

A Comparison of Approximation Methods for the Estimation of Probability Distributions on Parameters

H.T. Banks and Jimena L. Davis

Center for Research in Scientific Computation
Box 8205
North Carolina State University
Raleigh, North Carolina 27695-8205

October 10, 2005

Abstract

In this paper, we compare two computationally efficient approximation methods for the estimation of growth rate distributions in size-structured population models. After summarizing the underlying theoretical framework, we present several numerical examples as validation of the theory. Furthermore, we compare the results from a spline based approximation method and a delta function based approximation method for the inverse problem involving the estimation of the distributions of growth rates in size-structured mosquitofish populations. Convergence as well as sensitivity of the estimates with respect to noise in the data are discussed for both approximation methods.

1 Introduction and Theoretical Background

In this paper we consider approximation methods for a general class of estimation or inverse problems wherein the quantity of interest is a probability distribution. In particular we assume that we have a parameter ($q \in Q$) dependent system with model responses $x(q)$ describing in some manner a population of interest. For data or observations, we are given a set of values $\{z_l\}$ for the expected values

$$\mathcal{E}[x_l(q)|P] = \int_Q x_l(q)dP(q)$$

with respect to an unknown probability distribution P describing the distribution of the parameters q over the population under investigation. We wish to use this data to choose from a given family $\mathcal{P}(Q)$ the distribution P^* that gives a best fit of the underlying model to the data. Here we formulate an ordinary least squares (OLS) version of the problem, but this is not essential to our results and one could equally well use a weighted least squares, a maximum likelihood estimator, etc., approach. Thus we seek to minimize

$$J(P) = \sum_l |\mathcal{E}[x_l(q)|P] - z_l|^2$$

over $P \in \mathcal{P}(Q)$. Even for simple dynamics for x_l , this is in general an infinite dimensional optimization problem and approximations that lead to computationally tractable schemes are desirable. That is, it is useful to formulate methods to yield finite dimensional sets $\mathcal{P}^M(Q)$ over which to minimize $J(P)$. Of course, we wish to choose these methods so that “ $\mathcal{P}^M(Q) \rightarrow \mathcal{P}(Q)$ ” in some sense. In this case we shall use the Prohorov metric [3, 15] of weak star convergence of measures to assure the desired approximation results. A general theoretical framework is given in [3] with specific results on the approximations we use here given in [2, 12]. Briefly, ideas for the underlying theory are as follows. One argues continuity of $P \rightarrow J(P)$ on $\mathcal{P}(Q)$ with the Prohorov metric (this is trivial in the cases considered here). If Q is compact (again, easily established in our case) then it is known that $\mathcal{P}(Q)$ is a complete metric space and is also compact when taken with the Prohorov metric. Moreover, if the approximation families $\mathcal{P}^M(Q)$ are chosen so that elements $P^M \in \mathcal{P}^M(Q)$ can be found to approximate elements $P \in \mathcal{P}(Q)$ in the Prohorov metric, then well-posedness (existence, continuous dependence of estimates on data, etc.) can be obtained.

The data $\{z_l\}$ available (which, in general, will involve longitudinal or time evolution data) determines the nature of the problem. The most classical problem (which we shall refer to as a *Type I* problem) is one in which individual longitudinal data is available for members in the population. In this case there is a wide statistical literature (in the context of hierarchical modeling, mixing distributions, mixed or random effects, mixture models, etc.) [14, 17, 18, 19, 20, 26, 27, 28, 29, 30, 33, 34] which provides theory and methodology for estimating not only individual parameters but also population level parameters and allows one to investigate both intra-individual and inter-individual variability in the population and data. In what we shall refer to as *Type II* problems one has only *aggregate* or population level longitudinal data available. This is common in marine, insect, etc., *catch and release* experiments [11] where one samples at different times from the same population but cannot be guaranteed of observing the same set of individuals at each sample time. This type of data is also typical in experiments where the organism or population member being studied is sacrificed in the process of making a single observation (e.g., certain physiologically based pharmacokinetic (PBPK) modeling [13, 21, 31] and whole organism transport models [11]). In this case one may still have dynamic (i.e., time course) models for individuals, but no individual data is available. Finally, the third class of problems which we shall refer to as *Type III* problems involves dynamics which depend explicitly on the probability distribution P itself. In this case one only has dynamics (*aggregate dynamics*) for the expected value

$$\bar{x}(t) = \int_Q x(t, q)dP(q)$$

of the state variable. No dynamics are available for individual trajectories $x(t, q)$ for a given $q \in Q$. Such problems arise in viscoelasticity and electromagnetics as well as biology [3, 5, 6, 12, 23]. While the

approximations we discuss in this paper are applicable to all three types of problems, we shall illustrate the computational results in the context of size-structured marine populations where the inverse problems are of Type II.

Finally, we note that in the problems considered here, one can not sample directly from the probability distribution being estimated and this again is somewhat different from the usual case treated in some of the statistical literature, e.g., see [35, 36] and the references cited therein.

2 The Growth Rate Distribution Model and Inverse Problem

We now turn our attention to a specific application of the general ideas discussed above. The motivating application for this work is the estimation of growth rate distributions for size-structured mosquitofish populations. Mosquitofish have been used in the place of pesticides as a way to control mosquito populations in rice fields. Biologists desire to correctly predict the growth and decline of the mosquitofish population in order to determine the optimal densities of mosquitofish to use in rice fields to control mosquito populations. Thus, a mathematical model that accurately describes the mosquitofish population would be beneficial in this application, as well as in other problems involving population dynamics and age/size-structured data.

Based on data collected from rice fields, a reasonable mathematical model would have to predict two key features that are exhibited in the data: dispersion and bifurcation (a unimodal density becomes a bimodal density) of the population density over time [4, 9, 10]. The growth rate distribution model, developed in [4] and [9], captures both of these features in its solutions. This model is a modification of the Sinko-Streifer model (used for modeling age/size-structured populations) which takes into account that individuals have different characteristics or behaviors. The standard Sinko-Streifer model (SS) for size-structured mosquitofish populations is given by

$$\frac{\partial v}{\partial t} + \frac{\partial}{\partial x}(gv) = -\mu v, \quad x_0 < x < x_1, \quad t > 0 \quad (1a)$$

$$v(0, x) = \Phi(x) \quad (1b)$$

$$g(t, x_0)v(t, x_0) = \int_{x_0}^{x_1} K(t, \xi)v(t, \xi)\partial\xi \quad (1c)$$

$$g(t, x_1) = 0. \quad (1d)$$

In this model, $v(t, x)$ represents the size (given in numbers per unit length) or population density, where t represents time and x represents the length of the mosquitofish. The growth rate of the individual mosquitofish is given by $g(t, x)$, where

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

for each individual, so that all mosquitofish of a given size have the same growth rate in this model. We also note that $\mu(t, x)$ represents the mortality rate of the mosquitofish. The function $\Phi(x)$ represents the initial size density of the population, while K represents the fecundity kernel. The boundary condition at $x = x_0$ is the recruitment, or birth, rate, while the boundary condition at $x = x_1 = x_{max}$ ensures the maximum size of the mosquitofish is x_1 . The SS model cannot be used as is to model the mosquitofish population because it does not predict dispersion or bifurcation of the population in time under biologically reasonable assumptions [4, 9]. However, by modifying the SS model so that the individual growth rates of the mosquitofish vary across the population (instead of being the same for all individuals in the population), one obtains a model, known as the growth rate distribution (GRD) model, which does in fact exhibit both dispersal in time of the mosquitofish population and the development of a bimodal density from a unimodal density (see [9, 10]).

In the growth rate distribution (GRD) model, the population density $u(t, x; P)$, discussed in [4] and

developed in [9], is actually given by

$$u(t, x; P) = \int_G v(t, x; g) dP(g). \quad (2)$$

Here G is a collection of admissible growth rates, P is a probability measure on G , and $v(t, x; g)$ is the solution of (1) with growth rate g . This model assumes that the population is actually made up of collections of subpopulations, where individuals in the same subpopulation have the same growth rate. Based on work in [9], solutions to this model exhibit both dispersion and bifurcation of the population density in time. For our considerations in this paper, we assume that the admissible growth rates have the form

$$g(x; b, \gamma) = b(\gamma - x)$$

for $x_0 \leq x \leq \gamma$ and zero otherwise, where b is the intrinsic growth rate of the mosquitofish and $\gamma = x_1$ is the maximum size. This choice is based on work in [4], where the idea of other properties related to the growth rates varying among the mosquitofish is discussed. Under the assumption of varying intrinsic growth rates and maximum sizes, we in fact assume that b and γ are random variables taking values in the compact sets B and Γ , respectively. A reasonable assumption is that both are bounded closed intervals. Thus we take

$$G = \{g(\cdot; b, \gamma) | b \in B, \gamma \in \Gamma\}$$

so that G is also compact in, for example, $C[x_0, X]$ where $X = \max(\Gamma)$. Then $\mathcal{P}(G)$ is compact in the Prohorov metric and we are in the framework outlined above. In the first set of examples, we chose a growth rate parameterized by the intrinsic growth rate b with $\gamma = 1$, leading to a one parameter family of varying growth rates g among the individuals in the population. We also assume here that $\mu = 0$ and $K = 0$ in order to focus on only the distribution of growth rates; however, distributions could just as well be placed on μ and K .

We next introduce two different approaches that can be used in the inverse problem for the estimation of the distribution of growth rates of the mosquitofish. The first approach, which has been discussed and used in [9] and [10], involves the use of delta functions. We assume that the probability distributions \mathcal{P}^M placed on the growth rates are discrete corresponding to a collection G^M with the form $G^M = \{g_k\}_{k=1}^M$ where $g_k(x) = b_k(1 - x)$, for $k = 1, \dots, M$. Here the $\{b_k\}$ are a discretization of B . For each subpopulation with growth rate g_k , there is a corresponding probability p_k that an individual is in subpopulation k . The population density $u(t, x; P)$ in (2) is then approximated by

$$u(t, x; \{p_k\}) = \sum_k v(t, x; g_k) p_k,$$

where $v(t, x; g_k)$ is the subpopulation density from (1) with growth rate g_k . We denote this delta function approximation method as DEL(M), where M is the number of elements used in this approximation.

While it has been shown that DEL(M) provides a reasonable approximation to (2), a better approach might involve techniques that will provide a smoother approximation of (2) in the case of continuous probability distributions on the growth rates. Thus, as a second approach, we chose to use an approximation scheme based on piecewise linear splines. Here we have assumed that P is a continuous probability distribution on the growth rates. We approximate P' using piecewise linear splines, which leads to the following approximation for $u(t, x; P)$ in (2):

$$u(t, x; \{a_k\}) = \sum_k a_k \int_B v(t, x; g) l_k(b) db,$$

where $g(x; b) = b(1 - x)$, $p_k(b) = a_k l_k(b)$ is the probability density for an individual in subpopulation k and l_k represents the piecewise linear spline functions. This spline based approximation method is denoted by SPL(M,N), where M is the number of basis elements used to approximate the growth rate

probability distribution and N is the number of quadrature nodes used to approximate the integral in the formula above. We used the composite trapezoidal rule for the approximation of these integrals [32].

We were then able to use the approximation methods DEL(M) and SPL(M,N) in the inverse problem for the estimation of the growth rate distributions. The least squares inverse problem that we wish to solve is

$$\begin{aligned} \min_{P \in \mathcal{P}^M(G)} J(P) &= \sum_{i,j} |u(t_i, x_j; P) - \hat{u}(t_i, x_j)|^2 \\ &= \sum_{i,j} (u(t_i, x_j; P))^2 - 2u(t_i, x_j; P)\hat{u}(t_i, x_j) + (\hat{u}(t_i, x_j))^2, \end{aligned} \quad (3)$$

where $\hat{u}(t, x)$ is the data and $\mathcal{P}^M(G)$ is the finite dimensional approximation to $\mathcal{P}(G)$. When using DEL(M), the finite dimensional approximation $\mathcal{P}^M(G)$ to the probability measure space $\mathcal{P}(G)$ is given by

$$\mathcal{P}^M(G) = \left\{ P \in \mathcal{P}(G) \mid P' = \sum_k p_k \delta_{g_k}, \quad \sum_k p_k = 1 \right\},$$

where δ_{g_k} is the delta function with an atom at g_k . However, when using SPL(M), the finite dimensional approximation $\mathcal{P}^M(G)$ is given by

$$\mathcal{P}^M(G) = \left\{ P \in \mathcal{P}(G) \mid P' = \sum_k a_k l_k(b), \quad \sum_k a_k \int_B l_k(b) db = 1 \right\}.$$

Furthermore, we are able to note that this least squares inverse problem (3) becomes a quadratic programming problem [9, 10]. Letting \mathbf{p} be the vector that contains $p_k, 1 \leq k \leq M$, when using DEL(M) or $a_k, 1 \leq k \leq M$, when using SPL(M,N), we let \mathbf{A} be the matrix with entries given by

$$A_{km} = \sum_{i,j} v(t_i, x_j; g_k) v(t_i, x_j; g_m),$$

\mathbf{b} the vector with entries given by

$$b_k = - \sum_{i,j} \hat{u}(t_i, x_j) v(t_i, x_j; g_k),$$

and

$$c = \sum_{i,j} (\hat{u}(t_i, x_j))^2,$$

where $1 \leq k, m \leq M$. In the place of (3), we now minimize

$$F(\mathbf{p}) \equiv \mathbf{p}^T \mathbf{A} \mathbf{p} + 2\mathbf{p}^T \mathbf{b} + c \quad (4)$$

over $\mathcal{P}^M(G)$. We note when using DEL(M) we also had to include the constraint

$$\sum_k p_k = 1,$$

while when using SPL(M,N) we had to include the constraint

$$\sum_k a_k \int_B l_k(b) db = 1.$$

However, in both cases, we were able to include these constraints in the programming of these two inverse problems.

3 Numerical Results

We next present and discuss several computational examples based on simulated data and performed in MATLAB involving the estimation of the probability distribution P on the growth rates of the size-structured mosquitofish population in order to demonstrate the validity of the theoretical results discussed earlier. The particular examples used here are based on previous formulations discussed in [4] and [9].

3.1 Convergence of DEL(M) and SPL(M,N) Estimated Probability Distributions

In order to estimate the probability distribution P^* on the growth rates of the size-structured mosquitofish population, we first began by preparing simulated population density data independent of the two approximation methods DEL(M) and SPL(M,N) used in the inverse problem. Since we were only concerned at this point with the estimation of P^* , we let $\mu = 0$ and $K = 0$ in the Sinko-Streifer model. We then chose a true distribution P^* on the growth rates $g(x)$, where $g(x) = b(1 - x)$ and b is the intrinsic growth rate of the mosquitofish. More specifically, we assumed that the intrinsic growth rate b of the mosquitofish is a random variable with distribution P^* . This allowed us to generate a collection of growth rates $G_n = \{g_1, g_2, \dots, g_n\}$, where we took $n = 200$ in order to get a good approximation of

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

and $G_n \subset G$. Our simulated data $u_d(t, x; P^*)$ was then created by first computing the solution $v(t, x; g_i)$ of the Sinko-Streifer model for each individual g_i using the method of characteristics and then computing

$$u_d(t, x; P^*) = \int_{G_n} v(t, x; g) dP^*(g)$$

using the Gauss-Legendre integration method ([32]). We took 50 uniformly spaced time observations, where the time interval was $[0, 0.5]$. The range of size values (x_0, x_1) was normalized to $(0, 1)$ and 50 uniformly spaced size values were used in this range for our simulated data. For the initial size density, we used

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

For our first set of data, we placed a “truncated” Gaussian distribution on b with mean $\bar{b} = 4.5$ and variance $\sigma^2 = 0.25$, where $b \in [\bar{b} - 3\sigma, \bar{b} + 3\sigma]$. Using this range of values for the intrinsic growth rates allows us to capture approximately 99% of the intrinsic growth rates. However, to ensure that this “truncated” Gaussian distribution was indeed a true distribution, we had to scale the weights used in the Gauss-Legendre integration method to ensure that

$$\int_{\bar{b}-3\sigma}^{\bar{b}+3\sigma} f(b; \bar{b}, \sigma) db = 1, \quad (5)$$

where $f(b; \bar{b}, \sigma)$ is the probability density function corresponding to the probability distribution P . Using the set of data generated with this “truncated” Gaussian distribution, we then performed the inverse problem using SPL(M,N) and DEL(M) to find estimates of the “true” probability density and distribution (under the assumption that the true probability distribution P was unknown) and then compared these estimates to the known density and distribution. As discussed earlier, this inverse problem is simplified to a quadratic programming problem when using least squares, where we minimize $\mathbf{p}^T \mathbf{A} \mathbf{p} + 2\mathbf{p}^T \mathbf{b} + c$.

Overall, we found that both SPL(M,N) and DEL(M) produced good approximations of the true Gaussian growth rate distribution using the simulated data; however, we also found that in some cases SPL(M,N) produced bad estimates of the probability distribution when M and N were not chosen correctly. In Figure 1, we have the results from the inverse problem using SPL(15,200), DEL(15), and DEL(45). We see from the results in Figure 1 that the spline based approximation method converges in

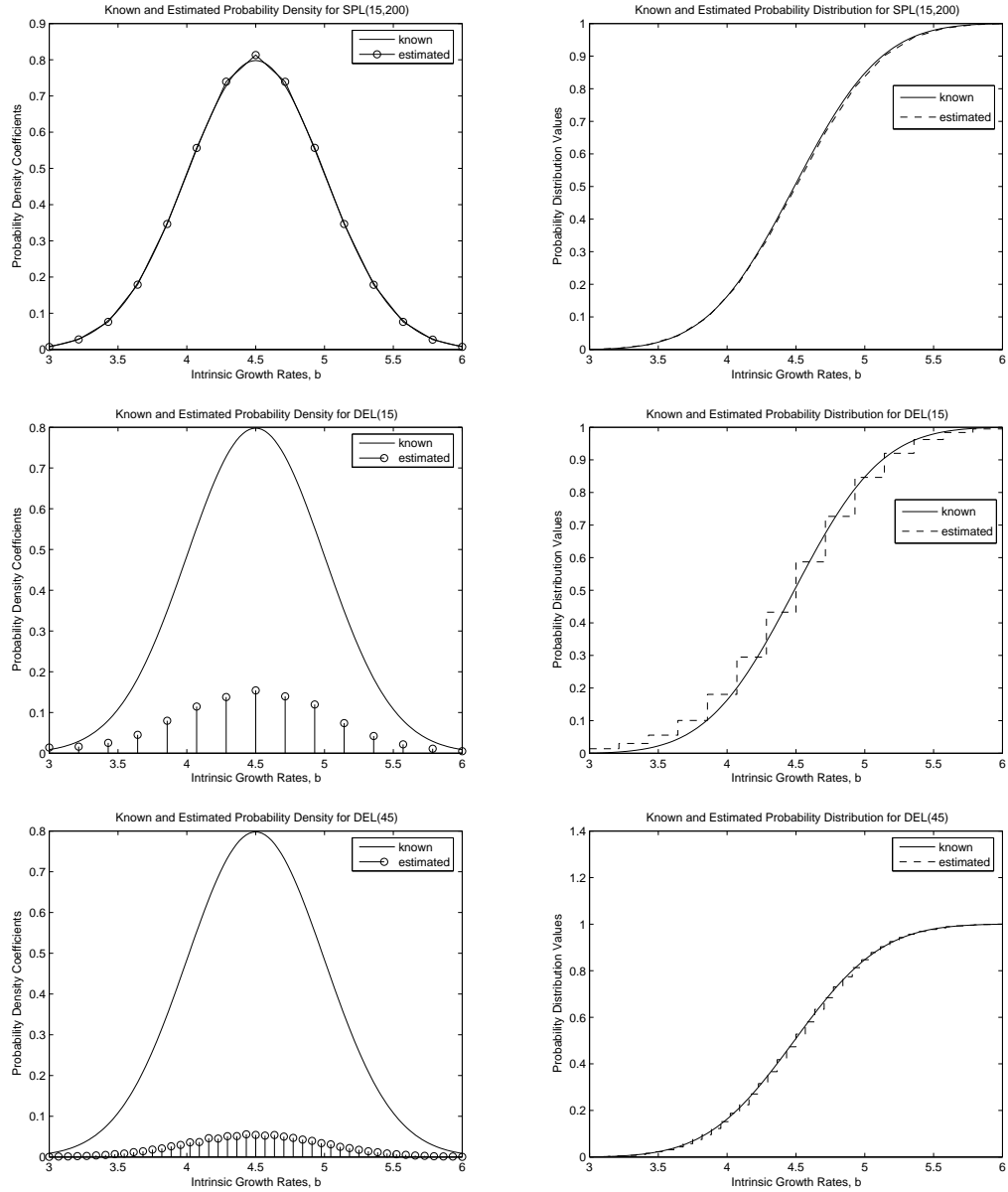


Figure 1: Estimates of probability densities and distributions for true gaussian distribution using a)SPL(15,200), b)DEL(15), c)DEL(45)

distribution much faster than the delta function approximation method for a given M , which we expect since the true distribution was smooth and continuous. For $M = 15$ and $N = 200$, the estimates of the probability distribution from SPL(M,N) have converged completely to the true distribution whereas the estimates of the probability distribution from DEL(M) have not quite converged. As M is increased, the estimated probability distributions from DEL(M) become better as seen in the results for $M = 45$. We also note that along with convergence in distribution, SPL(M,N) also provides convergence in density, whereas DEL(M) does not provide convergence in density. This may be attributed largely to the

difference in the constraints for these two methods. SPL(M,N) requires

$$\sum_k a_k \int_B l_k(b) db = 1,$$

where $p_k(b) = a_k l_k(b)$ is the probability density for an individual in subpopulation k and $l_k(b)$ represents the piecewise linear spline “functions.” On the other hand, DEL(M) requires

$$\sum_k p_k = 1,$$

where p_k denotes the probability coefficients and δ_{q_k} represents the delta functions. Since the true density was in fact smooth and continuous, one would not expect convergence in density when using DEL(M) because it is much cruder in its approximation of (5). We remark that this agrees fully with the underlying theory for convergence of distributions in the Prohorov metric wherein convergence of densities is not guaranteed.

We also used a second set of data in the inverse problem with a “truncated” Bi-Gaussian distribution on the intrinsic growth rates b . Based on previous work in [4], this type of distribution leads to data which exhibits both dispersion and bifurcation, which are two traits observed in actual mosquitofish field data [4, 9, 10]. In order to obtain a Bi-Gaussian distribution, we took the average of two Gaussian distributions, one with mean $\bar{b}_1 = 3.3$ and variance $\sigma_1^2 = 0.492$ and the second with mean $\bar{b}_2 = 5.7$ and variance $\sigma_2^2 = 0.492$. The simulated data was prepared in the same way as described above except with $b \in [\bar{b}_1 - 3\sigma_1, \bar{b}_2 + 3\sigma_2]$. The results for the inverse problem using this set of data are shown in Figure 2 for SPL(25,200), DEL(25), and DEL(85). Both methods do a good job of estimating the Bi-Gaussian probability distribution with the simulated data. Again, we see that SPL(M,N) converges to the true distribution faster than DEL(M). However, it should be noted that more basis elements (larger values of M) were required in both methods to achieve the same level of accuracy in approximating the Bi-Gaussian probability distribution in comparison to the Gaussian distribution. As mentioned earlier, the spline based approximation method results in both convergence in density and distribution for this example as well, while the delta function approximation method only results in convergence in distribution. We see that significantly more basis elements are required for full convergence of the approximations from DEL(M) to the true Bi-Gaussian probability distribution.

In the examples given above, SPL(M,N) produced good approximations to both the Gaussian and Bi-Gaussian densities and distributions. However, as we stated, obtaining good approximations when using SPL(M,N) was very much dependent on choosing M and N appropriately. In fact, we found if M and N were not chosen carefully, the estimates of the probability distributions from the inverse problem using SPL(M,N) were not very good as a result of the problem becoming unstable. However, by studying the condition number of the matrix A from the quadratic programming problem, we found that this behavior could be readily explained. There are several different ways in which the condition number $\kappa(A)$ of a matrix A can be described [32]. In terms of the norm $\|\cdot\|$ of a matrix, we have $\kappa(A) = \|A^{-1}\| \cdot \|A\|$. The condition number of a matrix A can also be defined as the ratio of the largest singular value to the smallest singular value in the singular value decomposition of the matrix A (see [32] for further discussion). What is of most importance is the information one learns from studying the condition number of a matrix. The matrix A is well-conditioned (well-behaved) if $\kappa(A)$ is relatively small. On the other hand, A is ill-conditioned (ill-behaved) if $\kappa(A)$ is relatively large. Thus, if $\kappa(A)$ is very large, meaning A is ill-conditioned, the inverse problem becomes unstable, which leads to poor approximations of the probability distribution P. We note that the discussion here is limited to SPL(M,N) based on the fact that the matrix A for the spline based method can become ill-conditioned for a given M based on the number of quadrature nodes used in the composite trapezoidal method used for integration purposes as discussed earlier. However, the matrix A in DEL(M) does not change for a given M due to the way in which the population density and A is obtained as discussed in the previous section. Since we must use a quadrature method to compute A for SPL(M,N), we expect the number of quadrature nodes N used to have a role in determining $\kappa(A)$. In fact, if N is chosen too small for a given M, meaning

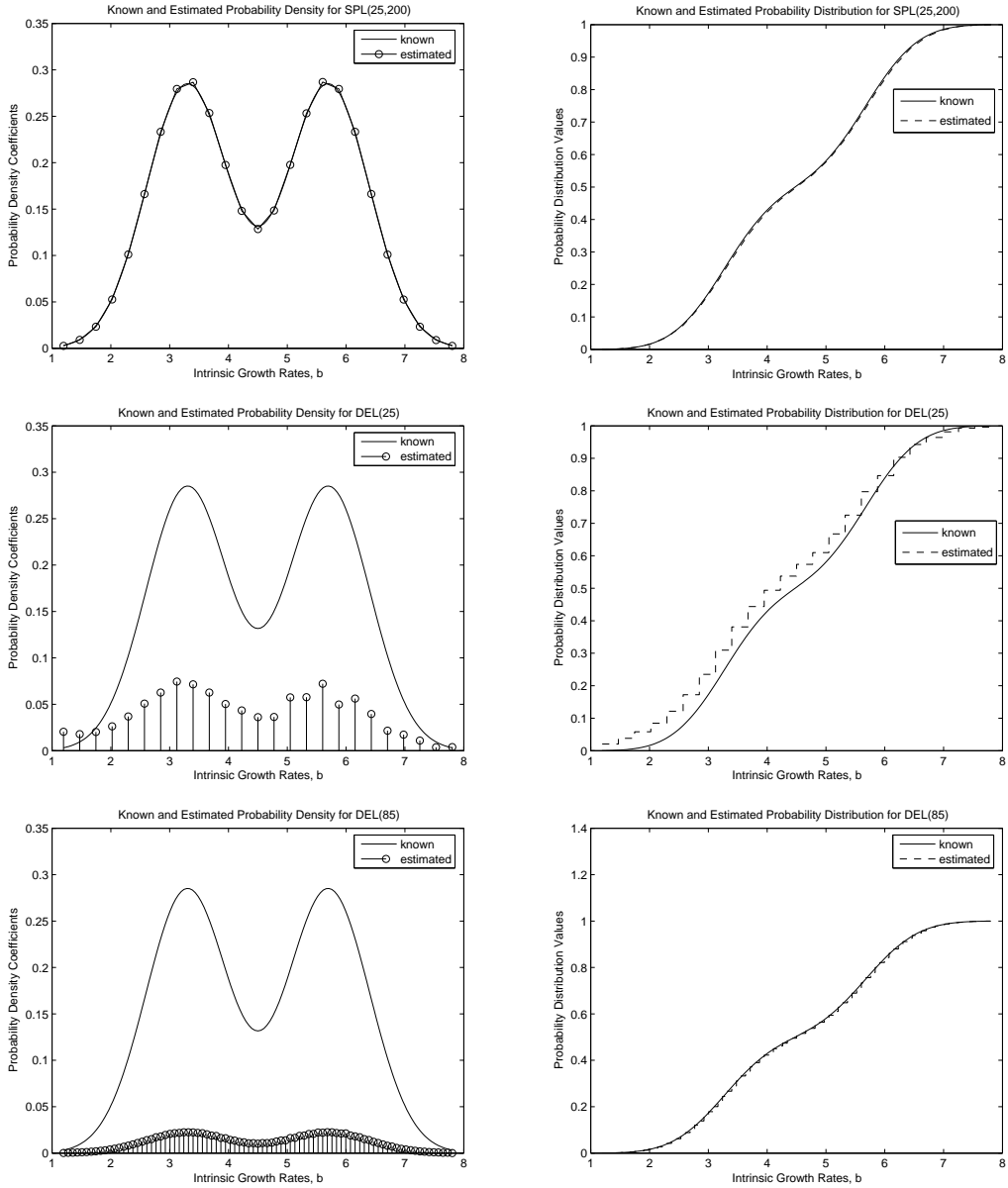


Figure 2: Estimates of probability densities and distributions for true bi-gaussian distribution using a)SPL(25,200), b)DEL(25), c)DEL(85)

the quadrature rule gives a very coarse approximation to the actual integration, then we expect $\kappa(A)$ to become larger, the problem to become unstable, and the estimates of the probability distribution to become poor.

To explore the validity of this argument, we computed the condition number of A for $M = 5, 25, 55, 75,$ and 95 and $N = 50, 100, 150, 200, 250,$ and 300 in both the Gaussian and Bi-Gaussian examples that were used above. The resulting condition numbers of A for the Gaussian example are given in Table 1. As we can see from the values in the table, when $M = 5$ and $25,$ $\kappa(A)$ is relatively small for all listed values of N . We found that the inverse problem produced good estimates of the Gaussian probability

M	N	$\approx \kappa(A)$
5	50,100,150,200,250,300	9.35
25	50,100,150,200,250,300	87
55	50	10^{17}
55	100,150,200,250,300	10^4
75	50	10^{19}
75	100,150,200,250,300	10^5
95	50	10^{18}
95	100,150,200,250,300	10^6

Table 1: Computed condition numbers of A for gaussian example when using SPL(M,N)

distribution when using SPL(5,N) and SPL(25,N). On the other hand, at $M = 55$, we began to see some significant differences in $\kappa(A)$, as shown by the values in Table 1. It can be noted that for $M = 55, 75$, and 95 , $\kappa(A)$ was very large when using SPL(M,50). However, there was a significant decrease in the condition number of A when using SPL(M,N) for $N = 100, 150, 200, 250, 300$ for these values of M. This difference in condition numbers was also evident in the estimates of the probability distribution from the inverse problem. For $M = 55, 75$, and 95 , the estimates when using SPL(M, 50) were worse than those obtained when using SPL(M,N) for all other listed values of N. We note the estimates of the Gaussian probability distribution became better as the value of N was increased for these fixed values of M.

M	N	$\approx \kappa(A)$
5	50,100,150,200,250,300	13.23
25	50,100,150,200,250,300	71
55	50	10^{16}
55	100,150,200,250,300	330
75	50	10^{17}
75	100,150,200,250,300	10^3
95	50	10^{19}
95	100,150,200,250,300	10^3

Table 2: Computed condition numbers of A for bi-gaussian example when using SPL(M,N)

For the Bi-Gaussian example, we also computed $\kappa(A)$ for the same values of M and N, and the results in this case are given in Table 2. The results obtained when using a “true” Bi-Gaussian probability distribution were very similar to the results obtained when using a “true” Gaussian probability distribution. At $M = 5$ and at $M = 25$, $\kappa(A)$ was relatively small (13.23 and 71, respectively, for each value of N). The inverse problem for these values of M was also stable, and the estimates of the probability distribution in both cases were good. However, when $M = 55, 75$, and 95 , SPL(M,50) results in large condition numbers for A. The estimates for the probability distribution using SPL(M,50) were very poor in comparison to the estimates obtained from SPL(M,N) for $M = 55, 75, 95$ and $N = 100, 150, 200, 250, 300$. When SPL(M,N) was used in the inverse problem for the estimation of the Bi-Gaussian probability distribution for $M = 55, 75, 95$ and for values of N greater than 50, we saw a decrease in the condition numbers of A (see values in Table 2) corresponding to better approximations of the probability distribution.

To summarize these computational results, in both the Gaussian and Bi-Gaussian case, when $N > M$, $\kappa(A)$ was relatively low, which resulted in good estimates of the growth rate probability distributions

when using SPL(M,N). However, when using SPL(M,N), for $M \sim N$, the condition number of A was very large, resulting in an ill-posed inverse problem and poor estimates of the probability distributions. For a fixed M, we observed that as the value of N increased, the condition number of A decreased, which agrees with results in [7] and [8]. Therefore, we have shown by these computational efforts that we can “regularize” the inverse problem when using SPL(M,N) by choosing proper ratios of M and N, which is known as “regularization by discretization balance.” By using a finer discretization in the quadrature method used in SPL(M,N), we were able to obtain better results in the inverse problem involving the estimation of growth rate distributions.

3.2 Sensitivity of DEL(M) and SPL(M,N) Estimated Probability Distributions

In the previous subsection, we discussed the results from two examples using simulated data in the estimation of probability distributions P on the growth rates of size-structured mosquitofish populations. However, data collected from an experiment is usually corrupted by noise, which can be a result of errors in collecting the data, errors in the instruments and techniques used, etc. Along with verifying that both SPL(M,N) and DEL(M) produce estimates which converge in distribution when simulated data with no noise is used in the inverse problem, we also wanted to be able to make some remarks about the sensitivity with respect to noisy data of the estimates of the probability distributions from the two approximation methods. Thus, we added random absolute noise to the simulated data used in the previous two examples in the following way:

$$\hat{u}(t, x; P^*) = u(t, x; P^*) + \eta \cdot \epsilon,$$

where η represents the noise level constant and ϵ represents normally distributed random values with mean 0 and variance 1. We then performed the inverse problem again using $\eta = 0.005, 0.025, 0.05$ corresponding to 1%, 5% and 10% absolute error, respectively, for both the Gaussian and Bi-Gaussian cases.

We begin by discussing the results of the inverse problem using the simulated data with a true “truncated” Gaussian distribution with the various noise level constants. Both approximation methods, SPL(M,N) and DEL(M), performed decently in the inverse problem with the varying percentages of absolute error. With only 1% absolute error, both DEL(M) and SPL(M,N), with M and N chosen appropriately, resulted in estimates that converged to the true growth rate probability distribution in very much the same manner as discussed above in the Gaussian example with no noise. The performance of these approximations methods was not greatly effected by the small amount of noise in the data. With a slightly larger percentage of absolute error in the simulated data, SPL(M,N) and DEL(M) were still able to produce good estimates of the probability distribution. However, the results from the inverse problem using $\eta = 0.025$ began to exhibit some small effects in the estimates obtained from both DEL(M) and SPL(M,N). For example, in Figure 3, the approximated probability distributions for DEL(25) and SPL(25,200) with 5% absolute error reveal slightly overestimated front tails. Moreover, SPL(25,200) with 5% absolute error resulted in small perturbations in the approximated density, which was very smooth when no noise was present in the data. With very noisy data, $\eta = 0.05$, SPL(M,N) and DEL(M) still perform fairly well. From the results for DEL(25) and SPL(25,200), shown in Figure 3, the noisier data resulted in only slightly poorer approximations in comparison to those obtained with 5% absolute noise. Moreover, the larger amount of noise produced more oscillatory behavior in the approximated probability densities for both DEL(M) and SPL(M,N), which would result in poorer approximations of the corresponding probability distributions.

We also tested the two approximation methods for sensitivity to error in the Bi-Gaussian data with the same percentages of absolute error. The results obtained from the inverse problem using DEL(M) and SPL(M,N) with noisy data with a “true” Bi-Gaussian distribution were very similar to those obtained when using noisy data with a “true” Gaussian distribution. Overall, the estimated probability distributions from DEL(M) and SPL(M,N) were not largely affected by the various amounts of noise added to the simulated data. Both methods were able to produce good approximations of the probability distributions in the presence of noise. With 1% absolute error in the data, the estimates of the growth rate distributions

from DEL(M) and SPL(M,N) did not change significantly from the estimates obtained when there was no noise in the data. We were still able to obtain convergence in distribution (with faster convergence when using SPL(M,N)) with both approximation methods. When the percentage of absolute error in the data was 5%, DEL(M) and SPL(M,N) still performed well by producing good estimates of the Bi-Gaussian probability distribution. The small amount of noise had some small effect on the estimates as seen in Figure 4; in fact, we see that the front tails in both the estimate from DEL(35) and SPL(35,200) for $\eta = 0.025$ are slightly larger than the tails for the “true” distribution. When even more noise is present in the data, the estimated probability distributions became slightly poorer for a fixed M and N. In Figure 4, the estimated probability densities and distributions are shown for DEL(35) and SPL(35) for data with 10% absolute error. It is clear from these plots that the estimates from DEL(M) and SPL(M,N) are indeed affected by the noisier data. As in the Gaussian case, we noticed some oscillatory behavior in the estimated probability densities from these two approximation methods as the amount of noise present in the data increased. Moreover, the front tails in the estimated probability distributions are overestimated, whereas the end tails are underestimated for both DEL(35) and SPL(35,200).

While we were still able to obtain convergence in distribution using both DEL(M) and SPL(M,N), with M and N chosen carefully, our computational results showed that as more noise was added to both the Gaussian and Bi-Gaussian data, the estimates of the growth rate distributions from both methods became relatively poorer for a fixed M and N. SPL(M,N) produced probability distribution estimates that converged faster in distribution than DEL(M) using both data with noise and without noise. This behavior, as stated earlier, was expected since the “true” probability distributions in these numerical examples are smooth and continuous. Although some of our computational evidence for SPL(M,N) also exhibited convergence in density as well as distribution, convergence in density is generally not guaranteed and is not supported by the theory underlying this work [1, 2, 3, 12].

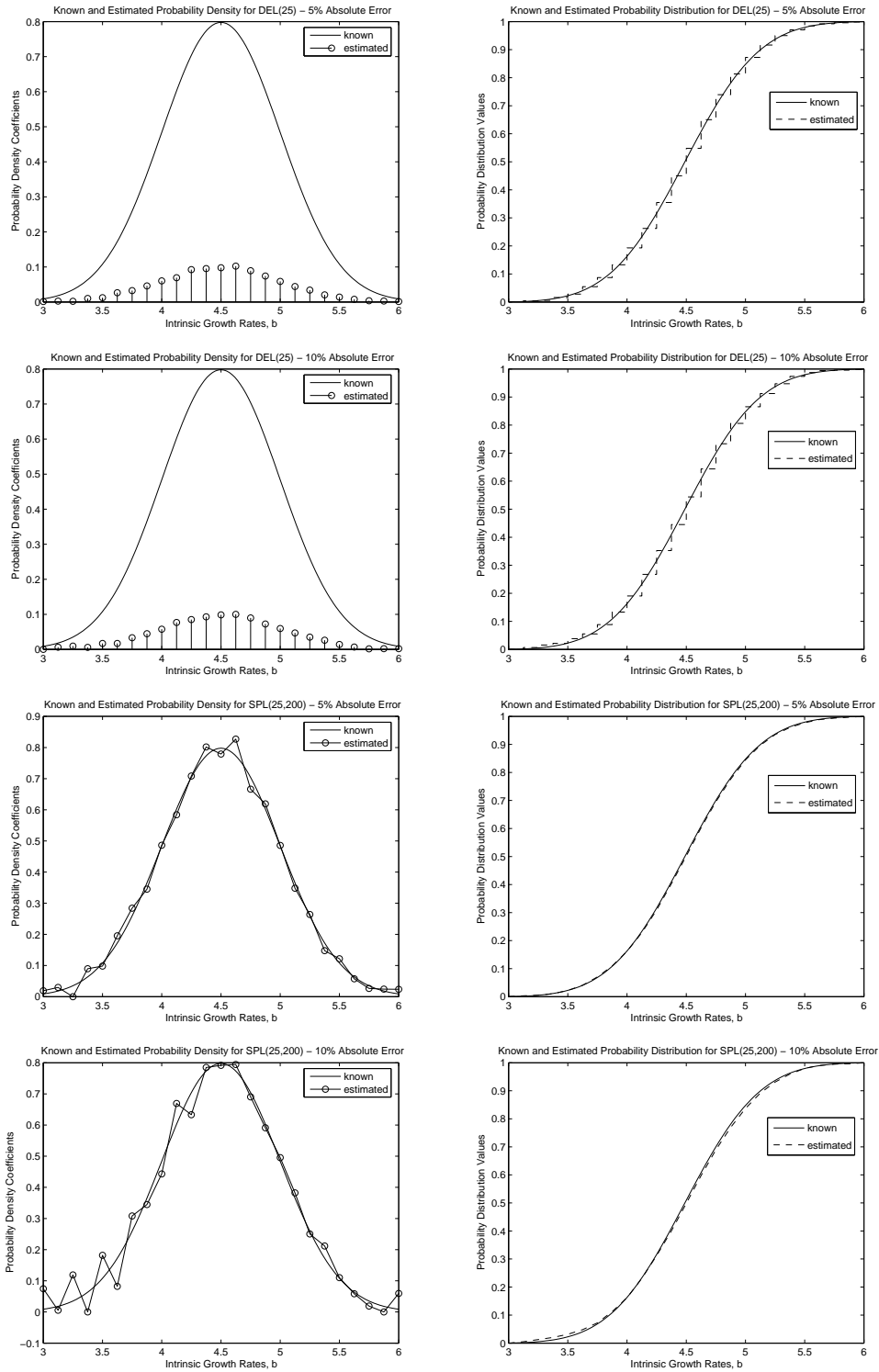


Figure 3: Estimates of probability densities and distributions for true gaussian distribution using a)DEL(25) with 5% absolute error, b)DEL(25) with 10% absolute error, c)SPL(25,200) with 5% absolute error, d)SPL(25,200) with 10% absolute error

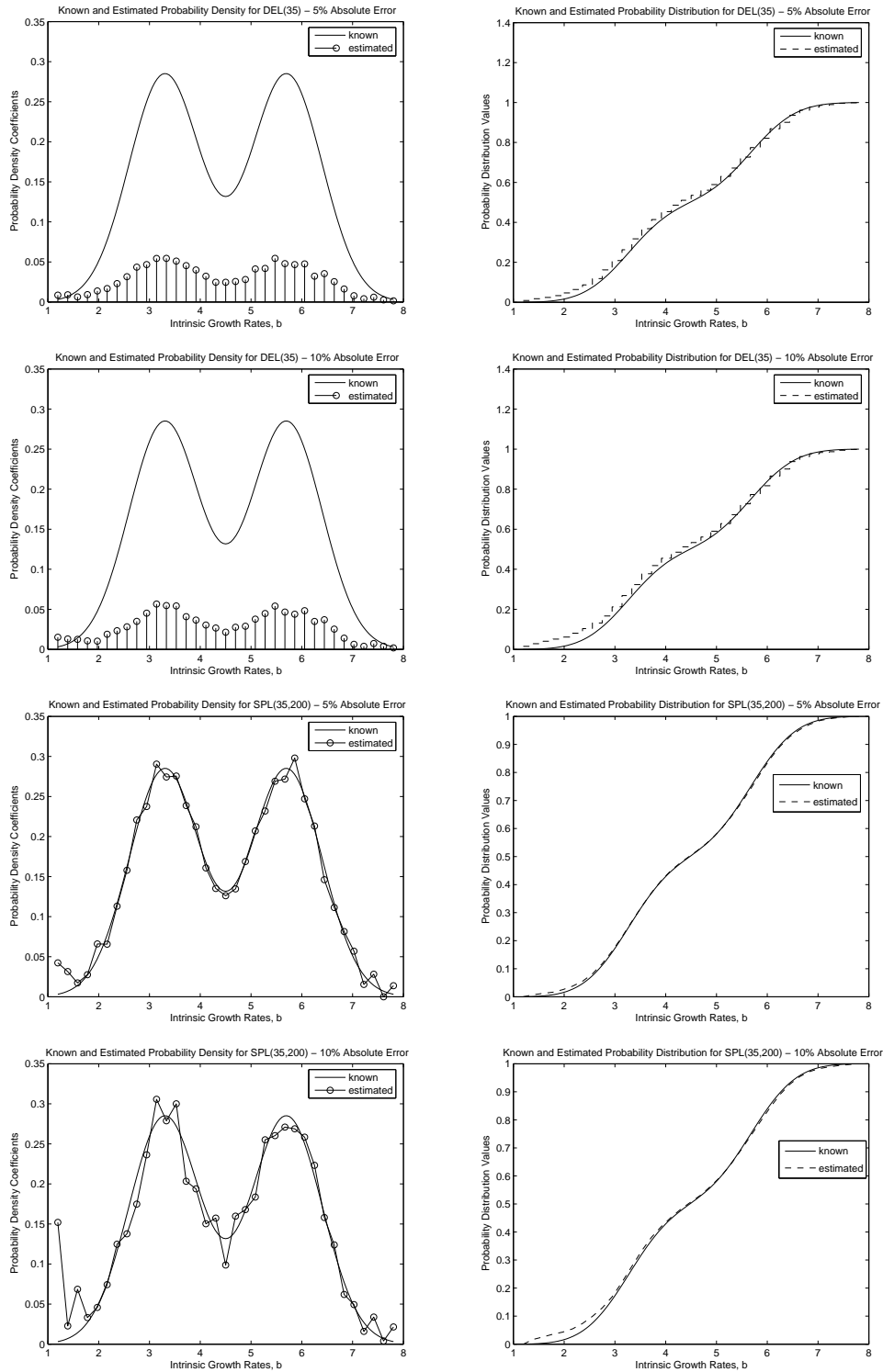


Figure 4: Estimates of probability densities and distributions for true bi-gaussian distribution using a)DEL(35) with 5% absolute error, b)DEL(35) with 10% absolute error, c)SPL(35,200) with 5% absolute error, d)SPL(35,200) with 10% absolute error

3.3 Sensitivity Analysis of Approximation Methods with Respect to Noisy Data

We next considered the sensitivity of the two approximation methods, DEL(M) and SPL(M,N), with respect to noisy data in order to make some comments about the uncertainty associated with the estimated growth rate distributions. We cannot physically observe the entire population; however, we can include some measures (e.g., confidence intervals) on the uncertainty of the estimates obtained from the two approximation methods when using only a sample from the population. We follow the standard statistical framework for asymptotic (as sample size $n \rightarrow \infty$) distributions for OLS estimators [19, 22, 25].

We began by considering the following statistical model for the mosquitofish population density

$$Y(\bar{x}_j) = f(\bar{x}_j, \theta_0) + \epsilon_j, \quad j = 1, \dots, n,$$

where $\{\bar{x}_j\}$ corresponds to $(t_l, x_m), l = 1, \dots, n_t, m = 1, \dots, n_x$ pairs (n_t corresponds to the number of time values, n_x corresponds to the number of size values used, and $n = n_t \cdot n_x$). We also note that

$$E[Y_j] = f(\bar{x}_j, \theta_0) \quad \text{and} \quad \text{var}[Y_j] = \sigma_0^2,$$

where $\theta_0 \approx \theta = \{p_k\}_{k=1}^M$ when using DEL(M) and $\theta_0 \approx \theta = \{a_k\}_{k=1}^M$ when using SPL(M,N) and $\epsilon_j \sim \mathcal{N}(0, \sigma_0^2)$. Here, θ_0 represents the “true” parameter value and σ_0^2 represents the “true” variance for the system (which are generally not known) and the θ are the parameters to be estimated for θ_0 . Furthermore, we note that

$$f(\bar{x}_j, \theta_0) \approx \sum_{k=1}^M p_k v(\bar{x}_j; g_k), \quad (6)$$

where $g_k(x) = b_k(\gamma - x)$, when considering DEL(M), while

$$f(\bar{x}_j, \theta_0) \approx \sum_{k=1}^M a_k \int_B v(\bar{x}_j; g) l_k(b) db \quad (7)$$

when considering SPL(M,N).

As stated previously, our goal is to quantify the uncertainty associated with the estimated growth rate distributions from the methods DEL(M) and SPL(M,N). We will make statements about the reliability of our estimates based upon standard errors. That is, we will compute confidence intervals corresponding to the estimated growth rate distributions. We note as $n \rightarrow \infty$,

$$\hat{\theta}_{OLS}(Y) \sim \mathcal{N}_M(\theta_0, \sigma_0^2 [\mathcal{X}^T(\theta_0) \mathcal{X}(\theta_0)]^{-1}) = \mathcal{N}_M(\theta_0, \Sigma_0),$$

where $\mathcal{X}(\theta) = \frac{\partial F}{\partial \theta}(\theta) = F_\theta(\theta)$ is the $n \times M$ sensitivity matrix with elements

$$\mathcal{X}_{jk}(\theta) = \frac{\partial f(\bar{x}_j, \theta)}{\partial \theta_k}.$$

We used the ordinary least squares estimator:

$$\text{minimize in } \theta \quad \sum_{j=1}^n (Y_j - f(\bar{x}_j, \theta))^2.$$

The estimates $\hat{\theta}$ for the growth rate distribution minimize

$$\sum_{j=1}^n (y_j - f(\bar{x}_j, \theta))^2$$

for a particular realization or data set $\{y_j\}$, and result from a quadratic programming problem as discussed earlier. As we have remarked, that along with θ_0 being unknown, σ_0 is also usually unknown. Thus, in order to compute the standard errors associated with $\hat{\theta}$, we must also estimate σ_0 . We use the following estimate for σ_0 :

$$\sigma_0^2 \approx \hat{\sigma}^2 = \frac{1}{n-M} \sum_{j=1}^n \left(y_j - f(\bar{x}_j, \hat{\theta}) \right)^2.$$

We then use these estimates to compute an estimate of the covariance matrix Σ_0 :

$$\Sigma_0 \approx \Sigma = \hat{\sigma}^2 \left[\mathcal{X}^T(\hat{\theta}) \mathcal{X}(\hat{\theta}) \right]^{-1}.$$

Moreover, we note that the standard errors for the growth rate distributions estimates $\hat{\theta}_k$ are given by

$$\text{SE}(\hat{\theta}_k) = \sqrt{\Sigma_{kk}}.$$

Before we present some results, we want to make a few comments about the covariance matrix Σ . We note that determining $\hat{\sigma}^2$ is very straightforward once we have $\hat{\theta}$. We simply multiply the residual at $\hat{\theta}$ by $\frac{1}{n-M}$ in order to compute $\hat{\sigma}^2$. We must also compute $\mathcal{X}(\hat{\theta})$, which can be more difficult in general when dealing with nonlinear systems. However, as we noted earlier, the elements of the $n \times M$ matrix $\mathcal{X}(\theta)$ are given by

$$\mathcal{X}_{jk}(\theta) = \frac{\partial f(\bar{x}_j, \theta)}{\partial \theta_k}.$$

These are actually the sensitivity elements associated with this system. We note for DEL(M) the entries in the sensitivity matrix $\mathcal{X}(\theta)$ are given by

$$\mathcal{X}_{jk}(\theta) = \frac{\partial f(\bar{x}_j, \theta)}{\partial \theta_k} = v(\bar{x}_j, g_k),$$

where θ_k (p_k or a_k) is the probability parameter associated with growth rate $g_k(x) = b_k(\gamma - x)$. For SPL(M,N), the entries in the sensitivity matrix $\mathcal{X}(\theta)$ are given by

$$\mathcal{X}_{jk}(\theta) = \frac{\partial f(\bar{x}_j, \theta)}{\partial \theta_k} = \int_B v(\bar{x}_j, g) l_k(b) db.$$

Since both (6) and (7) are linear in θ , then computing \mathcal{X} in this case is also very straightforward. Furthermore, the asymptotic distributional results given earlier are exact in this case (see [19, 22, 25]) as opposed to only being an approximation when $f(\bar{x}_j, \theta)$ is nonlinear in θ .

We present next some findings on the sensitivity of DEL(M) and SPL(M,N) using simulated data generated with a “true” Bi-Gaussian distribution. Table 3 displays the estimated probability density values and corresponding confidence intervals for DEL(9) and SPL(9,200) in the presence of 5% and 10% absolute error. In Figure 5, we see the confidence intervals corresponding to the estimated growth rate distributions with $\alpha = 0.05$ for DEL(9) and SPL(9,200) with simulated data with both 5% and 10% absolute error. The endpoints of the confidence intervals are given by

$$\hat{\theta} \pm t_{1-\alpha/2} \text{SE}(\hat{\theta}),$$

where $t_{1-\alpha/2}$ is a distribution value that is determined by the level of significance that is chosen [24]. After a level of significance is chosen, we determine the corresponding $t_{1-\alpha/2}$ value from a statistical table for the t-distribution. We chose to use $\alpha = 0.05$ for a significance level of 95%, which corresponds to $t_{1-\alpha/2} = 1.96$ when the number of degrees of freedom is large, i.e., $n \geq 30$. Based on the confidence intervals, we can make statements about the estimation procedure constructed from a realization of Y . If the resulting confidence intervals are relatively large in relation to $\hat{\theta}$, then we are not very confident

	DEL(9) - 5%	DEL(9) - 10%	SPL(9,200) - 5%	SPL(9,200) - 10%
p_1	0.1724 ± 0.0191	0.1750 ± 0.0196	0.0203 ± 0.0130	0.0460 ± 0.0252
p_2	0.1506 ± 0.0170	0.1487 ± 0.0174	0.0459 ± 0.0075	0.0452 ± 0.0145
p_3	0.1665 ± 0.0149	0.1644 ± 0.0152	0.2447 ± 0.0069	0.2404 ± 0.0133
p_4	0.1432 ± 0.0128	0.1386 ± 0.0131	0.2651 ± 0.0063	0.2506 ± 0.0121
p_5	0.0948 ± 0.0115	0.1005 ± 0.0117	0.0895 ± 0.0057	0.1121 ± 0.0110
p_6	0.1084 ± 0.0098	0.1058 ± 0.0100	0.2738 ± 0.0051	0.2593 ± 0.0098
p_7	0.0953 ± 0.0090	0.0946 ± 0.0092	0.2476 ± 0.0045	0.2509 ± 0.0088
p_8	0.0442 ± 0.0080	0.0482 ± 0.0082	0.0338 ± 0.0038	0.0397 ± 0.0075
p_9	0.0246 ± 0.0072	0.0242 ± 0.0073	0.0000 ± 0.0056	0.0038 ± 0.0109

Table 3: Estimated probability density values and confidence intervals for true bi-gaussian distribution for DEL(9) and SPL(9,200) with 5% and 10% absolute error

about the estimation procedure used to estimate θ_0 . In the results presented here, we can state that we are 95% confident that intervals constructed using the estimation procedures with DEL(M) and SPL(M,N) would “cover” θ_0 . We note that for the fixed value $M = 9$ the confidence intervals corresponding to $\alpha = 0.05$ when using DEL(M) are relatively small in relation to the estimated $\hat{\theta}$ for data with both

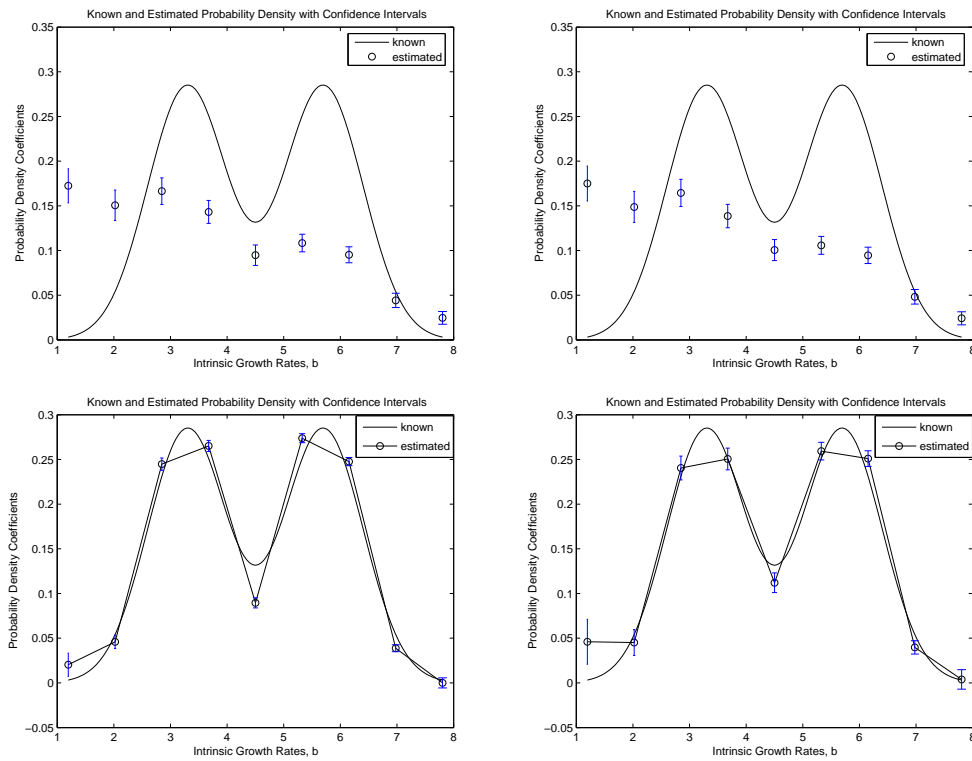


Figure 5: Estimates of probability densities with confidence intervals given a true bi-gaussian distribution using a)DEL(9) with 5% absolute error, b)DEL(9) with 10% absolute error, c)SPL(9,200) with 5% absolute error, d)SPL(9,200) with 10% absolute error

5% and 10% absolute error. Moreover, we note from Figure 5 that the resulting confidence intervals for DEL(9) with 5% and 10% absolute error are approximately the same. Thus, for $M = 9$, the delta function approximation method appears to be insensitive to noisy data. In comparison, we note for this fixed value $M = 9$ when using SPL(M,N) the resulting confidence intervals are relatively larger for data with 10% absolute error in comparison to those for data with 5% absolute error. Thus, the confidence associated with the estimator procedure based on SPL(M,N) appears to decrease as the percentage of absolute error increases. However, the confidence intervals are still relatively small in relation to the estimates $\hat{\theta}$. Based on these results, we would infer that for this fixed value of M, the spline based approximation method appears to be very slightly sensitive to very noisy data.

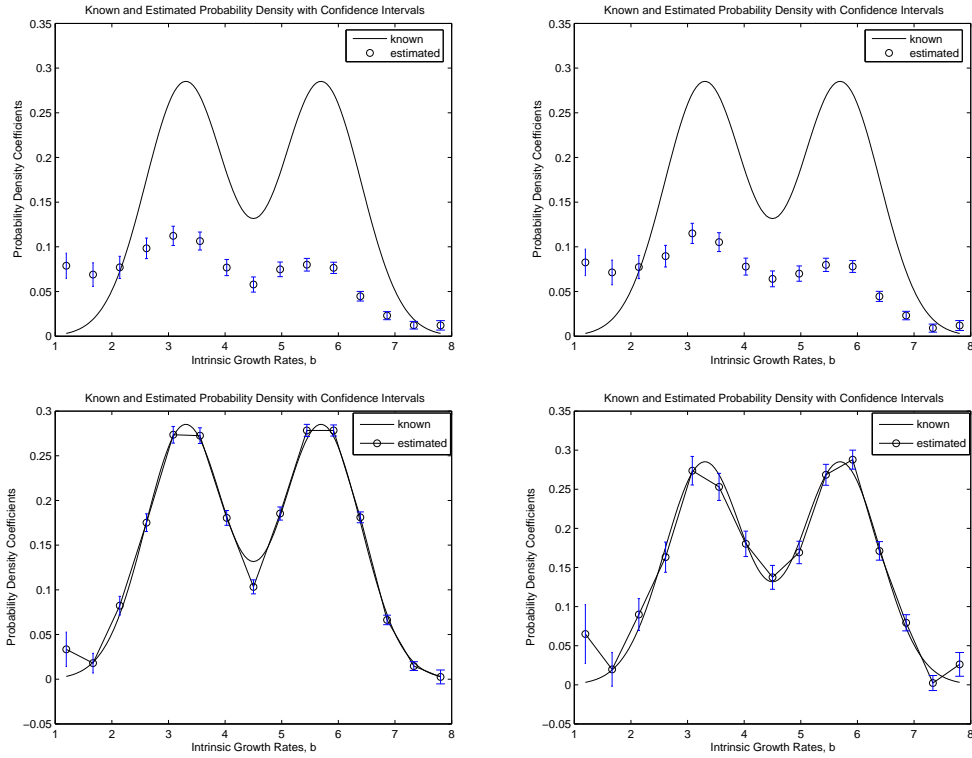


Figure 6: Estimates of probability densities with confidence intervals given a true bi-gaussian distribution using a)DEL(15) with 5% absolute error, b)DEL(15) with 10% absolute error, c)SPL(15,200) with 5% absolute error, d)SPL(15,200) with 10% absolute error

We note that the estimated probability density values and corresponding confidence intervals for DEL(15) and SPL(15,200) in the presence of 5% and 10% absolute error are given in Table 4. In Figure 6, we see the confidence intervals corresponding to the estimated growth rate distributions with $\alpha = 0.05$ for DEL(15) and SPL(15,200) with simulated data with both 5% and 10% absolute error. The endpoints of the confidence intervals are constructed in the same way as discussed above. We can again state that we are 95% confident that intervals constructed using the estimation procedures with DEL(M) and SPL(M,N) would “cover” θ_0 . From Figure 6 we see that the confidence intervals corresponding to $\alpha = 0.05$ when using DEL(M) for $M = 15$ are relatively small in relation to $\hat{\theta}$ for data with 5% and 10% absolute error much like those computed for $M = 9$. We also see in this case that the confidence intervals are approximately the same for both sets of data. We arrive at the same conclusion for $M = 15$ as we did for $M = 9$; that is, the delta function approximation method appears to be insensitive to noisy data. We now look at the results for $M = 15$ when using SPL(M,N) with data with 5% and 10%

	DEL(15) - 5%	DEL(15) - 10%	SPL(15,200) - 5%	SPL(15,200) - 10%
p_1	0.0788 ± 0.0141	0.0826 ± 0.0148	0.0335 ± 0.0192	0.0650 ± 0.0376
p_2	0.0689 ± 0.0132	0.0714 ± 0.0138	0.0180 ± 0.0110	0.0197 ± 0.0241
p_3	0.0771 ± 0.0123	0.0774 ± 0.0129	0.0823 ± 0.0104	0.0899 ± 0.0204
p_4	0.0983 ± 0.0115	0.0896 ± 0.0121	0.1752 ± 0.0099	0.1632 ± 0.0193
p_5	0.1123 ± 0.0107	0.1151 ± 0.0113	0.2735 ± 0.0093	0.2738 ± 0.0183
p_6	0.1064 ± 0.0101	0.1053 ± 0.0105	0.2725 ± 0.0088	0.2531 ± 0.0173
p_7	0.0768 ± 0.0091	0.0779 ± 0.0095	0.1805 ± 0.0083	0.1803 ± 0.0162
p_8	0.0579 ± 0.0085	0.0642 ± 0.0089	0.1033 ± 0.0078	0.1374 ± 0.0153
p_9	0.0748 ± 0.0082	0.0700 ± 0.0086	0.1854 ± 0.0073	0.1692 ± 0.0144
p_{10}	0.0800 ± 0.0071	0.0798 ± 0.0074	0.2783 ± 0.0068	0.2685 ± 0.0134
p_{11}	0.0765 ± 0.0063	0.0780 ± 0.0066	0.2783 ± 0.0062	0.2879 ± 0.0121
p_{12}	0.0447 ± 0.0054	0.0446 ± 0.0056	0.1811 ± 0.0061	0.1712 ± 0.0119
p_{13}	0.0230 ± 0.0045	0.0231 ± 0.0047	0.0663 ± 0.0053	0.0793 ± 0.0103
p_{14}	0.0123 ± 0.0043	0.0091 ± 0.0045	0.0147 ± 0.0049	0.0022 ± 0.0096
p_{15}	0.0121 ± 0.0053	0.0120 ± 0.0055	0.0025 ± 0.0078	0.0261 ± 0.0152

Table 4: Estimated probability density values and confidence intervals for true bi-gaussian distribution for DEL(15) and SPL(15,200) with 5% and 10% absolute error

absolute error. We also see for $M = 15$ as we saw for $M = 9$ that the resulting confidence intervals are larger when the data is noisier. As discussed in the case for $M = 9$, the spline based approximation method also appears to be slightly sensitive to noisy data. To summarize, we note based on the standard error analysis discussed in this section and computational results (those presented here as well as those obtained for $M = 5$) we can conclude that DEL(M) appears to be insensitive to noisy data. Moreover, we can state that we are confident about the estimated growth rate distributions obtained using this method. We also conclude that SPL(M,N) appears to be relatively insensitive to noisy data. Furthermore, we would feel certain about the estimated growth rate distributions obtained using SPL(M,N) with data with small amounts of noise; however, we would infer that larger amounts of noise in the data would lead to larger confidence intervals and less certainty in the associated estimated growth rate distributions obtained using SPL(M,N).

4 Computational Results for Inverse Problems with Field Data

In this section, we will present and discuss the results of the inverse problem for the estimation of growth rate distributions using the delta function approximation method, the spline based approximation method, and a parameterized ordinary least squares method. We use field data collected from rice paddies in the place of the simulated data. Since the actual growth rate distribution of the mosquitofish observed in the experiment is unknown, we must compare the field data to the estimated population data produced by the estimated growth rate distribution from each of these methods in order to compare the efficacy of these methods.

In these numerical simulations, we assume that the growth rate of the mosquitofish is now parameterized by both the intrinsic growth rate b and the maximum size γ , where $g(x) = b(\gamma - x)$. The collection of growth rates for the DEL method is now given by $G = \{g_{jk}\}$, for $j = 1, \dots, M_1$, and $k = 1, \dots, M_2$, where $g_{jk} = b_j(\gamma_k - x)$ with $\{b_j\}$ and $\{\gamma_k\}$ the discretizations of B and Γ , respectively. As earlier stated, in the GRD model mosquitofish are grouped together in the same subpopulation if they have the same growth rate g_{jk} . We assume here, as we also assumed in our earlier computations,

that $\mu = 0$ and $K = 0$, so that we can focus on the growth rate distribution only (mortality and fecundity were not thought to be important features of the experimental data of [4, 10]).

The field data that we are using in the inverse problem was collected in an experiment described in [10]. On June 28, 1992, four rice paddies were stocked with mosquitofish. In order to measure emigration, an outflow trap was placed on each paddy. Fifteen traps were used per paddy, and weekly measurements were taken. The length of the mosquitofish range from 0 to 40 mm, with the mosquitofish being grouped into size classes of 2 mm for a total of 20 size classes. The data for Day 195, Day 202, Day 209, and Day 216 are used in these simulations (see Figure 7). We define the size distribution frequency for size class i as $f_i = n_{m,i}/N_m$, where $n_{m,i}$ is the number of mosquitofish measured in size class i and N_m is the total number of mosquitofish measured. The total population of mosquitofish is divided into 512 subpopulations. We note that the discretizations for the intrinsic growth rates b and the maximum sizes γ are defined as

$$\begin{aligned} b_j &= 0.2 + \frac{1}{31} \cdot 4.8 \cdot (j - 1), \quad j = 1, 2, \dots, 32 \\ \gamma_1 &= \frac{16}{38}, \quad \gamma_2 = \frac{22}{38}, \quad \gamma_3 = \frac{24}{38} \\ \gamma_k &= \frac{16}{38} + \frac{1}{15} \cdot \frac{22}{38} \cdot (k - 1), \quad k = 4, \dots, 16. \end{aligned}$$

The Day 195 data is interpolated and used an approximation for the initial size density $\Phi(x)$. Since this data set is used as an approximation for $\Phi(x)$, it cannot be used in the estimation of the growth rate distributions. Therefore, we are left with only three data sets to use in the inverse problem.

We introduced the delta function approximation method, DEL(M), earlier when the growth rate is parameterized by b only. Since we are now considering a growth rate parameterized by b and γ , the approximated population density for $u(t, x; P)$ in (2) is now given by

$$u(t, x; \{p_{jk}\}) = \sum_{j,k} v(t, x; g_{jk}) p_{jk},$$

where $v(t, x; g_{jk})$ is the subpopulation density from (1) with growth rate g_{jk} and p_{jk} is the corresponding probability that an individual has growth rate g_{jk} . We will now use the notation DEL(M_1, M_2) to denote the delta function approximation method, where M_1 is the number of intrinsic rates b_j and M_2 is the number of maximum sizes γ_k used in the approximation.

The spline based approximation method, SPL(M,N), was also introduced earlier for the one parameter family of growth rates. For the two parameter family of growth rates that we are now using, the approximated population density for $u(t, x; P)$ in (2) is given by

$$u(t, x; \{a_{jk}\}) = \sum_{j,k} a_{jk} \int_B \int_\Gamma v(t, x; g) l_j(b) l_k(\gamma) d\gamma db,$$

where $g = g(x; b, \gamma) = b(\gamma - x)$ and $p_{jk}(b, \gamma) = a_{jk} l_j(b) l_k(\gamma)$ is the probability density for individuals in population subgroup jk with l_j, l_k piecewise linear spline functions. The notation that we employ here is SPL(M_1, N_1, M_2, N_2), where M_1 and M_2 are the number of basis elements used to approximate the growth rate probability distribution with respect to b and γ , and N_1 and N_2 represent the number of quadrature nodes used in the composite trapezoidal rule [16] for double integrals to approximate the integral in the expression above with respect to b and γ .

We next present the results of the approximation methods DEL(M_1, M_2) and SPL(M_1, N_1, M_2, N_2), which we used in the inverse problem for the estimation of the growth rate distribution for the field data. The addition of a second parameter in the growth rate g of the mosquitofish does not change the fact that

the least squares inverse problem that we want to solve

$$\begin{aligned} \min_{P \in \mathcal{P}^{M_1 \times M_2}(G)} J(P) &= \sum_{i,j} |u(t_i, x_j; P) - \hat{u}(t_i, x_j)|^2 \\ &= \sum_{i,j} (u(t_i, x_j; P))^2 - 2u(t_i, x_j; P)\hat{u}(t_i, x_j) + (\hat{u}(t_i, x_j))^2, \end{aligned}$$

where $\hat{u}(t, x)$ is the data and $\mathcal{P}^{M_1 \times M_2}(G)$ is the finite dimensional approximation to $\mathcal{P}(G)$, simplifies to a quadratic programming problem for an appropriately defined $F(\mathbf{p}) \equiv \mathbf{p}^T \mathbf{A} \mathbf{p} + 2\mathbf{p}^T \mathbf{b} + c$.

In Figure 7, we have the results from the inverse problem using DEL(32,16). These results were obtained in 514.3600 seconds, and the corresponding residual $J = 8.3169 \times 10^{-4}$. We see from the results shown in Figure 7 that the estimated population is a good fit to the field data. The two key features of the data, dispersion and bifurcation, are both exhibited in the estimated population. The corresponding estimated probability density and distribution are shown in Figure 8. While no useful information can be obtained from Figure 8a, the estimated probability distribution in Figure 8b appears to be Bi-Gaussian in both b and γ , which is expected based on results from [9] and [10].

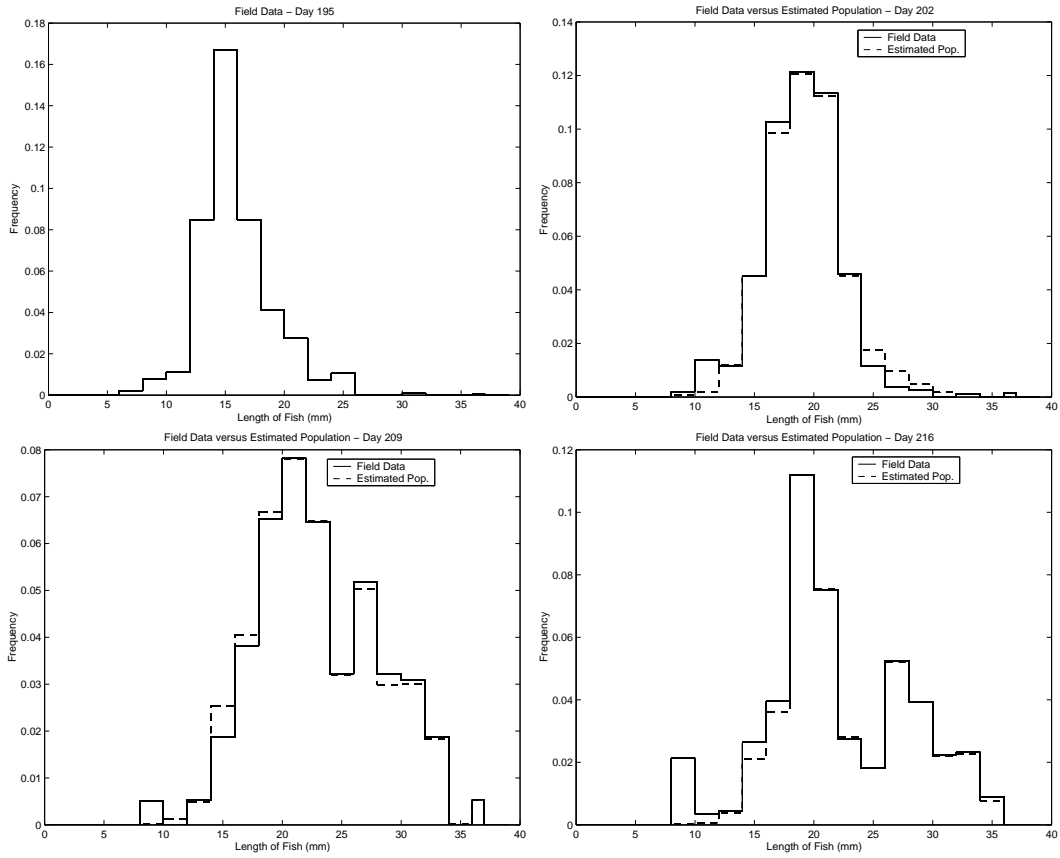


Figure 7: Field data versus estimated population for DEL(32,16)

The best results obtained using $\text{SPL}(M_1, N_1, M_2, N_2)$ for the estimation of the growth rate distribution of the field data are shown in Figure 9 for $\text{SPL}(5,35,9,35)$. We note that the results obtained using $\text{SPL}(5,35,5,35)$ and $\text{SPL}(5,35,7,35)$ were approximately the same as those obtained using $\text{SPL}(5,35,9,35)$ for Day 202 and Day 209. However, the estimated population data obtained using $\text{SPL}(5,35,9,35)$ gave a much better fit to the data for Day 216 than the results obtained using

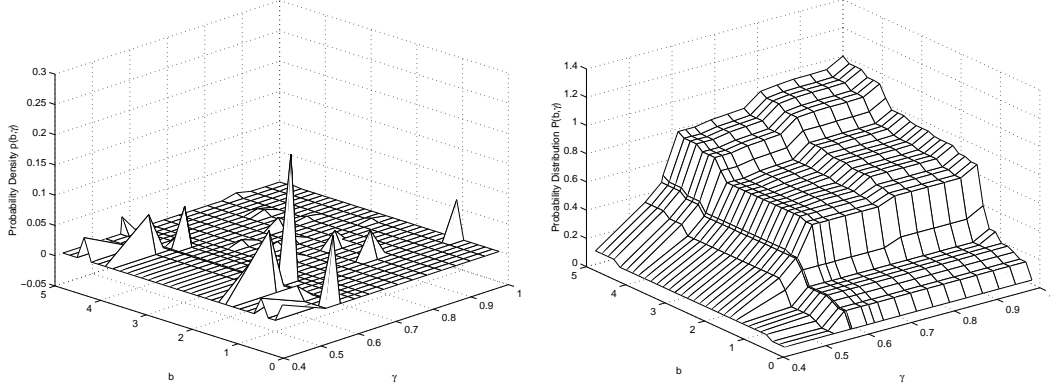


Figure 8: a) Estimated probability density and b) estimated probability distribution for DEL(32,16)

SPL(5,35,5,35) and SPL(5,35,7,35). The corresponding residual, J , for SPL(5,35,9,35) is 0.0054, and these results were obtained in 1.8822×10^3 seconds, or approximately 32 minutes. In comparison to the estimated population data produced by the estimated growth rate distribution from DEL(32,16), the estimated population data produced by the estimated growth rate distribution from SPL(5,35,9,35) does not give as good a fit to the field data, as is seen in Figure 9. The estimated probability density and distribution are shown in Figure 10. The resulting estimated probability distribution appears to be Bi-Gaussian in γ . In contrast to the results obtained in the previous examples with simulated data, the delta function approximation method does a better job of fitting the given field data in a more efficient way in comparison to the spline based approximation method.

Since the results from SPL(M_1, N_1, M_2, N_2) were not as good as those obtained from DEL(M_1, M_2), we tried one more method in the inverse problem with the field data. Based on previous work in [9] and [10] and our own numerical simulations with simulated data, we know that a Bi-Gaussian growth rate distribution results in population density data with the two key features of dispersion and bifurcation. The field data that we are using in these computations exhibit these features as well, so we suspect that the underlying growth rate distribution is Bi-Gaussian. With that in mind, instead of approximating the density with delta functions or piecewise linear splines, we chose to use a parametric Bi-Gaussian probability density function in the growth rate distribution (GRD) model. The estimated $\{p_{jk}\}$ and $\{a_{jk}\}$ found via the DEL(M_1, M_2) and SPL(M_1, N_1, M_2, N_2) methods, respectively, were not readily interpreted in terms of the actual mosquitofish growth rate and maximum size means and variances. However, by using an a priori Bi-Gaussian probability density function in the GRD model, we have in essence taken a standard parametric approach to the statistical inverse problem. The Bi-Gaussian probability density function p we choose is given by

$$p(b, \gamma; \bar{b}_1, \sigma_{b_1}^2, \bar{b}_2, \sigma_{b_2}^2, \bar{\gamma}_1, \sigma_{\gamma_1}^2, \bar{\gamma}_2, \sigma_{\gamma_2}^2) = \left(\frac{\exp\left\{-\frac{(b-\bar{b}_1)^2}{2\sigma_{b_1}^2}\right\}}{2\sqrt{2\pi\sigma_{b_1}^2}} + \frac{\exp\left\{-\frac{(b-\bar{b}_2)^2}{2\sigma_{b_2}^2}\right\}}{2\sqrt{2\pi\sigma_{b_2}^2}} \right) \times \left(\frac{\exp\left\{-\frac{(\gamma-\bar{\gamma}_1)^2}{2\sigma_{\gamma_1}^2}\right\}}{2\sqrt{2\pi\sigma_{\gamma_1}^2}} + \frac{\exp\left\{-\frac{(\gamma-\bar{\gamma}_2)^2}{2\sigma_{\gamma_2}^2}\right\}}{2\sqrt{2\pi\sigma_{\gamma_2}^2}} \right)$$

where we have assumed b and γ to be independent Bi-Gaussian random variables. The parameters (\bar{b}_1, \bar{b}_2) and $(\sigma_{b_1}^2, \sigma_{b_2}^2)$ represent the means and variances, respectively, of a Bi-Gaussian distribution on the intrinsic rates b , while the parameters $(\bar{\gamma}_1, \bar{\gamma}_2)$ and $(\sigma_{\gamma_1}^2, \sigma_{\gamma_2}^2)$ represent the means and variances of a

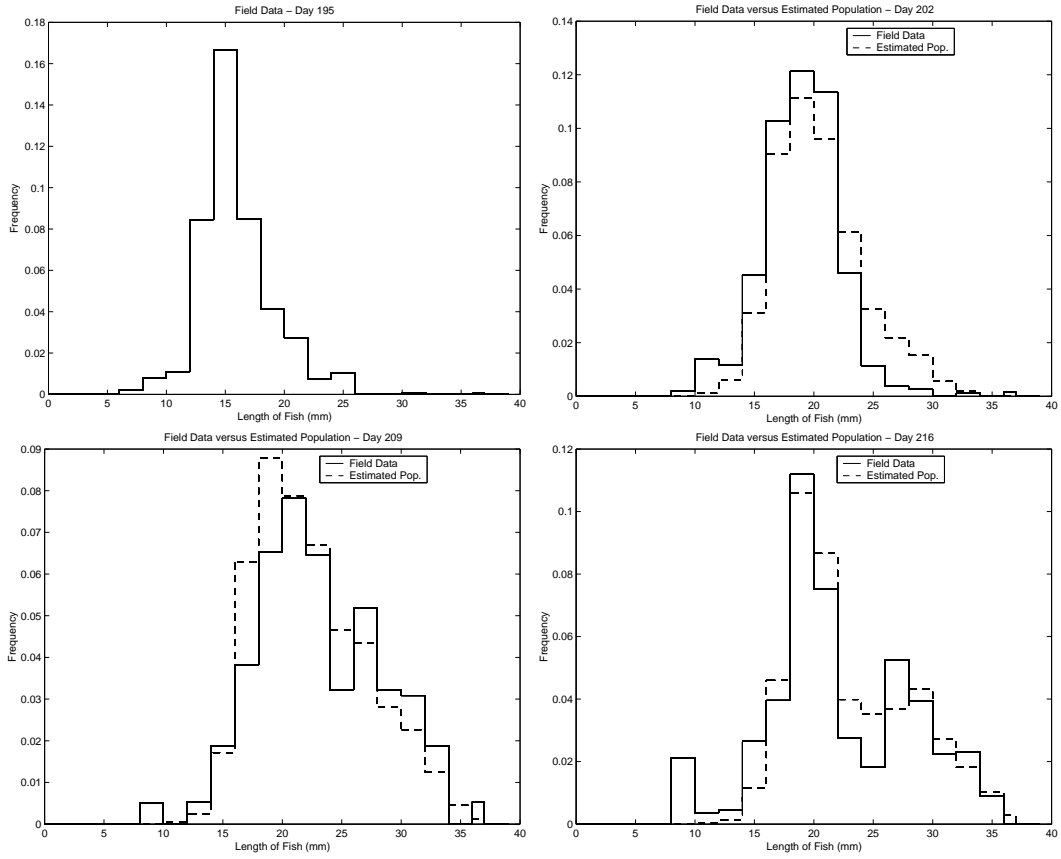


Figure 9: Field data versus estimated population for SPL(5,35,9,35)

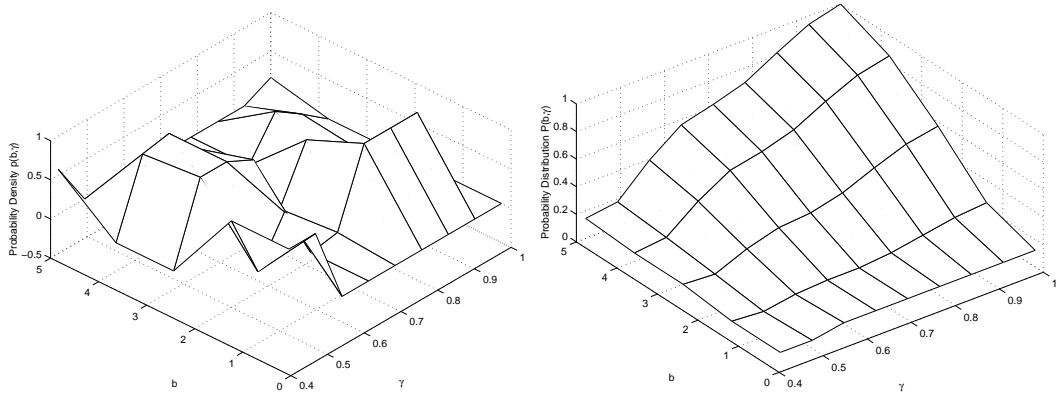


Figure 10: a) Estimated probability density and b) estimated probability distribution for SPL(5,35,9,35)

Bi-Gaussian distribution on the maximum sizes γ . We will define $\mathbf{q} = (\bar{b}_1, \sigma_{b_1}^2, \bar{b}_2, \sigma_{b_2}^2, \bar{\gamma}_1, \sigma_{\gamma_1}^2, \bar{\gamma}_2, \sigma_{\gamma_2}^2)$. Our third approach will not use an approximation to the GRD model, but instead use the Bi-Gaussian probability density function, given above, in the GRD model

$$u(t, x; \mathbf{q}) = \int_B \int_{\Gamma} v(t, x; g(b, \gamma)) p(b, \gamma; \mathbf{q}) d\gamma db,$$

where again $g(x; b, \gamma) = b(\gamma - x)$. We will denote this third approach as $\text{PAR}(M, N_1, N_2)$, where M is the number of parameters in \mathbf{q} and N_1 and N_2 represent the number of quadratures used in the composite trapezoidal rule [16] to approximate the double integral with respect to b and γ , respectively. We will estimate \mathbf{q} by solving the ordinary least squares problem

$$\min_{\mathbf{q} \in \mathbb{R}^M} J(\mathbf{q}) = \sum_{i,j} |u(t_i, x_j; \mathbf{q}) - \hat{u}(t_i, x_j)|^2,$$

where $\hat{u}(t, x)$ is the data. Once this least squares problem has been solved, we can use the optimal \mathbf{q} in the Bi-Gaussian probability density function to determine the population density $u(t, x; \mathbf{q})$.

The optimal results for the inverse problem using $\text{PAR}(8,35,35)$ are shown in Figure 11. The optimal parameters $\mathbf{q} = (1.9749, 0.0388 \times 10^{-3}, 3.9132, 0.2228 \times 10^{-3}, 0.5265, 0.7208, 0.0122, 0.0372)$, with a residual $J = 0.0074$, were determined in 20.5490 seconds. In comparison to the results obtained using $\text{DEL}(32,16)$ and $\text{SPL}(5,35,9,35)$, the estimated population density does not fit the field data as well as the estimated population density produced by $\text{DEL}(32,16)$ but is comparable to those obtained using $\text{SPL}(5,35,9,35)$. We note that the spline based approximation method does a better job of estimating the frequencies for the smaller size classes than the parameterized OLS technique. While the results from the spline based approximation method and the parameterized OLS method are very similar, the computational time required by $\text{PAR}(8,35,35)$ is much lower (compare the 32 minutes for $\text{SPL}(5,35,9,35)$ versus 21 seconds for $\text{PAR}(8,35,35)$) than the computational time required by $\text{SPL}(5,35,9,35)$. The estimated probability density and probability distribution generated by the optimal \mathbf{q} are shown in Figure 12. We clearly see for a fixed value of γ the probability distribution of b is Bi-Gaussian.

As we did previously, we would like to also present here some results on the uncertainty associated with the estimated growth distributions determined by the inverse problem. The treatment in this section with the field data is very similar to the treatment previously carried out with the simulated data. With respect to the optimal results given in this section, we will only be able to perform a statistical analysis for $\text{SPL}(5,35,9,35)$ and $\text{PAR}(8,35,35)$. We are unable to perform this analysis for $\text{DEL}(32,16)$ because the field data consists of 60 data points and the number of parameters determined by the $\text{DEL}(32,16)$ is 512; thus, the analysis in this case is invalid. Since we have explained in detail the underlying statistical model that we are considering, we will omit these details here and define functions and variables that are used with respect to $\text{SPL}(M_1, N_1, M_2, N_2)$ and $\text{PAR}(M, N_1, N_2)$. First, we note that $\{\bar{x}_i\}_{i=1}^n$ corresponds to $(t_l, x_m), l = 1, \dots, 3, m = 1, \dots, 20$ pairs since the field data that we use in the inverse problem consists of three days and twenty size classes (hence $n = 60$), $\theta = \{a_{jk}\}_{j,k=1}^{M_1 \times M_2}$ when using $\text{SPL}(M_1, N_1, M_2, N_2)$, and $\theta = \mathbf{q}$ when using $\text{PAR}(M, N_1, N_2)$. We also note that

$$f(\bar{x}_i, \theta) = \sum_{j=1}^{M_1} \sum_{k=1}^{M_2} a_{jk} \int_B \int_{\Gamma} v(\bar{x}_i; g) l_j(b) l_k(\gamma) d\gamma db,$$

where $\theta = \{a_{jk}\}$ when considering $\text{SPL}(M_1, N_1, M_2, N_2)$. When considering $\text{PAR}(M, N_1, N_2)$,

$$f(\bar{x}_i, \theta) = \int_B \int_{\Gamma} v(\bar{x}_i; g) p(b, \gamma; \mathbf{q}) d\gamma db.$$

We will define M from our previous analysis to be $M_1 \cdot M_2$ for $\text{SPL}(M_1, N_1, M_2, N_2)$ and M (the number of parameters in \mathbf{q}) for $\text{PAR}(M, N_1, N_2)$. For $\text{SPL}(M_1, N_1, M_2, N_2)$, the entries in the sensitivity matrix $\mathcal{X}(\theta)$ are given by

$$\mathcal{X}_{im}(\theta) = \frac{\partial f(\bar{x}_i, \theta)}{\partial \theta_m} = \int_B \int_{\Gamma} v(\bar{x}_i; g) l_j(b) l_k(\gamma) d\gamma db.$$

For the parameterized OLS method, the entries in the sensitivity matrix $\mathcal{X}(\theta)$ are given by

$$\mathcal{X}_{im}(\theta) = \frac{\partial f(\bar{x}_i, \theta)}{\partial \theta_m} = \int_B \int_{\Gamma} v(\bar{x}_i; g) \frac{\partial p(b, \gamma; \theta)}{\partial \theta_m} d\gamma db,$$

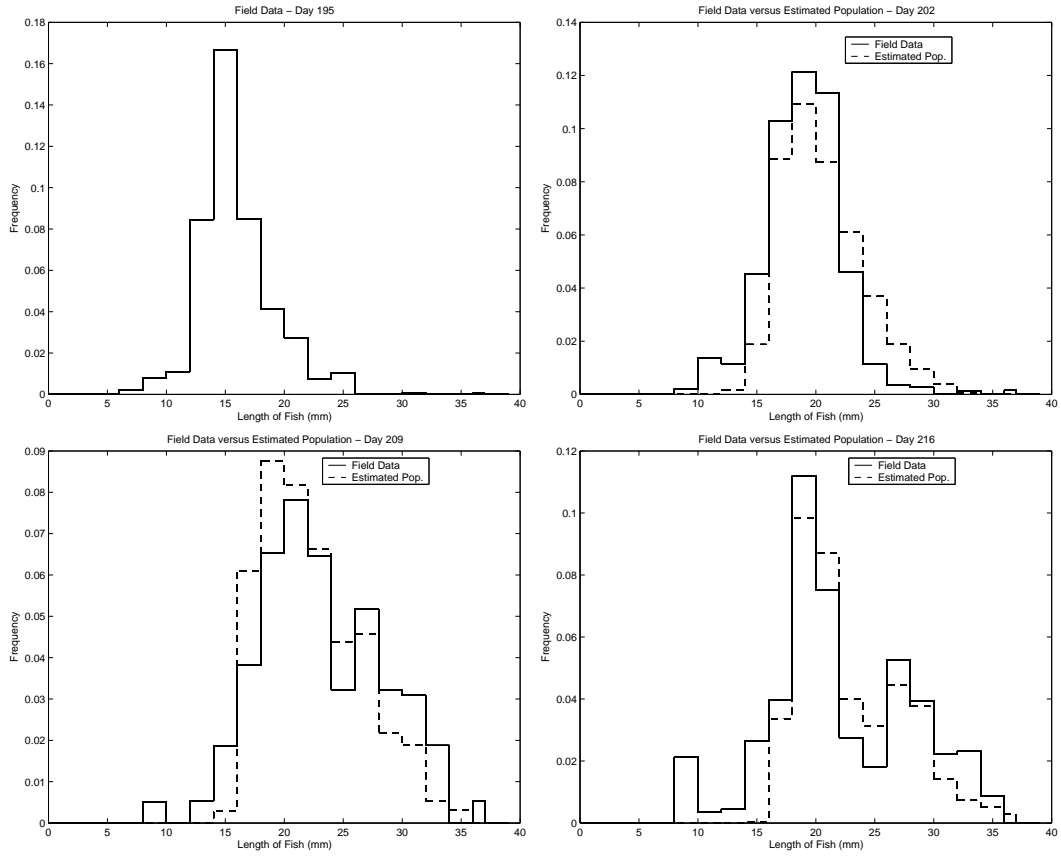


Figure 11: Field data versus estimated population for PAR(8,35,35)

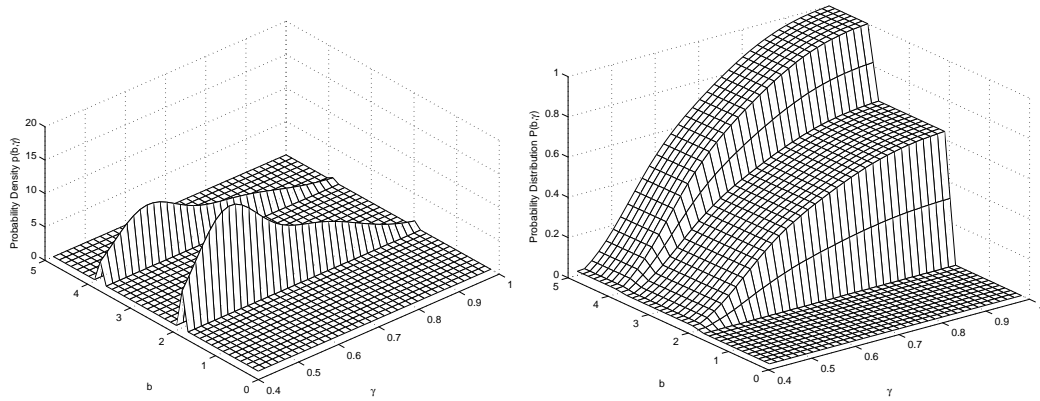


Figure 12: a) Estimated probability density and b) estimated probability distribution for PAR(8,35,35)

and we are able to analytically compute $\frac{\partial p(b,\gamma;\theta)}{\partial \theta_m}$, $m = 1, \dots, M$. Using these facts, we are able to estimate the covariance matrix Σ , which we use to determine the standard errors and confidence intervals for the $\hat{\theta}$. As we stated before, the endpoints of the confidence intervals are given by

$$\hat{\theta} \pm t_{1-\alpha/2} \text{SE}(\hat{\theta}),$$

where $t_{1-\alpha/2}$ is a distribution value given in the t-distribution statistical table (determined by the level of significance chosen [24]). As before, we also chose to use a significance level of 95% corresponding to $\alpha = 0.05$. From the statistical table for the t-distribution, $\alpha = 0.05$ corresponds to 1.96 when the number of degrees of freedom is large, i.e. $n \geq 30$ which is true in this case.

For SPL(5,35,9,35), the optimal value for a_{jk} and the corresponding confidence interval are given in Table 5. As clearly seen by the values given in this table, the confidence intervals are very large in comparison to the value of $\hat{\theta}$. Based on these confidence intervals, we are not very confident about estimated growth rate distributions obtained using SPL(5,35,9,35) in the inverse problem with this field data. Moreover, we found that as the quantity $M_1 \times M_2$ became larger (while still remaining smaller than n), $\mathcal{X}^T(\hat{\theta})\mathcal{X}(\hat{\theta})$ became nearly singular, resulting in a very ill-conditioned covariance matrix Σ . This, in turn, resulted in larger confidence intervals, which implied that the estimated growth rate distributions produced by the spline based approximation method were even more unreliable with respect to the field data. Furthermore, we were unable to compute confidence intervals for the (M_1, M_2) pairs (5,11) and (11,5) because the covariance matrix Σ did not exist in these cases ($\mathcal{X}^T(\hat{\theta})\mathcal{X}(\hat{\theta})$ was singular to machine accuracy).

	$\hat{\theta} \pm t_{1-\alpha/2}\text{SE}(\hat{\theta})$		$\hat{\theta} \pm t_{1-\alpha/2}\text{SE}(\hat{\theta})$		$\hat{\theta} \pm t_{1-\alpha/2}\text{SE}(\hat{\theta})$
α_{11}	$1.0000 \pm 0.4245 \times 10^5$	α_{12}	$1.0000 \pm 0.6404 \times 10^5$	α_{13}	$0.0000 \pm 0.7277 \times 10^5$
α_{14}	$0.0000 \pm 1.0735 \times 10^5$	α_{15}	$0.0000 \pm 3.1710 \times 10^5$	α_{16}	$0.0000 \pm 7.1776 \times 10^5$
α_{17}	$0.0000 \pm 8.7419 \times 10^5$	α_{18}	$0.0000 \pm 3.0799 \times 10^5$	α_{19}	$0.0000 \pm 0.2847 \times 10^5$
α_{21}	$1.0000 \pm 0.2878 \times 10^5$	α_{22}	$0.1525 \pm 0.3884 \times 10^5$	α_{23}	$0.7181 \pm 0.0999 \times 10^5$
α_{24}	$0.0000 \pm 0.4497 \times 10^5$	α_{25}	$0.0620 \pm 0.1099 \times 10^5$	α_{26}	$1.0000 \pm 0.5300 \times 10^5$
α_{27}	$1.0000 \pm 0.1159 \times 10^5$	α_{28}	$1.0000 \pm 0.1826 \times 10^5$	α_{29}	$0.0000 \pm 0.0721 \times 10^5$
α_{31}	$0.0000 \pm 0.1106 \times 10^5$	α_{32}	$1.0000 \pm 0.5037 \times 10^5$	α_{33}	$1.0000 \pm 0.4524 \times 10^5$
α_{34}	$0.8275 \pm 0.2581 \times 10^5$	α_{35}	$0.0000 \pm 0.0555 \times 10^5$	α_{36}	$0.6864 \pm 0.1384 \times 10^5$
α_{37}	$0.0000 \pm 0.1537 \times 10^5$	α_{38}	$0.1591 \pm 0.0115 \times 10^5$	α_{39}	$0.0000 \pm 0.0340 \times 10^5$
α_{41}	$0.0000 \pm 0.3470 \times 10^5$	α_{42}	$0.9352 \pm 0.2464 \times 10^5$	α_{43}	$1.0000 \pm 0.3213 \times 10^5$
α_{44}	$0.6130 \pm 0.2102 \times 10^5$	α_{45}	$0.0000 \pm 0.0946 \times 10^5$	α_{46}	$0.0000 \pm 0.0683 \times 10^5$
α_{47}	$0.4924 \pm 0.0181 \times 10^5$	α_{48}	$0.2598 \pm 0.0278 \times 10^5$	α_{49}	$0.0000 \pm 0.0011 \times 10^5$
α_{51}	$0.5786 \pm 0.7387 \times 10^5$	α_{52}	$0.0000 \pm 0.1237 \times 10^5$	α_{53}	$0.0000 \pm 0.1336 \times 10^5$
α_{54}	$0.0000 \pm 0.2196 \times 10^5$	α_{55}	$0.0000 \pm 0.1207 \times 10^5$	α_{56}	$0.0000 \pm 0.1020 \times 10^5$
α_{57}	$0.0000 \pm 0.0372 \times 10^5$	α_{58}	$0.0000 \pm 0.0172 \times 10^5$	α_{59}	$0.1682 \pm 0.0005 \times 10^5$

Table 5: Estimated values a_{jk} and confidence intervals for SPL(5,35,9,35) with field data

In Table 6, the optimal parameter values for \mathbf{q} with the corresponding confidence intervals are given for PAR(8,35,35). We see that the confidence intervals computed with respect to $\bar{b}_1, \sigma_{b_1}^2, \bar{b}_2, \sigma_{b_2}^2$, i.e. those associated with the individual growth rate b , are relatively large in comparison to the optimal value obtained from this method for these parameters. Therefore, we can not be very certain when reporting values for these particular parameters when using PAR(8,35,35). However, we see that the confidence intervals obtained for $\bar{\gamma}_1, \sigma_{\gamma_1}^2, \bar{\gamma}_2, \sigma_{\gamma_2}^2$, i.e., those associated with the maximum size γ , are very small in comparison to the optimal values obtained using this method. Based on this analysis, we feel more certain about the estimates obtained for these parameters because of the much smaller confidence intervals associated with these parameters.

Overall, we found that the estimated growth rate distributions obtained using DEL(32,16) produced the best fit to the field data in comparison to both SPL(5,35,9,35) and PAR(8,35,35). We also found the results produced by PAR(8,35,35) were very much comparable to those obtained using SPL(5,35,9,35) in a more efficient manner in terms of computational time. From the sensitivity analysis done on the

Parameter	$\hat{\theta} \pm t_{1-\alpha/2}\text{SE}(\hat{\theta})$
\bar{b}_1	1.9749 ± 395.3664
$\sigma_{b_1}^2$	$0.0388 \times 10^{-3} \pm 3.1450$
\bar{b}_2	3.9132 ± 139.7600
$\sigma_{b_2}^2$	$0.2228 \times 10^{-3} \pm 1.5226$
$\bar{\gamma}_1$	$0.5265 \pm 0.8471 \times 10^{-4}$
$\sigma_{\gamma_1}^2$	$0.0122 \pm 0.2982 \times 10^{-4}$
$\bar{\gamma}_2$	$0.7208 \pm 0.1426 \times 10^{-3}$
$\sigma_{\gamma_2}^2$	$0.0372 \pm 0.8201 \times 10^{-4}$

Table 6: Estimated parameters for Bi-Gaussian $p(b, \gamma)$ and confidence intervals for PAR(8,35,35) with field data

estimates obtained from SPL(5,35,9,35) and PAR(8,35,35), we observed that for this inverse problem the estimates from the spline based approximation method are not very reliable (based on the resulting large confidence intervals). We are more certain about the parameters related to γ when using the parameterized OLS than the parameters related to b . Overall, based on the fit-to-data and computational time, the delta function approximation provides the best estimates of the growth rate distribution for the field data in this example.

5 Concluding Remarks

In this paper we have presented computational and statistical comparisons of two “finite-element” type approximation schemes for estimation of probability distributions arising in problems with aggregate or population level observations. One approximation scheme, DEL, is at the level of the distribution being sought, while the other is at the level of the density (tacitly assumed to exist in the underlying theory [12]) associated with the distribution. Strengths and weaknesses are illustrated in several computational examples with simulated data in the context of functional growth rate estimation in size-structured population models. Finally, the most recently developed scheme, SPL, is used with experimental data for mosquitofish populations and compared to earlier findings with the DEL scheme. While the results here are specific to size-structured population models, the ideas are widely applicable to other areas of applications in biology as well as in viscoelastic materials and in electromagnetic interrogation of dielectric materials.

Acknowledgments

This research was supported in part (JLD) by the US Department of Energy Computational Science Graduate Fellowship under grant DE-FG02-97ER25308, in part (HTB) by the US Air Force Office of Scientific Research under grant FA9550-04-1-0220, in part (HTB) by the Joint DMS/NIGMS Initiative to Support Research in the Area of Mathematical Biology under grant 1R01GM67299-01, and in part (HTB and JLD) by the National Science Foundation under grant DMS-0112069 to the Statistical and Applied Mathematical Sciences Institute (SAMSI).

References

- [1] H.T. Banks, *Remarks on uncertainty assessment and management in modeling and computation*, Mathematical and Computer Modelling, **33** (2001), 39–47.
- [2] H.T. Banks and K.L. Bihari, *Modelling and estimating uncertainty in parameter estimation*, Inverse Problems, **17** (2001), 95–111.
- [3] H.T. Banks, D.M. Bortz, G.A. Pinter and L.K. Potter, *Modeling and imaging techniques with potential for application in bioterrorism*, CRSC-TR03-02, January, 2003; Chapter 6 in *Bioterrorism: Mathematical Modeling Applications in Homeland Security*, (H.T. Banks and C. Castillo-Chavez, eds.), Frontiers in Applied Math, **FR28**, SIAM, Philadelphia, PA, pp. 129–154, 2003.
- [4] 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, pp. 521–541, 1988.
- [5] H.T. Banks and N.L. Gibson, *Well-posedness in Maxwell systems with distributions of polarization relaxation parameters*, CRSC-TR04-01, NCSU, January, 2004; Applied Math. Letters, **18** (2005), 423–430.
- [6] H.T. Banks and N.L. Gibson, *Electromagnetic inverse problems involving distributions of dielectric mechanisms and parameters*, CRSC-TR05-29, August, 2005; Inverse Problems, submitted.
- [7] H.T. Banks and D.W. Iles, *A comparison of stability and convergence properties of techniques for inverse problems*, LCDS 86-3, Brown University, Providence, RI, January 1986.
- [8] H.T. Banks and D.W. Iles, *On compactness of admissible parameter sets: Convergence and stability in inverse problems for distributed parameter systems*, in Control Problems for Systems Described by Partial Differential Equations and Applications, (I. Lasiecka, R. Triggiani, eds.), LN in Control and Inf. Sci., Vol. 97, pp. 130–142, 1987.
- [9] H.T. Banks and B.G. Fitzpatrick, *Estimation of growth rate distributions in size-structured population models*, CAMS Tech. Rep. 90-2, University of Southern California, January, 1990; Quart. Appl. Math., **49** (1991), 215–235.
- [10] H.T. Banks, B.G. Fitzpatrick, L.K. Potter, and Y. Zhang, *Estimation of probability distributions for individual parameters using aggregate population data*, CRSC-TR98-6, (January, 1998); In Stochastic Analysis, Control, Optimization and Applications, (Edited by W. McEneaney, G. Yin and Q. Zhang), Birkhäuser Verlag, Basel, pp. 353–371, 1998.
- [11] H.T. Banks and K. Kunisch, *Estimation Techniques for Distributed Parameter Systems*, Birkhäuser, Boston, 1989.
- [12] H.T. Banks and G.A. Pinter, *A probabilistic multiscale approach to hysteresis in shear wave propagation in biotissue*, CRSC-TR04-03, N.C. State University, January, 2004; SIAM J. Multiscale Modeling and Simulation, **3** (2005), 395–412.
- [13] H.T. Banks and L.K. Potter, *Probabilistic methods for addressing uncertainty and variability in biological models: Application to a toxicokinetic model*, CRSC-TR02-27, N.C. State University, September, 2002; Math. Biosci., **192** (2004), 193–225.
- [14] S.L. Beal and L.B. Sheiner, *Estimating population kinetics*, CRC Critical Rev. in Biomed. Engr., **8** (1982), 195–222.
- [15] P. Billingsley, *Convergence of Probability Measures*, Wiley, New York, 1968.
- [16] Richard L. Burden and J. Douglas Faires, *Numerical Analysis* third edition, PWS Publishers, Boston, 1985.
- [17] M. Davidian and A.R. Gallant, *Smooth nonparametric maximum likelihood estimation for population pharmacokinetics, with application to quinidine*, J. Pharm. and Biopharm., **20** (1992), 529–556.

- [18] M. Davidian and A.R. Gallant, *The nonlinear mixed effects model with a smooth random effects density*, *Biometrika*, **80** (1993), 475–488.
- [19] M. Davidian and D.M. Giltan, *Nonlinear Models for Repeated Measurement Data*, Chapman & Hall, London, 1995.
- [20] M. Davidian and D.M. Giltan, *Nonlinear models for repeated measurement data: An overview and update*, *J. Agricultural, Biological, and Environ. Statistics*, **8** (2003), 387–419.
- [21] M.V. Evans, W.K. Boyes, P.J. Bushnell, J.H. Raymer and J.E. Simmons, *A physiologically based pharmacokinetic model for trichloroethylene (TCE) in Long-Evans rats*, preprint, 1999.
- [22] A.R. Gallant, *Nonlinear Statistical Models*, J. Wiley & Sons, New York, 1987.
- [23] N.L. Gibson, *Terahertz-Based Electromagnetic Interrogation Techniques for Damage Detection*, PhD Thesis, N. C. State University, Raleigh, August, 2004.
- [24] R.V. Hogg and E.A. Tanis *Probability and Statistical Inference* Prentice Hall, Upper Saddle River, 2001.
- [25] R.I. Jennrich, *Asymptotic properties of non-linear least squares estimators*, *Ann. Math. Statist.*, **40** (1969), 633–643.
- [26] B.G. Lindsay, *The geometry of mixture likelihoods: A general theory*, *Ann. Statist.*, **11** (1983), 86–94.
- [27] B.G. Lindsay, *Mixture Models: Theory, Geometry and Applications*, Vol. **5**, NSF-CBMS Regional Conf. Series in Probability and Statistics, Inst. Math. Stat., Haywood, CA, 1995.
- [28] B.G. Lindsay and M.L. Lesperance, *A review of semiparametric mixture models*, *J. Statist. Planning and Inference*, **47** (1995), 29–39.
- [29] A. Mallet, *A maximum likelihood estimation method for random coefficient regression models*, *Biometrika*, **73** (1986), 645–656.
- [30] L.B. Sheiner, B. Rosenberg and K. Melmon, *Modeling of individual pharmacokinetics for computer-aided drug dosage*, *Comp. Biomed. Res.*, **5** (1972), 441–459.
- [31] L.K. Potter, *Physiologically Based Pharmacokinetic Models for the Systemic Transport of Trichloroethylene*, Ph.D. Thesis, N.C. State University, August, 2001 www.lib.ncsu.edu.
- [32] A. Quarteroni, R. Sacco, and F. Saleri *Numerical Mathematics* Springer, New York, 2000.
- [33] A. Schumitzky, *Nonparametric EM algorithms for estimating prior distributions*, *Applied Mathematics and Computation*, **45** (1991), 141–157.
- [34] A. Schumitzky, *The nonparametric maximum likelihood approach to pharmacokinetic population analysis*, Proceedings of the Western Simulation Multiconference-Simulation in Health Care, Soc. for Computer Simulation, 1993.
- [35] G. Wahba, *Bayesian “confidence intervals” for the cross-validated smoothing spline*, *J. Roy. Statist. Soc. B*, **45** (1983), 133–150.
- [36] G. Wahba, *(Smoothing) splines in nonparametric regression*, Dept of Statistics-TR 1024, Univ. Wisconsin, September, 2000; In *Encyclopedia of Environmetrics*, (A. El-Shaarawi and W. Piegorsch, eds.), Wiley, Vol. 4, pp. 2099–2112, 2001.