Title
Individual heterogeneity in life history processes: Estimation and applications of demographic models to stage-structured arthropod populations

Author
Scranton, Katherine

Publication Date
2012

Peer reviewed|Thesis/dissertation

Individual heterogeneity in life history processes: Estimation and applications of demographic models to stage-structured arthropod populations

By

Katherine Scranton

A dissertation submitted in partial satisfaction of the

requirements for the degree of

Doctor of Philosophy

in

Environmental Science, Policy, and Management

in the

Graduate Division

of the

University of California, Berkeley

Committee in charge:

Associate Professor Perry de Valpine, Chair
Professor Wayne Getz
Professor Mary Powers

Fall 2012

Abstract

Individual heterogeneity in life history processes: Estimation and applications of demographic models to stage-structured arthropod populations

By

Katherine Scranton

Doctor of Philosophy in Environmental Science, Policy, and Management

University of California, Berkeley

Associate Professor Perry de Valpine, Chair

Life history variation is a general feature of natural populations. Most studies assume that local processes occur identically across individuals, ignoring any genetic or phenotypic variation in life history traits. In part, this is because a realistic treatment of individual heterogeneity results in very complex population models. Fitting models with individual heterogeneity to real data is further complicated by random effects in groups of the data, observations set at specific intervals, and the non-independence of data following a cohort of individuals through time. In this dissertation, I assume that individuals differ in the duration they spend in each developmental stage and also in the amount of time they live. Stage durations and survival times follow probability distributions with parameters specific to populations and stages. Parameters of these distributions may also include random effects when considering a subset of sampled populations and covariates such as temperature. In the first chapter I formulate a model and likelihood for variable development, using the time-to-event model framework. In the second chapter I use this model to ask whether field populations of herbivorous arthropods (*Tetranychus pacificus*) form host-associations on different cultivars of the same host species. In the third chapter I incorporate variable development with variable survival and ongoing reproduction in a stage-structured population model. I explore the ability of the approximate Bayesian computation framework to fit such a complex model to data, evaluating posterior distributions and model performance.

I first focus on variable development in a stage-structured population. I consider the distribution of maturation times (time from birth to adult) as time-to-event data, which are very common in other fields such as survival analysis, machine failure, and disease onset. Time-to-event data are also common in ecology but are rarely analyzed with sophisticated tools common in other fields. The main obstacles in applying time-to-event models to ecological data are the variation of natural systems and interval-censored data. I develop a mixed-effects Weibull model for interval-censored data on time-to-maturation of individuals in a cohort. I incorporate a fixed difference between types of cohorts and two levels of random effects. There is no available software for mixed-effects survival analysis for interval-censored data, so likelihood calculations with numerical integration of random

effects were programmed in R. A simulation study under different magnitudes of variation is used to evaluate the power of the likelihood ratio tests and the precision of parameter estimates. Differences in the mean scale as small as 5% can be detected with high power under low variation. Under higher variation, larger differences of 18% can be detected with 80% power. Omitting random-effects produces biased estimates of the Weibull parameters and highly inflated type I error rates in likelihood ratio tests. The methods developed in this paper for fitting hierarchical frailty models to interval censored data would be applicable to a wide range of ecological processes such as survival, oviposition, or onset of disease.

I apply the time-to-maturation model in a common garden experiment with Pacific spider mites (*Tetranychus pacificus*) to detect any significant intraspecific life history variation associated with the cultivar of original host plant (*Vitis vinifera*). Tetranychid mites in particular display large amounts of life history variation. They are prone to host-associated differentiation and forming host races on different host species, but the strength of their association with cultivar is less clear. To address this question I collect individuals from many field sites of two cultivars of grape, Zinfandel and Chardonnay. I then conduct a "common garden experiment" with bioassays of mites on bean plants (*Phaseolus lunatus*) in the laboratory. Assay populations are founded and sampled non-destructively with digital photography over 12 days to determine development times, survival times, and fecundity rates of individuals in single cohorts. Two classes of models are fit to the data: standard generalized linear mixed models and the time-to-event model for development. The time-to-event model allows for interval-censored data and random effects for the replicate sample sites and for the replicate assay populations. I employ a maximum likelihood approach to fit the time-to-event model and perform all likelihood ratio tests with randomized null distributions because of sample size concerns. Both models find a significant effect of cultivar on development time: individuals from Zinfandel grapes develop more quickly than those from Chardonnay grapes. There is a trend towards higher juvenile survival in populations from Zinfandel grapes and no difference in fecundity. I show that the time-to-event model provides more information about life history processes than standard GLMMs. Differences in development time in these lab populations may indicate field differences of these economically-important pests on grape crops. Growers and pest managers should consider the possibility of population-level differences on crops that are genetically similar and geographically close.

Including a time-to-event model of development in a stage-structured population model of ongoing dynamics increases the complexity of the model. Individual heterogeneity in development and survival, coupled with ongoing reproduction causes intractable likelihoods, making it impossible to use likelihood-based model inference. We lack a statistical framework flexible enough to analyze and fit such a stage-strucutred model to cohort data. The problem is further complicated by the non-independence caused by following the same individuals through time, with interval censoring. A potential solution is to use an approximate Bayesian computation (ABC) approach to model inference, replacing the likelihood calculations by comparing repeated simulations with an observed data set. I use the ABC framework to fit the stochastic stage-structured model to simulated cohort data with a sequential Monte Carlo (SMC) sampler to increase efficiency.

The computational methods are all programmed directly in R and C. I demonstrate a systematic way to select summary statistics and distance metrics using simulated data and the ROC curve from classification theory. I evaluate the performance of the ABC SMC algorithm, showing the posterior distributions of parameters and 95% credible intervals. I also show the computational performance of the algorithm in moving to the target posterior. I conclude that ABC shows great promise as a framework for parameter estimation and model inference of realistic stage-structured population models. My approach could be extended to include model selection, covariates or random effects on any of the demographic parameters, correlations between distributions of stage durations, or imperfect detection of individuals.

I dedicate this manuscript to my family. Especially to my mother, Dr. Joan Gluch, and her excellent metaphor for the strategy of writing a dissertation:
"What is the best way to eat an elephant? A spoonful at a time."

Acknowledgements

Foremost, I would like to thank my adviser Perry de Valpine for his unwavering support and inspiration. I would also like to thank my dissertation committee members Wayne Getz and Mary Powers; every conversation we have had has expanded my knowledge, exposed me to a new viewpoint, and made me a better researcher. I would also like to thank Nick Mills for his generosity with his time and support as I made my way in the world of entomology. Along with Nick Mills, I would also like to thank Menelaos Stavrinides for their expertise and patience in introducing me to the spider mite system in California. Thanks to the rest of the Mills lab especially Aviva Goldmann, Kevi Mace, and Kevin Yao for their support and assistance in the rearing of laboratory populations. The pest control managers in Lodi, including Larry Whitted, Charlie Starr, and Chris Storm were an invaluable resource in field collections.

I would also like to thank Jonas Knape for his assistance in formulating all of the population models, his expertise in C, and his assistance in implementing the sequential Monte Carlo sampler in R. Thanks to Karthik Ram and the rest of the de Valpine lab for helpful discussions about implementing the models and in preparing the manuscripts.

Lastly I would like to thank my amazing support network of friends and family. Your patience on the nights when I put my work first, your kindness when I cried out of stress or frustration, and your joy at my accomplishments has meant the world to me.

Contents

List of Figures

List of Tables

1 Chapter 1

1.1 Introduction

Ecological studies commonly yield time-to-event data, such as an individuals survival time, maturation time, or behavior durations (Zens and Peart, 2003; Ma and Bechinski, 2009; Wajnberg, 2006). A class of models, collectively known as Survival Analysis models or Failure Time models, have been developed in other fields for similar data, such as time to machine failure, disease onset, or time to death (Wong et al., 2005; Lindsey and Ryan, 1998). When exact event times are known, there are numerous parametric, semi-parametric and non-parametric approaches to model fitting and hypothesis testing (Hougaard, 2000). However, in practice an individual is rarely continuously monitored, but observed at regular times (known as "interval censoring") producing a range in which we know the event has occurred. Similar approaches can be taken to fit models to interval-censored data, although the calculations can become more complex (Goethals et al., 2009). Wider use of these methods in ecology would allow more accurate estimation of any time-to-event process from groups, giving some estimate of the amount of individual heterogeneity in the population.

One basic concept of survival analysis is that the time-to-event, T, is treated as a random variable with a probability density function f(t). Each density function has a corresponding Survival function (S(t)) and the hazard rate ($\lambda(t)$). The Survival function is defined as $S(t) = P(T > t)$, the probability of the event not having occurred yet, or the proportion of the population that has yet to experience the event. The hazard function is the instantaneous rate of occurrence where $\lambda(t)dt = P(t \leq T < t + dt \,|\, T \geq t)$. The three functions have one-to-one relationships such that $\lambda(t) = f(t)/S(t)$ (Sun, 2006).

These equations describe how the event times of individuals in one population are distributed. Fixed effects (covariates) that may affect the distribution of event times can be incorporated using the accelerated failure time (AFT) model and the Cox or proportional hazard (PH) model. Parametric AFT models assume a specific distribution for the probability density function, commonly using the Weibull or Gamma model. The covariates act directly on the (log) timescale, either accelerating or delaying the time-to-event (Sun, 2006). In a PH model, covariates act on the (log) hazard rate such that a change in the value of a covariate produces a proportional change in the non-parametric hazard function. PH models are used to assess risk relative to the covariates, not absolute risk (Hougaard, 2000). Other common approaches include proportional odds models, additive hazards models and linear transformation models (Sun, 2006). Many standard statistical software packages are capable of estimating parameters for these simple survival analysis models (Lindsey and Ryan, 1998)

Another basic concept of survival analysis models is the underlying assumption of individual heterogeneity in the population: individuals experience the event at slightly different times, according to the probability density function. Survival analysis models can be extended to include not only individual variation, but also variation between groups. These random-effect models stem from the idea that unknown, inherent traits would cause an individual to be more or less frail (more or less likely to survive) and could be shared

between members of a group, such as a family (Hougaard, 2000; Service, 2000). We expect the members of the group to be more similar to each other than to outsiders, but we have no reason to predict that they would be more or less frail than another group. These clustered random effects are known as shared frailty terms and enter the model in the same way as a covariate (Hougaard, 2000). Frailty terms are commonly assumed to follow a Gamma distribution, but other forms are also possible, including log-normal or positive stable distributions (Wintrebert et al., 2005). Frailty models have been used frequently in assessments of failure of medical devices or machines, disease onset in clinical trials and veterinary medicine, and for survival of related individuals, including human twins (Bellamy et al., 2004; Goethals et al., 2009; Wong et al., 2005).

Simpler fixed effects survival analysis models for exact event-time data are not new to ecology. Data from insect survival and development (Van Dooren et al., 2005; Ma, 2010; Ma and Bechinski, 2008) and plant mortality and residence time (Woodall et al., 2005; Ozinga et al., 2007) have been described by models with covariates that include individual characteristics such as size or age. Fire return interval data has also been modeled using survival analysis (Polakow and Dunne, 1999; Moritz et al., 2008). Survival analysis models have also been applied to behavioral data on insect dispersal time (Bishop et al., 2000), deer migration times (Fieberg and Delgiudice, 2008), parasitoid behavior during oviposition (Velema et al., 2005), time between attacks of pine weevils (He and Alfaro, 2000) and predator preference as evinced by prey survival (Schauber and Jones, 2006). These models are limited to considering covariates (fixed effects), excluding any random effects that may be due to replication within a trial or sampling within a site.

Random-effects (frailty) models for exact event-time data have entered the ecological literature with respect to mortality and survival. Frailty models with Gompertz hazard functions are commonly used to model senescence, but in this application the frailties represent individual heterogeneity, not an unexplained source of variation shared by members of a specific group (Service, 2000; Zens and Peart, 2003). Few examples of frailty models exist outside of the senescence literature in ecology, but shared frailty models have been developed for sage-grouse chick survival where frailties are grouped by brood (Aldridge and Boyce, 2007), larval survival of a geometrid moth with frailties grouped by habitat patch (tree) (Tanhuanpää et al., 2001), and scrub jay survival where frailties are grouped by relatedness: shared mother, father, natal territory (Fox et al., 2006). Wintrebert et al. (2005) used frailty models to estimate the joint probability of survival and breeding in interval-censored data on kittiwakes, but the frailties are shared between the two life history processes in the same individual, not shared by individuals of a group.

Frailty models can also be formulated for interval-censored data, in which the exact event time is unknown. Interval censoring is common in clinical trials and in observational studies of animals in which the subject or phenomenon is observed at semi-regular intervals (Lindsey and Ryan, 1998; Ma, 2010). Two observation times create a lower bound and upper bound for the exact event time. Right censoring and left censoring occur when the event occurs outside of the study period, either before the first observation (left) or after the last observation (right). Model estimation becomes increasingly difficult with interval-censoring due to more complex likelihood calculations (Lawless, 2003). Many

analyses sidestep this problem by imputing the exact event time as one bound or the midpoint of the interval, but this creates biased estimates of the underlying parameters (Goethals et al., 2009; Radke, 2003).

A handful of studies have analyzed mixed-effects frailty models for interval-censored data, fully accounting for the interval censoring instead of imputing an exact event time. Zuma (2007) analyzed a proportional hazards model with a Weibull baseline hazard and multiplicative Gamma frailties in an MCMC framework. Instead of using the interval data and corresponding likelihood, Zuma treated the exact event times as unobserved variables, adding dimensions, but simplifying the likelihood. Banerjee and Carlin (2004) analyzed interval-censored data on smoking cessation with Weibull and Gamma frailty models fitted in a Bayesian framework using MCMC. The authors assumed that several of the model parameters (including the frailties) were spatially correlated, so they used a multivariate conditionally autoregressive (MCAR) prior, with a precision matrix dependent on spatial relatedness (Banerjee and Carlin, 2004).

Bellamy et al. (2004) modeled the age at onset of asthma in a community-based study with a Weibull hazard and a normal frailties with additive effects. They directly calculated the likelihoods by approximating the integral over the frailty using quadrature, implemented as SAS macros. Wong et al. (2005) used a similar model in a Bayseian MCMC framework for the time-to-arrest of active dental carries. They simplified the Weibull by holding the rate parameter equal to 1, but added a level of frailty clustering (one for carries in the same individual, one for individuals in the same group). Most recently, Goethals et al. (2009) modeled time to infection in cows, using a Gamma frailty for the dependence of multiple infections in the same animal. They used a proportional hazard model where the frailty affects the hazard rate multiplicatively, and calculated likelihoods directly for several baseline hazard models (Weibull, Exponential, and log-Logistic) for likelihood ratio and AIC comparisons.

This study presents a general approach for maximum likelihood estimation of a mixed-effects time-to-event model for ecological applications. We directly calculate the likelihoods using quadrature to approximate the integrals over each random effect, assuming we have interval-censored data. To evaluate the performance of this model in parameter estimation and hypothesis tests, we present a simulation study exploring different levels of random variation and fixed differences between groups. We also compare results to those from a naïve fixed-effects model that ignores the shared sources of variation. Our example uses a Weibull model with two levels of hierarchical clustered normal frailties that act additively on the log-hazard rate. In this formulation the AFT and PH models are identical, illustrating the flexibility of our approach. There is no widely available statistical software for interval-censored mixed-effects survival analysis, so we programmed the calculations directly in R (R Development Core Team, 2010).

1.2. Methods

Suppose we have many source populations from which we sample individuals. Source populations are of two types: A and B. The two types may be experimental treatments or any other natural contrast. Many replicate cohorts from each source will be followed from

birth to maturation. The development times for individuals in the cohort will be affected by the fixed source type (A or B), the random variation between source populations of one type, and the random variation between replicates from one source. We wish to ask if we can detect a true difference between types A and B, while accounting for the two random effects: source and replicate. We assume that our event data is interval censored and that our observation time is long enough to capture all mature individuals, excluding the need to consider right-censored data. This model is similar to that analyzed by Bellamy et al. (2004), but includes another level of frailty clustering.

1.2.1 Methods: Model description

Suppose we have p source populations (of two types: A and B), each of which supplies m replicate populations with n individuals per replicate. D_{ijk} is the unobserved development time (time from egg or birth to adulthood) of the k^{th} individual in the j^{th} replicate from the i^{th} source population. We assume the development times follow a Weibull model with the probability density function and survivor function

$$f(t) = \lambda\gamma(\lambda t)^{\gamma-1} exp\left(-(\lambda t)^{\gamma}\right) \tag{1}$$

$$S(t) = exp\left(-(\lambda t)^{\gamma}\right) \tag{2}$$

The shape parameter, γ, forces the hazard function to be monotone decreasing for $\gamma < 1$, increasing for $\gamma > 1$ and constant for $\gamma = 1$. The rate parameter, λ, controls the position of the curve with respect to time and has a simple biological interpretation. It is easily shown that approximately 63% $(1 - e^{-1})$ of the individuals will have developed by $1/\lambda$ units (e.g. days, weeks, years). The Weibull can also be formulated in terms of the scale parameter $(1/\lambda)$, and we will interpret some of the results with respect to changes in scale. The frailties and covariates enter the model via the log of the rate parameter, yielding the modified survivor function

$$S_{ij}(t) = exp\left(-(\lambda t)^{\gamma}\, e^{\gamma(\epsilon_i + \nu_j + \beta X)}\right) \tag{3}$$

where $\epsilon \sim \mathcal{N}(0, \sigma_\epsilon^2)$ allows for variation between populations from different sources (source variation) and $\nu \sim \mathcal{N}(0, \sigma_\epsilon^2)$ allows for variation between replicate populations from the same source (replicate variation). β and X are vectors of coefficients and covariates, respectively. The only covariate in this case is the Indicator variable x_i for source type (type effect), which is 1 for source type A and 0 for source type B.

Interval-censored studies yield an interval of observation times, $[L_{ijk}, H_{ijk}]$, such that $L_{ijk} < D_{ijk} < H_{ijk}$ for each individual. The likelihood of an individual maturing in any interval $[L_{ijk}, H_{ijk}]$, conditional on values for the source effect and replicate effect, is

$$P[L_{ijk} < t < H_{ijk}] = \int_{L_{ijk}}^{H_{ijk}} f_{ij}(t)dt = S_{ij}(L_{ijk}) - S_{ij}(H_{ijk}) \tag{4}$$

This model now depends on five parameters, the set of which we will call θ, such that

$\theta : \{\lambda, \gamma, \beta, \sigma_\epsilon, \sigma_\nu\}$. We build the full likelihood starting at the scale of one individual. We move up in scale from one replicate to one source, and then finally to all sources in order to reach the full likelihood of our dataset. Every step up in scale corresponds with a product over all of the likelihoods at the smaller scale. Because we include random effects we will also need to integrate over all possible values of each random effect term at the appropriate scale. The likelihood contribution of the k^{th} individual in the j^{th} replicate from the i^{th} source population, conditional on values for both random effects is

$$Lik\,[\,[L_{ijk}, H_{ijk}\,] \mid \theta, \epsilon_i, \nu_j\,] = exp\left(-(\lambda L_{ijk})^\gamma\, e^{\gamma(\epsilon_i+\nu_j+\beta X)}\right) - exp\left(-(\lambda H_{ijk})^\gamma\, e^{\gamma(\epsilon_i+\nu_j+\beta X)}\right) \tag{5}$$

The likelihood contribution from the j^{th} replicate from the i^{th} source is the product of the likelihoods of the n individuals in that replicate, which can be calculated directly.

$$Lik[\,rep_{ij} \mid \theta, \epsilon_i, \nu_j\,] = \prod_{k=1}^{n} Lik[\,[L_{ijk}, H_{ijk}\,] \mid \theta, \epsilon_i, \nu_j\,] \tag{6}$$

Integrating over all possible replicate effects in (7) yields the marginal likelihood contribution for the j^{th} replicate from the i^{th} source, which has no closed form.

$$Lik[\,rep_{ij} \mid \theta, \epsilon_i\,] = \int Lik[\,rep_{ij} \mid \theta, \epsilon_i, \nu_j\,] f(\nu_j) d\nu_j \tag{7}$$

Moving up in scale to a source population, we calculate the likelihood of the i^{th} source as the product of likelihoods of all m replicates of a common source, sharing a source effect.

$$Lik[\,source_i \mid \theta, \epsilon_i\,] = \prod_{j=1}^{m} Lik[\,rep_{ij} \mid \theta, \epsilon_i\,] \tag{8}$$

Similarly to the likelihood construction for the j^{th} replicate, we can construct the marginal likelihood for the i^{th} source by integrating across all possible source effects, which again has no closed form.

$$Lik[\,source_i \mid \theta\,] = \int Lik[\,source_i \mid \theta, \epsilon_i\,] f(\epsilon_i) d\epsilon_i \tag{9}$$

The full likelihood of all data from all sources is the product across all p sources.

$$Lik[\,data \mid \theta\,] = \prod_{i=1}^{p} Lik[\,source_i \mid \theta\,] \tag{10}$$

Maximizing this likelihood (10) over the parameter set, θ, yields estimates of all model parameters. In equations (7) and (9) we encounter intractable integrals, which we estimate with numerical integration using quadrature. For each such likelihood, we have the general expression

$$\int Lik[\,data \mid \theta, \eta_i\,] f(\eta_i) d\eta_i \tag{11}$$

which we replace with

$$\sum_{i=1}^{100} Lik[\, data \mid \theta, \eta_i \,]\, f(\eta_i) \Delta\eta_i \tag{12}$$

where $\eta \sim \mathcal{N}(0, \sigma_\eta)$ and $\{\eta_1, \eta_2, ..., \eta_{100}\}$ is a sequence within the bounds $[-6\sigma_\eta, 6\sigma_\eta\,]$ evenly spaced by $\Delta\eta = (12/100)\,\sigma_\eta$. Substituting this estimation yields the full likelihood expression for a dataset

$$Lik[\, dataset \mid \theta\,] = \prod_{i=1}^{p} \left(\sum_{s=1}^{100} \left[\prod_{j=1}^{m} \left(\sum_{r=1}^{100} \left[\prod_{k=1}^{n} \left(\int_{L_{ijk}}^{H_{ijk}} f(t)dt \right) \right] f(\nu_r)\Delta\nu \right) \right] f(\epsilon_s)\Delta\epsilon \right) \tag{13}$$

We will refer to this model as the mixed-effects model, reflecting the fact that we include a fixed effect (type effect) and two random effects (source variation and replicate variation). In the simulation analysis we will compare this model to a naïve analysis of a fixed-effects model, which incorporates the type effect (covariate) into the Weibull model, but no random effects.

1.2.2 Methods: Simulations

We conducted a simulation study, varying levels of source and replicate variation to assess the precision and accuracy of parameter estimates and the power of the likelihood ratio tests. We compared the performances of the mixed-effects model to the naïve fixed-effects model of under a range of true differences between groups.

A simulated data set consisted of 200 trial cohorts: 20 source populations, each with 10 replicate cohorts. Of the 20 source populations, 10 were of type A and 10 were of type B. Each trial cohort consisted of 10 individuals whose development times were distributed according to a Weibull model, with constant shape and scale parameters and appropriate type, source, and replicate effects. The baseline shape parameter ($\gamma = 8$) and scale parameter ($1/\lambda = 8.33$) are consistent with development of small herbivorous arthropods (such as aphids or spider mites) with time units in days. Type effect (β) was constant for a data set, but varied between simulations in the range: {0, 0.05, 0.1, 0.2, 0.3, 0.4}. This range corresponds to decreases in the value of the true scale ($1/\lambda$) from 8.33 time-units to {8.33, 7.93, 7.54, 6.82, 6.17, 5.59}, representing decreases of {0% , 4.88%, 9.52%, 18.13%, 25.92%, 32.97%}. Source and replicate effects were drawn from independent normal distributions with mean zero and one of 3 levels of variation: zero, low ($\sigma = 0.05$), and high ($\sigma = 0.15$). We examined all combinations of source and replicate variation, except the two that paired zero variance with high variance, yielding 7 variation schemes.

Exact development times for the 10 individuals in a trial cohort were drawn from the appropriate Weibull distribution. A sampling interval of 2 time-units was imposed to create interval-censored data, instead of exact-time data. Mortality, observation error and reproduction were assumed to be absent. 200 data sets were simulated for each combination of type effect and source and replicate variation (42 scenarios). For each data set, likelihood ratio tests were conducted to determine if type effect was a significant factor

in each of two models: the mixed-effects model with source and replicate variation, and the naïve fixed-effects model. likelihood calculations were programmed directly in R using optim and the Nelder-Mead method to minimize the negative log-likelihood functions.

1.3 Results

The likelihood function behaved well across all combinations of type effect and source and replicate variation, consistently reaching a global maximum. True parameter values were consistently well estimated by the mixed-effects model and were always covered by the interquartile range (figure 1 for example). Increasing either the source or replicate variation increased the standard error of all parameter estimates, but did not bias the estimates. For true non-zero source and replicate variation, estimates of the source variation had consistently higher standard error than estimates of replicate variation.

The naïve fixed-effects model underestimated both Weibull parameters, resulting in large biases in the shape parameter (γ) and smaller biases in the rate parameter (λ) (figure 1). Biases increase dramatically with either source or replicate variation. Weibull curves resulting from the biased parameter values are wider than those produced by the true parameter values. Under zero variation, both the mixed-effects model and the fixed-effects model detected decreases in scale ($1/\lambda$) as small as 5% (decrease in scale from 8.33 to 7.93 time-units) with 100% power. When variation is included at any level, the fixed-effects model yields highly inflated type I error rates in the range of (0.5 - 0.8) (figure 2).

For the mixed-effects model, increasing either source variation or replicate variation decreased power over weak-medium type effects (figure 2), with changes in source variation affecting power more drastically than changes in replicate variation. Increasing replicate variation (moving left to right in the same row in figure 2) lowered power slightly for weak type effects ($\beta < 0.2$). In contrast, increasing source variation (moving down in the same column in figure 2) decreased power at all but the strongest type effect ($\beta = 0.4$). Type I error rates for likelihood ratio tests of the mixed-effects model did not always match nominal rates, but the differences were minor.

1.4 Discussion

Although survival analysis methods are being applied more frequently to diverse problems in ecology (Ma, 2010; Moritz et al., 2008; Marzolin et al., 2011) , we have lacked the methods to easily analyze mixed-effects survival models for interval-censored data. This gap will affect ecologists more strongly than other researchers whose systems are not as highly variable and complex. The framework we have developed in this paper for estimating survival models will hopefully motivate more ecologists to apply these sophisticated tools to their systems.

A strength of even the simplest survival time model is the assumption of individual heterogeneity in event time. It may be tempting to propose that this description of variation is sufficient in representing all of the variation in population data. However, the inflated type I error rates and biased parameter estimates clearly show how inadequate and inappropriate a fixed-effects survival model would be for most ecological data. The bias of

the naïve fixed-effects model in our study towards lower values of the shape parameter (γ) widens the Weibull curve to fit data that are actually distributed according to multiple narrow Weibull curves. The bias in the rate parameter (λ) contributes to the widening of the distribution, but to a lesser degree.

It may also be tempting to simplify the analysis by imputing an event time as the midpoint of the interval, or as one extreme. It has been well demonstrated that estimating an exact event time from an interval introduces bias in the parameter estimates and can lead the researchers to a different conclusion than they would reach by using the interval-censored data (Goethals et al 2009, Radke 2003). We do not attempt to replicate their results here in favor of exploring more levels of type effect and variation.

There are a few limitations to our analysis. Type I error rates did not always perfectly match nominal rates, and although we were satisfied the likelihood tests were valid for our simulations, in cases with fewer sources or replicates, we would urge consideration of randomization tests. Under the highest levels of variation we considered, the mixed-effects model could not reliably detect a type effect weaker than $\beta = 0.1$, which corresponds to a 9.5% decrease in the scale ($1/\lambda$), or approximately 0.79 time-units for our parameter choices. This type effect may seem weak or strong, depending on the organism and timescale. For example, the ability to detect, under high variability, a difference between two populations of 0.79 days (19 hours) in the development time of an arthropod pest would be incredibly powerful and useful to a pest manager or farmer. Choice of time unit and time scale will be very influential in any application to field data.

We have provided a framework for developing and analyzing a mixed-effects (shared frailty) time-to-event model for interval censored data in the context of individual development time. Any ecological time-to-event data could be analyzed in this framework, including oviposition, behavioral durations, or survival. Our model could easily be extended to incorporate additional sources of variation or other features of survival analysis models such as gender differences or a cure rate for individuals that die, migrate, or are otherwise lost to the population.

2 Chapter 2

2.1 Introduction

Life history variation is a characteristic feature of many natural systems (Uchmański, 1985). The scale and drivers of variation may vary, with considerable consequences for population dynamics (Bjornstad and Hansen, 1994; Souissi and Ban, 2001; van Noordwijk and de Jong, 1986). Variation at different scales may include phenotypic variation within a genotype, standing genetic variation within a population, and fixed differences between populations. Phenotypic or genetic differences between populations range from non-genetic polyphenisms to speciation (Diehl and Bush, 1984) and can be driven by host use (Dickey and Medina, 2010). Many Acarine mites, and spider mites (Family *Tetranychidae*) in particular, display significant amounts of life history variation and are prone to host-associated differentiation (Magalhães et al., 2007b). Host race formation is not commonly thought to occur between panmictic populations using different varieties of the same plant species. In this study we test whether mites are differentiated in ecologically important life history traits at small scales, between two cultivars in a landscape mosaic.

Evidence can be found to support the hypothesis that mites may evolve differences on hosts as similar as cultivars and at small geographic scales. Numerous studies have shown rapid adaptation to host species evinced by marked changes in life history processes such as survival, development and reproduction. *T. urticae* lines adapted to favorable (bean) or unfavorable (bean and cucumber) hosts each have higher fitness and survivorship than the other line on the "native" host. Lines adapted to the unfavorable hosts also have higher survivorship on novel marginal hosts (Gould, 1979). Populations of *T. urticae* adapted to hosts that impose high juvenile mortality (tomato and brocolli) show lower mortality, greater acceptance, and increased developmental rate (tomato only) than bean-adapted mites introduced to the unfavorable host plants (Fry, 1989). *T. urticae* populations reared on unfavorable hosts (cucumber, tomato, and pepper plants), show high variation in juvenile survival, longevity, and fecundity, and increases in mean trait values (excepting longevity) after adapting to novel favorable hosts (Magalhães et al., 2007a). Lines of *T. urticae* adapted to different hosts (tomato, arapidopsis, and bean) vary in fecundity and in feeding damage on novel hosts. Lines also differentially induce and respond to jasmonic acid plant defenses (Kant et al., 2008). *T. urticae* populations show adaptive plasticity in fecundity when comparing bean and tomato plant hosts (Agrawal et al., 2002).

Rapid adaptation to host species has also been demonstrated genetically, in addition to the above studies of ecological traits. Evidence from a genome-sequencing study of *Tetranychus urticae* shows that 24% of genes are differentially expressed upon host transfer from bean to a less favorable host (tomato or arapidopsis), with the most profound changes occurring in genes in the detoxification and pepsidase families (Grbić et al., 2011). Other genetic differences between spider mite lines adapted to different host plants are present in micro satellite markers (Nishimura et al., 2005), allozyme and nuclear ribosomal sequences (Navajas et al., 2000), and at the phosphoglucose isomerase locus (Gotoh et al., 1993). These genetic differences can become so extreme that lines of the same species (*T.*

kanzawai, *T. urticae*, and *O. gotohi*) become reproductively isolated (Gomi and Tetsuo, 1996; Navajas et al., 2000; Gotoh et al., 2007). Other herbivorous Acari species (*A. hystrix* and *N. paspalivorus*) are also reproductively isolated on host plant (Famah Sourassou et al., 2010; Skoracka, 2008). The above studies clearly demonstrate both life history and genetic differences between isolated mite populations adapted to different plant species. It is less clear what to expect from populations feeding on different cultivars (genotypes) of the same host species in a mosaic pattern on the landscape.

While cultivars of the same species are likely to be more similar than different species, they do vary in their response to herbivores and pathogens. Suitability of the host cultivar as measured by resistance to herbivory may provide a basis for mite adaptation. Plant species clearly differ in their resistance to herbivory by spider mites (Karban and English-Loeb, 1988). Similarly, levels of induced and constitutive resistance to *T. pacificus* vary widely among six cultivars of *V. vinifera* as well as among other *Vitis* species (English-Loeb et al., 1998). Abundances of several species of herbivorous and predatory mites were found to differ by cultivar (Merlot vs Verduzzo and Riesling vs Prosecco) in two *V. vinifera* vineyards about 3 ha large (Duso and Vettorazzo, 1999). *V. vinifera* cultivars, as well as other grape species, also vary in their constitutive resistance to pathogens such as grey mold emph(Botrytis cinerea) (Pezet et al., 2003) and powdery mildew (Gee et al., 2008; English-Loeb et al., 2005; English-Loeb and Norton, 2006). Genotypes of different *Vitis* species, including 13 cultivars of *V. vinifera* have genetic differences in the PR1 gene family that is involved in host resistance to pathogens (Li et al., 2011). This pattern of resistance varying by cultivar has also been shown in other systems, such as plant bug herbivory on strawberry (Rhainds and English-Loeb, 2003) and spider mite herbivory on cotton (Agrawal and Karban, 2000; Thaler and Karban, 1997)). This evidence suggests that we may expect cultivar-associated differences that are similar to host-associated differences in life history and genetics when lines are allowed to evolve in isolation, but provides no insight into expected differences in sympatric populations.

Vineyard herbivore populations live on a contiguous landscape in a mosaic of suitable host plants. Hosts vary widely, including vegetables in gardens, plants in residential landscaping, and crops in agricultural fields (Helle and Sabelis, 1985). Our study site lies in California grape crush district 11, which has 69,220 total standing acres of wine grapes (in 2009) including Zinfandel (18,800 acres), Chardonnay (13,563 acres), Cabernet Sauvignon (11,272 acres), and Merlot (7,497 acres) (California Department of Food and Agriculture and USDA National Agricultural Statistics Service, California Field Office, 2010). Fields of single cultivars are distributed across San Joaquin and Sacramento counties and are often adjacent to different cultivars. Spider mites disperse between host plants by crawling on leaves and in the soil and by ballooning on a thread of silk carried by wind, with median dispersal distances around 0.7 meters (Jung and Croft, 2001). The balance between the isolation caused by small scale dispersal and mixing from large scale dispersal events may determine the extent to which field populations adapt to host cultivar.

Our study investigates whether field populations of pacific spider mites (*Tetranychus pacificus*) sampled from different cultivars of grapevine (*Vitis vinifera*) display significant life history differences on a common host. We aim to detect specific differences in

developmental rate, juvenile survival rate, and fecundity between field populations from Zinfandel and Chardonnay cultivars in a "common garden" experiment on lima bean plants (*Phaseous luteus*) in the laboratory. Replicate sites are sampled from each cultivar and replicate bioassays are run from each site. Assay populations are non-destructively sampled using digital photography in order to count individuals over 12 days of density-independant dynamics. We analyze the bioassay data with standard generalized linear mixed models and with time-to-event models which are standard in other fields, but rarely applied to ecological data. Time-to-event models allow for variation between individuals in the rate at which individuals experience a discrete event, such as death or maturation. Fitting the time-to-event model is made more complicated by two levels of replication and the observations made at discrete time intervals. Since there is no widely available software for mixed-effects survival analysis with interval-censored data we use original code presented in Scranton and de Valpine (2012) for likelihood calculations.

2.2 Methods

The Pacific spider mite (*Tetranychus pacificus*) is an herbivorous arthropod pest in many agricultural systems, including the vineyards of California's central valley. This phytophagous mite feeds by piercing leaf cells with its mouthparts and sucking out the cell contents (Sabelis, 1986). *T. pacificus* has 5 distinct life stages including one larval stage and two nymphal stages (protonymph and deutonymph) and undergoes a quiescent period at each stage transition. Its demography (including development rate and fecundity) is highly temperature sensitive (Stavrinides and Mills, 2011) and highly variable (Benton et al, 2005).

T. pacificus were collected from vineyards (*Vitis vinifera*) near Lodi, CA in the San Joaquin Valley from mid-July through late-August 2009. Possible sites were chosen through discussions with local farm advisors and pest control advisors, and vineyards were sampled opportunistically due to the prompt spraying of miticide at outbreak sites. Mites were sampled from two different cultivars of grapevine: Zinfandel and Chardonnay. Over the sampling period, mites were collected from 8 Zinfandel vineyards and 3 Chardonnay vineyards (figure 3). At each site 10-15 leaves were clipped from infested grapevines and placed in paper bags labeled with identifying information. Sample bags were transported back to the lab in a cooler at 15°C.

2.2.1 Methods: Common garden experiment

Field samples were processed at the Oxford Tract Greenhouse and Insectary at UC Berkeley. If samples were not processed on the same day as collection, they were transferred to sealed Tupperware containers with moistened paper towels and kept in an incubator at 22°C with a 16L : 8D photocycle for less than 24 hours. 50-100 mated adult females were transferred from samples to greenhouse-grown, uninfested bean plants (*Phaseolus lunatus*) via leaf disc to found a *sample colony*.

All sample colonies were kept in a growth room on a 16L : 8D photocycle at 28°C constant temperature and 36% relative humidity. Each sample colony was maintained on 2

large bean plants in a 8 ft^3 cage on an elevated rack under fluorescent grow lights. Cages were made of plexiglass with large openings on three sides and the top, all covered with a mesh, fine enough to prevent movement of mites and sealed tightly at edges with plexiglass glue. Door edges and hinges were sealed with duct tape and precautions were taken to prevent mite movement when cage doors were opened to water plants. Cages were elevated on overturned pots, whose bases were ringed with stikem (Seabright Laboratories, 2009). Fresh plants were rotated in after 7-10 days to provide extra habitat and promote growth. Sample colonies were maintained for 14-21 days. The actual time that individuals spent on bean plants was long enough to discount any maternal effects from field conditions and short enough to prevent adaptation to the lima bean host. All field samples were processed using the same procedures and there were no systematic differences between cultivars in field collection date or assay date.

Individuals from each sample colony were used to found 4 *assay populations*. Two assay populations were initiated with a cohort of 10 eggs and two with a cohort of 5 mated adult females. Egg cohorts were initiated by directly transferring 10 nearly hatched eggs (with visible eyespots) to the assay population host plant with a paintbrush. To obtain an adult cohort of 5 newly-emerged, mated adult females, we removed approximately 10 females in the third (final) quiescent stage and 10 adult males from the sample colony and placed them on a large leaf disc surrounded by damp cotton in a petri dish. After 24 hours we transferred 5 adult females that had emerged and mated to a new, clean leaf disc and pinned the disc to the assay population host plant.

Each assay population host plant was a greenhouse bean plant approximately 3 weeks old. Plants were washed to protect against pest infestation, repotted, and trimmed to two large, paired leaves each. Leaves were trimmed to flat rectangular sections approximately 6 cm by 4 cm to control leaf surface area across all assay populations and to allow for clearly focused images. Trimmed, repotted plants were kept under grow lights in the growth room for 1-3 days before use to ensure they were healthy and free of pests.

Assay populations were maintained for 12 days on a plastic covered wooden rack under grow lights in the same growth room (16L : 8D photocycle at 28°C 36%) and hourly temperature recordings were kept. Temperature recordings showed reliably constant temperatures, with slight fluctuations that were not a significant factor in any of the analyses. Each assay population host plant sat in a small tray in one square of a large grid of stickem. Mite movement was further prevented by stikem around the lip of the tray and the top edge of the pot. No mites were noted in the stikem over the course of the experiment.

2.2.2 Methods: Non-destructive sampling

We chose to follow the assay populations in situ with a non-destructive sampling scheme. Non-destructive sampling of leaves to quantify leaf damage has been accomplished using photographic methods (Hargrove, 1988). Several recent studies have had success in using image processing tools, such as those available in MATHEMATICA or more specialized software (Škaloudová et al., 2006; Boese et al., 2008). Non-destructive sampling of herbivores in situ is more difficult because of our inability to physically manipulate the

leaf. Many studies still rely on image analysis with leaf damage as a proxy for abundance (Gilbert and Grègoire, 2003; Luedeling et al., 2009). We sampled leaves in situ with a digital camera (Nikon D5000) and 60 mm macro lens without touching the leaves. Every 48 hours, photographs were taken of each assay population. Four images (1 per quadrant) were taken of each side of each leaf for each assay population with 2 leaves. This summed to 16 images per assay population per sampling time. Mites were counted by eye with the aid of basic MATLAB image analysis and data organization tools. Individuals were categorized as eggs, immatures, or adults, and resulting data sets follow population abundance for each stage for 12 days.

2.2.3 Methods: Mixed-model analysis

The common garden experiment yielded 44 assay populations (11 sample colonies with 4 replicates each). Each of the 44 assay populations yielded a data set that tracked the abundance of a cohort over 12 days. Individuals were classified as either egg, immature, or adult from the images. Observations were made at 2 day intervals, yielding 7 observations per assay population. We highlighted data on development time and juvenile survival from the 22 egg-initiated assay populations and data on fecundity from the 22 adult-initiated assay populations. The goal of our analyses was to detect any difference in these 3 life history processes between assay populations from Zinfandel vineyards and assay populations from Chardonnay vineyards.

To obtain development intervals from egg-initiated assay populations, we highlighted the number of adults on each day. If the number of adults increased between observations, we inferred that immature individuals developed into adults sometime in the interval between observation times. Thus for each assay population, we formed a count of adults maturing in each observation interval. The two observation times created a lower bound and upper bound for the exact maturation time, which remained unknown. We first analyzed the development data with a generalized linear mixed model (GLMM). For the GLMM, we condensed the development intervals by grouping individuals into those who matured before day 6 and those who matured after day 6. We used a binomial regression to estimate the effect of cultivar on the probability of maturing to adulthood before day 6, with random effects for the sample colony and for the replicate assay populations. We used a second approach in analyzing the interval development data: a mixed-effect time-to-event model that allows incorporation of all development intervals.

Time-to-event models are a class of models from survival analysis with an underlying assumption of individual heterogeneity in the population: individuals experience the event at slightly different times, according to a probability density function (Hougaard, 2000). The specific probability density function can be chosen to fulfill specific assumptions in the data or to allow for a flexible shape in the probability of an individual experiencing the event. Time-to-event models can be extended to include covariates and random effects, or variation between groups (Hougaard, 2000). Models that include random effects (frailty models) stem from the idea that unknown, inherent traits would cause an individual to be more or less frail (in our case more or less likely to mature) and could be shared between members of a group (in our case an assay population or a field source) (Service, 2000).

Time-to-event models can be fit to data with exact event times with a number of existing software packages. Likelihood calculations are more complicated when data is interval-censored, as our development data was.

We fit a mixed effects time-to-event model to development times using a Weibull distribution to describe the probability of maturing as the individual ages. We included the cultivar as a fixed effect and included two random effects for the sample colony and for the replicate assay populations. We estimated the Weibull parameters (shape, γ, rate: λ) and tested the significance of the fixed cultivar effect (beta β the difference between cultivar sources in log lambda) while accounting for these two sources of random variation. Since there is no widely available software for mixed-effect time-to-event models with interval censored data we programmed likelihood calculations directly in R (R Development Core Team, 2010) using code from Scranton and de Valpine (2012).

To obtain juvenile survival rates from the egg-initiated assay populations we highlighted the number of individuals present on day 4 out of the initial 10 eggs in an egg-initiated assay population. We used a binomial regression to estimate the effect of cultivar on the probability of surviving to day 4, with random effects for the sample colony and for the replicate assay populations.

To obtain fecundity data from the adult-initaited assay populations, we counted the number of eggs present on day 2. We weighted the count by the number of surviving adult females on day 2 in that assay population. We again included the cultivar as a fixed effect and two random effects for the sample colony and for the replicate assay populations in a GLMM. We used a poisson regression for egg counts. All GLMM analyses were done in R using the lme4 package, version 0.999375-34, (Bates and Maechler, 2010).

For both the time-to-event model and the GLMMs, we fit models with and without the (null) cultivar effect and assessed the significance of the fixed effect with likelihood ratio tests. These tests use large sample approximations to the chi-squared distribution, but we had concerns about our small sample sizes so we performed randomization tests. We randomized the dataset by cultivar 1000 times to create a null distribution for each of the likelihood ratio test statistics (D), on which we base our main conclusions.

2.3 Results

The developmental rate of spider mite populations in our experiment differed significantly with source cultivar. The best-fit time-to-event model yielded a Weibull distribution for the development times of individuals from each cultivars (figure 4a). Each distribution depended on three parameters: a constant shape parameter (γ) and a rate parameter (λ) that was affected by the fixed cultivar effect (β) on a log scale. Maximum likelihood estimates of the Weibull parameters revealed that populations founded from mites on the Zinfandel cultivar developed more quickly than those from the Chardonnay grapevines (shape: $\hat{\gamma} = 8.090$, rate: $\hat{\lambda} = 0.1596$, fixed cultivar effect coefficient: $\hat{\beta} = 0.1299$). Likelihood ratio tests showed that the fixed cultivar effect was significant in the mixed-effects Weibull time-to-event model ($D = 5.499, p < 0.0110$ by randomization, $p = 0.0190$ by large sample chi-squared approximation). The estimated random effects show that assay populations varied randomly in development time, but different sample

colonies did not. Maximum likelihoods estimates of the standard deviation of each random effect revealed moderate variation between assay populations ($\hat{\sigma}_{assay} = 0.0702$) and zero variation between sample colonies $\hat{\sigma}_{source} = 2.98x10^{-5}$). Figure 4a depicts the range of variation due to these random effects with dashed lines depicting lower and upper bounds for the estimated Weibull distribution.

The GLMM for the proportion of individuals matured after 6 days found a significant effect of source cultivar ($D = 5.934, p < 0.0190$ by randomization, $p = 0.149$ by large sample chi-squared approximation, Fig 4b). Survival of juveniles after 4 days was, on average, 76.7% for Zinfandel and 63.3% for Chardonnay. Analysis with a GLMM found that this difference was marginally significant ($D = 4.246, p < 0.063$, by randomization, $p = 0.0393$, by large sample chi-squared approximation, Fig 4c). Females laid on average 10.8 eggs in the first two days of observation, independent of source cultivar ($D = 0.06923, p < 0.835$, by randomization, $p = 0.792$, by large sample chi-squared approximation, Fig 4d).

2.4 Discussion

Our study has demonstrated life history differences in spider mite populations associated with host cultivar. When reared on a common, favorable host, mites from a source population on Zinfandel grapes matured more quickly and likely experienced higher juvenile survival than individuals from Chardonnay source populations. Populations were not strictly isolated even while these associations developed, but were distributed in a mosaic on the landscape with limited movement. The host plants that drove these differences were also similar; different cultivars of the same plant species are associated with different life history traits.

Our analysis of the cultivar-associated life history differences using a mixed-effects time-to-event model provided additional detail and biological realism. Time-to-event models yielded information about the rate and shape of the distribution of development times over the entire time span of the study. The fitted model gave us insight into times at which individuals are most likely to mature and how likely individuals are to mature at the extremes, either very early or very late. Standard GLMM analysis only provided information about the development rate at one discrete time and ignored data from the larger time span of the study. Many likelihood tests for many discrete time points would be necessary to get some of the insight provided by time-to-event models.

Photographic sampling is common in observational studies of other species using camera traps, remote sensing or other aerial data. Most studies on herbivory that employ a photographic sampling scheme quantify leaf damage, not herbivore abundance. Our methods were limited by the difficulties of creating and testing feasible algorithms for automated counts. One main challenge is the lack of color contrast in the spider mite and lima bean system. Other ecological systems and advances in image processing and object-based image analysis may allow future studies to employ fully or semi-automated counting of individuals.

Our study did not attempt to address any evolutionary aspects of the cultivar-assocaited life history differences such as the timescale of adaptation to cultivar, the potential for speciation or the amount of gene flow between field populations. However,

numerous studies have found that tetranychid mites respond to selection on life history traits rapidly, within 6-15 generations (Fry, 1989; Magalhães et al., 2009; Tien et al., 2010). Adaptive phenotypic plasticity and host-assocaited differentiation can lead to evolutionary and ecological changes (Agrawal and Karban, 2000; Dres and Mallet, 2002) and sympatric speciation is common among phytophagous insects (Berlocher and Feder, 2002). These studies suggest that the capacity for rapid adaptation places some evolutionary questions in an ecologially relevant timescale. Future studies are needed to investigate this interplay between ecological and evolutionary dynamics, especially in these sympatric populations.

The scope of our study was also limited to populations on a common lab host (*Phaseolus lunatus*); we did not attempt to follow populations on the native grape cultivar. Management decisions could be affected depending on the degree to which our findings translate to field populations on native hosts. Pest managers and growers weigh the strength of outbreak, the speed of development (mostly relying on temperature as a proxy), and the harvest time of the crop against the cost of pesticide applications or release of natural enemies. Differences in development and survival in field populations would impact the decision to apply miticide and the timing of application. Our results may directly affect the way growers think about the speed of spider mite development in different vineyards.

Our results also highlight the need to investigate the impact of life history variation on on population growth rate on spider mite populations. Some empirical studies have found host-associated differences between spider mite populations in the intrinsic rate of increase (Gotoh and Gomi, 2003) and theoretical studies predict strong effects of life history variation on population dynamics (Bjornstad and Hansen, 1994; Vindenes et al., 2008). Our study was limited to specific demographic rates, not any metric of overall dynamics, but by establishing the presence of life history differences in sympatric populations on very similar host plants we have hopefully highlighted the importance of investigating the effects of life history variation on dynamics. A central tenet of population ecology is the importance of understanding the drivers of population dynamics and in many cases life history variation merits consideration as one of those drivers.

3 Chapter 3

3.1. Introduction

Many biological systems require complex models to accurately represent the dynamic processes and variability present in nature. In many cases, the complexity in the model results in a likelihood function that is analytically or computationally intractable. Commonly used likelihood-based methods cannot be used to evaluate these models, perform model selection, or estimate model parameters. In these cases we can turn to a more flexible statistical framework: approximate Bayesian computation (ABC). An ABC approach allows us to replace the likelihood function with repeated simulations from a model. By selecting simulations that are similar to the observed data, we can move from a prior distribution to the approximate posterior distribution using one of several established algorithms. The development of the ABC framework is a relatively recent advance (Tavare et al., 1997; Pritchard et al., 1999) that is occurring in a variety of fields (Beaumont, 2010).

ABC methods are used extensively in population genetics to infer demographic parameters from molecular variation in a sample (Csilléry et al., 2010). Parameters of interest include coalescence time (Tavare et al., 1997), crossing-over rates (Padhukasahasram et al., 2006), expansion times and migration rates (Hamilton et al., 2005), and mutation rates (Bazin et al., 2010). ABC methods are also commonly used in conservation genetics to estimate parameters that affect endangered populations such as inbreeding, effective population size, and minimum viable population size (Lopes and Boessenkool, 2010; Rabosky, 2009). The use of ABC is spreading beyond molecular genetics to a diverse group of fields such as epidemiology (Blum and François, 2008), systems biology (Barnes et al., 2011), veterinary medicine (Tinsley et al., 2012), human demography (Shriner et al., 2006), and psychology (Turner and Van Zandt, 2012).

ABC methods are slowly filtering into the field of ecology (Hartig et al., 2011; Beaumont, 2010). Jabot and Chave (2009) estimated immigration rates and regional diversity under the neutral model of biodiversity for tropical forest trees. ABC techniques have been used with simulated data to estimate parameters of classical theoretical models, such as Lotka-Volterra equations Toni et al. (2009) and continuous time models of host-parasite dynamics (Drovandi et al., 2011). Parameter estimation using simulated data has also been performed for mechanistic models such as a stochastic cellular automata model of range expansion (Rasmussen and Hamilton, 2012) and in adaptive management scenarios of deer herbivory control measures (Ramsey et al., 2012).

The ABC framework has the potential to fill a gap in our ability to fit models in population ecology: fitting models of stage-structured populations to variable cohort data. Many organisms develop through distinct life stages with highly variable stage durations, survival, and fecundity. Often individuals are grouped into size or stage classes as a proxy for their exact developmental state. Fitting models to data of highly variable stage-structured populations is a difficult problem and is most commonly solved by making simplifying assumptions in the model. Existing models, such as matrix models, can be fit with likelihood based approaches but lack individual heterogeneity in life history processes

and assume unrealistic stage durations (de Valpine, 2009; Birt et al., 2009). Likelihoods based on more realistic models with individual heterogeneity would be intractable. Using only the ability to simulate from our population models we can turn to the more flexible statistical framework of ABC to perform parameter estimation, model inference or model selection.

The simplest ABC algorithm is a rejection algorithm. Many simulations are generated from a model using parameter values drawn from a prior. Those simulations that match the data exactly or fall within some acceptable distance of the data are kept and the rest are thrown out. A threshold is introduced to delineate simulations that are "close" to the data from those that are "far." The specific parameter values that generate accepted simulations represent a sample from the posterior distribution. Such algorithms are not practical when it is unlikely that a randomly generated simulation may be accepted, due to the dimensions of the parameter set, the nature of the observed data, or large discrepancies between the prior and posterior. More sophisticated algorithms have been developed to address this high rejection rate (Appendix A) such as an ABC rejection sampler with regression adjustments (Beaumont et al., 2002), an ABC MCMC algorithm (Marjoram et al., 2003), and several different particle filter samplers (Cappé et al., 2004, 2007).

A large obstacle to implementing any ABC algorithm is the choice of summary statistics and distance metrics, which define the closeness of a simulation to the observed data. For those fields in which ABC has been used extensively, such as population genetics, summary statistics can be chosen from previous studies. For novel applications of ABC the choice of summary statistics and distance metrics is a difficult problem, with many authors conceding that there is no good rule (Wilkinson, 2008). In general, potential summary statistics can be identified from previously published studies or from the researcher's own knowledge of the system (Marjoram et al., 2003). Summary statistics should be sufficient (Beaumont, 2010) and independent (Wilkinson, 2008). There should be a homoskedastic relationship between the resulting distance metric and the parameter values (Blum and François, 2008).

Many different strategies are used to adhere to these general guidelines in choosing summary statistics. Some seek to ensure independence and reduce the number of summary statistics using linear discriminant analysis (Estoup et al., 2012), partial least-squares (Wegmann et al., 2009) or principal components analysis (Wilkinson, 2008). Other studies forgo the use of summary statistics and instead compare the raw data using distance metrics, such as the sum of absolute differences or sum of squared distances (Drovandi and Pettitt, 2011a; McKinley et al., 2009). However, this strategy merely transfers the issue of selection to the distance metrics.

The most common approach is to create training sets of simulated data for pilot ABC runs with different combinations of summary statistics, choosing the combination that produces the best results (Li and Jakobsson, 2012). Comparisons of results can be made directly using the estimated mean and variance of the posterior or by using the odds ratio of posteriors with and without each summary statistic (Joyce and Marjoram, 2008). Minimizing entropy or mean root sum of squared errors (MRSSE) of the accepted parameter sets from pilot ABC runs can also be used to select a combination of summary

statistics (Nunes and Balding, 2010). Jung and Marjoram (2011) use a genetic algorithm on their pilot ABC runs that "evolves" a set of weights for summary statistic by maximizing the "fitness:" the inverse of the sum of squared differences between each accepted parameter value and the estimated posterior mean. Fearnhead and Prangle (2012) use pilot ABC runs (with arbitrarily chosen summary statistics) to create informative priors and calculate summary statistics based on the posterior means of the parameters. They use non-linear regression of parameters on the simulated data to estimate the coefficients, which become the summary statistics for their ABC algorithm. This pilot run approach amounts to trial and error ABC until summary statistics are found that perform adequately, with extremely high computation costs.

Other approaches attempt to avoid extra computation steps by modifying the ABC algorithm. Hamilton et al. (2005) add a step that weights summary statistics according to the R^2 values from a regression adjustment (Beaumont et al., 2002). Drovandi et al. (2011) introduce an auxiliary model, whose parameters are estimated using maximum likelihood and used as the summary statistics for the more complex model. Blum (2010) add a forward stepwise selection step based on maximizing the marginal likelihood (evidence function) over all possible additional summary statistics.

In this paper, we present a novel application of the ABC framework to a model of stage-structured population dynamics. We use an ABC SMC algorithm with an approximated optimal backwards kernel in the weights and an adaptive threshold schedule for parameter inference. The stochastic stage-structured model explicitly includes individual heterogeneity in development and survival and allows for ongoing reproduction. Although the model is simple, fitting it to cohort data has not been successfully done in the past. Previous work has used cohort data with independent samples at each time and without reproduction (Read and Ashford, 1968; Hoeting et al., 2003). The estimation problem is considerably harder with ongoing reproduction and the temporal non-independance created by following unmarked individuals through time. We also develop a systematic method for evaluating potential summary statistics and distance metrics a priori, without costly computation steps. In the next section we provide specific descriptions of the population model, the ABC framework, the ABC SMC algorithm, and the methods for selecting summary statistics and distance metrics. In the third section we evaluate the performance of our estimation method with simulated data. We provide a discussion of the potential of ABC in population ecology and the obstacles we still face.

3.2. Methods

In this section we first formulate the specific problem we face in modelling the dynamics of a stage-structured cohort with individual heterogeneity in life history processes. We briefly touch on the general ABC framework, deferring to the existing reviews and comparisons of various ABC algorithms (Beaumont, 2010). In the third section we give a specific description of the ABC SMC algorithm and then provide a description of the methods used to select summary statistics and distance metrics

3.2.1 Methods: Stochastic stage-structured population model

A stage-structured population is a group of individuals of the same species who age through discrete life stages, experiencing different demographic rates in each stage. We expect individuals in the same stage to have inherent genetic differences in specific survival and fecundity rates and in the speed at which they develop through these stages. In addition, we would expect some phenotypic variability in the expression of those genotypes. These sources of individual heterogeneity in life history processes are important for population dynamics (de Valpine, 2009) and motivate the development of the model.

Data on stage-structured populations commonly results from studies of cohorts, or groups of individuals born at the same time and monitored for a certain amount of time. We follow the same individuals through time, so that observations are not temporally independent. Observation cannot be done continuously, so the exact times of stage transitions, births, and deaths are unknown. Instead the data is interval censored; individuals are observed at set points. Additionally, individuals are often unmarked, so we only have data on the number of individuals currently in each stage at each observation time.

We consider a population model for the demographic processes in a closed population. Each individual ages through 3 life stages: egg, immature, and adult. The duration of the egg stage (d_e) and the duration of the immature stage (d_i) follow Weibull distributions with scale and shape parameters $(1/\lambda_e, \gamma_e)$, $(1/\lambda_i, \gamma_i)$ respectively:

$$f(d_e) = \lambda_e \gamma_e (\lambda_e t)^{\gamma_e - 1} exp\left(-(\lambda_e t)^{\gamma_e}\right) \tag{14}$$

$$f(d_i) = \lambda_i \gamma_i (\lambda_i t)^{\gamma_i - 1} exp\left(-(\lambda_i t)^{\gamma_i}\right) \tag{15}$$

Adults experience mortality (m) such that survival time(d_s) follows an exponential distribution

$$f(d_s) = me^{-mt} \tag{16}$$

In order to simulate a population, we generate exact stage durations, exact survival times, and exact birth times of offspring for a cohort of individuals. These individuals with known birth times found a population which we follow for a period of time [0, T]. There is an exact stage duration for the egg and immature stages and an exact death time for each individual, according to equations (14), (15), and (16). Each individual reproduces at a constant rate (r) from maturation until death. As new eggs enter the population with exact birth times, they are also assigned stage durations and survival times. The bookkeeping for each individuals amounts to birth time, egg stage duration, immature stage duration, and survival time. The full stochastic stage-structured model has 6 parameters $\{\lambda_e, \gamma_e, \lambda_i, \gamma_i, m, r\}$.

In order to simulate a data set that would be observed from a population, we "sample" the exact population every 2 time steps. We count individuals according to stage, ignoring the exact ages. This creates a data set of the number of individuals in each stage at each sampling time, [0, 2, 4, ..., T].

Fitting a stage-structured population model to cohort data is made very difficult by

the individual heterogeneity in survival and stage durations, ongoing reproduction, temporal non-indenpendence of data, unmarked individuals, and interval censoring. Standard likelihood based methods cannot be used, so we turn to the ABC framework for model inference and parameter estimation.

3.2.2 Methods: Approximate Bayesian Computation

We assume that we observe some data, $\mathcal{D}$, generated by an underlying model, $\mathcal{M}(\theta)$. The set of model parameters, θ, have prior distribution $\pi(\theta)$. Our goal is to estimate $P(\theta|\mathcal{D})$, the posterior density of the parameters given the data, where $P(\theta|\mathcal{D}) \propto P(\mathcal{D}|\theta)\pi(\theta)$

An ABC approach replaces $P(\mathcal{D}|\theta)$, the likelihood of the observed data given model parameters, with comparisons of observed data and model simulations. A rejection sampler algorithm proceeds as follows. A candidate parameter vector, θ_i, is drawn from the prior distribution and a simulated data set, x_i, is generated from the model $\mathcal{M}(\theta_i)$. The simulated data set is compared to the observed data set by calculating a distance (see section 3.2.3 for details). θ_i is accepted if the distance is less than some threshold (ϵ). The posterior density is approximated by the distribution of accepted parameter values over many repeated draws of θ_i (Pritchard et al., 1999). Very often the rejection rate of a rejection sampler algorithm is very high, necessitating a more sophisticated algorithm (appendix A). We focus on the sequential Monte Carlo (SMC) sampler following Toni et al (2009), but refer the reader to the discussion in appendix A for details of other algorithms.

3.2.3 Methods: Summary statistics and distance metrics

All ABC algorithms rely on summary statistics and distance metrics to calculate the distance between an observed and simulated data set. We define a *summary statistic* as some value computed from one data set that may summarize or capture some information about the observation. We define a *distance metric* as some function of two data sets (or two summary statistics) that may capture some information about both observations. The distance metric may compare the actual data sets, $\rho(\mathcal{D}, x_i)$, or may compare summary statistics of the data, $\rho(S(\mathcal{D}), S(x_i))$. The objective is to choose distance metrics, summary statistics, and thresholds such that $P(\theta|\mathcal{D}) \approx P(\theta|\rho(S(\mathcal{D}), S(x_i)) < \epsilon)$.

Our goal is to select summary statistics and distance metrics without resorting to generating many pilot ABC runs with high computation costs. We propose the following procedure, guided by the idea that summary statistics and distance metrics should be necessary, but sufficient descriptors of the difference between two data sets. First we compile a large set of possible summary statistics and distance metrics by looking to ABC analyses of similar systems or similar data sets and by using the intuition and expertise of scientists familiar with our specific problem. Then we narrow down the field by explicitly quantifying the ability of each metric to track changes in each parameter. This second step does require the simulation of data sets, but the cost is negligible compared to one step of one ABC SMC run.

Among the few examples of ABC with temporal count or temporal frequency data,

distance metrics using the raw data are more common than summary statistics (table 1). Goodness of fit statistics include the square root of the sum of chi-square differences in allelic frequencies over all alleles in the sample (Sousa et al., 2009) and a scaled Freeman-Tukey statistic that compares observed and expected counts of mature microparasites marginalized over the sacrifice time of the host (Drovandi and Pettitt, 2011a). For a data set with many replicated leukocyte cell trajectories (but no random effects), the Kolmogorov-Smirnov distance between the simulated histogram and data was used as a distance metric (Liepe et al., 2012). Another chi-squared goodness of fit statistic on count data was used for an SEIR model of Ebola transmission, with 4 discrete classes (susceptible, exposed, infectious, and removed), analogous to our discrete stage classes (egg, immature, and adult) (McKinley et al., 2009). In fitting tuberculosis models to epidemiological data, (Sánchez et al., 2009) minimize a goodness of fit statistic (GF) of model outputs to data, where $GF = \sum \frac{(Obs-Exp)^2}{0.5(Obs+Exp)^2}$. McKinley et al. (2009) also define an envelope around the observed data with absolute count differences as a threshold for accepting simulated data. Similarly, Walker et al. (2010) use the Euclidean distance between simulated counts of reported cases of SARS and the actual data smoothed with a moving average filter. The average population size and the sample variance of population sizes were used as summary statistics for a stochastic cellular automata model of range expansion (Rasmussen and Hamilton, 2012). The Euclidean distance of the log-sig transformed differences was used as a single distance metric that incorporated all summary statistics.

We combined the examples from past studies with ideas motivated by both the literature and by the population model to come up with a list of potential summary statistics and distance metrics (table 2). For clarity, we will call each combination of summary statistic and distance metric a *distance function* and give a complete description of each metric in Appendix B. To select a subset from this group, we develop a way to quantify the ability of each distance function to track changes in each parameter. To correctly identify a change in a parameter value from θ to $(\theta + \Delta\theta)$, a distance function should classify two data sets as different if one was generated from θ and the other from $\theta + \Delta\theta$. Also the distance function should classify two data sets as the same if both were generated from θ. To quantify each distance function's ability to do this we turn to classification theory, specifically the receiver operating characteristic (ROC).

We use the ROC curve to show the performance of each distance function in discriminating between two groups of data. Group A are distances between data sets all generated with the same parameters θ. Group B are distances between pairs of data sets, where one is generated with parameters θ and one with parameters $(\theta + \Delta\theta)$. The ROC curve plots the proportion of distances in group A correctly identified as "close" versus the proportion of distances in group B that are incorrectly identified as "close." The area under the curve (AUC) is a statistic that summarizes the ROC curve. If the distance function were classifying distances at random, we would expect the AUC to be 0.5. A perfect distance function would always discriminate between two sets of distances at some threshold and would have an AUC of 1.

We quantified the ability of each proposed distance function to correctly identify

changes in parameter values using the AUC statistic. We chose a range of $\Delta\theta$'s around a baseline θ, varying only one parameter at a time. For each value of $\Delta\theta$, we plotted ROC curves using 1000 distances in groups A and B. We calculated the AUC statistic for each value of $\Delta\theta$ and plotted them with boxplots of the distances. Using these graphs and pairwise plots of distances, we identified uncorrelated metrics that discriminated well between groups of data sets simulated with different parameter values.

Commonly in ABC algorithms one distance value is calculated as the Euclidean distance over many summary statistics, creating a circular acceptance region (in 2D space). Summary statistics may be weighted differently to account for different variances or importance. Instead of using one overall measure of distance, we allow many distance functions with equal weighting. Thus for any simulated data set a vector of distances is calculated. Each element in the distance vector must be below a certain threshold for the particle to be accepted. This creates a rectangular acceptance region for a 2 dimensional distance vector (Pritchard et al., 1999).

3.2.4 Methods: ABC SMC algorithm

Instead of moving from the prior distribution to the posterior in one step, an SMC sampler introduces a number of intermediate steps. At the initial step (s=0), we draw a sample of parameters from the prior, $\{\theta_i\}^{(0)}$, which we now call *particles*. This set of particles has a corresponding set of simulated data sets $\{x_i\}^{(0)}$, distance vectors $\{d_i\}^{(0)}$, and weights $\{w_i\}^{(0)}$. Each data set x_i is simulated from the model. Each distance vector d_i is a J-dimensional set of distance values that are calculated with summary statistics and distance metrics. Again we refer to each combination of summary statistic and distance metric as a *distance function*. Using the systematic procedures outlined above, we select a set of J distance functions that we believe accurately track changes in parameter values and contain a sufficient amount of information to discriminate between "close" and "far" data sets. For j in J, the j^{th} distance function, $\rho_j(S(\mathcal{D}), S(x_i))$, is used to compare the observed data set and a simulated data set. We calculate a value, d_{ij}, the distance from the observed data set to the i^{th} model simulation using the j^{th} distance function. The same set of distance functions $\{\rho_j\}$ are used in each step.

In this initial step we keep all N particles and assign them equal weight such that $w_i = 1/N$ for $i = 1, ..., N$, but in subsequent steps we calculate weights (equation 17). For each of the next steps ($s = 1, ..., S$), we resample particles from the previous set $\{\theta_i\}^{(s-1)}$ by weights $\{w_i\}^{(s-1)}$. We perturb a sampled particle using a Gaussian kernel with a mean of zero and a variance equal to twice the sample variance of parameter values at the previous step. Again each data set x_i is simulated from the model. Again, each value in each distance vector, d_{ij}, is calculated using the distance functions, $\{\rho_j\}$.

Unlike the initial step, where all particles were kept, we keep only those particles that pass a test. For any particle θ_i we require every d_{ij} in j=1, ..., J, to be less than some threshold, ϵ_j. If $d_{ij} < \epsilon_j$ for all j=1, ..., J, then we keep θ_i. Weights are calculated according to the argument for a backwards perturbation kernel by Beaumont et al. (2009) (equation 17). We repeat this process until we have N new particles that pass our threshold test.

ABC SMC algorithm
All notation follows table 3
1. At s=0 initialize
    Repeat N times:
    1.1 Draw a particle $\theta_i \sim \pi(\theta)$
    1.2 Simulate data $x_i \sim \mathcal{M}(\theta_i)$
    1.3 Calculate the distance vector d_i for the set of J distance functions $\{\rho_j\}$:
        for $j = 1, ..., J$: $d_{ij} = \rho_j(S(\mathcal{D}), S(x_i))$
    1.4 Set weight $w_i = 1/N$
2. For each step s=1, ..., S:
    2.1 For each distance function reset the threshold:
        for $j = 1, ..., J$: $\epsilon_j = q_s^{th}$ quantile of $\{d_i\}_j^{(s-1)}$
    2.2 For each parameter reset the perturbation variance:
        for $l = 1, ..., L$: $\tau_l^2 = 2\text{Var}(\{\theta_i\}_l^{(s-1)})$
    2.3 Repeat until N particles are accepted
        2.3.1 Draw $\theta^* \sim \{\theta_i\}^{(s-1)}$ with probabilities $\{w_i\}^{(s-1)}$
        2.3.2 Perturb the particle to get $\theta_i \sim K(\theta|\theta^*)$
            if $\pi(\theta_i) = 0$ return to 2.3.1
        2.3.3 Simulate data $x_i \sim \mathcal{M}(\theta_i)$
        2.3.4 Calculate the distance vector d_i for the set of J distance functions $\{\rho_j\}$:
            for $j = 1, ..., J$: $d_{ij} = \rho_j(S(\mathcal{D}), S(x_i))$
        2.3.5 Reject particles with distances greater than the threshold:
            for $j = 1, ..., J$: if $d_{ij} \geq \epsilon_j$, return to 2.3.1
            otherwise, accept θ_i into $\{\theta_i\}^{(s)}$
        2.3.6 Calculate the weight for particle θ_i

$$w_i = \frac{\pi(\theta_i)}{\sum_{k=1}^{N}(w_k^{(s-1)} \prod_{l=1}^{L} \varphi[\tau_l^{-1}(\theta_{il}^{(s)} - \theta_{kl}^{(s-1)})])} \tag{17}$$

        return to 2.3.1
    2.4 Normalize the weights

We define all notation in table 3 and give details below on the adaptive thresholds and weighting scheme. In considering many distance functions, we must also consider many thresholds. The j^{th} distance function has a corresponding threshold ϵ_j, set to the value that would have included the particles with the smallest $q_s\%$ of distances in step (s-1) (Drovandi and Pettitt, 2011a). Tolerances are specific to the distance function so that we need not worry about scale or weighting of some dimensions over others. As s increases, each of the thresholds decreases and we move closer to the target posterior: $\epsilon_j^{(1)} > \epsilon_j^{(2)} > \cdots > \epsilon_j^{(S)} > 0$. We use a particle weighting scheme after Beaumont et al. (2009), with the formula explicitly given for the case of an L-dimensional parameter vector with independent Gaussian perturbation kernels (equation 17).

We evaluate algorithm performance in several ways. We fit the population model to simulated data, showing the posterior distributions. We compute 95% credible intervals for

the estimates of all parameters, where the probability of a value less than the interval and the probability of a value greater than the interval were both 0.025. We assess the convergence of the intermediate distributions to the posterior by comparing the distributions and by tracking the mean, median, and variance over all iterations. We evaluate the choice of threshold schedule using the effective sample size (ESS) at each step where

$$\mathrm{ESS}(\{w_i\}^{(s)}) = (\sum_{i=1}^{N}(w_i^{(s)})^2)^{-1} \tag{18}$$

Under equal particle weights, ESS=N. ESS values decrease as particles differentially contribute to the next step.

3.2.5 Methods: Algorithm settings

The ABC SMC algorithm was programmed in R (R Development Core Team, 2010). The population model was programmed and compiled in C and called from R to increase efficiency. We initiate a population with a cohort of 10 newly laid eggs (birth times equal 0). The shape parameter of egg stage duration and the shape parameter of immature stage duration were both fixed at 6 to reduce the number of parameter dimensions involved in applying a new method. Having fewer parameters allowed more simulations and a more extensive evaluation of the performance of our metric selection method and the ABC SMC algorithm itself. Once we have been satisfied as to estimation ability and model performance, we can extend the use of the method to real problems with more parameter dimensions to estimate. We used parameter values
$\{1/\lambda_e = 4, \gamma_e = 6, 1/\lambda_i = 3, \gamma_i = 6, m = 0.6, r = 5\}$ to simulate 1000 data sets and chose from them one representative data set as the "observed data" for evaluating our algorithms before trying them on many simulated data sets.. The number of particles (N) in each step was 10,000. Uninformative uniform priors were used for each parameter, with limits based on biological realism for a small, quickly developing, arthropod species. The quantiles used to determine thresholds were $\{0.9, 0.8, 0.7, 0.6, 0.6, 0.6, 0.5, 0.4, 0.4, 0.4\}$ for S=10 steps.

3.3.1 Results: Summary statistics and distance metrics

We evaluated 21 summary statistics and distance metrics for their ability to correctly classify two data sets as similar or different (table 2, Appendix B). We first chose a subset of metrics by AUC value. All metrics have AUC values of 0.5 when comparing data sets generated by the same parameters (zero change in parameter values). We select metrics for which AUC values increase as steeply as possible with increasing or decreasing parameter values. Next, using pairwise plots of metric distances, we eliminate metrics that are highly correlated with others in the subset, narrowing our list to 4: the sum of Chi-square differences in counts of individuals in each stage at each observation, the sum of squared differences in the number of adults at each observation, the cross-correlation function (with zero time lag) between the observed and simulated number of immatures over all observations, and the sum of earth mover distances between the distribution of individuals between stages at each observation (figure 5).

3.3.2 Results: Parameter estimation

The ABC SMC algorithm yielded posterior distributions for each parameter (figure 6). The 95% credible intervals contained the true value of each parameter. Distributions were all unimodal, but the variance varied widely between parameters, with much smaller variances for the distribution of egg scale and immature scale as opposed to fecundity and mortality.

We also evaluated the performance of the algorithm convergence to the posterior. Over the 10 resampling steps, the intermediate distributions shifted from the prior to a steady target posterior (figure 7). The mean and median of the parameter values approached the steady-state values, indicating these are likely to have converged to the desired posterior mean and median (figure 8). The variances for egg and immature scale also stabilized, while those for mortality and fecundity have nearly stabilized but show small downward trends even after 10 steps. ESS values decreased slightly from 10,000 to approximately 8500 (not shown), but did so smoothly, indicating an acceptable threshold schedule.

3.4. Discussion

Our study has shown that an ABC SMC algorithm is able to fit complex population models to variable stage-structured cohort data. The 95% credible intervals contained the true values of the parameters, although the variance of the posterior estimates depended on the parameter. Additionally we have shown that careful a priori investigation of summary statistics and distance metrics can greatly reduce the computation time by eliminating costly multiple ABC SMC runs.

This tool will allow ecologists to fit more realistic models to data, estimating parameters and making inference about processes that might be very important to dynamics. For example, knowledge about the shape of stage distributions, as opposed to a constant rate, may make it easier to detect real differences in distributions between treatment groups. Stage-structured models have long been used in conservation, where the ability to predict changes in real populations is vital. The ability to include realistic amounts of individual heterogeneity would provide more insight for managers evaluating specific conservation plans. Indeed, the flexibility of the ABC SMC algorithm makes it likely that it will perform well across many different systems with varying model assumptions.

The specific SMC sampler is one of many algorithms that increases the efficiency of a simple rejection algorithm (see Appendix A). All of these algorithms improve on the large inefficiencies in a rejection algorithm caused by the disparity between the target posterior and an uninformative prior. The SMC sampler in particular produces uncorrelated samples and does not become stuck in areas of low probability as an MCMC chain might. The algorithm eliminates particles that do not represent the posterior in favor of those that do, causing the particles to move more quickly from the prior to the posterior. The SMC algorithm also allows us to observe the intermediate distributions of parameter estimates, yielding information on the convergence to the posterior. Another advantage of an ABC SMC algorithm is the fact that it is independent of the specific simulation model. You can

apply the algorithm to different problems by plugging in a different simulation model.

One limitation of any ABC algorithm is the computation cost involved simulating the large number of data sets needed. This cost cannot be avoided, but it can be minimized by efficient code and in the choice of programming language. Another limitation of the ABC approach is the challenging issue of choosing summary statistics and distance metrics. The approach we have identified here avoids the huge computation time it would take to conduct separate ABC SMC runs with different subsets of summary statistics and distance metrics. Systematically investigating a large set of metrics and choosing one subset for use in the algorithm allows us to spend the computation time in the actual algorithm, producing more accurate posteriors.

The ABC framework is a particularly good fit for ecologists, who commonly represent their systems with a straightforward and intuitive stochastic model. Software packages for use in population genetics are already widely available (Purcell et al., 2012; Liepe et al., 2010). The ABC framework can easily incorporate many extensions important for ecologists, such as correlations between parameters (Drovandi and Pettitt, 2011a), model uncertainty (Wilkinson, 2008; Ratmann et al., 2009), and model selection (Toni et al., 2009; François and Laval, 2011; Grelaud et al., 2009). ABC provides an appealing alternative to likelihood based inference in situations where the likelihood is intractable or too computationally expensive to approximate and remains flexible enough to be useful in a wide range of systems.

4 Legends, Tables, and Figures

Table 1: Summary statistics and distance metrics used in previous studies of models or data sets similar to the model or data set we consider.

Table 2: The entire set of summary statistics and distance metrics under consideration.

Table 3: Notation used in the ABC SMC algorithm

Figure 1: Box and whisker plots of mixed-effects and fixed-effects model parameter estimates for a) rate parameter λ, b) fixed-effect coefficient β, c) shape parameter γ, d) random-effect standard deviations σ_{source} and σ_{rep}. True values are represented by dashed lines at $\lambda = 0.12$, $\beta = 0.3$, $\gamma = 8$ and 3 levels of standard deviation: $\sigma = 0$, 0.05, or 0.15. Hatched plots show parameter estimates for simulated data with low variance ($\sigma_{source} = 0, \sigma_{rep} = 0.05$). White plots show parameter estimates for simulated data with high variance ($\sigma_{source} = 0.05, \sigma_{rep} = 0.15$).

Figure 2: Power curves for four combinations of variation: a) $\sigma_{source} = 0.05, \sigma_{rep} = 0.05$, b) $\sigma_{source} = 0.05, \sigma_{rep} = 0.15$, c) $\sigma_{source} = 0.15, \sigma_{rep} = 0.05$, d) $\sigma_{source} = 0.15, \sigma_{rep} = 0.15$. The results from the mixed-effects models are represented by solid lines; the results from fixed-effects models by dashed lines. likelihood ratio tests were done at type I error rate $\alpha = 0.05$, represented by a star at $\beta = 0$.

Figure 3: Map of sampled vineyards in San Joaquin, Scaramento and Yolo Counties, surrounding the town of Lodi, California. Rivers are shown as solid lines. "C" markers denote fields of Chardonnay grapes and "Z" markers denote Zinfandel. The white square on the state map shows the location of our sites within California.

Figure 4: Results from the time-to-maturation model and generalized linear mixed model analysis. Throughout, data from Zinfandel source vineyards are in red, those from Chardonnay source vineyards are in green. a) Histograms of development times, grouped by source cultivar. Best-fit Weibull models for Zinfandel (solid red line) and Chardonnay (solid green line) source vineyards are shown, along with dashed lines that represent the extreme curves (0.05 and 0.95 quantile) under the estimated random effects variation for the sample colony and for the replicate assay populations. b) Binomial GLMM of the probability of individuals maturing before day 6. c) Binomial GLMM of the proability of individuals surviving through day 4, d) Poisson GLMM of the number of eggs laid per mature adult female by day 2.

Figure 5: Distance boxplots and AUC plots used to evaluate the ability of each distance function to track changes in each parameter: the scale parameter of egg stage duration, the scale parameter of immature stage duration, mortality, and fecundity, by column. Each row of plots represents the same distance function over all parameters. The first row uses the sum of χ^2 differences in counts of individuals in each stage at each time. The second row uses the sum of squared differences in the number of adults at each time. The third row

uses the cross-correlation function (with zero lag) between the counts of immatures at each time. The fourth row uses the Earth Mover Distance between the distributions at each time of the relative number of individuals in each stage. See Appendix B for further explanation of metrics. Boxplots show the distribution of distances between data sets simulated with baseline parameters and data sets simulated with slightly different parameters. The center boxplot in each plot shows the distribution of distances between data sets simulated from the same baseline parameters. Each plot tracks changes in one parameter while he other 3 are held constant. Distance scale is on the left axis. Open circle show the AUC statistic of the ROC curve for each pair of distributions of distances. AUC scale is on right axis.

Figure 6: Posterior distributions of parameters for a) the scale of egg stage duration, b) the scale of immature stage duration, c) mortality, d) fecundity. Red lines indicate the true parameter values that were used to simulate the "observed" data set. Solid lines show the median and dashed lines show the bounds of the 95% credible interval.

Figure 7: Intermediate distributions of parameter estimates as boxplots for each of the 4 parameters over 10 steps of the algorithm.

Figure 8: The changes in mean (green), median (blue), and variance (red) of parameter estimates for a) the scale of egg stage duration, b) the scale of immature stage duration, c) mortality, d) fecundity over the 10 steps of the algorithm. Plot c shows the scale of the mean and median of mortality estimates on the left axis and the scale of the variance of estimates on the right axis.

summary statistic	distance metric	source
data	envelope of absolute differences $\mathcal{X}^2$ goodness of fit	McKinley et al. (2009)
data	Kolmogorov-Smirnov distance	Liepe et al. (2012)
average pop. size sample var of pop. size	euclidean distance of log-sig transforms	Rasmussen and Hamilton (2012)
data	square root of the sum of $\mathcal{X}^2$	Sousa et al. (2009)
data	Freeman-Tukey statistic	Drovandi and Pettitt (2011a)
data smoothed with a moving average filter	euclidean distance	Walker et al. (2010)

Table 1: Literature summary statistics and distance metrics

summary statistic	distance metric
individuals in each stage at each time total individuals at each time	sum of $\mathcal{X}^2$
eggs at each time new eggs at each time immatures at each time adults at each time individuals in each stage at each time total individuals at each time	sum of squared differences
relative stage class distribution transformed stage class distribution	
relative stage class distribution	sum of Kullback-Leibler divergence sum of squared Kullback-Leibler divergence sum of Bhattacharyya distance sum of Hellinger distances sum of squared Hellinger distances sum of earth mover distance
eggs at each time immatures at each time adults at each time total individuals at each time	cross-correlation function
individuals in each stage at each time	ln $Pr[X = x_i]$ where $X \sim Pois(\mathcal{D})$

Table 2: All summary statistics and distance metrics

Symbol	Value	Definition
S	10	number of steps
s		indicates current step
N	10,000	number of particles at each step
i		indicates current particle
θ		particle (model parameter vector)
θ^*		intermediate particle
L	4	number of parameters
l		indicates current parameter
$\pi(\theta)$	Uniform	joint prior distribution of parameters
θ_{il}		l^{th} parameter value in the i^{th} particle
$\{\theta_i\}^{(s)}$		set of all particles at step s
$\{\theta_i\}_l^{(s)}$		set of the l^{th} parameter in all particles at step s
$K(\theta\|\theta^*)$	$\mathcal{N}(0,\tau^2)$	perturbation kernel - distribution of perturbed particles given an intermediate particle
τ^2		variance of perturbation kernel
$\mathcal{M}(\theta)$		population model
x		simulated data set
$\mathcal{D}$		observed data set
$S()$		summary statistic that may be used in a distance function
d		vector of distances
$\rho()$		distance function
J	4	number of distance functions
j		indicates current distance function
$\{d_i\}^{(s)}$		set of all distance vectors at step s
$\{d_i\}_j^{(s)}$		set of the j^{th} distance in all distance vectors at step s
$\{\rho_j()\}$		set of all distance functions
ϵ		threshold
q		fraction denoting the quantile
w		weight
$\{w_i\}^{(s)}$		set of all weights at step s
$\varphi()$	$\mathcal{N}(0,1)$	standard Normal distribution

Table 3: Algorithm notation

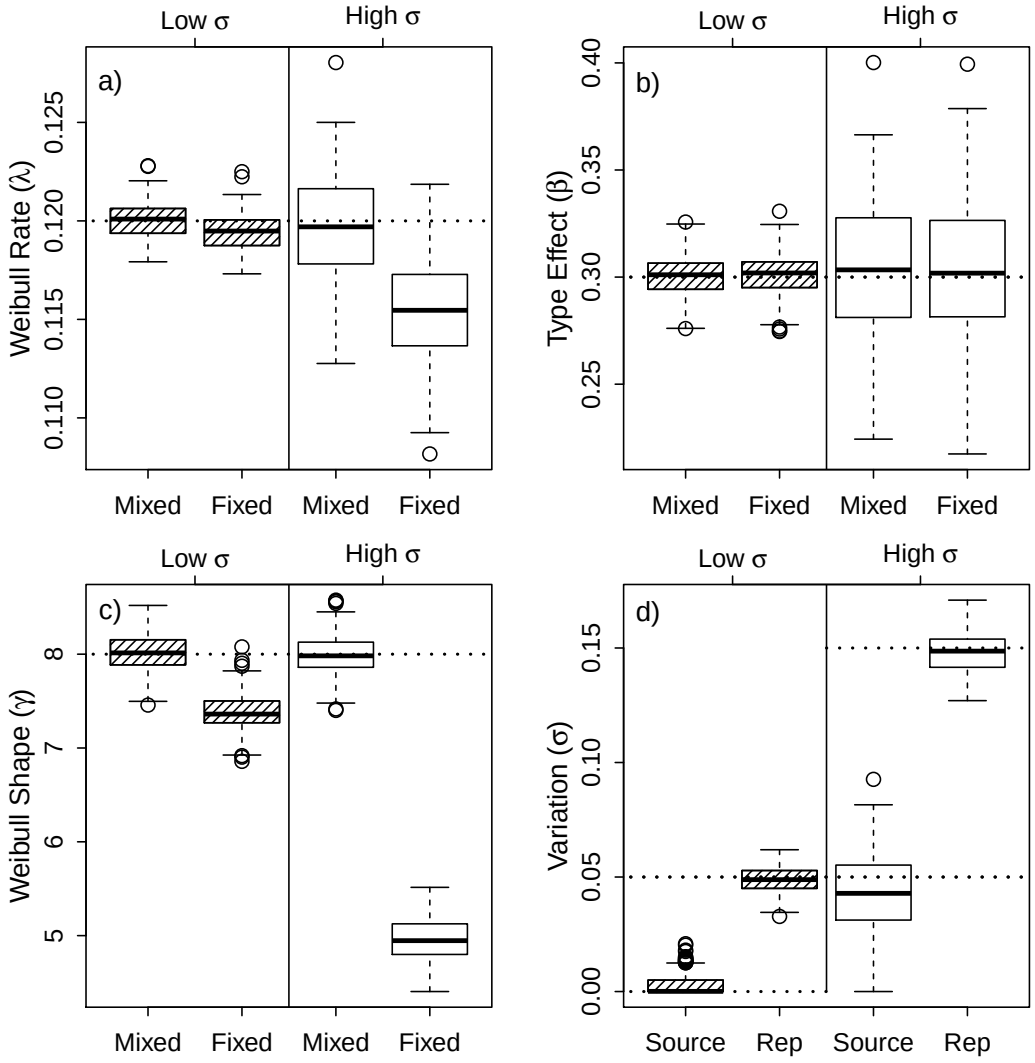

Figure 1: Model parameter estimates

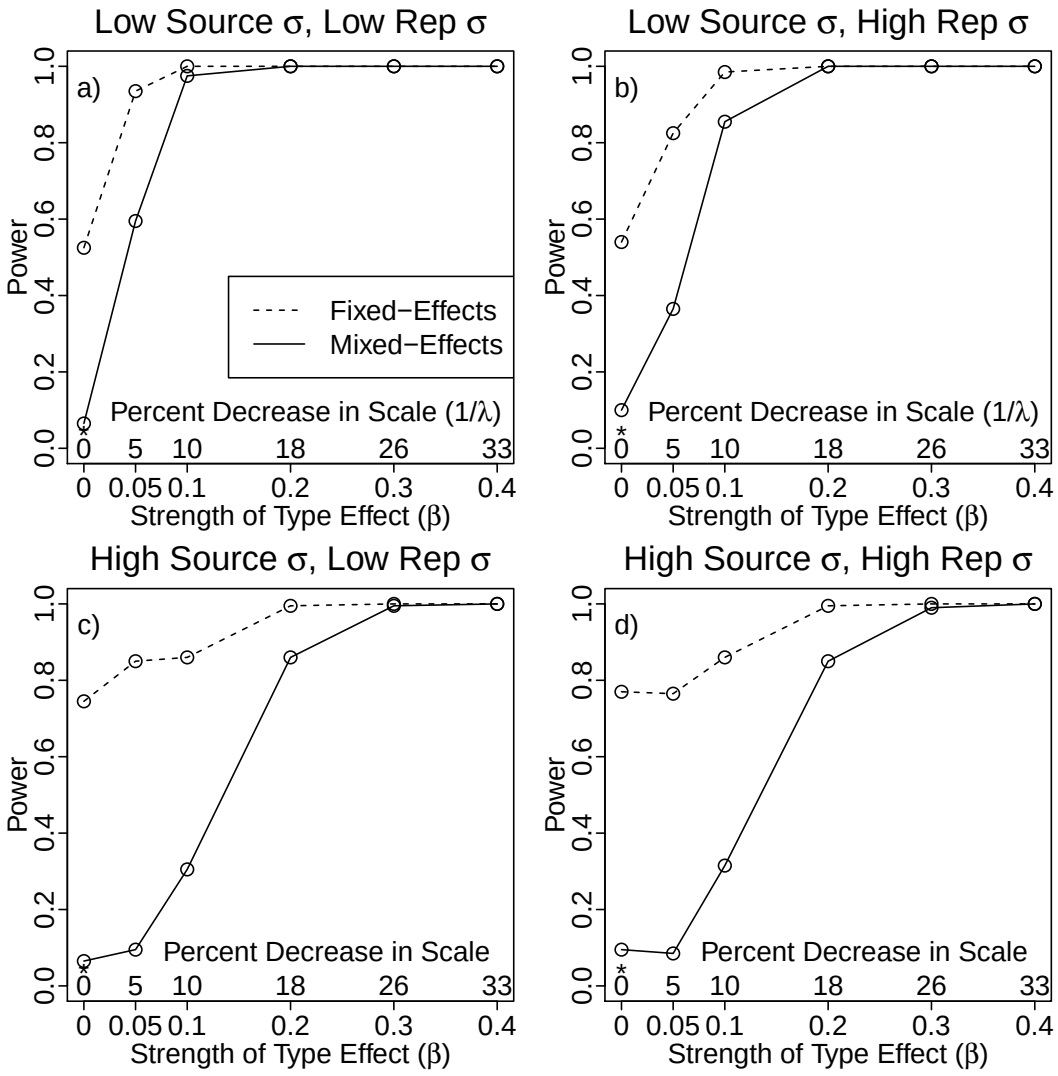

Figure 2: Power analysis

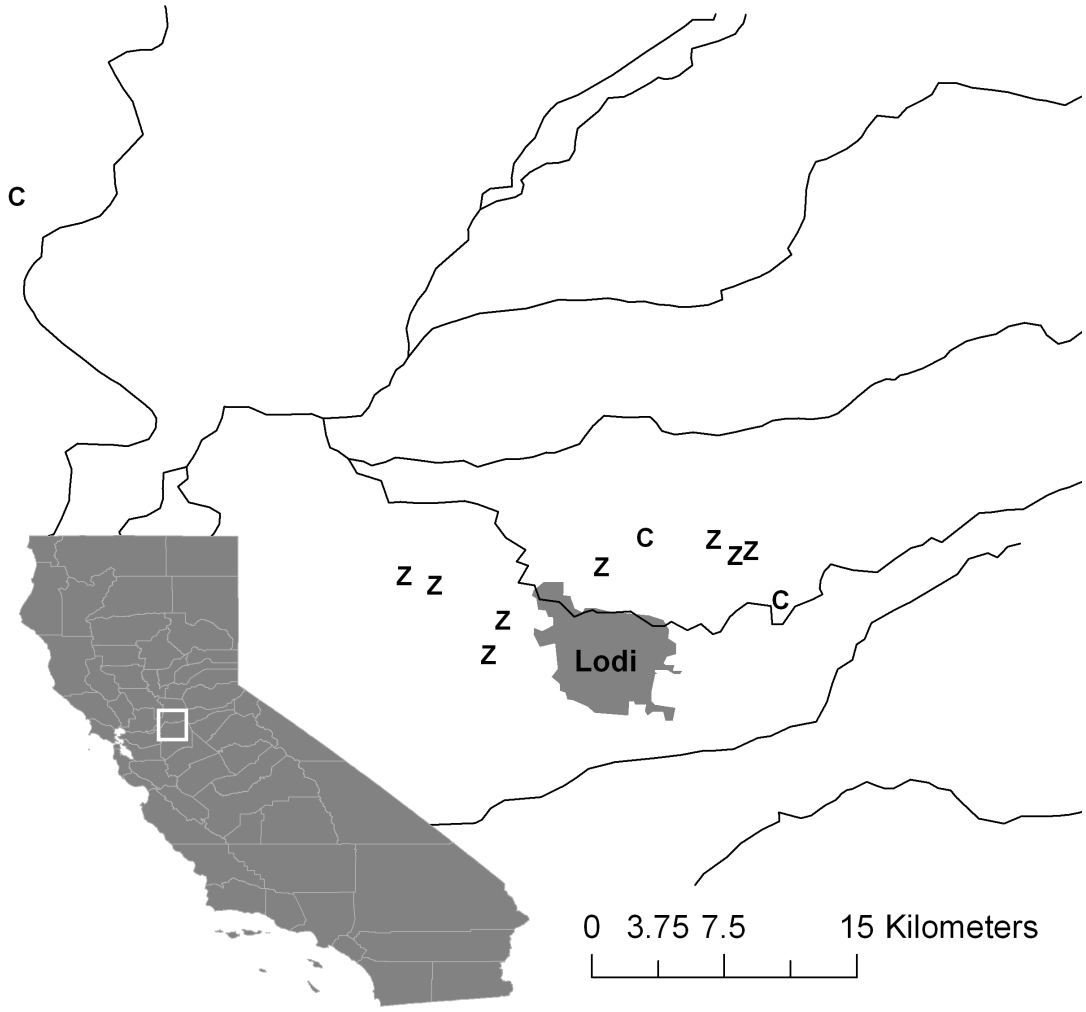

Figure 3: Site map

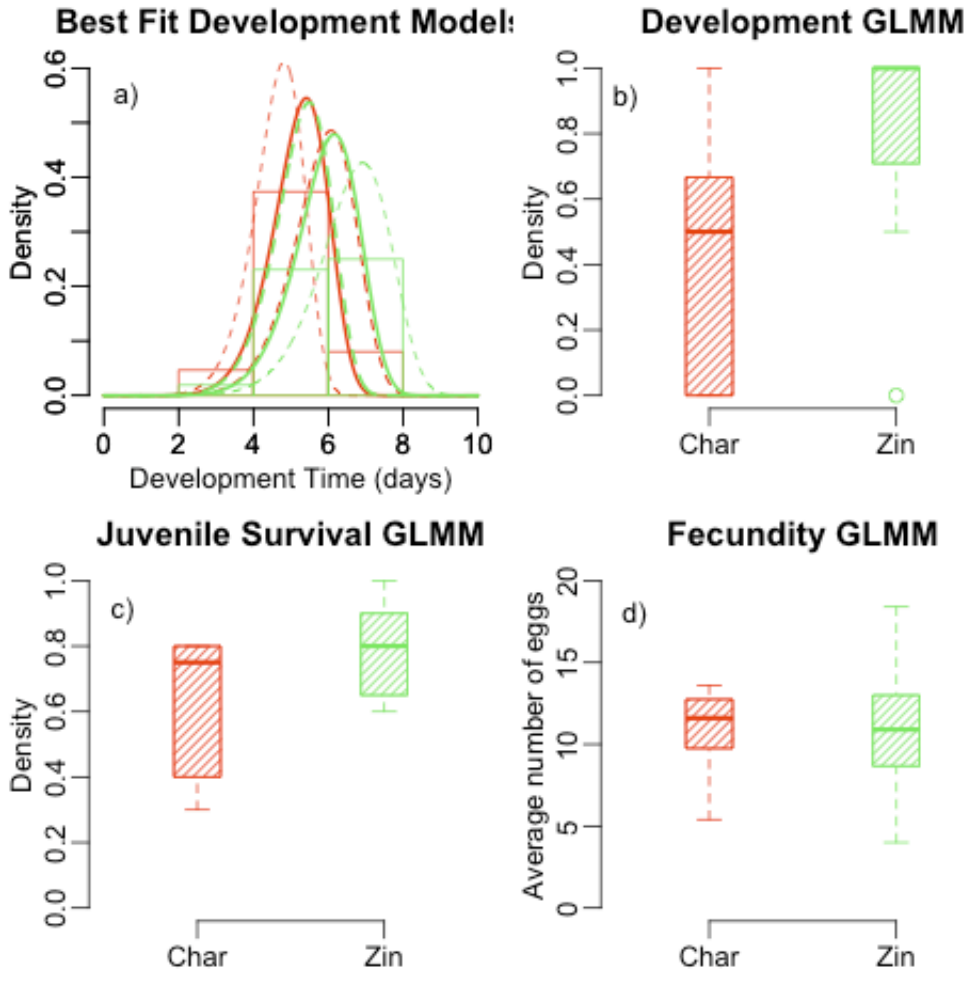

Figure 4: Demographic rate results

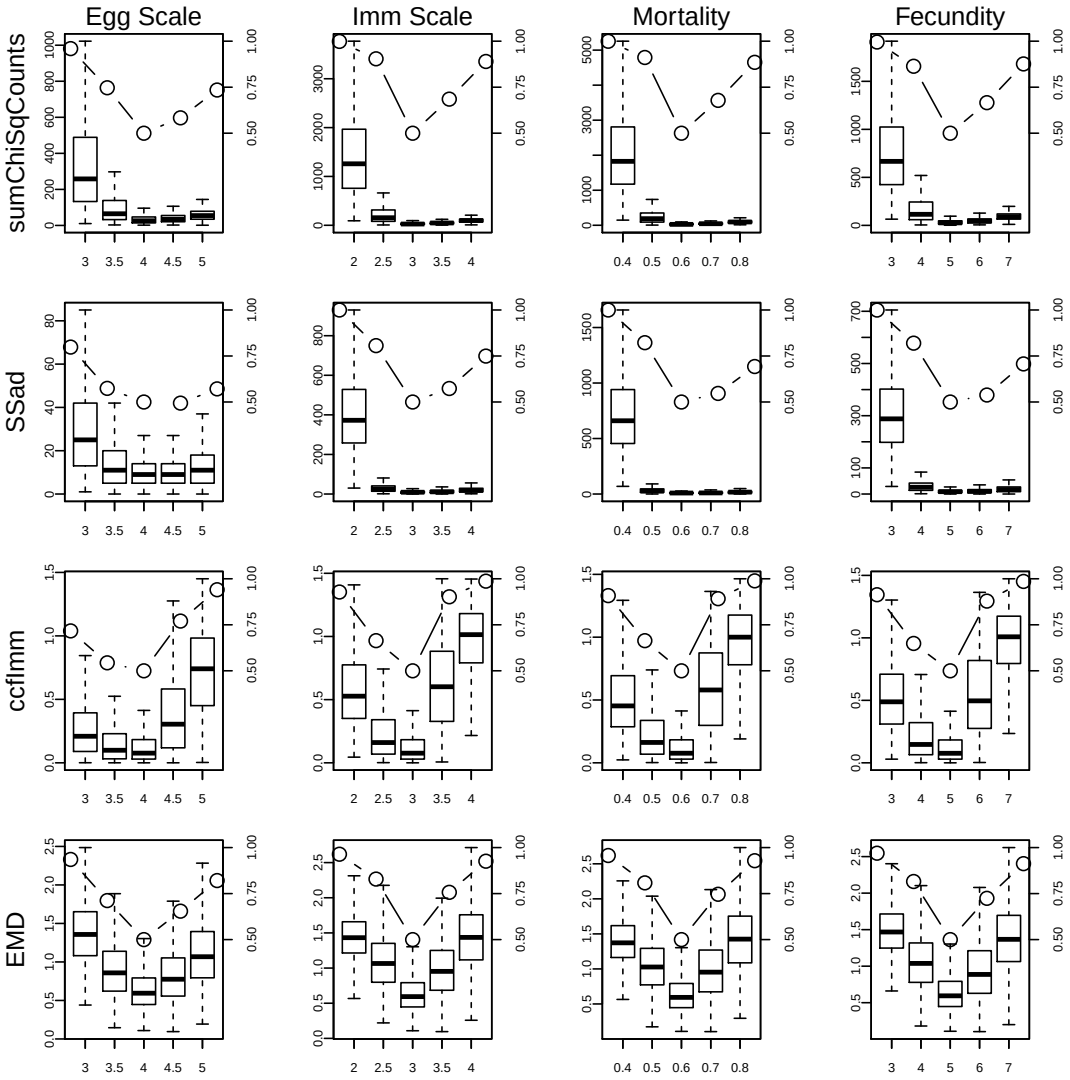

Figure 5: Distance metrics

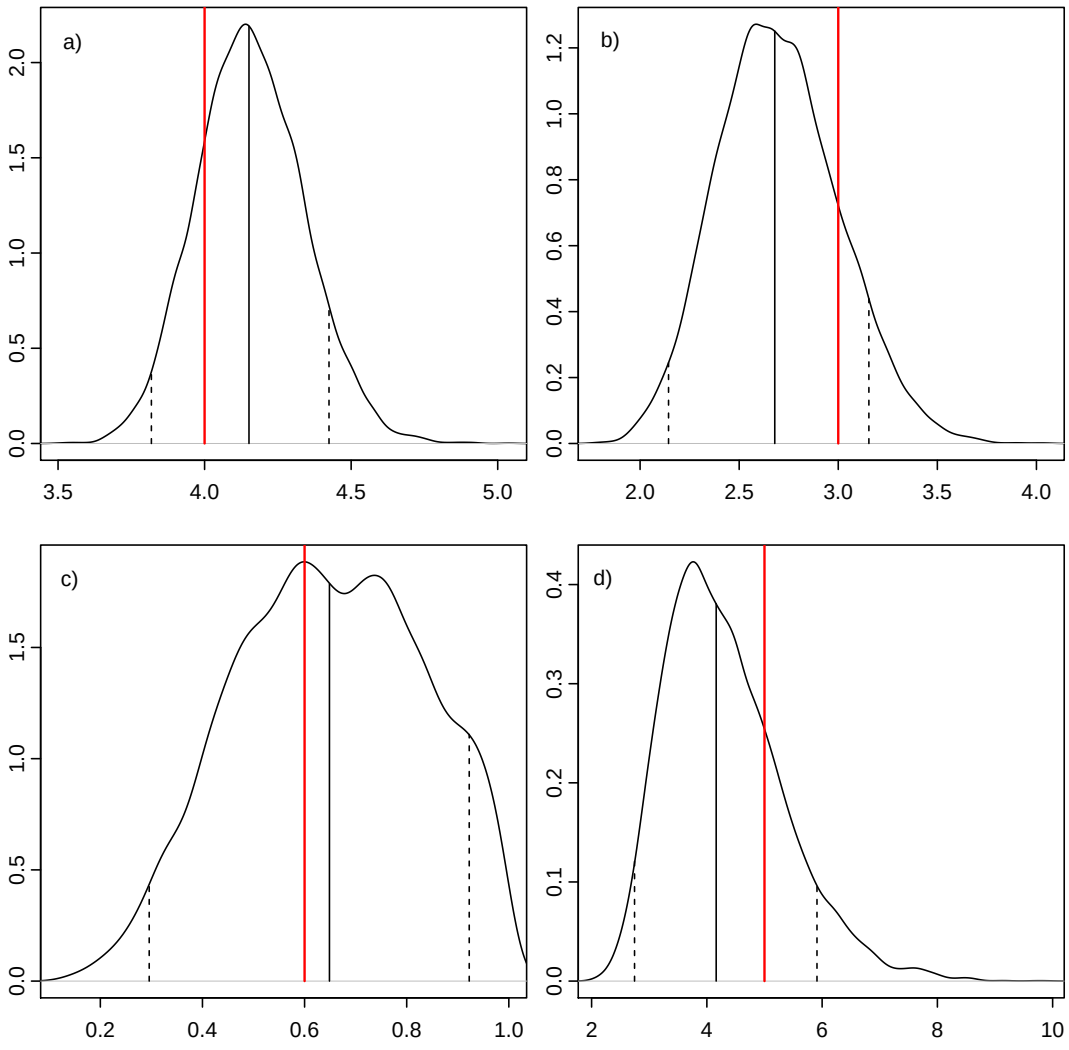

Figure 6: Posterior distributions

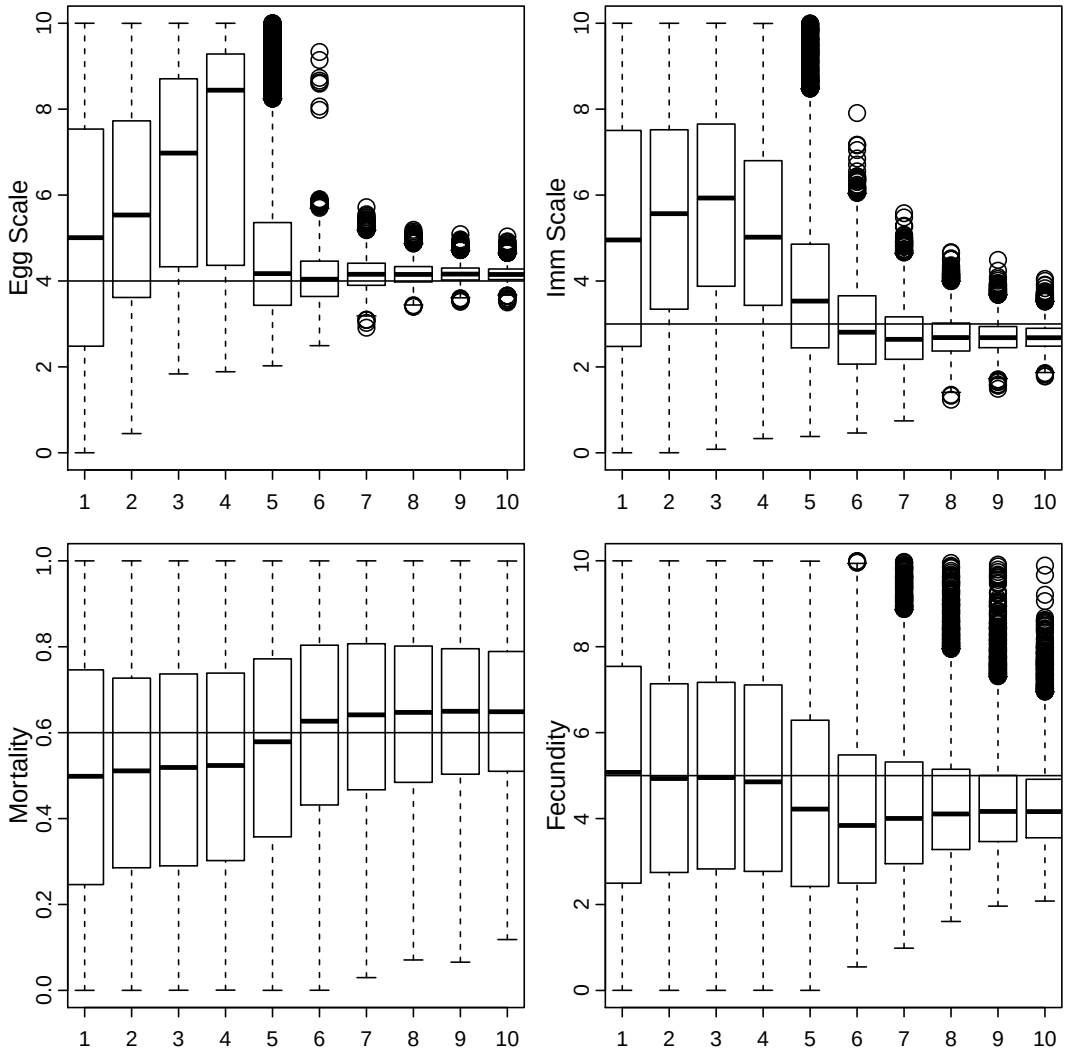

Figure 7: Intermediate distributions

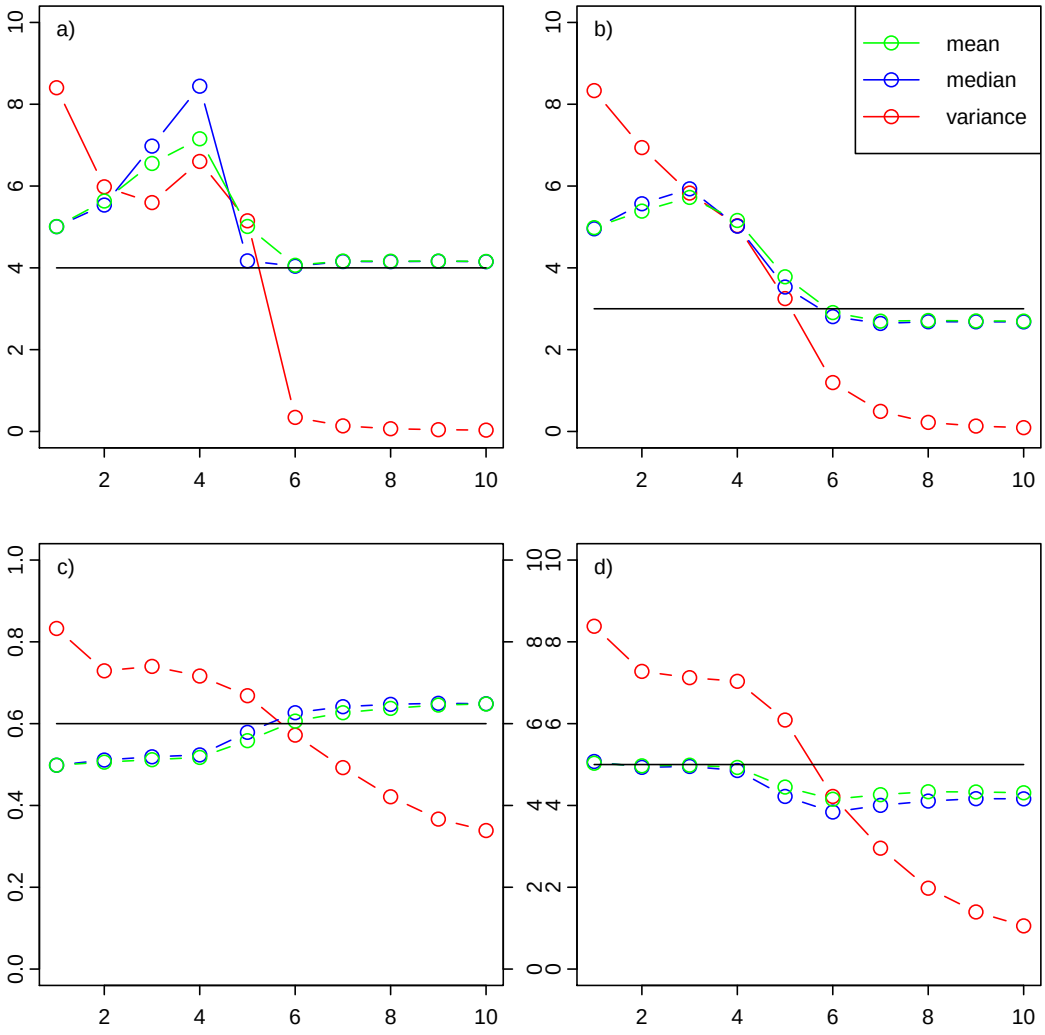

Figure 8: Parameter means, medians, and variances

References

Agrawal, A. A. and R. Karban, 2000. Specificity of constitutive and induced resistance: pigment glands influence mites and caterpillars on cotton plants. *Entomologia Experimentalis et Applicata* **96**:39–49.

Agrawal, A. A., F. Vala, and M. W. Sabelis, 2002. Induction of preference and performance after acclimation to novel hosts in a phytophagous spider mite: Adaptive plasticity? *The American Naturalist* **159**:553–565.

Aldridge, C. L. and M. S. Boyce, 2007. Linking occurrence and fitness to persistence: habitat-based approach for endangered greater sage-grouse. *Ecological Applications* **17**:508–526.

Banerjee, S. and B. P. Carlin, 2004. Parametric spatial cure rate models for Interval-Censored Time-to-Relapse data. *Biometrics* **60**:268–275.

Barnes, C. P., D. Silk, and M. P. H. Stumpf, 2011. Bayesian design strategies for synthetic biology. *Interface Focus* **1**:895–908.

Bates, D. and M. Maechler, 2010. Linear mixed-effects models using S4 classes. Available at http://lme4.r-forge.r-project.org/.

Bazin, E., K. J. Dawson, and M. A. Beaumont, 2010. Likelihood-free inference of population structure and local adaptation in a Bayesian hierarchical model. *Genetics* **185**:587–602.

Beaumont, M. A., 2010. Approximate Bayesian computation in evolution and ecology. In D. J. Futuyma, H. B. Shafer, and D. Simberloff, editors, *Annual Review of Ecology, Evolution, and Systematics, Vol 41*, volume 41, pages 379–406. Annual Reviews, Palo Alto.

Beaumont, M. A., J. Cornuet, J. Marin, and C. P. Robert, 2009. Adaptive approximate Bayesian computation. *Biometrika* **96**:983–990.

Beaumont, M. A., W. Zhang, and D. J. Balding, 2002. Approximate Bayesian computation in population genetics. *Genetics* **162**:2025–2035.

Bellamy, S. L., Y. Li, L. M. Ryan, S. Lipsitz, M. J. Canner, and R. Wright, 2004. Analysis of clustered and interval censored data from a community-based study in asthma. *Statistics in Medicine* **23**:3607–3621.

Berlocher, S. H. and J. L. Feder, 2002. Sympatric speciation in phytophagous insects: Moving beyond controversy? *Annual Review of Entomology* **47**:773–815.

Birt, A., R. M. Feldman, D. M. Cairns, R. N. Coulson, M. Tchakerian, W. Xi, and J. M. Guldin, 2009. Stage-structured matrix models for organisms with non-geometric development times. *Ecology* **90**:57–68.

Bishop, A. L., I. M. Barchia, and L. J. Spohr, 2000. Models for the dispersal in Australia of the arbovirus vector, culicoides brevitarsis kieffer (Diptera: ceratopogonidae). *Preventive Veterinary Medicine* **47**:243–254.

Bjornstad, O. N. and T. F. Hansen, 1994. Individual variation and population dynamics. *Oikos* **69**:167–171.

Blum, M. G. B., 2010. Approximate Bayesian computation: A nonparametric perspective. *Journal of the American Statistical Association* **105**:1178–1187.

Blum, M. G. B. and O. François, 2008. Non-linear regression models for approximate Bayesian computation. *Statistics and Computing* **20**:63–73.

Blum, M. G. B. and V. C. Tran, 2010. HIV with contact tracing: A case study in approximate Bayesian computation. *Biostatistics* **11**:644–660.

Boese, B. L., P. J. Clinton, D. Dennis, R. C. Golden, and B. Kim, 2008. Digital image analysis of zostera marina leaf injury. *Aquatic Botany* **88**:87–90.

Bortot, P., S. G. Coles, and S. A. Sisson, 2007. Inference for stereological extremes. *Journal of the American Statistical Association* **102**:84–92.

California Department of Food and Agriculture and USDA National Agricultural Statistics Service, California Field Office, 2010. California grape acreage report: 2009 crop.

Cappé, O., S. Godsill, and E. Moulines, 2007. An overview of existing methods and recent advances in sequential Monte Carlo. *Proceedings of the IEEE* **95**:899 –924.

Cappé, O., A. Guillin, J. M. Marin, and C. P. Robert, 2004. Population Monte Carlo. *Journal of Computational and Graphical Statistics* **13**:907–929.

Csilléry, K., M. G. B. Blum, O. E. Gaggiotti, and O. François, 2010. Approximate Bayesian computation (ABC) in practice. *Trends in Ecology & Evolution* **25**:410–418.

de Valpine, P., 2009. Stochastic development in biologically structured population models. *Ecology* **90**:2889–2901.

Del Moral, P., A. Doucet, and A. Jasra, 2006. Sequential Monte Carlo samplers. *Journal of the Royal Statistical Society: Series B (Statistical Methodology)* **68**:411–436.

Del Moral, P., A. Doucet, and A. Jasra, 2012. An adaptive sequential Monte Carlo method for approximate Bayesian computation. *Statistics and Computing* **22**:1009–1020.

Dickey, A. M. and R. F. Medina, 2010. Testing host-associated differentiation in a quasi-endophage and a parthenogen on native trees. *Journal of Evolutionary Biology* **23**:945–956.

Diehl, S. R. and G. L. Bush, 1984. An evolutionary and applied perspective of insect biotypes. *Annual Review of Entomology* **29**:471–504.

Dres, M. and J. Mallet, 2002. Host races in plant-feeding insects and their importance in sympatric speciation. *Philosophical Transactions of the Royal Society B: Biological Sciences* **357**:471–492.

Drovandi, C. C. and A. N. Pettitt, 2011a. Estimation of parameters for macroparasite population evolution using approximate Bayesian computation. *Biometrics* **67**:225–233.

Drovandi, C. C. and A. N. Pettitt, 2011b. Likelihood-free Bayesian estimation of multivariate quantile distributions. *Computational Statistics & Data Analysis* **55**:2541–2556.

Drovandi, C. C., A. N. Pettitt, and M. J. Faddy, 2011. Approximate Bayesian computation using indirect inference. *Journal of the Royal Statistical Society Series C-Applied Statistics* **60**:317–337.

Duso, C. and E. Vettorazzo, 1999. Mite population dynamics on different grape varieties with or without phytoseiids released (Acari : Phytoseiidae). *Experimental and Applied Acarology* **23**:741–763.

English-Loeb, G., R. Karban, and M. Walker, 1998. Genotypic variation in constitutive and induced resistance in grapes against spider mite (Acari : Tetranychidae) herbivores. *Environmental Entomology* **27**:297–304.

English-Loeb, G. and A. Norton, 2006. Lack of trade-off between direct and indirect defence against grape powdery mildew in riverbank grape. *Ecological Entomology* **31**:415–422.

English-Loeb, G., A. P. Norton, D. Gadoury, R. Seem, and W. Wilcox, 2005. Tri-trophic interactions among grapevines, a fungal pathogen, and a mycophagous mite. *Ecological Applications* **15**:1679–1688.

Estoup, A., E. Lombaert, J. Marin, T. Guillemaud, P. Pudlo, C. P. Robert, and J. Cornuet, 2012. Estimation of demo-genetic model probabilities with approximate Bayesian computation using linear discriminant analysis on summary statistics. *Molecular Ecology Resources* **12**:846–855.

Famah Sourassou, N., R. Hanna, I. Zannou, G. Moraes, K. Negloh, and M. W. Sabelis, 2010. Morphological variation and reproductive incompatibility of three coconut-mite-associated populations of predatory mites identified as *Neoseiulus paspalivorus* (Acari: Phytoseiidae). *Experimental and Applied Acarology* **53**:323–338.

Fearnhead, P. and D. Prangle, 2012. Constructing summary statistics for approximate Bayesian computation: Semi-automatic approximate Bayesian computation. *Journal of the Royal Statistical Society Series B-Statistical Methodology* **74**:419–474.

Fieberg, J. and G. Delgiudice, 2008. Exploring migration data using Interval-Censored Time-to-Event models. *Journal of Wildlife Management* **72**:1211–1219.

Fox, G. A., B. E. Kendall, J. W. Fitzpatrick, and G. E. Woolfenden, 2006. Consequences of heterogeneity in survival probability in a population of Florida scrub-jays. *The Journal of Animal Ecology* **75**:921–927.

François, O. and G. Laval, 2011. Deviance information criteria for model selection in approximate Bayesian computation. *Statistical Applications in Genetics and Molecular Biology* **10**:33.

Fry, J. D., 1989. Evolutionary adaptation to host plants in a laboratory population of the phytophagous mite *Tetranychus urticae* Koch. *Oecologia* **81**:559–565.

Gee, C. T., D. M. Gadoury, and L. Cadle-Davidson, 2008. Ontogenic resistance to uncinula necator varies by genotype and tissue type in a diverse collection of vitis spp. *Plant Disease* **92**:1067–1073.

Gilbert, M. and J. Grègoire, 2003. Visual, semi-quantitative assessments allow accurate estimates of leafminer population densities: an example comparing image processing and visual evaluation of damage by the horse chestnut leafminer cameraria ohridella (Lep., gracillariidae). *Journal of Applied Entomology* **127**:354–359.

Goethals, K., B. Ampe, D. Berkvens, H. Laevens, P. Janssen, and L. Duchateau, 2009. Modeling interval-censored, clustered cow udder quarter infection times through the shared gamma frailty model. *Journal of Agricultural, Biological, and Environmental Statistics* **14**:1–14.

Gomi, K. and G. Tetsuo, 1996. Host plant preference and genetic compatibility of the kanzawa spider mite *Tetranychus kanzawi* Kishida (Acari : Tetranychidae). *Applied Entomology and Zoology* **31**:417–425.

Gotoh, T., S. Abe, Y. Kitashima, and S. Ehara, 2007. Divergence in host range and reproductive compatibility in three strains of *Oligonychus gotohi* Ehara (Acari : Tetranychidae). *International Journal of Acarology* **33**:7–13.

Gotoh, T., J. Bruin, M. W. Sabelis, and S. B. J. Menken, 1993. Host race formation in *Tetranychus urticae*: genetic differentiation, host plant preference, and mate choice in a tomato and a cucumber strain. *Entomologia Experimentalis et Applicata* **68**:171–178.

Gotoh, T. and K. Gomi, 2003. Life-history traits of the kanzawa spider mite *Tetranychus kanzawai* (Acari: Tetranychidae). *Applied Entomology and Zoology* **38**:7–14.

Gould, F., 1979. Rapid host range evolution in a population of the phytophagous mite *Tetranychus urticae* Koch. *Evolution* **33**:791–802.

Grbić, M., T. Van Leeuwen, R. M. Clark, S. Rombauts, P. Rouzé, V. Grbić, E. J. Osborne, W. Dermauw, P. C. T. Ngoc, F. Ortego, P. Hernández-Crespo, I. Diaz, M. Martinez, M. Navajas, É. Sucena, S. Magalhães, L. Nagy, R. M. Pace, S. Djuranović, G. Smagghe, M. Iga, O. Christiaens, J. A. Veenstra, J. Ewer, R. M. Villalobos, J. L. Hutter, S. D.

Hudson, M. Velez, S. V. Yi, J. Zeng, A. Pires-daSilva, F. Roch, M. Cazaux, M. Navarro, V. Zhurov, G. Acevedo, A. Bjelica, J. A. Fawcett, E. Bonnet, C. Martens, G. Baele, L. Wissler, A. Sanchez-Rodriguez, L. Tirry, C. Blais, K. Demeestere, S. R. Henz, T. R. Gregory, J. Mathieu, L. Verdon, L. Farinelli, J. Schmutz, E. Lindquist, R. Feyereisen, and Y. Van de Peer, 2011. The genome of *Tetranychus urticae* reveals herbivorous pest adaptations. *Nature* **479**:487–492.

Grelaud, A., C. P. Robert, J. Marin, F. Rodolphe, and J. Taly, 2009. ABC likelihood-free methods for model choice in Gibbs random fields. *Bayesian Analysis* **4**:317–335.

Hamilton, G., M. Currat, N. Ray, G. Heckel, M. Beaumont, and L. Excoffier, 2005. Bayesian estimation of recent migration rates after a spatial expansion. *Genetics* **170**:409–417.

Hargrove, W. W., 1988. A photographic technique for tracking herbivory on individual leaves through time. *Ecological Entomology* **13**:359–363.

Hartig, F., J. M. Calabrese, B. Reineking, T. Wiegand, and A. Huth, 2011. Statistical inference for stochastic simulation models - theory and application. *Ecology Letters* **14**:816–827.

He, F. and R. I. Alfaro, 2000. White pine weevil attack on white spruce: A survival time analysis. *Ecological Applications* **10**:225–232.

Helle, W. and M. Sabelis, 1985. Spider mites: their biology, natural enemies, and control. Elsevier.

Hoeting, J. A., R. L. Tweedie, and C. S. Olver, 2003. Transform estimation of parameters for stage-frequency data. *Journal of the American Statistical Association* **98**:503–514.

Hong, M., M. F. Bugallo, and P. M. Djuric, 2010. Joint model selection and parameter estimation by population Monte Carlo simulation. *IEEE Journal of Selected Topics in Signal Processing* **4**:526–539.

Hougaard, P., 2000. Analysis of multivariate survival data. Springer.

Jabot, F. and J. Chave, 2009. Inferring the parameters of the neutral theory of biodiversity using phylogenetic information and implications for tropical forests. *Ecology Letters* **12**:239–248.

Joyce, P. and P. Marjoram, 2008. Approximately sufficient statistics and Bayesian computation. *Statistical Applications in Genetics and Molecular Biology* **7**:26.

Jung, C. and B. A. Croft, 2001. Aerial dispersal of phytoseiid mites (Acari: Phytoseiidae): Estimating falling speed and dispersal distance of adult females. *Oikos* **94**:182–190.

Jung, H. and P. Marjoram, 2011. Choice of summary statistic weights in approximate Bayesian computation. *Statistical Applications in Genetics and Molecular Biology* **10**:45.

Kant, M. R., M. W. Sabelis, M. A. Haring, and R. C. Schuurink, 2008. Intraspecific variation in a generalist herbivore accounts for differential induction and impact of host plant defences. *Proceedings of the Royal Society B: Biological Sciences* **275**:443 –452.

Karban, R. and G. M. English-Loeb, 1988. Effects of herbivory and plant conditioning on the population dynamics of spider mites. *Experimental & Applied Acarology* **4**:225–246.

Klinkenberg, D. and H. Nishiura, 2011. The correlation between infectivity and incubation period of measles, estimated from households with two cases. *Journal of Theoretical Biology* **284**:52–60.

Lawless, J. F., 2003. Statistical models and methods for lifetime data. Wiley-Interscience.

Li, S. and M. Jakobsson, 2012. Estimating demographic parameters from large-scale population genomic data using approximate Bayesian computation. *BMC Genetics* **13**:22.

Li, Z. T., S. A. Dhekney, and D. J. Gray, 2011. PR-1 gene family of grapevine: a uniquely duplicated PR-1 gene from a Vitis interspecific hybrid confers high level resistance to bacterial disease in transgenic tobacco. *Plant Cell Reports* **30**:1–11.

Liepe, J., C. Barnes, E. Cule, K. Erguler, P. Kirk, T. Toni, and M. P. H. Stumpf, 2010. ABC-SysBio – approximate Bayesian computation in Python with GPU support. *Bioinformatics* **26**:1797–1799.

Liepe, J., H. Taylor, C. P. Barnes, M. Huvet, L. Bugeon, T. Thorne, J. R. Lamb, M. J. Dallman, and M. P. H. Stumpf, 2012. Calibrating spatio-temporal models of leukocyte dynamics against *in vivo* live-imaging data using approximate Bayesian computation. *Integrative Biology* **4**:335–345.

Lindsey, J. C. and L. M. Ryan, 1998. Tutorial in biostatistics methods for interval-censored data. *Statistics in Medicine* **17**:219–238.

Lopes, J. S. and S. Boessenkool, 2010. The use of approximate Bayesian computation in conservation genetics and its application in a case study on yellow-eyed penguins. *Conservation Genetics* **11**:421–433.

Luciani, F., S. A. Sisson, H. Jiang, A. R. Francis, and M. M. Tanaka, 2009. The epidemiological fitness cost of drug resistance in *Mycobacterium tuberculosis*. *Proceedings of the National Academy of Sciences of the United States of America* **106**:14711–14715.

Luedeling, E., A. Hale, M. Zhang, W. J. Bentley, and L. C. Dharmasri, 2009. Remote sensing of spider mite damage in california peach orchards. *International Journal of Applied Earth Observation and Geoinformation* **11**:244–255.

Ma, Z. S., 2010. Survival analysis approach to insect life table analysis and hypothesis testing: with particular reference to russian wheat aphid (Diuraphis noxia (Mordvilko)) populations. *Bulletin of Entomological Research* **100**:315–324.

Ma, Z. S. and E. J. Bechinski, 2008. A survival-analysis-based simulation model for russian wheat aphid population dynamics. *Ecological Modelling* **216**:323–332.

Ma, Z. S. and E. J. Bechinski, 2009. Accelerated failure time (AFT) modeling for the development and survival of russian wheat aphid, Diuraphis noxia (Mordvilko). *Population Ecology* **51**:543–548.

Magalhães, S., E. Blanchet, M. Egas, and I. Olivieri, 2009. Are adaptation costs necessary to build up a local adaptation pattern? *BMC Evolutionary Biology* **9**:182.

Magalhães, S., J. Fayard, A. Janssen, D. Carbonell, and I. Olivieri, 2007a. Adaptation in a spider mite population after long-term evolution on a single host plant. *Journal of Evolutionary Biology* **20**:2016–2027.

Magalhães, S., M. R. Forbes, A. Skoracka, M. Osakabe, C. Chevillon, and K. D. McCoy, 2007b. Host race formation in the Acari. *Experimental and Applied Acarology* **42**:225–238.

Marjoram, P., J. Molitor, V. Plagnol, and S. Tavare, 2003. Markov chain Monte Carlo without likelihoods. *Proceedings of the National Academy of Sciences* **100**:15324–15328.

Marzolin, G., A. Charmantier, and O. Gimenez, 2011. Frailty in state-space model: application to actuarial senescence in the dipper. *Ecology* **92**:562–567.

McKinley, T., A. R. Cook, and R. Deardon, 2009. Inference in epidemic models without likelihoods. *The International Journal of Biostatistics* **5**:24.

Moritz, M. A., T. J. Moody, L. J. Miles, M. M. Smith, and P. Valpine, 2008. The fire frequency analysis branch of the pyrostatistics tree: sampling decisions and censoring in fire interval data. *Environmental and Ecological Statistics* **16**:271–289.

Navajas, M., A. Tsagkarakov, J. Lagnel, and M.-J. Perrot-Minnot, 2000. Genetic differentiation in *Tetranychus urticae* (Acari : Tetranychidae): polymorphism, host races or sibling species? *Experimental and Applied Acarology* **24**:365–376.

Nishimura, S., N. Hinomoto, and A. Takafuji, 2005. Gene flow and spatio-temporal genetic variation among sympatric populations of *Tetranychus kanzawai* (Acari: Tetranychidae) occurring on different host plants, as estimated by microsatellite gene diversity. *Experimental and Applied Acarology* **35**:59–71.

Nunes, M. A. and D. J. Balding, 2010. On optimal selection of summary statistics for approximate Bayesian computation. *Statistical Applications in Genetics and Molecular Biology* **9**:34.

Ozinga, W. A., S. M. Hennekens, J. H. J. Schamine, N. A. C. Smits, R. M. Bekker, C. Rmermann, L. Klime, J. P. Bakker, and J. M. Groenendael, 2007. Local aboveground persistence of vascular plants: Lifehistory tradeoffs and environmental constraints. *Journal of Vegetation Science* **18**:489–497.

Padhukasahasram, B., J. D. Wall, P. Marjoram, and M. Nordborg, 2006. Estimating recombination rates from single-nucleotide polymorphisms using summary statistics. *Genetics* **174**:1517–1528.

Pezet, R., O. Viret, C. Perret, and R. Tabacchi, 2003. Latency of *Botrytis cinerea* Pers.: Fr. and biochemical studies during growth and ripening of two grape berry cultivars, respectively susceptible and resistant to grey mould. *Journal of Phytopathology* **151**:208–214.

Polakow, D. A. and T. T. Dunne, 1999. Modelling fire-return interval T: stochasticity and censoring in the two-parameter Weibull model. *Ecological Modelling* **121**:79–102.

Pritchard, J. K., M. T. Seielstad, A. Perez-Lezaun, and M. W. Feldman, 1999. Population growth of human Y chromosomes: A study of Y chromosome microsatellites. *Molecular Biology and Evolution* **16**:1791–1798.

Purcell, K., N. Ling, and C. Stockwell, 2012. Evaluation of the introduction history and genetic diversity of a serially introduced fish population in New Zealand. *Biological Invasions* **14**:2057–2065.

R Development Core Team, 2010. R: A Language and Environment for Statistical Computing. R Foundation for Statistical Computing, Vienna, Austria.

Rabosky, D. L., 2009. Heritability of extinction rates links diversification patterns in molecular phylogenies and fossils. *Systematic Biology* **58**:629–640.

Radke, B. R., 2003. A demonstration of interval-censored survival analysis. *Preventive Veterinary Medicine* **59**:241–256.

Ramsey, D. S. L., D. M. Forsyth, C. J. Veltman, S. J. Nicol, C. R. Todd, R. B. Allen, W. J. Allen, P. J. Bellingham, S. J. Richardson, C. L. Jacobson, and R. J. Barker, 2012. An approximate Bayesian algorithm for training fuzzy cognitive map models of forest responses to deer control in a New Zealand adaptive management experiment. *Ecological Modelling* **240**:93–104.

Rasmussen, R. and G. Hamilton, 2012. An approximate Bayesian computation approach for estimating parameters of complex environmental processes in a cellular automata. *Environmental Modelling & Software* **29**:1–10.

Ratmann, O., C. Andrieu, C. Wiuf, and S. Richardson, 2009. Model criticism based on likelihood-free inference, with an application to protein network evolution. *Proceedings of the National Academy of Sciences* **106**:10576–10581.

Read, K. and J. R. Ashford, 1968. A system of models for life cycle of a biological organism. *Biometrika* **55**:211–221.

Rhainds, M. and G. English-Loeb, 2003. Variation in abundance and feeding impact of tarnished plant bug (Hemiptera: Miridae) for different cultivars of strawberry: Role of flowering phenology and yield attributes. *Journal of Economic Entomology* **96**:433–440.

Rieckermann, J., J. Anta, A. Scheidegger, and C. Ort, 2011. Assessing wastewater micropollutant loads with approximate Bayesian computations. *Environmental Science & Technology* **45**:4399–4406.

Sánchez, M. S., J. O. Lloyd-Smith, B. G. Williams, T. C. Porco, S. J. Ryan, M. W. Borgdorff, J. Mansoer, C. Dye, and W. M. Getz, 2009. Incongruent HIV and tuberculosis co-dynamics in Kenya: Interacting epidemics monitor each other. *Epidemics* **1**:14–20.

Schauber, E. M. and C. G. Jones, 2006. Comparative predation on naturally occurring gypsy moth (Lepidoptera: lymantriidae) pupae and deployed freeze-dried pupae. *Environmental Entomology* **35**:293–296.

Scranton, K. and P. de Valpine, 2012. A mixed-effects model of variable development time for interval-censored cohort data. *Manuscript submitted for publication* .

Seabright Laboratories, 2009. Stikem special hold fast.

Service, 2000. Heterogeneity in individual mortality risk and its importance for evolutionary studies of senescence. *The American Naturalist* **156**:1–13.

Shriner, D., Y. Liu, D. C. Nickle, and J. I. Mullins, 2006. Evolution of intrahost HIV-1 genetic diversity during chronic infection. *Evolution* **60**:1165–1176.

Sisson, S. A., Y. Fan, and M. M. Tanaka, 2007. Sequential Monte Carlo without likelihoods. *Proceedings of the National Academy of Sciences* **104**:1760–1765.

Sisson, S. A., Y. Fan, and M. M. Tanaka, 2009. Correction for sisson et al., sequential Monte Carlo without likelihoods. *Proceedings of the National Academy of Sciences* **106**:16889–16889.

Škaloudová, B., V. Křivan, and R. Zemek, 2006. Computer-assisted estimation of leaf damage caused by spider mites. *Computers and Electronics in Agriculture* **53**:81–91.

Skoracka, A., 2008. Reproductive barriers between populations of the cereal rust mite *Abacarus hystrix* confirm their host specialization. *Evolutionary Ecology* **22**:607–616.

Souissi, S. and S. Ban, 2001. The consequences of individual variability in moulting probability and the aggregation of stages for modelling copepod population dynamics. *Journal of Plankton Research* **23**:1279 –1296.

Sousa, V. C., M. Fritz, M. A. Beaumont, and L. Chikhi, 2009. Approximate Bayesian computation without summary statistics: The case of admixture. *Genetics* **181**:1507–1519.

Sun, J., 2006. The statistical analysis of interval-censored failure time data. Springer.

Tanaka, M. M., A. R. Francis, F. Luciani, and S. A. Sisson, 2006. Using approximate Bayesian computation to estimate tuberculosis transmission parameters from genotype data. *Genetics* **173**:1511–1520.

Tanhuanpää, M., K. Ruohomaki, and E. Uusipäikka, 2001. High larval predation rate in non-outbreaking populations of a geometrid moth. *Ecology* **82**:281–289.

Tavare, S., D. J. Balding, R. C. Griffiths, and P. Donnelly, 1997. Inferring coalescence times from DNA sequence data. *Genetics* **145**:505–518.

Thaler, J. S. and R. Karban, 1997. A phylogenetic reconstruction of constitutive and induced resistance in gossypium. *The American Naturalist* **149**:1139–1146.

Tien, N. S. H., M. W. Sabelis, and M. Egas, 2010. The maintenance of genetic variation for oviposition rate in two-spotted spider mites: Inferences from artificial selection. *Evolution* **64**:2547–2557.

Tinsley, M., F. I. Lewis, and F. Bruelisauer, 2012. Network modeling of BVD transmission. *Veterinary Research* **43**:11.

Toni, T., G. Jovanovic, M. Huvet, M. Buck, and M. P. H. Stumpf, 2011. From qualitative data to quantitative models: Analysis of the phage shock protein stress response in *Escherichia coli*. *BMC Systems Biology* **5**:69.

Toni, T., Y. Ozaki, P. Kirk, S. Kuroda, and M. P. H. Stumpf, 2012. Elucidating the *in vivo* phosphorylation dynamics of the ERK MAP kinase using quantitative proteomics data and Bayesian model selection. *Molecular Biosystems* **8**:1921–1929.

Toni, T. and M. P. H. Stumpf, 2010. Simulation-based model selection for dynamical systems in systems and population biology. *Bioinformatics* **26**:104–110.

Toni, T., D. Welch, N. Strelkowa, A. Ipsen, and M. P. H. Stumpf, 2009. Approximate Bayesian computation scheme for parameter inference and model selection in dynamical systems. *Journal of The Royal Society Interface* **6**:187–202.

Turner, B. M. and T. Van Zandt, 2012. A tutorial on approximate Bayesian computation. *Journal of Mathematical Psychology* **56**:69–85.

Uchmański, J., 1985. Differentiation and frequency distributions of body weights in plants and animals. *Philosophical Transactions of the Royal Society of London Series B-Biological Sciences* **310**:1–75.

Van Dooren, T. J. M., T. Tully, and R. Ferrière, 2005. The analysis of reaction norms for age and size at maturity using maturation rate models. *Evolution* **59**:500.

van Noordwijk, A. J. and G. de Jong, 1986. Acquisition and allocation of resources: Their influence on variation in life history tactics. *The American Naturalist* **128**:137–142.

Velema, H.-P., L. Hemerik, M. S. Hoddle, and R. F. Luck, 2005. Brochosome influence on parasitisation efficiency of Homalodisca coagulata (Say) (Hemiptera: cicadellidae) egg masses by Gonatocerus ashmeadi (Girault) (Hymenoptera: mymaridae). *Ecological Entomology* **30**:485–496.

Vindenes, Y., S. Engen, and B. Sæther, 2008. Individual heterogeneity in vital parameters and demographic stochasticity. *The American Naturalist* **171**:455–467.

Wajnberg, E., 2006. Time allocation strategies in insect parasitoids: from ultimate predictions to proximate behavioral mechanisms. *Behavioral Ecology and Sociobiology* **60**:589–611.

Walker, D. M., D. Allingham, H. W. J. Lee, and M. Small, 2010. Parameter inference in small world network disease models with approximate Bayesian computational methods. *Physica A: Statistical Mechanics and its Applications* **389**:540–548.

Wegmann, D., C. Leuenberger, and L. Excoffier, 2009. Efficient approximate Bayesian computation coupled with Markov Chain Monte Carlo without likelihood. *Genetics* **182**:1207–1218.

Wilkinson, R. D., 2008. Approximate Bayesian computation (ABC) gives exact results under the assumption of model error. *arXiv:0811.3355* .

Wintrebert, C. M., A. Zwinderman, E. Cam, R. Pradel, and J. van Houwelingen, 2005. Joint modelling of breeding and survival in the kittiwake using frailty models. *Ecological Modelling* **181**:203–213.

Wong, M., K. Lam, and E. Lo, 2005. Bayesian analysis of clustered interval-censored data. *Journal of Dental Research* **84**:817 –821.

Woodall, C., P. Grambsch, and W. Thomas, 2005. Applying survival analysis to a large-scale forest inventory for assessment of tree mortality in minnesota. *Ecological Modelling* **189**:199–208.

Zens, M. and D. Peart, 2003. Dealing with death data: individual hazards, mortality and bias. *Trends in Ecology & Evolution* **18**:366–373.

Zuma, K., 2007. A bayesian analysis of correlated Interval-Censored data. *Communications in Statistics - Theory and Methods* **36**:725–730.

A Appendix: ABC algorithms

Simple rejection sampler algorithms often have very high rejection rates, due to model complexity or stochasticity, or high disparity between the prior and posterior. This issue can be solved by modifying the algorithm to include a regression adjustment (Beaumont et al., 2002), by using ABC in markov chain monte carlo (MCMC) methods (Marjoram et al., 2003), or by employing a particle filter or sequential Monte Carlo approach (Cappé et al., 2004, 2007).

Local linear regression adjustments to the proposed parameter values are made by weighting parameter sets by the distance between their simulated data and the observed data set (Beaumont et al., 2002). Weighted linear or quadratic regressions of the parameter value on the distance yield an intercept estimate that corresponds to the posterior mean of the parameters (Blum and François, 2008). Posterior densities are approximated by kernel density estimation with the weighted parameters (Beaumont et al., 2002). Regression adjustments do not universally improve parameter estimation but do achieve better performance with a homoskedastic relationship between distance and parameter (Blum and François, 2008). Improvements to local linear regression include neural network modelling and importance sampling (Blum and Tran, 2010).

An MCMC approach can also be applied in the ABC framework, where proposed changes to a parameter value are accepted with a probability proportional to some distance from the observed data based on summary statistics (Marjoram et al., 2003). MCMC improves on the inefficiencies of a rejection algorithm, but high rejection rates along with the correlated nature of accepted samples may trap the chain in areas of low probability (Sisson et al., 2007). ABC MCMC algorithms are used to estimate transmission rate, doubling time, reproductive value of a strain of tuberculosis (Tanaka et al., 2006) and the size of ellipsoid imperfections in the manufacturing of clean steel (Bortot et al., 2007).

Algorithms based on the sequential Monte Carlo (SMC) approach improve on the "brute-force" rejection algorithm (Del Moral et al., 2006, 2012). In a SMC algorithm, particles are sampled from a prior distribution and propagated through intermediate distributions until they resemble a sample from the target posterior distribution (Sisson et al., 2007). This propagation can occur through parameter space or through a time series, as in a population Monte Carlo (PMC) algorithm (Cappé et al., 2004). Any SMC algorithm depends on a weighted resampling that favors particles that generate simulations "closer" to the observed data and a set of decreasing thresholds for particle acceptance (Beaumont et al., 2009). The choice of weights can be problematic, with biased posteriors arising from the backwards kernel in the partial rejection control algorithm (Beaumont et al., 2009; Sisson et al., 2009). Using an approximation to the optimal backwards kernel (Beaumont et al., 2009) has a higher computation cost (Del Moral et al., 2012), but may not substantially slow computation time of the overall algorithm (Toni et al., 2009). Using an an MCMC kernel for the parameter transition kernel allows the use of the exact backwards kernel and avoids the extra computation cost (Drovandi and Pettitt, 2011a; Del Moral et al., 2012)

Sequential Monte Carlo algorithms have been used for parameter inference in fields

such as epidemiology, bioinformatics, and cellular biology. Partial rejection control algorithms have been used to estimate generation times and incubation time of measles (Klinkenberg and Nishiura, 2011) and to infer a complex transmission and evolution model of drug resistance in tuberculosis (Luciani et al., 2009). Population Monte Carlo algorithms have been used to estimate population parameters and spatial dynamics in species range expansion (Rasmussen and Hamilton, 2012), and in signal processing with white and colored noise (Hong et al., 2010). ABC SMC algorithms have been used for inference of multivariate quantile distributions (Drovandi and Pettitt, 2011b), models of (bioinformatics) influenza dynamics and signalling pathways (Toni and Stumpf, 2010), stochastic environmental models of wastewater micropollutant loads (Rieckermann et al., 2011), microphage shock response (Toni et al., 2011), an SEIR model for Ebola transmission (McKinley et al., 2009), and phosphorylation pathways (Toni et al., 2012).

B Appendix: Summary statistics and distance metrics

We evaluated 21 combinations of summary statistics and distance metrics based on the literature and our knowledge of population dynamics (table 2). We define the observed data set to which we fit our model as the *true data set* and each model simulation as a *simulated data set*. We describe the summary statistics and distance metrics in detail below.

The main summary statistic we used was the untransformed data: the counts of individuals at each time. We used the total number of individuals over the 5 observation times, but we also separated the individuals into stage for 3 different summary statistics: the counts of eggs, immatures, adults. Additionally we pooled this stage count data together into another summary statistic with 15 data points: counts of 3 separate stages at 5 observation times. We also used the change in the number of eggs between observation times.

Another summary statistic that was used extensively was the relative stage class distributions. At each time we calculated the proportion of individuals in each stage: egg, immature, and adult. These proportions summed to 1 across the 3 stages at each single observation time. The summary statistic then yielded 15 data points: three proportions at 5 observation times. We also used the arcsin transform of the relative stage distributions.

The first distance metric we investigated was the $\mathcal{X}^2$ differences between observed (simulated data set) and expected (true data set). We took the sum of the chi square differences across all data points in the summary statistic. We used the $\mathcal{X}^2$ metric in combination with two summary statistics: the number of individuals in each stage at each time (15 data points) and the total number of individuals at each time (5 data points).

The second distance metric we included was the sum of squared differences between the true data set and a simulated data set. We used the sum of squares metric in combination with all 6 untransformed data summary statistics: the eggs at each time, the new eggs at each time, immatures at each time, adults at each time, total individuals at each time, and the individuals in each stage at each time. We also used the sum of squared differences of relative stage distributions and arcsin transformed stage distributions.

We based another group of distance metrics on the comparison of distributions. Statistics and probability theory contain many methods for quantifying the difference between two distributions: a true distribution P and a second distribution Q. We applied these methods to quantify the difference between the relative stage distributions of the true data set (P) and a simulated data set (Q) at each observation time. The first distance metric in this group was the Kullback-Leibler divergence, which measures the difference between distributions as

$$D_{KL}(P,Q) = \sum_i ln\left(\frac{P(i)}{Q(i)}\right) P(i) \tag{19}$$

where $i = 3$ for the 3 stage classes. The sum of Kullbeck-Liebler divergence took the sum of $D_{KL}(P,Q)$ over all observation times where the sum of squared Kullbeck-Liebler divergence took the sum of squared values of $D_{KL}(P,Q)$ over all observation times. One issue with the Kullback-Leibler divergence is that it is only defined if the $P(i)$ is nonzero. Thus it

provided no information when there are zero individuals in a stage in the true data set.

The next probability-based distance metric we used was the Bhattacharyya distance,

$$D_B(P,Q) = -ln\left(\sum_i \sqrt{P(i)Q(i)}\right) \tag{20}$$

where i=3 for the number of stages. The sum of the Bhattacharyya distance took the sum of $D_B(P,Q)$ over the 5 observation times.

The Hellinger distance was also used, where

$$D_H(P,Q) = \frac{1}{\sqrt{2}}\sqrt{\sum_i \left(\sqrt{P(i)} - \sqrt{Q(i)}\right)^2} \tag{21}$$

where $i = 3$ stages. The sum of Hellinger distances took the sum of $D_H(P,Q)$ over the 5 observation times. The squared Hellinger distance is often used in statistics, so we also took the sum of $(D_H(P,Q))^2$ over the 5 observation times.

We also used the earth mover's distance (EMD) or Wasserstein metric, a measure of the distance between two probability distributions developed in computer science and mathematics. Often analogies to piles of dirt are used to describe the EMD. It is the minimum cost of turning one pile of dirt into another; the amount of dirt that needs to be moved, weighted by the distance it needs to be moved. We wished to find the cost of turning the relative stage distribution of a simulated data set into the relative stage distribution of the true data set at any observation time. We "moved" the mass of the simulated data set's relative stage distribution until it matched that of the true data set, keeping track of the mass. The distance between the egg and immature stages and between the immature and adult stages was 1 and the distance between the egg and adult stages was 2. The sum of the earth mover distance summed the EMD over the relative stage distributions at each of 5 observation points.

We also used a more traditional population dynamics metric: the cross-correlation. Using the untransformed data as a time series, we used the cross-correlation function with zero time lag to find the cross-correlation between counts of individuals in the true data set and in a simulated data set. We used this distance metric to find cross-correlations in the number of eggs, the number of immatures, the number of adults, and the total number of individuals. We take the distance as 1 - the cross-correlation value. One limitation of this metric is that it does not distinguish between the absolute numbers of individuals in each data set. For example, the cross-correlation function would find zero distance between a cohort of 5 eggs that remain 5 eggs for the entire observation time, and 200 eggs that remain at 200 eggs.

The final distance metric we evaluated was a heuristic use of Poisson log probability, treating the observed count as the Poisson expected value and the simulated count as the Poisson datum,

$$lnPr(x_i|\mathcal{D}) = \sum_{y\in\mathcal{D},k\in x_i} (-y + k/y) \tag{22}$$

where k is a count of individuals in one stage at one time in x_i, a simulated data set, and y

is a count of individuals in the same stage at the same time in $\mathcal{D}$, the true data set.