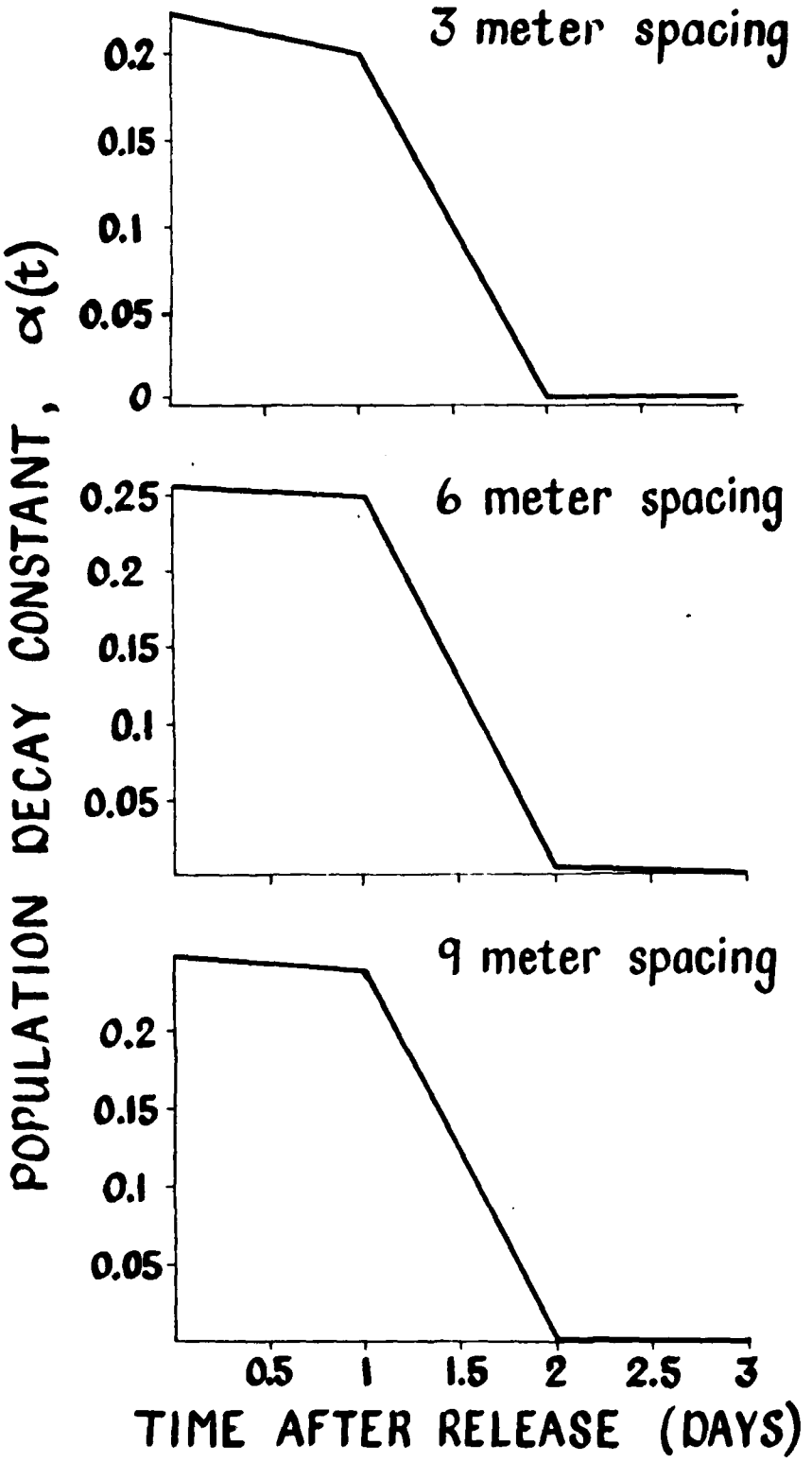

3 meter spacing
6 meter spacing
9 meter spacing
POPULATION DECAY CONSTANT, α(t)
0.2
0.15
0.1
0.05
0
0.25
0.2
0.15
0.1
0.05
0.2
0.15
0.1
0.05
0.5
1
1.5
2
2.5
3
TIME AFTER RELEASE (DAYS)

FIGURE 2

REFERENCES

[1] H. T. Banks, J. M. Crowley, and K. Kunisch, Cubic spline approximation techniques for parameter estimation in distributed systems, IEEE Trans. Auto. Control 28 (1983), 773-786.

[2] 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.

[3] H. T. Banks, P. L. Daniel and P. Kareiva, Estimation techniques for transport equations, Proc. Int'l. Conf. on Mathematics in Biology and Medicine (Bari, July 18-22, 1983), to appear.

[4] H. T. Banks and P. Kareiva, Parameter estimation techniques for transport equations with application to population dispersal and tissue bulk flow models, J. Math. Biology 17 (1983), 253-273.

[5] N. Draper and H. Smith, Applied Regression Analysis, Wiley and Sons, New York, 1966.

[6] R.J.V. Joyce, Insect flight in relation to problems of pest control, in R.C. Rainey, (ed.) Insect Flight, Royal Ent. Soc. Syms. 7 (1976), Blackwell, Oxford.

[7] P. Kareiva, Experimental and mathematical analysis of movement: quantifying the influence of plant spacing and quality on foraging discrimination, Ecol. Mon. 52 (1982), 261-282.

[8] P. Kareiva, Local movement in herbivorous insects: applying a passive diffusion model to mark-recapture field experiments, Oecologia 57 (1983), 322-327.

[9] P. Kareiva, Influence of vegetation texture on herbivore populations: resource concentration as herbivore movement, in Denno & McClure (eds.), Variable Plants and Herbivores in Natural and Managed Systems, Academic Press, New York, 1983.

[10] S. Kauffman, R. Shymko and K. Trabert, Control of sequential compartment formation in Drosophila, Science 199 (1978), 259-270.

[11] S. Levin, The role of theoretical ecology in the description and understanding of populations in heterogeneous environments, Amer. Zoologist 21 (1981), 865-875.

[12] A. Okubo, Diffusion and Ecological Problems: Mathematical Models, Biomathematics, Vol. 10, Springer, New York, 1980.

[13] S. Risch, D. Andow and M. Altieri,Agroecosystem diversity and pest control: data, tentative conclusions and new research directions, Env. Entomology, 12 (1983) 625-629.

[14] S. I. Rubinow, Introduction to Mathematical Biology, John Wiley and Sons, New York, 1975.

[15] M. H. Schultz, Spline Analysis, Prentice-Hall, Englewood Cliffs, 1973.

[16] T.R.E. Southland, Ecological Methods, Second Edition, Chapman and Hall, London, 1978.

[17] R. Stinner, C. Barfield, J. Stimac and L. Dohse, Dispersal and movement of insect pests, Ann. Rev. Entomology 28 (1983) 319-335.