Bias Adjustment in Bayesian Estimation of Bird Nest Age-specific Survival Rates

Jing Cao and Chong Z. He*

Department of Statistics, University of Missouri-Columbia,
Columbia, Missouri 65211, U.S.A.

*e-mail: chong@stat.missouri.edu

SUMMARY. The populations of many North American landbirds are showing signs of declining. Gathering information on breeding productivity allows early detection of unhealthy populations and helps develop good habitat-management practices. In this paper, we study the performance of the Bayesian model (He, 2003) for age-specific nest survival rates with irregular visits. We find that the estimates are satisfactory except for the age-one survival rate. Usually the more days skipped between two visits, the more serious the underestimation of the age-one survival rate. We investigated the problem and developed three approaches to adjust for the underestimation bias. The simulation results show that the three approaches can significantly improve the estimation of the age-one survival rate. Finally the adjustment is illustrated through an analysis of the Missouri Dickcissel data set.

KEY WORDS: Age-specific survival; Bayesian estimation; MCMC; Bird nest survival rate; Double interval censored data; Hierarchical prior; Noninformative prior.
1 Introduction

Birds are valued and highly visible components of natural ecosystems, and they serve as indicators of environmental quality. Consequently, many efforts have been directed toward measuring and monitoring the condition of bird fauna. Researchers discover that the populations of many North American landbirds are showing signs of declining. The information on breeding productivity not only provides an early warning signal that allows detection of unhealthy populations but also helps identify habitat conditions associated with successful and failed breeding attempts. Such information is also critical for developing habitat- and land-management practices that will maintain healthy bird populations.

There is a large literature on the method of estimating nest survival rates. The Mayfield (1961, 1975) method and its modifications (Miller and Johnson, 1978; Johnson, 1979; Hensler and Nichols, 1981; Bart and Robson, 1982) have been the standard methods. Its key assumption is that the daily survival rate is constant regardless of the nest’s age, which is unrealistic in many cases. Pollock and Cornelius (1988) proposed a discrete survival model that allows survival rate to be age-specific, but it requires that the nests are revisited daily. They also suggested that the nest life span be grouped into longer intervals so that the number of parameters is small enough to avoid instability problems. Heisey and Nordheim (1990) pointed out that this grouping may lead to unacceptable bias in the daily survival rate estimates. Heisey and Nordheim (1995) used the bivariate approach to model both the age of the nest at recruitment and the age at failure in the irregular visit case. The likelihood is maximized using the EM algorithm. This algorithm is often slow to converge and may produce unstable estimates when the nest life span is relatively long.

He, Sun, and Tra (2001) presented a Bayesian model for age-specific nest survival rates when the nests are revisited daily. The Bayesian approach does not need the grouping of data or the structural assumptions about unknown parameters. In reality, the schedule of daily visits is hard to maintain because of bad weather or schedule conflicts, etc. Furthermore, biologists believe that the act of observation can influence the nest’s survival. For example, the direct disturbance of the nests may cause the parent birds to abandon the nests or inform predators of the locations of nests they might have otherwise overlooked (Cornelius, 1993).
He (2003) extended the Bayesian model, which accommodates any mixture of observing schedule, to handle the case of general irregular visits. Because this model gives the estimates of daily survival rates, it is a great help for the biologists to study bird nest survival patterns.

In practice, biologists are trying to visit the nests as often as possible to ensure a good estimate based on the Mayfield methods. However, they would like to obtain sufficient information on the nest survival rate while trying to disturb the nests as little as possible. Once He’s method is available, a natural question to ask is how many days can be safely skipped between visits without invalidating the estimation. To answer this question, we have tried different visiting schedules with He’s model in the simulation study. Accidentally, we find that the age-one survival rate could be underestimated by He’s method under some irregular visiting schedules. The underestimation becomes more serious as more days are skipped between visits, and it does not disappear by simply increasing the sample size.

In this paper, we show why and when the underestimation of age-one survival rate occurs and propose three approaches to adjust the bias. All the approaches improve the estimate significantly and hence help to make the irregular visiting schedule more feasible in reality. The intensive simulation study searching the appropriate visiting schedule will be reported in another paper to be submitted to a wildlife journal. In Section 2, we briefly introduce the basic framework of the Bayesian model (He, 2003). In Section 3, we show the simulation results to illustrate the underestimation of survival rate at age one, and then we investigate the problem and summarize the conditions for the underestimation. In Section 4, three adjustment approaches are proposed and compared. In Section 5, the approaches are applied to the Missouri Dickcissel data set. Finally, we present a discussion of the results in Section 6.

## 2 Basic Framework of the Bayesian Model

In nest survival studies, biologists search for active nests in a marked area and record the date of first encounter and revisits, as well as the nests’ condition (failed, still active, or succeeded). Revisits stop once the nest has failed (destroyed or abandoned) or succeeded (at least one egg hatches or one baby bird fledges, as defined in the study). Note that the data
are left-truncated because nests that failed before they would be discovered are not included in the data. With irregular visits, most of the data are double-interval censored, i.e., neither the origin (date incubation started) nor the outcome date is known exactly.

Let \( J \) be the number of days in the nesting period, \( n \) be the number of observed nests (sample size) and \( k \) designate the \( k \)th observed nest. Define \( u_k, z_k \) and \( t_k \) to be the age at first encounter, the number of days under observation and the age at outcome of the \( k \)th nest respectively; then we have \( t_k = u_k + z_k - 1 \). Define

\[
y_k = \begin{cases} 
1, & \text{if the } k \text{th nest is a success}, \\
0, & \text{otherwise}.
\end{cases}
\]

With irregular visits, we know \( z_k \) only up to an interval \([Z_{L_k}, Z_{R_k}]\), where \( Z_{L_k} \) is the lower boundary and \( Z_{R_k} \) is the higher boundary of \( z_k \). Similarly, we know \( u_k \) only up to \([U_{L_k}, U_{R_k}]\). Define \( y = (y_1, \ldots, y_n)' \), \( Z = (Z_{L_1}, \ldots, Z_{L_n})' \), \( Z_R = (Z_{R_1}, \ldots, Z_{R_n})' \), \( U_L = (U_{L_1}, \ldots, U_{L_n})' \), and \( U_R = (U_{R_1}, \ldots, U_{R_n})' \). The observed data is of the form \((y, U_L, U_R, Z_L, Z_R)\).

Let \( N \) be the number of discoverable nests (including both observed nests and truncated nests). Because we cannot observe \( u = (u_1, \ldots, u_n)' \) and \( t = (t_1, \ldots, t_n)' \) directly with irregular visits, they enter the model as latent variables. To simplify the likelihood function, He (2003) also introduced \( g = (g_{ij}, 1 \leq j < i \leq J)' \) as latent variables, where \( g_{ij} \) is the number of nests that failed at age \( j \) before otherwise being encountered at age \( i \) \((1 \leq j < i \leq J)\). Note that \( g_{ij} \) is unknown because of truncation and \( \sum_{j \geq i > j \geq 1} g_{ij} = N - n \).

Let

\[
\delta_i = P(u = i \mid u \leq J) \quad \text{and} \quad q_i = P(t = i, y = 0).
\]

Here \( \delta_i \) is the conditional encounter probability at age \( i \) given the nest is discoverable (a nest is discoverable if \( u \leq J \); a nest is observed if \( u \leq J \) and \( u \leq t \)), and \( q_i \) is the failure rate at age \( i \). Note that \( \delta_1 + \cdots + \delta_J = 1 \) and \( 0 \leq \delta_1 + \cdots + q_J \leq 1 \). Let \( q_{J+1} = 1 - (q_1 + \cdots + q_J) \), then \( q_{J+1} \) is the total survival rate. Because the ranges of these parameters are irregular, He (2003) has reparameterized \( \delta_i \) and \( q_j \) into discrete hazard rates so that the ranges of new parameters are unit intervals. Define the hazard rates corresponding to \( \delta_i \) and \( q_j \) by

\[
\alpha_i = \frac{\delta_i}{\delta_1 + \cdots + \delta_J} \quad \text{for} \quad i = 1, \ldots, J - 1, \quad \xi_j = \frac{q_j}{q_1 + \cdots + q_{J+1}} \quad \text{for} \quad j = 1, \ldots, J.
\]
and $\alpha_J = \xi_{J+1} = 1$. Note that the daily survival rates are

$$s_i = P(t > i | t > i - 1) = 1 - \xi_i \quad \text{for } i = 1, \ldots, J.$$  

With independent beta priors for $\alpha_i$ and $\xi_j$ and a noninformative prior for $N$, the full conditional posterior distributions are all standard ones, such as beta, negative binomial, multinomial, and discrete distributions. Gibbs sampling is done easily without using any other complicated sampling methods. The likelihood function, posterior density, and full conditional densities are given in the Appendix.

### 3 Underestimation Bias

#### 3.1 Simulation Background

In many nest survival studies, biologists define a nest to be a success if at least one baby bird fledges. The nesting period usually consists of two stages: incubation stage and nestling stage. Grant, Shaffer, Madden, and Pietz (2004) presented a survival pattern for Vesper Sparrow as shown in Figure 1. In their study, the nests were aged on their first encounter. The survival pattern obtained from aged nests is relatively reliable. In the simulation study, we use this survival pattern for the illustration. The results are similar with other survival patterns.

The geometric encounter distribution has often been used in simulation studies (Bromaggin, 1991; Heisey and Nordheim, 1990, etc). We use geometric probability distribution (with parameter $p = 0.4$) to set the encounter rates. It is worth mentioning that the higher the age one conditional encounter rate, the better the estimation (Cornelius 1993). We strongly recommend that the biologists obtain as many observations encountered at age one as possible.

Assuming $J$ is 21, we generate $u_k$ and $t_k$ independently until we have $n$ nests with $u_k \leq t_k$. If $u_k > t_k$, the nest is discarded. Then we get the double-interval censored data $[U_{L_k}, U_{R_k}]$ and $[Z_{L_k}, Z_{R_k}]$ by fitting $u_k$ and $t_k$ into the predetermined irregular visiting schedule. We repeat this process 1,000 times and generate 1,000 samples of size $n$. In our simulation,
3.2 Explanation of the Underestimation

Figure 1 shows the average of the Bayesian estimates of age-specific survival rates over 1,000 samples along with their square root of mean squared error (\( \sqrt{\text{mse}} \)) under different visiting schedules. We can see that the more days skipped between visits, the larger the underestimation bias at age one. The \( \sqrt{\text{mse}} \) of the estimated survival rate at age one also becomes larger. Meanwhile, all the other daily survival rate estimates are relatively accurate and stable. One serious consequence of the underestimation is that the total survival rate, which some biologists are more concerned with, is significantly underestimated. Simulations with different values of the true survival rates reveal the same problem. So if we can adjust the underestimation of age-one survival rate, He’s Bayesian model would work well even when several days are skipped between visits. This would be advantageous for biologists because they cannot only save money and other resources but also reduce the effect of observing disturbances, which in turn could bring in more accurate information.

To explore the cause of the underestimation, we need to know how the Bayesian computation works. In each Gibbs sampling iteration, the encounter hazard rates (\( \alpha_i \)) and the survival hazard rates (\( \xi_j \)) are sampled from their full conditionals, and then latent variables \( u_k \) and \( t_k \) (\( k = 1, \ldots, n \)) are sampled. So in the raw data, we may know \( u_k \) only to an interval; but in each iteration, it is the sampled point estimate of \( u_k \) that is involved further in the estimation. Finally, the parameter \( N \) and latent variables \( g \) are sampled.

For the convenience of the analysis, we introduce a parameter, \( a \), which controls the visiting schedule. If the visit is made daily, \( a = 1 \); if the visit is made every two days, \( a = 2 \) and so on. Let \( S_{ij} \) be the set of the successful nests encountered between age \( i \) and \( j \). Then

\[
S_{ij} = \{ \text{nests} | (y_k = 1, U_{L_k} = i, U_{R_k} = j, Z_{L_k} = J - j + 1, Z_{R_k} = J - i + 1), k = 1, \ldots, n \}.
\]

We traced the Gibbs sampling cycles and found the size of \( S_{11} \) (number of sampled \( u_k=1 \) and \( y_k=1 \)) is too small under some irregular visiting schedules. For example, with \( J = 21 \), \( n = 300 \) and \( a = 3 \) (visit every three days), we cannot directly observe \( S_{11} \). Instead, \( S_{12} \) is observed, in which \( S_{11} \) and \( S_{22} \) are not distinguishable because they have the same data.
form (see Figure 2 for illustration). In each Gibbs sampling cycle, $S_{11}$ and $S_{22}$ are sampled out of $S_{12}$, according to a binomial distribution (see Appendix A.4.(12)). Once we sampled a small value of age-one encounter rate $\alpha_1$, it makes the probability of sampling $S_{11}$ out of $S_{12}$ very small, which may lead to an empty set of $S_{11}$. It further makes the number of nests encountered at age one ($m_1$) small, and a small $m_1$ makes $\alpha_1$ smaller in return (A.4.(9)). Meanwhile, a smaller $\alpha_1$ leads to a larger $N$ and larger numbers of truncated nests failed at age one ($g_{i1}, i > 1$). This causes the underestimation of the age-one survival rate (A.4.(10)). Figure 3 shows the mechanism of the underestimation.

The underestimation seldom occurs with daily visits because $S_{11}$ can be observed and usually it is not empty. Thus $m_1$ is always greater than zero, which prevents $\alpha_1$ from being sampled very small. For the irregular visits, if the visiting schedule is mixed (i.e., different number of days are skipped between visits), $S_{11}$ often is observable and not empty. This is why we did not see any underestimation bias in He (2003) and He, Sun, and Tra (2001). Unobservable $S_{11}$ is more likely to appear under systematic irregular visits, where a fixed number of days are skipped between visits. When $S_{11}$ is not observable, it has to be sampled in each Gibbs sampling cycle. Once $S_{11}$ is sampled empty, then $m_1$ has no greater-than-zero lower boundary. So $\alpha_1$ can be sampled very close to zero and trapped there, which eventually causes the underestimation of the age-one survival rate.

### 3.3 Conditions for the Underestimation

From the above analysis, we conclude that the empty set $S_{11}$ is related to the underestimation of age-one survival rate. The irregular visiting schedule is just a necessary, rather than, a sufficient condition of the underestimation. So if $S_{11}$ is not empty in an irregular visit schedule, there is no underestimation bias. Take $J = 20$ for example: when $a \leq 3$, we can easily show that $S_{11}$ can be observed; when $a = 4$ (visit every four days), all the successful nests encountered between age one and age three belong to one set $S_{13}$. In other words, $S_{11}$, $S_{22}$, and $S_{33}$ are not observable; $S_{13}$ is observed instead. As we can see from Figure 5, the estimate of age-one survival rate is not underestimated for $a \leq 3$; the underestimation occurs at $a = 4$, where $S_{11}$ cannot be observed. Additionally, the $\sqrt{mse}$ of the estimate is
much larger at $a = 4$ than those in the other cases because of the underestimation. This implies that He’s model works well with irregular visits as long as $S_{11}$ is not empty in the data set. Otherwise, we need to adjust the underestimation bias.

Recall that $S_{ij}$ is the set that includes the successful nests encountered between age $i$ and age $j$. Let $f$ be the smallest number that $S_{1f}$ is observable for a given visiting schedule. It is when $f$ is greater than one that we need to adjust for the underestimation. Table 1 shows the relationship between $J$, $a$, and $f$, and when the adjustment is needed under systematic irregular visits. Using our last example of $J = 20$ (Figure 4), when $a = 2$, $f = \text{mod}(20 - 1, 2) = 1$, so $S_{11}$ is observable; when $a = 4$, $f = \text{mod}(20 - 1, 4) = 3$, then $S_{11}$ is not observable, and we have to adjust for the underestimation bias.

4 Adjustment Approaches

Once we understand the cause of the underestimation, it is natural to propose the following three approaches.

**Approach I (Change-one Approach):** Because the underestimation of age-one survival rate is directly caused by the emptiness of set $S_{11}$, we propose to reassign one of the nests in $S_{1f}$ to $S_{11}$. The estimation will be based on the changed data. For example, in a data set with $J = 20$ and $a = 4$, we have 10 successful nests encountered between age one and three (i.e. 10 nests in $S_{13}$). We reassign one of the nests in $S_{13}$ to $S_{11}$, then $S_{13}$ of size 10 is changed to $S_{11}$ of size 1 and $S_{13}$ of size 9, and we use the new data set for the estimation. This approach is an *ad hoc* approach. In practice, biologists always try to make the observations as early as possible, and it’s reasonable to assume that there is at least one observation from $S_{1f}$ that belongs to $S_{11}$.

**Approach II (Iteration Approach):** The second approach involves iteration. In the first step, we apply He’s model to the original data. In the second step, $S_{1f}$ in the original data is partitioned into $S_{ii}$, $i = 1, \cdots, f$, according to the proportion of their estimated encounter rates. In the third step, He’s model is applied to the data from step 2. Repeat the last two steps until the relative difference of the estimated age-one survival rates in the two consecutive rounds is less than a fixed small value (in our simulation, we use 0.01 and
0.03 respectively, and they have shown little difference in the results). This approach is an implementation of the EM algorithm in essence. In our last example, suppose in the first step, the estimated encounter rates for the first three days are 0.3, 0.18, 0.12, respectively. Then the $S_{13}$ of size 10 becomes $S_{11}$ of size 5, $S_{22}$ of size 3 and $S_{33}$ of size 2. The updated data set is used for the second round of estimation and so on. From the trace of the iteration, we find that the underestimation of age-one survival rate is mostly adjusted at the second round of estimation, and the estimates in the iterations afterwards just fluctuate around the estimate from the second round. So we do not need to set the convergence limit to be very small; it takes longer time with little to gain.

**Approach III (Two-stage Prior Approach):** In He’s model (2003), the priors for the survival hazard rates are independent beta distributions with parameter 0.5. Distribution Beta(0.5, 0.5) has its mean at 0.5, and the true value of age-one hazard rate $\xi_1$ is usually much less than 0.5. If there is no data in $S_{11}$, the prior will be dominant for the Bayesian estimation of $\xi_1$, which leads to the overestimation of $\xi_1$ (i.e., underestimation of age-one survival rate). In the third approach, we adjust the bias through the hierarchical priors for the survival hazard rates:

$$
\log \frac{\xi_j}{1 - \xi_j} = \mu + \varepsilon_j, \quad j = 1, \ldots, J
$$

$$
\varepsilon_j \sim N(0, \delta),
$$

$$
\mu \sim N(u, s),
$$

$$
\delta \sim IG(a, b),
$$

where $N(u, s)$ denotes the normal distribution with mean $u$ and variance $s$, and $IG(a, b)$ denotes the inverse gamma distribution with density function

$$
[\delta | a, b] = \frac{b^a}{\Gamma(a)} \frac{1}{\delta^{a+1}} e^{-\frac{b}{\delta}}, \quad \delta > 0.
$$

Define $logit(\xi) = (logit(\xi_1), \cdots, logit(\xi_J))$, then the covariance structure of $logit(\xi)$ given $\delta$
Note that the hierarchical priors we proposed here are positively correlated. Because the logit transformation of $\xi_j$ is a strictly-increasing one-to-one transformation, the priors for $\xi = (\xi_1, \ldots, \xi_J)$ also are positively correlated. It is reasonable to believe that the survival rates in the nesting period are positively autocorrelated. So even though there are no data in $S_{11}$, the covariance structure of the hierarchical priors enable the estimation of age-one hazard rate to “borrow information” from the rest of the data.

If researchers have well-grounded beliefs about the parameters in the prior distribution, it is wise to use the informative priors. If there is little information known about the survival curve, the choice of the prior parameters' values are guided by a desire to provide as little information as possible and to ensure that the posterior distribution reflects trends in the data. In this case, the prior on $\mu$ is a zero-centered normal with a relative large variance ($u = 0, s = 10$). Constant prior is used for $\delta$ ($a = -1, b = 0$). The corresponding full conditionals are given in the Appendix. Note that there is no closed form for the full conditional of $\text{logit}(\xi_j)$, but it can be easily shown it is log-concave, which means it can be sampled efficiently with adaptive rejection sampling (Gilks & Wild, 1992, 1993).

Next we apply the three approaches to the simulated data. Figure 5 shows the original and adjusted estimates and their $\sqrt{mse}$ for $J = 21$ and $a = 3$. Figure 6 is for $J = 20$ and $a = 4$. As we can see from the plots, all the three approaches successfully adjust the underestimation, and they work almost equally well. More specifically, the iteration approach takes a little more time to compute, and it requires the encounter rate at age one to be quite high if the sample size is small to moderate. The change-one approach works well even if the encounter rate at age one is not very high. The two-stage prior approach gives relatively smoother estimates that may not reach the peak or the valley of the survival curve, but it has the smallest variation. Note that the adjustments have improved not only the accuracy of the estimate but also its stability.
The underestimation bias of the total survival rate is automatically corrected after the adjustment of the underestimation of age-one survival rate. For $J = 20$ and $a = 4$, the true total survival rate in the simulation study is 0.27055. The average of the Bayesian estimates (over 1000 samples) of the rate without the adjustment is 0.22336 with $\sqrt{mse}$ of 0.05487. The estimation results from the change-one approach, iteration approach, and two-stage prior approach are 0.25270 (0.03009), 0.24764 (0.03386), and 0.26831 (0.02477), respectively. They have been much improved compared to the one without the adjustment.

5 Application to the Missouri Dickcissel Data

The data set is from McCoy’s study (1996), which was partly funded by the Missouri Department of Conservation. The project was to study whether avian abundance and composition, avian nesting success, and vegetation characteristics differ between cool-season grass and warm-season grass fields in northern Missouri. The study areas were located in three counties (Knox, Macon, and Linn). The same areas were sampled during the years 1993-1995. During the observation period, teams of searchers walked abreast (approximately 1m apart) actively scanning for nests and flushing birds until nesting plots had been systematically traversed. They marked the nests and revisited them every 2-3 days to determine the outcomes. The success of a nest was defined as the nest producing at least one fledging young bird.

For the Dickcissel, $J = 20$. The number of observed nests is 88. Because the visiting schedule is a mixed one, we have one observed nest in $S_{11}$ and three nests in $S_{13}$. So there is no need to apply the adjustment to the original data set. To illustrate our methods, we assume that the nest in $S_{11}$ was missed, which was quite possible in practice. Then we have a sample size $n = 87$ and no data in $S_{11}$, and the data is referred as the reduced data. Because the estimate of age-one encounter rate is quite low (about 7.8%), the iteration approach is not applicable because it assigns zero nest out of the three in $S_{13}$ to $S_{11}$. The other two approaches are shown for the adjustment. There are four sets of estimates in Figure 7. The estimation based on the complete data serves as the baseline for comparison. The other three are the estimates with and without the adjustment based on the reduced data.
respectively. Figure 7 shows that the age-one survival rate is underestimated based on the reduced data with the empty $S_{11}$, and the two adjustment approaches can eliminate the bias as well as decrease the variation of the estimate. Note that the estimates from the two-stage prior approach on the reduced data are even smoother and more stable than those from the complete data. But it may be over smoothed. This is because the estimation based on the complete data is based on He’s method (He, 2003), which has independent beta priors for $\xi$, while the two-stage approach has the correlated hierarchical priors.

6 Discussion

In this article, we have discussed the condition under which the Bayesian model (He, 2003) underestimates the age-one survival rate with irregular visits. We then presented three approaches to adjust the underestimation bias. In practice, one should first check if there is at least one successful nest encountered at age one in the data set. If yes, apply He’s model without the adjustment; otherwise, apply one of the adjustment approaches. Both the simulation study and real data application show that the adjustment approaches can significantly improve the estimation.

Another way to adjust the underestimation bias is to use the informative prior. Suppose we know that the age-one survival rate is around 90% (then the hazard rate is around 10%), then $\text{Beta}(0.1, 0.9)$ can be used instead of $\text{Beta}(0.5, 0.5)$ as the prior for age-one hazard rate. The problem is that the estimate of age-one hazard rate is quite sensitive to the informative beta prior, and it’s not easy to give an accurate informative prior in practice. Additionally, suppose $S_{1f}$ ($f > 1$) is observable in the study. If we provide the informative prior for the age-one hazard rate only and noninformative prior for the rest of hazard rates, then $S_{11}$ is overly preferred to $S_{ii}$ ($i = 2, \cdots, f$) in Gibbs sampling, which may lead to the overestimation of the age-one survival rate and underestimation of the survival rates between age 2 and $f$.

Recall that $f$ is the smallest number such that $S_{1f}$ is observable. In practice, $S_{1f}$ often is not empty. However, if the researchers started the search late and missed the early stage of the incubation, $S_{1f}$ could be observable but empty in the data. In this case, no matter whether $f = 1$ or $f > 1$, the age-one survival rate is always underestimated. The first two
adjustment approaches are no longer valid, but the two-stage prior approach still works.

From Figure 1 and Figure 4, we see that the $\sqrt{mse}$ of the Bayesian estimates with daily visits is larger than that with irregular visits except for the first few days that are influenced by the underestimation. It may seem counter intuitive, but we will see why. In the simulation study, each sample has 300 observations. With $J = 20$, there are 39 parameters to be estimated ($19 \alpha_i$'s and $20 \xi_j$'s), which means we have less than 10 observations for the estimation of each parameter. In this case, a sample of 300 observations is not large enough. With irregular visits, we have observations with larger range on $u_k$ (age at first encounter) and $z_k$ (number of days under observation) than those with daily visits. For a small sample, the enlarged range allows the model to “borrows strength” from the neighbors of the underlining $u_k$ and $z_k$. Usually, as the range increases, the variance of the estimates decreases and the bias of the estimates increases. The trade-off between the variance and the bias determines the value of $a$ (number of days between visits) that yields the smallest $\sqrt{mse}$. With $n = 300$ and the survival pattern shown in this paper, the smallest $\sqrt{mse}$ is reached when $a = 3$ or $a = 4$ (visit every 3 or 4 days). Meanwhile, we have run the simulation with $n = 20,000$. With the sufficiently large sample, the $\sqrt{mse}$ corresponding to daily visits is then the smallest.

Our simulation study also shows that the higher the conditional encounter rate at age one, the better the estimation. As for the question of how many days we could skip between visits without invalidating the estimation, the answer may vary with the encounter-survival pattern and the sample size. With the pattern demonstrated in this paper, the simulation results show that if we have a relatively large sample size, for example $n = 300$, we could skip as many as four days ($a = 5$) and still have satisfactory estimates; however, if the sample size is small (less than 100), no more that two days should be skipped.

For future study, we would like to extend the model by adding some important covariates such as treatment, study area, year, and environmental variables. For a large study area, there may be spatial relationships between different sites. We need to account for these in the model. Finally, we want to use techniques such as Bayes factor or DIC for the model selection.
References


of Wildlife Management 42, 471-476.

*Biometrics* 55, 553-559.


APPENDIX

A.1 Likelihood Function. Let \( \alpha = (\alpha_1, \cdots, \alpha_{J-1})' \) and \( \xi = (\xi_1, \cdots, \xi_J)' \). Then the likelihood function for \( (N, \alpha, \xi) \) based on the latent data \((u, t, g)\) and observed data \((n, y; Z_L, Z_R, U_L, U_R)\) is

\[
L(N, \alpha, \xi) \propto \frac{N!}{\prod_{j \geq i \geq 1} g_{ij}!} \left\{ \prod_{i=1}^{J-1} \alpha_i^{m_i} (1 - \alpha_i)^{l_i} \right\} \left\{ \prod_{j=1}^{J} \xi_j^{m_j} (1 - \xi_j)^{l_j} \right\},
\]

where

\[
m_i = \begin{cases} 
  n_1^S + \sum_{j=1}^{J} n_j^F, & \text{if } i = 1, \\
  n_i^S + \sum_{j=1}^{J-i+1} n_j^F + \sum_{j=1}^{i-1} g_{ij}, & \text{if } i = 2, \cdots, J - 1;
\end{cases}
\]

\[
l_i = \sum_{l=i+1}^{J} \left( n_l^S + \sum_{j=1}^{J-l} n_j^F + \sum_{j=1}^{l-1} g_{ij} \right), \quad \text{if } i = 1, \cdots, J - 1;
\]

\[
m_j^* = \begin{cases} 
  \sum_{l=1}^{J} n_{i,j-l+1}^F + \sum_{i=j+1}^{J} g_{ij}, & \text{if } j = 1, \cdots, J - 1, \\
  \sum_{l=1}^{J} n_{l,j-l+1}^F, & \text{if } j = J;
\end{cases}
\]

\[
l_j^* = \begin{cases} 
  n_j^S + \sum_{l=j+1}^{J} \left( \sum_{i=l+1}^{J} g_{il} \right) + \sum_{i+l-2 \geq j} n_{il}^F, & \text{if } j = 1, \cdots, J - 1, \\
  n_j^S, & \text{if } j = J,
\end{cases}
\]

and \( n^S = \#\{y_k = 1 : k = 1, \cdots, n\} \) (the total number of successful nests), \( n_i^S = \#\{(y_k, u_k) = (1, i) : k = 1, \cdots, n\} \) (the number of successful nests encountered at age \( i \)), and \( n_{ij}^F = \#\{(y_k, z_k, u_k) = (0, j, i) : k = 1, \cdots, n\} \) (the number of failed nests encountered at age \( i \) and observed \( j \) days).

A.2 Priors. The ranges of the components of \( \alpha \) and \( \xi \) are all unit intervals; it is natural to assume independent Beta priors for \( \alpha_i \) and \( \xi_j \). The joint prior density of \((\alpha, \xi)\) is

\[
\pi(\alpha, \xi) \propto \left\{ \prod_{i=1}^{J-1} \alpha_i^{c_i-1} (1 - \alpha_i)^{c_i^*-1} \right\} \left\{ \prod_{j=1}^{J} \xi_j^{d_j-1} (1 - \xi_j)^{d_j^*-1} \right\},
\]

where \( c_i, c_i^*, d_j, \) and \( d_j^* \) are positive constants.

Usually we don’t have enough prior information on \( N \), and a noninformative prior is proper in this case

\[
\pi(N) \propto \frac{1}{N}.
\]
A.3 Posterior Density. The joint posterior of \((N, \alpha, \xi, u, t, g)\) given the observed data \((n, y, Z, U, Z_R, U, U_R)\) is

\[
[N, \alpha, \xi, u, t, g \mid \text{data}] \propto \frac{(N - 1)!}{\prod_{i \geq j \geq 1} \gamma_i} \left\{ \prod_{i=1}^{J-1} \alpha_i^{m_i+c_i-1}(1-\alpha_i)^{l_i+c_i'-1} \right\} \left\{ \prod_{j=1}^{J} \xi_j^{m_j+d_j-1}(1-\xi_j)^{l_j+d_j'-1} \right\}.
\] (8)

A.4 Full Conditional Densities in He’s Model. Given \((\xi, u, t, g, N; \text{data})\), \(\alpha_1, \cdots, \alpha_{J-1}\) are independent and

\[
(\alpha_i \mid \xi, u, t, g, N; \text{data}) \sim \text{Beta}(m_i + c_i, l_i + c_i), \quad i = 1, \cdots, J - 1.
\] (9)

Given \((\alpha, u, t, g, N; \text{data})\), \(\xi_1, \cdots, \xi_J\) are independent and

\[
(\xi_j \mid \alpha, u, t, g, N; \text{data}) \sim \text{Beta}(m_j^* + c_i, l_j^* + c_i), \quad j = 1, \cdots, J.
\] (10)

Given \((\alpha, \xi, g, N; \text{data})\), \((u_1, t_1), \cdots, (u_n, t_n)\) are independent and

\[
[u_k, t_k \mid \alpha, \xi, g, N; \text{data}] = [t_k \mid u_k, \alpha, \xi, g; \text{data}] [u_k \mid N, \alpha, \xi, g; \text{data}].
\] (11)

Here,

\[
P(u_k = i \mid \alpha, \xi, g, N; \text{data})
\] (12)

\[
e \begin{cases} 
\sum_{h=U_{L_k}}^{U_{R_k}} \left\{ \alpha_h \prod_{l=1}^{i-1} (1-\alpha_l) \sum_{j=\min(J, i+Z_{R_k}-1)}^{\min(J, i+Z_{R_k}-1)} \xi_j \prod_{l=1}^{j-1} (1-\xi_l) \right\}, & \text{if } y_k = 0, \quad U_{L_k} \leq i \leq U_{R_k}, \\
\sum_{h=U_{L_k}}^{U_{R_k}} \left\{ \alpha_h \prod_{l=1}^{i-1} (1-\alpha_l) \right\}, & \text{if } y_k = 1, \quad U_{L_k} \leq i \leq U_{R_k}.
\end{cases}
\]

and

\[
P(t_k = j \mid u_k, \alpha, \xi, g, N; \text{data})
\] (13)

\[
e \begin{cases} 
\frac{\xi_j \prod_{l=1}^{j-1} (1-\xi_l)}{\sum_{j_1=\min(J, u_k+Z_{R_k}-1)}^{\min(J, u_k+Z_{R_k}-1)} \xi_j \prod_{l=1}^{j_1-1} (1-\xi_l)}, & \text{if } y_k = 0, \quad \min(J, u_k + Z_{L_k} - 1) \leq j \leq \min(J, u_k + Z_{R_k} - 1), \\
1, & \text{if } y_k = 1, \quad j = J.
\end{cases}
\]
The conditional distribution of \((g, N)\) given \((\alpha, \xi, u, t; \text{data})\) is

\[
[N, g \mid \alpha, \xi, u, t; \text{data}] = [g \mid N, \alpha, \xi, u, t; \text{data}] [N \mid \alpha, \xi, u, t; \text{data}].
\]  

(14)

Here

\[
[N \mid \alpha, \xi, u, t; \text{data}] \propto (N - 1)! (1 - A)^{N-n}/(N-n)!, \quad \text{for } N = n, n + 1, \cdots;
\]

(15)

\[
[g \mid N, \alpha, \xi, u, t; \text{data}] \propto \prod_{1 \leq j < i \leq J} p_{ij}^{g_{ij}}/ \prod_{1 \leq j < i \leq J} g_{ij}!, \quad \text{for } g_{ij} \geq 0, \sum_{1 \leq j < i \leq J} g_{ij} = N - n.
\]

(16)

That is, \((N \mid \alpha, \xi, u, t; \text{data}) \sim \text{Negative Binomial } (n, A)\) and \((g \mid N, \alpha, \xi, u, t; \text{data}) \sim \text{Multinomial } (N - n, p)\), where,

\[
A = 1 - \sum_{j=1}^{J-1} \xi_j (1 - \alpha_j) \left\{ \prod_{l=1}^{j-1} (1 - \alpha_l)(1 - \xi_l) \right\}
\]

(17)

and

\[
p_{ij} = \begin{cases} 
\frac{\alpha_i \xi_j}{1 - \alpha_i} \prod_{l=1}^{i-1} (1 - \alpha_l) \prod_{h=1}^{j-1} (1 - \xi_h), & \text{if } J - 1 \geq i > j \geq 1, \\
\frac{\xi_i \xi_j}{1 - \alpha_i} \prod_{l=1}^{i-1} (1 - \alpha_l) \prod_{h=1}^{j-1} (1 - \xi_h), & \text{if } J - 1 \geq j > i \geq 1.
\end{cases}
\]

(18)

### A.5 Full Conditional Density in the Prior Approach.

Let \(v_j = \logit(\xi_j)\), given the other parameters and data, the conditional posterior of \(v_j\) is

\[
[v_j \mid \text{others}; \text{data}] \propto \left( \frac{e^{v_j}}{1 + e^{v_j}} \right)^{m_j^*} \left( \frac{1}{1 + e^{v_j}} \right)^{t_j^*} e^{-\frac{(v_j - \mu)^2}{a^2}}, \quad j = 1, \cdots, J.
\]

(19)

Given the other parameters and data, the conditional posterior of \(\mu\) is

\[
(\mu \mid \text{others}; \text{data}) \sim \text{N}(L^{-1}G, L^{-1}),
\]

(20)

where

\[
L = \frac{J}{\delta} + \frac{1}{s}, \quad G = \frac{\sum_{j=1}^{J} v_j}{\delta} + \frac{u}{s}.
\]

Given the other parameters and data, the conditional posterior of \(\delta\) is

\[
(\delta \mid \text{others}; \text{data}) \sim \text{IG} \left( \frac{J}{2} + a, \frac{1}{2} \sum_{j=1}^{J} (v_j - \mu)^2 + b \right).
\]

(21)
Table 1: Conditions for the Underestimation of Age-one Survival Rate

<table>
<thead>
<tr>
<th>mod($J - 1, a$)</th>
<th>$f$</th>
<th>$S_{1f}$</th>
<th>$(y, U_L, U_R, Z_L, Z_R)$</th>
<th>Need Adjustment</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>1</td>
<td>$S_{11}$</td>
<td>$(1, 1, 1, J, J)$</td>
<td>No</td>
</tr>
<tr>
<td>2</td>
<td>2</td>
<td>$S_{12}$</td>
<td>$(1, 1, 2, J - 1, J)$</td>
<td>Yes</td>
</tr>
<tr>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
<td>$\vdots$</td>
</tr>
<tr>
<td>$a - 1$</td>
<td>$a - 1$</td>
<td>$S_{1a-1}$</td>
<td>$(1, 1, a - 1, J - a + 2, J)$</td>
<td>Yes</td>
</tr>
<tr>
<td>0</td>
<td>$a$</td>
<td>$S_{1a}$</td>
<td>$(1, 1, a, J - a + 1, J)$</td>
<td>Yes</td>
</tr>
</tbody>
</table>

NOTE: mod($J - 1, a$) = remainder of ($J - 1)/a
Figure 1. Bayesian estimates of age-specific survival rates and their $\sqrt{mse}$ under different visiting schedules with $J = 21$. — true age-specific survival rates; ······ visit daily; - - - - visit every two days; -- visit every three days; — — visit every four days.
1. First encounter at age one

\[ (Y, U_L, U_R, Z_L, Z_R) = (1, 1, 20, 21) \]

2. First encounter at age two

\[ (Y, U_L, U_R, Z_L, Z_R) = (1, 1, 20, 21) \]

3. First encounter at age three

\[ (Y, U_L, U_R, Z_L, Z_R) = (1, 3, 17, 19) \]

Figure 2. Illustration of the visiting schedule for successful nests with \( J = 21 \) and \( a = 3 \) (the dark cells denote the days visited).

Note that the last observation day must be on or after age \( J \) for the successful nests. For example, the successful nests encountered at age one (we don’t know it in practice) have the first visiting schedule, where the last observation is made at age 22.

\[
\begin{align*}
Z_R &= \min(\text{actual number of days under observation}, J) = \min(22 - 1 + 1, J) = 21 \\
Z_L &= \text{smallest number of days under observation possibly} = 20 - 1 + 1 = 20 \\
U_R &= J - Z_L + 1 = 21 - 20 + 1 = 2 \\
U_L &= J - Z_R + 1 = 21 - 21 + 1 = 1
\end{align*}
\]

Similarly, we can show the data for the successful nests encountered at age two and age three. As we can see, the successful nests encountered at age one and age two have the same data, while age three has different data.
Figure 3. Mechanism of the underestimation

\[ \text{small } S_{i1} \rightarrow \text{small } m_1 \]
\[ \text{trap of small } \alpha_1 \]
\[ \text{small } \alpha_1 \]
\[ \text{large } N \rightarrow \text{large } g_{i1} \]
\[ \text{overestimate } \xi_1 \]
\[ \text{underestimate } 1 - \xi_1 \]

age-one survival rate = 1 - \xi_1
Figure 4. Bayesian estimates of age-specific survival rates and their $\sqrt{mse}$ under different visiting schedules with $J = 20$. —— true age-specific survival rates; · · · · visit daily ($a = 1$); ——— visit every two days ($a = 2$); ——— visit every three days ($a = 3$); ——— visit every four days ($a = 4$).
Figure 5. Adjusted Bayesian estimates of age-specific survival rates and their $\sqrt{\text{mse}}$ with $J = 21$ and $a = 3$ by different adjustment approaches. —— true age-specific survival rates; ······· estimates by He’s model; ····· estimates by the change-one approach; ——— estimates by the iteration approach; ——— estimates by the two-stage prior approach.
Figure 6. Adjusted Bayesian estimates of age-specific survival rates and their $\sqrt{mse}$ with $J = 20$ and $a = 4$ by different adjustment approaches. —— true age-specific survival rates; ······ estimates by He’s model; - - - - estimates by the change-one approach; -- -- estimates by the iteration approach; — — estimates by the two-stage prior approach.
Figure 7. Bayesian estimates of age-specific survival rates and their standard deviations from the Missouri Dickcissel data. —— estimates by He’s model with the original data; ······· estimates by He’s model with the reduced data; - - - - estimates by applying the change-one approach to the reduced data; - - - estimates by applying the two-stage prior approach to the reduced data.