Published in final edited form as:

Environ Ecol Stat. 2016 December ; 23(4): 585–603. doi:10.1007/s10651-016-0355-x.

A Bayesian mixture model for missing data in marine mammal growth analysis

Mary E. Shotwell[1], **Wayne E. McFee**[2], and **Elizabeth H. Slate**[3]

[1]Department of Computer Information Systems, Middle Tennessee State University, Murfreesboro, TN 37132, USA

[2]Center for Coastal Environmental Health and Behavioral Research, National Ocean Service, Fort Johnson, Charleston, SC 29412, USA

[3]Department of Statistics, Florida State University, Tallahassee, FL 32306, USA

Abstract

Much of what is known about bottle nose dolphin (*Tursiops truncatus*) anatomy and physiology is based on necropsies from stranding events. Measurements of total body length, total body mass, and age are used to estimate growth. It is more feasible to retrieve and transport smaller animals for total body mass measurement than larger animals, introducing a systematic bias in sampling. Adverse weather events, volunteer availability, and other unforeseen circumstances also contribute to incomplete measurement. We have developed a Bayesian mixture model to describe growth in detected stranded animals using data from both those that are fully measured and those not fully measured. Our approach uses a shared random effect to link the missingness mechanism (i.e. full/partial measurement) to distinct growth curves in the fully and partially measured populations, thereby enabling drawing of strength for estimation. We use simulation to compare our model to complete case analysis and two common multiple imputation methods according to model mean square error. Results indicate that our mixture model provides better fit both when the two populations are present and when they are not. The feasibility and utility of our new method is demonstrated by application to South Carolina strandings data.

Keywords

Gibbs sampler; Growth; Necropsy sampling; Selection bias; *Tursiops truncatus*

✉Mary E. Shotwell, Mary.Shotwell@mtsu.edu.

NOAA disclaimer

This publication does not constitute an endorsement of any commercial product or intend to be an opinion beyond scientific or other results obtained by the National Oceanic and Atmospheric Administration (NOAA). No reference shall be made to NOAA, or this publication furnished by NOAA, to any advertising or sales promotion which would indicate or imply that NOAA recommends or endorses any proprietary product mentioned herein, or which has as its purpose an interest to cause the advertised product to be used or purchased because of this publication

1 Introduction

1.1 Motivation

The Marine Mammal Protection Act (MMPA) initiated in 1972 prohibits the take of marine mammals in U.S. waters with few exceptions, limiting the sources of data on marine mammals. Stranding events provide opportunities to study aspects of marine mammals that may not otherwise be permitted. For the state of South Carolina, staff and volunteers from the National Oceanic and Atmospheric Administration (NOAA), South Carolina Department of Natural Resources, Coastal Carolina University, and volunteer designees comprise the Marine Mammal Stranding Network (MMSN). Through this program, data have been collected on marine mammal anatomy, physiology, diet, health status, and causes of mortality for hundreds of stranded animals. Unfortunately, few stranded animals, including bottlenose dolphins (*Tursiops truncatus*), are detected before major decomposition occurs, limiting the amount and quality of information available from each animal. Thus it is important to obtain and utilize as much information as possible on animals in good to fair condition.

In response to a 1994 amendment to the MMPA, the Marine Mammal Health and Stranding Response Program (MMHSRP) was formed. One goal of the MMHSRP is to model bottlenose dolphin growth for the southeastern U.S. coastal population. Growth curve comparisons between regions can help to distinguish biological or ecological differences in populations. Although several growth models exist, the Gompertz model (in various forms) has wide use in marine mammal studies, including those for the Gulf of Mexico (Mattson et al. 2006; Turner et al. 2006), eastern Florida (Stolen et al. 2002), and South Carolina (McFee et al. 2010) populations. Given observed age and total body mass on each of N animals, $\{(\text{age}_i, w_i), i = 1, 2, \ldots, N\}$, the Gompertz model as used by Turner et al. (2006) is

$$w_i = a \cdot \exp(-b \cdot \exp(-k \cdot \text{age}_i)) + \varepsilon_i$$

$$\varepsilon_i \sim N(0, \sigma^2), \ \text{independently}$$

$$a > 0, b > 0, k > 0, \ \text{age}_i > 0,$$

where a is the upper asymptotic total body mass, b quantifies a shift of the model on the age axis, and k is the growth rate.

To use this model for body mass growth, both age and total body mass are needed for each animal. However, not all stranded animals the South Carolina MMSN detects can be fully measured. When possible, animals are transported to the laboratory for further analysis, including total body mass. Due to the limited number of volunteers, equipment availability, or other factors, it is more feasible to transport a smaller animal than a larger adult animal for laboratory measurements. This type of missingness presents a selection bias in which smaller animals have a higher tendency to be completely measured than larger animals, causing an underestimate of animal mass when a growth curve is fit to only measured animals. In addition, pending weather conditions, accessibility to transportation (boat, truck, etc.), and even time of day (tidal influences, hours of daylight left) can pose constraints on the thoroughness of strandings assessments. All animals regardless of size are equally susceptible to this random missingness. Growth curve parameters are not expected to be biased due to this type of missingness; however, the decrease in sample size would degrade precision.

These two missing data mechanisms are difficult or impossible to correct for at the sampling stage. The current practice is to omit animals with missing data from the analysis (‘complete case analysis’). With the already small sample sizes of strandings datasets, it is imperative that all available data be used. The goals of this study are: (1) Develop a statistical model for growth in bottlenose dolphins off SC utilizing two groups–animals fully measured and not fully measured, to compensate for the omission of larger animals in the complete case analysis; (2) Evaluate model performance through a simulation study, applying the model to generated data based on real data characteristics; (3) Compare the developed model results to those from complete case and multiple imputation methods performed by Shotwell et al. (2010); (4) Apply the model to real data.

1.2 Missing data considerations

Handling missing data by complete case analysis can be costly not only in terms of decreased statistical power and imprecise variance estimates, but also the parameter estimates will be biased unless the complete cases represent a random sample of the focal population (Rubin 1976). Since random chance is rarely the reason for missingness, several missing data methods have been established to deal with other known sources for missingness.

Established missing data methods often rely on assumptions about the missing data mechanism. Rubin (1976) established mechanisms based on whether missingness depends on the observed data, unobserved data, or neither. Missing completely at random (MCAR) refers to the case when missingness does not depend on either the observed data or the unobserved data. If MCAR holds, the complete case analysis would be valid in the sense that parameter estimates are unbiased. A less restrictive case is the missing at random (MAR) mechanism, in which the missingness depends on the observed data, while in the least restrictive missing not at random (MNAR) mechanism, missingness depends on the unobserved data. In strandings data, the lack of total body mass measurement depends on total body mass itself; however, total body mass can be predicted by the observed total body length, making MAR methods possible.

1.3 Available methods

Several established missing data methods could help to compensate for the selection bias present in marine mammal strandings data. Previously, regression and propensity score multiple imputation (MI) methods were examined to assess the effect of missing data on Gompertz growth parameters (Shotwell et al. 2010). Both methods are straightforward to implement, and application to strandings data suggested possible bias in growth parameter estimates obtained under complete case analysis. However, Carpenter et al. (2006) have shown that MI is not robust to misspecification of the imputation model, and resulting estimates can therefore be inconsistent.

Inverse probability weighting (IPW) methods (Horvitz and Thompson 1952) weight the observed data by the inverse of the propensity score (i.e. missingness probability). Carpenter et al. (2006) and Robins and Rotnitzky (1995) developed more complex missing data models using IPW as the foundation, which have led to further development of weighted estimating equations (Robins et al. 1994, 1995; Rotnitzky et al. 1998) and concepts of double robustness (Robins and Rotnitzky 2001; Laan and Robins 2003). Methods utilizing IPW have been shown to have high variance inherently (Kang and Schafer 2007), and are known to be less efficient than methods based on likelihood (Clayton et al. 1998; Carpenter et al. 2006).

Likelihood methods are commonly used and can be readily executed if the likelihood of interest is fairly simple. However, as the model becomes progressively more complex and when data are missing, likelihood methods can become computationally intensive and cumbersome. The expectation–maximization (EM) algorithm (Dempster et al. 1977) was developed for these cases, although it is not always possible to calculate the expectation of the E step in closed form. In some cases the maximization of the M step requires a series of iterative steps, which can result in achieving several maxima (Little and Schluchter 1985). Dong and Peng (2013) provide a recent review of these missing data methods.

Bayesian methods are especially valuable if prior information is known about distributions or clinical relevance of parameters. Prior distributions can range from highly informative, so that the investigator can emphasize relevant field knowledge, to uninformative or relatively “flat”, in which case the resulting estimates rely heavily on the data. When the posterior distribution cannot be derived in closed form, Markov chain Monte Carlo (MCMC) methods can be used to obtain realizations from the posterior distributions. Gibbs sampling (Geman and Geman 1984) is an MCMC method that uses draws from the successive full conditional distributions that, upon convergence, characterize the posterior distribution. Missing data does not create difficulty for this method, since a value for the missing element can be drawn from the posterior distributions at each step. These advantages make the Bayesian method attractive for the purposes of this study.

2 Methods

2.1 Data

All detected stranded bottlenose dolphins (*T. truncatus*) considered freshly stranded or moderately decomposed along the South Carolina coastline from 1993 to 2007 were included in this study. A total of 91 animals (55 female, 20 male, 16 unknown sex) were necropsied at the stranding site or the NOAA lab facility in Charleston, SC. Morphometric measurements (Hoffman 1991), including total body length (cm), girths (cm), blubber depths (cm), and total body mass (kg) were taken during each necropsy when possible. A total of 18 animals did not have total body length measured, leaving a sample of 73 animals for these analyses. Organ mass measurements (g), tissue samples, and bodily fluid samples were also part of each necropsy. Stranding site information, such as location, condition of the animal, signs of human interaction, and date were recorded. Age in years was determined by mid-longitudinal, stained and mounted thin tooth sections according to established protocols (Hohn et al. 1989).

2.2 Model

A goal of this study is to model total body mass (kg) as a function of age (years) for all stranded, detected bottlenose dolphins off South Carolina. It is suspected that animals not fully measured tend to have larger body sizes, and hence may have different Gompertz growth parameters than those measured.

For animal i, $i = 1, 2, \ldots, N$, let w_i denote the total body mass, age_i be the age, length_i be the animal length, and let m_i be an indicator of whether the animal is fully measured, i.e. $m_i = 1(0)$ according to whether w_i is observed(unobserved). Our model is as follows:

$$w_i|\mathbf{Q}, z_i{=}j{\sim}N(\mu_j(\text{age}_i), \sigma^2) \quad j{=}0, 1$$

$$\mu_j(\text{age}_i){=}a_j\ \exp(-b_j\ \exp(-k_j\text{age}_i))$$

$$z_i|\mathbf{Q}{\sim}\text{Bern}(p_i) \quad (1)$$

$$m_i|\mathbf{Q}{\sim}\text{Bern}(p_i)$$

$$\text{logit}(p_i){=}\beta_0{+}\beta_1\text{length}_i,$$

where $\mathbf{Q} = \{\beta_0, \beta_1, a_0, b_0, k_0, a_1, b_1, k_1, \sigma\}$ and all conditional distributions hold independently $i = 1, 2, \ldots, N$. Here, conditional on the parameters $\mathbf{Q}$, the total body mass, w_i, has a distribution that is the mixture of two normal components, each having a Gompertz mean. The latent indicator z_i has the same distribution as the observed measurement status, m_i. It is the latent z_i, however, and not m_i, that determines the component of the normal mixture governing w_i, thereby enabling information from all animals, fully measured and not, to contribute to estimation of both components of the normal mixture. Thus z_i, or equivalently p_i, is a shared random effect linking the growth model for w_i to the missingness model.

The probability that total body mass for animal i is observed, p_i, is modeled as a linear function of length on the logit scale. This model choice follows from the work of Shotwell et al. (2010), who investigated available predictors for the probability of missing total body mass and concluded that only total body length was significant.

The prior distributions for the parameters in **Q** were set as follows. The logistic regression parameters (β_0, β_1) have independent normal priors with mean zero and large variance (100). The prior on a_1 is a non-informative log-normal distribution with large variance (log-Normal(0,100)), since $a_1 > 0$. Because unmeasured animals are expected to have greater mean asymptotic total body mass than measured animals, we require $a_0 \gg a_1$ and parameterize as $a_0 = a_1 + \delta$ for $\delta \geq 0$. We take $\delta \sim U[0, 50]$, as values over 50kg would imply unrealistically large differences in total body mass between the two groups. Additionally, given our knowledge that measured animals have asymptotic total body mass approximately 175kg (males and females combined, Turner et al. 2006), δ larger than 50kg corresponds to unrealistically large masses for unmeasured animals. For marine mammals, the Gompertz shift parameter b tends to range from 0 to 2, while the slope k ranges between 0 and 1 (see Turner et al. 2006; Stolen et al. 2002). To be relatively noninformative, we expanded these ranges and set b_0, $b_1 \sim U(0, 5)$ and k_0, $k_1 \sim U(0, 2)$. Following the suggestion of Gelman (2006), the prior distribution for σ was taken as $U(0, 5)$.

2.3 Estimation

Model fitting and estimation were performed using R (©2006) v. 2.4.0 software. A Gibbs sampler was designed for sampling from the joint posterior distribution. The full conditional posterior distributions are given in the Appendix. Since many of these are nonstandard distributions, the Metropolis–Hastings (M–H) algorithm was used to draw samples. The proposal distributions in the M–H algorithms were normal with a mean equal to the previous draw's value for that parameter (or initial value at the start). The variance of the proposal distribution was optimized to ensure that the acceptance rate was no less than 10% (see Tanner 1996 for review).

The number of burn-in iterations for analysis of the strandings data was determined by evidence of convergence. Two chains were run simultaneously, and the variance between and within chains was used to calculate the Gelman–Rubin statistic R (Gelman and Rubin 1992) for all model parameters. A mean Gelman–Rubin statistic of 1.10 or less (mean of the means across the number of replications) was a criterion for convergence. Trace plots were also examined to ensure mixing of the two chains over the iterations.

The number of burn-in iterations for the simulation study was determined using data generated under scenario $a_0 \gg a_1$ (described in Sect. 2.4.2) with $N = 100$ animals and 50 run, and corresponding Rs were assessed. If convergence critera were not met, an additional 10,000 iterations were run, with the process continuing until convergence was met. This iterative process continued through 5 generated datasets, using the previous number of iterations needed for convergence as the burn-in for the next dataset. Correlation plots for each parameter were examined, and thinning was performed using every 5, 10, 20, 30, 40, and 50 iterations. Convergence occurred within 210,000 burn-in iterations. Posterior inference was based on a sample of 7000 obtained by taking a subsequent 210,000 iterations and thinning by 30.

Initial simulation runs with the small sample size of $N = 50$ revealed that convergence criteria were not uniformly met with 210,000 burn-in iterations. For these simulations, we used 420,000 burn-in iterations.

2.4 Simulation scheme

All simulations were performed using R (©2006) v. 2.4.0 software. Simulations were designed to achieve several goals, studying the effects on parameter estimation in model (1) of changes in sample size, the percentage of missing data, and the difference in the underlying values of the Gompertz parameters in $\mu_0(\cdot)$ and $\mu_1(\cdot)$.

2.4.1 Data generation—Populations of dolphins of size N were generated according to the sequence animal age, then total body length, then measurement status and, lastly, total body mass. Observed data from wild SC bottlenose dolphins (described in Sect. 2.1) were used to ensure the simulated data were realistic. Simulated animal ages were randomly drawn from the fit of the exponential distribution to the ages reported by NOAA investigators for the observed SC wild dophins; this fit had mean 10.1 years. Total body lengths were then drawn according to a Gompertz function of age consistent with prior analyses of Turner et al. (2006) and Shotwell et al. (2010). More specifically, the total body length for animal i was drawn from a normal distribution with standard deviation 4 cm and mean $250\cdot\exp(-0.75\cdot\exp(-0.35\cdot\text{age}_i))$. Each animal i was then assigned a probability p_i of being fully measured based on the logistic regression in (1) and the desired missingness percentage (i.e. fraction not fully measured), as described below. The indicator for fully measured, m_i, and the indicator determining the Gompertz mean for total body mass, z_i, were assigned (independently) according to a Bernoulli(p_i) distribution. Finally, given age_i and z_i, total body mass was generated from the corresponding normal distribution with Gompertz mean and standard deviation $\sigma = 4$ kg.

A desired missingness percentage q was achieved by determining appropriate values for β_0 and β_1 in (1) as follows: First, the maximum probability of measurement was fixed at $p_{\max} = 0.95$, corresponding to the length of $L_{\min} = 95$ cm, approximately the smallest length observed among SC dolphins. Second, a minimum probability of measurement, $p_{\min}$, was proposed corresponding to a length of $L_{\max} = 275$ cm, approximately the largest length observed among SC dolphins. Using these values in model (1) determined candidate values for β_0 and β_1. Five data sets of size $N = 100$ animals were then generated using these candidate values, and the average percentage of missingness, $\hat{q}_{\text{av}}$, was calculated (upon generation of the measurement indicators m_i). The proposed value of $p_{\min}$ was adjusted and the remainder of the procedure repeated until $|\hat{q}_{\text{av}} - q| \leq 0.5\%$, at which point the current values of β_0 and β_1 were retained for data generation.

2.4.2 Sample size, missingness, growth differences—The sample size, N, percentage of missing data, q, and difference in μ_0 and μ_1 were studied using a $3 \times 3 \times 3$ factorial scheme. Sample sizes of $N = 50$, 100, and 500 were used, reflecting plausible strandings dataset sizes. Percentages of missing data examined were $q = 20$, 30, and 40 since such large percentages are present in real data. And three growth scenarios were examined: (1) no difference in growth between measured and not measured groups, i.e. $(a_0, b_0, k_0) = (a_1, b_1, k_1)$, denoted as "$a_0 = a_1$"; (2) growth in the not measured group reaches a larger mean asymptotic total body mass at a quicker rate, with the initial fast period of growth sooner than those that are measured, denoted as "$a_0 \gg a_1$"; (3) same as scenario 2, but the magnitudes of difference are decreased (“$a_0 > a_1$”). Scenario 3 was used to investigate the model fit when the two different groups are less well distinguished.

The Gompertz parameters used to generate total body masses for the animals not measured varied, while the Gompertz parameters for the measured animals remained constant over the simulations ($a_1 = 160$, $b_1 = 2.0$, $k_1 = 0.175$). Turner et al. (2006) found the standard error of a to be approximately between 11 and 20, however, differences larger than 10 kg may be unreasonable. Also, using corresponding standard errors for b and k (Turner et al. 2006), the largest differences scenario ($a_0 \gg a_1$) had $a_0 = 170$, $b_0 = 1.7$, and $k_0 = 0.25$. The midpoints of the differences ($a_0 = 165$, $b_0 = 1.85$, $k_0 = 0.21$) were used for the scenario with decreased differences ($a_0 > a_1$). The standard deviation used in generating total body lengths and masses was specified as 4, resulting in mass and length distributions similar to those seen in the real data.

2.5 Determination of simulation size

The number of replications, i.e. the number of generated data sets under one scenario, was determined so that credible intervals for the mean differences $a_0 - a_1$, $b_0 - b_1$, and $k_0 - k_1$ had specified width.

Using $a_0 - a_1$ for illustration, it was assumed that the estimated mean difference d_a between a_0 and a_1 had mean δ_a and variance σ_a^2 The distribution of $\bar{d}_a$, the mean of the estimated means over m generated data sets, is then approximately normally distributed with mean δ_a and variance σ_a^2/m. Under these assumptions, m was selected so that, for specified margin of error E and coverage probability $1 - \alpha$, $\Pr(-E < \bar{d}_a - \delta_a < E) = 1 - \alpha$. Thus, $m = (z_{\alpha/2}\,\sigma_a/E)^2$, where $z_{\alpha/2}$ is the $\alpha/2$ critical value from the standard normal distribution.

Approximate values for the variances of the difference in Gompertz parameters, σ_a^2, σ_b^2 and σ_k^2 were determined as the mean of the posterior means from fitting the model to 100 data sets, each of size 100 animals generated with 50 missing total body mass under the scenario $a_0 \gg a_1$. The burn-in and sample iterations with thinning were as described in Sect. 2.3. Using $\alpha = 0.05$ and $E = 2$, 5 and 0.01 for the differences in a, b and k, respectively, the total number of replications m was determined to be 35, 23, and 15, respectively. The more extreme situation of 50% missing total body mass was used to be conservative, driving larger estimated variances and hence larger m, helping to ensure that all simulation scenarios would produce $1 - \alpha$ credible intervals within the desired margin of error. Based on these results, 50 replications for each of the 27 data generation scenarios was deemed sufficient.

2.6 Comparison analyses

In addition to our new mixture model, both regression and propensity score multiple imputation methods were implemented for each generated dataset, using the approach of Shotwell et al. (2010). Total body length was used to predict total body mass under regression multiple imputation, and to predict the probability of being measured under propensity score multiple imputation. Five completed datasets were imputed, and results combined for estimation of one Gompertz curve.

2.7 Simulation evaluation

Upon convergence, the posterior means for all parameters were computed for each simulation run. The sample mean and variance of these posterior means were computed across the simulations and recorded. These estimated parameters were compared to the true parameter settings to compute parameter-specific mean square errors.

Because the data were simulated, every total body mass, whether simulated as fully measured or not, could be compared to a mean total body mass at that same age from the fitted Gompertz model. The model mean square error was computed between the true simulated total body mass w_i and the predicted total body mass $\hat{w}_i$ as

$${}^{1}\!/_{N} \sum_{i=1}^{N} (\hat{w}_i - w_i)^2.$$

The model mean square errors for the complete case and multiple imputation analyses were calculated using their respective estimated *a, b*, and *k* in the Gompertz model. In the corresponding MSE calculation for our new Bayesian model, $\hat{w}_i$ was computed as the mixture $(1 - \tilde{z}_i)\,\tilde{\mu}_0(\text{age}_i) + \tilde{z}_i\,\tilde{\mu}_1(\text{age}_i)$ where the tilde indicates the posterior mean.

3 Results

3.1 Simulation

3.1.1 Bias—Tables 1, 2 and 3 display the relative bias and mean square errors for each parameter. Bias was largest in the small sample size scenarios ($N = 50$). All three Gompertz parameter estimates for the measured animal group (a_1, b_1, k_1) had small relative biases with the exception of k_1 for one scenario ($a_0 > a_1$, 40% missing). Relative biases for a_0, b_0, and k_0 were moderate, with an exception for k_0 when the growth groups were equal with 20% missing data. Increasing sample size to 100 animals greatly reduced relative bias to 0–5%, and to nearly zero for $N = 500$. The exceptions were β_0 and β_1, with relative biases ranging from approximately −5 to 50% for $N = 50$ and reducing to about −10 to 10% for $N = 500$.

Few trends in relative bias across percentages of missing data were evident. The relative bias in β_0, β_1, and b_0 tended to decrease as percentage of missing data increased. This was apparent in sample sizes of 50 and 100 animals, but not in sample sizes of 500 animals.

There were several changes in relative bias according to the magnitude of difference between μ_0 and μ_1. Relative biases for a_0, β_0, β_1, and b_0 were lowest under the scenario $a_0 \gg a_1$ and increased as the two growth groups approached equality. The relative bias for k_0 was small except when the growth groups were equal with sample size $N = 50$. All three parameters within μ_1 had less relative bias when μ_0 differed from μ_1, as did σ. When the growth groups were equal, σ was consistently underestimated by 5–10%.

3.1.2 Parameter-specific MSE—Since mean square error is a relative measure, it cannot be directly compared across parameters. However, in the case of this model, it may be warranted to compare MSEs of the growth parameters (a_0, b_0, k_0) to those of (a_1, b_1, k_1), since they share a similar interpretation. For parameters β_1 and k_1, the estimated mean square errors were all less than 0.01, constituting very small changes over the simulation scenarios. Trends in estimated MSE described here are representative of the remaining parameters.

Sample sizes of $N = 50$ produced the largest mean square errors, and as sample size increased, estimated mean square errors decreased. At the smallest sample size, the range of MSEs for a_0 was wider than that found for a_1, ranging from approximately 40–450 and 12–265, respectively. The same holds true for the range of estimates comparing b_0 with b_1 and k_0 with k_1. For all parameters, estimated MSEs decreased to much smaller values at $N = 100$, with the least amount of change occurring for σ. At $N = 500$, all MSEs were estimated to be small or close to zero.

As was the case with relative bias, MSE did not display obvious trends with percentage of missing data. Estimated MSEs for β_0 and b_0 tend to decrease as percentage of missing data increases, while those for a_1 and b_1 increase.

Differences in estimated MSEs between the growth groups decreased as sample size increased. For β_0, σ, a_0, b_0, and k_0, the scenario with $a_0 \gg a_1$ resulted in the lowest MSEs. Overall, the scenarios in which $a_0 \neq a_1$ tended to result in estimated MSEs less than those found when there were no differences between the groups, with the exception of a_1 and b_1 when $N = 50$. The MSEs for k_0 when $N = 50$ again showed that equal growth groups produced results unlike the other scenarios.

3.1.3 Model MSE—Figure 1 contains plots of the estimated mean square error across percentage of missing data within each Gompertz mean scenario and sample size. As the percentage of missing data increased, the estimated MSE increased for the new, complete case, and regression multiple imputation models. The propensity score multiple imputation model, however, had the largest MSE when the percentage of missing data was 30%.

Complete case analysis performed best for the scenarios with $N = 50$ and $a_0 = a_1$, and performed similarly to the new model when $a_0 > a_1$. For all other scenarios, the new model outperformed the complete case, regression MI, and propensity MI in terms of model MSE. Complete case performed second best in these other scenarios, and propensity MI performed consistently worst in every scenario.

The sample variances of the MSEs for each model were largest for the smallest sample size, and decreased as the sample size increased. The new model MSE sample variance tended to increase as percentage of missing data increased, whereas no apparent trend was seen in the complete case results. The variances of the regression and propensity multiple imputation methods were generally larger than those found for either the new or complete case methods, with the propensity MI methods consistently having the largest variances.

3.2 Application

The results from fitting model (1) to the SC MMSN data are given in Table 4. The complete cases totaled 37 animals, while the new method additionally utilized age and lengths for those animals missing total body mass, totaling 73 animals.

The probability of being measured, based on the logistic regression parameter estimates and the minimum/maximum total body length, ranged from 0.20 to 0.83. The mean asymptotic total body mass between the two growth groups differed by an estimated 49 kilograms. Animals in the measured group had a slightly higher estimated growth rate ($k_1 = 0.287$) than the unmeasured group ($k_0 = 0.256$), with the fast growth period occurring later ($b_1 = 2.317$; $b_0 = 1.727$).

Complete case analysis resulted in estimates of $a = 172.5$ (se = 10.4), $b = 2.10$ (se = 0.2), and $k = 0.23$ (se = 0.04). Propensity score multiple imputations results ($a = 171.17$ (10.1), $b = 2.00$ (0.2), $k = 0.26$ (0.05)) most closely resembled complete case results out of the methods investigated. Regression multiple imputation resulted in a lower mean asymptotic total body mass ($a = 163.74$ (7.7)) than complete case analysis, with lower shift ($b = 1.91$ (0.2)) and faster growth ($k = 0.27$ (0.04)).

The new model is a mixture model, with different growth structures for the fully and partially measured groups in the population. However, in keeping with the prevalent use and interpretability of the Gompertz model, investigators may wish to have one unifying Gompertz curve that represents the fitted new model. Toward this end, estimates of β_0 and β_1 can be used with total body length of animal i to estimate the probability animal i is measured ($\hat{p}_i$), via

$$\text{logit}(\hat{p}_i)=\hat{\beta}_0+\hat{\beta}_1\text{length}_i.$$

The weighted average total body mass, $\overline{w}_i$, is then computed from means $\hat{\mu}_0$ and $\hat{\mu}_1$ weighted by $\hat{p}_i$ as follows:

$$\overline{w}_i=(1-\hat{p}_i)\hat{\mu}_0(\text{age}_i)+\hat{p}_i\hat{\mu}_1(\text{age}_i).$$

The left panel in Fig. 2 displays the observed data and the two components $\hat{\mu}_0$ and $\hat{\mu}_1$ of the new model fit, together with the computed values of $\overline{w}_i$ denoted by solid circles. The Gompertz curve closest to the $\overline{w}_i$ (by least squares) is shown in the right panel ($a = 171.01$, $b = 2.20$, $k = 0.28$). Standard errors for the estimates of a, b, and k, respectively, were 0.47, 0.02, and less than 0.01, although these standard errors ignore the uncertainty of the estimates of the parameters in **Q**. The mean asymptotic total body mass for the unifying summary curve is similar to that found by Shotwell et al. (2010) for complete case ($a = 172.5$, $b = 2.10$, $k = 0.23$) and propensity score analysis. However, both b and k are estimated to be larger than those estimated in the other analyses.

4 Discussion

Because data for modeling the growth of bottlenose dolphins is severely limited, it is important to use all available information from strandings. By postulating that animals not fully measured have a separate growth curve constrained to have greater asymptotic total mass and allowing information to be shared between the fully measured and not fully measured animals, our model provides a way to accommodate information from both of these groups, unlike complete case analysis. The application uses all available information and hence produces a growth curve that better reflects observed *and* unobserved total body masses.

Overall, parameter estimation was not as sensitive to percentage of missing data as to sample size and difference in growth parameters. The percentage of missing data, however, may affect the strength of the logistic regression and hence the ability to estimate β_0 and β_1. Generally both relative bias and MSE decreased as sample size increased and as the growth groups increasingly differed. It is reasonable to expect such a trend, in that the more distinct the means are, the better the estimation.

The new model performed best among the examined models, with lower mean model MSE than the two multiple imputation methods in all simulation scenarios, and lower model MSE than complete case in most scenarios. However, when the two groups had the same mean growth pattern, we observed underestimation of the error variance suggesting an anticipated overfitting. Another weakness of the new model is less precise estimation of (the greater number of) parameters when the sample size is small. Nonetheless, when the growths differ in small datasets, or no differences occur in large datasets, the new model is preferred over complete case and multiple imputation methods.

Analysis of the South Carolina MMSN data revealed a large difference in asymptotic total body mass for the measured and unmeasured components of the mixture. The complete case asymptotic total body mass is closer to a_0 than a_1, consistent with the underlying assumption (of complete case analysis) that the measured animals are representative of the unmeasured animals. The low estimate for a_1 may emphasize how rarely large animals are indeed measured and how, when measured, these animals may have a large influence on the asymptote in complete case analysis. The estimated growth rate is slightly faster for measured than unmeasured animals, while the initial growth spurt occurs later. The unmeasured group achieves larger sizes with the fastest growth period occurring sooner, but the overall growth rate is slightly lower than the measured group.

The complete case analysis from Shotwell et al. (2010) may underestimate the initial growth in total body mass, while overestimating asymptotic total body mass. In comparison to the new model, the complete case model may weight the largest overall total body mass too heavily, while not weighting the larger total body masses for the younger ages enough. The regression multiple imputation curve follows closely the unweighted mean of the two new method growth curves (i.e., the midpoint curve), suggesting its usefulness when missingness is not due to animal size. The propensity multiple imputation curve most closely follows the total body mass curve of the unmeasured animals, with poorer fit in the juvenile to adult ages. The uncertainty at these middle ages may be due to the logistic regression fit, in which the probability an animal is measured is close to 0.50 in this age range.

It is possible the animal with a total body mass greater than 240 kg is part of the offshore population. It was not excluded in the analysis, however, since other animals in the dataset may also be offshore animals, as definitive distinction is difficult. Offshore animals tend to be larger in body size than the resident population (Mead and Potter 1995), and hence would be fully measured less frequently. A benefit of our mixture model is that it could be adapted in future work to separate the offshore and resident populations with two different underlying growth scenarios. Some applications may call for an additional model extensions to heterogeneous variances for the mixture components or variance as a function of age.

Investigators are advised to focus interpretation on the Gompertz growth parameters in the two components of the mixture. However, the weighted average curve discussed in Sect. 4 and shown in Fig. 2 for the South Carolina MMSN data provides a unified curve that does adjust for bias due to missingness.

Acknowledgments

We thank members of the NOS Coastal Stranding Assessments Program operating under Section 109(h) of the Marine Mammal Protection Act. Many thanks to stranding volunteers and staff members, the SCDNR, and Coastal Carolina University for stranding participation.

This project was funded in part by the National Institute of General Medical Sciences (NIH/NIGMS 1T32GM074934) training grant ‘Biostatistics Training for Basic Biomedical Research’ and NSF grant DMS-0604666 (Slate).

Appendix

Full conditional posterior distributions for all model parameters.

$$P(\beta_0|Q_{-\beta_0},\mathbf{w},\mathbf{m},\mathbf{z}) \propto \prod_{i=1}^{n} p_i^{z_i}(1-p_i)^{1-z_i}\prod_{i=1}^{n} p_i^{m_i}(1-p_i)^{1-m_i}\frac{1}{10\sqrt{2\pi}}\exp\left(-\frac{\beta_0^2}{200}\right)$$

$$P(\beta_1|Q_{-\beta_1},\mathbf{w},\mathbf{m},\mathbf{z}) \propto \prod_{i=1}^{n} p_i^{z_i}(1-p_i)^{1-z_i}\prod_{i=1}^{n} p_i^{m_i}(1-p_i)^{1-m_i}\frac{1}{10\sqrt{2\pi}}\exp\left(-\frac{\beta_1^2}{200}\right)$$

$$P(\delta|Q_{-a_0},\mathbf{w},\mathbf{m},\mathbf{z}) \propto \frac{1}{50}\prod_{i=1}^{n}\phi(w_i|\mu_0(\text{age}_i),\sigma^2)^{1-z_i}$$

$$P(b_0|Q_{-b_0},\mathbf{w},\mathbf{m},\mathbf{z}) \propto \frac{1}{5}\prod_{i=1}^{n}\phi(w_i|\mu_0(\text{age}_i),\sigma^2)^{1-z_i}$$

$$P(k_0|Q_{-k_0}, \mathbf{w}, \mathbf{m}, \mathbf{z}) \propto \frac{1}{2}\prod_{i=1}^{n}\phi(w_i|\mu_0(\text{age}_i), \sigma^2)^{1-z_i}$$

$$P(a_1|Q_{-a_1}, \mathbf{w}, \mathbf{m}, \mathbf{z}) \propto \prod_{i=1}^{n}\phi(w_i|\mu_0(\text{age}_i), \sigma^2)^{1-z_i}\phi(w_i|\mu_1(\text{age}_i), \sigma^2)^{z_i} \times \frac{1}{10a_1\sqrt{2\pi}}\text{exp}\left(-\frac{(\ln a_1)^2}{200}\right)$$

$$P(b_1|Q_{-b_1}, \mathbf{w}, \mathbf{m}, \mathbf{z}) \propto \frac{1}{5}\prod_{i=1}^{n}\phi(w_i|\mu_1(\text{age}_i), \sigma^2)^{z_i}$$

$$P(k_1|Q_{-k_1}, \mathbf{w}, \mathbf{m}, \mathbf{z}) \propto \frac{1}{2}\prod_{i=1}^{n}\phi(w_i|\mu_1(\text{age}_i), \sigma^2)^{z_i}$$

$$P(\sigma|Q_{-\sigma}, \mathbf{w}, \mathbf{m}, \mathbf{z}) \propto \frac{1}{5}\prod_{i=1}^{n}\phi(w_i|\mu_0(\text{age}_i), \sigma^2)^{1-z_i}\phi(w_i|\mu_1(\text{age}_i), \sigma^2)^{z_i}$$

$$P(z_i{=}1|Q, w_i, m_i){=}\frac{\phi(w_i|\mu_1(\text{age}_i), \sigma^2)p_i}{\phi(w_i|\mu_0(\text{age}_i), \sigma^2)(1-p_i){+}\phi(w_i|\mu_1(\text{age}_i), \sigma^2)p_i}$$

$$P(w_i|Q, m_i, z_i){=}\phi(w_i|\mu_0(\text{age}_i), \sigma^2)^{1-z_i}\phi(w_i|\mu_1(\text{age}_i), \sigma^2)^{z_i}$$

References

Carpenter JR, Kenward MG, Vansteelandt S. A comparison of multiple imputation and doubly robust estimation for analyses with missing data. J R Stat Soc. 2006; 169(3):3798–3813.

Clayton D, Spiegelhalter D, Dunn G, Pickles A. Analysis of longitudinal binary data from multi-phase sampling (with discussion). J R Stat Soc Series B Stat Methodol. 1998; 60(1):71–87.

Dempster AP, Laird NM, Rubin DB. Maximum likelihood from incomplete data via the EM algorithm. J R Stat Soc Series B Stat Methodol. 1977; 39(1):1–38.

Dong Y, Peng CJ. Principled missing data methods for researchers. Springer Plus. 2013; 2(1):222. [PubMed: 23853744]

Gelman A. Prior distributions for variance parameters in hierarchical models. Bayesian Anal. 2006; 1:1–19.

Gelman A, Rubin DB. Inference from iterative simulation using multiple sequences. Stat Sci. 1992; 7(4):457–472.

Geman S, Geman D. Stochastic relaxation, Gibbs distributions, and the Bayesian restoration of images. IEEE Trans Pattern Anal Mach Intell. 1984; 6(6):721–741. [PubMed: 22499653]

Hoffman, RJ. History, goals and achievements of the regional marine mammal stranding network. In: Reynolds, JE., Odell, DK., editors. nOAA Technical Report NMFS. Vol. 98. Marine Mammal Strandings in the United States; 1991. p. 7-17.http://www.nmfs.noaa.gov/publications.htm

Hohn A, Scott M, Wells R, Sweeney J, Irvine A. Growth layers in teeth from free-ranging, known-age bottlenose dolphins. Mar Mamm Sci. 1989; 5:315–342.

Horvitz DG, Thompson DJ. A generalization of sampling without replacement from a finite universe. J Am Stat Assoc. 1952; 47:663–685.

Kang JDY, Schafer JL. Demystifying double robustness: a comparison of alternative strategies for estimating a population mean from incomplete data. Stat Sci. 2007; 22(4):523–539.

Little RJA, Schluchter MD. Maximum likelihood estimation for mixed continuous and categorical data with missing values. Biometrika. 1985; 72(3):497–512.

Mattson MC, Mullin KD, Ingram GW, Hoggard W. Age structure and growth of the bottlenose dolphin (Tursiops truncatus) from strandings in the Mississippi Sound region of the north-central Gulf of Mexico from 1986 to 2003. Mar Mamm Sci. 2006; 22(3):654–666.

McFee W, Schwacke J, Stolen M, Mullin K, Schwacke L. Investigation of growth phases for bottlenose dolphins using a Bayesian modeling approach. Mar Mamm Sci. 2010; 26:67–85.

Mead JG, Potter CW. Recognizing two populations of the bottlenose dolphin (Tursiops truncatus) off the Atlantic coast of North America: Morphological and ecological considerations. Int Mar Biol Res Inst: IBI Reports. 1995; 5:31–44.

MMPA. Marine Mammal Commission. 4340 East-West Highway, Bethesda, MD 20910: 2007. The Marine Mammal Protection Act (MMPA) of 1972 as Amended 2007. (1972) http://www.nmfs.noaa.gov/pr/pdfs/laws/mmpa.pdf

Robins JM, Rotnitzky A. Semiparametric efficiency in multivariate regression models with missing data. J Am Stat Assoc. 1995; 90:122–129.

Robins, JM., Rotnitzky, A. Comment on "Inference for semiparametric models: some questions and an answer,". In: Bickel, PJ., Kwon, J., editors. Stat Sin. Vol. 11. 2001. p. 920-936.

Robins JM, Rotnitzky A, Zhao LP. Estimation of regression coefficients when some regressors are not always observed. J AmStat Assoc. 1994; 89:846–866.

Robins JM, Rotnitzky A, Zhao LP. Analysis of semiparametric regression models for repeated outcomes in the presence of missing data. J Am Stat Assoc. 1995; 90:106–121.

Rotnitzky A, Robins JM, Scharfstein DO. Semiparametric regression for repeated outcomes with ignorable nonresponse. J Am Stat Assoc. 1998; 93:1321–1339.

Rubin DB. Inference and missing data. Biometrika. 1976; 63:581–592.

Shotwell ME, McFee W, Slate EH. Estimating Gompertz growth curves from marine mammal strandings in the presence of missing data. Int J Ecol Econ Stat. 2010; 19:32–46. [PubMed: 24812444]

Stolen MK, Odell DK, Barros NB. Growth of bottlenose dolphins (Tursiops truncatus) from the Indian River Lagoon system, Florida, USA. Mar Mamm Sci. 2002; 18(2):348–357.

Turner JP, Clark LS, Haubold EM, Worthy GAJ, Cowan DF. Organ weights and growth profiles in bottlenose dolphins (Tursiops truncatus) from the Northwestern Gulf of Mexico. Aquat Mamm. 2006; 32(1):46–57.

Van der Laan, MJ., Robins, JM. Unified Methods for Censored Longitudinal Data and Causality. New York: Springer; 2003.

Biographies

Mary E. Shotwell is Lecturer in the Jones College of Business, Department of Computer Information Systems at Middle Tennessee State University. She received her M.S. in Applied Marine Science at the University of Cape Town and Ph.D. in Biostatistics at the Medical University of South Carolina. Her teaching focuses on applied statistics often for students with little to no statistics background. Her enthusiasm for adapting to changes in the field has been integral in helping to develop a Business Intelligence and Analytics emphasis in her department. Dr. Shotwell's research interests have broadened from environmental statistics to statistics in education. Her paper titled "Student performance and success factors in learning Business Statistics in online vs. on-ground classes using a web-based assessment platform" was published in the Journal of Statistics Education in 2015. Her past research on bottlenose dolphins (Tursiops truncatus) has won her best student oral presentations at the Southeast and Mid-Atlantic Marine Mammal Symposium in 2009, the South Carolina Chapter of the American Statistical Association in 2008, and the R. L. Anderson Student Paper Award at the Southern Regional Council on Statistics Summer Research Conference in 2008. Dr. Shotwell took on the role of Faculty Advisor to the Association of Information Technology Professionals in 2013, developing mock interviews, resume workshops, and other exercises for students to prepare for the workforce. She also actively participates in several groups promoting women in science, including DigiGirlz (sponsored by Nissan) and the Boys & Girls Club of Smyrna, TN.

Wayne E. McFee has over 23 years of experience as Research Wildlife Biologist and Program Lead for the Coastal Marine Mammal Assessments Program at NOAA's National Ocean Service, Center for Coastal Environmental Health and Biomolecular Research in Charleston, South Carolina. He serves on numerous Federal working groups related to dolphin/human interaction issues, marine debris, and life history of dolphins. His research includes issues related to marine mammal strandings, marine animal health, human interaction with cetaceans, life history of cetaceans, and natural resource damage assessments related to natural and man-made disasters.

Elizabeth H. Slate is Duncan McLean and Pearl Levine Fairweather Professor of Statistics at Florida State University. Her research interests in biostatistics include Bayesian modeling, longitudinal data analysis and recurrent events, with applications in oral health, disease biomarkers and other health and biology areas.

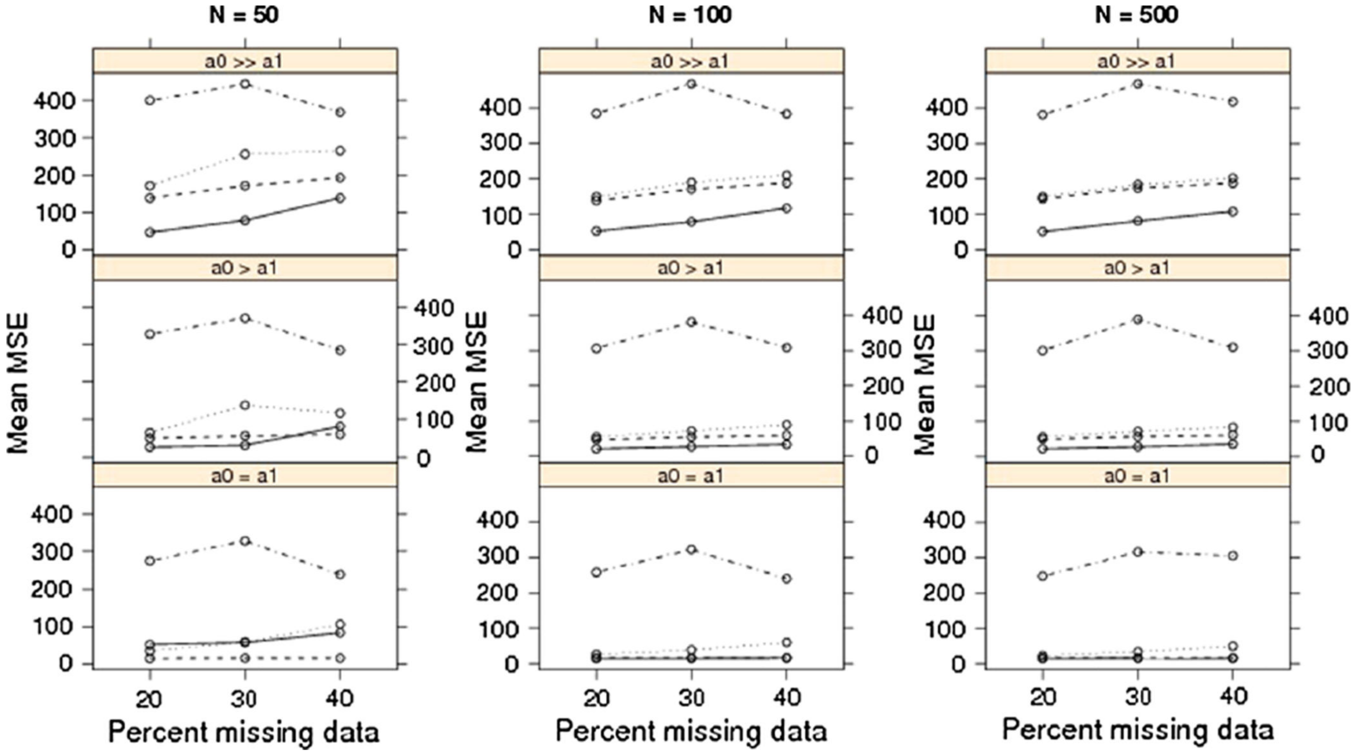

Fig. 1.

Comparison of missing data methods. The mean MSE across the replications for the new method is shown in the *solid line*, complete case method in the *dashed line*, regression MI method in the *dotted line*, and propensity MI method in the mixed *dash-dot line*. The median standard deviation of these MSE values is 69.3 for propensity MI and in the range 2–6 for the other methods

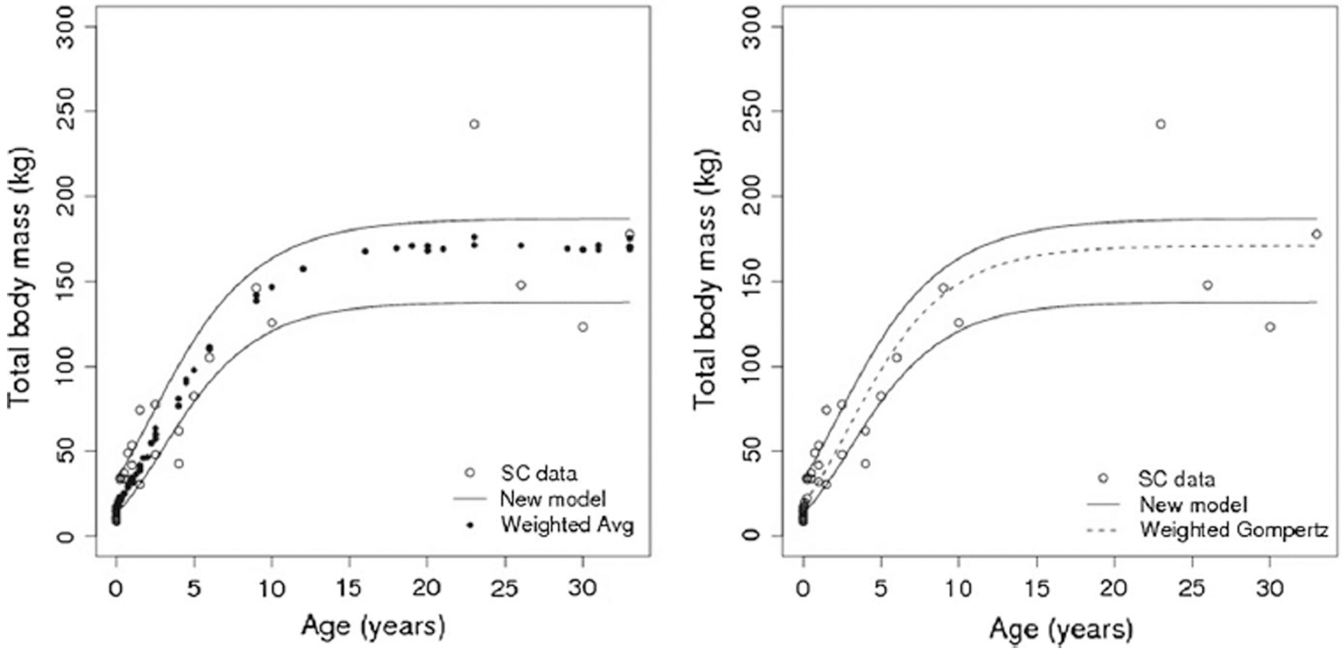

Fig. 2.

Weighted average and subsequent Gompertz curve from the new mixture model. The South Carolina data are represented by *hollow circles* and the calculated weighted averages are represented by *solid circles*. The two Gompertz curves resulting from the mixture model are shown in *solid lines*, while the Gompert curve fit to the weighted averages is shown in the *dashed line*

Table 1

Parameter summaries for sample size $N = 50$ across percentage of missing data and level of difference in Gompertz curves (0 = unobserved, 1 = observed)

% Missing data	Parameter	Growth scenario								
		$a_0 \gg a_1$			$a_0 > a_1$			$a_0 = a_1$		
		True	**RB**	**MSE**	**True**	**RB**	**MSE**	**True**	**RB**	**MSE**
20 %	β_0	4.148	0.6	4.5	4.148	27.9	7.8	4.148	43.7	10.1
	β_1	−0.013	−3.4	0.00	−0.013	35.9	0.00	−0.013	40.7	0.00
	a_0	170	1.0	52.7	165	3.7	78.8	160	12.1	448.5
	b_0	1.70	12.1	0.19	1.85	9.8	0.25	2.00	23.1	0.81
	k_0	0.25	4.6	0.001	0.21	0.5	0.002	0.175	232.8	0.271
	a_1	160	0.4	12.1	160	0.4	12.7	160	0.0	5.6
	b_1	2.00	0.1	0.003	2.00	−0.5	0.004	2.00	0.3	0.002
	k_1	0.175	−0.1	0.00	0.175	−0.6	0.00	0.175	0.4	0.00
	σ	4	5.3	0.14	4	4.9	0.19	4	−4.8	0.19
30 %	β_0	4.520	11.2	6.0	4.520	20.1	5.7	4.520	24.5	10.5
	β_1	−0.017	10.8	0.00	−0.017	19.9	0.00	−0.017	20.0	0.00
	a_0	170	1.9	42.7	165	4.1	93.7	160	9.5	282.0
	b_0	1.70	0.6	0.05	1.85	3.7	0.12	2.00	14.4	0.42
	k_0	0.25	−3.0	0.001	0.21	−4.3	0.002	0.175	84.4	0.087
	a_1	160	1.0	29.7	160	0.3	13.8	160	−0.8	5.9
	b_1	2.00	0.1	0.003	2.00	−0.6	0.004	2.00	−0.7	0.004
	k_1	0.175	−1.2	0.00	0.175	0.7	0.00	0.175	0.8	0.00
	σ	4	3.4	0.13	4	4.3	0.13	4	−7.9	0.28
40 %	β_0	4.872	−0.9	2.3	4.872	15.8	6.5	4.872	20.0	9.4
	β_1	−0.020	−3.6	0.00	−0.020	15.1	0.00	−0.020	16.7	0.00
	a_0	170	3.7	140.4	165	6.3	352.0	160	10.0	398.4
	b_0	1.70	−0.4	0.02	1.85	6.0	0.18	2.00	7.0	0.17
	k_0	0.25	−5.8	0.001	0.21	−7.2	0.009	0.175	38.8	0.106

% Missing data	Parameter	Growth scenario								
		$a_0 \gg a_1$			$a_0 > a_1$			$a_0 = a_1$		
		True	**RB**	**MSE**	**True**	**RB**	**MSE**	**True**	**RB**	**MSE**
	a_1	160	2.2	90.7	160	−0.2	265.2	160	−1.3	26.0
	b_1	2.00	−0.4	0.005	2.00	1.8	0.011	2.00	−0.3	0.003
	k_1	0.175	−2.6	0.00	0.175	17.9	0.00	0.175	2.7	0.00
	σ	4	4.6	0.18	4	7.9	0.16	4	−6.3	0.18

Relative bias (RB) was calculated as the difference of the estimated and true parameter values, divided by the true value. Mean square error (MSE) was computed by summing the squared bias and variance

Table 2

Parameter summaries for sample size $N = 100$ across percentage of missing data and level of difference in Gompertz curves (0 = unobserved, 1 = observed)

% Missing data	Parameter	Growth scenario								
		$a_0 \gg a_1$			$a_0 > a_1$			$a_0 = a_1$		
		True	**RB**	**MSE**	**True**	**RB**	**MSE**	**True**	**RB**	**MSE**
20 %	β_0	4.148	13.2	2.9	4.148	21.6	4.5	4.148	15.4	3.5
	β_1	−0.013	15.7	0.00	−0.013	28.6	0.00	−0.013	16.9	0.00
	a_0	170	0.3	8.5	165	1.4	19.6	160	5.1	84.0
	b_0	1.70	2.1	0.04	1.85	5.2	0.08	2.00	−0.9	0.08
	k_0	0.25	1.2	0.001	0.21	0.0	0.00	0.175	−1.0	0.007
	a_1	160	0.4	4.9	160	0.0	4.6	160	−0.7	4.1
	b_1	2.00	−0.5	0.001	2.00	−0.3	0.001	2.00	0.4	0.002
	k_1	0.175	−0.9	0.00	0.175	0.0	0.00	0.175	1.5	0.00
	σ	4	2.8	0.09	4	2.4	0.11	4	−8.0	0.23
30 %	β_0	4.520	13.5	2.5	4.520	6.5	2.7	4.520	5.2	3.0
	β_1	−0.017	14.3	0.00	−0.017	7.6	0.00	−0.017	2.7	0.00
	a_0	170	0.5	4.7	165	1.1	9.6	160	4.7	90.7
	b_0	1.70	0.5	0.01	1.85	1.3	0.03	2.00	−0.5	0.03
	k_0	0.25	−0.2	0.000	0.21	−1.7	0.000	0.175	−7.8	0.000
	a_1	160	0.1	6.6	160	0.4	7.5	160	−0.9	4.6
	b_1	2.00	0.1	0.001	2.00	0.0	0.001	2.00	0.5	0.002
	k_1	0.175	0.6	0.00	0.175	−0.4	0.00	0.175	2.0	0.00
	σ	4	2.6	0.12	4	3.1	0.12	4	−8.3	0.20
40 %	β_0	4.872	1.4	2.0	4.872	3.0	2.0	4.872	8.5	3.7
	β_1	−0.020	−0.3	0.00	−0.020	2.0	0.00	−0.020	7.3	0.00
	a_0	170	0.7	5.1	165	1.5	14.8	160	3.8	55.8
	b_0	1.70	−1.3	0.01	1.85	0.5	0.02	2.00	−1.6	0.01

% Missing data	**Parameter**	**Growth scenario**								
		$a_0 \gg a_1$			$a_0 > a_1$			$a_0 = a_1$		
		True	**RB**	**MSE**	**True**	**RB**	**MSE**	**True**	**RB**	**MSE**
	k_0	0.25	−2.5	0.000	0.21	−2.3	0.000	0.175	−6.3	0.000
	a_1	160	0.4	10.5	160	0.3	11.2	160	−1.3	8.2
	b_1	2.00	0.0	0.002	2.00	0.2	0.002	2.00	0.7	0.003
	k_1	0.175	−0.2	0.00	0.175	0.1	0.00	0.175	3.1	0.00
	σ	4	4.0	0.12	4	5.3	0.13	4	−7.0	0.17

Relative bias (RB) was calculated as the difference of the estimated and true parameter values, divided by the true value. Mean square error (MSE) was computed by summing the squared bias and variance

Table 3

Parameter summaries for sample size $N = 500$ across percentage of missing data and level of difference in Gompertz curves (0 = unobserved, 1 = observed)

% Missing data	Parameter	Growth scenario								
		$a_0 \gg a_1$			$a_0 > a_1$			$a_0 = a_1$		
		True	RB	MSE	True	RB	MSE	True	RB	MSE
20 %	β_0	4.148	−0.7	0.5	4.148	3.5	0.6	4.148	−6.2	0.6
	β_1	−0.013	−2.0	0.00	−0.013	4.9	0.00	−0.013	−9.3	0.00
	a_0	170	0.0	1.2	165	−0.1	1.7	160	1.7	9.1
	b_0	1.70	0.7	0.00	1.85	0.7	0.01	2.00	1.1	0.01
	k_0	0.25	0.7	0.000	0.21	0.2	0.000	0.175	−2.6	0.000
	a_1	160	0.0	1.0	160	0.0	0.7	160	−0.5	1.1
	b_1	2.00	0.0	0.000	2.00	0.1	0.000	2.00	0.0	0.000
	k_1	0.175	0.0	0.00	0.175	0.6	0.00	0.175	1.1	0.00
	σ	4	0.7	0.03	4	2.2	0.04	4	−4.9	0.06
30 %	β_0	4.520	1.2	0.5	4.520	1.0	0.3	4.520	4.8	0.7
	β_1	−0.017	−0.1	0.00	−0.017	0.8	0.00	−0.017	5.2	0.00
	a_0	170	0.1	0.9	165	0.0	1.2	160	1.4	5.8
	b_0	1.70	−0.1	0.00	1.85	0.4	0.00	2.00	0.5	0.01
	k_0	0.25	0.2	0.000	0.21	0.4	0.000	0.175	−1.8	0.000
	a_1	160	−0.1	0.8	160	0.3	1.5	160	−0.9	2.7
	b_1	2.00	−0.1	0.000	2.00	0.1	0.000	2.00	−0.1	0.001
	k_1	0.175	0.1	0.00	0.175	−0.4	0.00	0.175	1.4	0.00
	σ	4	1.6	0.04	4	2.0	0.04	4	−4.7	0.06
40 %	β_0	4.872	−1.4	0.4	4.872	0.0	0.5	4.872	2.6	0.9
	β_1	−0.020	−3.6	0.00	−0.020	0.0	0.00	−0.020	2.1	0.00
	a_0	170	0.0	1.1	165	0.1	1.4	160	1.1	4.3
	b_0	1.70	0.7	0.00	1.85	0.7	0.00	2.00	0.1	0.00

% Missing data	Parameter	Growth scenario								
		$a_0 \gg a_1$			$a_0 > a_1$			$a_0 = a_1$		
		True	RB	MSE	True	RB	MSE	True	RB	MSE
	k_0	0.25	0.8	0.000	0.21	0.0	0.000	0.175	−1.9	0.000
	a_1	160	0.1	1.6	160	−0.1	1.8	160	−1.0	3.8
	b_1	2.00	0.3	0.000	2.00	−0.2	0.000	2.00	0.0	0.001
	k_1	0.175	0.1	0.00	0.175	0.3	0.00	0.175	1.6	0.00
	σ	4	1.8	0.05	4	2.7	0.06	4	−5.0	0.08

Relative bias (RB) was calculated as the difference of the estimated and true parameter values, divided by the true value. Mean square error (MSE) was computed by summing the squared bias and variance

Table 4

Posterior means and standard deviations (SD) for model (i) parameters from the fit to the SC MMSN data

Parameter	Mean	SD
β_0	2.869	0.675
β_1	−0.015	0.004
a_0	186.797	3.192
b_0	1.727	0.063
k_0	0.256	0.025
a_1	137.734	3.141
b_1	2.317	0.093
k_1	0.287	0.017
σ	4.973	0.027