Estimation of Dynamic Rate Parameters in Insect Populations Undergoing Sublethal Exposure to Pesticides

H. T. Banks*, John E. Banks†, Lara K. Dick*, John D. Stark‡

October 22, 2006

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

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

‡Department of Entomology
Ecotoxicology Program
Washington State University
Puyallup, Washington 98371

Corresponding author:
H. T. Banks
Center for Research in Scientific Computation
Box 8205
North Carolina State University
Raleigh, North Carolina 27695-8205
htbanks@ncsu.edu
fax: 919-515-1636
Final manuscript submitted in LATEX

Abstract

With newer, more environmentally friendly and, subsequently less lethal, pesticides in use, evaluating efficacy of a pesticide now requires more than simply counting deaths after treatment. A discrete, age-structured matrix model that incorporates a species' life history traits (such as birth rate, death rate and fecundity) has previously been used by ecologists. This model will be presented and discussed along with an alternative continuous, age-structured model which offers significant advantage in considering sublethal damage. We use this continuous model to estimate time-dependent mortality parameters in an ordinary least-squares technique. Confidence intervals are given and results from tests for statistical significance of added parameters are presented.

Key Words: Population models; Leslie; Sinko-Streifer; McKendrick-VonFoerster; time-dependent parameters

1 Introduction

Motivated by environmental concerns, entomologists need to understand the effects of new classes of pesticides and insecticides that are less lethal than their predecessors. These new pesticides act as growth regulators, causing scientists to be concerned with survival, sublethal damage and recovery delays exhibited by species targeted by the pesticide. Determining the efficacy of a pesticide requires information about life history traits [B, C, Ca, CL, L, LW, SBV, SW, W, WS]. In this paper, we will describe and compare models that utilize these traits, such as birth rate, death rate and fecundity, to describe population growth and death.

Our comparison of models is in the context of aphid growth and mortality in populations subjected to varying levels of pesticide treatment. Aphids reproduce on day five or six of their life span, and since two day old aphids and ten day old aphids will exhibit different life traits, they provide an excellent subject for age-dependent effects. Thus, in order to understand how a pesticide affects growth and reproduction, one must know something about the individual's age class. We first employ Leslie age-structured matrix projection models, and discuss simulations of population growth, compare different exposure methods, report various results, and summarize some previously documented conclusions. As we shall see, these matrix models, while now widely accepted in the ecology community, are not adequate to treat sublethal damage that is prevalent in new classes of insecticides. We then turn to continuous age-structured models which to date are not widely used in the community. It would be optimistic to expect these models, which are more conceptually challenging, to be readily embraced by the community. As Caswell [Ca] has noted, "It took 25 years for ecologists to adopt matrix population models. This was partly because matrix algebra was perceived as an advanced and esoteric branch of mathematics." But in our efforts below, we believe we make a compelling case for the adoption of these models by researchers. We support our computational inverse problem efforts with statistical tests to offer additional evidence that the time-dependent rates obtained with the continuous models we propose are indeed a significant improvement over constant rates obtained with either the standard Leslie matrix models or the continuous models.

2 Discrete Age-Structured Models

2.1 Experimental Procedure and Data Description

In 1995, Stark and Wennergren [SW] argued that demographic toxicological analysis estimating the total effect of insecticides on populations was best completed using a stage-structured approach. They argued that demographic toxicology is an ecotoxicological technique that incorporates life table parameters in the context of toxicology. It is a process that allows one to compare life table parameters for unexposed populations with those for populations exposed to various concentrations of a pollutant. Moreover, they explained that because of stricter regulations placed on insecticides requiring them to be more environmentally friendly, recently developed insecticides are slower acting and produce a greater degree of *sublethal effects* than earlier ones. Hence, the use of life tables combining both lethal and sublethal effects is expected to be beneficial in evaluating efficacy of insecticides.

Specifically, Stark and Wennergren studied the effects of the neem insecticide, Margosan-O on the pea aphid, *Acyrthosiphon pisum*, by exposing both neonates and adults in separate sets of experiments. Twenty-one pots of broad bean, *Vicia faba*, were planted in rectangular pots 10 cm wide, 10 cm high, and grown in a greenhouse. When the plants reached around 25 cm in height, they were thinned to five per pot and treated with a water control or one of six concentrations of Margosan-O (10, 20, 40, 80, 100 ppm). Three pots of broad bean were treated for each concentration. The plants were treated completely on both the upper and under surfaces of all leaves and allowed to dry.

For the exposure of the neonates, one young apterous adult female aphid was placed in a clip cage attached to the leaves of the treated plants. There were ten clip cages per plant, thus ten aphids per pot and thirty per concentration. Because of the structure of the clip cages, the aphids were in constant contact with the surface of the leaves, and hence, chronically exposed. After 24 hours, all aphids were removed from the cages except one 1st instar. Thus these neonate aphids were exposed at birth to a specific treatment level and were then studied with mortality and reproduction recorded daily throughout their life-span.

Exposure of adult aphids was completed in a manner similar to that of the neonates. One young apterous adult female was placed in a clip cage attached to the leaves of untreated plants. There were ten clip cages per plant, for a total of 210 aphids. Again, all aphids except one 1st instar were removed from the cages. These aphids were kept in the cages until one day after they reached the adult stage. The aphid reproduces on day five or six of their life-span, thus they were kept in the cages until day seven. The adult aphids were relocated to a separate set of clip cages on broad bean plants that had been treated with the various concentrations of Margosan-O and the control. These adults were then studied with mortality and reproduction recorded daily. For a more detailed explanation of the experimental procedure see [SW].

Mortality and reproduction were recorded for exposure of both neonates and adults and thus the data collected is the population density at age, x, and time, t. Assuming we know the percent mortality and percent of loss of reproduction, we can calculate various

parameters (defined below) that are of interest. Stark and Wennergren used the intrinsic rate of increase (r_m) as a bioassay parameter because it is based on survivorship (L). These values can be found using the experimental data. All formulas pertaining to the Leslie model and life table parameters can be found in Table 1, and are discussed in more detail throughout the next sections.

Table 1: Formulas and descriptions for life table parameters

Parameter	**Physical Description**	**Formula**
x	Age Interval in Days	N/A
t	Number of Days Experiment was Run	N/A
$D(x)$	Initial Adult Population	N/A
$d(x)$	Initial Neonate Population	N/A
$N(t)$	Population Density	N/A
$mort$	Percent Mortality	N/A
$loss$	Percent Loss of Reproduction	N/A
$S(x)$	Surviving Number of Adults Alive	$\frac{D(x)(100-mort)}{100}$
$s(x)$	Surviving Number of Neonates Alive	$\frac{d(x)(100-loss)(100-mort)}{100^2}$
$L(x)$	Probability of Survival to Next Age Interval	$\frac{S(x)}{S(0)}$
$M(x)$	Number of Offspring Per Surviving Female	$\frac{s(x)}{S(x)}$
$G(x)$	Probability to Survive and Grow	$\frac{L(x+1)}{L(x)}$
$F(x)$	Fecundity	$\frac{M(x)+G(x)*M(x+1)}{2}$
R_0	Net Reproductive Rate	$\sum_x L(x)M(x)$
T	Mean Generation Time	$\frac{\sum_x xL(x)M(x)}{\sum_x L(x)M(x)}$
r_m	Intrinsic Rate of Increase	$\frac{lnR_0}{T}$
r_i	Instantaneous Rate of Increase	$\frac{ln\left(\frac{N(t)}{N(0)}\right)}{t}$
b	Intrinsic Birth Rate	$\frac{1}{\sum_x L(x)e^{-r_m x}}$
d	Intrinsic Death Rate	$b - r_m$

2.2 Model Description

An age-structured matrix projection model known as the Leslie matrix model [C, Ca] was used to investigate the population structure. The Leslie matrix model is a discrete model that uses age distributions to project population densities as time progresses. As discussed in [K], it was introduced by Whelpotn in 1936 and formalized by Bernardelli in 1941, Lewis in 1942, and Leslie [L] in 1945 and 1948. The basic structure of the Leslie matrix is given in

$$L = \begin{bmatrix} F_0 & F_1 & F_2 & \dots & F_{\omega-1} \\ G_0 & 0 & \dots & \dots & \dots \\ 0 & G_1 & 0 & \dots & \dots \\ \dots & \dots & \ddots & \ddots & \dots \\ 0 & 0 & 0 & G_{\omega-2} & 0 \end{bmatrix}. \tag{1}$$

Here F represents the fecundity, G represents the probability to survive and grow into the next stage, and ω represents the number of stages. Fecundity and survivorship formulas for the pea aphid experiments are given in Table 1. They are based on percent mortality and percent loss reproduction values which were determined through calculations with the experimental data.

If one knows the age distribution or population density at age x and time t, one can project to time $t+1$. Therefore, population growth can be found via repeated matrix multiplications as shown in equation (2). In vector matrix notation the model can be written as $n_{t+1} = Ln_t$, where $n_{j,t}$ is the population density at age j and time t, and L is the Leslie matrix; thus for ω stages we have

$$\begin{bmatrix} n_{0,t+1} \\ n_{1,t+1} \\ \vdots \\ n_{\omega-1,t+1} \end{bmatrix} = \begin{bmatrix} F_0 & F_1 & F_2 & \dots & F_{\omega-1} \\ G_0 & 0 & \dots & \dots & \dots \\ 0 & G_1 & 0 & \dots & \dots \\ \dots & \dots & \ddots & \ddots & \dots \\ 0 & 0 & 0 & G_{\omega-2} & 0 \end{bmatrix} \begin{bmatrix} n_{0,t} \\ n_{1,t} \\ \vdots \\ n_{\omega-1,t} \end{bmatrix}. \tag{2}$$

Using our previous results, we can solve for the net reproductive rate (R_0) and mean generation time (T) given in Table 1 that are then used to calculate two population-level ecotoxicological endpoints: intrinsic rate of increase (r_m) and direct instantaneous growth rate (r_i). Both measurements are of interest to ecologists. The rate (r_m) allows one to determine both the intrinsic birth and death rate, and the stable age distribution, all of which provide more insight into the population dynamics. The rate (r_i) is determined using the results from the Leslie simulation.

2.3 Simulations and Results

To provide a basis for comparison of our results, we used MATLAB to recompute all of the Leslie simulations previously carried out by one of the co-authors (J. Stark) of this paper. Experimental data for the control experiment (no pesticide treatment) with aphids exposed as neonates are given in Table 2. This data was used, along with an assumed percent mortality and percent loss reproduction equivalent to zero, to solve for the parameters in Table 1. Once survivorship and fecundity were determined, we were able to develop the Leslie matrix as used by Stark and Wennergren.

Table 2: Control data for neonate exposure

Age Class	**0**	**1**	**2**	**3**	**4**	**5**	**6**	**7**
Total Alive	30	30	30	30	30	30	30	30
Total Young	0	0	0	0	0	0	9	118
Age Class	**8**	**9**	**10**	**11**	**12**	**13**	**14**	**15**
Total Alive	30	30	30	30	30	30	29	28
Total Young	232	209	202	220	246	179	181	149
Age Class	**16**	**17**	**18**	**19**	**20**	**21**	**22**	**23**
Total Alive	28	28	28	28	28	28	27	27
Total Young	133	116	100	54	46	27	7	5
Age Class	**24**	**25**	**26**	**27**	**28**	**29**	**30**	**31**
Total Alive	26	26	26	20	13	9	5	2
Total Young	0	0	0	0	0	0	0	0

Similar experimental data for all six concentrations of aphids exposed both as neonates and as adults was used to develop a Leslie matrix for each case. Percent mortality and percent loss reproduction values, which determine survivorship and fecundity, for each concentration can be found in Table 3. From this table we see that the neonates were much more susceptible to higher concentrations of Margosan-O, than were the adults.

Table 3: Percent mortality and percent loss reproduction for neonate and adult exposure

Concentration	**Percent Mortality**		**Percent Loss Reproduction**	
	Neonate	Adult	Neonate	Adult
10ppm	7	17	59	23
20ppm	30	24	85	40
40ppm	70	52	97	50
60ppm	80	55	99	54
80ppm	95	55	100	65
100ppm	98	62	—	70

Survivorship graphs for the aphids in each age class, for both neonate and adult exposure, are plotted in Figure 1. Survivorship of aphids exposed to Margosan-O as neonates was reduced as concentration of Margosan-O increased. For an exposure of 80ppm, the maximum longevity of the neonates after exposure was five days compared with 32 days for the control. For aphids exposed as adults, there was a slight reduction of survivorship, but the effect was much less than in individuals exposed as neonates. For an exposure of

100ppm, the maximum longevity of the aphids was 24 days compared with 32 days from the control. These results are supported by [RS]: the aphid population growth is significantly inhibited in a concentration-dependent manner by the introduction of Margosan-O. It should also be noted that under different concentrations, the effect of the insecticide declines as older aphids are exposed. This is supported by our simulations and the results presented in [WS].

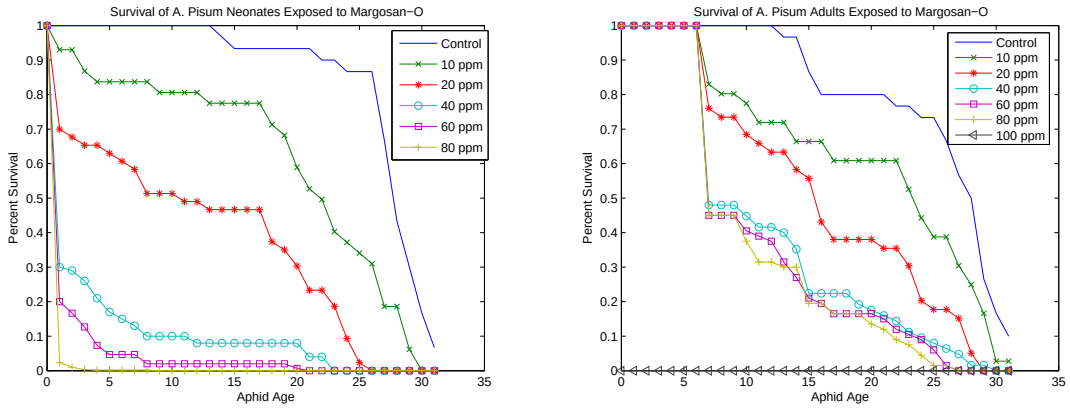

Figure 1: Survival of neonate exposure (left) and adult exposure (right)

The relationship between concentration and fecundity is similar to that of concentration and survivorship. Fecundity graphs for the aphids in each age class, both neonate and adult exposure, are given in Figure 2. Clearly, aphids exposed to Margosan-O as neonates undergo a huge decrease in their fecundity rates. For as little an exposure as 40ppm, the fecundity of the aphids is negligible. Maximum fecundity was around 80% for the control compared with a zero fecundity for pesticide concentrations higher than 40ppm. For aphids exposed as adults, while fecundity was reduced in a concentration dependent manner, the effect of the pesticide was not as prominent as for neonate exposure. Maximum fecundity was around 20% for an exposure of 100ppm compared with a maximum fecundity of around 90% for the control.

Once survivorship and fecundity were determined, simulations for each exposure method and concentration level were run beginning with day 0, and starting with 30 aphids in age class 0, 20 aphids in age class 1, 10 aphids in age class 2, 10 aphids in age class 3, 30 aphids in age class 4 and 0 aphids in the rest of the age classes. These aphids make up our initial age distribution vector that was then multiplied by the Leslie matrix to obtain our population projections. All simulations were run for 96 days.

For both exposure methods, the population growth for the control experiment appears to be exponential (for related graphs, see [BBDS]), which is what one would expect. However, once exposed to the pesticide, the population growth does not remain exponential. For neonate exposure, a concentration of 10ppm still allows exponential growth, a concentration

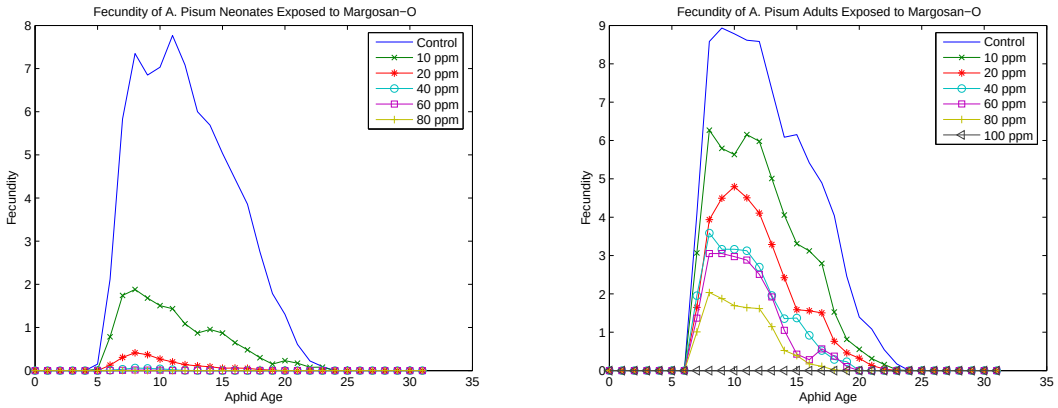

Figure 2: Fecundity of neonate exposure (left) and adult exposure (right)

of 20ppm exhibits an eventually linear growth, and all higher concentrations produce an exponential decline in population. The population level at day 96 for the control is 1.58×10^{18} compared with 0 for an exposure of 80ppm. For adult exposure, population growth remains exponential for all concentrations, however population maximums are reduced in a concentration dependent manner. The population level at day 96 for the control is 5.46×10^{17} compared with 4.82×10^{7} for an exposure of 80ppm.

Using the results from the simulation, we are able to calculate the net reproductive rate (R_0) and mean generation time (T) for all concentrations and both means of exposure (again, numerous graphs are given in [BBDS]). According to [SW], the net reproductive rate is the number of females that replace the average female over a generation. We found that R_0 decreases as a function of Margosan-O concentration. For neonate exposure, the number of females that replaced an average female over a generation was 74.4 for the control compared with 6.6×10^{-7} for 80ppm exposure. For adult exposure, the net reproductive rate was 81.9 for the control and 4.51 for 80ppm exposure. The reduction of R_0 was not as substantial with adult exposure, but the effect of the chemical can still be seen. Mean generation time (T) also declined with increased pesticide concentration. The more pesticide to which the aphids were exposed, the lower the mean generation time. As might be expected, the aphids exposed from birth were affected more than the aphids exposed as adults for all pesticide concentrations.

Once we obtained the values of R_0 and T, we were able to determine the intrinsic rate of increase (r_m) for each exposure case. As described in [W, WS], the intrinsic rate of increase is a measure of the ability of a population to increase exponentially in an unlimited environment. Thus r_m is taken as an important measurement of population dynamics. This rate is defined by $r_m = \frac{\ln R_0}{T}$. For neonate exposure, r_m decreased in a concentration dependent manner, becoming negative at 40ppm. A negative r_m implies population extinction. Interestingly, for adult exposure, r_m does not appear to be significantly affected by any amount

of pesticide (see [BBDS]).

Ecologists have studied the differences between instantaneous rate of increase (r_i) and the intrinsic rate of increase (r_m). As explained in [WS], the instantaneous rate of increase is a measure of the ability of a population to increase exponentially over time. This rate is given by $r_i = r_i(t) = \frac{1}{t}\ln\left(\frac{N(t)}{N(0)}\right)$. We see that there is a r_i value for each day the simulation is run. Since we need a single value $\hat{r_i}$ of r_i to compare to r_m, we obtain an $\hat{r_i}$ value (Table 4) by finding the value equivalent to the population growth rate once the stable age distribution has been reached, which is just determining the asymptote (Figure 8 of [BBDS]). As with the r_m values, a negative $\hat{r_i}$ implies population extinction. Although an asymptote is not reached for neonate exposure of 80 ppm, we are not surprised by these results.

The authors of [WS] suggested that $\hat{r_i}$ could be used as a substitute for r_m as a population endpoint measurement. They reported no significant difference ($p = 95\%$) between the two values. A p-value of 95% means (using conventional terminology) one is 95% confident that there is no difference between r_m and $\hat{r_i}$. Our calculated $\hat{r_i}$ values are listed with the corresponding r_m values in Table 4. As mentioned above, our $\hat{r_i}$ values should closely approximate our r_m values; however, this does not always seem to be the case. One-way analysis of variance (ANOVA) was not performed on this data set, so we are not certain of the significance of the difference between r_m and $\hat{r_i}$. Nevertheless, ecologists would like to use the $\hat{r_i}$ values instead of the r_m values because r_m is much more difficult to calculate.

Table 4: Intrinsic versus instantaneous rate of increase

Concentration	**Neonate Exposure**		**Adult Exposure**	
	$\hat{r_i}$	r_m	$\hat{r_i}$	r_m
Control	.393	.321	.381	.321
10ppm	.229	.199	.322	.280
20ppm	.016	.012	.274	.247
40ppm	-.328	-.373	.208	.193
60ppm	-.627	-.717	.186	.175
80ppm	-1.768	-1.778	.138	.131
100ppm	—	—	.098	.092

With an r_m value, or theoretically an $\hat{r_i}$ value, one can calculate the final parameters of interest. Birth rate decreased as concentration of pesticide increased for both neonate and adult exposure, but not surprisingly, the decrease was much more significant for the neonate. Death rate increased as concentration of pesticide increased for neonate exposure, while, an increase in pesticide concentration did not appear to significantly affect the death rate for adult exposure.

Figures and tables containing computed values of the previously mentioned calculated

parameters can be found in [BBDS].

2.4 Time Dependence

There is reason to believe that dealing with sublethal damage due to insecticide exposure motivates a need for age/time dependent parameters such as mortality, birth and fecundity rates. Indeed, it does not seem reasonable to assume that aphids will exhibit the same death rate one day after the spray and twenty days after the spray. In order to obtain time or age dependent parameters from the Leslie model, one must consider some of the input parameters in Table 1 to be time/age dependent. Since we are concerned with the death rate, we will consider percent mortality, $mort(x)$ to be age dependent.

The resulting formulation of the Leslie matrix detailed in Section 2.4 of [BBDS] is much more complicated than the formulation obtained when one considers all of the parameters to be constant. According to Wood [Wo2], "Although it is relatively easy to collect population data structured by life-history stage, the corresponding birth and death rates are all but impossible to measure directly." We also find that obtaining the sublethal parameters of interest is a daunting task when using the Leslie model. In particular, using parameter estimation techniques as we do in subsequent sections with a continuous model to estimate time-dependent rates would be extremely difficult if not impossible with the modified Leslie formulation.

In order to run simulations using the time-dependent rate formulation, one would in essence have to compute multiple different Leslie simulations. Wood [Wo3] comments on this idea of a time dependent parameter Leslie model, noting "while representing an advance on previous methods, the technique has several problems: time-dependent rates are only obtained by considering data in small sections and performing an analysis for each one." He continues on to explain why such an approach is not a good idea, observing that "there are fundamental demographic instabilities associated with estimating time and stage-dependent mortality rates from stage-structured data. The most important of these instabilities is that, even with perfect data, any small error in estimated mortality rates in one stage causes increasingly large errors in the mortality estimates for subsequent stages." Wood remarks that the only way to get past this instability is to "make restrictive assumptions about underlying population dynamics," such as assuming one has constant parameters. We turn in the next section to an alternative approach employing a partial differential equation (PDE) population model which readily allows for time/age-dependent parameters.

In his book [Ca], Caswell states that although they can be more difficult to work with, PDE's can describe the vital rates in detail, because they neither lump individuals into discrete classes nor divide time into intervals. We use the Sinko-Streifer or McKendrick-VonForester continuous age-structured population model. Ohman and Wood [OW] praise this model, remarking that "this method has been shown to be robust to moderate levels of sampling error and to avoid the demographic instabilities common to some other methods as well as to generate both time- and stage-dependent mortalities with confidence limits."

3 Continuous Age-Structured Model

In this section, we present an alternative approach to obtaining information on sublethal damage from insecticide treatments. In doing so, we wish to extrapolate information from the Leslie models and incorporate it into a continuous age-structured model. McKendrick and VonForester (see [K] for a historical summary) worked with age structured models [M, VF] and were the first (1959) to develop the well-known partial differential equation, continuous-time model, which was later (1967) popularized by Sinko and Streifer [SS]. The Sinko-Streifer (SS) model made allowances for size/age dependent growth rates. In this seminal effort, Sinko and Streifer worked with size-structured populations to describe population dynamics.

The Sinko-Streifer population model involves growth and mortality with size and/or age dependent parameters. We use the age-structured population model in an attempt to describe the population dynamics of aphids exposed to different concentrations of Margosan-O. Under specific assumptions on recruitment, we can derive the exact solution of the model using the method of characteristics, considering the population structure as a function of age (x) and time (t). We then utilize information gathered from Leslie to run simulations.

3.0.1 Model Assumptions

Before discussing the formulation of the specific model for our investigation , we will discuss the underlying assumptions for a general class of age/size structured models as presented in [BT]. First, the age growth rates of same age (x) individuals are assumed the same and are denoted by $g(x,t)$. Thus

$$\frac{dx}{dt} = g(x,t).$$

Second, we assume that there is a "minimum" and a "maximum" age, or equivalently

$$x_0 \leq x \leq x_{max}.$$

Third, the death rate (denoted by μ) is assumed to be proportional to the age density, while the recruitment rate (R) is assumed to be cumulatively proportional to the population age densities. Finally, we assume that the population is sufficiently large so as to allow us to use a continuous model to approximate the population densities.

3.0.2 Sinko-Streifer Model

With these assumptions in mind, we write the model

$$\frac{\partial u}{\partial t} + \frac{\partial}{\partial x} gu = -\mu u \text{ for } x \in (x_0, x_{max}) \text{ and } t \in (0, t_f), \tag{3}$$

with a boundary condition

$$R(t) = g(x_0,t)u(x_0,t) = \int_{x_0}^{x_{max}} k(\xi,t)u(\xi,t)d\xi \tag{4}$$

and an initial condition

$$u(x,0) = \Phi(x).$$

Here $u(x,t)$ represents the population density at age x and time t, $g(x,t)$ represents the individual growth or aging rate, $\Phi(x)$ represents the initial age density, and k represents the fecundity rate. We also have x_{max} equal to the maximum age, while x_0 is the initial age. As stated in the assumptions, R is the recruitment rate and μ is the death rate.

3.0.3 Solution Via the Method of Characteristics

Under certain conditions on the model (3)–(4), one can use classical techniques to reduce the partial differential equation to a family of ordinary differential equations (ODEs). We do this using the method of characteristics discussed in [BT] under the assumption that the recruitment R in (4) is known. We note that for our specific problem, the growth or aging rate is given by $g(x,t) = 1$. If an aphid in age class x is alive at time t, she will always move into the next age class $x+1$ at time $t+1$. We can readily compute the exact solution $u(x,t)$ of (3) for the case of constant growth rate ($g(x,t) = 1$) and death rate ($\mu(x,t) = \mu_0$). The general form of the solution has two parts, one initial condition driven and the other boundary condition driven. Since $g(x,t)$ is constant, these conditions depend solely on the relationship between age (x) and time (t). The initial condition driven solution is found when $t \leq G(x)$ and is given by

$$u(x,t) = \Phi(X(0;x,t))e^{-\int_0^t [g_x(\xi,X(\xi;x,t))+\mu(\xi,X(\xi;x,t))]d\xi}, \tag{5}$$

while the boundary condition driven solution holds when $t > G(x)$ and is given by

$$u(x,t) = \frac{R(T(x_0;x,t))}{g(T(x_0;x,t),x_0)} e^{-\int_{T(x_0;x,t)}^t [g_x(\xi,X(\xi;x,t))+\mu(\xi,X(\xi;x,t))]d\xi}. \tag{6}$$

Here $G(x)$ is defined as $G(x) = T(x;0,x_0)$, where $(x,G(x))$ is the characteristic curve (i.e., the solution of $\frac{dx}{dt} = g$) that passes through $(x_0,0)$. We must find the solution for x and t along characteristic curves: $X(0;x,t)$ is the characteristic curve evaluated at $t = 0$ that passes through the point (x,t), and $T(x_0;x,t)$ is the characteristic curve evaluated at $x = x_0$ that passes through the point (x,t). Figure 3 depicts a graphical schematic of the generic solution domain.

In order to compute the exact solution for our special case of $g = 1$ and μ constant, we begin by computing: $X(t;\hat{x},\hat{t})$, $T(t;\hat{x},\hat{t})$ and $G(x) = T(x;x_0,0)$. We first consider $X(t;\hat{x},\hat{t})$ noting that $\frac{dx}{dt} = 1$ and $x(\hat{t}) = \hat{x}$. Thus $x(t) = t + c \Rightarrow c = \hat{x} - \hat{t}$. Using the previous results we have

$$X(t;\hat{x},\hat{t}) = t + \hat{x} - \hat{t}.$$

Next solving for T, we find

$$T(t;\hat{x},\hat{t}) = x - \hat{x} + \hat{t}.$$

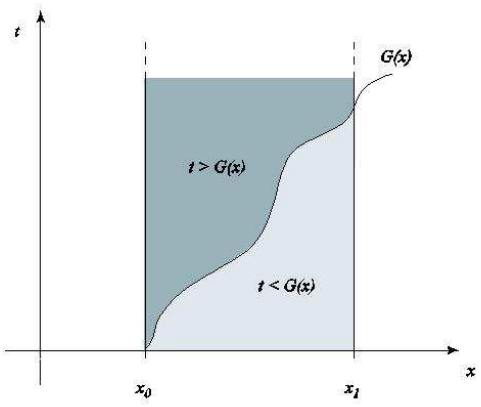

Figure 3: Characteristics

Finally, we solve for G by substitution:

$$G(x) = T(x; x_0, 0) = x - x_0.$$

Completing our special solution, we now substitute $X(t; \hat{x}, \hat{t})$, $T(t; \hat{x}, \hat{t})$ and $G(x) = T(x; x_0, 0)$ into equations (5) and (6). We obtain the *initial condition driven solution* (i.e., $t \leq G(x)$)

$$\begin{aligned} u(x,t) &= \Phi(X(0;x,t))e^{-\int_0^t [g_x(\xi,X(\xi;x,t))+\mu(\xi,X(\xi;x,t))]d\xi} \\ &= \Phi(x-t)e^{-\int_0^t (0+\mu_0)d\xi} \\ &= \Phi(x-t)e^{-\mu_0\xi|_0^t} \\ &= \Phi(x-t)e^{-\mu_0 t}, \end{aligned}$$

and the *boundary condition driven solution* ($t > G(x)$)

$$\begin{aligned} u(x,t) &= \frac{R(T(x_0;x,t))}{g(T(x_0;x,t),x_0)} e^{-\int_{T(x_0;x,t)}^t [g_x(\xi,X(\xi;x,t))+\mu(\xi,X(\xi;x,t))]d\xi} \\ &= R(x_0-x+t)e^{-\int_{x_0-x+t}^t (0+\mu_0)d\xi} \\ &= R(x_0-x+t)e^{-\mu_0\xi|_{x_0-x+t}^t} \\ &= R(x_0-x+t)e^{(x_0-x)\mu_0}. \end{aligned}$$

Thus under the special assumptions of $g = 1$ and constant mortality μ_0, we obtain an exact analytic solution using the method of characteristics. Note that this requires that the recruitment rate R be known or at least be represented by some known (observable) time dependent function $R(t)$.

3.1 Simulations and Results

In order to adapt the SS model to our specific problem, we have to first determine parameter values. We determined intrinsic death rates for each concentration and exposure type when running the Leslie simulations. The actual values that are substituted into our solutions are located in Table 5. Our age intervals range from age $x_0 = 1$ to age $x_{max} = 32$ and are measured in increments corresponding to days. (Note: In Section 2, we used Leslie with a zero age class (i.e., $x_0 = 0$). Here with SS we take $x_0 = 1$. This in no way affects our ability to compare results between the two models.) The simulations can be run for as many days as desired, but all simulations begin with day $t_0 = 0$ to facilitate comparison with the Leslie simulations. There were technical reasons for starting Leslie with a zero class pertaining to the manner in which our experimental data was given. With all of the previous parameters as known values, we are only left with determining our initial age distribution function and our recruitment function.

Table 5: Intrinsic death rates for neonate and adult exposures

Concentration	**Neonate Death Rate μ_0**	**Adult Death Rate μ_0**
Control	.0576	.0581
10ppm	.0516	.0540
20ppm	.0810	.0509
40ppm	.37388	.0597
60ppm	.7174	.0619
80ppm	1.7778	.0757

The initial age distribution function was determined using the initial data from the Leslie simulations. These values: 30 aphids in age class 0 (i.e., $x_0 = 1$), 20 aphids in age class 1, 10 aphids in age class 2, 10 aphids in age class 3, 30 aphids in age class 4 and 0 aphids in the rest of the age classes were placed into a vector and then linear interpolation was used to obtain our initial age distribution function Φ.

Linear interpolation uses piecewise linear splines to represent a linear polynomial between two adjacent data points, thereby producing a piecewise smooth, continuous curve for discrete samples or data. The representation used to determine the initial age distribution function $\Phi(x)$ is given by

$$\Phi(x) = \sum_{j=1}^{6} \Phi_j l_j(x),$$

with Φ_j representing the discrete values gathered from Leslie, and $l_j(x)$ representing the linear spline basis elements known as 'hat' functions. The value of $\Phi(x)$ for $x > 6$ was taken as zero since no individuals greater than six days old were in the initial population.

This same interpolation method was employed to determine the recruitment functions $R(t)$ for each exposure method and each concentration level. The first entry in each Leslie simulation represents the number of individuals being born into age class $x_0 = 1$ at each time increment, thus these entries can be thought of as the recruitment rate for the specific data set. The values from the first entry in each Leslie simulation were obtained, placed into a recruitment vector, and then linear interpolation was used to calculate the recruitment functions. Examples of the resulting functions are plotted in Figures 4 through 6.

Since we now have the necessary information to solve SS, we are able to run simulations for each exposure method and concentration level for a total of twelve SS simulations. Our SS simulation model for the case of $g = 1$ and constant $\mu = \mu_0$ is given by

$$u(x,t) = \Phi(x-t)e^{-\mu_0 t} \text{ for } t \leq G(x) \quad (7)$$

$$u(x,t) = R(x_0 - x + t)e^{(x_0-x)\mu_0} \text{ for } t > G(x). \quad (8)$$

3.1.1 Neonate Exposure

Results from neonate exposure for the control and 20ppm concentrations are given in Figure 4. We note that for the control the maximum population level essentially corresponds with the maximum population level determined with the Leslie simulations (all comparable results with Leslie simulations are depicted in the graphs of Figure 4 of [BBDS]). For 20ppm, as in the controls (upper right of Figure 4), we observe that the population level increase is not exponential in time. With Leslie for 20ppm (Figure 4 of [BBDS]), population growth was linear. We see that the SS simulation eventually produces linear growth in time at all age levels, but begins with some fluctuations (lower right of Figure 4). These fluctuations were not seen in the Leslie simulations. We do however, still discover that the concentration increase has greatly affected the maximum population level.

With a neonate exposure of 40ppm (Figure 5), we see a low recruitment rate and a low maximum population growth rate. By time $t = 30$, the population completely disappears. These results are similar to the Leslie results. The main difference between the two models, is that since the SS model is continuous, the population level does not remain at 100 individuals for all of day 1. After a few time increments, the aphids have begun to drastically die out. In comparing the recruitment rate with the population structure, we see that there are a few individuals that survive the initial dose and age. The ones that survive, age a little, but eventually die out.

With the results for neonate exposure of 80ppm (Figure 5) we have hardly any recruitment and a very small population growth. The effect of Margosan-O is substantial. Although these results are comparable with the Leslie results, the Leslie model is not readily able to determine the small levels of aging. Since we are working with a continuous model when using SS, we are able to determine population dynamics not attainable with the Leslie models.

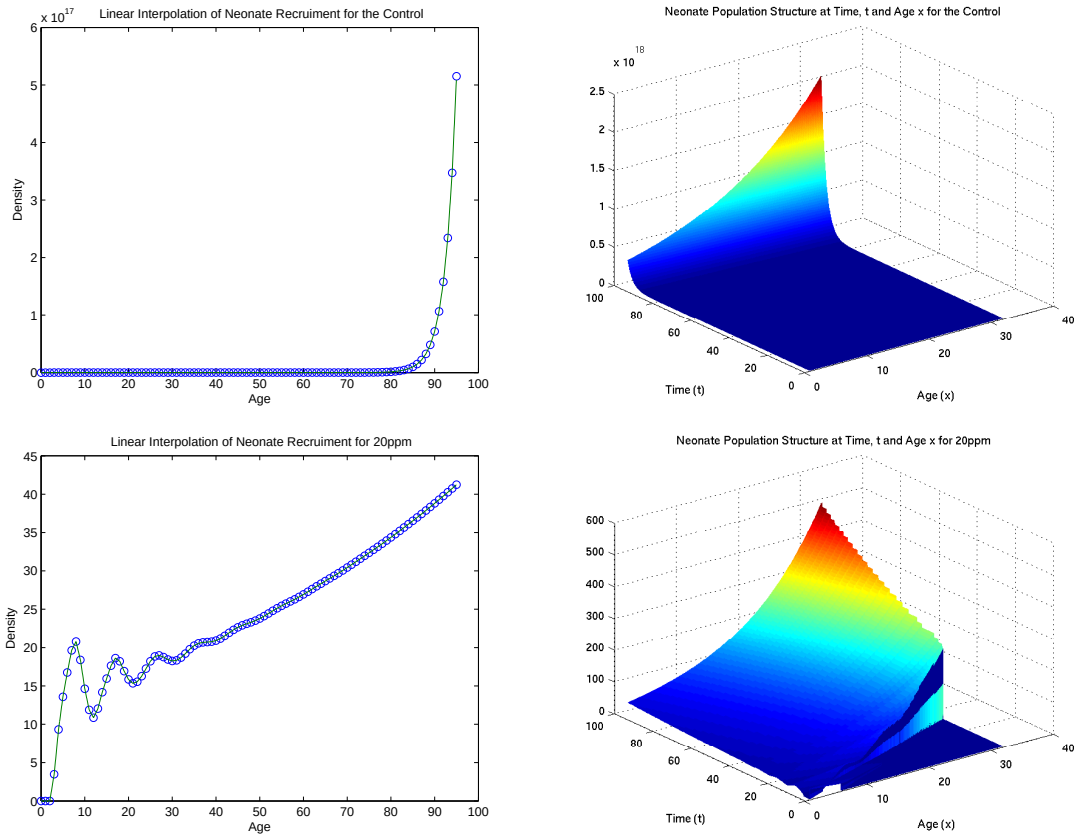

Figure 4: Neonate recruitment and population structure for the control and 20ppm

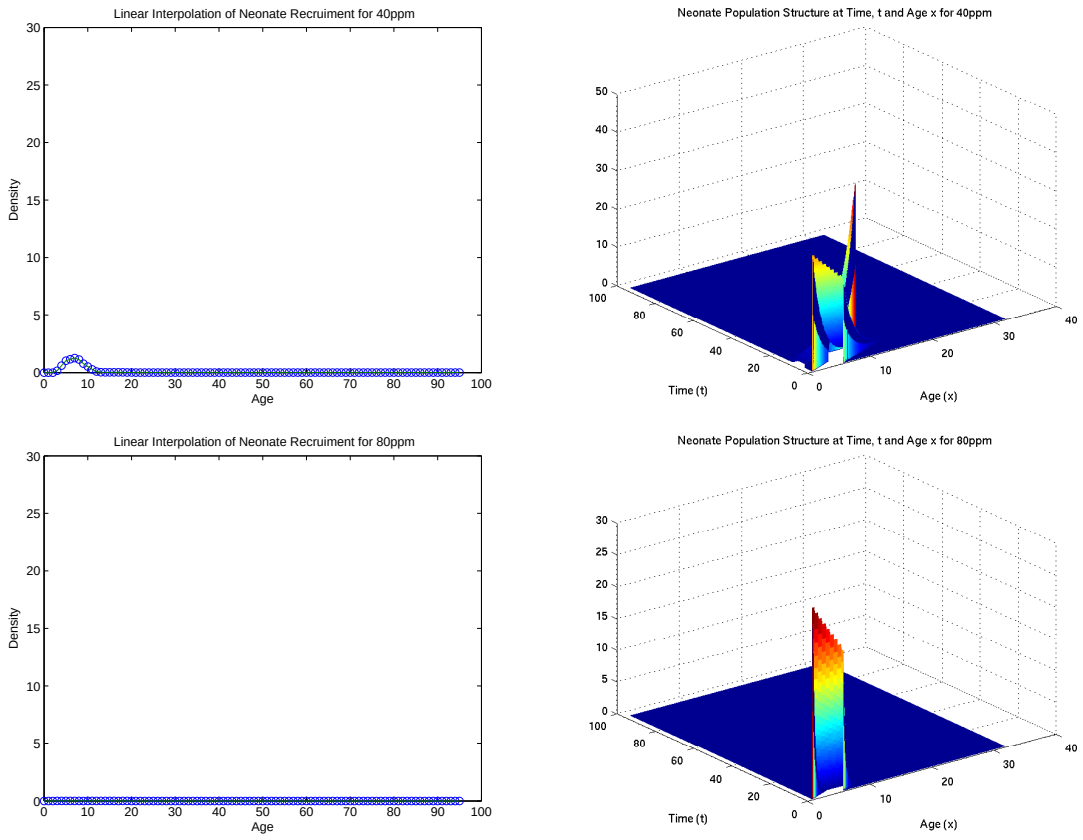

Figure 5: Neonate recruitment and population structure for 40ppm and 80ppm

3.1.2 Adult Exposure

We next summarize the Sinko-Streifer adult exposure simulations. Results from adult exposure for the control and 80ppm concentrations are graphed in Figure 6.

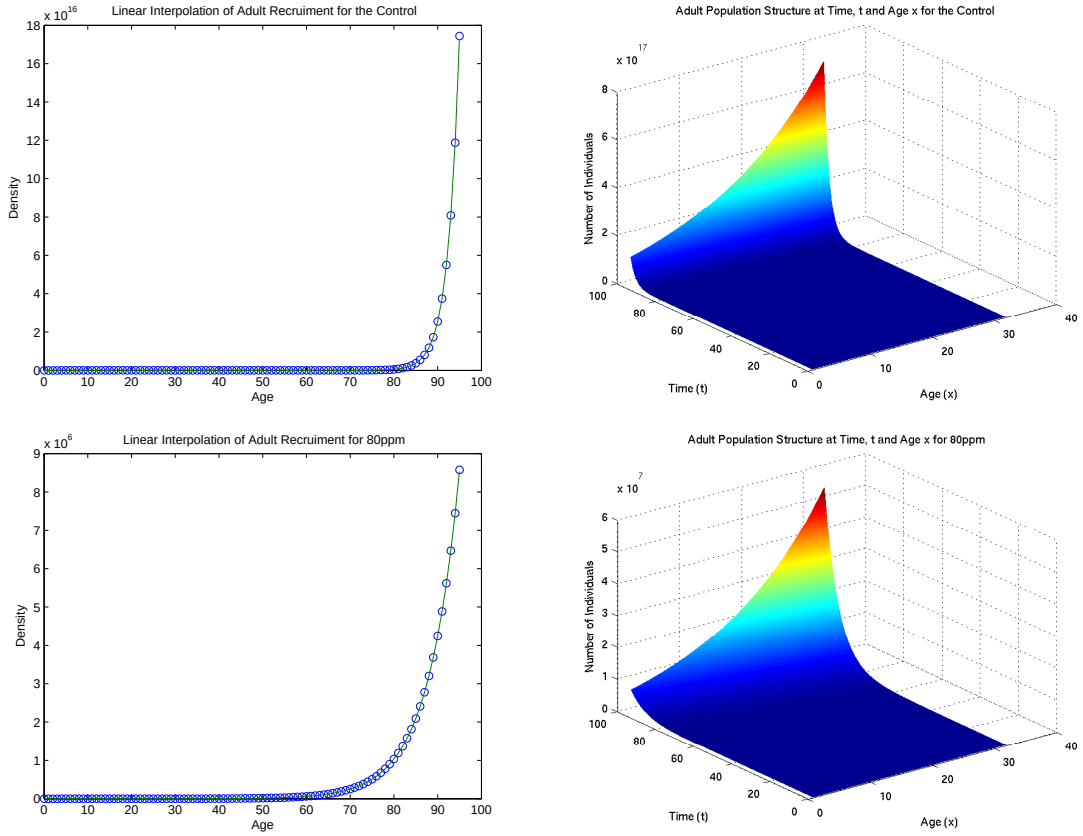

Figure 6: Adult recruitment and population structure for the control and 80 ppm

We first remark that the maximum population level essentially agrees with the maximum population level determined with the Leslie simulations. As we observed with the neonate exposure, these results lead us to believe that the SS model is producing reliable results from which conclusions can be drawn. There does not appear to be a substantial difference between the SS and the Leslie simulation results in these cases. As with Leslie, we notice that an increase in the concentration level produces a decrease in the maximum population levels, but not sufficiently to counteract exponential growth.

The SS simulations for adult exposure appear to compare very closely to the Leslie simulations in the case of adults. On the other hand, the SS model is a better model when dealing with neonate exposure since sublethal effects can be seen with the continuous model. For example, with 20ppm neonate exposure (Figure 4) we are able to observe fluctuations

of population growth within the age classes. It is not possible to discern these sublethal effects using the Leslie model. Moreover, the 80ppm exposure simulation reveals that the population actually recovers from the initial dosage. This recovery does not last long, but since Leslie is a discrete model, this potential recovery is not captured in our Leslie simulations. These observations suggest that SS is a potentially more useful model.

3.1.3 Comments

Since we have used results from the Leslie formulation to determine our parameters for Sinko Streifer, the values we obtain using the formulas from Table 1 are identical to those we calculated using Leslie. The only difference is the value of r_i, the instantaneous rate of increase. We calculate r_i using the population totals for each day. Since we are dealing with a continuous model, r_i has not been calculated. It may be possible to extrapolate the formula used to determine r_i in the Leslie model to our SS model, but this has not been done. If we use each time increment as a "day" and compute r_i using population totals from each time increment, we find that these r_i values are much lower than they were with the Leslie simulations. They no longer are comparable to r_m, the intrinsic rate of increase.

4 Parameter Estimation

4.1 Introduction

Inverse problems in science and engineering are quite simply problems where the solution is known, but not necessarily the underlying dynamics. Applied mathematicians often "work backward" on problems with which they have experimental data, but not a mathematical model with appropriate parameter values to describe the data. Their goal is to develop a model that best represents the experimental data while containing some measure of fidelity with the physical or biological phenomenon under investigation. Based on our early comparisons, we suspect that the SS model may provide us with a better representation of our data than did the Leslie model. Since we have experimental data with which to compare results, we performed an inverse problem to estimate parameters in the SS model. We used the resulting estimated parameter values to carry out simulations that we then compared to our experimental data. Throughout the presentation below, we discuss ramifications of using parameter estimates from Leslie in SS and also discuss limitations of the data with which we are working.

4.2 Estimation of the Death Rate

Recall the Sinko-Streifer model from equation (3). Both the initial condition driven solution (equation (5)) and the boundary condition driven solution (equation (6)) depend on values of μ, the death rate. Death rates computed from our Leslie results (see Table 1) were used in the SS model to carry out all SS simulations reported on in Section 3.1. Since our first goal is to compare the ability of the Leslie model to describe the experimental data with

that of the SS model, we would like to estimate all parameters in SS independent of those in Leslie. In order to obtain a death rate specifically for SS, we estimate μ using an ordinary least squares (OLS) inverse problem technique. We seek value(s) of μ that will provide us with an accurate estimate of the population dynamics.

To estimate parameters with the OLS technique, we minimize the criterion

$$J(\mu) = \sum_{i=1}^{N_x} \sum_{j=1}^{N_t} \left| U_{SS}(x_i, t_j; \mu) - U_{i,j}^d \right|^2, \tag{9}$$

where $U_{SS}(x_i, t_j; \mu)$ is the solution obtained by the method of characteristics and evaluated at age x_i and time t_j for a given value of μ, and $U_{i,j}^d$ is our experimental data. In order to compute the value of μ_0 which minimizes the above cost function, we used the MATLAB routine *fminsearch*, which is a Nelder-Mead direct search method. Direct search methods [Ke] use values of J taken from a set of parameter sample points to continue the sampling. Direct search methods do not require approximate gradient information.

Before we can discuss results, we must discuss the experimental data with which we are working. Recall the experimental procedure detailed in Section 2.1. The experiment was designed to track one generation of aphids, namely the generation sprayed with pesticide. Table 2 contains an example of the type of data we have for all exposure levels for both neonates and adults. We know the amount being born into the population each day, but do not subsequently follow the lives of these newborns. Thus, when we convert our experimental data into matrix form, we obtain information in the following form:

$$\begin{array}{c} \\ \text{Age} \end{array} \begin{array}{c} \text{Time} \\ \begin{bmatrix} \star & * & * & \dots & * \\ 0 & \star & 0 & \dots & 0 \\ 0 & 0 & \star & 0 \dots & 0 \\ \vdots & 0 & \dots & \ddots & 0 \\ 0 & 0 & \dots & \dots & \star \end{bmatrix} \end{array} \tag{10}$$

with $\star$ representing the generation initially sprayed with pesticide and $*$ representing the individuals born into the population each day, which for our SS model represents our recruitment rate. On the other hand, our SS simulations result in a "full" matrix, i.e., values of $u(x,t)$ for all (x,t). Ideally, we would like to have data to compare with the full simulation matrix. However, this is not the case. We only have data along one characteristic (i.e, along $(x,t) = (x, x - x_0)$) with which to perform the inverse problem, namely the main diagonal that follows the initial population as it ages. Our results are based on measurements along only one characteristic to estimate the optimal death rate μ_0. We discuss these and other implications in the following sections.

4.3 Estimating Constant Death Rates

The results of the OLS problem for estimating μ for neonate exposure are summarized in Table 6 and in Figure 7. We note that all graphs for the inverse problem results represent

calculations starting at $t_0 = 1$, corresponding to the data point x_0 for the population value at t_0 in the criterion from (9). We see that the estimated death rates are considerably lower than those calculated with the Leslie model. One of the purposes of performing this inverse problem is to give us a means with which to compare results from the Leslie and SS simulations. The values of the estimated death rates increase in a treatment concentration dependent manner similar to the increases seen with the Leslie model. However, the actual values are much smaller.

Table 6: OLS results of estimation of μ for neonate exposure

Concentration	**μ from Leslie**	**OLS/SS μ**	**Estimated Cost**
Control	.0576	.0128	1,110.2
10ppm	.0516	.0277	571.5
20ppm	.0810	.0396	297.6
40ppm	.3739	.0951	114.6
60ppm	.7174	.2034	122.6
80ppm	1.7780	.7936	3.3

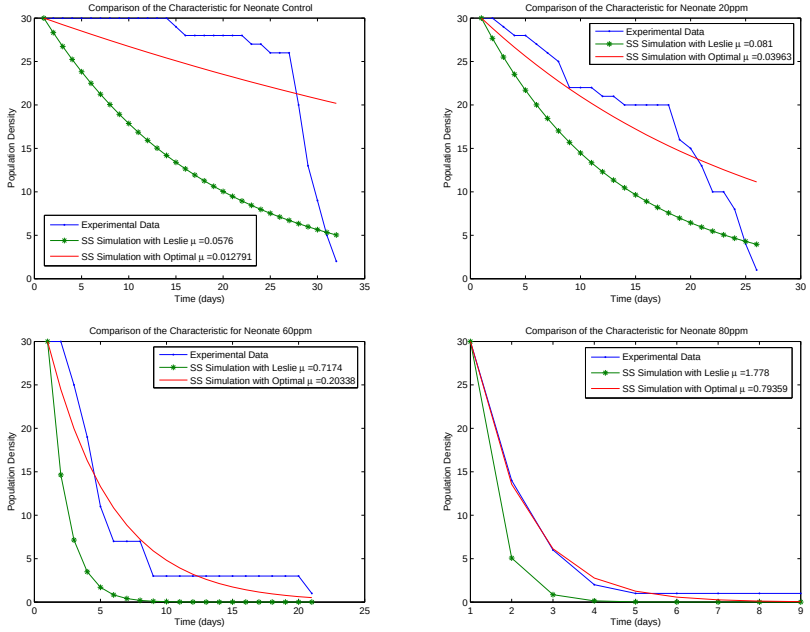

Figure 7: Comparisons of experimental data to SS simulations using Leslie μ and the optimal OLS μ for neonate exposure

In studying the graphs in Figure 7, we see that the SS model with the optimal μ values appears to better fit the data than the SS model with the μ obtained from Leslie. This is true for all exposure levels. Thus, in determining a death rate independent of Leslie, we obtain a death rate that better predicts our experimental data.

The results of the inverse problem for estimating μ for adult exposure are given in Table 7 and in Figure 8. These results are similar to those we observed for neonate exposure. Again, the values appear to be exposure treatment concentration dependent, but on a much smaller level than the death rates obtained from the Leslie model. Moreover, Figure 8 contains results similar to those for neonate exposure. The optimal μ appears to be a better estimation of death rate than the death rate corresponding to the Leslie formulation.

Table 7: OLS results of estimation of μ for adult exposure

Concentration	**μ from Leslie**	**OLS/SS μ**	**Estimated Cost**
Control	.0581	.0173	766.3
10ppm	.0540	.0251	768.3
20ppm	.0509	.0337	587.4
40ppm	.0597	.0420	919.0
60ppm	.0619	.0417	737.0
80ppm	.0757	.0463	765.4

One possible criterion to use in determining how well our model fits our experimental data is the cost associated with each of the inverse problems. The cost or error values for neonate exposure can be seen in Table 6 and values for adult exposure can be seen in Table 7. These cost values will provide us with a means of statistical comparison (to be discussed later) when we add more degrees of freedom to our model. It is of value to note that although the majority of the optimal rates yield models that only approximate the experimental data, the optimal death rate corresponding to a neonate exposure of 80ppm results in a cost of only three. This observation is consistent with Figure 7. The graphs of the experimental data and the SS simulation run with the optimal value for 80ppm, lie almost on top of each other. A possible explanation of this involves the fact that at an exposure of 80ppm, the experimental data contains deaths from day one to day two mirroring the mathematical effects of SS.

Comparing the values (Tables 6 and 7) of the Leslie obtained μ to our SS estimated μ using the SS model may not be a particularly meaningful comparison. There does not appear to be an acceptable correlation between the death rates obtained from Leslie and those obtained through our OLS problem. We believe that using the Leslie simulations to predict parameters for sublethal effects may not be as reliable as using SS.

To obtain a better understanding of our results and their dependence on the amount of experimental data available, we decided to simulate the inverse problem using more "data". As stated before, the inverse problem was carried out using data from only one

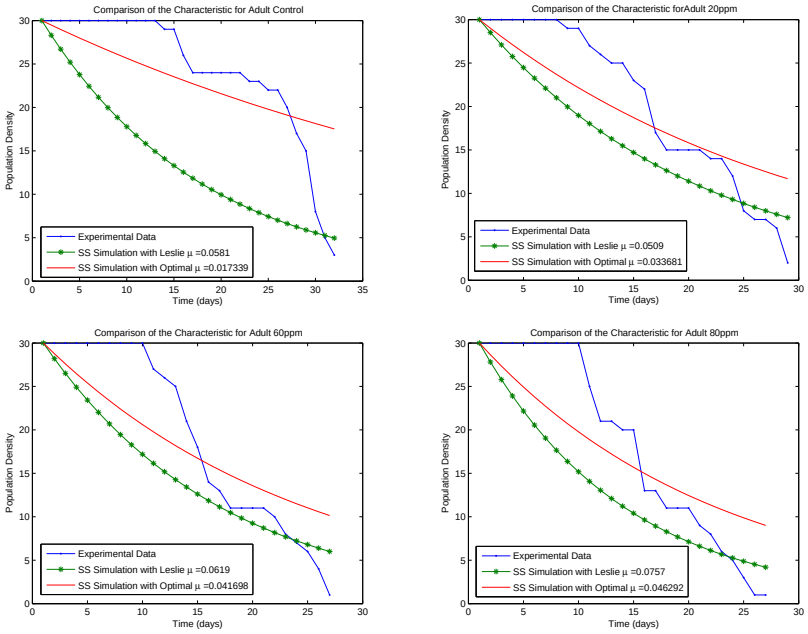

Figure 8: Comparisons of experimental data to SS simulations using Leslie μ and the optimal OLS μ for adult exposure

characteristic. That means that we determined the optimal OLS values for death rate based on only a limited set of data points, ranging from 9 to 32 depending on exposure level and method. We hypothesize that given more experimental data, we would obtain more informative results. A second characteristic of *simulated* observations was added to our experimental data. We created this simulated data using SS simulations run with a systematic assortment of death rates that increase as concentration increases (See Tables 8 and 9). The first characteristic with non-zero densities resulting from the boundary condition driven solution was added to our experimental data to create a two characteristic data set. Results for the inverse problem run with these two characteristics for neonate exposure can be found in Table 8, and results for adult exposure in Table 9.

These results suggest that when the second characteristic, dependent on recruitment, is added to our experimental data, we are sometimes able to match the death rates used to create the simulated data. For neonate exposure, we find a match in the optimized μ for the first three concentrations. The second three concentrations are so high that there are no aphids being born into the population, i.e., the recruitment is equal to zero. This causes the addition of another characteristic to be only the addition of zeros. Thus, our optimized values with two characteristics are almost identical to those computed with only one characteristic. For adult exposure, we find that no matter what death rate we use to simulate the second characteristic, we are able to recover that same value with the inverse

Table 8: OLS results of estimation of μ for neonate exposure with two characteristics

Concentration	**Two Characteristics**		**One Characteristic**
	Simulated Data μ	**OLS/SS μ**	**OLS/SS μ**
Control	.2	.2	.0128
10ppm	.3	.3	.0277
20ppm	.4	.4	.0396
40ppm	.5	.0951	.0951
60ppm	.6	.2033	.2034
80ppm	.7	.7937	.7936

Table 9: OLS results of estimation of μ for adult exposure with two characteristics

Concentration	**Two Characteristics**		**One Characteristic**
	Simulated Data μ	**OLS/SS μ**	**OLS/SS μ**
Control	.1	.1	.0173
10ppm	.2	.2	.0251
20ppm	.3	.3	.0337
40ppm	.4	.4	.0420
60ppm	.5	.5	.0417
80ppm	.6	.6	.0463

problem. This suggests, as one might expect, that in some cases of interest, additional experimental data would be most helpful in estimating death rates.

4.3.1 Implications

The above results serve to illustrate to us the importance of having more experimental data if one insists on models with constant death rate. Data along one characteristic does not appear to be enough to compute optimal constant values of the death rates unless the concentration of pesticide is so high that it immediately kills off the population. Through an analysis of the results obtained from our OLS problem, we see a sizable discrepancy between the constant death rate calculated from the Leslie simulation and the constant optimal death rate with SS; however neither model with these constant death rates yields a very good fit to the experimental data.

We suspect that a main reason for the lack of good fit to our experimental data lies in the fact that when evaluating the data there appears to be a *time dependency* in the death rate. Consider equation (7). The initial condition driven solution incorporates the death rate into its first calculation, producing deaths from day one to day two. However,

our experimental data shows that for the neonate control, there are no deaths until day 15, and for the adult control, there are no deaths until day 13. This pattern follows for each concentration level and exposure method, except for neonate exposure of 80ppm. For 80ppm, deaths do occur from day one to day two, allowing SS to more closely match the data. Recall that Leslie gave us a constant death rate, which we then used in our SS simulations. This suggests that perhaps death rate should not be treated as a constant, but rather should be time-dependent. This dependence would vary with exposure level and method.

Considering the nature of our experiments with the varying levels of application, we suspect sublethal damage and hence a possible need for time dependent mortalities. Standard Leslie formulations do not readily permit us to consider death rate as a time dependent parameter. However time dependent death rates are easily accommodated in SS. Since SS is much more flexible, we believe it is more desirable to use for our attempts to model the experimental data.

4.4 Estimating Time Dependent Death Rates

In general, a species' death rate may change as individuals move through different life stages. Thus, even for the control experiments, there is little reason for us to believe that the aphids have a constant death rate. Moreover, in the cases with treatment, in addition to aging, the aphids are also being exposed to pesticide. We should expect to see improved results if we allow μ to be time dependent, especially if we believe sublethal damage occurs during pesticide exposure. Therefore we next consider simulations with $\mu = \mu(t)$ in the SS model.

We take μ as a time dependent, piecewise constant variable by assuming a parametric representation of the time dependent function given by

$$\mu(t) \approx \mu^M(t) = \sum_{j=1}^{M} a_j \chi_j^M(t),$$

where

$$\chi_j^M = \begin{cases} 1 & [t_{j-1}^M, t_j^M) \\ 0 & \text{elsewhere,} \end{cases}$$

with time partitions or nodes $\{t_j^M\}$. We are thus seeking to estimate the M nodal values $\{a_j\}_{j=1}^M$ for $\mu^M(t)$ that provide a good fit to the experimental data.

Since our death rate is no longer constant, we must return to equation (5) to derive a SS solution for this case. However, our observations only contain data along one characteristic, and thus it is only necessary for us to derive a solution for the initial condition driven

solution (t≤G(x)). Along the characteristic $X(0;x,t)$ we obtain

$$\begin{aligned} u^M(x,t) &= \Phi(X(0;x,t))e^{-\int_0^t [g_x(\xi,X(\xi;x,t))+\mu^M(\xi,X(\xi;x,t))]d\xi} \\ &= \Phi(x-t)e^{\int_0^t \mu^M(\xi)d\xi} \\ &= \Phi(x-t)e^{-\int_0^t \sum_{j=1}^M a_j \chi_j^M(\xi)d\xi} \\ &= \Phi(x-t)e^{-\sum_{j=1}^M a_j \int_0^t \chi_j^M(\xi)d\xi}. \end{aligned}$$

Thus our SS model with time dependent death rate with, for example, $M = 4$ is given by

$$u^M(x,t) = \Phi(x-t)e^{-\int_0^t \mu^M(\xi)d\xi} \tag{11}$$

with

$$e^{\int_0^t \mu^M(\xi)d\xi} = \begin{cases} e^{-a_1 t} & [0, t_1^M) \\ e^{(a_2-a_1)t_1^M} e^{-a_2 t} & [t_1^M, t_2^M) \\ e^{(a_3-a_2)t_2^M} e^{(a_2-a_1)t_1^M} e^{-a_3 t} & [t_2^M, t_3^M) \\ e^{(a_4-a_3)t_3^M} e^{(a_3-a_2)t_2^M} e^{(a_2-a_1)t_1^M} e^{-a_4 t} & [t_3^M, t) \end{cases}$$

Table 10: OLS results of estimation of $\mu(t)$ for adult exposure

	Estimated μ Constant	**Estimated μ 2 Nodes**	**Estimated μ 4 Nodes**
Control	.0173	[.0027, .0599]	[0, .0194, .0106, .1755]
10ppm	.0251	[.0082, .0802]	[.0019, .0291, .0250,.2448]
20ppm	.0337	[.0134, .1118]	[0, .0437, .0752, .1735]
40ppm	.0420	[.0172, .1651]	[0, .0579, .1178, .2422]
60ppm	.0417	[.0159, .1600]	[0, .0520, .1313, .1472]
80ppm	.0463	[.0192, .1795]	[0, .0649, .1249, .2666]

	Cost Constant	**Cost 2 Nodes**	**Cost 4 Nodes**
Control	766.3	307.0	95.6
10ppm	768.3	245.2	50.9
20ppm	587.4	105.5	40.7
40ppm	919.0	217.3	79.1
60ppm	737.0	163.3	72.6
80ppm	765.4	162.1	49.5

To solve the inverse problem using our new model from equation (11), we again used the MATLAB routine fminsearch. For adult and neonate exposure, our MATLAB program provided the results given in Table 10 and Table 11, respectively, for several different

values of M (results for a complete set of values of M between 1 and 6 can be found in [BBDS]). An important modeling question is how large we should choose M (i.e., what level of parametrization is most appropriate?). Theoretical arguments related to increased degrees of freedom guarantee that we should get at least as good a value of the cost function each time we add parameters through partition refinements (i.e., model refinement through refinements of partitions for parameter discretization) to our original parametrization (e.g., see [BK]). With the statistical tests we employ here, we can readily compare results for refinements of our parametrization. Thus we can look at comparisons of one (i.e., μ constant) versus any number of additional parameters, two versus four parameters, etc. We chose Nelder-Mead for optimization because of its simplicity and its consistent performance record in numerous other optimization problems.

Table 11: OLS results of estimation of $\mu(t)$ for neonate exposure

	Estimated μ Constant	**Estimated μ 2 Nodes**	**Estimated μ 4 Nodes**
Control	.0128	[0, .0501]	[0, .0070, 0, .1972]
10ppm	.0277	[.0098, .0870]	[.0180, .0010, .0613, .2031]
20ppm	.0396	[.0253, .0897]	[.0270, .0328, .0345, .2799]
40ppm	.0951	[.1057, .0385]	[.0956, .1350, 0, .1010]
60ppm	.2034	[.2127, .0300]	[.1987, .2693, 0, .0394]
80ppm	.7936	[.7959, 0]	[.7895, .8492, 0, 0]

	Cost Constant	**Cost 2 Nodes**	**Cost 4 Nodes**
Control	1110.2	557.7	128.5
10ppm	571.5	113.8	32.4
20ppm	297.6	131.3	37.6
40ppm	114.6	80.3	66.2
60ppm	122.6	99.8	93.6
80ppm	3.3	.967	.917

Table 11 contains the MATLAB program results for neonate exposure. In this case, computations for comparisons of our simulations with neonate exposure also produced cost function values corresponding to theoretical predictions for all exposure levels and all nodal values.

4.5 Comparisons

We next compare our simulations and results for different types of exposure. Simulations run using Leslie's constant μ, SS's optimal constant μ and SS's optimal time dependent

$\mu^M(t)$ with $M = 4$ nodal values are compared with our experimental data in Figures 9 (adult exposure) and 10 (neonate exposure).

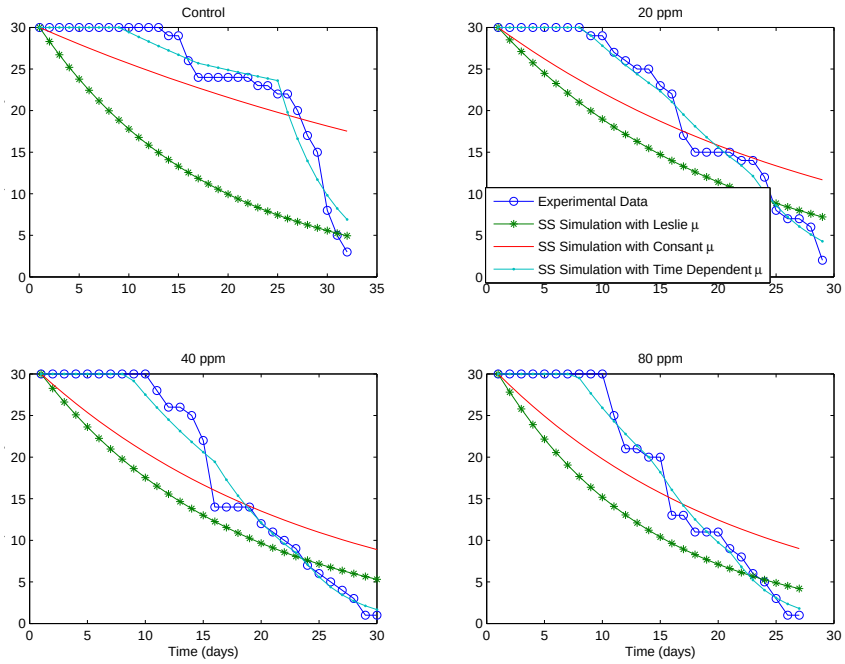

Figure 9: Comparisons of experimental data to SS simulations using Leslie μ, the optimal constant μ and the optimal $\mu^M(t)$ with 4 nodes obtained for adult exposure

For adult exposure with four nodal values of $\mu^M(t)$ (Figure 9), we see that the simulation trajectories with $\mu^M(t)$ appear to closely match our experimental data and provide a striking improvement over models with constant mortality rates. Moreover, corresponding figures (see [BBDS]) for refinements (as reported in Tables 10 and 11) reveal that we obtain a better approximation to our experimental data each time we increase the number of nodal values for $\mu^M(t)$ (i. e., refine the time partitions). In every case, solutions with the time dependent death rate provide a better approximation to our experimental data than do solutions with the constant death rates from Leslie or solutions from our initial optimization of SS with constant mortality.

With neonate exposure $M = 4$ (Figure 10), we again see different results using $\mu^M(t)$ for different exposure levels. There appears to be less of a need for a time dependent death rate as exposure levels increase. For the control and for 20ppm exposure, the time dependent μ^M appears to do a better job approximating the experimental data. However for an exposure level of 40ppm and especially 80ppm, the time dependent μ^M does not appear to model the experimental data too differently from the optimal constant μ. For 80ppm exposure, in

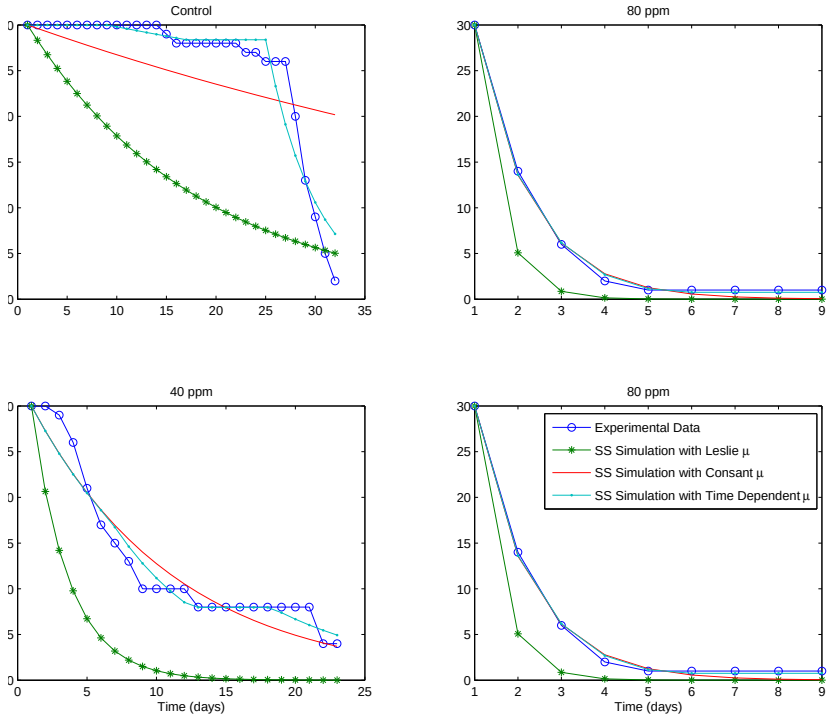

Figure 10: Comparisons of experimental data to SS simulations using Leslie μ, the optimal constant μ and the optimal $\mu^M(t)$ with 4 nodes obtained for neonate exposure

day 2 the population experiences drastic deaths that appear to continue as the population declines. Therefore it may be reasonable to treat death rate as a constant value for those exposure levels that exhibit substantial immediate deaths. In these cases sublethal damage may not be an important factor in the population dynamics.

4.5.1 Leslie Considerations

One may ask how the SS Leslie trajectory compares with the actual Leslie matrix model output. The Leslie model output involves only the percent mortality ($mort$) value, which is not the death rate. The Leslie simulation is run, output received and finally used to compute the death rate. Therefore we do not expect that the output of SS using the Leslie mortality to be similar to that of the actual Leslie model output. For sake of comparison, graphs of these simulations are given in Figures 11 and 12. Henson [H] discusses the relationship between Leslie and the SS PDE models (see also [U]). She points out that "the Leslie model of choice in a given approximation is often low dimensional, and is hence too coarse

a discretization to bear much realization to the associated PDE."

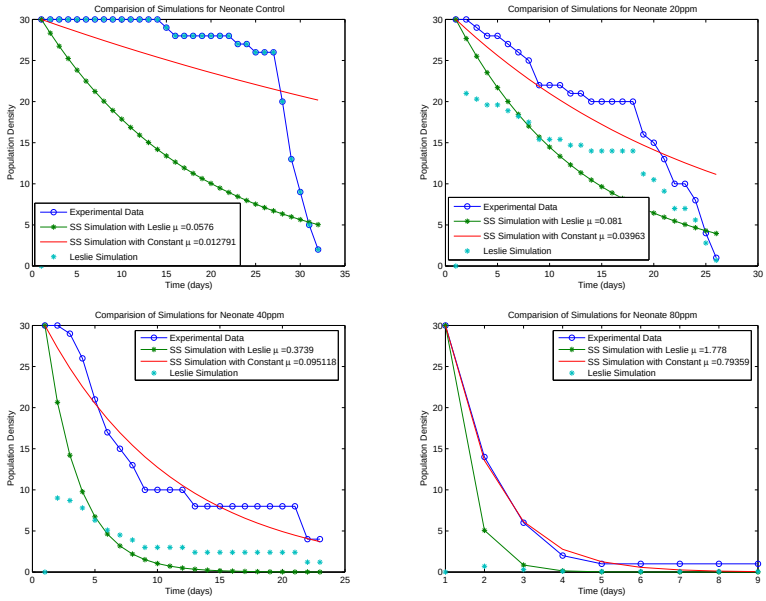

Figure 11: Comparison of Leslie and SS simulation for neonate exposure

The figures support and illustrate this statement. The Leslie model does not do a very good job of estimating the experimental data. One may say that for the control, the Leslie model is exactly the data. This is true because $mort = 0$ for both neonate and adult control. The Leslie model uses the experimental data in its calculations; therefore, for the lower pesticide concentrations it appears to approximate the data. However for all higher concentration levels, the Leslie model does not come close to estimating the true values. The Leslie model performs even worse than the SS model with the Leslie computed death rate. This analysis leads us to again conclude that the discrete Leslie model is not an adequate population model for these sublethal dynamics. The SS PDE model clearly does a much better job approximating the data.

Overall, our results strongly suggest that death rates in sublethal insecticide treatments should not be thought of as a constant value unless the exposure levels are so high that immediate death is certain. This only occurs for neonate exposure with high concentration levels, thus making a decision as to how to consider death rate (constant vs. non-constant) relatively straight forward. It is rather clear that we should use time dependent death rates when we wish to characterize the effects of sublethal damage on the aphids.

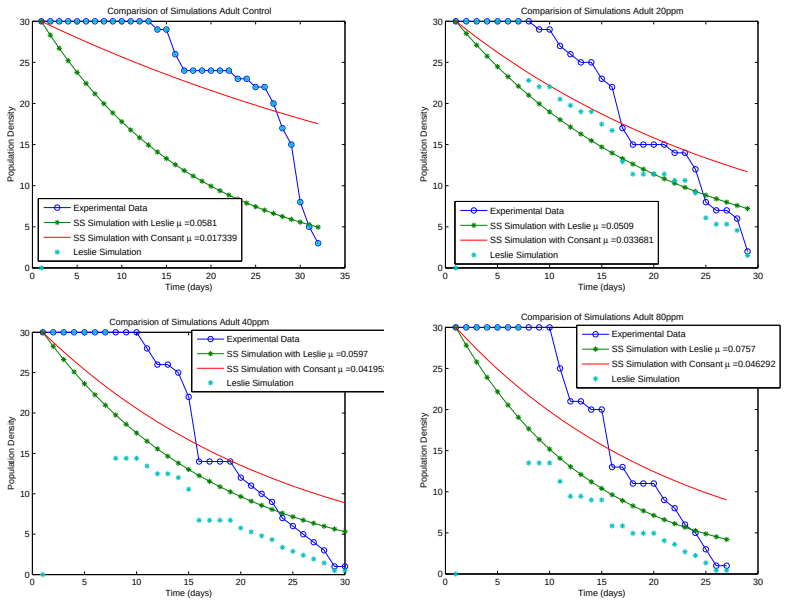

Figure 12: Comparisons of Leslie and SS simulation for adult exposure

4.6 Determination of Confidence Intervals

Since our estimated parameters are obtained using experimental data with inherent measurement error, the values have an uncertainty associated with them that can be quantified in a standard statistical framework [DG, SeWi, WMS] involving confidence intervals. Here we use a classical large sample size formulation for nonlinear systems as briefly summarized in the appendix. We determine confidence intervals for each of our estimated parameters. The "parameter" of interest is the death rate, $\mu^M(t)$, which we represent through a vector of coefficients

$$\mu^M(t) \Leftrightarrow \{a_k\}_{k=1}^M = \bar{a}, \tag{12}$$

where M is the number of nodal values of $\mu^M(t)$.

As part of our computation of standard errors and hence confidence intervals, we need to compute an approximation, $\hat{\sigma}^2$, for the variance in the measurement or observation process and a sensitivity matrix, $F_{\bar{a}}$. Both of these values are computed using the experimental data and the model solution. Recall that we have experimental data along (x, t), where $t = x - x_0$, and thus we will consider $x = t + x_0$. The model solution with which we are concerned is $U(t; \bar{a}) = U_{SS}(t+x_0, t; \bar{a})$, where $U_{SS}(x, t; \bar{a})$ is the solution of the Sinko-Streifer

equation with $\mu(t)$ approximated by

$$\mu(t) \approx \mu^M(t) = \sum_{k=1}^{M} a_k \chi_k^M(t), \tag{13}$$

where

$$\chi_k^M = \begin{cases} 1 & [t_{k-1}^M, t_k^M) \\ 0 & \text{elsewhere.} \end{cases}$$

Even though we begin simulations at $t_0 = 1$, we consider parameter approximations on $[0, t_{max}]$ and use partitions $t_k^M = \frac{t_{max} k}{M}$. Here and below we suppress the notation indicating that U_{SS} depends on M. The variance is approximated by [DG, SeWi, WMS]

$$\sigma^2 \approx \hat{\sigma}^2 = \frac{1}{n-M} \sum_{j=1}^{n} \left(U(t_j, \bar{a}) - U_j^d \right)^2,$$

with n representing the number of data points, M representing the number of nodal values of $\bar{a}$, and U_j^d representing the experimental data. The $n \times M$ sensitivity matrix $F_{\bar{a}}$ is computed in the following manner:

$$F_{\bar{a}} = \left((F_{\bar{a}})_{jk} \right) = \left(\frac{\partial U(t_j, \bar{a})}{\partial a_k} \right) = \begin{pmatrix} \frac{\partial U(t_1)}{\partial a_1} & \cdots & \frac{\partial U(t_1)}{\partial a_M} \\ \vdots & \ddots & \vdots \\ \frac{\partial U(t_n)}{\partial a_1} & \cdots & \frac{\partial U(t_n)}{\partial a_M} \end{pmatrix}.$$

As explained in the appendix, we assume a statistical model for the observation process given by: $Y_j = U(t_j, \bar{a}) + \epsilon_j$, with $\bar{a} \in R^M$ where $E[\epsilon_j] = 0$ and $var[\epsilon_j] = \sigma^2$. The formula for the confidence intervals ($k = 1, \ldots, M$) is given in [WMS] as

$$[a_k - t_{1-\frac{\alpha}{2}} SE(a_k), a_k + t_{1-\frac{\alpha}{2}} SE(a_k)],$$

where $t_{1-\frac{\alpha}{2}}$ is a t–distribution value obtained from a statistical chart. The standard errors, $SE(a_k)$, are determined by the formula

$$SE(a_k) = \sqrt{\Sigma_{kk}},$$

with Σ_{kk} the diagonal elements of the covariance matrix, Σ, which is approximated by

$$\Sigma \approx \hat{\sigma}^2 \left[F_{\bar{a}}^T F_{\bar{a}} \right]^{-1}.$$

We use the above derivations to compute confidence intervals for a typical case in our problem. Recall, for example, the SS solution when $M = 4$. We have

$$U(t; \bar{a}) = \begin{cases} \Phi(x_0) e^{-a_1 t} & [0, t_1^M) \\ \Phi(x_0) e^{(a_2 - a_1) t_1^M} e^{-a_2 t} & [t_1^M, t_2^M) \\ \Phi(x_0) e^{(a_3 - a_2) t_2^M} e^{(a_2 - a_1) t_1^M} e^{-a_3 t} & [t_2^M, t_3^M) \\ \Phi(x_0) e^{(a_4 - a_3) t_3^M} e^{(a_3 - a_2) t_2^M} e^{(a_2 - a_1) t_1^M} e^{-a_4 t} & [t_3^M, t). \end{cases}$$

Thus, the sensitivity matrix, $F_{\bar{a}}$, is given by

$$F_{\bar{a}} = \left(\frac{\partial U(t_j, \bar{a})}{\partial a_k} \right)$$

where

$$\begin{aligned}
\frac{\partial U}{\partial a_1} &= \begin{cases} -tU(t:\bar{a}) & [0, t_1^M) \\ -t_1^M U(t:\bar{a}) & [t_1^M, t) \end{cases} \\
\frac{\partial U}{\partial a_2} &= \begin{cases} 0 & [0, t_1^M) \\ (t_1^M - t)U(t:\bar{a}) & [t_1^M, t_2^M) \\ (t_1^M - t_2^M)U(t:\bar{a}) & [t_2^M, t) \end{cases} \\
\frac{\partial U}{\partial a_3} &= \begin{cases} 0 & [0, t_2^M) \\ (t_2^M - t)U(t:\bar{a}) & [t_2^M, t_3^M) \\ (t_2^M - t_3^M)U(t:\bar{a}) & [t_3^M, t) \end{cases} \\
\frac{\partial U}{\partial a_4} &= \begin{cases} 0 & [0, t_3^M) \\ (t_3^M - t)U(t:\bar{a}) & [t_3^M, t) \end{cases}.
\end{aligned}$$

Once we obtain $F_{\bar{a}}$, we can compute the approximate standard errors SE as outlined above. To complete the confidence interval calculations, a level of significance must be chosen. We choose $\alpha = .05$ which corresponds to a significance level of 95%. (When the number of data points is large, $n \geq 30$, the corresponding t-value is 1.96). Thus we are confident that 95% of the time our estimation procedure will yield confidence intervals that cover the nodal values of the "true" values of $\mu^M(t)$.

As usual in asymptotic (large n) arguments, the above formulation tacitly assumes the existence of a "true" $\mu^M(t)$. As in most applications of the theory, this may not be exactly true. Moreover, there are other fundamental assumptions (see the appendix) such as independence of errors and constant variance under which the arguments are carried out. For the data used here, we did plot residuals vs. time and residuals vs. estimated models and the findings suggest that for our problem the assumptions of independence and constant variance for the errors ϵ_j are valid. It is ordinarily not known in any specific application the size of n under which such asymptotic results will yield valid approximations. For sample size $n \approx 30$ as in most of the data sets used here, the estimation of μ^M, for $M = 1, \ldots, 4$, entailed efforts involving a range of 30 sample points per parameter for $M = 1$ to ≈ 7 sample points per parameter for $M = 4$. Although it is not clear whether these conditions are sufficient to lead to fully reliable approximations, use of these asymptotic tools in evaluating the estimation results and hypothesis testing detailed below is probably a reasonable, albeit possibly optimistic, first measure of uncertainty to consider.

The exceptions to these remarks were for the neonate exposures at the 40, 60 and 80 ppm levels where the corresponding numbers of data points were $n = 23, 21$ and 9, respectively

(see [BBDS]). In these cases one has higher t-values and hence significantly larger confidence intervals as well as smaller values of n. Consequently, one might place less relevance in the related confidence intervals and in the corresponding p-values reported in the statistical comparison tests in the next section.

The above method was utilized to determine confidence intervals for all exposure levels and methods with $\bar{a}^M$ for several values of M. The results for neonate exposure are given for $M = 1, 2, 4$ in Table 12 and for $M = 4$ in Figures 13 and 14. The corresponding results for adult exposure are in Table 13 and Figures 15 and 16. (Similar findings for other values of M can be found in [BBDS].) The tables contain the actual confidence intervals, while the figures contain the population densities computed using the optimal $\mu^M(t)$ and the "upper and lower bounds" on $\mu^M(t)$, and the error bars on $\mu^M(t)$. The "upper bounds" on $\mu^M(t)$ are the SS simulations computed using the highest death rate resulting from the confidence interval calculations (highest value of error bar) while the "lower bounds" on $\mu^M(t)$ are the SS simulations computed using the lowest death rate (lowest value of error bar) resulting from the confidence interval calculations. High death rates produce the "lower band" of population simulations; low death rates produce the "high band" of simulations.

Table 12: Confidence intervals for neonate exposure: $\bar{a_k} \pm t_{1-\frac{\alpha}{2}} SE(a_k)$

	Control	**10ppm**	**20ppm**
M=1	.0128± .0188	.0277± .0223	.0296± .0255
M=2	0 ± .0080	0 ± .0046	.0253± .0080
	.0501± .0044	.0870± .0029	.0897± .0041
M=4	0 ± .0188	.0180± .0068	.0270± .0108
	.0070± .0043	.0010± .0028	.0328± .0051
	0 ± .0053	.0613± .0039	.0345± .0068
	.1972± .0095	.2031± .0079	.2799± .0104

	40ppm	**60ppm**	**80ppm**
M=1	.0951± .0582	.2034± .5027	.7174± 3.5263
M=2	.1057± .0141	.2127± .0346	.7959± .0153
	.0385± .0043	.0300± .0054	0 ± .0010
M=4	.0956± .0257	.1987± .0479	.7895± .0598
	.1350± .0160	.2693± .0535	.8492± .0235
	0 ± .0245	0 ± .1108	0 ± .0282
	.1010± .0203	.0394± .0654	0 ± .0150

Table 13: Confidence intervals for adult exposure: $\bar{a_k} \pm t_{1-\frac{\alpha}{2}} SE(a_k)$

	Control	**10ppm**	**20ppm**
M=1	.0173± .0179	.0251± .0228	.0337± .0428
M=2	.0027± .0061	.0082± .0058	.0134± .0049
	.0599± .0037	.0802± .0041	.1118± .0039
M=4	0 ± .0102	.0019± .0075	0 ± .0074
	.0194± .0039	.0291± .0030	.0437± .0036
	.0106± .0052	.0250± .0043	.0752± .0061
	.1755± .0093	.2448± .0096	.1735± .0122

	40ppm	**60ppm**	**80ppm**
M=1	.0420± .0415	.0417± .0417	.0463± .0477
M=2	.0172± .0069	.0159± .0072	.0129± .0074
	.1651± .0064	.1600± .0066	.1795± .0072
M=4	0 ± .0096	0 ± .0128	.7895± .0106
	.0579± .0048	.0520± .0055	0 ± .0048
	.1178± .0094	.1313± .0110	.0649± .0097
	.2422± .0262	.1472± .0238	.1249± .0239

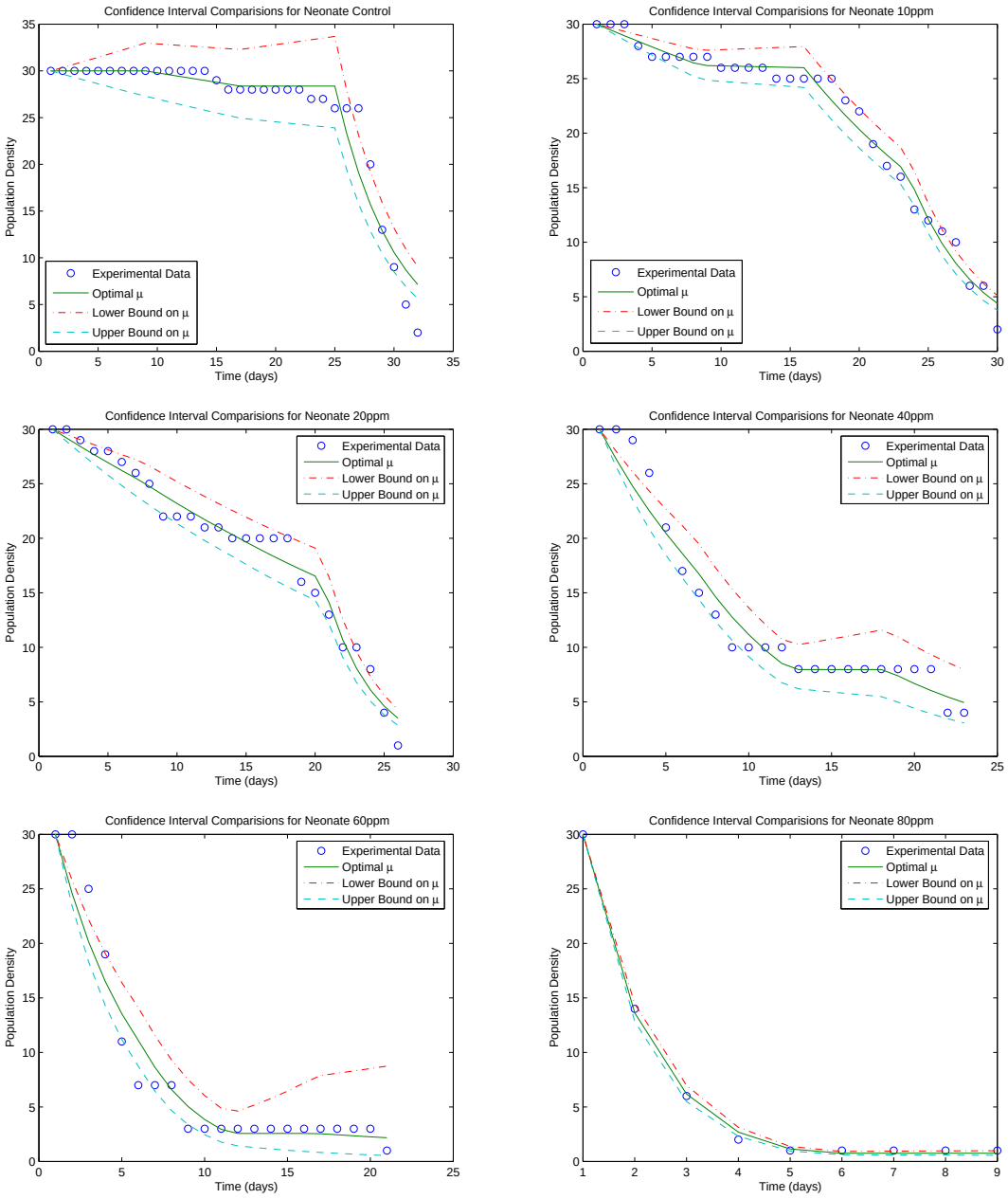

Figure 13: Population densities for neonate exposure run with upper and lower bounds of confidence intervals (M=4)

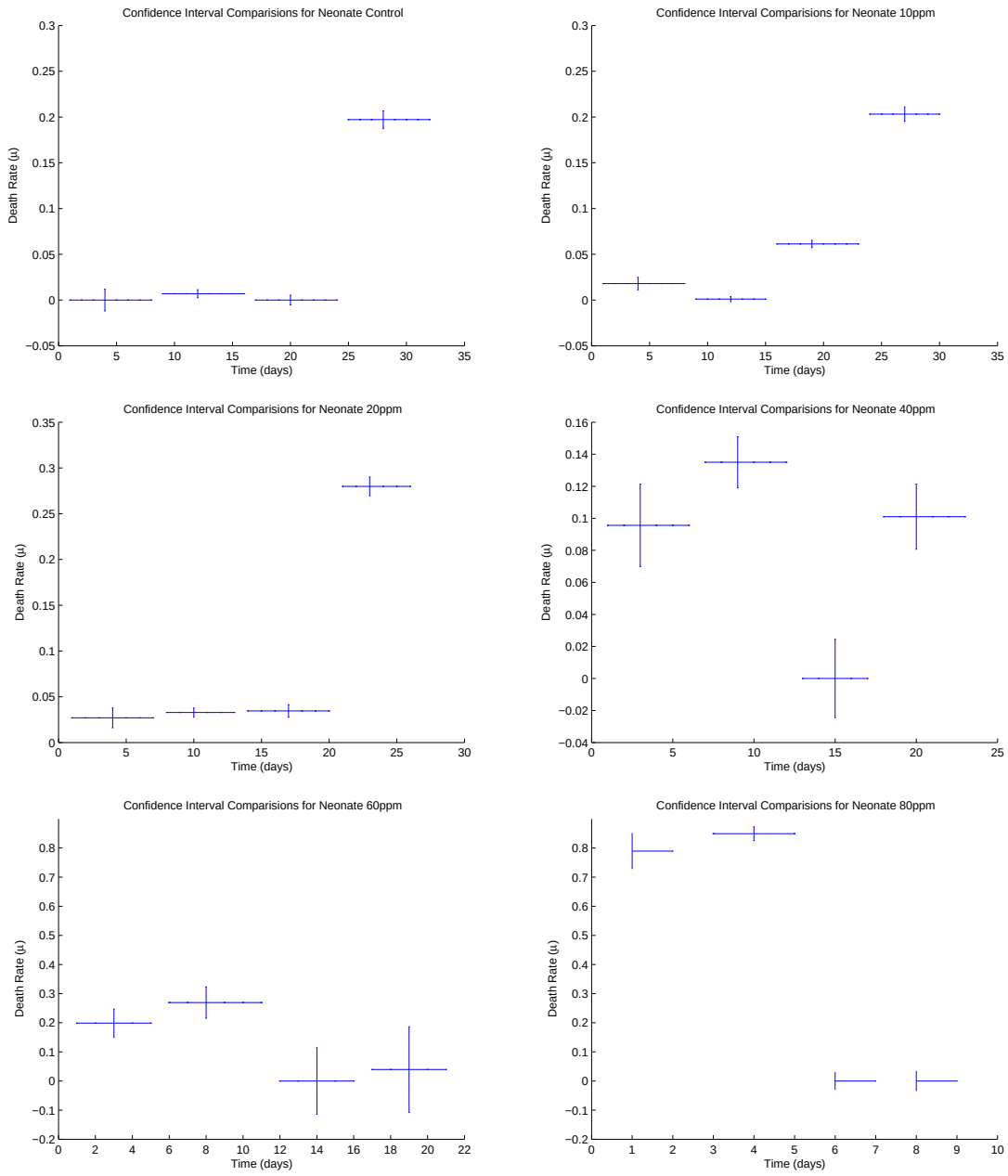

Figure 14: Confidence intervals on μ^M for M=4 neonate exposure (note scaling in graphs differs with exposure levels)

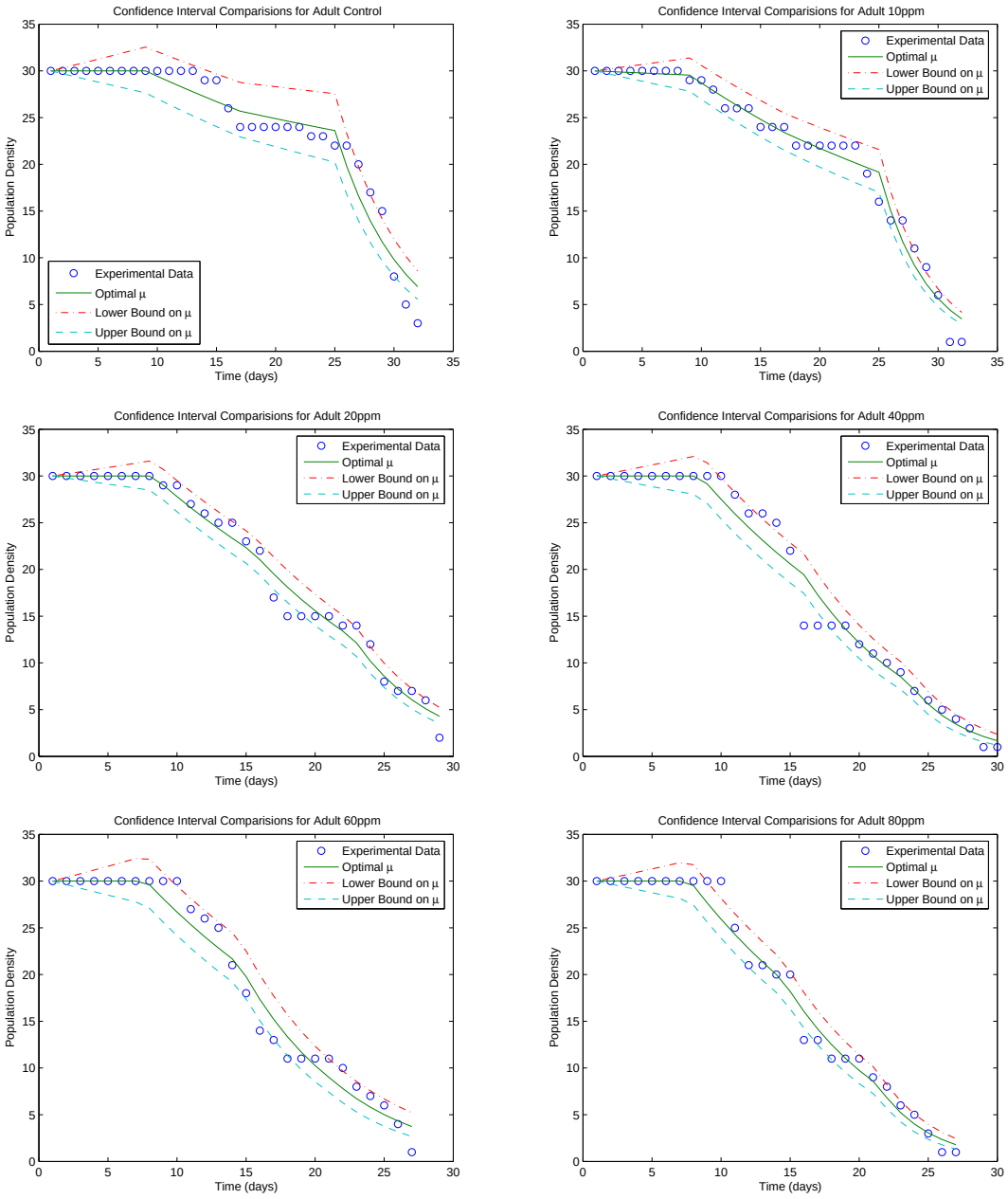

Figure 15: Population densities for adult exposure run with upper and lower bounds of confidence intervals (M=4)

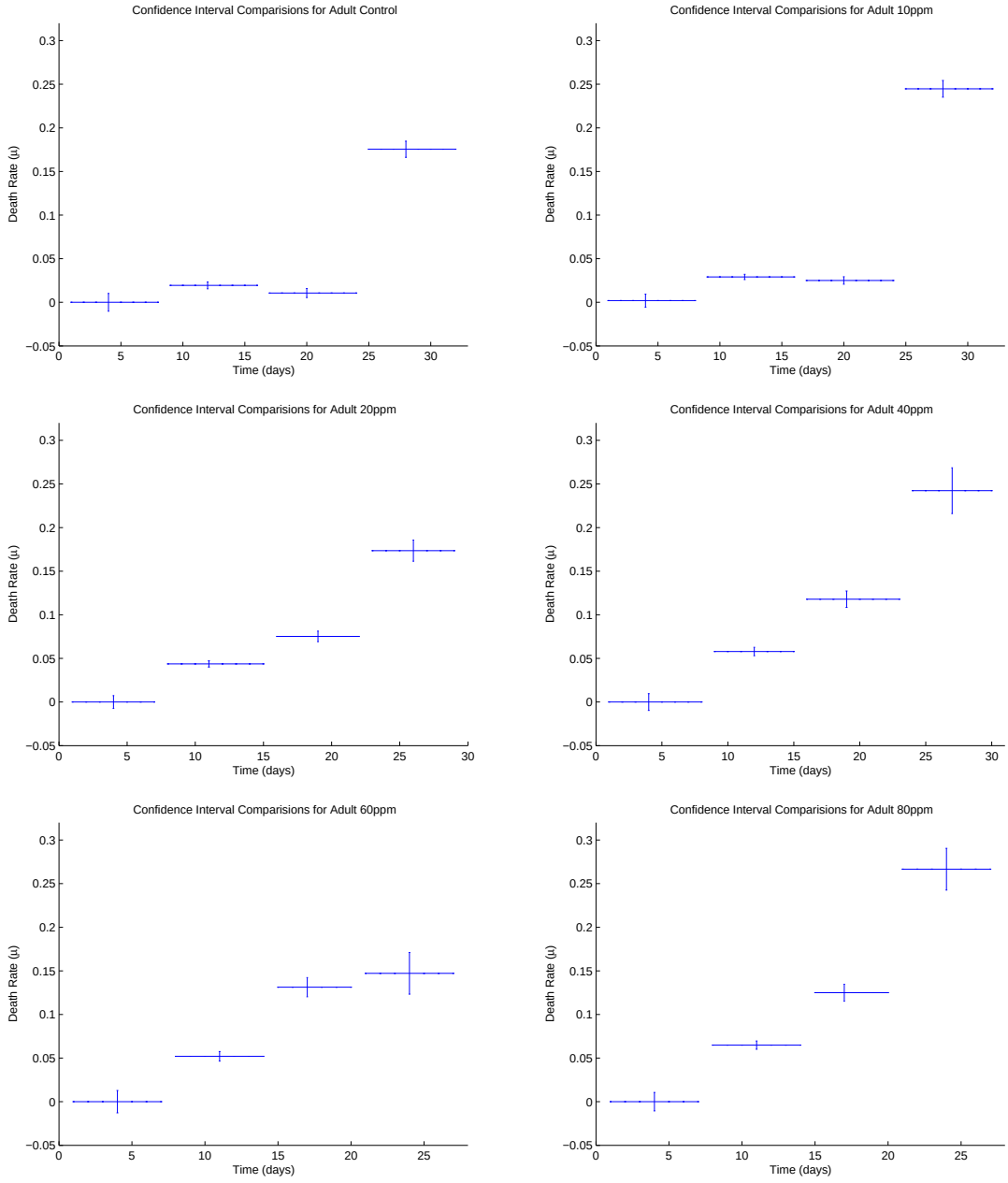

Figure 16: Confidence intervals on μ^M for M=4 adult exposure

4.7 Model Comparisons

The previous discussions depicted graphically our empirical findings about the consideration of death rate as a time dependent parameter. Examination of our figures provides qualitative evidence that the model with $\mu^M(t)$ does a better job representing our experimental data than does the model with constant μ. However, we would like some quantitative measure of our success in modeling the data with the time dependent $\mu^M(t)$ versus the optimal constant μ. Moreover, we would like information on the important question of over-parametrization in such inverse problems. We turn to statistical tests for model comparisons to assist with this analysis.

We used a statistical method known as hypothesis testing to assist us in better understanding our results at the various levels of parametrization. This method, as developed in [BF1, BF2, F], is detailed in [BK] and its formulation for use here is summarized in the appendix below. We expect that increasing the degrees of freedom of our death rate through parameter partition refinements ($M_2 > M_1$ corresponding to $\theta_{M_i} = (a_1, a_2, \ldots, a_{M_i})^T$) from our optimal $\mu^M(t)$ (see (12) and (13)) will provide us with a reduced cost function value (i.e., a reduced residual sum of squares in the minimized OLS). We wish to ascertain if this reduced cost signifies a truly better fit, statistically speaking, to our experimental data, or if the reduced cost is simply a reflection of reduction due to increased degrees of freedom. For a precise statement of the statistical framework, see the appendix. As indicated in the appendix, the test statistic values one obtains are then used with χ^2 tables. One determines the degrees of freedom and then compares the test statistic value with threshold values in the table to determine either the appropriate significance levels α or, more commonly, one uses the test statistic to compute the associated $p - values$ (see Chapter 8 of [CB] and Chapter 3 of [Leh]).

The calculated χ^2 test statistics and corresponding p-values for neonate exposure at various exposure levels can be found in Table 14 for comparisons of parameter refinements involving $M = 4$ vs. $M = 1$, $M = 2$ vs. $M = 1$ and $M = 4$ vs. $M = 2$. We see that most of our test statistics yield very small p-values. (A similar table for adult exposures yields $p < .001$ for *all* the corresponding comparisons.) However, as seen in Table 14, as neonate exposure levels increase in the comparison $M = 4$ vs. $M = 2$, we obtain larger p-values and might be tempted to conclude that additional parameters for $\mu^M(t)$ are not significant for some exposure levels and additional number of nodal points. Recall (see the remarks in the previous section) however that that $n \approx 30$ for all data sets except for neonate exposure at the $40, 60$ and 80 ppm levels where the values of n are smaller and consequently the asymptotic theory on which the statistical tests are based is less likely to provide reliable results.

Even if we conclude in the light of the weak statistical support that additional nodal values are not significant compared to a constant death rate in the higher exposure levels for two versus four nodal values, we might argue that we are justified in determining that *in general* additional nodal values in $\mu^M(t)$ provide us with a better fit to the experimental data except for the high exposure levels. We note that, weak statistical support not with standing, higher exposure levels for the neonates likely result in high numbers of immediate

Table 14: $\chi^2(s)$ test statistics and corresponding p-values for neonate exposure

Nodes	**Control**	**10ppm**	**20ppm**	**40ppm**	**60ppm**	**80ppm**
4 vs. 1	244.4486	499.7015	179.8627	16.7996	6.5043	23.2973
	$p < .001$	$p < .001$	$p < .001$	$p < .001$	$p = .01$	$p < .001$
2 vs. 1	31.6944	120.7690	32.9080	9.8128	4.8108	21.6565
	$p < .001$	$p < .001$	$p < .001$	$p < .002$	$p = .028$	$p < .001$
4 vs. 2	106.887	75.3998	64.8609	4.8973	1.3779	.4821
	$p < .001$	$p < .001$	$p < .001$	$p = .087$	$p = .502$	$p = .786$

deaths in the population due to the pesticide effects so that subsequent constant in time death rates are more likely to be acceptable. That is, it may be true that at $M = 4$, for the neonates we have over parameterized the death rates in the case of higher exposure levels.

The fact that we obtain small p-values and hence significance at a high confidence level for both population classes and for most exposure levels supports our arguments that generally the death rate should be considered as a time dependent parameter. We have shown that the method used to determine the constant death rate using the Leslie simulations does not produce values that are reasonable when compared with experimental data. We would also argue that constant mortality rates in general do not provide an appropriate framework needed to understand sublethal effects when used in the SS model.

5 Discussion and Interpretation of Findings

While the exact mechanisms for death due to *Margosan-O* exposure are still being investigated by entomologists, there is sufficient understanding to attempt to develop and test models against experimental data and draw some conclusions about proposed mechanisms of pesticide induced death. Insecticides generally affect neonates and adults somewhat differently and at different rates. From previous studies [NCSU], entomologists know that each adult female aphid produces 6 to 8 neonates (nymphs) a day up to a total of 100 nymphs. In addition, one finds that it takes a newborn pea aphid or neonate τ days to become an adult aphid where τ can range from 5 to 7 days. During this period the neonates molt approximately daily until they reach adulthood. It is suspected that *Margosan-O* can act on neonates and adults in several ways. First, the surfactant that causes the insecticide to adhere to plants' leaves can coat an aphid's exoskeleton upon contact, causing it to effectively smother to death very shortly after exposure. In the 1994 experiments [RS], Stark and Rangus concluded that *Margosan-O* also affects the reproduction rate of the adult pea aphids. A further possible cause of death in the neonates in the presence of insecticide is mutation and subsequent death during molting. We expect this aspect of damage to be *Margosan-O* concentration dependent. Other sublethal effects such as hormesis, general increase in death rate because of increased susceptibility to environmental threats, etc.,

might also possibly be seen in concentration dependent death rates.

The death rates depicted in Figures 14 and 16 clearly exhibit some of the suspected features of *Margosan-O* exposure. First there is a marked difference in the estimated death rate functions between adults and neonates.

In particular, in Figure 16 where we plot piecewise constant death rates for $M = 4$ obtained from fitting the model to adult data, we note the generally monotone increasing (in time) nature of the death rate with rates generally being slightly increased as compared to those in the control group. This pattern was observed in fits with other M values (see [BBDS]) and suggests that sublethal damage occurs during initial exposure that eventually produces higher death rates in adults. The results for neonates (see Figure 14) suggest a similar sublethal effect at low dose exposure levels. However, at high ($60 - 80$ ppm) dose levels the death rate graphs suggest that the neonate populations are killed off early as evidenced by initially high death rates which drop off to 0 as time increases and the population is completely depleted.

Neonate rates exhibit the clear possibility of molting-related deaths in early periods ($\leq$ 5 to 10 days after exposure) of growth. This is not present in adults rates where the adults appear to be relatively unaffected with only slightly elevated death rates during the early days ($\approx$ 7 to 15 days) after exposure at day 7.

Finally, the experimental protocol and data do not allow one to use the model to make any conclusions about suspected pesticide effects on offspring of exposed adults.

6 Concluding Remarks

In this paper we have presented a continuum (in time and age) modeling approach with a simple computational and statistical methodology to investigate the presence of variable rates in population dynamics. We have demonstrated its utility by application to a specific problem of understanding sublethal effects of pesticides in pea aphid populations, specifically detailing improvements it provides over the standard discrete matrix model approach.

The approach to estimation of death rates we have taken in this paper is mechanistically completely non-parametric in nature. That is, we do not assume any specific shape for the function $\mu(t)$ in our work and hence we do not assume or propose a priori any specific mechanisms of death rate. While our findings did suggest functions that might represent sublethal damage mechanisms, this was not initially one of our goals. Rather, our effort is a "proof-of-concept" approach that attempts to establish that more sophisticated population models with time dependent parameters can offer improvements over a currently popular Leslie model approach, the traditional LC_{50} quantitative analyses or models with constant death rates. In future investigations, one might assume a priori parameterized shape functions in continuum models to represent hypothesized death, growth or fecundity mechanisms. This most likely will lead to the need for different types of data and the use of mathematical models to suggest experimental design.

The mathematical model and and associated statistical methodology for model analysis/comparison we employ is quite general and can be applied to a much broader class of

problems in ecology and toxicology. For example, we treat recruitment by using results from the Leslie model and linear interpolation to determine the recruitment rate $R(t)$. In a more general approach, one might rather use the fecundity rates estimated from data by solving the integral equation (4) for the fecundity kernel k. Of course this cannot be calculated in a traditional forward manner. That is, k becomes a parameter to be estimated as part of the inverse problem procedures. One would essentially treat recruitment in a fundamentally different way writing the PDE in "weak" or "variational form", in order to solve the now coupled system (3)–(4). Finite element or finite difference techniques could then be used to solve this system in the context of the inverse algorithm iterations.

While we used OLS estimators, our approach can also be used in the context of maximum likelihood estimators (MLE) or even generalized least squares (GLS) estimators with more complex variance structure in the statistical model for the observation process (see [DG]). Moreover, the linearization-based large sample theory used for the asymptotic standard error/confidence interval computations are among the simplest approaches available. More sophisticated (but also usually more computationally demanding) methods such as the bootstrap and the jackknife [E, SeWi] could be used. We have not done so here since our data does not contain significant outliers (see Chapter 5 of [SeWi]), apparent non-independence in sampling or nonconstant variance in the longitudinal data. The asymptotic estimates we obtain, while possibly conservative, illustrate sufficiently the ideas we are attempting to discuss.

Acknowledgments

This work was supported in part (HTB and LKD) by the US Air Force Office of Scientific Research under grant AFOSR FA9550-04-1-0220, and in part (LKD) by a CRSC/Lord Corporation Fellowship. The authors are grateful to a referee whose comments and questions on an earlier version of this manuscript resulted in a significantly improved version. They are also most grateful to Dr. Marie Davidian for her comments and suggestions on certain aspects of the statistical methods used in this manuscript.

References

[BBDS] H. T. Banks, J. E. Banks, L. K. Dick, and J. D. Stark, *Estimation of Dynamic Rate Paramters in Insect Populations Undergoing Sublethal Exposure to Pesticides*, CRSC Tech. Rep. 05-22, May, 2005.

[BF1] H. T. Banks and B. G. Fitzpatrick, Inverse problems for distributed systems: statistical tests and ANOVA, LCDS/CCS Rep. 88-16, July, 1988; *Proc. International Symposium on Math. Approaches to Envir. and Ecol. Problems*, Springer Lecture Note in Biomath.,81: 262-273, 1989.

[BF2] H. T. Banks and B. G. Fitzpatrick, Statistical methods for model comparison in parameter estimation problems for distributed systems, CAMS Tech. Rep. 89-4, September, 1989, University of Southern California; *J. Math. Biol.*, 28:501-527, 1990.

[BK] H. T. Banks and K. Kunsich, *Estimation Techniques for Distributed Parameter Systems*, Birkhauser, Boston, 1989.

[BN] H. T. Banks and H. K. Nguyen, Sensitivity of dynamical system to Banach space parameters, CRSC Tech Rep., CRSC-TR05-13, N.C. State University, February, 2005; *J. Math Anal. Appl.*, in press.

[BT] H. T. Banks and H. T. Tran, *Mathematical and Experimental Modeling of Physical and Biological Processes*, to appear.

[B] L. C. Birch, The intrinsic rate of natural increase of an insect population, *Journal of Animal Ecology*, 17: 15-26, 1948.

[C] J. R. Carey, *Applied Demography for Biologists with Special Emphasis on Insects*, Oxford University Press, Oxford, England, 1993.

[CL] J. R. Carey and P. Liedo, Mortality dynamics of insects: General principals derived from aging research on the mediterranean fruit fly (diptera: Tephritidae), *American Entomologist*, 45(1): 49-55, 1999.

[CB] G. Casella and R. L. Berger, *Statistical Inference,* Duxbury, California, 2002.

[Ca] H. Caswell, *Matrix Population Models*, Sinauer Associates, Inc. Publishers, Sunderland, Massachusetts, 2001.

[DG] M. Davidian and D. Giltinan, *Nonlinear Models for Repeated Measurement Data*, Chapman & Hall, London, 1998.

[D] L. K. Dick, *Prediction of life history traits in invertebrate species exposed to pesticide*, Master Thesis, North Carolina State University, Raleigh, 2005.

[E] Bradley Efron, *The Jackknife, the Bootstrap and Other Resampling Plans*, CBMS Vol. 38, Society for Industrial and Applied Mathematics, Philadelphia, 1982.

[F] B. G. Fitzpatrick, *Statistical methods in parameter identification and model selection*, Ph.D. Thesis, Division of Applied Mathematics, Brown University, Providence, RI, 1988.

[G] A. R. Gallant, *Nonlinear Statistical Models*, John Wiley & Sons, Inc., New York, 1987.

[H] S. M Henson, Leslie matrix models as "stroboscopic snapshots" of McKendrick PDE models, *Journal of Mathematical Biology*, 37: 309-328, 1998.

[J] R. I. Jennrich, Asymptotic properties of non-linear least squares estimators., *Ann. Math. Statist.*, 40: 633-643, 1969.

[Ke] C. T. Kelley, *Iterative Methods for Optimization*, Society for Industrial and Applied Mathematics, Philadelphia, 1999.

[K] Mark Kot, *Elements of Mathematical Ecology*, Cambridge University Press, Cambridge, 2001.

[LW] J. Ladin and U. Wennergren, Population growth and structure in a variable environment, *Oecologia*, 93: 394-405, 1993.

[Leh] E.L. Lehmann, *Elements of Large-Sample Theory*, Springer-Verlag, New York, 1999.

[L] P. H. Leslie, On the use of matrices in certain population mathematics, *Biometrika*, 33: 183-212, 1945.

[M] A. G. McKendrick, Applications of mathematics to medical problems, *Proceedings of the Edinburgh Mathematical Society*, 40: 98-130, 1926.

[NCSU] *The NCSU Entomology Homepage,* http://ipm.ncsu.edu/AG271/forages/pea_aphid.html.

[OW] M. D. Ohman and S. N. Wood, Mortality estimation for plankonic copepods: *Pseudocalanus newmani* in a temperate fjord, *Limnology and Oceeanogrpahy*, 41: 126-135, 1996.

[RS] T. M. Rangus and J. D. Stark, Lethal and sublethal effects of the neem insecticide formulation, 'Margosan-O', on the pea aphid, *Pesticide Science*, 41: 155-160, 1994.

[SeWi] G. A. F. Seber and C. J. Wild, *Nonlinear Regression,* John Wiley & Sons, Inc., New York, 1989.

[SS] J. W. Sinko and W. Streifer, A new model for age-size structure of a population, *Ecology*, 48(6): 910-918, 1967.

[SBV] J. D. Stark, J. E. Banks, and R. Vargas, How risky is risk assessment: The role that life history strategies play in susceptibility of species to stress, *PNAS*, 101: 732-736, 2004.

[SW] J. D. Stark and U. Wennergren, Can population effects of pesticides be predicted from demographic toxicological studies?, *Journal of Economic Entomology*, 88:1089-1096, 1995.

[U] G. Uribe, *On the relationship between continuous and discrete models for size-structured population dynamics*, Doctoral dissertation, University of Arizona, Tucson, 1993.

[VF] H. Von Foerster, Some remarks on changing populations, in *The Kinetics of Cellular Proliferation*, F. Stohlman Jr, ed., Grune & Stratton, New York, 1959, pp. 382–407.

[WMS] D. D. Wackerly, W. Mendenhall III, and R. L. Scheaffer, *Mathematical Statistics with Applications*, Duxbury Thompson Learning, USA, 2002.

[W] W. K. Walthall and J. D. Stark, Comparison of two population-level ecotoxicological endpoints: The intrinsic (r_m) and instantaneous (r_i) rates of increase, *Environmental Toxicology and Chemistry*, 16: 1068-1073, 1997.

[WS] U. Wennergren and J. D. Stark, Modeling long-term effects of pesticides on populations: Beyond just counting dead animals, *Ecological Applications*, 10: 295-302, 2000.

[Wo] S. N. Wood, Inverse problems and structured-population dynamics, in *Structured-Population Models in Marine, Terrestrial, and Freshwater Systems*, S. Tuljapurkar and H. Caswell, eds., Chapman and Hall, New York, 1997, pp. 555–586.

[Wo2] S. N. Wood, Spline models of biological population dynamics: How to estimate mortality rates for stage structured populations with dimorphic life histories, *Journal of Mathematics Applied in Medicine & Biology*, 11: 61-78, 1992.

[Wo3] S. N. Wood, Obtaining birth and mortality patterns from structured population trajectories, *Ecological Monographs*, 64(1): 23-44, 1994.

Appendix

We assume that n scalar longitudinal observations are represented by the statistical model

$$Y_j \equiv f_j(\theta_0) + \epsilon_j, \;\; j = 1, 2, \ldots, n, \tag{14}$$

where $f_j(\theta_0)$ is the model for the observations in terms of the state variables and $\theta_0 \in \mathbb{R}^M$ is a "set" of theoretical "true" parameter values (assumed to exist in a standard statistical approach). We assume for our statistical model of the observation or measurement process (14) that the errors ϵ_j, $j = 1, 2, \ldots, n$, are independent identically distributed (*i.i.d.*) random variables with mean $E[\epsilon_j] = 0$ and constant variance $var[\epsilon_j] = \sigma_0^2$, where of course σ_0^2 is unknown. We then have that the observations Y_j are *i.i.d.* with mean $E[Y_j] = f_j(\theta_0)$ and variance $var[Y_j] = \sigma_0^2$.

We consider estimation of parameters using an ordinary least squares (OLS) approach. Thus we seek to use data $\{y_j\}$ for the observation process $\{Y_j\}$ with the model to seek a value $\hat{\theta}^n$ that minimizes

$$J_n(\theta) = \sum_{j=1}^{n} |f_j(\theta) - y_j|^2. \tag{15}$$

Since Y_j is a random variable, we have that the estimator $\hat{\theta}^n_{OLS}$ is also a random variable with a distribution called the *sampling distribution*. Knowledge of this sampling distribution provides uncertainty information (e.g., standard errors) for the numerical values of $\hat{\theta}^n$ obtained using a specific data set $\{y_j\}$ (i.e., a realization of $\{Y_j\}$) when minimizing $J_n(\theta)$.

Under reasonable assumptions on the statistical model above and on smoothness and regularity (the smoothness requirements for model solutions are readily verified using continuous dependence results for differential equations in our example; the regularity requirements involve, among others, conditions on how the observations are taken as sample size increases, i.e., $n \to \infty$), the standard nonlinear regression approximation theory ([DG], [G], [J], and Chapter 12 of [SeWi]) for asymptotic (as $n \to \infty$) distributions can be invoked and standard formulations for the computation of standard errors and confidence intervals as outlined in Section 4.6. Since the computations for confidence intervals and also the model comparison tests outlined below thus depend on asymptotic limit distribution theory, one must interpret the findings as sometimes crude indicators of uncertainty inherent in the inverse problem findings. Nonetheless, it is useful to consider the formal mathematical requirements underpinning these techniques. Among the more readily checked hypotheses are those of the statistical model requiring that the errors ϵ_j, $j = 1, 2, \ldots, n$, are independent identically distributed (*i.i.d.*) random variables with mean $E[\epsilon_j] = 0$ and constant variance $var[\epsilon_j] = \sigma_0^2$. After carrying out the estimation procedures, one can readily plot the residuals vs. time and the residuals vs. the resulting estimated model values. A random pattern for the first is strong support for validity of the independence assumption while a non increasing, random pattern for the latter suggests the assumption of constant variance may be reasonable for the data (measurements) used in the inverse problem calculations. The underlying assumption that the sampling size n must be large (recall the theory is asymptotic in that it holds as $n \to \infty$) is not so readily "verified" and is often ignored. Indeed the asymptotic theories are often used in a very heuristic underlying manner to give a loose feeling for the uncertainty involved in the estimates and the level of parametrization used in approximating the underlying mathematical model.

To investigate possible over-discretization of the functional parameters being estimated, we consider the generic problem of minimizing the OLS functional J_n of (15) over $\Theta_{M_1} \subset \Theta_{M_2} \subset \mathbb{R}^{M_2}$ and define

$$\theta^*_{M_i} = \text{argmin}\{J_n(\theta) | \theta \in \Theta_{M_i}\}, \quad i = 1, 2.$$

Let Θ_{M_1} be defined by

$$\Theta_{M_1} = \{\theta \in \Theta_{M_2} | K\theta = k_0\}, \tag{16}$$

where K is a given $s \times M_2$ matrix of full rank $s < M_2$ and k_0 is a given constant in $\mathbb{R}^s$. Next formulate the test statistic

$$U_n = \frac{n(J_n(\theta^*_{M_1}) - J_n(\theta^*_{M_2}))}{J_n(\theta^*_{M_2})}. \tag{17}$$

Then again under reasonable assumptions on the statistical model above (essentially the same as outlined above) and on smoothness and regularity (the smoothness requirements for

model solutions are again readily verified using continuous dependence results for differential equations in our example; the regularity requirements involve as above, among others, conditions on how the observations are taken as sample size increases, i.e., $n \to \infty$), one can use the nonlinear theory developed in [BF2, F] to argue that for $n \to \infty$, the distribution for U_n can be approximated by a chi-square distribution $\chi^2(s)$ with s degrees of freedom. We use this approximate distribution to test hypotheses about goodness of fit of models to data when one allows increased degrees of freedom in parametrized functions to be estimated. To this end we define the null hypothesis $H_0 : \theta_0 \in \Theta_{M_1}$ with the alternative $H_A : \theta_0 \notin \Theta_{M_1}$. Then we interpret rejection of the null hypothesis as evidence that the fit-to-data over the larger constraint set Θ_{M_2} gives a statistically significant better fit to the data even in the context of increased degrees of freedom.

To carry out the tests, we first carry out the optimizations of J_n over Θ_{M_1} and Θ_{M_2}, obtaining $\theta^*_{M_1}$ and $\theta^*_{M_2}$, respectively. We then compute the test statistic U_n of (17) and use this to compute standard p-values (see p. 397 of [CB]) which allows one to accept or reject H_0 at any desired confidence level. To use this test with the examples of Section 4.7, we identify the parameters $\theta_i \in \Theta_{M_i}$ with the vector $(a_1, a_2, \ldots, a_{M_i})^T$ of the coefficients $\{a_k\}_{k=1}^{M_i}, i = 1, 2$, of death rate representations (12) and (13) for discretization levels $M_2 > M_1$. One possibility is that we do this so that M_2 is always an integer multiple of M_1 and the discretizations are through uniform partitions of the time interval $[0, t_{max}]$. Thus we can readily compare model fits with $M_2 = lM, M_1 = M$ for any integer l by choosing $k_0 = 0$ in (16) and appropriately defining the constraint matrix K. For example, for $M_2 = 2M,\ M_1 = M$, defining the $M \times 2M$ matrix

$$K = K_{M_2:M_1} = K_{2M:M} = \begin{pmatrix} 1 & -1 & 0 & \ldots & \ldots & \ldots & 0 \\ 0 & 0 & 1 & -1 & 0 & \ldots & 0 \\ \vdots & & & & & & \vdots \\ 0 & \ldots & \ldots & \ldots & 0 & 1 & -1 \end{pmatrix},$$

with rank $s = M_2 - M_1 = M$ in (16) corresponds to the constraints $a_1 = a_2,\ a_3 = a_4, \ldots,\ a_{2M-1} = a_{2M}$ and allows us to compare discretizations for $2M$ and M. The choice $M_2 = M,\ M_1 = 1$ and the $(M-1) \times M$ constraint matrix

$$K = K_{M_2:M_1} = K_{M:1} = \begin{pmatrix} 1 & -1 & 0 & \ldots & \ldots & \ldots & 0 \\ 0 & 1 & -1 & 0 & \ldots & \ldots & 0 \\ \vdots & & & & & & \vdots \\ \vdots & \ldots & \ldots & 0 & 1 & -1 & 0 \\ 0 & \ldots & \ldots & \ldots & 0 & 1 & -1 \end{pmatrix}$$

with rank $s = M - 1$ corresponds to $a_1 = a_2 = a_3 = \cdots = a_M$ and allows us to compare fits with arbitrary piecewise constant death rates $\mu^M(t)$ to fits with constant μ.