

Estimation in Time-Delay Modeling of Insecticide-Induced Mortality

H.T. Banks,^{*} John E. Banks[†] and Sarah Lynn Joyner^{*}

October 26, 2008

^{*}Center for Research in Scientific Computation
North Carolina State University
Raleigh, North Carolina 27695-8212

[†]Environmental Science
Interdisciplinary Arts and Sciences
University of Washington
Tacoma, Washington 98402-3100

Abstract

We present a mathematical and statistical computational framework for inverse problems involving delay or hysteretic differential equations. We demonstrate efficacy of the methodology in the context of models for insect maturation and mortality due to insecticide exposure.

Key words: Inverse problems, ordinary and generalized least squares, delay-differential equation models, insect populations, insecticide exposure, time-varying mortality and maturation rates.

1 Introduction

We present an inverse problem framework for studying systems with distributed temporal delays. In particular, we discuss a general mathematical and statistical framework including implementable approximation ideas for inverse problems involving measure dependent dynamical systems. Such dynamical systems arise in the study of multiscale phenomena in diverse fields such as immunological population dynamics including HIV modeling, viscoelasticity of polymers and rubber, and polarization in dielectric materials (see [5, 6] and the references therein). We develop these ideas in the context of new models for problems arising in investigating insecticide-induced mortality for insect populations.

The widespread use of chemical pesticides motivates mathematical modeling of their effects on insect populations in hopes of understanding the relationship between vital rates and insecticide dosage. Mortality and reproductive rates are the two most important parameters to consider, and differential equation models are well-suited for the study of this relationship. Their capacity to incorporate time-varying parameters is critical to successful modeling of population dynamics, as evidenced by previous efforts [3, 4]. Previous models, however, have not accounted for reproductive effects of the insecticides, including the variability in time at which juvenile insects become adults. To this end, we propose a coupled time-delay or functional differential equation (FDE) model for populations exposed to insecticides.

Our previous work on the subject of insect mortality due to exposure to chemical pesticides began with sets of data taken as part of a study of neonate and adult pea aphids [24]. Modeling efforts began with a size-structured partial differential equation model, where it was seen that time-varying parameters provided a very good model fit to data as compared to the fits to data provided by constant parameter models [3, 12]. We subsequently hypothesized that the model's success was due entirely to the time-varying parameters and that the size structure in this particular case was not essential. To assess the validity of this hypothesis, we developed an ordinary differential equation model and, using the same data sets, were able to obtain excellent model fits [4]. However, because the data gave information for one generation of aphids only for separate populations of adults and neonates, this model did not incorporate reproductive effects, multiple generations or subclass structures involving transition from neonate to adult. In anticipation of multiple generation field data becoming available in the new future, we turned to more complex model structures to enhance our ability to investigate such data.

In this paper we present a time delay or functional differential equation model with time dependent parameters that might be used to investigate mixed neonate/adult multi-generational populations. We show how to formulate such models as special cases of a class of abstract differential equations with function space parameters (including probability densities) which are readily approximated by finite element systems. We formulate inverse problems in the form of both ordinary and generalized least squares frameworks, and carry out computations (including asymptotic error analysis involving approximate sampling distributions) with simulated noisy data to demonstrate both efficiency and efficacy of the proposed methodologies.

2 A Functional Differential Equation Model

Ideally, a mathematical model that is suitable for field data with mixed populations would consider reproductive effects and would also account for multiple generations, containing neonates and adults and their interconnectedness. This consideration suggests the need at the minimum for a coupled system of equations describing two separate age classes. Additionally, due to individual differences within the insect population, it is biologically unrealistic to assume that all neonate aphids born on the same day reach the adult age class at the same time. In fact, the age at which the insects reach adulthood varies from as few

as five to as many as seven days [15, 24]. Hence we must include a term in our model to account for this variability, leading us to develop a coupled delay differential equation model for the insect population dynamics. In this section we examine the delay between birth and adulthood for neonate pea aphids and present a first mathematical model that treats this delay as a random variable. For a careful derivation of models with similar structure in HIV progression at the cellular level, see [5].

Let $A(t)$ and $N(t)$ denote the number of adults and neonates, respectively, in the population at time t . We lump the mortality due to insecticide into one time-varying parameter $p_A(t)$ for the adults, $p_N(t)$ for the neonates, and denote by $d_A(t)$ and $d_N(t)$ the time-varying background or natural mortalities for adults and neonates respectively. We let $b(t)$ be the time-varying rate at which neonates are born into the population.

We suppose that there is a time delay for maturation of a neonate to adult life stage. We further assume that this time delay varies across the insect population according to a probability distribution $P(\tau)$ for $\tau \in [-T_N, 0]$ with corresponding density $m(\tau) = \frac{dP(\tau)}{d\tau}$. Here we tacitly assume an upper bound on T_N for the maturation period of neonates into adults. Thus, we have that $m(\tau)$, $\tau < 0$, is the probability per unit time that a neonate who has been in the population $-\tau$ time units becomes an adult. Then the rate at which such neonates become adults is $N(t + \tau)m(\tau)$. Summing over all such τ 's, we obtain that the rate at which neonates become adults is $\int_{-T_N}^0 N(t + \tau)m(\tau)d\tau$. Using the biological knowledge that the maturation process varies between five and seven days (i.e., m vanishes outside $[-7, -5]$), we obtain the functional differential equation (FDE) (see [1, 2, 5, 6, 11, 17, 18, 19, 21, 22, 25, 26, 27] for the widespread interest and use of such systems) system

$$\begin{aligned} \frac{dA}{dt}(t) &= \int_{-7}^{-5} N(t + \tau)m(\tau)d\tau - (d_A(t) + p_A(t))A(t) \\ \frac{dN}{dt}(t) &= b(t)A(t) - (d_N(t) + p_N(t))N(t) - \int_{-7}^{-5} N(t + \tau)m(\tau)d\tau \\ A(\theta) &= \Phi(\theta), \quad N(\theta) = \Psi(\theta), \quad \theta \in [-7, 0] \\ A(0) &= A^0, \quad N(0) = N^0, \end{aligned} \tag{1}$$

where m is now a probability density kernel which we have assumed has the property $m(\tau) \geq 0$ for $\tau \in [-7, -5]$ and $m(\tau) = 0$ for $\tau \in (-\infty, -7) \cup (-5, 0]$.

3 FDE Formulation and Approximation

For general inverse problem algorithms, one first needs good simulation methods. The system of functional differential equations described in (1) can be simulated using an algorithm first developed by Banks and Kappel for the linear time-invariant case [11] and later extended by Banks and Rosen to nonlinear and time-dependent systems [2, 14]. To use this algorithm, we first convert the system to an abstract evolution equation (AEE) and then approximate the solutions in a space spanned by piecewise linear splines. Thus we can numerically calculate the generalized Fourier coefficients of the approximate solution in the spline basis representation and recover an approximation to the solution of (1).

Let

$$x(t) = (A(t), N(t))^T$$

and

$$x_t(\tau) = x(t + \tau), -7 \leq \tau \leq 0. \quad (2)$$

We define the Hilbert space $Z \equiv \mathbb{R}^2 \times L^2(-7, 0; \mathbb{R}^2)$ with inner product

$$\|(\eta, \varphi)\|_Z = \left(|\eta|^2 + \int_{-7}^0 |\varphi(\theta)|^2 d\theta \right)^{1/2}, (\eta, \varphi) \in Z,$$

and let $z(t) = (x(t), x_t) \in Z$. Then our system (1) can be written as

$$\begin{aligned} \frac{dx}{dt}(t) &= L(t, x(t), x_t) \text{ for } 0 \leq t \leq T, \\ (x(0), x_0) &= (\Phi(0), \Phi) \in Z, \quad \Phi \in \mathcal{C}(-7, 0; \mathbb{R}^2), \end{aligned} \quad (3)$$

where $T < \infty$ and for $\eta = (\psi^0, \zeta^0)^T \in \mathbb{R}^2$ and $\varphi = (\psi, \zeta)^T \in \mathcal{C}(-7, 0; \mathbb{R}^2)$

$$L(t, \eta, \varphi) = \begin{bmatrix} -d_A(t) - p_A(t) & 0 \\ b(t) & -d_N(t) - p_N(t) \end{bmatrix} \eta + \begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix} \int_{-7}^{-5} \varphi(\tau) m(\tau) d\tau. \quad (4)$$

We now define a linear operator $\mathcal{A} : \mathcal{D}(\mathcal{A}) \subset Z \rightarrow Z$ with domain

$$\mathcal{D}(\mathcal{A}) = \{(\eta, \varphi) \in Z \mid \varphi \in H^1(-7, 0; \mathbb{R}^2) \text{ and } \eta = \varphi(0)\} \quad (5)$$

by

$$\mathcal{A}(t, \eta, \varphi) = (L(t, \eta, \varphi), \dot{\varphi}). \quad (6)$$

Note that while \mathcal{A} depends on t , $\mathcal{D}(\mathcal{A})$ does not. Then the delay system (1) can be formulated as

$$\begin{aligned} \dot{z}(t) &= \mathcal{A}(t)z(t) \\ z(0) &= z_0, \end{aligned} \quad (7)$$

where $z_0 = ((A^0, N^0)^T, (\Phi, \Psi)^T)$.

Define Z^N to be an approximating piecewise linear spline subspace of Z , Π^N as the orthogonal projection of Z onto Z^N , and $\mathcal{A}^N(t)$ as the approximating operator for $\mathcal{A}(t)$ given by $\mathcal{A}^N(t) = \Pi^N \mathcal{A}(t) \Pi^N$. Then the problem given in (7) is approximated by the finite dimensional problem

$$\begin{aligned} \dot{z}^N(t) &= \mathcal{A}^N(t)z^N(t), \quad t \geq 0, \\ z^N(0) &= \Pi^N z_0. \end{aligned} \quad (8)$$

We fix the basis Z_1^N for a special case of Z^N corresponding to the partition $t_j^N = -j(7/N)$ for $j = 0, \dots, N$. Then the basis is defined by

$$\hat{\beta}^N = (\beta^N(0), \beta^N) \text{ where } \beta^N = (e_0^N, e_1^N, \dots, e_N^N) \otimes \mathbb{I}_2, \quad (9)$$

and the e_j^N 's are piecewise linear splines defined by

$$e_j^N(t_i^N) = \delta_{ij} \text{ for } i, j = 0, 1, \dots, N.$$

When $\mathcal{A}^N(t)$ is restricted to Z_1^N , we have a matrix representation $A_1^N(t)$ of $\mathcal{A}^N(t)$. Define $w^N(t)$ so that $z^N(t) = \hat{\beta}^N w^N(t)$. Then solving for $z^N(t)$ in the finite dimensional system (8) is equivalent to solving for $w^N(t)$ in the linear system

$$\begin{aligned} \dot{w}^N(t) &= A_1^N(t)w^N(t) \\ w^N(0) &= w_0^N, \end{aligned} \tag{10}$$

where $\hat{\beta}^N w_0^N = \Pi^N z_0$. We note that having obtained w^N , the product $\hat{\beta}^N w^N$ converges uniformly in t to the solution of (7), $z(t) = (x(t), x_t)$.

3.1 Convergence of Solutions

According to mathematical theory (see [14] for proofs), the solutions of (8) converge to the solution of (7) as the number N of uniformly spaced elements in $[-7, 0]$ goes to infinity. When we perform an inverse problem using the delay differential equation model, we need to know how large to take our N so that we obtain accurate estimates of our parameters. To find the N at which solutions are nearly converged, we carry out the forward problem for a given set of parameters and for increasing values of N and plot the solutions.

For this forward problem, we require knowledge of the parameter functions $d_A(t)$, $p_A(t)$, $d_N(t)$, $p_N(t)$, and $b(t)$ as well as the probability density $m(\tau)$. For realistic solutions, such as a non-negative population for all time, we use our previously developed ordinary differential equation model [4] to estimate mortality parameters for a given level of insecticide exposure and adapt these parameters for use in the FDE model presented here. That is, in [4] we obtained reasonable “best” parameters using experimental data sets in a ordinary least squares inverse problem formulation. We use the general form of these functions now for our study of the corresponding delay systems. For reasons that will subsequently be obvious, we call these “true” parameters and denote them by $\Theta_0 = (d_{A0}, p_{A0}, d_{N0}, p_{N0}, b_0, m_0)$. It is important to note that each of these six time-varying parameter functions will be approximated using piecewise linear splines, so as the number of nodes taken for each function increases, we enlarge the corresponding nodal parameter space (methods to investigate the general problem of appropriate level of discretization for the parameter functions are discussed in [4, 13, 20]). The nodal values for all of the parameter functions are summarized in Table 1.

Based on previous experience, we take five equally spaced nodes for each of the functions except for $m(\tau)$, which we define on $[-7, -5]$ by

$$m(\tau) = \sum_{i=0}^2 a_i \ell_i(t) \tag{11}$$

where the $\ell_i(t)$'s are hat functions defined on the partition of $[-7, -5]$ with $t_i = -7 + i$, $i = 0, 1, 2$ and $a_0 = a_2 = 0$, $a_1 = 1$. That is, the “true” density kernel is the single hat function $\ell_1(t)$.

For our chosen exposure level, we have that the minimum length of time the aphids survive after introduction of the insecticide is 15 days. Thus we use only the first 15 days

Table 1: Nodal values for the parameter functions $\Theta_0 = (d_{A0}, p_{A0}, d_{N0}, p_{N0}, b_0, m_0)$ for (1).

Parameter Function	Region of Support	# of Equally Spaced Nodes	Nodal Values
$d_A(t)$	$[0, 15]$	5	$(0, 0, .041437, .022689, 0)$
$p_A(t)$	$[0, 15]$	5	$(0, .010763, .012490, .11679, .0098291)$
$d_N(t)$	$[0, 15]$	5	$(0, 0, 0, 0, .034625)$
$p_N(t)$	$[0, 15]$	5	$(0, .19273, .29144, 0, .073119)$
$b(t)$	$[0, 15]$	5	$(0, 6.4828, 4.1277, 1.0357, 0)$
$m(\tau)$	$[-7, -5]$	3	$(0, 1, 0)$

of data to estimate all four mortality parameters, and we likewise define the birth rate on the time interval 0 to 15 days. This rate $b(t)$ is estimated by interpolating the reproduction data for the adult age class at the given exposure level at five equally spaced nodes.

Having set the parameters, we use the Matlab integrator `ode15s` to solve the system with $N = 16, 32, 64, 128, 256, 512, 1024$. The solutions are plotted together in Figure 1.

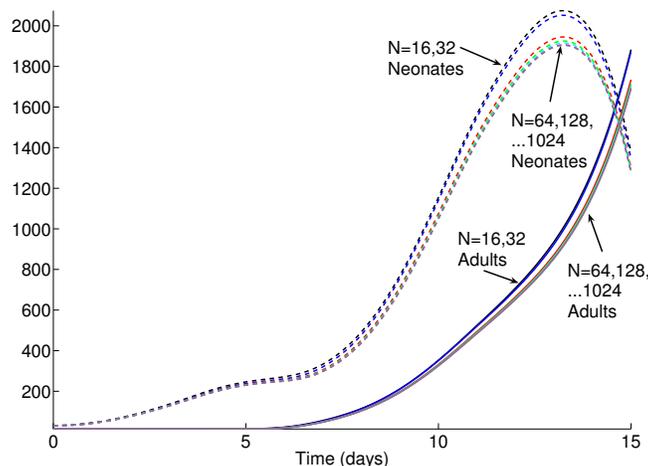


Figure 1: Solutions to delay system for increasing N .

At $N = 16$ and $N = 32$ elements, the solutions are not converged, but they begin to approach the limit at $N = 64$, and by $N = 256$, the curves are nearly indistinguishable. However, one important consideration is the computational time required for one iteration of the forward problem. The times required for the Matlab integrator to solve the system for each N are shown in Table 2. Time is not a factor through $N = 128$, but the computational time increases by a factor of 6 when we increase the number of elements to 256. We thus take $N = 128$ to be the number of elements needed to obtain converged solutions.

Table 2: Computational time required to solve the delay system for increasing N .

N	Time (s)	N	Time (s)
16	0.787528	256	24.521670
32	0.985615	512	200.121309
64	1.317057	1024	11973.085994
128	4.219158		

4 Inverse Problem and Algorithm Analysis

We wish to eventually estimate the time-varying parameters $d_A(t)$, $p_A(t)$, $d_N(t)$, $p_N(t)$, and $b(t)$ as well as the probability density $m(\tau)$ from experimental data. As a first step in this direction, we demonstrate the numerical and statistical accuracy of our model and associated inverse problem algorithms with simulated “data.” To do this, we simulate a solution at the converged level with the known parameters used in the convergence analysis. We then add noise to this solution and finally carry out an inverse problem using the noise-added solution as data to attempt to recover the original parameters. That is, we test our methods on *simulated data* to demonstrate the efficacy of our mathematical and statistical methodology.

4.1 Ordinary Least Squares

With the six time-varying parameter functions defined as in Table 1 of Section 3.1, we set $\Theta = \Theta_0 = (d_{A0}, p_{A0}, d_{N0}, p_{N0}, b_0, m_0)$ and solve the delay system on the time interval $t = 0$ to $t = 15$ days. We then generate “data” to use in ordinary least squares (OLS) inverse problems by adding noise to the solution according to the *statistical model*

$$y_i = x^N(t_i, \Theta_0) + \varepsilon_i \quad (12)$$

where each component of ε_i is taken from a normal distribution with mean zero and variance one, and Θ_0 is a set of “true” parameters, tacitly assumed to exist in standard statistical treatments. That is, we are first testing our methodology with so-called “*constant variance*” noise (see [3] for full discussions) for which an ordinary least squares formulation is appropriate.

We optimize over a constrained parameter space using the Matlab routine *fmincon* to minimize the OLS cost functional

$$J^N(\Theta) = \frac{1}{2n} \sum_{i=0}^n |y_i - x^N(t_i, \Theta)|^2 \quad (13)$$

where

$$y_i = (x_A^N(t_i, \Theta_0), x_N^N(t_i, \Theta_0)) + (\varepsilon_i^A, \varepsilon_i^N)$$

Table 3: Constraints on delay system parameters.

Parameter	Lower Bound	Upper Bound	Other
d_A	0	1	
p_A	0	1	
d_N	0	1	
p_N	0	1	
b	0	8	
$m(\tau)$	0	—	$\int_{-7}^{-5} m(\tau) d\tau = 1$

and $x^N(t_i, \Theta)$ is the mathematical model solution using parameters Θ with N finite elements in the delay interval $[-7, 0]$. For the optimization, we take $N = 128$ in (13), while the simulated data was generated using $N = 256$ in (12). The constraints imposed upon the parameters are summarized in Table 3. Note that in the case of upper or lower bounds, the bound can simply be applied to all spline coefficients that define the piecewise linear function for each parameter. Additionally, the constraint on the integral value of $m(\tau)$ is implemented as a linear constraint using the trapezoidal rule.

Because our parameter space consists of a total of 28 parameters for nodal values with “true” values as given in Table 1, we begin by setting $\Theta = m$ and holding the other five parameter functions fixed at their “true” values $d_{A0}, p_{A0}, d_{N0}, p_{N0}, b_0$ for our minimization of (13). Once we see that the estimated values for this one parameter function provide a better fit to the data than the true parameter function, we enlarge our set of parameter functions to $\Theta = (b, m)$ and hold the other four parameter functions fixed at $d_{A0}, p_{A0}, d_{N0}, p_{N0}$. We carry out the inverse problem to estimate this Θ , and after obtaining new estimates, we enlarge the parameter space again by one parameter function and optimize over the new set of parameters Θ . We continue in this manner until Θ is the set of all six time-varying parameter functions. This method of gradually increasing the dimension of the parameter space allows us to see how well the inverse problem technique is performing as we increase the difficulty for the inverse problem by increasing the number of parameters to be estimated. As the number of degrees of freedom increases, we expect in general that if we perform the minimization well, the minimal value of the least squares cost functional should decrease: we indeed observe that this is the case, as is seen in Table 4.

Although the value of $J(\hat{\Theta})$ decreases as we increase the number of parameters, we do not necessarily obtain accurate estimates of the true parameters. As we see in Figure 2, when $\Theta = (d_A, p_A, d_N, p_N, b, m)$, the estimates for $m(\tau)$ and b are very close to the true parameters, but the estimated mortality rates differ drastically from the true parameters. This discrepancy can be attributed to the fact that both the background mortality and insecticide-induced mortality rates have the same region of support. These two mortality rates cannot be estimated independently; in fact, it is the result of their sum that is optimized by the routine rather than the individual parameter functions. Thus it is not surprising to

Table 4: Minimal values of (13) for increasing size parameter space.

Θ	# of Parameters	$J^{128}(\hat{\Theta})$
m	3	4.0534
(b, m)	8	3.7009
(p_N, b, m)	13	3.5224
(d_N, p_N, b, m)	18	3.5224
(p_A, d_N, p_N, b, m)	23	3.4628
$(d_A, p_A, d_N, p_N, b, m)$	28	3.4628

obtain inaccurate individual results for d_A , p_A , d_N , and p_N , but we expect the estimated sums $d_A(t) + p_A(t)$ and $d_N(t) + p_N(t)$ to be accurate approximations to the sums of the true parameter functions $d_{A0} + p_{A0}$ and $d_{N0} + p_{N0}$. This accuracy of sums is in fact the case, as we see in Figure 3.

However, it is unlikely that in practice we would estimate both the background and insecticide-induced mortality rates at the same time. Rather, we should expect to have control data which we could use in an inverse problem to estimate d_A and d_N in a manner similar to that for the ordinary differential equation model (see [4]). In this case, we would set $p_A = 0$ and $p_N = 0$, reducing the delay system (1) to

$$\begin{aligned} \dot{A}(t) &= \int_{-7}^{-5} N(t + \tau)m(\tau)d\tau - d_A(t)A(t) \\ \dot{N}(t) &= b(t)A(t) - d_N(t)N(t) - \int_{-7}^{-5} N(t + \tau)m(\tau)d\tau. \end{aligned} \tag{14}$$

Having estimated the background mortality rates, we can take these parameter values to be known and optimize over $\Theta = (p_A, p_N, b, m)$ to estimate parameters for cases where we do have longitudinal data during exposure to insecticide.

To verify that the estimates for the insecticide-induced mortality rates are accurate when we use this approach, we take $d_A = d_{A0}$ and $d_N = d_{N0}$ and find the parameters $\Theta = (p_A, p_N, b, m)$ that minimize (13). When we compare the true mortality parameters to the optimal mortality parameters, we see that the results (depicted in Figure 4) are much more accurate than when we optimize over all four mortality rates. We can examine the results numerically by comparing the infinity norm of the difference between the true and optimal parameters for both cases, and we see that the norm of the error is much smaller when we hold the background mortality rates at the true parameter values. We note that the minimized value of the least squares cost functional in this case is 3.4628, which is the same as when we optimize over all six parameter functions (see Table 4). The plots of the true and optimal parameters p_A and p_N are shown in Figure 4, and the norm comparisons are detailed in Table 5.

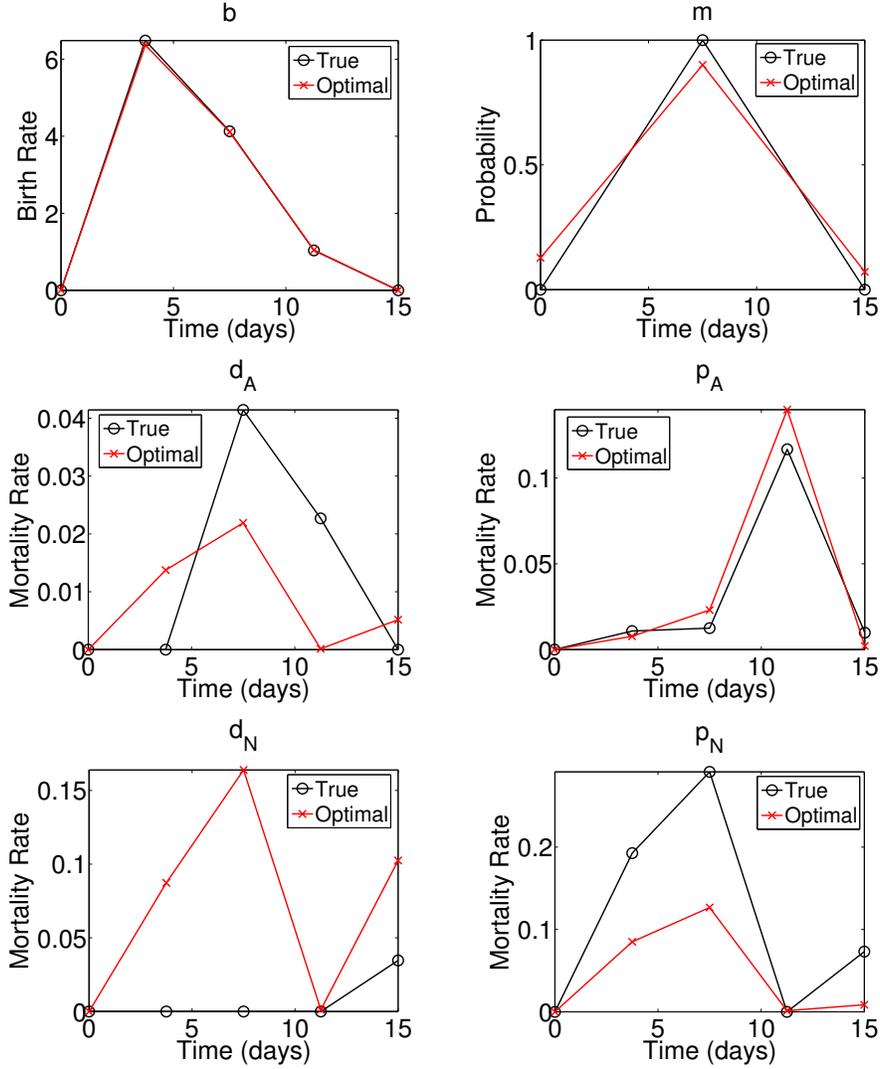


Figure 2: Comparison of estimated and true parameter functions.

Table 5: Error comparisons for insecticide-induced mortality rates.

Norm	$\Theta = (d_A, p_A, d_N, p_N, b, m)$	$\Theta = (p_A, p_N, b, m)$	Ratio
$\ \hat{p}_A^5 - p_{A0}^5\ _\infty$	2.3089×10^{-2}	1.0726×10^{-2}	2.1526
$\ \hat{p}_N^5 - p_{N0}^5\ _\infty$	1.6473×10^{-1}	2.0460×10^{-2}	8.0513
$J^N(\hat{\Theta})$	3.4628	3.4628	1

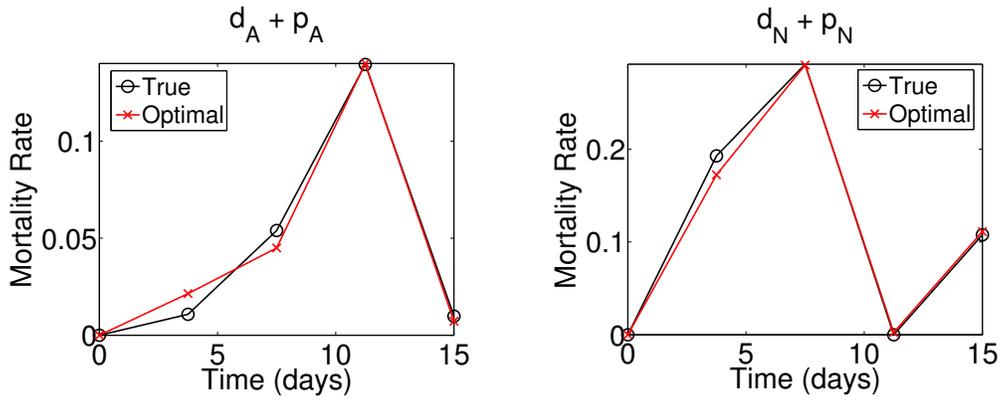


Figure 3: Comparison of sums of estimated and true mortality rates.

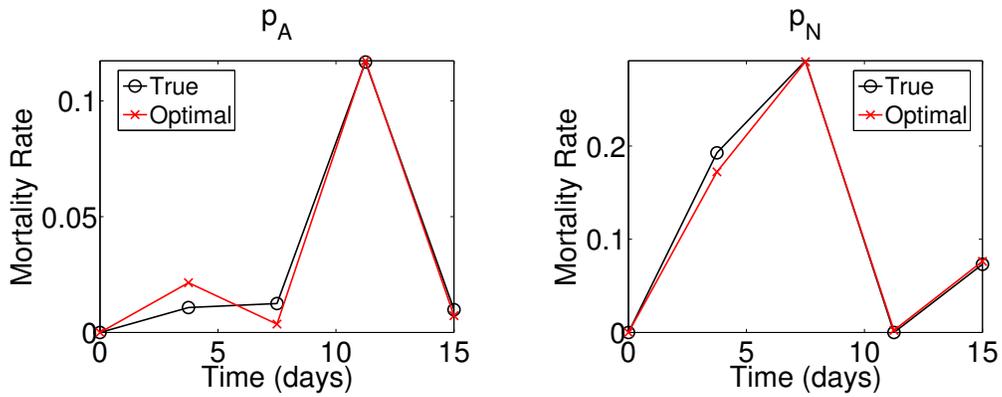


Figure 4: Comparison of true and optimal insecticide-induced mortality rates when background mortality rates are held at true values.

5 Parameter Estimation Error Analysis and Generalized Least Squares

We next present a discussion of statistical error analysis that should be included in reporting on any inverse problem findings. We do this in the context of a generalized least squares (GLS) inverse problem formulation which is appropriate when dealing with data containing relative noise (see the discussions in [7]). This type of data is frequently encountered when the data involves population counts (e.g., the error in counting 10 individuals vs. that in counting 1000 is likely to depend on the level of the count itself).

We consider a statistical model with “*nonconstant variance*” noise (or model response dependent relative error)

$$\vec{Y}_j = \vec{f}(t_j, \Theta) \left(1 + \frac{\nu}{100} \mathcal{E}_j\right), \quad (15)$$

where $\vec{f}(t_j, \Theta) \in R^m$ is our model solution using parameters Θ at time t_j , and we take $\nu = 5$, producing 5% relative error for our computational examples. To generate data, we obtain a realization $\{\vec{y}_j\}_{j=1}^n$ of the measurement random process $\{\vec{Y}_j\}_{j=1}^n$ using a realization $\{\epsilon_j\}_{j=1}^n$ of the error process $\{\mathcal{E}_j\}_{j=1}^n$. We then calculate an estimate $\hat{\Theta}$ of the true parameters Θ_0 using a generalized least squares (GLS) procedure, which for our problem involves finding solutions $\hat{\Theta}$ to the normal equations (see [7] for a more detailed discussion)

$$\sum_{j=1}^n [\vec{y}_j - \vec{f}(t_j, \hat{\Theta})]^T W_j^{-1}(\hat{\Theta}) \nabla \vec{f}(t_j, \hat{\Theta}) = 0, \quad (16)$$

where the $m \times m$ matrix W_j is given by

$$W_j(\hat{\Theta}) = \text{diag}[f_1^2(t_j, \hat{\Theta}), f_2^2(t_j, \hat{\Theta}), \dots, f_m^2(t_j, \hat{\Theta})]. \quad (17)$$

We will begin by estimating only the nodal values for the probability density $m(\tau)$ that accounts for the probability with which neonates become adults; subsequently we will consider estimating more parameters simultaneously.

5.1 Parameter Estimation for $m(\tau)$

The generalized least squares (GLS) procedure involves an iterative algorithm that consists of minimizing steps even though the GLS is not a minimization procedure (again see [7] for details).

We first use the OLS algorithm (even though with non-constant variance this is not an algorithm of choice—see [7]) to obtain an initial guess of $\hat{\Theta}$ for the iterative procedure. Having obtained the estimate $\hat{\Theta}_{\text{OLS}}$, we can carry out the GLS procedure to find the estimate $\hat{\Theta}_{\text{GLS}}$. This latter estimate is found using the iterative process described in [7] and [16] as well as

numerous other places. We begin by setting $\hat{\Theta}^{(0)} = \hat{\Theta}_{\text{OLS}}$ and the counter k to $k = 0$. Given $\hat{\Theta}^{(k)}$, we then form the weights $W_j(\hat{\Theta}^{(k)})$ of (17) and re-estimate $\hat{\Theta}$ by solving

$$\hat{\Theta}^{(k+1)} = \arg \min_{\Theta} \sum_{j=1}^n [\vec{y}_j - \vec{f}(t_j, \Theta)]^T W_j^{-1}(\hat{\Theta}^{(k)}) [\vec{y}_j - \vec{f}(t_j, \Theta)]. \quad (18)$$

Setting $k = k + 1$, we recompute the weighting matrices W_j and continue to iterate until two consecutive estimates for $\hat{\Theta}_{\text{GLS}}$ are sufficiently close.

We note that even though the GLS estimate is *not* defined by a minimization, the above iterative procedure was formulated by minimizing (over Θ) for a given $\tilde{\Theta}$ the quantity

$$\sum_{j=1}^n [\vec{y}_j - \vec{f}(t_j, \Theta)]^T W_j^{-1}(\tilde{\Theta}) [\vec{y}_j - \vec{f}(t_j, \Theta)]$$

and then updating the weights $W_j(\tilde{\Theta})$ after each iteration in the hope of obtaining a solution to the normal equations (16).

Because the GLS procedure is initialized with results from an ordinary least squares (OLS) parameter estimation, in this section we present both OLS and GLS results for estimating the nodal values for $m(\tau)$ and compute standard errors and confidence intervals in both cases. We shall see that even though the OLS formulas for standard errors are not correct when the data has nonconstant variance, it is not always easy to discern from the estimates and errors that the incorrect statistical model and corresponding error formulae were used in computing them. Indeed, estimates and standard errors computed with OLS formulae may appear perfectly reasonable even when used with inverse problem results employing nonconstant variance data. There are, however, post analysis residual tests that one may use to test correctness of statistical model assumptions after carrying out inverse problem calculations. Discussions and illustrative examples can be found in [7].

5.1.1 Ordinary Least Squares (OLS) Results

Our system requires that we use the vector form of the OLS procedure. In this case, the problem consists of finding the estimate

$$\hat{\Theta} = \hat{\Theta}_{\text{OLS}} = \arg \min_{\Theta} \sum_{j=1}^n [\vec{y}_j - \vec{f}(t_j, \Theta)]^T \hat{V}^{-1} [\vec{y}_j - \vec{f}(t_j, \Theta)], \quad (19)$$

where the $m \times m$ matrix \hat{V} is the estimate of the variance and is computed by

$$\hat{V} = \text{diag} \left(\frac{1}{n-p} \sum_{j=1}^n [\vec{y}_j - \vec{f}(t_j, \hat{\Theta})] [\vec{y}_j - \vec{f}(t_j, \hat{\Theta})]^T \right). \quad (20)$$

These estimates are coupled, and to solve the coupled system, we follow the iterative process outlined in [7]. This consists of alternating between computing \hat{V} using (20) with the current value of $\hat{\Theta}$ and then using this \hat{V} in (19) to compute a new value of $\hat{\Theta}$.

Carrying out the OLS procedure, we obtained the following estimates for nodal values a_i for $m(\tau)$ (see (11)), defined using 3 equally spaced nodes over the interval $[-7, -5]$:

$$\hat{m} \sim (\hat{a}_1, \hat{a}_2, \hat{a}_3) = (8.4906 \times 10^{-2}, 9.4126 \times 10^{-1}, 3.2575 \times 10^{-1}).$$

The model fit to the noise-added data is depicted in Figure 5. The estimate for the variance (note in our example $m = 2$) is

$$\hat{V} = \begin{bmatrix} 9.8967 \times 10^2 & 0 \\ 0 & 2.2676 \times 10^3 \end{bmatrix}.$$

Before carrying out the GLS procedure, we can compute the covariance matrix, standard errors, and 95% confidence intervals for these parameters. The covariance matrix (see [7]) is computed by

$$\hat{\Sigma}_{\text{OLS}}^n = \left(\sum_{j=1}^n D_j^T(\hat{\Theta}) \hat{V}^{-1} D_j(\hat{\Theta}) \right)^{-1}, \quad (21)$$

where, given an m -dimensional system with p parameters to estimate, the $m \times p$ matrix $D_j(\Theta)$ is given by

$$D_j(\Theta) = \nabla_{\Theta} \vec{f}(t_j, \Theta) \equiv \begin{bmatrix} \frac{\partial f_1(t_j, \Theta)}{\partial \theta_1} & \frac{\partial f_1(t_j, \Theta)}{\partial \theta_2} & \dots & \frac{\partial f_1(t_j, \Theta)}{\partial \theta_p} \\ \vdots & \vdots & & \vdots \\ \frac{\partial f_m(t_j, \Theta)}{\partial \theta_1} & \frac{\partial f_m(t_j, \Theta)}{\partial \theta_2} & \dots & \frac{\partial f_m(t_j, \Theta)}{\partial \theta_p} \end{bmatrix}. \quad (22)$$

In our case, we obtain the matrix

$$\hat{\Sigma}_{\text{OLS}}^n = \begin{bmatrix} 1.3887 \times 10^{-1} & -1.1554 \times 10^{-1} & 9.7153 \times 10^{-2} \\ -1.1554 \times 10^{-1} & 9.7919 \times 10^{-2} & -8.3291 \times 10^{-2} \\ 9.7153 \times 10^{-2} & -8.3291 \times 10^{-2} & 7.1430 \times 10^{-2} \end{bmatrix}.$$

We compute the standard errors [7] by taking the square root of the diagonal elements. We compute $SE_i = SE(a_i)$ as

$$SE_1 = 3.7265 \times 10^{-1}, \quad SE_2 = 3.1292 \times 10^{-1}, \quad SE_3 = 2.6726 \times 10^{-1}.$$

Finally, the 95% confidence intervals on the nodal values a_i are calculated using the standard errors in the usual manner (see [7]) and are found to be

$$\begin{aligned} \hat{a}_1 &: [-6.5145 \times 10^{-1}, 8.2127 \times 10^{-1}], \\ \hat{a}_2 &: [3.2293 \times 10^{-1}, 1.5596 \times 10^0], \\ \hat{a}_3 &: [-4.9554 \times 10^{-1}, 5.6069 \times 10^{-1}]. \end{aligned}$$

These results are summarized in Table 6.

5.1.2 Generalized Least Squares (GLS) Results

Using the GLS procedure with the OLS estimate from the previous section as an initial guess, we obtain the GLS estimate for $m(\tau)$

$$\hat{m} \sim (\hat{a}_1, \hat{a}_2, \hat{a}_3) = (5.4731 \times 10^{-2}, 9.7263 \times 10^{-1}, 1.5001 \times 10^{-19}).$$

With these estimated nodal values, we obtain the model fit to the noisy data depicted in Figure 5. A qualitative comparison of the OLS and GLS fits reveals little discernable difference in the fits to the data.

As with the OLS procedure, we can compute standard errors and confidence intervals for the GLS estimate. The variance is obtained using the formula

$$\hat{\sigma}_{\text{GLS}}^2 = \frac{1}{n-p} \sum_{j=1}^n [\vec{y}_j - \vec{f}(t_j, \hat{\Theta})]^T W_j^{-1}(\hat{\Theta}) [\vec{y}_j - \vec{f}(t_j, \hat{\Theta})], \quad (23)$$

giving us in this case the variance estimate

$$\hat{\sigma}_{\text{GLS}}^2 = 2.8339 \times 10^{-3}.$$

We compute the covariance matrix [7] by

$$\hat{\Sigma}_{\text{GLS}}^n = \hat{\sigma}_{\text{GLS}}^2 \left(\sum_{j=1}^n D_j^T(\hat{\Theta}) W_j^{-1}(\hat{\Theta}) D_j(\hat{\Theta}) \right)^{-1}, \quad (24)$$

where the matrix D_j is defined as in (22). This calculation for our example yields the covariance matrix

$$\hat{\Sigma}_{\text{GLS}}^n = \begin{bmatrix} 6.4283 \times 10^{-2} & -4.2979 \times 10^{-2} & 3.0996 \times 10^{-2} \\ -4.2979 \times 10^{-2} & 3.0887 \times 10^{-2} & -2.3494 \times 10^{-2} \\ 3.0996 \times 10^{-2} & -2.3494 \times 10^{-2} & 1.8675 \times 10^{-2} \end{bmatrix}.$$

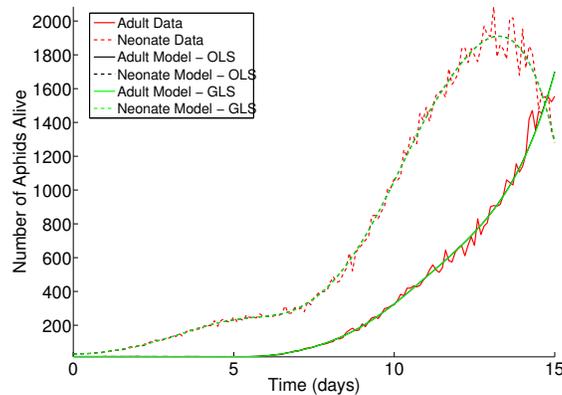


Figure 5: Model fits to noisy data using OLS and GLS estimates for nodal values for $m(\tau)$.

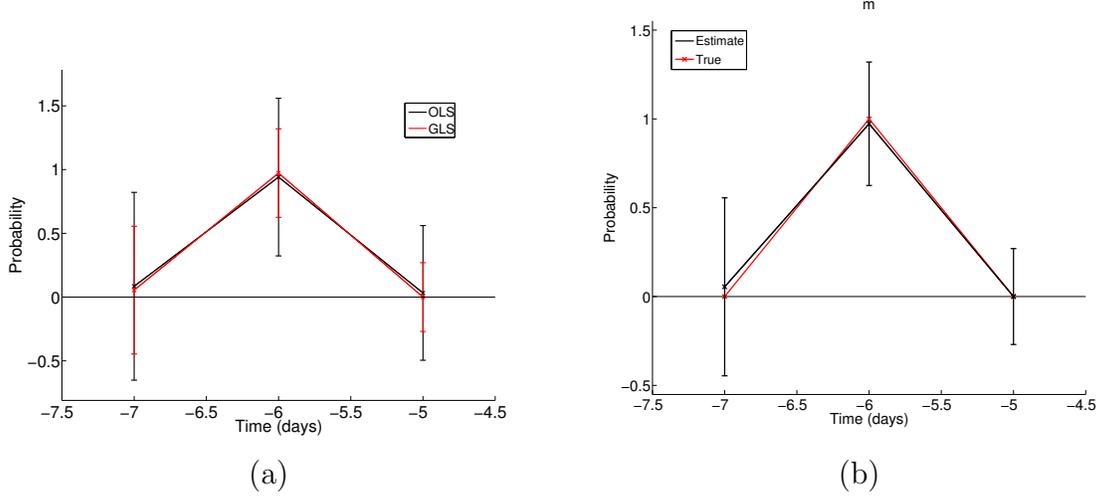


Figure 6: (a) OLS and GLS estimates for $m(\tau)$ with respective 95% confidence intervals; (b) True values for $m(\tau)$ and GLS estimate with 95% confidence intervals.

Table 6: Summary of statistics for OLS and GLS optimizations for $m(\tau)$.

Node	True Value	OLS		
		Estimate	Std. Err.	Conf. Bound
a_1	0	8.4906×10^{-2}	3.7265×10^{-1}	$\pm 7.3636 \times 10^{-1}$
a_2	1	9.4126×10^{-1}	3.1292×10^{-1}	$\pm 6.1833 \times 10^{-1}$
a_3	0	3.2575×10^{-2}	2.6726×10^{-1}	$\pm 5.2811 \times 10^{-1}$
Node	True Value	GLS		
		Estimate	Std. Err.	Conf. Bound
a_1	0	5.4731×10^{-1}	2.5354×10^{-1}	$\pm 5.0100 \times 10^{-1}$
a_2	1	9.7263×10^{-1}	1.7575×10^{-1}	$\pm 3.4728 \times 10^{-1}$
a_3	0	1.5001×10^{-19}	1.3666×10^{-1}	$\pm 2.7004 \times 10^{-1}$

To find the standard error for each nodal value, we take the square roots of the diagonal elements of the covariance matrix and find

$$SE_1 = 2.5354 \times 10^{-1}, \quad SE_2 = 1.7575 \times 10^{-1}, \quad SE_3 = 1.3666 \times 10^{-1}.$$

Finally, the 95% confidence intervals are:

$$\begin{aligned} \hat{a}_1 &: [-4.4627 \times 10^{-1}, 5.5573 \times 10^{-1}], \\ \hat{a}_2 &: [6.2536 \times 10^{-1}, 1.3199 \times 10^0], \\ \hat{a}_3 &: [-2.7004 \times 10^{-1}, 2.7004 \times 10^{-1}]. \end{aligned}$$

The GLS estimate and these 95% confidence intervals along with the OLS estimate and its corresponding 95% confidence intervals are depicted in Figure 6(a). We see that though

the estimates are similar, the confidence intervals are somewhat smaller (and correct because the data possessed nonconstant variance) for the GLS estimate. A plot of the true values for m along with the GLS estimate and confidence intervals can be seen in Figure 6(b), and we note that the true values are covered by the confidence intervals. Table 6 contains a summary of the OLS and GLS estimates for $m(\tau)$ along with the corresponding standard errors and confidence intervals.

5.2 Parameter Estimation for $b(t)$

Before estimating more than one parameter function at a time, we estimate the birth rate function $b(t)$ alone to examine our estimation error. Using the GLS procedure, we find

$$\hat{b} \sim (\hat{b}_1, \hat{b}_2, \hat{b}_3, \hat{b}_4, \hat{b}_5) = (0, 6.4692, 4.0572, 1.0466, 0),$$

and the estimate for the variance is $\hat{\sigma}_{\text{GLS}}^2 = 2.8499 \times 10^{-3}$. The covariance matrix is computed to be

$$\hat{\Sigma}_{\text{GLS}}^n = \begin{bmatrix} 7.6726 \times 10^{-2} & -4.8492 \times 10^{-2} & 3.2513 \times 10^{-2} & 1.5779 \times 10^{-3} & -7.1586 \times 10^{-4} \\ -4.8492 \times 10^{-2} & 3.3508 \times 10^{-2} & -2.4441 \times 10^{-2} & -6.3227 \times 10^{-4} & 2.9072 \times 10^{-4} \\ 3.2513 \times 10^{-2} & -2.4441 \times 10^{-2} & 1.9535 \times 10^{-2} & 3.8603 \times 10^{-6} & -1.2863 \times 10^{-4} \\ 1.5779 \times 10^{-3} & -6.3227 \times 10^{-4} & 3.8603 \times 10^{-6} & 2.6248 \times 10^{-4} & -7.7175 \times 10^{-5} \\ -7.1586 \times 10^{-4} & 2.9072 \times 10^{-4} & -1.2863 \times 10^{-4} & -7.7175 \times 10^{-5} & 7.5633 \times 10^{-5} \end{bmatrix},$$

and taking the square roots of the diagonal elements we obtain the standard errors

$$SE_1 = 2.7699 \times 10^{-1}, SE_2 = 1.8305 \times 10^{-1}, SE_3 = 1.3977 \times 10^{-1}, \\ SE_4 = 1.6201 \times 10^{-2}, SE_5 = 8.6967 \times 10^{-3}.$$

Using these standard errors, we obtain the 95% confidence intervals for each node:

$$\hat{b}_1 : [-5.4734 \times 10^{-1}, 5.4734 \times 10^{-1}], \\ \hat{b}_2 : [6.1075, 6.8309], \\ \hat{b}_3 : [3.7810, 4.3333], \\ \hat{b}_4 : [1.0146, 1.0786], \\ \hat{b}_5 : [-1.7185 \times 10^{-2}, 1.7185 \times 10^{-2}].$$

The estimate for $b(t)$ is plotted in Figure 7 along with the confidence intervals for each node and the true values for the function. As was the case when we estimated m , the estimate is very close to the true nodal values, which are covered by the confidence intervals. The results for the GLS procedure are summarized in Table 7.

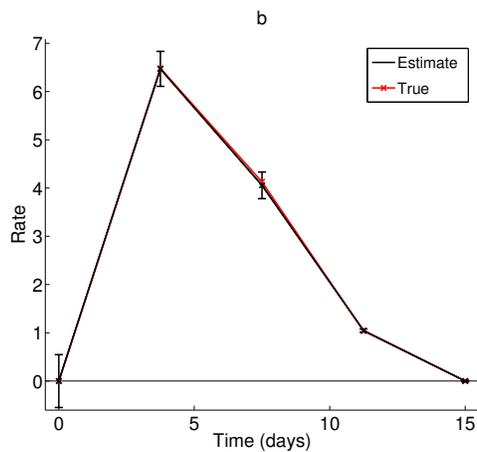


Figure 7: True values and GLS estimate for $b(t)$ with 95% confidence intervals for each node.

Table 7: Summary of statistics for GLS estimate for $b(t)$.

Node	True	Estimate	Std. Err.	Conf. Bound
b_1	0	0	2.7699×10^{-1}	$\pm 5.4734 \times 10^{-1}$
b_2	6.4828	6.4692	1.8305×10^{-1}	$\pm 3.6171 \times 10^{-1}$
b_3	4.1277	4.0572	1.3977×10^{-1}	$\pm 2.7618 \times 10^{-1}$
b_4	1.0357	1.0466	1.6201×10^{-2}	$\pm 3.2013 \times 10^{-2}$
b_5	0	0	8.6967×10^{-3}	$\pm 1.7185 \times 10^{-2}$

5.3 Estimating Two Parameter Functions

Having obtained estimates for one parameter at a time, we now consider our accuracy in simultaneously estimating two parameter functions. In this section, we consider estimation of both $m(\tau)$ and $b(t)$ and compare the standard errors and confidence intervals in this case with those obtained for the estimates of the two functions separately.

We used the GLS procedure to estimate nodes for $m(\tau)$ and $b(t)$. We obtained the estimates

$$\hat{m} \sim (\hat{a}_1, \hat{a}_2, \hat{a}_3) = (5.1543 \times 10^{-2}, 9.7423 \times 10^{-1}, 2.4161 \times 10^{-21})$$

and

$$\hat{b} \sim (\hat{b}_1, \hat{b}_2, \hat{b}_3, \hat{b}_4, \hat{b}_5) = (1.5229 \times 10^{-19}, 6.4980, 4.0881, 1.0466, 0).$$

The variance estimate is $\hat{\sigma}_{\text{GLS}}^2 = 2.8687 \times 10^{-3}$, and the standard errors and confidence bounds are summarized in Table 8. Figures 8(a) and 8(b), respectively, depict plots of m and b comparing the true and the estimated nodal values. For each function, the estimate is quite close to the true value, and the true values fall well within the 95% confidence intervals at each node.

In Figure 9, we see that while the estimates for m are similar regardless of whether the GLS procedure is performed for m alone or for m and b simultaneously, the confidence intervals are smaller in the former case. That is, an increase in the degrees of freedom apparently results in a less accurate estimate. However, the opposite is true for b . The estimates are again similar, but the confidence intervals are larger when this function alone is estimated than when it is estimated simultaneously with m . Thus, for this example at least, no general pattern arises for accuracy as a function of number of degrees of freedom.

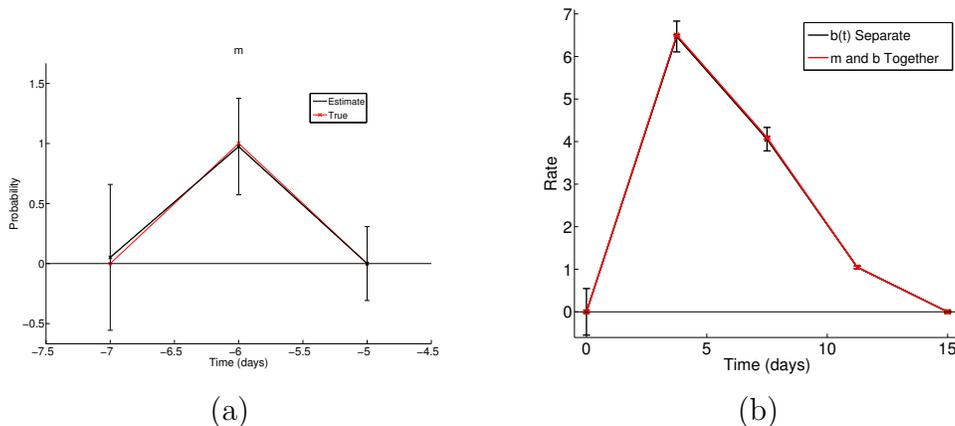
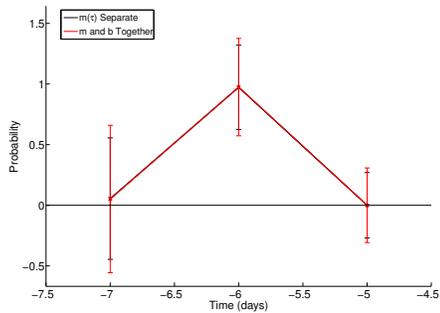


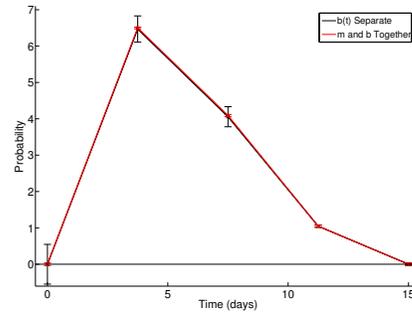
Figure 8: Estimates along with confidence intervals for (a) $m(\tau)$ and (b) $b(t)$, estimated simultaneously, compared with true values.

Table 8: GLS estimates with standard errors and confidence intervals for $m(\tau)$ and $b(t)$ estimated simultaneously.

Function	Node	True	Estimate	Std. Err.	Conf. Bound
$m(\tau)$	a_1	0	5.1543×10^{-2}	3.0740×10^{-1}	$\pm 6.0742 \times 10^{-1}$
	a_2	1	9.7423×10^{-1}	2.0315×10^{-1}	$\pm 4.0142 \times 10^{-1}$
	a_3	0	2.4161×10^{-21}	1.5572×10^{-1}	$\pm 3.0770 \times 10^{-1}$
$b(t)$	b_1	0	1.5229×10^{-19}	1.6998×10^{-2}	$\pm 3.3588 \times 10^{-2}$
	b_2	6.4828	6.4980	1.1031×10^{-2}	$\pm 2.1796 \times 10^{-2}$
	b_3	4.1277	4.0881	1.4496×10^{-2}	$\pm 2.8645 \times 10^{-2}$
	b_4	1.0357	1.0466	1.3529×10^{-2}	$\pm 2.6732 \times 10^{-2}$
	b_5	0	0	2.1415×10^{-2}	$\pm 4.2316 \times 10^{-2}$



(a)



(b)

Figure 9: Estimates and confidence intervals for (a) $m(\tau)$ and (b) $b(t)$ estimated separately and simultaneously.

5.4 Estimating Mortality Rates

Of our four mortality rate functions, we illustrate estimation of only the insecticide-induced rates p_A and p_N because in all likelihood, the background or natural mortality rates d_A and d_N will be known (e.g., previously estimated from control data). In this section, we first report on simultaneous estimation of the nodal values for the two mortality functions, holding the other parameter functions at their true values, for a total of ten parameters to be estimated. We compute standard errors and confidence intervals, then estimate these two parameters along with m and b to give a total of 18 nodal values to estimate: three for m and five each for b , p_N , and p_A . The error statistics are again computed and compared to those for prior sets of estimated parameters.

5.4.1 Estimation of Mortality Rates Alone

Table 9: GLS estimates with standard errors and confidence intervals for p_N and p_A estimated simultaneously.

Function	Node	True	Estimate	Std. Err.	Conf. Bound
$p_N(t)$	p_N^1	0	1.0300×10^{-2}	3.3025×10^{-1}	$\pm 6.3281 \times 10^{-1}$
	p_N^2	1.9273×10^{-1}	1.8511×10^{-1}	2.1914×10^{-1}	$\pm 4.3303 \times 10^{-1}$
	p_N^3	2.9144×10^{-1}	3.0341×10^{-1}	1.7357×10^{-1}	$\pm 3.4298 \times 10^{-1}$
	p_N^4	0	0	4.6126×10^{-16}	$\pm 9.1144 \times 10^{-16}$
	p_N^5	7.3119×10^{-2}	6.3624×10^{-2}	5.0588×10^{-2}	$\pm 9.9962 \times 10^{-2}$
$p_A(t)$	p_A^1	0	0	9.7503×10^{-2}	$\pm 1.9267 \times 10^{-1}$
	p_A^2	1.0763×10^{-2}	1.1982×10^{-2}	1.1292×10^{-1}	$\pm 2.2313 \times 10^{-1}$
	p_A^3	1.2490×10^{-2}	1.1946×10^{-2}	2.7209×10^{-2}	$\pm 5.3764 \times 10^{-2}$
	p_A^4	1.1679×10^{-1}	1.2177×10^{-1}	2.6070×10^{-3}	$\pm 5.1514 \times 10^{-3}$
	p_A^5	9.8291×10^{-3}	0	8.9004×10^{-4}	$\pm 1.7587 \times 10^{-3}$

First, we used the GLS procedure to find simultaneous estimates for the nodal values for the neonate and adult mortality rates, $p_N(t)$ and $p_A(t)$. We found the variance to be $\hat{\sigma}_{\text{GLS}}^2 = 2.8907 \times 10^{-3}$, then computed the covariance matrix using (24); we took the square root of the diagonal elements to find the standard errors and 95% confidence intervals as given in Table 9. Plots of the estimates for the two mortality rate functions with the confidence intervals along with their true values can be seen in Figure 10.

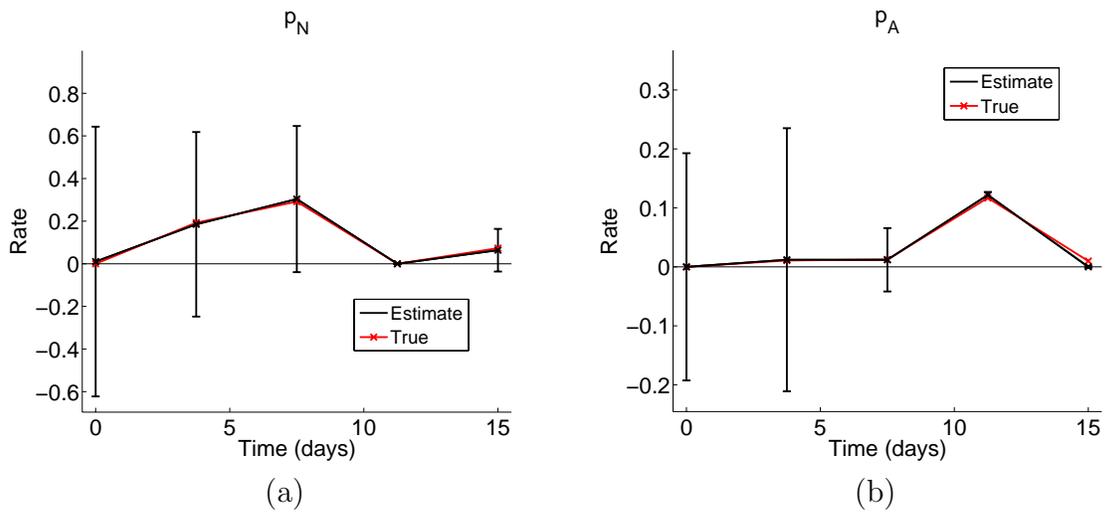


Figure 10: Simultaneous GLS estimates and 95% confidence intervals for (a) p_N and (b) p_A compared with true values for each function.

5.4.2 Estimation of Mortality Rates with b and m

Carrying out the GLS procedure to estimate nodal values for m , b , p_N , and p_A , we obtained the estimates given in Table 10. The variance was calculated using (23) to be $\hat{\sigma}_{\text{GLS}}^2 = 2.9541 \times 10^{-3}$. We computed the covariance matrix and took the square root of the diagonal elements to obtain the standard errors as reported in Table 10 along with the 95% confidence bounds that define the confidence intervals depicted in Figure 11. We see that the confidence intervals for several of the nodal values are quite large; however, the estimates are very similar to the true values, and the true nodal values fall well within the confidence intervals.

Table 10: GLS estimates with standard errors and confidence intervals for m , b , p_N , and p_A estimated simultaneously.

Function	Node	True	Estimate	Std. Err.	Conf. Bound
$m(\tau)$	a_1	0	8.2082×10^{-2}	6.6512×10^{-1}	± 1.3143
	a_2	1	9.5896×10^{-1}	2.8742×10^{-1}	$\pm 5.6795 \times 10^{-1}$
	a_3	0	8.0288×10^{-21}	2.5807×10^{-1}	$\pm 5.0994 \times 10^{-1}$
$b(t)$	b_1	0	0	6.9643×10^{-1}	± 1.3761
	b_2	6.4828	6.5983	1.9400×10^{-1}	$\pm 3.8334 \times 10^{-1}$
	b_3	4.1277	4.0703	1.1193×10^{-1}	$\pm 2.2118 \times 10^{-1}$
	b_4	1.0357	1.0630	7.3863×10^{-1}	± 1.4595
	b_5	0	6.4386×10^{-2}	6.1981×10^{-2}	$\pm 1.2247 \times 10^{-1}$
$p_N(t)$	p_N^1	0	1.6873×10^{-2}	2.9105×10^{-12}	$\pm 5.7511 \times 10^{-12}$
	p_N^2	1.9273×10^{-1}	1.9735×10^{-1}	1.2656	± 2.5007
	p_N^3	2.9144×10^{-1}	2.9413×10^{-1}	2.1905	± 4.3284
	p_N^4	0	1.3745×10^{-6}	1.2190	± 2.4087
	p_N^5	7.3119×10^{-2}	1.2305×10^{-1}	2.3855	± 4.7137
$p_A(t)$	p_A^1	0	2.7934×10^{-20}	3.3582×10^{-3}	$\pm 6.6357 \times 10^{-3}$
	p_A^2	1.0763×10^{-2}	1.0810×10^{-2}	1.8177×10^{-3}	$\pm 3.5918 \times 10^{-3}$
	p_A^3	1.2490×10^{-2}	0	6.1569×10^{-3}	$\pm 1.2166 \times 10^{-2}$
	p_A^4	1.1679×10^{-1}	1.2124×10^{-1}	2.9085×10^{-2}	$\pm 5.7473 \times 10^{-2}$
	p_A^5	9.8291×10^{-3}	0	2.6058×10^{-2}	$\pm 5.1490 \times 10^{-2}$

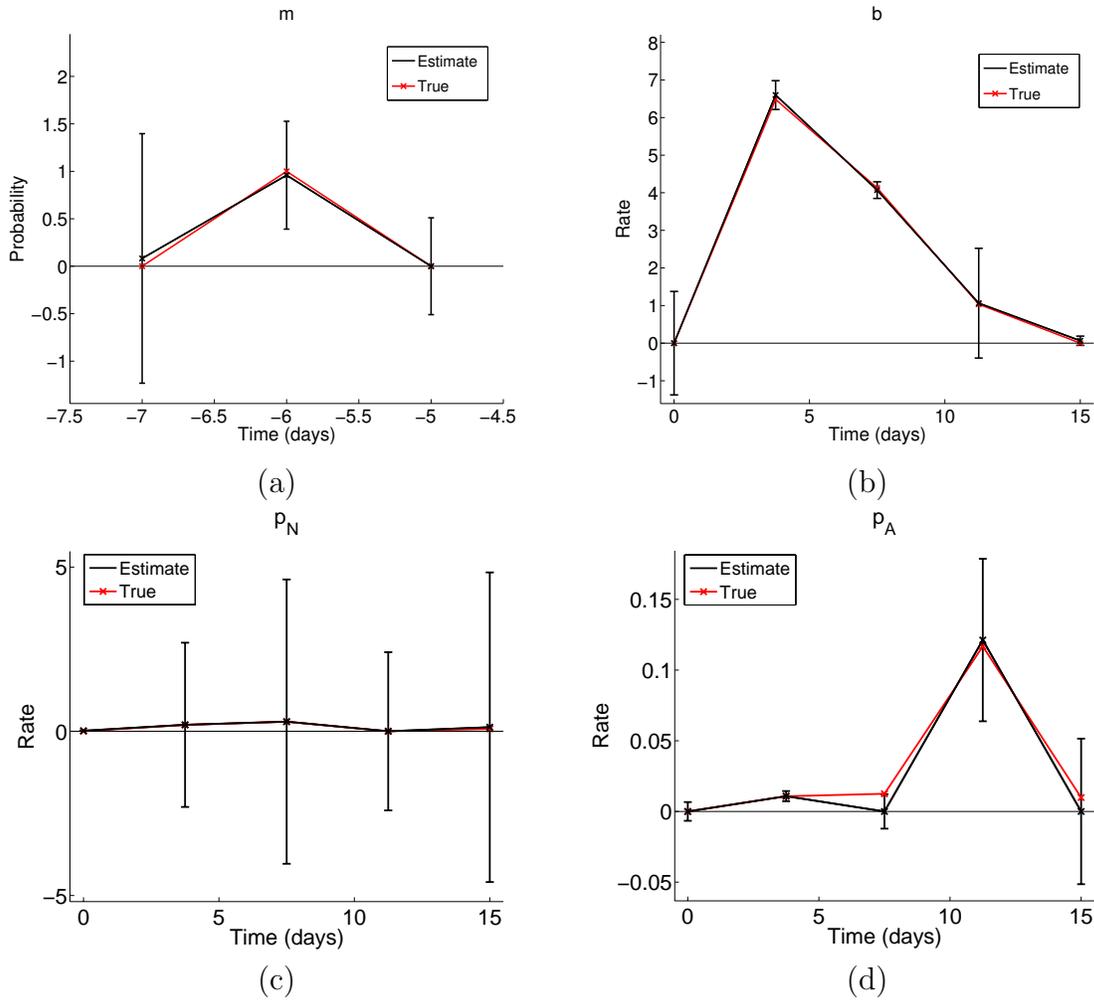


Figure 11: Simultaneous GLS estimates and 95% confidence intervals for (a) m , (b) b , (c) p_N , and (d) p_A with true values for each function. Note differing scales for p_N and p_A .

6 An Alternative Sensitivity Analysis

The above approach to sensitivity with respect to functional parameters involved first discretizing both the delay system (approximating it by a high dimensional ODE system) and then representing the time varying coefficients and the kernel m in terms of nodal values relative to a fixed set of basis elements—in this case piecewise linear spline elements. All estimation, sensitivity, standard error, etc., calculations were then carried out for the resulting finite dimensional state and finite dimensional parameter problems and ultimately employed existing nonlinear asymptotical statistical theory [23] for sampling distributions for the estimated nodal values. This “*discretize first and then optimize*” is often encountered in the control literature. An alternative approach, “*optimize in a function space setting and then discretize for computation*,” would require as a *first step* the computation of sensitivity (or gradients) of the original delay system (1) solutions with respect to its function space parameters (especially the probability kernel m). To compute these one might develop a set of sensitivity equations which could then be a basis for an (as yet undeveloped) infinite dimensional nonlinear asymptotic statistical framework (see [8, 9] for some initial efforts in this direction) for the computation of functional “confidence bands” for the estimated functions. This of course would involve infinite dimensional function space versions of the sensitivity and covariance matrices that are fundamental in the finite dimensional asymptotic theory for sampling distributions.

For example, one might consider the sensitivity of solutions of (1) with respect to the probability kernel m and follow the theoretical arguments of [10] to derive an equation (which will also be a time delay system) for this sensitivity. For the problem of this paper, one is particularly concerned with sensitivity to m because it is the one for which we have the least physical information. One knows only that neonate aphids become adults at some time between 5 and 7 days of age, but the shape of the associated probability distribution is completely unknown. Different toxicants will likely yield different shaped probability kernels, corresponding to different modes of action. Furthermore, the shape of the density could also be relevant to insect predator-prey systems, in which later instars often enjoy greater motility/dispersal abilities. In this case, quicker development times can alter predator-prey dynamics. Thus one might desire to investigate how the model output is affected by changes in the shape of this parameter function. We shall discuss this briefly here, referring the interested reader to [20] for a complete derivation.

Using the representation (3)–(4) for (1), and noting the explicit dependence of $L = L(t, x, x_t, m)$ on the probability density m , one can readily derive the delay equation for the sensitivity $y(t) = \frac{\partial x}{\partial m}(t, m; q - m)$ of x with respect to m , which is defined as the directional derivative of x with respect to m in the direction $q - m$, where q is another probability distribution with the same properties as m . We find that y satisfies

$$\begin{aligned} \dot{y}(t) &= \frac{\partial L}{\partial x}(t, x, x_t, m)y(t) + \frac{\partial L}{\partial x_t}(t, x, x_t, m)y_t + \frac{\partial L}{\partial m}(t, x, x_t, m; q - m) \\ y(\tau) &= 0, \quad \tau \in [-7, 0], \end{aligned} \quad (25)$$

where L is defined in (4). The function y_t is the delay sensitivity function and is defined in

the same way as the delay function x_t is defined in (2). The quantities $\partial L/\partial x$, $\partial L/\partial x_t$, and $\partial L/\partial m$ are the Frechet derivatives of L with respect to x and x_t , and finally the directional derivative of L with respect to m in the direction $q - m$, respectively.

In [20], rigorous arguments are given to establish that the corresponding delay differential equation for the sensitivity y is given by

$$\begin{aligned} \dot{y}(t) &= \begin{bmatrix} -d_A(t) - p_A(t) & 0 \\ b(t) & -d_N(t) - p_N(t) \end{bmatrix} y(t) + \begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix} \int_{-7}^0 y_t(\tau) m(\tau) d\tau \\ &+ \begin{bmatrix} 0 & 1 \\ 0 & -1 \end{bmatrix} \int_{-7}^0 x_t(\tau) [q(\tau) - m(\tau)] d\tau, \\ y(0) &= (0, 0)^T. \end{aligned} \tag{26}$$

It [20] it is also shown how one approximates this equation in a manner similar to that of the approximations (8)–(10) for (1). This will be useful for example if one prefers to consider the estimation problems in a function space setting before complete discretization.

7 Concluding Remarks

We have demonstrated the ready use of inverse problem and uncertainty analysis methodologies associated with delay systems such as that formulated here. This was done with simulated “noisy data”. Future work on pest control methodologies with insecticides will depend critically upon obtaining new data that is collected through experiments designed to support and revise the model. We anticipate the collection of data in both the laboratory and the field, keeping track of mortality and reproduction for not only the first generation of insects, but also their offspring. This will permit use of inverse problem techniques to find parameters for the delay models that provide the best fit to this data, producing models that will effectively capture the lethal as well as sublethal aspects of population dynamics. Sensitivity analysis on the density kernel m will be instrumental in determining the appropriate form of the probability density. Once model verification and validation are completed, we would also hope to determine model based optimal dosage levels of insecticides for several different environments.

Acknowledgements

This research was supported in part by the U.S. Air Force Office of Scientific Research under grant AFOSR-FA9550-08-1-0147 and in part by the National Institute of Allergy and Infectious Disease under grant NIAID 9R01AI071915-05. The first two authors are grateful to the Helen Riaboff Whiteley Center at Friday Harbor Laboratories on San Juan Island, WA, for lodging and use of their facilities during preparation of this paper. Finally, the first author is thankful for the hospitality of the Radon Institute for Computational and Applied

Mathematics (RICAM) in Linz, Austria, during a visit in which writing of this manuscript was completed.

References

- [1] H. T. Banks, Approximation of nonlinear functional differential equation control systems, *J. Optim. Theory Applic.*, **29** (1979), 383–408.
- [2] H. T. Banks, Identification of nonlinear delay systems using spline methods, *Proc. Intl. Conf. on Nonlinear Phenomena in Math. Sciences* (V. Lakshmikantham, Ed.), Academic Press, New York, (1982), 47–55.
- [3] H. T. Banks, J. E. Banks, L. K. Dick and J. D. Stark, Estimation of dynamic rate parameters in insect populations undergoing sublethal exposure to pesticides, *Bulletin of Math Biology*, **69** (2007), 2139–2180.
- [4] H. T. Banks, J. E. Banks, S. L. Joyner and J. D. Stark, Dynamic models for insect mortality due to exposure to insecticides, *Mathematical and Computer Modelling*, **48** (2008), 316–332.
- [5] H. T. Banks, D. M. Bortz and S. E. Holte, Incorporation of variability into the mathematical modeling of viral delays in HIV infection dynamics, *Math. Biosciences*, **183** (2003), 63–91.
- [6] H. T. Banks, D. M. Bortz, G. A. Pinter and L. K. Potter, Modeling and imaging techniques with potential for application in bioterrorism, Chapter 6 in *Bioterrorism: Mathematical Modeling Applications in Homeland Security*, H.T. Banks and C. Castillo-Chavez, eds., Frontiers in Applied Mathematics, SIAM, Philadelphia, 2003, pp. 129–154.
- [7] H. T. Banks, M. Davidian, J. R. Samuels Jr., and K. L. Sutton, An inverse problem statistical methodology summary, Center for Research in Scientific Computation Technical Report, CRSC-TR08-01, NCSU, January, 2008; in *Mathematical and Statistical Estimation Approaches in Epidemiology*, (G. Chowell, et. al., eds.), Springer, New York, to appear.
- [8] H. T. Banks and J. L. Davis, A comparison of approximation methods for the estimation of probability distributions on parameters, CRSC-TR05-38, NCSU, October, 2005; *Applied Numerical Mathematics*, **57** (2007), 753–777.
- [9] H. T. Banks and J. L. Davis, Quantifying uncertainty in the estimation of probability distributions with confidence bands, CRSC-TR07-21, NCSU, December, 2007; *Math. Biosci. Engr.*, **5** (2008), 647–667.

- [10] H. T. Banks, S. Dediu and H. K. Nguyen, Sensitivity of dynamical systems to parameters in a convex subset of a topological vector space, *Math. Biosci. and Engineering*, **4** (2007), 403–430.
- [11] H. T. Banks and F. Kappel, Spline approximations for functional differential equations, *J. Differential Equations*, **34** (1979), 496–522.
- [12] J. E. Banks, L. K. Dick, H. T. Banks and J. D. Stark, Time-varying vital rates in ecotoxicology: selective pesticides and aphid population dynamics, *Ecological Modelling*, **210** (2008), 155–160.
- [13] H. T. Banks and K. Kunisch, *Estimation Techniques for Distributed Parameter Systems*, Birkhauser, Boston, 1989.
- [14] H. T. Banks and I. G. Rosen, Spline approximations for linear nonautonomous delay systems, *J. Mathematical Analysis and Applications*, **96** (1983), 226–268.
- [15] C. C. Carter, T. N. Hunt, D. L. Kline, T. E. Reagan and W. P. Barney, Insect and related pests of field crops: pea aphid, Center for Integrated Pest Management, http://ipm.ncsu.edu/AG271/forages/pea_aphid.html, May, 1996.
- [16] M. Davidian and D. Giltinan, *Nonlinear Models for Repeated Measurement Data*, Chapman & Hall, London, 1998.
- [17] J. K. Hale, *Functional Differential Equations*, Springer-Verlag, New York, NY, 1971.
- [18] J. K. Hale, *Theory of Functional Differential Equations*, Springer-Verlag, New York, NY, 1977.
- [19] J. K. Hale and S. M. Verduyn Lunel, *Introduction to Functional Differential Equations*, Springer-Verlag, New York, NY, 1993.
- [20] S. L. Joyner, *Dynamic Models for Insect Mortality Due to Exposure to Insecticides*, M.S. Thesis, North Carolina State University, Raleigh, March, 2008.
- [21] M. A. Krasnoselskii and A. V. Pokrovskii, *Systems with Hysteresis*, Springer-Verlag, New York, NY, 1989.
- [22] I. D. Mayergoyz, *Mathematical Models of Hysteresis*, Springer-Verlag, New York, NY, 1991.
- [23] G. A. F. Seber and C. J. Wild, *Nonlinear Regression*, John Wiley & Sons, New York, NY, 1989.
- [24] J. D. Stark and U. Wennergren, Can population effects of pesticides be predicted from demographic toxicological studies? *J. Economic Entomology*, **88**, (1995), 1089–1096.

- [25] A. Visintin, *Differential Models of Hysteresis*, Springer-Verlag, New York, NY, 1994.
- [26] J. Warga, Functions of relaxed controls, *SIAM J. Control*, **5** (1967), 628–641.
- [27] J. Warga, *Optimal Control of Differential and Functional Equations*, Academic Press, New York, NY, 1972.