Population model for the decline of Homalodisca vitripennis (Hemiptera: Cicadellidae) over a ten-year period

H. T. Banks, John E. Banks, Natalie G. Cody, Mark S. Hoddle and Annabel E. Meade

Center for Research in Scientific Computation, N. C. State University, Raleigh, NC, United States; Department of Entomology, University of California, Riverside, CA, United States; Undergraduate Research Opportunities Center, California State University, Seaside, CA, United States

Abstract
The glassy-winged sharpshooter, Homalodisca vitripennis (Germar), is an invasive pest which presents a major economic threat to grape industries in California, because it spreads a disease-causing bacterium, Xylella fastidiosa. In this note we develop a time and temperature dependent mathematical model to analyze aggregate population data for H. vitripennis from a 10-year study consisting of biweekly monitoring of H. vitripennis populations on unsprayed citrus, during which H. vitripennis decreased significantly. This model was fitted to the aggregate H. vitripennis time series data using iterative reweighted least squares (IRWLS) with assumed probability distributions for certain parameter values. Results indicate that the H. vitripennis model fits the phenological and temperature data reasonably well, but the observed population decrease may possibly be attributed to factors other than the abiotic effect of temperature. A key factor responsible for this decline but not analyzed here could be biotic, for example, potentially parasitism of H. vitripennis eggs by Cosmocomoidea ashmeadi. A biological control program targeting H. vitripennis utilizing the mymarid egg parasitoid Cosmocomoidea ashmeadi (formerly Gonatocerus ashmeadi (Girault)) is described.

Article history
Received 2 July 2018
Accepted 2 May 2019

Keywords
Population model; data fitting; inverse problems; weighted least squares; generalized least squares; biological control; Homalodisca vitripennis; Gonatocerus ashmeadi; Cosmocomoidea ashmeadi; density dependence

2010 Mathematics Subject Classifications
92D25; 34A55

1. Introduction
The glassy-winged sharpshooter (GWSS), Homalodisca vitripennis (Germar) (Hemiptera: Cicadellidae), is a xylem-feeding leafhopper native to the Southeastern United States and Northeastern Mexico [12]. H. vitripennis is a notorious agricultural pest, because it vectors a xylem-dwelling bacterium, Xylella fastidiosa, that causes a variety of often lethal plant diseases (e.g. Pierce’s disease of grapes) [24, 25].

H. vitripennis has exhibited high invasion potential and has become a significant pest outside of its native range. Transportation of live plants infested with H. vitripennis resulted in this pest establishing in Hawaii, the Cook Islands, Easter Island, and in French Polynesia, and its population densities reached extraordinarily high levels [15, 16]. Around 1990, GWSS invaded California and represented a major economic threat to the wine, table, and...
raisin grape industries, because it was implicated in a significant increase of the lethal and incurable grape malady, Pierce’s disease [12, 25]. This disease cost approximately $104 million per year in crop damage and resources devoted to mitigating this threat [26]. The *H. vitripennis* population in CA has exhibited a significant decline. In this note we develop a time and temperature dependent mathematical model to analyze aggregate population data for *H. vitripennis* from a 10-year study (during which *H. vitripennis* decreased significantly) consisting of biweekly monitoring of *H. vitripennis* populations on unsprayed citrus. There are a number of abiotic and biotic factors that can be implicated in regulating GWSS population numbers, including temperature [1, 25], rainfall [17], host plant [14, 22], parasitism [24], etc. - with, for example, *Cosmocomoidea ashmeadi* known to be a primary parasitoid [27].

Female *H. vitripennis* oviposit eggs in ‘mass’ consisting of approximately 10 individual eggs which are deposited side-by-side under the epidermis on the underside of leaves. *C. ashmeadi* is a solitary endoparasitoid that has high levels of host specificity. Female parasitoids lay individual eggs inside a *H. vitripennis* egg. The developing parasitoid larva kills the host egg by consuming the contents. It then pupates within the host egg and emerges as a winged adult capable of flight. *C. ashmeadi* was accidentally introduced into California probably as parasitized *H. vitripennis* eggs that were infesting leaves of imported plant material that originated from the native range of this pest [24]. Since its discovery in California, significant effort has been invested in mass rearing and releasing this natural enemy for the biological control of *H. vitripennis* [19].

In the present initial investigation we develop a mathematical model to investigate whether abiotic density independent factors, such as temperature, could be the primary cause of the observed decrease in *H. vitripennis* populations over time.

### 2. Aggregate population data sets

Commencing 13 March 2002 and ending 8 August 2012, biweekly surveys were made of ten Eureka lemon trees in an unsprayed mixed variety citrus plot in the Agricultural Operations Facility at the University of California Riverside (UCR Ag. Ops). Field surveys were conducted at approximately 1200 to 1300 h for each sample date. Experimental trees were divided into north, south, west, and east quadrants, and visually searched for 30 s each. The total number of adult *H. vitripennis* infesting each citrus tree was recorded. Following visual counts for adults, each tree was searched for five minutes and all *H. vitripennis* egg masses that were found were removed from trees, labelled by tree, and were taken to the lab where the total number of egg masses and individual eggs collected on each sampling date were recorded.

Labeled egg masses were placed in ventilated vials and maintained at 25°C in a temperature cabinet for the emergence of parasitoids. Parasitoid emergence was checked weekly for up to three weeks post collection (this allowed adequate time for parasitoids to complete development and to emerge). The number of parasitoids that emerged was recorded, and sex and species determinations were made. The average realized percent parasitism (total parasitoid emergence divided by the sum of total parasitoid emergence and total GWSS nymph emergence) over all samples was 36%. The samples taken in 2002 had the largest average percent parasitism at 82%, and 2009 had the smallest average percent parasitism
Figure 1. Population phenology data for egg (A) and adult (B) *H. vitripennis*, *H. vitripennis* egg parasitoids reared in the laboratory (C) over March 2002 to August 2012, and (D) the minimum and maximum average monthly temperatures in Riverside, California in citrus at Agricultural Operations, University of California, Riverside.

at 10%. The vast majority of parasitoids that emerged were *C. ashmeadi*, 4714 individuals or 84% of all emerged parasitoids. The second most common parasitoid was *Ufens* sp. (Hymenoptera: Trichogrammatidae) (500 individuals or 9% of emerged parasitoids), followed by *Cosmocomoidea* (formerly *Gonatocerus* morrilli Howard (384 individuals or 7% of emerged parasitoids). These population data are plotted in Figure 1. Until other factors can be ruled out, it is unclear whether lab reared parasitoids can be associated with a primary cause of the 10-year decline of *H. vitripennis* egg and adult populations. The first four years of this data (March 2006 to August 2006) were published in [18], but the full 10-year data set has not yet been published, and none of these data have been mathematically modelled.

It is highly unlikely that the biweekly removal of egg masses and the parasitoid-infected eggs significantly affected either the GWSS population or the parasitoid population. The number of egg masses removed each sampling period was relatively low in comparison to the availability on the trees, especially with the five minute time limit. In addition, only a small part of each experimental tree was sampled, leaving the top of and bottom areas of each tree undisturbed. The experimental Eureka lemon trees were in a large plot of over 100 unsprayed citrus trees that all had high GWSS infestations. *H. vitripennis* adults travel long distances to find optimal host plants [22] and readily fly from tree to tree in this experimental plot, so any removed egg masses would have been readily replaced by GWSS from neighboring trees.
Figure 2. Population phenology data for adult and egg *H. vitripennis* are plotted along with the maximum yearly egg and adult populations, respectively. The average monthly temperatures in Riverside, CA in citrus at Agricultural Operations, University of California, Riverside are plotted along with the maximum monthly average temperature for each year. Average monthly temperatures are calculated by taking the average of the mean of the minimum average monthly temperatures and the maximum average monthly temperatures.

2.1. Temperature data sets

The daily minimum and maximum temperatures collected from 1 January 2002 to 31 December 2012 in Riverside, CA were downloaded from an automated weather station, CIMIS 44, located at UCR Ag. Ops and used to calculate monthly averages (Figure 1(D)). For the mathematical simulations, we assume a daily sinusoidal curve in temperature fluctuation in which the maximum temperature occurred at approximately 1500 h, and the minimum temperature occurred at approximately 0300 h daily (personal observation from data). Average monthly temperatures during the study period are presented with *H. vitripennis* egg and adult population data collected over that same period in Figure 2.

*Homalodisca vitripennis* are strong fliers and travel long distances to find optimal host plants [22]. Thus, it is reasonable to assume that each data point represents a different subsample of the total population, making this an *aggregate* [7] data set (i.e. each sample may possibly consist of differing individuals from the overall population). Consequently, *H. vitripennis* adults, *H. vitripennis* eggs, and the parasitoids were assumed to have been sampled longitudinally from the aggregate population. Therefore, parameters regulating population fluctuation are likely described by some probability distribution (see Chapter 5 in [7]). This is vital and changes the inverse problem when comparing aggregate population data to a mathematical model meant to describe an individual population.

3. Mathematical model

*H. vitripennis* is known to begin life as an egg and undergo five nymphal instars in which they are incapable of sexual reproduction before developing into sexually mature adults. Thus, we use the continuous-time age-structured population growth model described by Kapur [20] to model the *H. vitripennis* individual population. For simplicity, we model the
egg stage, adult stage, and combined nymphal stage in

\[
\begin{align*}
\frac{dx_e}{dt} &= b(T_t-270)x_A - (r_e(T_t-270) + d_e)x_e \\
\frac{dx_N}{dt} &= r_e(T_t-270)x_e - (r_N(T_t-270) + d_N)x_N \\
\frac{dx_A}{dt} &= r_N(T_t-270)x_N - d_A(T_t-270)x_A.
\end{align*}
\]

where \( t \) is time in days and \( T_t-270 \) is temperature in °C at time \( t-270 \) days. The egg and adult time series data in Figures 2 indicates that the \( H. \) vitripennis population in Riverside, California exponentially increases and decreases with temperature, which is known to significantly affect GWSS developmental and reproductive biology [1, 25]. Thus, we use temperature-dependent birth and developmental rates. Eggs, \( x_e \), can die at rate \( d_e \) or hatch and become nymphs at rate \( r_e(T_t-270) \); nymphs, \( x_N \), can die at rate \( d_N \) or develop into sexually mature adults at rate \( r_N(T_t-270) \); and adults, \( x_A \), give birth to more eggs at rate \( b(T_t-270) \) and die at rate \( d_A(T_t-270) \).

The temperature dependent rates, \( b(T_t-270) \), \( r_e(T_t-270) \), \( r_N(T_t-270) \), and \( d_A(T_t-270) \), are calculated using data from Pilkington, Lewis, Jeske, and Hoddle [25]. Developmental rates, \( r_e(T) \) and \( r_N(T) \), are defined by the Lactin model

\[
\begin{align*}
r_e(T) &= e^{\rho_e T} - e^{(\rho_e T_{ue} - T)/\Delta_e} + \lambda_e \\
r_N(T) &= e^{\rho N T} - e^{(\rho N T_{uN} - T)/\Delta_N} + \lambda_N,
\end{align*}
\]

with \( T \) being temperature in °C and parameter values listed in Table 1. Pilkington et al. [25] uses life tables to find the net female maternity, \( R_0(T) \), and mean generation time, \( T_c(T) \).

| \( \theta \) | Definition | Chosen Values | Units | Source |
| --- | --- | --- | --- | --- |
| \( r_e(T) \), \( r_N(T) \) | Developmental rate of eggs/nymphs into nymphs/adults | – | 1/day | [21] |
| \( T_{ue}, T_{uN} \) | Temperature when \( r_e/N(T) = \lambda \) | 37, 39.28 | °C | [25] |
| \( \rho_e, \rho_N \) | Fitted in Logan Model | 0.0073, 0.0022 | – | [25] |
| \( \Delta_e, \Delta_N \) | Fitted in Logan Model | 0.5, 1.99 | – | [25] |
| \( \lambda_e, \lambda_N \) | Determines upper temperature threshold | -1.087, -1.03 | 1/day | [25] |
| \( R_0(T) \) | Number of daughters produced/female | – | egg/adult | [13] |
| \( a_0 \) | Quadratic parameter (constant) | -701.19 | – | [25] |
| \( a_1 \) | Quadratic parameter | 57.48 | – | [25] |
| \( a_2 \) | Quadratic parameter | -1.113 | – | [25] |
| \( T_c(T) \) | Mean generation time | – | day | [13] |
| \( b_0 \) | Quadratic parameter (constant) | -1616.2 | – | [25] |
| \( b_1 \) | Quadratic parameter | 147.92 | – | [25] |
| \( b_2 \) | Quadratic parameter | -3.0264 | – | [25] |
| \( d_e \) | Death rate of eggs | 0.05 | 1/day | [25] |
| \( d_N \) | Death rate of nymphs | 0.02 | 1/day | [25] |
| \( d_A(T) \) | Death rate of adults | – | 1/day | [25] |
| \( \mu_0 \) | Quadratic parameter (constant) | 0.089 | – | [25] |
| \( \mu_1 \) | Quadratic parameter | 0.0067 | – | [25] |
| \( \mu_2 \) | Quadratic parameter | \( 1.33 \times 10^{-4} \) | – | [25] |

Note: Bolded expressions are temperature-dependent rates.
at different temperatures, and fits each of these parameters to a quadratic curve. We let 
\[
b(T) = \frac{R_0(T)}{T_c(T)} = \frac{a_0 + a_1 T + a_2 T^2}{b_0 + b_1 T + b_2 T^2}
\]  
(6)
is the average number of offspring produced per female per day at given temperature \( T \). Finally, temperature-dependent GWSS adult death rate is calculated by inverting the mean adult longevities (emergence as an adult to death) from Pilkington et al. [25] and fitting them to a quadratic curve,

\[
d_A(T) = \mu_0 + \mu_1 T + \mu_2 T^2.
\]  
(7)

Since birth, developmental, and death rates are nonnegative, negative values of \( b(T) \), \( r_e(T) \), \( r_N(T) \), and \( d_A(T) \) are set to zero. The sex ratio of \( H. \ vitripennis \) is assumed approximately even and unaffected by temperature [25].

\( H. \ vitripennis \) have a lack of effective natural enemies, abundant host plants, and no significant competitors in California [16, 17, 23, 25], so we assume that most GWSS deaths can be attributed to unfavourable temperatures. In a controlled laboratory study, \( H. \ vitripennis \) spend on average 5 to 14 days as eggs and 34 to 65 days as nymphs [25]. Thus, we assume that if a GWSS egg has not developed in 20 days then it will die, so we set the GWSS egg death rate to be \( d_e = 1/20 = 0.05 \). Similarly, we assume that if a GWSS nymph has not developed in 50 days then it will die, so we set the GWSS nymph death rate be \( d_N = 1/50 = 0.02 \).

In Figure 2, egg and adult GWSS population data tend to increase and decrease with the temperature, and maximum yearly population counts occur during the hottest parts of the year. If we calculate the temperature-dependent birth and developmental rates, \( r_e(T_t) \), \( r_N(T_t) \), and \( b(T_t) \), at the current daily temperature \( T_t \), then the model populations \( x_e \), \( x_N \), and \( x_A \) begin to increase during the hottest time of the year, and the population counts do not reach their yearly maximum until a few months later. This behaviour, which is presented alongside the data in Figure 3, is not consistent with the data. The average time between the maximum yearly model population and the next maximum yearly data population, as plotted in Figure 3, is 269 days, which we round to 270 days. Thus, temperature-dependent rates \( r_e(T_t-270) \), \( r_N(T_t-270) \), \( b(T_t-270) \), and \( d_A(T_t-270) \) are calculated using temperatures from 9 months or 270 days in the past. With this added delay, the maximum yearly model populations, \( x_e \), \( x_N \), and \( x_A \), occur during the hottest
Figure 3. Population phenology data for the egg and adult *H. vitripennis* populations are plotted along with the maximum yearly egg and adult populations, respectively. Numerical solutions to the GWSS egg (1) and adult (3) population equations are calculated with no added temperature delay; the temperature-dependent rates, \( r(T_t), r_N(T_t) \), and \( b(T_t) \), are calculated using temperatures from the current day. The maximum yearly egg and adult populations from the model are also plotted. The adult death rate parameter \( \mu_0 \) is set to \( \mu_0 = 0.080 \) so that spikes in the population are visible, and other parameter values are taken from Table 1.

Figure 4. Population phenology data for the egg and adult *H. vitripennis* populations are plotted along with the maximum yearly egg and adult populations, respectively. Numerical solutions to the egg (1) and adult (3) population equations (calculated with the temperature delay of 270 days) and the maximum yearly egg and adult populations from the model are also plotted. The adult death rate parameter \( \mu_0 \) is set to \( \mu_0 = 0.012 \) so that spikes in the population are visible, and other parameter values are taken from Table 1.

time of the year. This matches up with the behaviour of the phenology population data presented in Figure 2. We plot the observable model populations \( x_e \) and \( x_A \) with the 270 day temperature delay and the corresponding *H. vitripennis* egg and adult population data in Figure 4.
We have temperature data beginning on 1 January 2002, so with a 270 day delay we cannot model populations earlier than 28 September 2002. Thus, the first fifteen dates in the time series cannot be modelled, bringing the total number of usable *H. vitripennis* egg and adult data points from 544 to 514, which is still sufficient for our purposes (usually 30–40 data points [11] are suggested per parameter being estimated).

4. Aggregate parameter estimation methodology

4.1. Aggregate model and parameter estimation methodology

Equations (1)–(3) model individual populations, where all individual glassy-winged sharpshooters are assumed to have the same growth, birth, and death rates, i.e. the same parameter values from Table 1 that determine these rates. However, this assumption does not apply to the aggregate data which consists of frequent sampling of the changing (over time) populations with possibly different individuals at each time point having different parameter values (e.g. $\rho_e$ or $\lambda_e$ or $a_0$, etc., varying over a range of values $\theta \in \mathcal{G}$). Thus, using the individual model we must formulate an aggregate model with which we can compare the aggregate data (see Chapter 5 of [7] for a more complete discussion).

Consider the 3-dimensional mathematical model or dynamical system,
\[
\frac{d\tilde{x}}{dt}(t) = \tilde{g}(t, \tilde{x}(t), \theta),
\]
\[
\tilde{x}(t_0) = \tilde{x}_0,
\]
where $\tilde{x} = [x_e, x_N, x_A]^T$ as defined in (1), (2), and (3), and $\theta$ is a parameter chosen from Table 1 (e.g. $\theta = \rho_e$ or $\lambda_e$ or $a_0$, etc.). There is an aggregate population vector $\tilde{u} = [u_e, u_N, u_A]^T$ corresponding to the individual population vector $\tilde{x}$ and given by
\[
\tilde{u}(t; P) = \int_{\mathcal{G}} \tilde{x}(t; \theta) \, dP(\theta),
\]
where $P = P(\theta)$ is now a random variable, $\tilde{x}(t; \theta)$ is the solution to (8), $\mathcal{G}$ is the collection of admissible parameter values for $\theta$, and $P$ is a probability measure on $\mathcal{G}$. Note that $\tilde{u}$ is the expected value of $\tilde{x}$, which is also a random vector. Under the assumption that the probability distribution, $P$, possesses a density, $p$, the population count is given by
\[
\tilde{u}(t; P) = \int_{\mathcal{G}} \tilde{x}(t; \theta) p(\theta) \, d\theta,
\]
where the density $P' = dP/d\theta = p(\theta)$.

Now that we have an aggregate model which can be compared to the data, we follow techniques from Banks, Hu, and Thompson [7] and Banks, Bekele-Maxwell, Everett, Stephenson, Shao, and Morgenstern [10] to estimate parameters in our mathematical model. For the inverse problem, consider the 3-dimensional dynamical system, defined in (9) to be estimated using the data. We are interested in determining the probability density $P = p(\theta)$ which gives the best fit of the underlying model to the aggregate data. However, this parameter estimation problem involves an infinite dimensional parameter space (the space $\mathcal{P}(\mathcal{G})$ of probability measures defined on the set $\mathcal{G}$). Instead of
using a specific probability density function in the aggregate model, we use a family of
finite approximations \( P^M(\mathcal{G}) \). Based on [3, 6, 11], we are guaranteed convergence in the
Prohorov metric. Thus, the finite dimensional approximation \( P^M(\mathcal{G}) \) to the probability
measurespace \( P(\mathcal{G}) \) is defined using linear splines given by

\[
P^M(\mathcal{G}) = \left\{ P \in P(\mathcal{G}) \left| \frac{dP}{d\theta} = p(\theta) = \sum_{k=0}^{M} \Theta_k^M \Phi_k^M(\theta), \sum_{k=0}^{M} \Theta_k^M \int_{\mathcal{G}} \Phi_k^M(\theta) d\theta = 1 \right. \right\},
\]

where the chosen and known piecewise linear splines are represented by \( \Phi_k^M(\theta) \), and the
M-dependent (unknown) spline coefficients are \( \Theta_k^M \) for \( k = 0, \ldots, M \). Thus, in order to
estimate \( P \) in equation (9), we only need to estimate the set of \( M+1 \) spline coefficients,

\[
\{ \Theta_0^M, \Theta_1^M, \ldots, \Theta_M^M \}, \tag{10}
\]

where \( P^M \) is the finite approximation of the probability distribution \( P \) and \( \{ \Theta_k^M \}_{k=0}^{M} \) are the
\( M+1 \) spline coefficients of the probability density function which uniquely defines \( P^M \). If
we assume \( \mathcal{G} = [\theta_l, \theta_u] \) is a closed bounded interval of possible values for the parameters, then

\[
\tilde{u}(t; P^M) = \sum_{k=0}^{M} \Theta_k^M \int_{\theta_l}^{\theta_u} \tilde{x}(t; \theta) \Phi_k^M(\theta) d\theta, \tag{11}
\]

is the new approximate 3-dimensional dynamical system to consider in the inverse problem.

Define the \( m \)-dimensional aggregate observation process

\[
\tilde{f}(t; P^M) = C \tilde{u}(t; P^M) \tag{12}
\]

where

\[
C = \begin{bmatrix} 1 & 0 & 0 \\ 0 & 0 & 1 \end{bmatrix} \tag{13}
\]

is an \( m \times 3 \) matrix with \( m = 2 \), since our data set consists of only egg and adult aggregate
populations. Let \( \tilde{u}_j = (u_1^j, u_2^j) \) represent the \textit{H. vitripennis} egg and adult population data
or observations collected at time \( t_j \) for \( j = 1, \ldots, n \). We note that there is some uncertainty
between the actual phenomenon, which is represented through the data, and in the above
observation process. This uncertainty is accounted for in the statistical model (again, see
Chapter 3 of [7] where it is explained that both a mathematical model and a statistical model
are required to carry out an inverse problem properly with uncertainty in observations)

\[
\begin{bmatrix} U_1^j \\ U_2^j \\ \tilde{U}_j \end{bmatrix} = \begin{bmatrix} f_1(t_j; P^M) \\ f_2(t_j; P^M) \\ \tilde{f}(t_j; P^M) \end{bmatrix} + \begin{bmatrix} \tilde{f}^1_j(t_j; P^M_0) \\ \tilde{f}^2_j(t_j; P^M_0) \end{bmatrix} \odot \begin{bmatrix} \epsilon^1_j \\ \epsilon^2_j \end{bmatrix},
\]

where \( P^M_0 \) is the nominal probability density approximation. The values \( \tilde{\gamma} = (\gamma_1, \gamma_2),
\gamma_i \geq 0 \) in the error term are weighting factors representing the dependency of error on the
dynamics, and the model itself corresponds to the choices of data (see Section 4.3 below for further comments on the choice of \( \vec{y} \)).

Note that \( \circ \) is the Hammond or Schur product of component wise multiplication of two vectors. The \( n \times 1 \) random error vectors \( \vec{E}_1 \) and \( \vec{E}_2 \) respectively are assumed to be independent and identically distributed (i.i.d.) with mean zero, \( \text{Var}(\vec{E}_1) = \sigma_{01}^2 \) and \( \text{Var}(\vec{E}_2) = \sigma_{02}^2 \). The corresponding realizations (for the random variables \([U_j^1, U_j^2]^T\)) are

\[
\begin{bmatrix}
    u_j^1 \\
    u_j^2 \\
    \vec{u}_j
\end{bmatrix} = \begin{bmatrix}
    f_1(t_j; \hat{P}_0^M) \\
    f_2(t_j; \hat{P}_0^M) \\
    \vec{f}_j(t_j; \hat{P}_0^M)
\end{bmatrix} + \begin{bmatrix}
    \epsilon_j^1 \\
    \epsilon_j^2 \\
    \epsilon_j
\end{bmatrix}
\]

This multiplicative structure of the observational error in the above statistical model exists, because often in biological models the size of the observation error is proportional to the size of the observations. A rather thorough discussion of these issues along with concrete examples is given in Chapter 3 of [7]. For \( \gamma \geq 0 \), a generalized least squares method or an iterative reweighted weighted least squares (IRWLS) method [10] used below is appropriate to perform the inverse problem. In order to estimate \( \hat{P}^M \approx \hat{P}_0^M \), we want to minimize the distance between the collected data and aggregate mathematical model, where the observables are weighted according to their variability and, for each observable, the observations over time are weighted unequally (once again, we refer the reader to [7, 10] for a more detailed discussion and relevant examples).

The iterative reweighted weighted least squares estimate, \( \hat{P}^M \), is numerically determined by iteratively solving the following system:

\[
\hat{P}^M = \arg \min_{\hat{P}_M \in \mathbb{P}^M(\mathcal{G})} \sum_{j=1}^{n} \left[ \vec{u}_j - \vec{f}(t_j; \hat{P}_M^M) \right]^T \hat{V}_j^{-1} \left[ \vec{u}_j - \vec{f}(t_j; \hat{P}_M^M) \right]
\]  
(14)

\[
\hat{V}_j = \text{diag} \left\{ \frac{1}{n - (M + 1)} \hat{W}_j \left( \sum_{j=1}^{n} \left[ \vec{u}_j - \vec{f}(t_j; \hat{P}_M^M) \right] \left[ \vec{u}_j - \vec{f}(t_j; \hat{P}_M^M) \right]^T \hat{W}_j^{-1} \right)_{ii} \right\}
\]

\[
= \text{diag} \left\{ \frac{1}{n - (M + 1)} \left( f_1(t_j; \hat{P}_M^M)^{2\gamma_1} \sum_{j=1}^{n} \frac{[u_{1,j} - f_1(t_j; \hat{P}_M^M)]^2}{f_1(t_j; \hat{P}_M^M)^{2\gamma_1}} + f_2(t_j; \hat{P}_M^M)^{2\gamma_2} \right) \right\}
\times \sum_{j=1}^{n} \frac{[u_{2,j} - f_2(t_j; \hat{P}_M^M)]^2}{f_2(t_j; \hat{P}_M^M)^{2\gamma_2}}
\]  
(15)

for \( n \) data points. The matrix \( \hat{W}_j = \text{diag}(f_1^{2\gamma_1}(t_j; \hat{P}_M^M), f_2^{2\gamma_2}(t_j; \hat{P}_M^M)) \) is made up of squared error weights, and \( \hat{V}_j \) is the estimated covariance matrix at data point \( j = 1, \ldots, n \). We use the following iterative procedure [5, 7]:

1. Estimate \( \hat{P}^{M(0)} \) using (10) with \( \hat{V}_j = I \). Set \( l = 0 \).
2. Compute weight matrices \( \hat{W}_j^{(l)} = \text{diag}(f_1^{2\gamma_1}(t_j; \hat{P}^{M(l)}), f_2^{2\gamma_2}(t_j; \hat{P}^{M(l)})) \).
3. Solve for \( \hat{V}_j^{(l)} \) using \( \hat{P}^{M(l)} \) and \( \hat{W}_j^{(l)} \) in equation (15).


(4) Estimate $\hat{P}^{(l+1)}$ using $V_j^{(l)}$ in equation (14).
(5) Set $l := l + 1$ and return to step (2). Terminate when two successive estimates for $\hat{P}^M$ are sufficiently close.

Note that this is not the same as taking the derivative of the argument in the right side of (14) and setting it equal to zero. We iteratively want to minimize

$$J(P^M) = \sum_{j=1}^{n} \left[ \bar{u}_j - \hat{f}(t_j; P^M) \right]^T V_j^{-1} \left[ \bar{u}_j - \hat{f}(t_j; P^M) \right].$$

The estimated variances for each observable $\sigma_{01}^2$ and $\sigma_{02}^2$ are approximated by the following:

$$[\hat{\sigma}_{01}^2, \hat{\sigma}_{02}^2] = \frac{1}{n - (M + 1)} \sum_{j=1}^{n} \text{diag}(\hat{V}_j).$$

If we assume $\gamma = (\gamma_1, \gamma_2) = (0, 0)$, then our statistical model is called an absolute error model and an ordinary least squares method is appropriate for parameter estimation. However, we believe it is more biologically realistic to assume the observation error is proportional to the size of the observed quantity i.e. a relative error model for our data sets and models investigated here.

4.2. Uncertainty quantification: standard errors

In order to quantify uncertainty in estimating the linear spline coefficients $\Theta_k^M$ for the aggregate model in equation (11), the standard errors and confidence intervals can be computed using standard asymptotic theory for the IRWLS spline coordinate estimators $\Theta_k^M$ [4, 5, 7]. For each time point $t_j, j = 1, \ldots, n$, the $m \times (M + 1)$ sensitivity matrix for the $i = 1, \ldots, m$ observations is

$$\left( D_j(\hat{P}^M) \right)_{ik} = \frac{\partial f_i(t_j; \hat{P}^M)}{\partial \Theta_k^M} = \int_{\theta_i}^{\theta_u} x_i(t_j; \theta) P_k^M(\theta) d\theta$$

where $x_1 = x_e$ and $x_2 = x_A$, and $f_1 = u_e$ and $f_2 = u_A$ is defined in equations (12). The $(M + 1) \times (M + 1)$ Fisher Information Matrix,

$$\mathcal{F} \approx \left( \sum_{j=1}^{n} D_j^T(\hat{P}^M) \hat{V}_j^{-1} W_j D_j(\hat{P}^M) \right)^{-1},$$

is approximated using the sensitivity matrices $D_j(\hat{P}^M)$, the covariance matrices $\hat{V}_j$ from equation (15), and the matrices of error weights $\hat{W}_j = \text{diag}([f_1^{2\gamma_1}(t_j; \hat{P}^M), f_2^{2\gamma_2}(t_j; \hat{P}^M)])$. The estimated standard errors for each probability density coefficients are

$$\text{SE}(\Theta_k^M) = \sqrt{\mathcal{F}_{kk}^{-1}}.$$
4.3. Difference-based methods and modified residuals

We use a second order difference-based method [9] to determine the correct statistical model ($\vec{\gamma}$ value). Another method often implemented consists of performing an inverse problem with some $\vec{\gamma}$ value and computing the modified residuals,

$$M_i^j = \frac{u_i^j - f_i(t_j; \hat{PM})}{f_i(t_j; \hat{PM})^\gamma_i},$$  \hspace{1cm} (21)

for each observable $i$ at time $t_j$, $j = 1, \ldots, n$. The plots of $M_i^j$ vs. $t_j$ should be randomly scattered around the horizontal axis. If an undesired non-random shape is present (e.g. a fan or inverted fan shape), then a different $\gamma_i$ value is chosen and the process is repeated until a $\gamma_i$ value produces the desired random scatter plot. However, this method does not consider both the mathematical model and statistical model misspecifications; it determines the correct statistical model under the tacit assumption that one has a correct mathematical model. It is also time consuming as it might take several attempts of performing an inverse problem (each of which may be computationally expensive) and plotting the modified residuals until a good statistical model is chosen.

We follow [9] and first apply the second order difference-based method directly to the data to determine the correct $\gamma_i$ value, which is both computationally economical as well as time efficient and independent of any assumed correct mathematical model. We first calculate the following pseudo measurement errors

$$\hat{\epsilon}_j^i = \begin{cases} 
\frac{1}{\sqrt{2}}(u_{j+1}^i - u_j^i) & \text{for } j = 1 \\
\frac{1}{\sqrt{6}}(u_{j-1}^i - 2u_j^i + u_{j+1}^i) & \text{for } j = 2, \ldots, n - 1 \\
\frac{1}{\sqrt{2}}(u_{j-1}^i - u_{j+1}^i) & \text{for } j = n.
\end{cases}$$

for each observable $i$ at time $t_j$, $j = 1, \ldots, n$. Next, we calculate the modified pseudo errors

$$\eta_j^i = \frac{\hat{\epsilon}_j^i}{|u_j^i - \hat{\epsilon}_j^i|^\gamma_i}$$  \hspace{1cm} (22)

for each observable $i$ at time $t_j$, $j = 1, \ldots, n$ for different values of $\gamma_i$ ranging between 0 and 2 and plot these modified pseudo errors versus time. By inspecting these plots, we find that $\vec{\gamma} = (\gamma_1, \gamma_2) = (0.3, 0.6)$ produce the most visually randomized scatter plots (again, see [7] for methodical discussions). The unweighted modified pseudo errors and the modified pseudo errors with the chosen weights $\vec{\gamma} = (\gamma_1, \gamma_2) = (0.3, 0.6)$ are plotted in Figure 5. We use these values in the statistical model to perform the inverse problem and compute modified residuals in (21). If the modified residual plots are not randomly distributed around the horizontal axis, then the error must be due to mathematical model misspecification, implying another iteration of the modelling process is needed. Before we perform the inverse problem, we must select a parameter from Table 1 over which we assume the populations are distributed.
4.4. Local sensitivity analysis for parameter selection

Usually we choose to estimate parameters to which the model observations are most sensitive. However, the aggregate observations, \( f(t; P^M) = [u_e, u_A]^T \) from (12) and (13), depend explicitly on the parameter chosen to estimate. Thus, we perform sensitivity analysis on the individual population model, \( x_e \) and \( x_A \) defined in equations (1) and (3). The local sensitivity of \( f \) to a parameter \( \theta = \hat{\theta} \) is

\[
\chi(\hat{\theta}, t) = \frac{\partial f(t; \hat{\theta})}{\partial \theta}
\]

(23)

where \( t \) is time, and \( f = x_e \) or \( f = x_A \). Many of the parameters have different orders of magnitude, so in order to better compare the sensitivities of these parameters, it is useful to observe the normalized sensitivities,

\[
\chi_n(\hat{\theta}, t) = \frac{\partial f(t; \hat{\theta})}{\partial \theta} \frac{\hat{\theta}}{f(t; \hat{\theta})}.
\]

(24)

The sensitivities depend on our initial parameter value choice, \( \hat{\theta} \). The individual populations \( x_e \) and \( x_A \) die out almost immediately with parameter values from Table 1, making
the sensitivities zero. Thus, in order to observe the sensitivities of a fluctuating population model, we decrease the death rate of GWSS adults by letting $\mu_0 = 0.013$ (originally this value is set at $\mu_0 = 0.089$). At this value, the individual GWSS population model still decreases, but remains nonzero throughout the time period. (See the GWSS egg population $x_e$ plotted in Figure 6.) All other parameters are fixed at values in Table 1.

The individual populations $x_e$ and $x_A$ are temperature-dependent, so we use real temperature data downloaded from an automated weather station, CIMIS 44, located at UCR Ag. Ops in Riverside, CA. We use a fourth order Runge Kutta method and the complex step method [8] to numerically evaluate the sensitivities at the times and temperatures from the data set. The sensitivities of $x_e$ and $x_A$ to parameters $T_{\text{dec}}$, $T_{\text{deg}}$, $T_{\text{wN}}$, $\rho_e$, $\rho_N$, $\lambda_e$, $\lambda_N$, $a_0$, $a_1$, and $a_2$ are positive, because as these parameters increase either the birth rate or one of the developmental rates increases, so the GWSS populations also increase. As these rates increase, the egg and adult populations should also increase. Similarly, the sensitivities of $x_e$ and $x_A$ to parameters $\Delta_e$, $\Delta_N$, $b_0$, $b_1$, and $b_2$ are negative, because as these parameters increase either the birth rate or one of the developmental rates decreases, so the GWSS populations also decrease. The sensitivities of $x_e$ and $x_A$ to the death rate parameters, $d_e$, $d_N$, $\mu_0$, $\mu_1$, and $\mu_2$ are negative, because if these parameters increase, the GWSS populations tend to decrease. The normalized sensitivities are either positive and linearly increasing with time, or negative and linearly decreasing with time on a log scale. We plot the sensitivities and normalized sensitivities of $x_e$ to $\lambda_e$ and $b_2$ in Figure 6 to illustrate this behaviour.
Figure 7. Log plot of maximum absolute sensitivities $\max_t |\chi(\hat{\theta}, t)|$ (A) and normalized sensitivities, $\max_t |\chi_n(\hat{\theta}, t)|$ (B) of the egg observation, $x_e$, and the adult observation, $x_A$, to each parameter. Sensitivities and normalized sensitivities are evaluated at parameter values from Table 1 with $\mu = 0.013$. Maximum sensitivities are in order from largest to smallest.

For each parameter, we find the sensitivity and normalized sensitivity of $x_e$ and $x_A$ with the largest magnitude over time, take the absolute value of each of these values, and plot them in Figure 7. According to these plots, $x_e$ and $x_A$ are most sensitive to the parameter $\rho_N$, $\rho_e$, and $\lambda_N$. The observations have the largest normalized sensitivity to the parameters $a_1$, $\lambda_N$, and $a_2$, followed by the parameters $a_0$, $b_1$, $b_2$, $b_0$, and $\lambda_e$. Thus, we choose the egg developmental rate parameters $\rho_e$ and $\lambda_e$, the nymph developmental rate parameters $\rho_N$ and $\lambda_N$, and the birth rate parameters $a_1$ and $a_2$ to estimate in the inverse problem. Each of the birth rate parameters ($a_0$, $a_1$, $a_2$, $b_0$, $b_1$, $b_2$) have approximately the same effect on the birth rate equation within an appropriately chosen range of values, so the results when estimating $a_0$, $b_0$, $b_1$, and $b_2$ are redundant and not shown.

5. Results
5.1. Estimating the probability density of egg developmental rate parameters

First, we assume a probability distribution over parameter $\rho_e$, which is part of the egg developmental rate equation. At different closed intervals the parameter $\rho_e$ still converges and exhibits the same behaviour, so for these results we choose to let $\rho_e$ vary over the interval $[0.001, 0.060]$ which permits convergence to a centred value of the distribution for $\rho_e$. Next, we hold $\rho_e$ fixed and assume a probability distribution over parameter $\lambda_e$, which is
Figure 8. Temperature-dependent developmental rate of *H. vitripennis* eggs from equation (4) for different values of $\rho_e$ (A) and $\lambda_e$ (B). Note that $\rho_e$ and $\lambda_e$ are parameters in the egg developmental rate equation. Parameter values not stated in the plot legends are taken from Table 1.

Also part of the egg developmental rate equation. At different closed intervals the parameter $\lambda_e$ still converges and exhibits the same behaviour, so we let $\lambda_e$ vary over the interval $[-1.1, 1.0]$ (Figure 8).

Figure 9. Estimated probability density coefficients of the parameter $\rho_e$ (A) for $M = 20, 24, 28, \text{ and } 32$ linear spline functions; corresponding *H. vitripennis* egg (B) and adult (C) aggregate model observations at the estimated probability density with $M = 32$ linear spline functions plotted alongside the population phenology data; the nymph (D) aggregate model solution at the estimated probability density with $M = 32$ linear spline function (there is no GWSS nymph data); and modified residual errors for the egg (E) and adult (F) aggregate model observations.
We run the two separate inverse problems with $M = 20, 24, 28,$ and 32 spline functions to estimate the probability distributions of $\rho_e$ and $\lambda_e$. The resulting probability densities of each inverse problem are plotted in Figures 9 and 10 along with the model simulations of the aggregate egg $u_e$, nymph $u_N$, and adult $u_A$ GWSS populations and the egg and adult modified residual errors for $M = 32$. The model simulations and residual errors when $M = 20, 24,$ and 28 are not graphed, because they look very similar to the solutions when $M = 32$. The optimal cost value $J^*$ from equation (16), the estimated variances of the eggs and adults from equation (17), and the condition number of the Fisher Information Matrix from equation (19) are listed in Table 2.

The estimated probability distributions of $\rho_e$ are clustered around $\rho_e \approx 0.045$, and the estimated probability distributions of $\lambda_e$ are centred around $\lambda_e \approx 0.4$. These estimates are much larger than the parameter values originally chosen and listed in Table 1 of $\rho_e = 0.0073$ and $\lambda_e = -1.087$. In both cases, the probability distributions are multimodal, and the estimated variances of the GWSS eggs listed in Table 2 are much larger and different from the results of the other inverse problems (see Tables 3 and 4). Thus, the results of these inverse problems are more unstable. Furthermore, it usually takes a GWSS egg at least 5 days to develop [25], but at the estimated values, the egg developmental time is as short as 12 h at optimal temperatures.
Table 2. Optimal cost value ($J^*$), estimated variances for eggs and adults respectively ($\hat{\sigma}^2_{01}$ and $\hat{\sigma}^2_{02}$), and the condition number of the Fisher Information Matrix, $\mathcal{F}$ when computing the inverse problem for $\rho_e$ and when computing the inverse problem for $\lambda_e$.

| Parameter | M   | $J^*$ | $\hat{\sigma}^2_{01}$ | $\hat{\sigma}^2_{02}$ | $\kappa (\mathcal{F})$ |
|-----------|-----|-------|------------------------|------------------------|------------------------|
| $\rho_e$  | 20  | 470   | $1.5315 \times 10^{14}$ | $2.2941 \times 10^6$  | $3.2142 \times 10^7$  |
|           | 24  | 462   | $8.9131 \times 10^6$  | $2.2951 \times 10^6$  | $1.4175 \times 10^7$  |
|           | 28  | 454   | $1.0975 \times 10^7$  | $2.4357 \times 10^6$  | $9.6961 \times 10^7$  |
|           | 32  | 446   | $1.2252 \times 10^7$  | $2.4755 \times 10^6$  | $1.7532 \times 10^7$  |
| $\lambda_e$ | 20  | 470   | $2.4062 \times 10^7$  | $2.2850 \times 10^6$  | $5.6114 \times 10^{10}$|
|           | 24  | 462   | $2.0745 \times 10^7$  | $2.2936 \times 10^6$  | $7.3654 \times 10^{10}$|
|           | 28  | 454   | $3.1495 \times 10^7$  | $2.4443 \times 10^6$  | $1.2657 \times 10^{11}$|
|           | 32  | 446   | $1.4986 \times 10^{14}$ | $2.5269 \times 10^6$  | $4.1252 \times 10^{11}$|

Table 3. Optimal cost value ($J^*$), estimated variances for eggs and adults respectively ($\hat{\sigma}^2_{01}$ and $\hat{\sigma}^2_{02}$), and the condition number of the Fisher Information Matrix, $\mathcal{F}$ when computing the inverse problem for $\rho_N$ and when computing the inverse problem for $\lambda_N$.

| Parameter | M   | $J^*$ | $\hat{\sigma}^2_{01}$ | $\hat{\sigma}^2_{02}$ | $\kappa (\mathcal{F})$ |
|-----------|-----|-------|------------------------|------------------------|------------------------|
| $\rho_N$  | 20  | 470   | $7.4063 \times 10^6$  | $2.3947 \times 10^6$  | $7.1455 \times 10^{10}$|
|           | 24  | 462   | $7.5147 \times 10^6$  | $2.4165 \times 10^6$  | $1.4790 \times 10^{11}$|
|           | 28  | 454   | $7.9515 \times 10^6$  | $2.6431 \times 10^6$  | $1.1532 \times 10^{11}$|
|           | 32  | 446   | $8.0000 \times 10^6$  | $2.6118 \times 10^6$  | $1.6746 \times 10^{11}$|
| $\lambda_N$ | 20  | 470   | $7.3988 \times 10^6$  | $2.2458 \times 10^6$  | $9.1025 \times 10^{10}$|
|           | 24  | 462   | $7.5104 \times 10^6$  | $2.2779 \times 10^6$  | $8.2667 \times 10^{10}$|
|           | 28  | 454   | $7.8326 \times 10^6$  | $2.4395 \times 10^6$  | $6.8128 \times 10^{11}$|
|           | 32  | 446   | $7.8712 \times 10^6$  | $2.4081 \times 10^6$  | $8.3738 \times 10^{11}$|

Table 4. Optimal cost value ($J^*$), estimated variances for eggs and adults respectively ($\hat{\sigma}^2_{01}$ and $\hat{\sigma}^2_{02}$), and the condition number of the Fisher Information Matrix, $\mathcal{F}$ when computing the inverse problem for $a_1$ and when computing the inverse problem for $a_2$, respectively.

| Parameter | M   | $J^*$ | $\hat{\sigma}^2_{01}$ | $\hat{\sigma}^2_{02}$ | $\kappa (\mathcal{F})$ |
|-----------|-----|-------|------------------------|------------------------|------------------------|
| $a_1$     | 20  | 470   | $7.4784 \times 10^6$  | $2.6766 \times 10^6$  | $5.3184 \times 10^{12}$|
|           | 24  | 462   | $7.9731 \times 10^6$  | $2.8816 \times 10^6$  | $6.5348 \times 10^{12}$|
|           | 28  | 454   | $7.9557 \times 10^6$  | $2.9131 \times 10^6$  | $1.0785 \times 10^{12}$|
|           | 32  | 446   | $8.4654 \times 10^6$  | $3.1037 \times 10^6$  | $1.2305 \times 10^{12}$|
| $a_2$     | 20  | 470   | $7.3870 \times 10^6$  | $2.6407 \times 10^6$  | $1.1917 \times 10^{11}$|
|           | 24  | 462   | $8.1032 \times 10^6$  | $3.0002 \times 10^6$  | $1.1224 \times 10^{11}$|
|           | 28  | 454   | $7.9753 \times 10^6$  | $2.9774 \times 10^6$  | $9.6515 \times 10^{10}$|
|           | 32  | 446   | $8.5984 \times 10^6$  | $3.2434 \times 10^6$  | $2.0637 \times 10^{11}$|

5.2. Estimating the probability density of nymph developmental rate parameters

We assume a probability distribution over parameter $\rho_N$, which is part of the nymph developmental rate equation. At different closed intervals the parameter $\rho_N$ still converges and exhibits the same behaviour, so we arbitrarily let $\rho_N$ vary over the interval [0.002, 0.005]. Next, we hold $\rho_N$ fixed and assume a probability distribution over parameter $\lambda_N$, which is also part of the nymph developmental rate equation. At different closed intervals the parameter $\lambda_N$ still converges and exhibits the same behaviour, so we arbitrarily let $\lambda_N$ vary over the interval [−1.04, −0.96]. The temperature-dependent nymph developmental rate is plotted in Figure 11 for different values of $\rho_N$ and $\lambda_N$, respectively.

We run the two separate inverse problems with $M = 20, 24, 28, \text{ and } 32$ spline functions to estimate the probability distributions of $\rho_N$ and $\lambda_N$. The resulting probability densities
Figure 11. Temperature-dependent developmental rate of *H. vitripennis* nymphs from equation (5) for different values of $\rho_N$ (A) and $\lambda_N$ (B). Note that $\rho_N$ and $\lambda_N$ are parameters in the nymph developmental rate equation. Parameter values not stated in the plot legends are taken from Table 1.

of each inverse problem are plotted in Figures 12 and 13 along with the model simulations of the aggregate egg $u_e$, nymph $u_N$, and adult $u_A$ GWSS populations and the egg and adult modified residual errors for $M = 32$. The model simulations and residual errors when $M = 20, 24, \text{ and } 28$ are not graphed, because they look very similar to the solutions when

Figure 12. Estimated probability density coefficients of the parameter $\rho_N$ (A) for $M = 20, 24, 28, \text{ and } 32$ linear spline functions; corresponding *H. vitripennis* egg (B) and adult (C) aggregate model observations at the estimated probability density with $M = 32$ linear spline functions plotted alongside the population phenology data; the nymph (D) aggregate model solution at the estimated probability density with $M = 32$ linear spline function (there is no GWSS nymph data); and modified residual errors for the egg (E) and adult (F) aggregate model observations.
5.3. Estimating the probability density of birth rate parameters

Next, we assume that the glassy-winged sharpshooter eggs and adults have a probability distribution over the parameters $a_1$ and $a_2$, which are part of the birth rate equation. We let $a_1$ vary over the interval $[57.4, 58.2]$ while fixing $a_2$, and then we let $a_2$ vary over the interval $[-1.14, -1.08]$ while fixing $a_1$. At different closed intervals the parameters $a_1$ and $a_2$ still converge and exhibit the same behaviour, so we choose these intervals arbitrarily. The temperature-dependent birth rate is plotted in Figure 14 for different values of $a_1$ and $a_2$, respectively. Since these two parameters will be estimated separately and independently,
Figure 14. Temperature-dependent birth rate of *H. vitripennis* from equation (6) for different values of $a_1$ (A) and $a_2$ (B). Note that $a_1$ and $a_2$ are parameters in the birth rate equation. Parameter values not stated in the plot legends are taken from Table 1.

whenever $a_1$ is varied, $a_2$ is held fixed at its value from Table 1, and whenever $a_2$ is varied, $a_1$ is held fixed at its value from Table 1.

We run the two separate inverse problems with $M = 20, 24, 28,$ and $32$ spline functions to estimate the probability distributions of $a_1$ and $a_2$. The resulting probability densities

![Estimated Probability Distribution of $a_1$ and Resulting Aggregate Model Solutions](image)

Figure 15. Estimated probability density coefficients of the parameter $a_1$ (A) for $M = 20, 24, 28,$ and $32$ linear spline functions; corresponding *H. vitripennis* egg (B) and adult (C) aggregate model observations at the estimated probability density with $M = 32$ linear spline functions plotted alongside the population phenology data; the nymph (D) aggregate model solution at the estimated probability density with $M = 32$ linear spline function (there is no GWSS nymph data); and modified residual errors for the egg (E) and adult (F) aggregate model observations.
Figure 16. Estimated probability density coefficients of the parameter $a_2$ (A) for $M = 20, 24, 28, \text{and } 32$ linear spline functions; corresponding $H. \text{vitripennis}$ egg (B) and adult (C) aggregate model observations at the estimated probability density with $M = 32$ linear spline functions plotted alongside the population phenology data; the nymph (D) aggregate model solution at the estimated probability density with $M = 32$ linear spline function (there is no GWSS nymph data); and modified residual errors for the egg (E) and adult (F) aggregate model observations.

of each inverse problem are plotted in Figures 15 and 16 along with the model simulations of the aggregate egg $u_e$, nymph $u_N$, and adult $u_A$ GWSS populations and the egg and adult modified residual errors for $M = 32$. The model simulations and residual errors when $M = 20, 24, \text{and } 28$ are not graphed, because they look very similar to the solutions when $M = 32$. The optimal cost value $J^*$ from equation (16), the estimated variances of the eggs and adults from equation (17), and the condition number of the Fisher Information Matrix from equation (19) are listed in Table 4. The estimated probability distributions of $a_1$ are clustered around $a_1 \approx 58.05$, and the estimated probability distributions of $a_2$ are centred around $a_2 \approx -1.09$. The probability distributions in both cases appear to be approaching a single value for both parameters. Based on data from [25], GWSS adults lay, on average, one egg per five days, but at the estimated values GWSS lay as many as 5 eggs per day at optimal temperatures.

6. Discussion

The goal of this study is to understand what led to the $H. \text{vitripennis}$ population decrease during a 10-year study. Using only the biotic effect of temperature, our results are stable and appear to follow the trends in the data reasonably well. Although assuming a probability distribution over the estimated parameters was necessary due to the aggregate nature of the data, most of the parameters appear to approach a single value as $M$, the number of
linear spline functions used to estimate the probability density, approaches infinity. While we do not form a ‘best value’ of $M$ to use in our approximations of the probability measures, we found consistently that $M = 32$ leads to decreasing values of the weighted least squares measure $J^*$ while at the same time leads to only slight increases in the ill-conditioning of the Fisher Information Matrix $\kappa(F)$, suggesting a ‘sweet spot’ for the reported values of the parameters.

We obtain inferior results when we estimate the egg developmental rate parameters $\rho_e$ and $\lambda_e$, chiefly when fitting the GWSS egg population model to the data. The aggregate egg populations are estimated to be zero (see Figures 9(B–C)), and the egg observation variance is estimated to be $\sigma^2_1 \approx 2 \times 10^7$ which is much larger compared to the other inverse problems ($\sigma^2_1 \approx 8 \times 10^6$). Furthermore, while other estimated probability distributions are centred around a single value, the probability distributions of the egg developmental rate parameters are spread over a large range of values ($0.03 < \rho_e < 0.06$ and $-0.1 < \lambda_e < 1.0$) without any signs of convergence as $M$ increases. Thus, the results when estimate the egg developmental rate appear to be more uncertain than when estimating the nymphal development rate and the birth rate.

We obtain much better results when estimating birth rate parameters, and we obtain the best results when estimating the nymphal developmental rates. The observation variances ($\sigma^2_1 \approx 7.8 \times 10^6$ for eggs and $\sigma^2_1 \approx 2.5 \times 10^6$ for adults) when estimating the nymphal developmental rate parameters are smaller than when estimating birth rate parameters ($\sigma^2_1 \approx 8 \times 10^6$ for eggs and $\sigma^2_1 \approx 3 \times 10^6$ for adults), because more variance is captured in the model. Furthermore, the probability distributions of the nymphal developmental rate parameters $\lambda_N$ and $\rho_N$ exhibit consistent unimodal behaviour as $M$ increases (see Figures 12(A) and 13(A)), while the probability distributions of the birth rate parameters have several tightly clustered modes (see Figures 15(A) and 16(A)).

In all cases, the estimated parameters are larger than the parameter values originally estimated in controlled lab experiments [25] (see Table 1), indicating that this population of $H. vitripennis$ have exceptionally high developmental and birth rates. In the non-aggregate model described in (1)–(3), using all of the chosen parameter values from Table 1 and real temperature data, the GWSS popoulations die out almost immediately. Thus, in order to fit the nonzero data, it follows that the inverse problem estimates this population to have larger parameter values and more robust birth and developmental rates. In other words, this population of $H. vitripennis$ is laying more eggs, developing more quickly from eggs to nymphs and nymphs to adults, or some combination of these than we anticipated. This does not necessarily rule out parasitism as the cause of the GWSS decline, but there is some disconnect between GWSS populations reared in the lab at constant temperatures and GWSS populations in the wild. In a study conducted by Bahar, Soroka, and Dosdall [2], populations of the insect $Plutella xylostella$ had shorted developmental times and higher adult longevity compared with populations reared under constant temperatures. Thus, there may be similar difference between lab-reared GWSS populations from which initial parameter values were gathered, and GWSS in the wild.

7. Conclusion

This study takes a first look at aggregate population phenology data for $H. vitripennis$ collected over the course of 10 years at the Agricultural Operations Facility at the University of
California Riverside. The decline in *H. vitripennis* population observed in this data could potentially be attributed to many factors. The abiotic effect of temperature is investigated here and does a reasonable job describing the data. The phenology data fit the model best when the birth and developmental rates are much larger than originally estimated in lab studies.

For our future work, we intend to add the effect of parasitism to the mathematical model, since this population of *H. vitripennis* are known to have been attacked by *C. ashmeadi*. We intend to analyze the parasitoid and GWSS data together, and test for statistically significantly better fits to our data.

**Disclosure statement**

No potential conflict of interest was reported by the authors.

**Funding**

This research was supported in part by the Air Force Office of Scientific Research (HTB) under grant number AFOSR FA9550-15-1-0298, in part by the National Science Foundation (AEM) under grant number RTG/DMS - 1246991. The GWSS survey work was supported, in part, by funds awarded to MSH by the California Department of Food and Agriculture Pierce's Disease Research Program and the University of California Pierce's Disease Research Program.

**References**

[1] A.K. Al-Wahaibi and J.G. Morse, *Homalodisca coagulata* (*Hemiptera: Cicadellidae*) embryonic development at constant temperatures, Fla. Entomol. 86(4) (2003), pp. 477–478.

[2] M.H. Bahar, J.J. Soroka, and L.M. Dox dall, Constant versus fluctuating temperatures in the interactions between *Plutella xylostella* (*Lepidoptera: Plutellidae*) and its larval parasitoid *Diadegma insulare* (*Hymenoptera: Ichneumonidae*), Environ. Entomol. 41(6) (2012), pp. 1653–1661.

[3] H.T. Banks and K.L. Bihari, Modelling and estimating uncertainty in parameter estimation, Inverse Probl. 17 (2001), pp. 95–111.

[4] H.T. Banks and J.L. Davis, Quantifying uncertainty in the estimation of probability distributions with confidence bands, Math. Biosci. Eng. 5(4) (2008), pp. 647–667.

[5] H.T. Banks and H.T. Tran, *Mathematical and Experimental Modeling of Physical and Biological Processes*, CRC Press, Chapman & Hall, Boca Raton, FL, 2009.

[6] H.T. Banks, D.M. Bortz, G.A. Pinter. Modeling and imaging techniques with potential for application in bioterrorism, in *Bioterrorism: Mathematical Modeling Applications in Homeland Security*, SIAM, Philadelphia, PA, 2003, pp. 129–154.

[7] H.T. Banks, S. Hu, and W.C. Thompson, *Modeling and Inverse Problems in the Presence of Uncertainty*, CRC Press, Chapman & Hall, Boca Raton, FL, 2014.

[8] H.T. Banks, L. Bekele-Maxwell, L. Bociu, M. Noorman, and K. Tillman, The complex-step method for sensitivity analysis of non-smooth problems arising in biology, Eurasian J. Math. Comput. Appl. 3 (2015), pp. 16–68.

[9] H.T. Banks, J. Catenacci, and S. Hu, Use of difference-based methods to explore statistical and mathematical model discrepancy in inverse problems, J. Inverse Ill-posed Probl. 24(4) (2016), pp. 413–433.

[10] H.T. Banks, K. Bekele-Maxwell, R.A. Everett, L. Stephenson, S. Shao, and J. Morgenstern, Dynamic modeling of problem drinkers undergoing behavioral treatment, Bull. Math. Biol. 79(6) (2017), pp. 1254–1273.

[11] H.T. Banks, K.B. Flores, I.G. Rosen, E.M. Rutter, M. Sirlanci, and W. Clayton, *Thompson. The prohorov metric framework and aggregate data inverse problems for random PDEs*, Commun. Appl. Anal. 22(3) (2018), pp. 415–446.
[12] California Department of Food and Agriculture, Pierce’s disease control program, 2018. Available at http://www.cdfa.ca.gov/pdcp/. Online; accessed 9 March 2018.

[13] J.R. Carey, *The multiple decrement life table: A unifying framework for cause-of-death analysis in ecology*, Oecologia 78(1) (1989), pp. 131–137.

[14] W. Chen, R.A. Leopold, and M.A. Boetel, *Host plant effects on development and reproduction of the glassy-winged sharpshooter, Homalodisca vitripennis (Homoptera: Cicadellidae)*, Environ. Entomol. 39(5) (2010), pp. 1545–1553.

[15] J. Grandgirard, M.S. Hoddle, J.N. Petit, G.K. Roderick, and N. Davies, *Engineering an invasion: Classical biological control of the glassy-winged sharpshooter, homalodisca vitripennis, by the egg parasitoid gonatocerus ashmeadi in Tahiti and Moorea, French Polynesia*, Biol. Invasions 10(2) (2008), pp. 135–148.

[16] J. Grandgirard, M.S. Hoddle, J.N. Petit, G.K. Roderick, and N. Davies, *Classical biological control of the glassy-winged sharpshooter, homalodisca vitripennis, by the egg parasitoid gonatocerus ashmeadi in the society, Marquesas and Austral archipelagos of French Polynesia*, Biol. Control 48(2) (2009), pp. 155–163.

[17] M.S. Hoddle, *The potential adventive geographic range of glassy-winged sharpshooter, Homalodisca coagulata and the grape pathogen Xylella fastidiosa: Implications for California and other grape growing regions of the world*, Crop Prot. 23(8) (2004), pp. 691–699.

[18] M.S. Hoddle, *Are glassy-winged sharpshooter populations regulated in California? Long-term phenological studies and construction of multi-cohort life tables for the glassy-winged sharpshooter in citrus*, California Department of Food and Agriculture Symposium Proceedings, 2006 Pierce’s Disease Research Symposium. November 27–29, 2006, San Diego, California, pp.73–75.

[19] N. Irvin, J. Suarez-Espinoza, and M.S. Hoddle, *Functional response of Gonatocerus ashmeadi and the “new association” parasitoid G. tuberculifemur attacking eggs of Omalodiscas vitripennis*, Environ. Entomol. 38(6) (2009), pp. 1634–1641.

[20] J.N. Kapur, *A continuous time age structured population growth model*, Appl. Math. Model. 3(6) (1979), pp. 455–458.

[21] D.J. Lactin, N.J. Holliday, D.L. Johnson, and R. Craigen, *Improved rate model of temperature-dependent development by arthropods*, Environ. Entomol. 24(1) (1995), pp. 68–75.

[22] R.F. Mizell, C. Tipping, P.C. Andersen, B.V. Brodbeck, W.B. Hunter, and T. Northfield, *Behavioral model for Homalodisca vitripennis (Homoptera: Cicadellidae): Optimization of host plant utilization and management implications*, Environ. Entomol. 37(5) (2008), pp. 1049–1062.

[23] J. Petit, M.S. Hoddle, J. Grandgirard, G.K. Roderick, and N. Davies, *Successful spread of a biocontrol agent reveals a biosecurity failure: Elucidating long distance invagic pathways for the Gonatocerus ashmeadi in French Polynesia*, BioControl 54(4) (2009), pp. 485–495.

[24] L.J. Pilkington and M.S. Hoddle, *Reproductive and developmental biology of Gonatocerus ashmeadi (Hymenoptera: Mymaridae), an egg parasitoid of Homalodisca coagulata (Homoptera: Cicadellidae)*, Biol. Control 37(3) (2006), pp. 266–275.

[25] L.J. Pilkington, M. Lewis, D. Jeske, and M.S. Hoddle, *Calculation and thematic mapping of demographic parameters for Homalodisca vitripennis (Homoptera: Cicadellidae) in California*, Ann. Entomol. Soc. America 107(2) (2014), pp. 424–434.

[26] K. Tumber, J. Alston, and K. Fuller, *Pierce’s disease costs California $104 million per year*, Calif. Agric. 68(1) (2014), pp. 20–29.

[27] D.B. Vickerman, M.S. Hoddle, S. Triapitsyn, and R. Stouthamer, *Species identity of geographically distinct populations of the glassy-winged sharpshooter parasitoid Gonatocerus ashmeadi: Morphology, DNA sequences, and reproductive compatibility*, Biol. Control 31(3) (2004), pp. 338–345.