

ESTIMATION OF GROWTH RATE DISTRIBUTIONS IN SIZE STRUCTURED POPULATION MODELS

By

H. T. BANKS (*University of Southern California, Los Angeles, California*)

AND

B. G. FITZPATRICK (*University of Tennessee, Knoxville, Tennessee*)

Abstract. We propose models for size-structured populations which allow growth rates to vary with individuals (growth rate distribution across all possible individual growth rates). A theoretical framework for the estimation of the growth rate distribution from data of sized population densities is developed. Numerical examples are presented to demonstrate feasibility of the ideas.

1. Introduction. In this paper we develop an inverse problem methodology for a class of size-structured population models which allow for "stochastic dispersion" in single cohort pulses of population. The models formulated and studied here are based on ideas discussed in [BBKW] which entail models wherein growth rates may vary across individuals of the population as well as with size and time. Simulation studies were presented in [BBKW] to demonstrate that such ideas could lead to population densities that exhibit dispersion and bimodality. Here we give a rigorous theoretical development of inverse problem methods and demonstrate the efficacy of these methods with computational examples.

As noted in [BBKW], our efforts on such problems were initiated in collaboration with biologists in studies related to the introduction of mosquito fish into rice fields, in attempts to control mosquito populations without chemicals. For such control problems, of course, it is very important to have models which predict the evolution of the populations accurately. In the paper [BVWLKRC], the mosquitofish population was modeled using the Sinko-Streifer (also called McKendrick-Von Foerster) model for size structured population density evolution. The Sinko-Streifer model,

Received January 30, 1990.

This research was supported in part under grants NSF DMS-8818530 and AFOSR-F49620-86-C-0111. Part of the research was performed while both authors were at the Center for Control Sciences, Division of Applied Mathematics, Brown University and while the first author was a visiting scientist at ICASE, NASA Langley Research Center, Hampton, VA, which is operated under NASA Contracts Nos. NAS 1-18107 and NAS 1-18605.

©1991 Brown University

hereafter referred to as (SS), is given by

$$\frac{\partial v}{\partial t} + \frac{\partial}{\partial x}(gv) = -\mu v \quad x_0 < x < x_1, \quad t > 0, \quad (1.1)$$

$$v(0, x) = \Phi(x), \quad (1.2)$$

$$g(t, x_0)v(t, x_0) = \int_{x_0}^{x_1} k(t, \xi)v(t, \xi) d\xi, \quad (1.3)$$

$$g(t, x_1) = 0. \quad (1.4)$$

The parameter g represents the growth rate of an individual, so that each individual grows according to $dx/dt = g(t, x)$. This individual growth equation defines the characteristics of (1.1). Problems arise in applying the model (SS) to observed data, however, because the data often exhibit features that the model cannot predict. One such phenomenon is a dispersion in size as time progresses. Figure 1.1 illustrates this dispersion in size for a particular mosquitofish population.

In fact, Figure 1.1 also shows another interesting phenomenon: the population begins with a unimodal density and actually develops into a bimodal density. One problem with (SS), as discussed in [B4], [BBKW], is that, for the model to achieve dispersion, we must have $\partial g/\partial x > 0$ (i.e., the characteristics must spread apart). This condition is unrealistic for many size structured biological systems of interest: it seems unreasonable for all larger individuals to grow at a faster rate than any smaller ones. Furthermore, even if the spread of the characteristics were biologically feasible, the transition from a unimodal to a bimodal density is qualitatively inconsistent with solutions of the Sinko-Streifer model.

It is our goal here to "modify" the model (SS) in order that it may exhibit features present in the data, and, further, to develop approximation ideas to facilitate estimation of model parameters from the data. In Sec. 2, we discuss a modeling approach to achieving dispersion: "mixing" densities that obey (1.1)–(1.4), only with differing growth rates. We use as the population density the function

$$u(t, x) = \int_G v(t, x; g) dP(g),$$

where G is a collection of growth rates and P is a probability measure on G , as suggested in [BBKW]. As shall be seen in Sec. 5, this model is rich enough to exhibit the phenomena of interest, namely dispersion and development of two modes from one. We discuss briefly this approach as it relates to some of the standard models in current use. Sec. 3 recalls some of the pertinent background on abstract least squares identification theory (with approximation), numerical solution and parameter identification for the Sinko-Streifer model, and weak convergence of probability measures. In Sec. 4, we bring the techniques of the previous section to bear on the identification problem for the proposed model. The goal is to prove convergence of parameter estimates for the approximate problems to those of the original problem. Numerical examples are discussed in Sec. 5. It is demonstrated in this section that our growth rate distribution model can predict the phenomena of interest in the data. We illustrate our identification techniques with computer-generated examples.

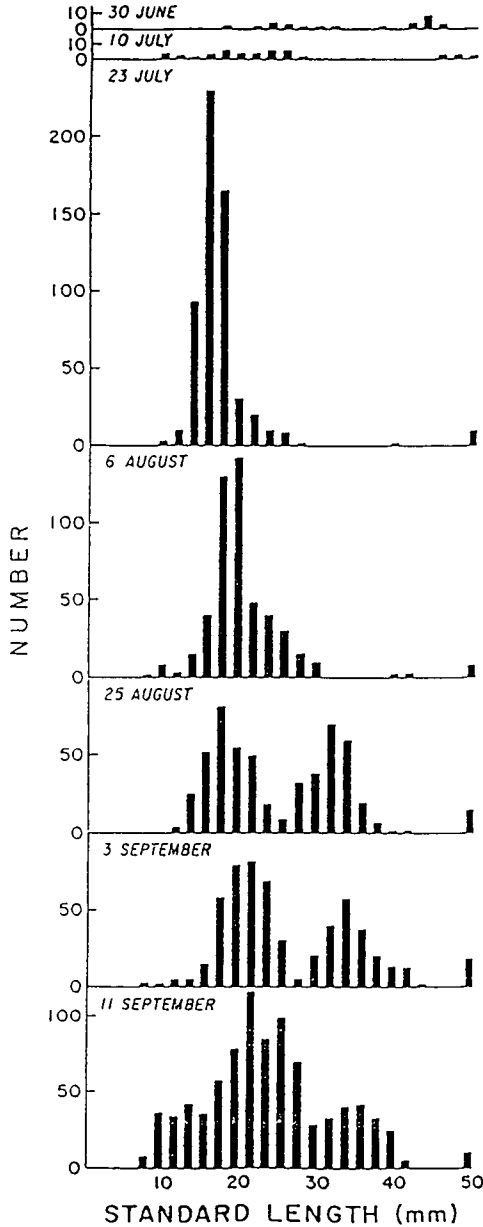


FIG. 1.1.

2. Modeling considerations. One modeling approach which can achieve dispersion is to treat the population under consideration as being composed of several subpopulations, each with its own growth rate. We denote these growth rates by g_1, \dots, g_M . For each i between 1 and M , we let p_i denote the proportion of individuals in the total population with growth rate g_i . These proportions are assumed to be independent of time. At time $t = 0$, we have an initial population density of $\Phi(x)$,

so that the subpopulations have initial densities $p_i \Phi(x)$, $1 \leq i \leq M$. We denote the subpopulation density by $\hat{v}(t, x; g_i)$, and we assume that, for each growth rate g_i , \hat{v} satisfies (1.1)–(1.4), with initial condition $p_i \Phi(x)$. A scaled version of $\hat{v}(t, x; g_i)$ is $v(t, x; g_i)$, which, for each i , satisfies (1.1)–(1.4) with initial condition $\Phi(x)$. To obtain the density u for the total population, we sum the subpopulation densities:

$$u(t, x) = \sum_{i=1}^M \hat{v}(t, x; g_i).$$

The density u can also be computed with respect to the functions v by

$$u(t, x) = \sum_{i=1}^M v(t, x; g_i) p_i.$$

This formula for u suggests a natural generalization, taking a population density of the form

$$u(t, x; P) = \int_G v(t, x; g) dP(g), \quad (2.1)$$

where G is a collection of growth rates for the subpopulations, and $P[A]$ is the proportion of the entire population which has a growth rate lying in $A \subset G$. We note that P is a probability measure, by virtue of the fact that $P[G] = 1$, but the model as derived is not a stochastic one. The fact that the total population is a mixture of subpopulations with different growth rates is embodied in the distribution P of growth rates. This model clearly differs from (SS) in that the individuals are no longer assumed to be identical. We have assumed here that the measure P does not depend on time. Thus, we have excluded some long term genetic effects. When applying the model to populations over a small number of generations, though, this lack of time dependence should not be a crucial feature.

It is interesting to note that one can also arrive at this type of model using stochastic assumptions. In the following derivation we assume that there are initially N individuals in the population. These individuals have initial sizes which are independent random variables, having densities $\Phi_j(x)$, $1 \leq j \leq N$, and growth rates which are independent, identically distributed random functions with distribution $P(dg)$. Let $X_1(t), \dots, X_N(t)$ be the sizes of the individuals at time t . We denote by $y(t, x, g)$ the solution of the initial value problem $dy/dt = g$, with $y(0) = x$. Suppose for the moment that individual j is known to have growth rate g . We define $F_{j,g}(a, b, t)$ by

$$F_{j,g}(a, b, t) = \Pr[a \leq X_j(t) \leq b].$$

It is clear that

$$\begin{aligned} F_{j,g}(y(t, a, g), y(t, b, g), t) &= F_{j,g}(a, b, 0) \\ &= \int_a^b \Phi_j(x) dx. \end{aligned}$$

Hence, $dF_{j,g}(y(t, a, g), y(t, b, g), t)/dt = 0$. Denoting by $v_j(t, x; g)$ the density of the distribution $F_{j,g}$, we see that

$$\begin{aligned} \frac{d}{dt}F_{j,g}(y(t, a, g), y(t, b, g), t) &= \frac{d}{dt} \int_{y(t, a, g)}^{y(t, b, g)} v_j(t, x; g) dx \\ &= \int_{y(t, a, g)}^{y(t, b, g)} \frac{\partial}{\partial t} v_j(t, x; g) dx \\ &\quad + g(t, y(t, b, g)) v_j(t, y(t, b, g); g) \\ &\quad - g(t, y(t, a, g)) v_j(t, y(t, a, g); g). \end{aligned}$$

Put $x = y(t, a, g)$ and $\Delta x = y(t, b, g) - y(t, a, g)$, so that $x + \Delta x = y(t, b, g)$. Rewriting the above equation, we have

$$\int_x^{x+\Delta x} \frac{\partial}{\partial t} v_j(t, \xi; g) d\xi + g(t, x + \Delta x) v_j(t, x + \Delta x; g) - g(t, x) v_j(t, x; g) = 0.$$

Dividing by Δx , and then letting $\Delta x \rightarrow 0$ ($b \rightarrow a$), we see that v_j satisfies the equation

$$\frac{\partial}{\partial t} v_j + \frac{\partial}{\partial x} (g v_j) = 0,$$

with initial data Φ_j . As we have assumed so far that g is known, the function $v_j(t, x; g)$ is the conditional density for the size of individual j , given the growth rate g . The density for the size of individual j is then given by

$$u_j(t, x) = \int_G v_j(t, x; g) dP(g).$$

Now, we can compute the expected number of individuals who have sizes between a and b at time t :

$$\begin{aligned} E[N_a^b(t)] &= \sum_{j=1}^N \Pr[a \leq X_j(t) \leq b] \\ &= \sum_{j=1}^N \int_a^b u_j(t, x) dx \\ &= \int_b^a \sum_{j=1}^N u_j(t, x) dx \\ &= \int_b^a \int_G \sum_{j=1}^N v_j(t, x; g) dP(g) dx \\ &= \int_b^a \int_G v(t, x; g) dP(g) dx, \end{aligned}$$

where the function v satisfies the equation

$$v_t + (g v)_x = 0,$$

with initial data $\Phi = \sum_{j=1}^N \Phi_j$. We define the density for the total expected population $u(t, x)$ by

$$u(t, x) = \int_G v(t, x; g) dP(g).$$

We have not yet, however, taken into account births or deaths. We model deaths by a death rate which is proportional to the size of the population, a model which leads to

$$\frac{d}{dt}u(t, x) = -\mu u(t, x).$$

This type of model is widely used: see, for example, [MD]. We assume here that this distribution of growth rates, P , does not vary in time, so that the proportion of individuals with growth rate g is independent of time. This in turn implies that v must satisfy the equation

$$\frac{d}{dt}v(t, x; g) = -\mu v(t, x; g).$$

Births into the population are modeled using a fecundity rate k . We consider for a moment the case of P having finite support, which is small in number when compared to the number of individuals in the population. Then $\int_a^b v(t, x; g) dx$ can be thought of in a formal sense as the expected number of individuals with sizes between a and b at time t , given they all have growth rate g . Now, the rate at which "newborns" or "recruits" enter the population is commonly modeled as (see [MD])

$$g(x_0)v(t, x_0; g) = \int_{x_0}^{x_1} k(t, \xi)v(t, \xi; g) d\xi.$$

This model assumes recruits have the same growth rate as their parents. We now have a conditional birth law, which gives us all the components of (1.1)–(1.4).

Of course, this is not the only differential equation model that exhibits dispersion. One such model is the equation

$$u_t + (gu)_x = (\mathcal{D}u)_x,$$

which contains a Fickian flux term. However, this type of model is based on gradient driven movement, which does not make sense for size structured populations: it seems unlikely that individuals would grow (or shrink) because there are fewer individuals at larger (or smaller) sizes. Thus, this model is an unlikely choice for the populations of interest.

Another model which is widely used in biological problems is the Fokker-Planck equation:

$$u_t + (gu)_x = (\mathcal{D}u)_{xx}.$$

The solution of this equation is known to be the density for a random process $X(t)$ which obeys the stochastic differential equation

$$dX = g dt + (2\mathcal{D})^{1/2} dW,$$

where $W(t)$ is a standard Brownian motion. This model is also derived in [O] as an approximation to the density of a random process $Y(t)$ which obeys a Markov

transition law. One feature of the process $X(t)$ is that it is nowhere differentiable in t . Of course, it is not clear that individuals grow according to a stochastic differential equation or a Markov transition law. However, this model may be a reasonable alternative for populations exhibiting dispersion in time. In future studies, we hope to compare the Fokker-Planck model to the growth rate distribution model proposed here.

3. Theoretical background. In the previous section, we discussed some motivation for the use of the model (2.1) with (1.1)–(1.4). We now begin the study of the inverse problem of determining the parameter P from observed data. In order to formulate the estimation problem, we shall use the least squares cost functional

$$J(P) = \sum_{k=1}^n \int_{x_0}^{x_1} [\hat{u}(t_k, x) - u(t_k, x; P)]^2 dx, \quad (3.1)$$

where $\hat{u}(t, x)$ is the observed density at time t and size x , and $u(t, x; P)$ is the (P -dependent) density from (2.1). To estimate P , we minimize J over an appropriate collection of probability measures. Solving this problem for a given set of data requires approximation of the infinite-dimensional system (1.1)–(1.4), and, if the collection of probability measures of interest is infinite dimensional, approximation of this parameter space as well. Furthermore, to discuss convergence and continuous dependence (on data) of the approximate parameters, we shall need a topology on the space of probability measures. In this section, we recall relevant results on abstract least squares found in [B2] and [BK], an approximation theory for the Sinko-Streifer model from [B1], [BBKW], and [BM], and finally the theory of weak convergence of probability measures, as discussed in [B3] and [EK].

We consider the problem of minimizing

$$J(q) = |Cu(q) - z|_Z^2 \quad (3.2)$$

over $q \in Q$, where $z \in Z$ is an observation, $u(q) \in \mathcal{H}$ is the parameter dependent system state, and C is a mapping from \mathcal{H} to Z . We also consider a sequence of approximations \mathcal{H}^N and Q^M to \mathcal{H} and Q , respectively. The approximate problem, then, is to minimize

$$J^N(q) = |Cu^N(q) - z|_Z^2 \quad (3.3)$$

over $q \in Q^M$. Conditions which guarantee convergence of solutions $\hat{q}^{N,M}$ of (3.3) to solutions of \hat{q} of (3.2) and their continuous dependence on observations are given in [B1] and [BK]. We restate the convergence theorem here (for details of continuous dependence on observations, see the discussions on “method stability” in [B1], [BK]).

THEOREM 3.1. Consider the following assumptions:

- (i) The sets Q, Q^M lie in a metric space \tilde{Q} , with Q, Q^M compact in the \tilde{Q} topology.
- (ii) If $q^M \in Q^M$ for each M , there exists $q^* \in Q$ and a subsequence q^{M_k} such that $q^{M_k} \rightarrow q^*$ in the \tilde{Q} topology.

- (iii) For any $q \in Q$, there exists a sequence $\{q^M\}$ such that $q^M \in Q^M$ and $q^M \rightarrow q$ in the \tilde{Q} topology.
- (iv) For any $q^k \rightarrow q$ in \tilde{Q} , we have $Cu^N(q^k) \rightarrow Cu(q)$ in Z as $N, k \rightarrow \infty$.
- (v) For each N , the mapping $q \rightarrow J^N(q)$ is continuous.

Then, under (i)–(v), we have that there exist minimizers $\hat{q}^{N,M}$ of J^N over Q^M . Furthermore, any subsequence of $\{\hat{q}^{N,M}\}$ has a convergent subsequence, and if \bar{q} is the limit of such a subsequence, then it is a minimizer of J over Q .

We now recall assumptions and results for the Sinko-Streifer model (1.1)–(1.4). These are based on the work in [BM].

(H1) We are given a sequence, \mathcal{H}^N , of finite-dimensional subspaces of $H^1(x_0, x_1)$. We take $\mathcal{H} = L^2(x_0, x_1)$ as the state space. The orthogonal projections $\Pi^N: \mathcal{H} \rightarrow \mathcal{H}^N$ satisfy

$$\Pi^N \phi \xrightarrow{H^1} \phi \quad \text{as } N \rightarrow \infty, \quad \forall \phi \in H^1(x_0, x_1).$$

Many commonly used approximating families (e.g., linear or cubic splines; see [S]) satisfy the hypothesis (H1).

(H2) The functions μ and k are known functions with

$$\mu \in L^2(x_0, x_1) \cap L^\infty(x_0, x_1), \quad \text{and} \quad k \in L^2([0, T] \times [x_0, x_1]),$$

with $|k(t, \cdot)|_{L^2} \leq \nu_2 < \infty$ for all $t \in [0, T]$.

Since the emphasis here is on the estimation of P , we shall treat μ and k as known. This is for simplicity's sake, and our arguments can be readily modified to accommodate estimation and approximation of μ and k (e.g., see [BM], [BBKW] for details on estimation of μ and k).

(H3) We have given an admissible set $G \subset H^1(x_0, x_1)$ of growth rate functions. This set G is a compact set (in H^1) of functions satisfying $g(x) > 0$ for $x \in [x_0, x_1)$, $g(x_1) = 0$, $g(x_0) \geq \nu_1$, for some positive constant ν_1 .

We remark that assuming time independence is also merely for simplicity; the arguments can be easily extended. It should furthermore be noted that we have not included approximation of G : the reason for this will be apparent in the next section.

For numerical purposes, it is convenient to consider the weak form of the Sinko-Streifer equation,

$$\langle v_t, \phi \rangle + \langle \mu v, \phi \rangle - \langle g v, D\phi \rangle - \phi(x_0)R(t, v) = 0, \quad \text{for all } \phi \in \mathcal{H}, \quad (3.4)$$

where the functional R is defined by

$$R(t, \psi) = \int_{x_0}^{x_1} k(t, \xi) \psi(\xi) d\xi.$$

If $v: [0, T] \rightarrow \mathcal{H}$ satisfies (3.4), with $v(0) = \Phi \in \mathcal{H}$, we refer to v as a weak solution of (1.1)–(1.4). The assumption (H1) provides an obvious approach to approximation by consideration of weak solutions v^N in the space \mathcal{H}^N :

$$\langle v_t^N, \phi \rangle + \langle \mu v^N, \phi \rangle - \langle g v^N, D\phi \rangle - \phi(x_0)R(t, v^N) = 0, \quad \text{for all } \phi \in \mathcal{H}^N. \quad (3.5)$$

The initial condition for the approximate problem is $v^N(0) = \Pi^N \Phi$.

Under the assumptions above, we have (e.g., see [BM]) that, for each $t \in [0, T]$,

$$v^N(t, \cdot; g^M) \xrightarrow{L^2} v(t, \cdot; g) \quad \text{as } M, N \rightarrow \infty,$$

whenever $g^M \rightarrow g$ in G , and that $g \rightarrow v(t, \cdot; g)$ is a continuous mapping from G into $L^2(x_0, x_1)$, for each t . These facts will be of importance in the convergence proofs, as we shall see in the next section.

The final portion of review concerns the theory of weak convergence of probability measures. We recall only the theorems and facts we need for our identification problem; a complete account can be found in [B3] and [EK].

Suppose S is a complete metric space with metric d . We denote by $\mathcal{P}(S)$ the space of probability measures on the Borel subsets of S . For any closed set F in S and $\varepsilon > 0$, we define

$$F^\varepsilon = \{x \in S : \inf_{y \in F} d(x, y) < \varepsilon\}.$$

Clearly $F \subset F^\varepsilon$. Also, if $P, P' \in \mathcal{P}(S)$, we put

$$\rho(P, P') = \inf\{\varepsilon > 0 : P[F] \leq P'[F^\varepsilon] + \varepsilon, \text{ for all closed } F \text{ in } S\}.$$

It is known that ρ is a metric on $\mathcal{P}(S)$, and that, in fact, $\mathcal{P}(S)$ is a complete metric space under ρ . Furthermore, if S is compact, then $\mathcal{P}(S)$ is also compact. Convergence of probability measures with respect to ρ is characterized in the following theorem.

THEOREM 3.2. Let (S, d) be a complete metric space, and let $\{P_n\} \subset \mathcal{P}(S)$, $P \in \mathcal{P}(S)$. The following are equivalent:

- (a) $\rho(P_n, P) \rightarrow 0$ as $n \rightarrow \infty$.
- (b) $\int_S f dP_n \rightarrow \int_S f dP$ for all bounded, uniformly continuous functions $f: S \rightarrow \mathfrak{R}$.
- (c) $P_n[A] \rightarrow P[A]$ for all Borel sets A with $P[\partial A] = 0$.

This theorem tells us that the ρ topology is that of convergence in distribution. It also coincides with the weak star topology on $\mathcal{P}(S)$ as a subset of the dual space of the space of bounded, uniformly continuous functions on S . Henceforth, we shall denote convergence of P_n to P with respect to ρ by $P_n \xrightarrow{\text{dist}} P$.

In our subsequent discussions, we shall make use of the following theorem in defining approximation spaces $\mathcal{P}^M(S)$ for $\mathcal{P}(S)$.

THEOREM 3.3. Suppose S is a complete, separable metric space. If \hat{S} is a countable, dense set in S , then the set

$$\widehat{\mathcal{P}} = \left\{ \sum_{i=1}^M a_i \delta_{s_i} : a_i \geq 0, \sum_{i=1}^M a_i = 1, M \in \mathcal{Z}^+, s_i \in \hat{S} \right\}$$

is a dense subset of $\mathcal{P}(S)$.

With these abstract results and with an approximation scheme for (SS) at hand, we now proceed to the problem of estimating the growth rate distribution P in (2.1) by least squares techniques.

4. Estimation and approximation of the growth rate distribution. Recall that we are interested in minimizing the least squares cost functional (3.1):

$$J(P) = \sum_{k=1}^n \int_{x_0}^{x_1} [\hat{u}(t_k, x) - u(t_k, x; P)]^2 dx.$$

The approach here is to use the approximation ideas for (SS) recalled above, together with the weak convergence theory for probability measures, to verify that the assumptions of Theorem 3.1 hold for our problem. We begin with the following theorem.

THEOREM 4.1. Suppose $v^N(t, \cdot; g) \xrightarrow{L^2} v(t, \cdot; g)$ for each t , uniformly in g in G , and that, for each t , $g \rightarrow v(t, \cdot; g)$ is a continuous mapping of G into $L^2(x_0, x_1)$, where G is as in assumption (H3). Suppose $P^M \xrightarrow{\text{dist}} P$ in $\mathcal{P}(G)$. Then

$$\int_G v^N(t, \cdot; g) dP^M(g) \xrightarrow{L^2} \int_G v(t, \cdot; g) dP(g),$$

as $N, M \rightarrow \infty$, for each t .

Proof. For each fixed t we have

$$\begin{aligned} & \left| \int_G v^N(t, \cdot; g) dP^M(g) - \int_G v(t, \cdot; g) dP(g) \right|_{L^2} \\ & \leq \left| \int_G (v^N(t, \cdot; g) - v(t, \cdot; g)) dP^M(g) \right|_{L^2} \\ & \quad + \left| \int_G v(t, \cdot; g) dP^M(g) - \int_G v(t, \cdot; g) dP(g) \right|_{L^2} \\ & \equiv \text{I} + \text{II}. \end{aligned}$$

Then using Jensen's inequality and Tonelli's theorem for the first term, we find

$$\begin{aligned} \text{I}^2 &= \int_{x_0}^{x_1} \left[\int_G (v^N(t, x; g) - v(t, x; g)) dP^M(g) \right]^2 dx \\ &\leq \int_{x_0}^{x_1} \int_G [v^N(t, x; g) - v(t, x; g)]^2 dP^M(g) dx \\ &= \int_G \int_{x_0}^{x_1} [v^N(t, x; g) - v(t, x; g)]^2 dx dP^M(g). \end{aligned}$$

Suppose $\varepsilon > 0$. Choose N_0 so that $N \geq N_0$ implies

$$\int_{x_0}^{x_1} [v^N(t, x; g) - v(t, x; g)]^2 dx < \varepsilon \quad \text{for all } g \in G.$$

Then for every M we have

$$\int_G \int_{x_0}^{x_1} [v^N(t, x; g) - v(t, x; g)]^2 dx dP^M(g) < \int_G \varepsilon dP^M(g) = \varepsilon.$$

Thus $\text{I} \rightarrow 0$ as $N \rightarrow \infty$, uniformly in M .

Considering the second term we have

$$\Pi^2 = \int_{x_0}^{x_1} \left[\int_G v(t, x; g) dP^M(g) - \int_G v(t, x; g) dP(g) \right]^2 dx.$$

We see as above, that

$$\int_{x_0}^{x_1} \left[\int_G v(t, x; g) dP(g) \right]^2 dx \leq \int_{x_0}^{x_1} \int_G v(t, x; g)^2 dP(g) dx,$$

where $v(t, \cdot; \cdot) \in L^2([x_0, x_1] \times G, dx \times dP(g))$. Hence, for each $\varepsilon > 0$, there exists $v_\varepsilon \in C([x_0, x_1] \times G)$ such that (e.g., see Theorem 3.14 of [R])

$$\int_{x_0}^{x_1} \int_G [v_\varepsilon(x, g) - v(t, x; g)]^2 dP(g) dx < \varepsilon^2.$$

Furthermore, we have

$$\begin{aligned} \Pi &= \left| \int_G v(t, \cdot; g) dP(g) - \int_G v(t, \cdot; g) dP^M(g) \right|_{L^2} \\ &\leq \left| \int_G v(t, \cdot; g) dP(g) - \int_G v_\varepsilon(\cdot, g) dP(g) \right|_{L^2} \\ &\quad + \left| \int_G v_\varepsilon(\cdot, g) dP(g) - \int_G v_\varepsilon(\cdot, g) dP^M(g) \right|_{L^2} \\ &\quad + \left| \int_G (v_\varepsilon(\cdot, g) - v(t, \cdot; g)) dP^M(g) \right|_{L^2} \\ &\equiv A + B + C. \end{aligned}$$

Suppose $\delta > 0$ is given and put $\varepsilon = \delta/4$. By definition of v_ε , $A < \varepsilon = \delta/4$. Also, v_ε is continuous on the compact set $[x_0, x_1] \times G$, so there exists M_ε such that $|v_\varepsilon(x, g)| \leq M_\varepsilon$. Since $P^M \xrightarrow{\text{dist}} P$, we have

$$\int_G v_\varepsilon(x, g) dP^M(g) \rightarrow \int_G v_\varepsilon(x, g) dP(g).$$

Hence, we may use the dominated convergence theorem to conclude that $B \rightarrow 0$ as $M \rightarrow \infty$. Choose M_0 such that $M \geq M_0$ implies $B < \varepsilon = \delta/4$. For each $M \geq M_0$ we have

$$\begin{aligned} C^2 &= \int_{x_0}^{x_1} \left[\int_G (v_\varepsilon(x, g) - v(t, x; g)) dP^M(g) \right]^2 dx \\ &\leq \int_{x_0}^{x_1} \int_G [v_\varepsilon(x, g) - v(t, x; g)]^2 dP^M(g) dx \\ &= \int_G \left\{ \int_{x_0}^{x_1} [v_\varepsilon(x, g) - v(t, x; g)]^2 dx \right\} dP^M(g). \end{aligned}$$

Now, $\int_{x_0}^{x_1} [v_\varepsilon(x, g) - v(t, x; g)]^2 dx$ is a continuous function of g , so we have that for $M \rightarrow \infty$

$$\begin{aligned} & \int_G \int_{x_0}^{x_1} [v_\varepsilon(x, g) - v(t, x; g)]^2 dx dP^M(g) \\ & \rightarrow \int_G \int_{x_0}^{x_1} [v_\varepsilon(x, g) - v(t, x; g)]^2 dx dP(g) < \varepsilon^2. \end{aligned}$$

Choose $M_1 \geq M_0$ such that $M \geq M_1$ implies

$$\int_G \int_{x_0}^{x_1} [v_\varepsilon(x, g) - v(t, x; g)]^2 dx dP^M(g) < 4\varepsilon^2.$$

Therefore $M \geq M_1$ implies $C < 2\varepsilon = \delta/2$, and hence $M \geq M_1$ implies $\text{II} = A + B + C < \delta$, so that $\text{II} \rightarrow 0$ as $M \rightarrow \infty$, and the theorem is proved.

This theorem is, in essence, a verification of hypothesis (iv) of Theorem 3.1. We now consider hypothesis (v).

THEOREM 4.2. Suppose $P^M \xrightarrow{\text{dist}} P$, and that for each N and each $t, g \rightarrow v(t, \cdot; g)$ is a continuous map $G \rightarrow L^2(x_0, x_1)$. Then, for each $N, J^N(P^M) \rightarrow J^N(P)$, where

$$J^N(P) = \sum_{k=1}^n \int_{x_0}^{x_1} [\hat{u}(t_k, x) - u^N(t_k, x; P)]^2 dx, \quad (4.1)$$

and

$$u^N(t, x; P) = \int_G v^N(t, x; g) dP(g).$$

Proof. Fix N and fix k with $1 \leq k \leq n$. Then we have

$$\begin{aligned} & \left| |\hat{u}(t_k, \cdot) - u^N(t_k, \cdot; P^M)|_{L^2}^2 - |\hat{u}(t_k, \cdot) - u^N(t_k, \cdot; P)|_{L^2}^2 \right| \\ & \leq \int_{x_0}^{x_1} |[\hat{u}(t_k, x) - u^N(t_k, x; P^M)]^2 - [\hat{u}(t_k, x) - u^N(t_k, x; P)]^2| dx \\ & \leq \int_{x_0}^{x_1} |u^N(t_k, x; P^M) - u^N(t_k, x; P)| \\ & \quad \times |2\hat{u}(t_k, x) - u^N(t_k, x; P^M) - u^N(t_k, x; P)| dx \\ & \leq |u^N(t_k, \cdot; P^M) - u^N(t_k, \cdot; P)|_{L^2} \\ & \quad \times \{2|\hat{u}(t_k, \cdot)|_{L^2} + |u^N(t_k, \cdot; P^M)|_{L^2} + |u^N(t_k, \cdot; P)|_{L^2}\}. \end{aligned}$$

We shall show that the first factor goes to zero, while the second factor is bounded above, for all M . Clearly the only term in the second factor of concern is the middle

one. For this term we find

$$\begin{aligned} |u^N(t_k, \cdot; P^M)|_{L^2}^2 &= \int_{x_0}^{x_1} \left[\int_G v^N(t_k, x; g) dP^M(g) \right]^2 dx \\ &\leq \int_{x_0}^{x_1} \int_G [v^N(t_k, x; g)]^2 dP^M(g) dx \\ &= \int_G \int_{x_0}^{x_1} [v^N(t_k, x; g)]^2 dx dP^M(g) \\ &\rightarrow \int_G \int_{x_0}^{x_1} [v^N(t_k, x; g)]^2 dx dP(g), \end{aligned}$$

since $\int_{x_0}^{x_1} [v^N(t_k, x; g)]^2 dx$ is a continuous function on the compact set G . Hence the second factor is bounded.

Furthermore, the first factor goes to zero as $M \rightarrow \infty$, by virtue of the same argument used in the proof of Theorem 4.1. Thus the proof is complete.

To verify hypotheses (i), (ii), and (iii) of Theorem 3.1 for our problem, we first observe that since G is compact in $H^1(x_0, x_1)$, $Q = \tilde{Q} = \mathcal{P}(G)$ is a compact metric space. Furthermore, since G is separable, choosing $\hat{G} = \{g_1, g_2, \dots\}$ dense in G , we have the collections $\mathcal{P}^M(\hat{G}) = \{\sum_{i=1}^M a_i \delta_{g_i} : \sum a_i = 1, a_i \geq 0\}$ dense in $Q = \mathcal{P}(G)$. Hence (i), (ii), and (iii) of Theorem 3.1 are satisfied. We summarize the results in the following theorem.

THEOREM 4.3. Suppose the hypotheses (H1), (H2), and (H3) hold for the system (1.1)–(1.4). Then there exist minimizers $\hat{P}^{N,M}$ of J^N over $\mathcal{P}^M(\hat{G})$. Moreover, any subsequence of this sequence of minimizers has a subsequence which converges to a minimizer of J over $\mathcal{P}(G)$.

5. Numerical experiments. In this section we report on our tests of the methods described above in estimating growth rate distributions P from size-structured population density data. For these computational tests we used simulated data and concentrated on estimation of P , letting $\mu = k = 0$ in (1.1)–(1.4) for simplicity (conceptually, estimation of μ and k in addition to P poses no added difficulties—see [BM], [BBKW]; of course, computational complexity increases but not to a level infeasible for currently available computing equipment and software). The examples discussed here in our presentation on inverse problem techniques are similar in nature to those used in the simulation studies presented in [BBKW].

The simulated data used in our tests were prepared in the following way. We are given an initial population density function Φ and we choose a “true” growth rate distribution $P^* \in \mathcal{P}(G)$. We then generate a random sample of growth rates g_1, g_2, \dots, g_L with distribution P^* . For each growth rate g_i , we compute the solution $v(t, x; g_i)$ of the (SS) system (1.1)–(1.4) with $\mu = k = 0$. We then define $\hat{u}(t, x) = \frac{1}{L} \sum_{i=1}^L v(t, x; g_i)$. By the law of large numbers, we have that for L sufficiently large, $\hat{u}(t, x)$ will be a good approximation for $u(t, x; P^*) = \int_G v(t, x; g) dP^*(g)$. We thus use this “noisy” data as our “observed data” in the inverse problems for $J(P)$ and $J^N(P)$ discussed in Sec. 4, assuming, of course, in

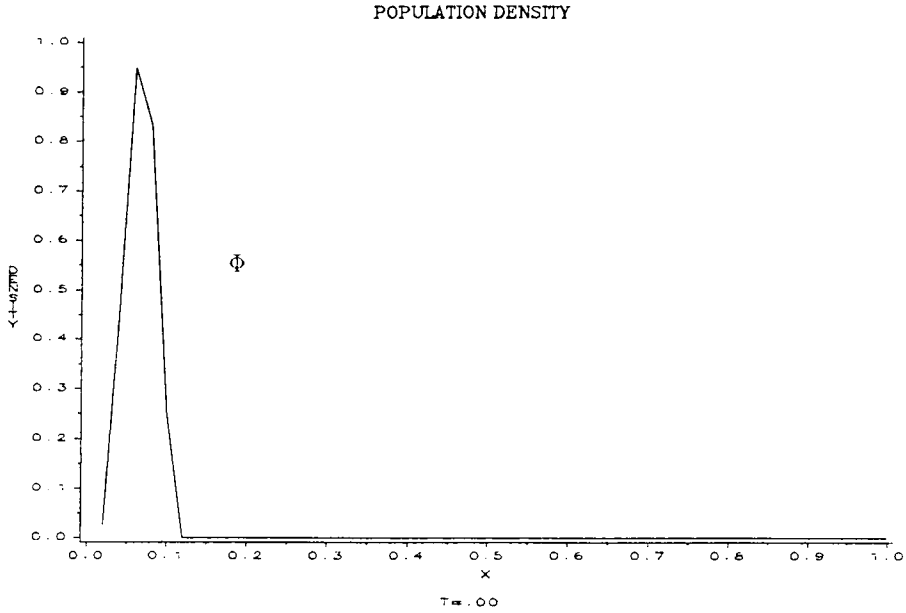


FIG. 5.1.

carrying out our tests that we do not know the distribution P^* that was used in generating the data.

In our tests we took $L = 200$ for our sample sizes and used "data" corresponding to a grid of $n = 50$ time observations, uniformly spaced in $[0, T] = [0, .2]$. The size range (x_0, x_1) was normalized to $(0, 1)$ and in each example we used the initial data

$$\Phi(x) = \begin{cases} \sin^2(10\pi x) & x \leq 0.1, \\ 0 & .1 < x, \end{cases}$$

which is depicted in Figure 5.1.

In all our examples, the sets G are defined by a parametrized family of growth rates

$$G = \{g(x) = b(1 - x) : .001 \leq b \leq 200\}$$

which is a compact subset of $H^1(0, 1)$ and is easily seen to satisfy hypothesis (H3) of Sec. 3. To define $P^* \in \mathcal{P}(G)$, we assume that the parameter b is stochastic, which yields in a natural way a distribution P^* . In our first example we choose for b a Gaussian distribution with mean $\hat{\mu} = 4.5$ and a variance $\sigma^2 = .25$ (in actuality, we used a Gaussian "truncated" to fit the restrictions $b \in [.001, 200]$). The evolution in time of the corresponding density $\hat{u}(t, x)$ is depicted in Figs. 5.2, 5.3, and 5.4 for $t = .06, .12, .18$, respectively.

In the approximate inverse problem for (4.1) with (3.5), for the state approximation spaces \mathcal{H}^N we chose spaces generated by standard piecewise linear splines (we used $N = 32$ in our computations) so that hypothesis (H1) is readily verified. To approximate $\mathcal{P}(G)$, we chose $\hat{G} = \{g \in G : b \text{ is rational in } [.001, 200]\}$. We then

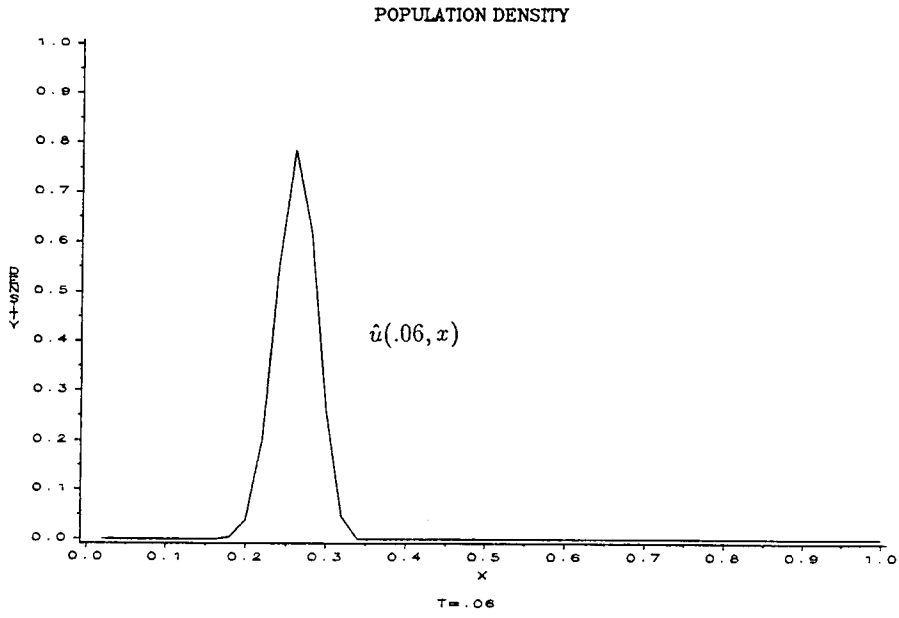


FIG. 5.2.

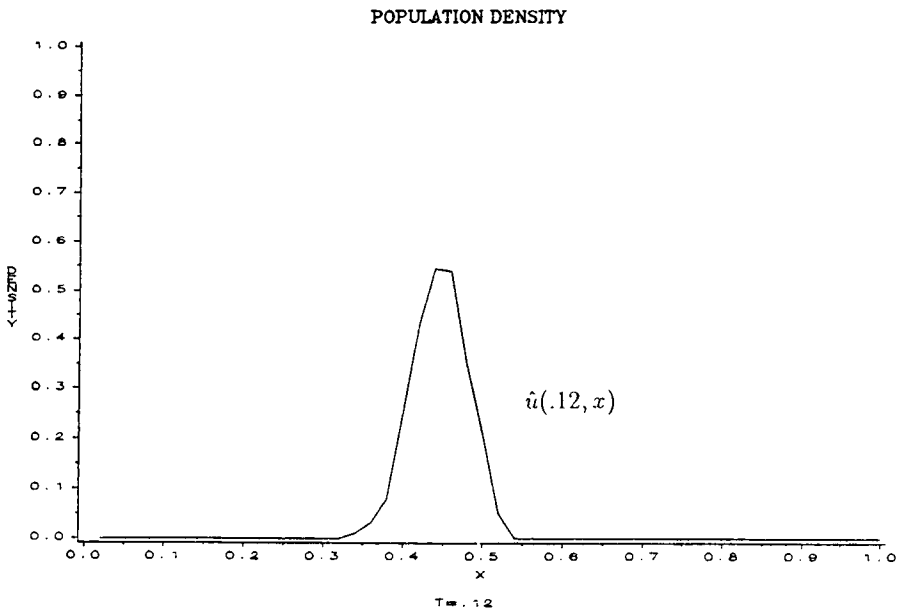


FIG. 5.3.

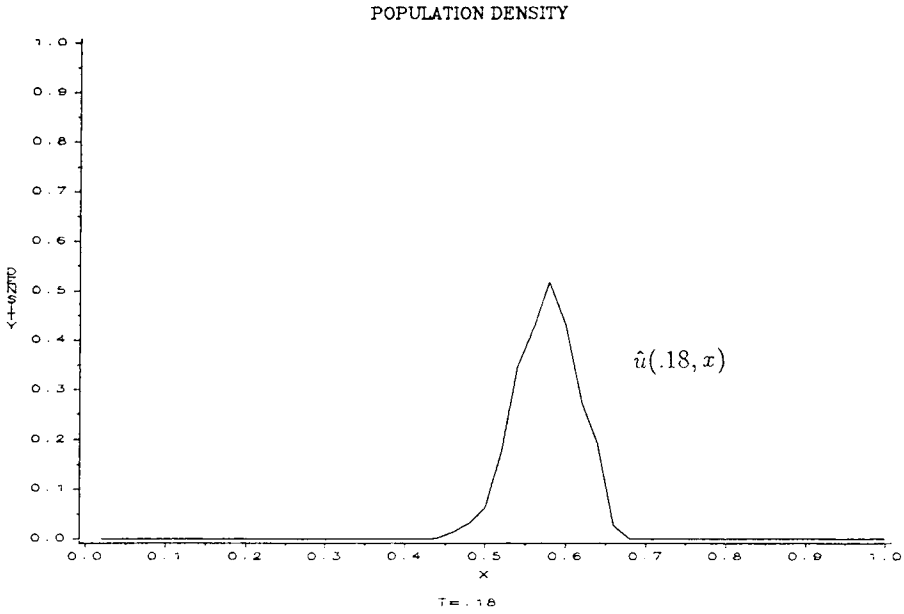


FIG. 5.4.

defined subsets of rationals $B^M = \{b_1^M, \dots, b_M^M\}$ of $[.001, 200]$ and took

$$\mathcal{P}^M(\hat{G}) = \left\{ P^M = \sum_{i=1}^M a_i \delta_{g_i^M} : g_i^M \text{ is a point mass at } b_i^M \in B^M, a_i \geq 0 \text{ and } \sum a_i = 1 \right\}.$$

In our tests we used $M = 9$ and $M = 21$ and for the sets B^M we chose $B^9 = \{b_i^9 = 4.5 + .3(i - 5) : i = 1, 2, \dots, 9\}$ and $B^{21} = \{b_i^{21} = 4.5 + .15(i - 11) : i = 1, 2, \dots, 21\}$.

We note that for these problems the cost functional (4.1) is given by

$$\begin{aligned} J^N(P^M) &= \sum_{k=1}^n \int_{x_0}^{x_1} [\hat{u}(t_k, x) - u^N(t_k, u; P^M)]^2 dx \\ &= \sum_{k=1}^n \int_{x_0}^{x_1} [\hat{u}(t_k, x) - \sum_{i=1}^M a_i v^N(t_k, x; g_i^M)]^2 dx, \end{aligned}$$

where the (a_1, \dots, a_M) satisfy $a_i \geq 0$, and $\sum_{i=1}^M a_i = 1$. Observe then that J^N is a quadratic function of the a_i . Hence, we can easily solve this minimization problem using Lagrange multipliers. All that is required is to solve the linear system $\partial J^N / \partial a_m = 1, m = 1, \dots, M$, and then scale to obtain $\sum a_i = 1$. Of course, we must also check the positivity constraint. The fact that we have reduced a nonlinear least squares problem in g to a linear problem in P is a very nice computational feature of this model.

In Figs. 5.5 and 5.6, we compare the “true” distribution P^* with the minimizing distributions $\hat{P}^{N, M}$ of Theorem 4.3 for $n = 32, M = 9, M = 21$, respectively.

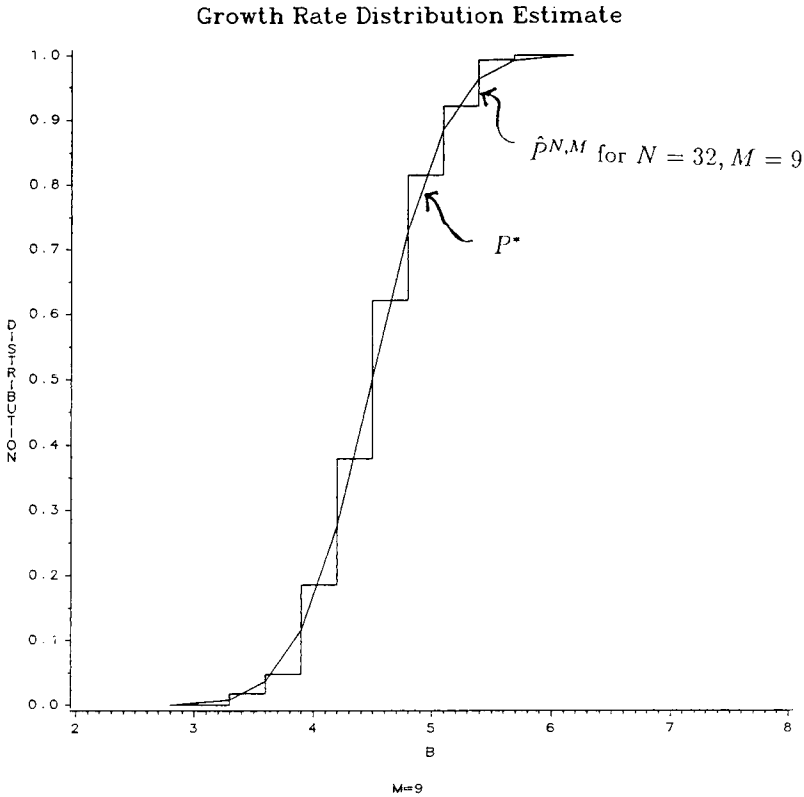


FIG. 5.5.

We also carried out numerical tests with simulated data exhibiting the bimodality observed in the mosquito fish data shown in Fig. 1.1. For these tests, data was generated as described above except in this case we assumed that the “true” growth rate distribution P^* was bimodal in nature. The set G was taken as above but now we assumed that b was a truncated bimodal Gaussian obtained by averaging two truncated Gaussians with means $\bar{\mu}_1 = 3.3$, $\bar{\mu}_2 = 5.7$, and variances $\sigma_1^2 = \sigma_2^2 = .25$. Using the same initial population density Φ given in Fig. 5.1, we generated “noisy” data using samples as described above. Graphs of the resulting data $\hat{u}(t, x)$ are shown in Fig. 5.7–5.9 (corresponding to $t = .06, .12, .18$, respectively) where the bimodal features of the population density can be clearly seen. The sets $P^M(\hat{G})$ and B^M were defined as above for these computations. In Fig. 5.10 we compare the resulting $\hat{P}^{N,M}$ for $N = 32$, $M = 21$ with the “true” distribution P^* .

6. Concluding remarks. The growth rate distribution added to the Sinko-Streifer model provides a new mechanism to capture features present in field data. We have seen through numerical experiments for inverse problems that dispersion and bimodality can be modeled accurately with this approach. Furthermore, the inverse problem can be solved quite simply by Lagrange multipliers. We have yet, however, to study field data with these methods, as this task will require more computational

Growth Rate Distribution Estimate

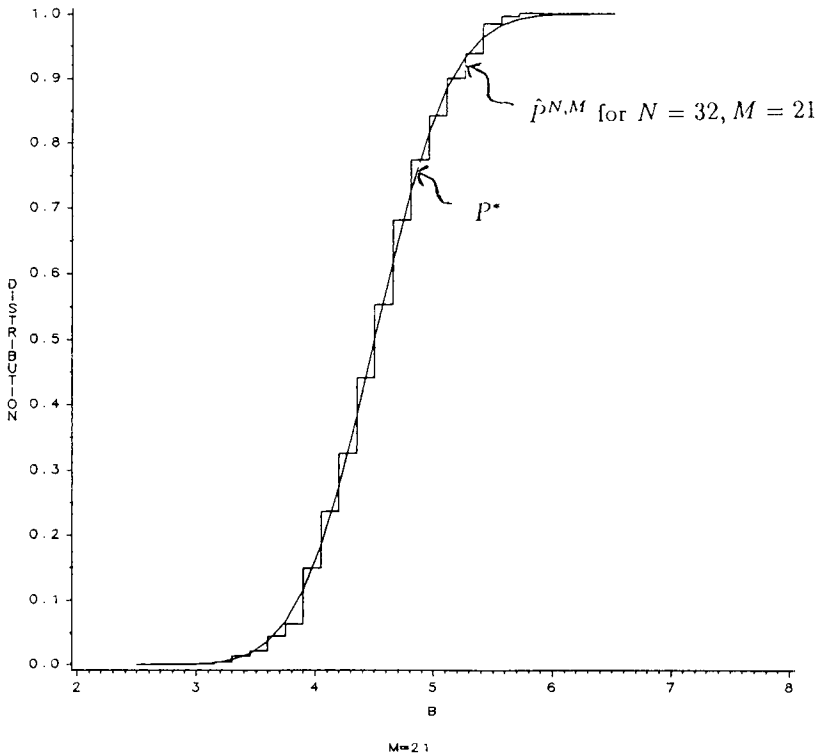


FIG. 5.6.

effort since we must incorporate the estimation of μ and k . Our presentation here was to demonstrate feasibility of our approach. We believe that these ideas afford promising alternatives to current approaches.

As we mentioned above, another question of interest involves comparing this growth rate distribution model to the Fokker-Planck diffusion model. In our continuing efforts to understand mechanisms as represented by features of field data, we are currently pursuing investigations related to the use of the statistical techniques of [BF] to compare these two models.

Extensions of our ideas must be considered: some involve the use of more general collections of growth rates (a one parameter family of growth rates makes for easy implementation, but at the expense of a highly simplified model of individual growth). Others entail the use of higher-order approximating elements for the distribution. We are currently studying the possibility of using linear spline approximations for the growth rate distribution. In the case of a smooth "true" distribution, one would expect faster convergence for piecewise linear functions over piecewise constant ones. The implementation of linear splines would not seem to introduce many more difficulties, but associated theoretical problems are currently under investigation.

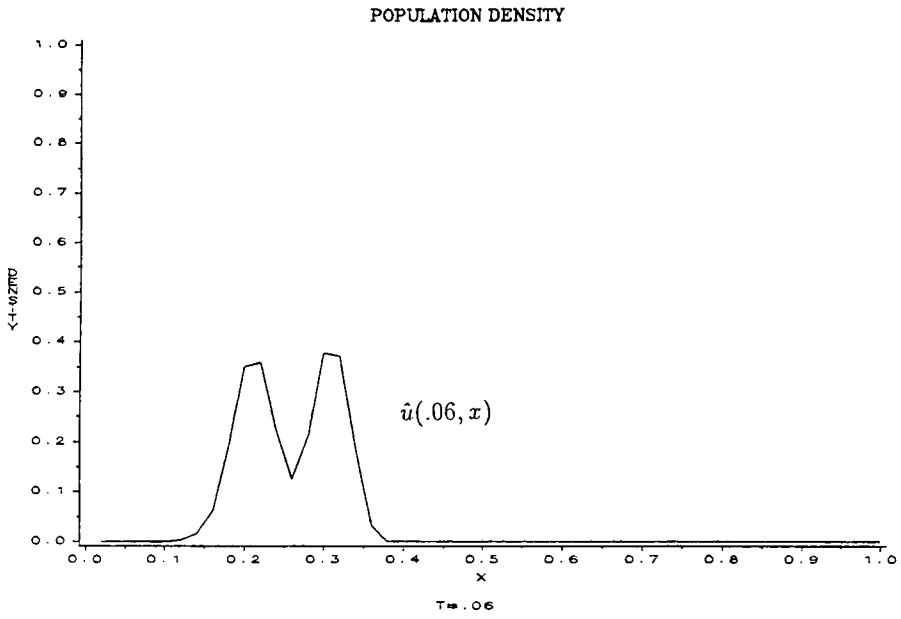


FIG. 5.7.

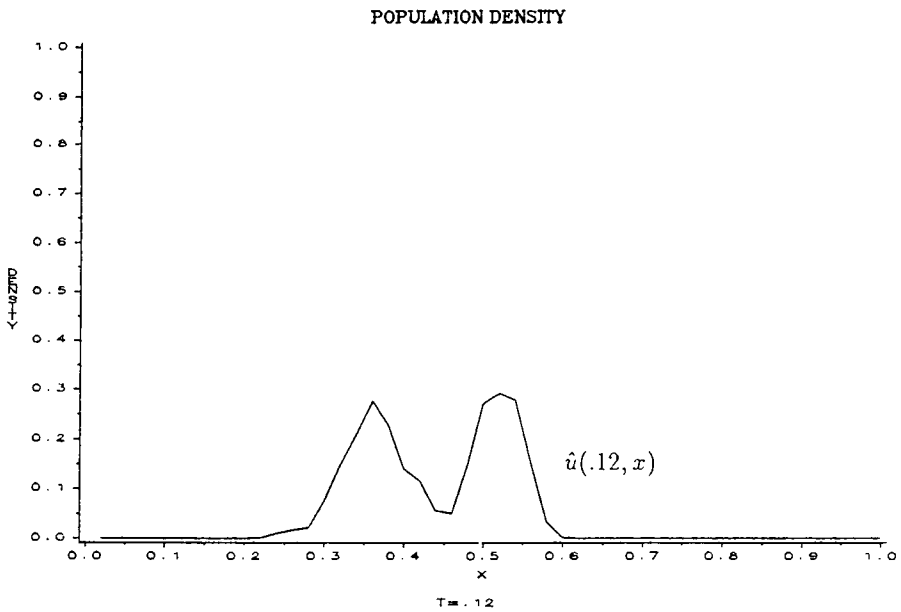


FIG. 5.8.

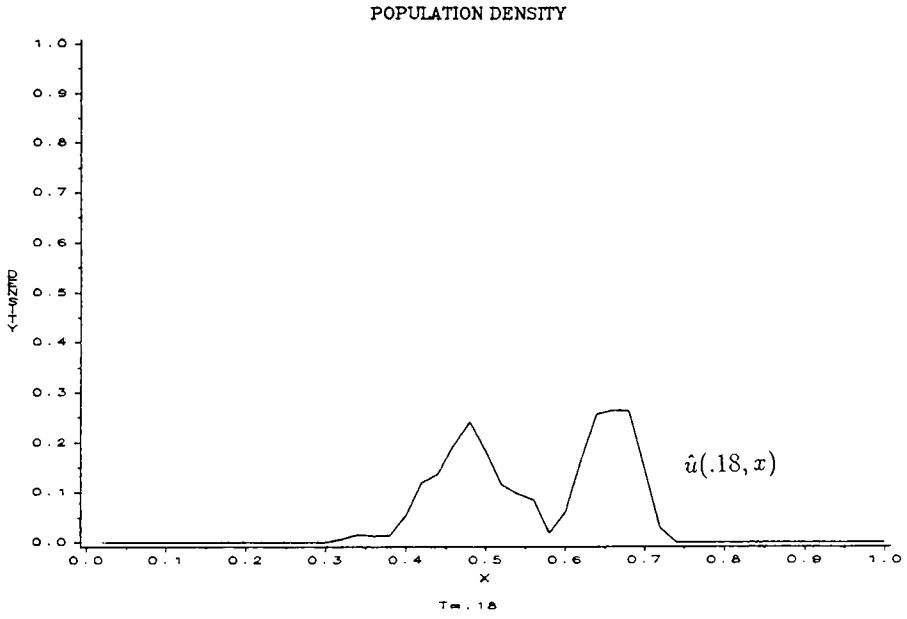


FIG. 5.9.

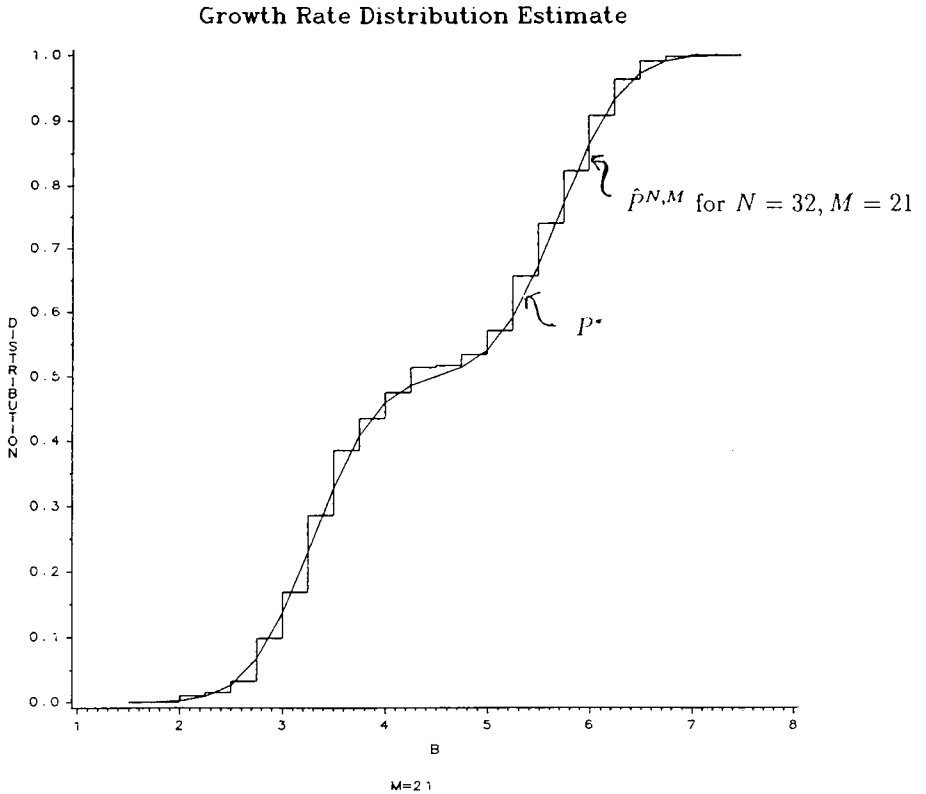


FIG. 5.10.

REFERENCES

- [B1] H. T. Banks, *On a variational approach to some parameter estimation problems*, Distributed Parameter Systems, Springer Lecture Notes in Control and Information Sciences, Vol. 75, 1985, pp. 1–23
- [B2] H. T. Banks, *Computational techniques for inverse problems in size structured stochastic population models*, LCDS-CCS Report 87-41, Brown University; Proceedings IFIP Conference on Optimal Control of Systems Governed by Partial Differential Equations, Santiago de Compostela, Spain (July 6–9, 1987), Springer Lecture Notes in Control and Information Sciences, Vol. 114, 1988, pp. 3–10
- [BBKW] H. T. Banks, L. W. Botsford, F. Kappel, and C. Wang, *Modeling and estimation in size structured population models*, LCDS-CCS Report 87-13, Brown University; Proceedings 2nd Course on Mathematical Ecology, (Trieste, December 8–12, 1986) World Press, Singapore, 1988, pp. 521–541
- [BF] H. T. Banks and B. G. Fitzpatrick, *Statistical methods for model comparison in parameter estimation problems for distributed systems*, CAMS Rep. 89–4, September 1989, University of Southern California; *J. Math. Biol.* **28**, 501–527 (1990)
- [BK] H. T. Banks and K. Kunisch, *Estimation Techniques for Distributed Parameter Systems*, Birkhäuser, Boston, 1989
- [BM] H. T. Banks and K. A. Murphy, *Quantitative modeling of growth and dispersal in population models*, LCDS Report 86–4, Brown University, Proc. Symp. on Math. Biology, Springer Lecture Notes in Biomath, Vol. 71, 1987, pp. 98–109
- [B3] P. Billingsley, *Convergence of Probability Measures*, Wiley, New York, 1968
- [B4] L. W. Botsford, *Effect of individual growth rates on expected behavior of the northern California Dungeness crab (*Cancer magister*) fishery*, *Can. J. Fish. Aquat. Sci.* **41**, 99–107 (1984)
- [BVWLKRC] L. W. Botsford, B. Vandracek, T. Wainwright, A. Linden, R. Kope, D. Reed, and J. J. Cech, *Population development of the mosquitofish, *Gambusia Affinis*, in rice fields*, *Envir. Biol. of Fishes* **20**, 143–154 (1987)
- [EK] S. N. Ethier and T. G. Kurtz, *Markov Processes: Characterization and Convergence*, Wiley, New York, 1986
- [MD] J. A. J. Metz and O. Diekmann, *The Dynamics of Physiologically Structured Populations*, Springer Lecture Notes in Biomathematics, Vol. 68, 1986
- [O] A. Okubo, *Diffusion and Ecological Problems: Mathematical Models*, Springer-Verlag, New York, 1980
- [R] W. Rudin, *Real and Complex Analysis*, McGraw-Hill, New York, 1966
- [S] M. H. Schultz, *Spline Analysis*, Prentice Hall, Englewood Cliffs, 1973