



Bayesian Estimation of Survival from Mark-Recapture Data

Author(s): David Poole

Reviewed work(s):

Source: *Journal of Agricultural, Biological, and Environmental Statistics*, Vol. 7, No. 2 (Jun., 2002), pp. 264-276

Published by: [International Biometric Society](#)

Stable URL: <http://www.jstor.org/stable/1400699>

Accessed: 08/06/2012 06:24

Your use of the JSTOR archive indicates your acceptance of the Terms & Conditions of Use, available at
<http://www.jstor.org/page/info/about/policies/terms.jsp>

JSTOR is a not-for-profit service that helps scholars, researchers, and students discover, use, and build upon a wide range of content in a trusted digital archive. We use information technology and tools to increase productivity and facilitate new forms of scholarship. For more information about JSTOR, please contact support@jstor.org.



International Biometric Society is collaborating with JSTOR to digitize, preserve and extend access to *Journal of Agricultural, Biological, and Environmental Statistics*.

<http://www.jstor.org>

Bayesian Estimation of Survival From Mark–Recapture Data

David POOLE

An understanding of survival patterns is a fundamental component of animal population biology. Mark–recapture models are often used in the estimation of animal survival rates. Maximum likelihood estimation, via either analytic solution or numerical approximation, has typically been used for inference in these models throughout the literature. In this article, a Bayesian approach is outlined and an easily applicable implementation via Markov chain Monte Carlo is described. The method is illustrated using 13 years of mark–recapture data for fulmar petrels on an island in Orkney. Point estimates of survival are similar to the maximum likelihood estimates (MLEs), but the posterior variances are smaller than the corresponding asymptotic variances of the MLEs. The Bayesian approach yields point estimates of 0.9328 for the average annual survival probability and 14.37 years for the expected lifetime of the fulmar petrels. A simple modification that accounts for missing data is also described. The approach is easier to apply than augmentation methods in this case, and simulations indicate that the performance of the estimators is not significantly diminished by the missing data.

Key Words: Bayesian inference; Capture–recapture; Fulmar petrels; Markov chain Monte Carlo; Maximum likelihood; Missing data; Survival estimation.

1. INTRODUCTION

Mark–recapture models are widely used to estimate animal abundances and survival rates in open populations, i.e., populations that are subject to additions and deletions over time. The basic Jolly–Seber model (Jolly 1965; Seber 1965) is applicable in this situation. A key assumption of the original Jolly–Seber formulation is that the probability of capture is the same for previously marked animals as for unmarked animals. There are also modifications of the Jolly–Seber model (Buckland 1980, 1982; Pollock 1981) that relax or avoid this assumption in various ways.

David Poole is Senior Technical Staff Member, AT&T Labs-Research, 180 Park Avenue, Florham Park, NJ 07932 (E-mail: poole@research.att.com).

©2002 American Statistical Association and the International Biometric Society
Journal of Agricultural, Biological, and Environmental Statistics, Volume 7, Number 2, Pages 264–276

In many studies, new animals are captured, marked, and released in each sample. On later visits, previously marked animals are recaptured and more unmarked animals may be marked and released. Sometimes, previously marked individuals can be identified without recapture. In these cases, the initial capture of unmarked animals and the resighting of previously marked animals involve separate sampling procedures, and the assumption of equally likely capture for animals in both groups does not hold.

Cormack (1964) developed a model for estimating survival in experiments of this type. In contrast with the Jolly–Seber formulation, the Cormack model conditions on previously marked animals, and it therefore employs the lesser restriction that all previously marked animals alive at a given sampling occasion are equally likely to be reidentified on that occasion. In this way, the number of marked animals constitutes the population under study, and it must be assumed that these animals are representative of the wider population in terms of survival rate. Because the likelihood for the Cormack model is contained within the Jolly–Seber likelihood, the basic Cormack model is generally referred to as the Cormack–Jolly–Seber (CJS) model.

The CJS model provides estimates of time-dependent capture and survival probabilities. Cormack (1964) obtained analytic maximum likelihood estimates of these parameters and their asymptotic variances. Sandland and Kirkwood (1981) and Clobert (1981) reduced the number of parameters in the CJS model by allowing for equality constraints among the survival and/or capture probabilities, while Jolly (1982) did the same for the full Jolly–Seber model. Unlike the CJS model, analytic maximum likelihood solutions do not exist for a reduced-parameter Cormack model, so Sandland and Kirkwood (1981) used numerical methods to maximize the likelihood.

In this article, a Bayesian approach to the analysis of mark–recapture data is described and applied to the fulmar petrel (*Fulmaris glacialis*) data analyzed by Cormack (1964). The Bayesian estimates are compared with the maximum likelihood estimates (MLEs) obtained by Cormack (1964). Markov chain Monte Carlo (MCMC) methods are used to obtain samples from the joint posterior distribution of the model parameters. The methodology is then extended to account for missing data caused by the absence of one or more sampling occasions during the experiment. For Cormack-type models, this extension is simpler and easier to apply than more standard methods for missing data such as the EM algorithm or MCMC data augmentation. A brief simulation that illustrates the technique for missing data is included.

Bayesian approaches to mark–recapture estimation have been used in the past (cf., Gaskell and George 1972; Castledine 1981; Rodrigues, Bolfarine, and Galvão Leite 1988; Smith 1988, 1991; George and Robert 1992; Ananda 1997; Madigan and York 1997; Kinas and Bethlem 1998; Fienberg, Johnson, and Junker 1999, for applications involving the size of a closed population). Vounatsou and Smith (1995) and Brooks, Catchpole, Morgan, and Barry (2000) applied a Bayesian approach to survival estimation from ring-recovery data. Dupuis (1995) used a Bayesian method to estimate survival and movement transition probabilities from data collected in a number of strata.

2. BAYESIAN ESTIMATION VIA MARKOV CHAIN MONTE CARLO

Under the CJS model framework, the recapture history of each animal is a sequence of independent Bernoulli events from the time of initial marking until the end of the experiment. At each sampling occasion following initial capture, the individual is resighted or not with probability that depends on survival rates and the chance of recapture. Using the notation of Cormack (1964), the likelihood for the basic model can be written in the form

$$L \propto \prod_{j=1}^{n-1} (\chi_j)^{c_j} (\phi_j)^{v_j} (p_{j+1})^{a_{j+1}} (1 - p_{j+1})^{v_j - a_{j+1}}, \quad (2.1)$$

where n is the number of sampling periods (usually years), ϕ_j is the probability that an animal alive at time j will survive until time $(j + 1)$, χ_j is the probability that an animal known to be alive at time j will not be seen after time j , and p_j is the probability that a marked animal alive at time j will be sighted at time j . We need to assume that the population is closed during the time required to obtain the sample at time j . χ_j is a function of the sighting and survival probabilities and can be written as

$$\chi_j = 1 - \sum_{k=j}^{n-1} \left\{ \left(\frac{p_{k+1}}{1 - p_{k+1}} \right) \prod_{i=j}^k \phi_i (1 - p_{i+1}) \right\}. \quad (2.2)$$

The mark–recapture data enter the likelihood (2.1) as follows: a_j is the number of previously marked animals sighted at time j , c_j is the number of marked animals seen for the last time at time j , and v_j is the number of marked animals known to be alive after time j , i.e., sighted at least once after time j . The numbers of animals marked for the first time in each sample (denoted b_j) are regarded as fixed quantities. Cormack (1964) derived analytic maximum likelihood estimates for all sighting (p_j) and survival (ϕ_j) probabilities except the survival over the last period (ϕ_{n-1}) and the probability of sighting at the last occasion (p_n), which are confounded. (Their product does have an explicit MLE.)

2.1 METROPOLIS–HASTINGS IMPLEMENTATION

MCMC methods (e.g., Gilks, Richardson, and Spiegelhalter 1996) are commonly used to obtain Bayesian estimates in problems involving many parameters. For the CJS model, we construct a Markov chain whose stationary distribution is the joint posterior distribution of the sighting and survival probabilities.

We construct the Markov chain using a Metropolis–Hastings algorithm (Metropolis, Rosenbluth, Rosenbluth, Teller, and Teller 1953; Hastings 1970). Suppose that we wish to sample from $\pi(X)$. If X_t is the state of the chain at time t , then the next state X_{t+1} is chosen as follows:

1. Sample a candidate point X^* from a proposal distribution $q(\cdot \mid X_t)$, where the form of $q(\cdot \mid X_t)$ is essentially arbitrary.

2. Accept the candidate point X^* with probability

$$A(X_t, X^*) = \min \left\{ 1, \frac{\pi(X^*)q(X_t | X^*)}{\pi(X_t)q(X^* | X_t)} \right\}.$$

$A(X_t, X^*)$ is called the acceptance probability of X^* and the fraction on the right is often referred to as the Hastings ratio.

3. If the candidate point is accepted, then the next state is $X_{t+1} = X^*$; otherwise, $X_{t+1} = X_t$.

This process is repeated at each step after the chain is initialized using a sensible starting value X_0 . Under mild conditions, a chain constructed in this way can be shown to converge to the target density $\pi(X)$. When X is multidimensional, say $X = (X_1, X_2, \dots, X_n)$, the algorithm above can be applied to each component X_i in turn and is then known as the single-component (or componentwise) Metropolis–Hastings algorithm. This is the approach taken for the multiparameter CJS model.

In implementations of this type, it is common to run the chain for an initial period of time (the burn-in period) and to store every k th iteration ($k > 1$) thereafter until some prespecified stopping time. This is referred to as thinning the chain, and it reduces dependence among the sample of points from the posterior distribution. More details regarding the implementation and diagnostics are given in Section 3.

2.2 CASE I: COMPLETE DATA

If a complete set of mark–recapture data is available for the n sampling years, the posterior distribution for the basic CJS model is

$$\begin{aligned} \pi(\phi_1, \dots, \phi_{n-1}, p_2, \dots, p_n \mid \text{data}) \\ \propto L(\text{data} \mid \phi_1, \dots, \phi_{n-1}, p_2, \dots, p_n) p(\phi_1, \dots, \phi_{n-1}, p_2, \dots, p_n), \end{aligned}$$

where $L(\cdot \mid \cdot)$ is the Cormack likelihood in (2.1) and $p(\cdot)$ is the prior distribution of the survival and sighting probabilities. We can apply a componentwise Metropolis–Hastings algorithm using the following proposal distribution in each case. For a given value of ϕ_j , propose a candidate ϕ_j^* by transforming to the logit scale, adding a uniform increment centered at the current value, and then transforming back to the probability scale. This leads to

$$\phi_j^* = \frac{\phi_j e^U}{1 + \phi_j (e^U - 1)},$$

where $U \sim \text{uniform}(-k, k)$ for some constant $k > 0$. The conditional distribution of ϕ_j^* given ϕ_j , say $q(\phi_j^* \mid \phi_j)$, then has the form

$$q(\phi_j^* \mid \phi_j) = \frac{1}{2k\phi_j^*(1 - \phi_j^*)},$$

with support $S = \{\phi^* : -k < \log(\phi^*/(1 - \phi^*)) - \log(\phi/(1 - \phi)) < k\}$. It follows that

$$\frac{q(\phi_j | \phi_j^*)}{q(\phi_j^* | \phi_j)} = \frac{\phi_j^*(1 - \phi_j^*)}{\phi_j(1 - \phi_j)}.$$

Let $\phi_j \sim \text{beta}(\alpha, \beta)$ with density

$$p(\phi_j) \propto \phi_j^{\alpha-1}(1 - \phi_j)^{\beta-1}$$

be the prior distribution of ϕ_j . Then the acceptance probability of ϕ_j^* is

$$A(\phi_j, \phi_j^*) = \min \left\{ 1, \frac{(\phi_j^*)^{v_j+\alpha}(1 - \phi_j^*)^\beta \prod_{i=1}^j (\chi_i^*)^{c_i}}{(\phi_j)^{v_j+\alpha}(1 - \phi_j)^\beta \prod_{i=1}^j (\chi_i)^{c_i}} \right\}, \quad (2.3)$$

where χ_i^* is calculated using (2.2) with ϕ_j^* rather than ϕ_j . Note that the analytic form of the beta prior $p(\phi_j)$ is similar to that of the proposal distribution. This leads to the convenient expression of the Hastings ratio in (2.3). The beta distribution is a sensible and flexible choice for a prior on a probability since it has support on the $(0, 1)$ line segment only. It also possesses convenient conjugacy properties in binomial problems. The uniform prior is a special case when $\alpha = \beta = 1$. The derivation above assumes independence of the prior distributions on each ϕ_j .

For the sighting probabilities p_j , the Metropolis–Hastings algorithm works in exactly the same way. A candidate p_j^* is selected using the same proposal mechanism as before. If $p_j \sim \text{beta}(\gamma, \delta)$ is the prior distribution, the acceptance probability can be shown to be

$$A(p_j, p_j^*) = \min \left\{ 1, \frac{(p_j^*)^{a_j+\gamma}(1 - p_j^*)^{v_{j-1}-a_j+\delta} \prod_{i=1}^{j-1} (\chi_i^*)^{c_i}}{(p_j)^{a_j+\gamma}(1 - p_j)^{v_{j-1}-a_j+\delta} \prod_{i=1}^{j-1} (\chi_i)^{c_i}} \right\}, \quad (2.4)$$

where χ_i^* is calculated using p_j^* rather than p_j .

2.3 CASE II: MISSING DATA

Now consider the case where mark–recapture data are missing in year m , where $1 < m < n$, and relabel the subsequent sampling occasions with the subscripts m through $n - 1$ to reflect the fact that there are $n - 1$ (not n) sampling occasions in the experiment.

In many applications of MCMC, missing data can be regarded as extra parameters to be estimated along with the regular parameters. This is an effective approach in many cases, but it cannot be applied here due to the high degree of interdependence among the data. For example, we must have $\sum b_j = \sum c_j$, and for each increment in the number of animals seen at time j , a previous time must exist for which the count of animals sighted at a later time is incremented. In the context of MCMC, the proposal of a value for a missing datum (e.g., the number of previously marked animals sighted in a year in which there was no census) will necessarily result in complex changes to other data that enter the likelihood function. The EM algorithm (e.g., Dempster, Laird, and Rubin 1977) is commonly used to calculate posterior modes in missing data problems but would also be complicated and difficult to

apply in this case. In addition, with EM, it is difficult to obtain the posterior variances and correlations that are immediately available from sampling-based approaches.

Another option is to circumvent the problem of missing data by explicitly modifying the priors and/or the likelihoods in the model. For the basic CJS model, we interpret the survival from time $m - 1$ to time $m + 1$ as a two-period survival rate. Thus, ϕ_{m-1} , which would usually represent survival between times $m - 1$ and m , now represents the survival probability from time $m - 1$ to $m + 1$. The data require no modification other than noting that there are only $n - 1$ samples in the experiment. Because the method is Bayesian, we can alter the prior distribution on ϕ_{m-1} to reflect the fact that it represents a survival over two periods. In so doing, the Bayesian approach is easily able to accommodate this extra information. This is especially relevant if the prior distributions are strongly nonuniform and place high prior mass on specific survival probabilities (large values, e.g.) and is less informative if the priors are flat.

If the independent prior distributions on the survival rates ϕ_j ($j \neq m - 1$) are $\text{beta}(\alpha, \beta)$, as before, then it follows that the prior on ϕ_{m-1} should be distributed as the product of two independent $\text{beta}(\alpha, \beta)$ random variables. If $\alpha = \beta = 1$, this product prior has the form

$$p(\phi_{m-1}) = -\log(\phi_{m-1}), \quad 0 < \phi_{m-1} < 1, \quad (2.5)$$

and the resulting modification to the Hastings ratio is simple.

As an extension, if there are $d \geq 2$ consecutive missing samples, then (2.5) generalizes to

$$p(\phi_d) = (-\log(\phi_d))^d / d!, \quad 0 < \phi_d < 1,$$

where ϕ_d is the survival over the d missing periods.

When $\alpha \neq 1$ or $\beta \neq 1$, the derivation of the distribution of the product is more difficult. If we assume that $\phi_m = \phi_{m-1}$, then the prior on ϕ_{m-1} is distributed as the square of a single $\text{beta}(\alpha, \beta)$ random variate rather than as the product of two independent variates, and it then has the simple form

$$p(\phi_{m-1}) \propto \phi_{m-1}^{\alpha/2-1} (1 - \phi_{m-1}^{1/2})^{\beta-1}, \quad 0 < \phi_{m-1} < 1,$$

for all values of α and β . Using the usual Metropolis–Hastings scheme, the acceptance probability for ϕ_{m-1}^* turns out to be

$$\begin{aligned} & A(\phi_{m-1}, \phi_{m-1}^*) \\ &= \min \left\{ 1, \frac{(\phi_{m-1}^*)^{v_{m-1}+\alpha/2} (1 - (\phi_{m-1}^*)^{1/2})^{\beta-1} (1 - \phi_{m-1}^*) \prod_{i=1}^{m-1} (\chi_i^*)^{c_i}}{(\phi_{m-1})^{v_{m-1}+\alpha/2} (1 - \phi_{m-1}^{1/2})^{\beta-1} (1 - \phi_{m-1}) \prod_{i=1}^{m-1} (\chi_i)^{c_i}} \right\}. \end{aligned} \quad (2.6)$$

The Hastings ratios remain as in (2.3) for all the other ϕ_j , $j \neq m - 1$, and as in (2.4) for all the sighting probabilities.

3. APPLICATION: FULMAR PETREL DATA

Cormack (1964) illustrated his model using 13 years of mark–recapture data on fulmar petrels on a small island in Orkney. Each summer between 1950 and 1962, zoologists from the University of Aberdeen visited the island and marked new birds by tagging the legs with colored rings while the birds were in their nests. The rings were visible during flight, so recapture on subsequent visits did not require the birds to be caught again; they were simply observed. In this way, the probabilities of initial capture (tagging) and subsequent recapture (observation in flight) in a given year were not the same. The data are shown in Table 1, using the notation from Section 2. Further details on the fulmar study are given in Carrick and Dunnet (1954) and Dunnet, Anderson, and Cormack (1963).

The Bayesian analysis of the same data set used independent uniform priors for all sighting and survival probabilities. The Markov chain was run for 250,000 iterations and the first 50,000 of these were discarded as the burn-in sample. In order to obtain approximately independent points in the posterior sample, every 200th iteration for the remaining length of the chain was recorded. This yielded a sample of size 1,000 from the target posterior distribution. The entire simulation took approximately 10 minutes. Time series plots of the componentwise posterior samples revealed that convergence and mixing of the chain had been achieved very quickly. More formally, the `gibbsit` diagnostic software (Raftery and Lewis 1996) confirmed that the burn-in time and total length of the chain were sufficient. For example, `gibbsit` indicated that a run of length 150,000 would have achieved acceptable convergence. In addition, further runs using longer chains and greater burn-in times yielded no significant improvement. For each parameter, the size of the bound on the uniform increment in the proposal distribution was chosen so that an overall acceptance rate of approximately 35% was obtained. Gilks et al. (1996) suggest an acceptance rate of 20–40% for most Metropolis–Hastings algorithms.

The results are shown in Table 2. The posterior means and modes were similar to the medians in nearly every case. Note that the posterior medians were close to Cormack’s

Table 1. Mark–Recapture Data for Fulmar Petrels Used by Cormack (1964)

Year	No. captured b_j	No. resighted a_j	No. seen for last time c_j	No. known alive v_j
1950	11	—	1	10
1951	66	4	7	69
1952	28	36	4	93
1953	2	43	3	92
1954	4	54	4	92
1955	51	63	10	133
1956	13	69	9	137
1957	5	99	18	124
1958	19	85	10	133
1959	8	51	4	137
1960	26	102	16	147
1961	3	133	34	116
1962	18	116	134	0

Table 2. Bayesian Posterior Medians and Maximum Likelihood Estimates for the Fulmar Data. The last two columns are the posterior and asymptotic maximum likelihood standard deviations for the survival probabilities

Year	Survival $\hat{\phi}_j$		Disappearance $\hat{\chi}_j$		Sighting \hat{p}_j		SD ($\hat{\phi}_j$)	
	Bayes	ML	Bayes	ML	Bayes	ML	Bayes	ML
1950	0.8959	0.9697	0.1630	0.0909	—	—	0.0940	0.0970
1951	0.9333	0.9287	0.1065	0.1000	0.3906	0.3750	0.0360	0.0400
1952	0.9636	0.9735	0.0822	0.0625	0.5023	0.5056	0.0296	0.0387
1953	0.9582	0.9619	0.0767	0.0667	0.4467	0.4453	0.0302	0.0412
1954	0.9552	0.9593	0.0770	0.0690	0.5684	0.5695	0.0314	0.0361
1955	0.9595	0.9664	0.0975	0.0877	0.6609	0.6646	0.0278	0.0346
1956	0.9419	0.9419	0.1129	0.1098	0.4942	0.4897	0.0333	0.0400
1957	0.8560	0.8546	0.1762	0.1731	0.6786	0.6830	0.0393	0.0400
1958	0.9438	0.9444	0.1048	0.0862	0.6591	0.6633	0.0327	0.0374
1959	0.9584	0.9662	0.0778	0.0678	0.3682	0.3670	0.0289	0.0361
1960	0.9019	0.9028	0.1293	0.1250	0.7185	0.7183	0.0330	0.0316
1961	0.8653	—	0.2615	0.2500	0.8723	0.8770	0.0786	—
1962	—	—	—	—	0.8645	—	—	—

maximum likelihood estimates for the base parameters of the model (ϕ_j and p_j) except for the first year, where the variance of the survival estimate was much higher than for subsequent years. The Bayesian and maximum likelihood estimates were also similar for the probability of disappearance (χ_j).

Note that the posterior standard deviations for survival were generally less than the corresponding asymptotic standard deviations, sometimes by a factor of almost $2^{1/2}$. The Bayesian posterior variance was almost half the asymptotic maximum likelihood variance in such cases (see also Figures 1 and 2).

The Bayesian approach yields estimates of the final-year sighting and survival probabilities (p_n and ϕ_{n-1}) for which maximum likelihood solutions do not exist. In Table 2, the posterior medians of these two quantities are very similar. Point estimates for these final-year parameters will necessarily be similar because they appear as a product in the likelihood and cannot be separated. The estimates obtained here can be thought of as estimates of the square root of the product of p_n and ϕ_{n-1} and are the Bayesian alternative to a maximum likelihood estimate of $\phi_{n-1}p_n$, which does exist. In general, the individual estimates of p_n and ϕ_{n-1} are not particularly informative.

Cormack (1964) defined the average survival as the geometric mean of the individual ϕ_j 's and the life expectancy of an individual when they enter the population as $-1/\log(\text{average survival})$. He obtained estimates of the average survival of 0.9420 (0.0100) and expected lifetime of 16.7 (3.0) years, where the quantities in parentheses are the asymptotic standard deviations. The Bayesian posterior distributions, excluding the final-year survival, yielded estimates of 0.9284 (0.0115) for average survival and 13.5 (2.0) years for expected lifetime. In the Bayesian case, the quantities in parentheses are the posterior standard deviations. Cormack noted that the precision of both estimates could be improved if the high-variance initial year was omitted from the analysis. In this case, he obtained 0.9378 (0.0075) for average survival and 15.58 (1.93) years for expected lifetime. The corresponding Bayesian

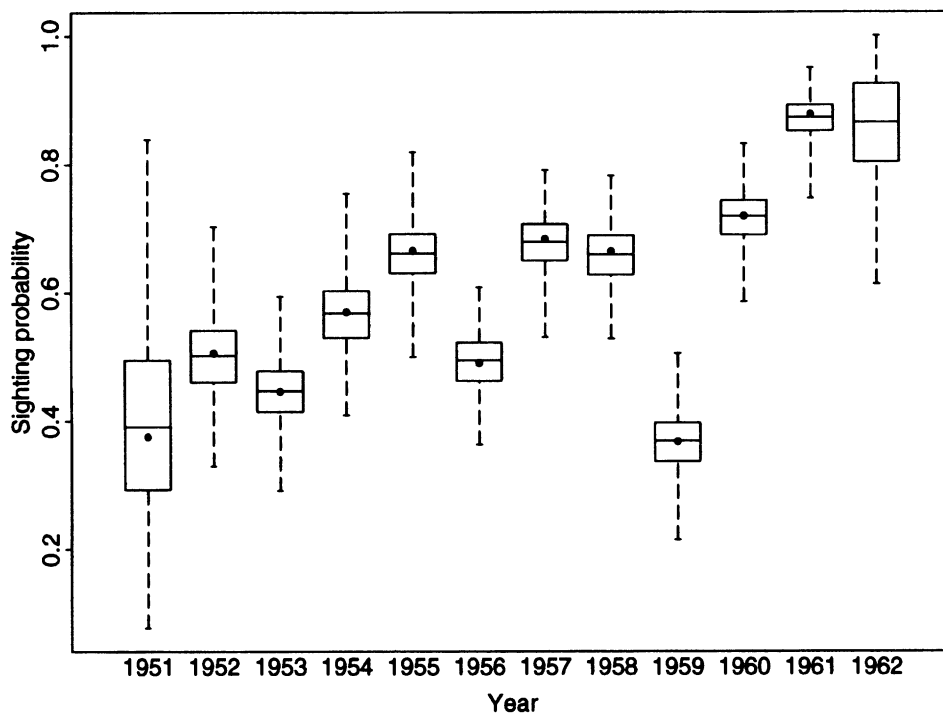


Figure 1. Boxplots of the Posterior Distributions of the Sighting Probabilities for the Fulmar Data. The horizontal lines in the boxes are the posterior medians and the black dots are the maximum likelihood estimates. The boxes represent the interquartile range, while the whiskers extend to the minimum and maximum sampled points.

estimates were 0.9328 (0.0073) for average survival and 14.37 (1.63) for expected lifetime. Removal of the initial year brings the maximum likelihood and Bayesian estimates into very close agreement. The Bayesian posterior distributions of the average survival and life expectancy, excluding the initial year, are shown in Figure 3.

3.1 A SIMULATION FOR MISSING DATA

To examine the results of the procedure when data are missing (see Section 2.3), a small simulation was performed. Assuming true values for all the sighting and survival probabilities given by the Bayesian posterior medians in Table 2, a test set of capture–recapture histories, using the numbers of newly marked birds in the fulmar petrel dataset (the b_j 's in Table 1), was generated. Recall that the numbers of animals seen for the first time in each year are constants in the Cormack model, so they need to be specified in order to generate data from the model. All the sightings in 1958 were then omitted to simulate the lack of a survey in that year.

For all years in which there were simulated data, the method produced accurate estimates of the true survival and sighting probabilities. Of specific interest is the result for the missing year. The true probability of survival between 1957 and 1959 was 0.81 (since

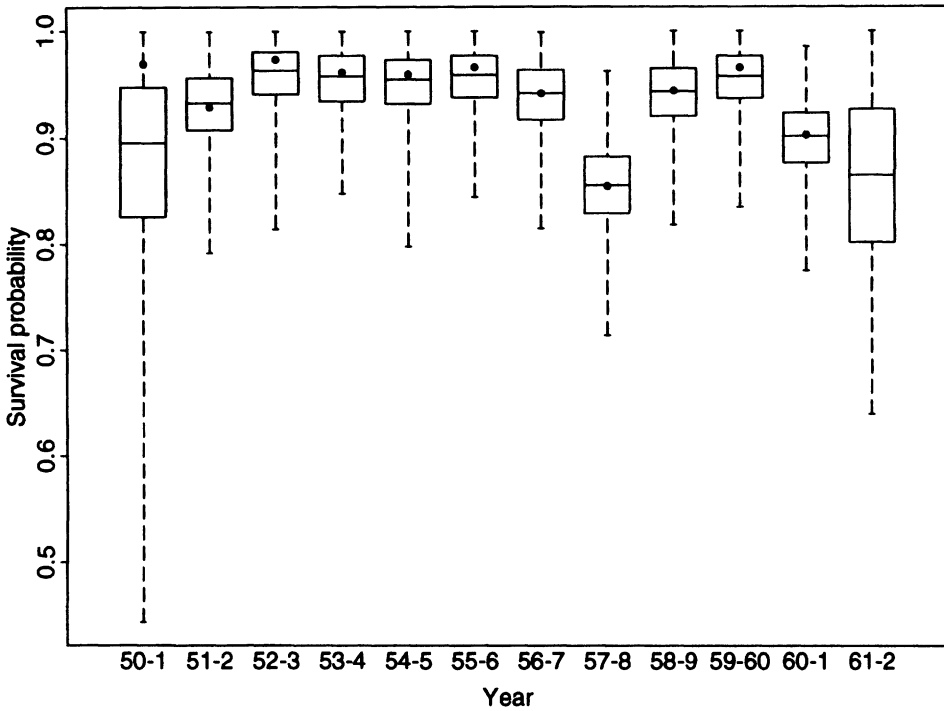


Figure 2. Boxplots of the Posterior Distributions of the Survival Probabilities for the Fulmar Data. The horizontal lines in the boxes are the posterior medians and the black dots are the maximum likelihood estimates. The boxes represent the interquartile range, while the whiskers extend to the minimum and maximum sampled points.

$0.81 \approx 0.8560 \times 0.9438$, from Table 2). The posterior median for this 2-year parameter was 0.78 with 95% probability interval of $[0.70, 0.88]$. These estimates (both the median and the bounds) were considerably lower than any estimates for single-year survivals in Table 2, and the true value fell roughly in the middle of the 95% interval. The amended prior distribution assisted the procedure in obtaining an accurate estimate.

4. DISCUSSION

The Bayesian analysis of the fulmar petrel data yielded point estimates very similar to those of maximum likelihood. For simplicity in this article, independent uniform priors for all the sighting and survival probabilities were used. However, the Bayesian approach allows plenty of scope for other prior distributions. For example, one may choose to place priors on survival that suggest high rates, and this may help reduce the posterior variance when the sighting probabilities are small. In addition, whenever prior correlations, especially among survival probabilities, are known to exist, it should be possible to improve the results by explicitly incorporating these correlations into the analysis. However, it is not clear how feasible it would be to implement such a strategy in practice.

Zeh et al. (2000) applied the Bayesian procedure to photo-identification data for the Bering–Chukchi–Beaufort Seas stock of bowhead whales. These data contain omissions of

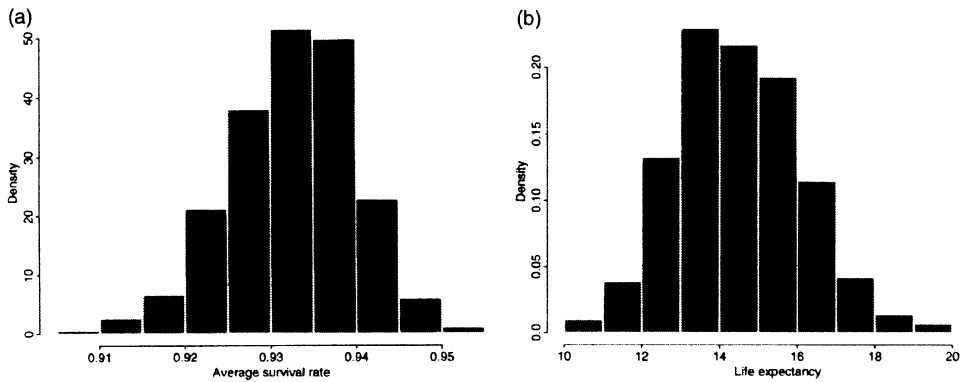


Figure 3. Posterior Distributions of (a) the Average Survival Rate and (b) the Life Expectancy in Years for the Fulmar Dataset.

exactly the type described here, so it is important to check that missing data are accounted for in a reliable manner. In simulations based on the fulmar petrel data, the modified method performed well. The approach is easily extended to account for more than one missing year, although the reliability of results will naturally decrease as the availability of real data diminishes.

George and Robert (1992) used the Gibbs sampler to estimate a population size using a mark–recapture model. In their example, the full conditional distributions were of a standard form, enabling direct sampling. In the Cormack model, the full conditionals are too complex to allow direct sampling, hence, the use of Metropolis–Hastings approaches here. For the fulmar petrel dataset and for the whale data in Zeh et al. (2000), the code that performs the simulations was custom-written specifically for the purpose at hand. However, it should be possible to make use of software packages such as BUGS (Spiegelhalter, Thomas, Best, and Gilks 1994) for problems of this type.

Pollock (1991) noted that Bayesian methods are suited to many mark–recapture studies. Prior information can sometimes be available from a previous study or from earlier in the same study. In general, mark–recapture methods take place over a period of time, so it is appealing to update a posterior distribution as more data become available. Empirical Bayes methods might also be useful. Many mark–recapture models also contain large numbers of nuisance parameters that are difficult to handle using classical methods; Monte Carlo Bayesian methods allow us to perform analyses of such models in a formal yet computationally easier manner than existing classical approaches.

Another advantage of the approach here is that the technical machinery of the Metropolis–Hastings algorithm changes very little among different capture–recapture models; one can simply select the priors and likelihoods of choice and then monitor the various diagnostics (e.g., Gelman 1996; Raftery and Lewis 1996) to ensure adequate mixing and convergence of the Markov chain. The approach developed here could easily be applied to more complex models than that of Cormack (1964) or to reduced models. One possibility is when a single survival rate ϕ , independent of time, is assumed to apply throughout the sampling period.

The sighting probabilities remain dependent on time. In this case, the Cormack likelihood (2.1) reduces to

$$L = \prod_{j=1}^{n-1} (\chi_j)^{c_j} (\phi)^{v_j} (p_{j+1})^{a_{j+1}} (1 - p_{j+1})^{v_j - a_{j+1}}, \quad (4.1)$$

where

$$\chi_j = 1 - \sum_{k=j}^{n-1} \left\{ \left(\frac{p_{k+1}}{1 - p_{k+1}} \right) \phi^{k-j+1} \prod_{i=j}^k (1 - p_{i+1}) \right\}. \quad (4.2)$$

Appropriate and simple modifications to the Hastings ratios can then be made. In this case, the missing data problem is approached in a slightly different way. Instead of specifying a new prior for the 2-year period, the constant survival ϕ would apply twice between the previous and succeeding years, regardless of its prior distribution. The missing data would thus be accounted for via a direct modification to the likelihood function. The method used for obtaining posterior inference in the reduced model remains exactly the same as for the full model.

ACKNOWLEDGMENTS

This research was carried out in part while the author was a graduate student at the University of Washington and was supported by the North Slope Borough, Alaska, and the National Oceanic and Atmospheric Administration (through the National Marine Mammal Laboratory and the Alaska Eskimo Whaling Commission). The author expresses thanks to the editor, the associate editor, and two referees for very useful comments and to Judy Zeh, Adrian Raftery, Julian Besag, and Simon Byers for their helpful practical suggestions.

[Received September 2000. Accepted July 2001.]

REFERENCES

- Ananda, M. M. A. (1997), "Bayesian Methods for Mark-Resighting Surveys," *Communications in Statistics—Theory and Methods*, 26, 685–697.
- Brooks, S. P., Catchpole, E. A., Morgan, B. J. T., and Barry, S. C. (2000), "On the Bayesian Analysis of Ring-Recovery Data," *Biometrics*, 56, 951–956.
- Buckland, S. T. (1980), "A Modified Analysis of the Jolly-Seber Capture-Recapture Model," *Biometrics*, 36, 419–435.
- (1982), "A Mark-Recapture Survival Analysis," *Journal of Animal Ecology*, 51, 833–847.
- Carrick, R., and Dunnet, G. M. (1954), "Breeding of the Fulmar *Fulmarus glacialis*," *Ibis*, 96, 356–370.
- Castledine, B. J. (1981), "A Bayesian Analysis of Multiple-Recapture Sampling for a Closed Population," *Biometrika*, 68, 197–210.
- Clobert, J. (1981), "Etude de la Dynamique d'une Population d'Etourneauux Sansonnets (*Sturnis vulgaris*, L.)," doctoral dissertation, Université Catholique de Louvain, Louvain-la-Neuve, Belgique.
- Cormack, R. M. (1964), "Estimates of Survival From the Sighting of Marked Animals," *Biometrika*, 51, 429–438.
- Dempster, A. P., Laird, N. M., and Rubin, D. B. (1977), "Maximum Likelihood From Incomplete Data via the EM Algorithm" (with discussion), *Journal of the Royal Statistical Society, Series B*, 39, 1–38.

- Dunnet, G. M., Anderson, A., and Cormack, R. M. (1963), "A Study of the Survival of Adult Fulmars With Observations on the Pre-Laying Exodus," *British Birds*, 56, 2–18.
- Dupuis, J. A. (1995), "Bayesian Estimation of Movement and Survival Probabilities From Capture–Recapture Data," *Biometrika*, 82, 761–772.
- Fienberg, S. E., Johnson, M. S., and Junker, B. W. (1999), "Classical Multilevel and Bayesian Approaches to Population Size Estimation Using Multiple Lists," *Journal of the Royal Statistical Society, Series A*, 162, 383–405.
- Gaskell, T. J., and George, B. J. (1972), "A Bayesian Modification of the Lincoln Index," *Journal of Applied Ecology*, 41, 377–384.
- Gelman, A. (1996), "Inference and Monitoring Convergence," in *Markov Chain Monte Carlo in Practice*, eds. W. R. Gilks, S. Richardson, and D. J. Spiegelhalter. London: Chapman and Hall, pp. 131–143.
- George, E. I., and Robert, C. P. (1992), "Capture–Recapture Estimation via Gibbs Sampling," *Biometrika*, 79, 677–683.
- Gilks, W. R., Richardson, S., and Spiegelhalter, D. J. (eds.) (1996), *Markov Chain Monte Carlo in Practice*, London: Chapman and Hall.
- Hastings, W. K. (1970), "Monte Carlo Sampling Methods Using Markov Chains and Their Applications," *Biometrika*, 57, 97–109.
- Jolly, G. M. (1965), "Explicit Estimates From Capture–Recapture Data With Both Death and Immigration—Stochastic Model," *Biometrika*, 52, 225–247.
- (1982), "Mark–Recapture Models With Parameters Constant in Time," *Biometrics*, 38, 301–321.
- Kinas, P. G., and Bethlem, C. B. P. (1998), "Empirical Bayes Abundance Estimation of a Closed Population Using Mark–Recapture Data, With Application to Humpback Whales, *Megaptera novaeangliae*, in Abrolhos, Brazil," *Report of the International Whaling Commission*, 48, 447–450.
- Madigan, D., and York, J. C. (1997), "Bayesian Methods for Estimation of the Size of a Closed Population," *Biometrika*, 84, 19–31.
- Metropolis, N., Rosenbluth, A. W., Rosenbluth, M. N., Teller, A., and Teller, E. (1953), "Equations of State Calculations by Fast Computing Machines," *Journal of Chemistry and Physics*, 21, 1087–1091.
- Pollock, K. H. (1981), "Capture–Recapture Models Allowing for Age-Dependent Survival and Capture Rates," *Biometrics*, 37, 521–529.
- (1991), "Modeling Capture, Recapture, and Removal Statistics for Estimation of Demographic Parameters for Fish and Wildlife Populations: Past, Present, and Future," *Journal of the American Statistical Association*, 86, 225–238.
- Raftery, A. E., and Lewis, S. M. (1996), "Implementing MCMC," in *Markov Chain Monte Carlo in Practice*, eds. W. R. Gilks, S. Richardson, and D. J. Spiegelhalter. London: Chapman and Hall, pp. 115–130.
- Rodrigues, J., Bolfarine, H., and Galvão Leite, J. (1988), "A Bayesian Analysis in Closed Animal Populations From Capture Recapture Experiments With Trap Response," *Communications in Statistics, Part B—Simulation and Computation*, 17, 407–430.
- Sandland, R. L., and Kirkwood, G. P. (1981), "Estimation of Survival in Marked Populations With Possibly Dependent Sighting Probabilities," *Biometrika*, 68, 531–541.
- Seber, G. A. F. (1965), "A Note on the Multiple-Recapture Census," *Biometrika*, 52, 249–259.
- Smith, P. J. (1988), "Bayesian Methods for Multiple Capture–Recapture Surveys," *Biometrics*, 44, 1177–1189.
- (1991), "Bayesian Analyses for a Multiple Capture–Recapture Model," *Biometrika*, 78, 399–408.
- Spiegelhalter, D. J., Thomas, A., Best, N. G., and Gilks, W. R. (1994), *BUGS: Bayesian Inference Using Gibbs Sampling* (Ver. 0.30). Cambridge: Medical Research Council Biostatistics Unit.
- Vounatsou, P., and Smith, A. F. M. (1995), "Bayesian Analysis of Ring-Recovery Data via Markov Chain Monte Carlo Simulation," *Biometrics*, 51, 687–708.
- Zeh, J. E., Poole, D., Miller, G. W., Koski, W. R., Baraff, L., and Rugh, D. J. (2000), "Survival of Bowhead Whales, *Balaena mysticetus*, Estimated From 1981–98 Photo-Identification Data," Paper SC/52/AS19 presented to the IWC Scientific Committee, June 2000.