Biometrics 62, 142–149
March 2006

DOI: 10.1111/j.1541-0420.2005.00399.x

An Extension of the Cormack–Jolly–Seber Model for Continuous Covariates with Application to *Microtus pennsylvanicus*

S. J. Bonner* **and C. J. Schwarz****

Department of Statistics and Actuarial Science, Simon Fraser University,
8888 University Drive, Burnaby, British Columbia V5A 1S6, Canada
**email:* sbonner@stat.sfu.ca
***email:* cschwarz@stat.sfu.ca

Summary. Recent developments in the Cormack–Jolly–Seber (CJS) model for analyzing capture–recapture data have focused on allowing the capture and survival rates to vary between individuals. Several methods have been developed in which capture and survival are functions of auxiliary variables that may be discrete, constant over time, or apply to the population as a whole, but the problem has not been solved for continuous covariates that vary with both time and individual. This article proposes a new method to handle such covariates by modeling changes over time via a diffusion process and using logistic functions to link the variable to the CJS capture and survival rates. Bayesian methods are used to estimate the model parameters. The method is applied to study the effect of body mass on the survival of the North American meadow vole, *Microtus pennsylvanicus*.

Key words: Capture–recapture; Continuous time-dependent individual covariates; Cormack–Jolly–Seber; *Microtus pennsylvanicus*.

1. Introduction

Capture–recapture methods have been used by ecologists to study different animal populations that are difficult to enumerate and follow over time. After trapping, marking, and releasing individuals on one or more capture occasions, inference about the population is obtained by comparing the numbers of marked and unmarked individuals captured at subsequent times. The quantities of interest in many capture–recapture experiments are the animals' survival rates.

The basis of most models developed to study survival from capture–recapture data is the Cormack–Jolly–Seber (CJS) model (Cormack, 1964; Jolly, 1965; Seber, 1965). In its simplest form, the CJS model assigns probabilities to each possible capture history in terms of two sets of parameters. These are the capture probabilities, the probabilities that animals alive at one capture occasion are actually captured, and the survival probabilities, the probabilities that animals alive at one capture occasion are still alive at the next. Although these probabilities are allowed to change over time, the CJS model imposes the assumption that the capture and survival probabilities at a single capture occasion are the same for all animals in the population. Other assumptions of the model are that capture occasions are instantaneous events, that no animals are killed in the capture process, that any emigration is permanent, and that individuals behave independently of one another (Williams, Nichols, and Conroy, 2002, p. 422).

In recent years, statisticians and ecologists have developed different extensions of the CJS model that allow the capture and survival rates to vary as functions of environmental and individual covariates. A complete review of these methods is given by Pollock (2002), though in general there are two competing strategies: the use of multistate models and the use of generalized linear models (GLM).

In a multistate model, the capture and survival rates are allowed to vary independently between animals in a finite number of states defined by one or more factors. The basic multistate model is the Arnason–Schwarz model (Arnason, 1973; Schwarz, Schweigert, and Arnason, 1993), which was originally developed to allow the rates to change by geographic location, though it may be used with any discrete covariate. The main advantage of a multistate model is that it can incorporate covariates that are both time dependent and unique to each individual. This is accomplished by using a Markov chain to describe the movement of individuals between states as well as to account for unobserved values of the covariate in the model likelihood. A disadvantage is that multistate models cannot incorporate continuous covariates. While some have used the multistate approach simply by categorizing continuous covariates (e.g., Nichols et al., 1992), this may lead to a loss of information that obscures the underlying relationship between the covariate and survival.

In the GLM-based approach, the survival and capture rates depend on linear combinations of the covariates through prespecified link functions. This strategy is described by Lebreton et al. (1992) who stress the logistic link because it satisfies the constraint that the estimated probabilities lie between 0 and 1. The clear advantage of this model over the multistate model is that both continuous and discrete covariates may be accounted for in the linear predictor. Unfortunately, current models cannot include time-dependent covariates because of

the difficulties that arise when an individual is not captured and the value of the covariate is not observed. To include such covariates in the GLM approach would require some way of modeling the distribution of the missing values of the covariate on each capture occasion (Pollock, 2002). This is exactly what we address.

In this article, we introduce a stochastic model to describe changes in a continuous covariate over time that is based on the Wiener process. This model is then used to account for the unobserved covariate values in the model likelihood. We use logistic functions to incorporate information from the covariate into the CJS model and a Bayesian approach based on the Metropolis–Hastings (MH) algorithm to estimate the model parameters. As an example, the model is applied to study the effect of body mass on the survival of the North American meadow vole, *Microtus pennsylvanicus*.

2. Methods

2.1 *Basic Notation*

Study parameters

k = number of capture occasions
t_s = time of capture occasion s, $s = 1, \ldots, k$
$\Delta_s = t_{s+1} - t_s$ time between capture occasions s and $s+1$
n = number of animals captured during the experiment

Observed data

$$\omega_{is} = \begin{cases} 1 \text{ if individual } i \text{ is captured at occasion } s \\ 0 \text{ otherwise} \end{cases}$$

$\omega_i = (\omega_{i1}, \ldots, \omega_{ik})$ capture history for individual i
z_{is} = covariate for individual i at time s (note that z_{is} is missing if $\omega_{is} = 0$)

Summary variables

a_i = first occasion that individual i is captured
d_i = last occasion that individual i is part of the study population (this is unobserved but known to be not less than the last time individual i was captured)

CJS model parameters

$p_{is} = p(z_{is})$ probability that animal i with covariate z_{is} alive at capture occasion s is captured
$\phi_{is} = \phi(z_{is})$ probability that animal i with covariate z_{is} in the study population at occasion s is still part of the population at occasion $s+1$

2.2 *Basic Cormack–Jolly–Seber Model*

The original CJS model is based on the assumption that the capture and survival rates at each capture occasion are the same for all animals. That is,

$$\phi_{is} = \phi_s \quad \text{and} \quad p_{is} = p_s.$$

Because it is impossible to know exactly when an animal entered the population, either by birth or immigration, the model assigns probabilities to each capture occasion conditional upon the animal's first release. For example, the probability that an animal is captured at times 2 and 4 in a study with $k = 5$ (i.e., $\omega_i = (0, 1, 0, 1, 0)$) is modeled as

$$\phi_2(1 - p_3)\phi_3 p_4(\phi_4(1 - p_5) + (1 - \phi_4)).$$

The likelihood is then produced by multiplying these contributions for each animal. Estimates of the survival and capture rates may then be produced using a variety of methods including computing maximum likelihood estimates analytically (Williams et al., 2002, p. 424) or by iterative methods like the expectation-maximization (EM) algorithm (Van Deusen, 2002), and through Bayesian inference (Poole, 2002).

2.3 *Modeling Continuous Covariates*

To include the effects of a time-dependent covariate, it is necessary to develop a model that describes the distribution of the covariate when an animal is not captured and the covariate is not observed. The motivating idea behind our model is that animals living in the same area should react in a similar manner to changes in the environment. Thus, differences in the value of the covariate between consecutive capture occasions should be similar for all individuals in the population. As an illustration, consider the body masses of animals forming a population living in a single region. In times when food is plentiful there will be little competition between animals and we would expect them all to gain mass. When food is scarce and competition increases, we would expect them to lose mass. Of course, there will be some individual variation due to unmeasured factors and random chance.

Our model begins with three assumptions about the behavior of the covariate over time:

1. Differences in the value of the covariate between any two times are normally distributed across the population.
2. The mean rate of change between capture occasions s and $s+1$ is the same for all animals, denoted by μ_s.
3. The rate of variance in the process is a constant over all time, denoted by σ^2.

In continuous time these conditions are satisfied by a Wiener process with time-dependent drift, $\mu(t)$, such that $\int_{t_s}^{t_{s+1}} \mu(t)\,dt = \mu_s$. Denoting the process by $X(t)$, $t \geq 0$, it is defined by three properties:

(i) X(0) = 0
(ii) for any $t_1 < t_2$ between the first and last capture occasions $X(t_2) - X(t_1) \sim N(\int_{t_1}^{t_2} \mu(t)\,dt, \sigma^2(t_2 - t_1))$
(iii) for any $t_1 < t_2 < t_3 < t_4$ between the first and last capture occasions $X(t_2) - X(t_1)$ and $X(t_4) - X(t_3)$ are independent.

A full description of the theory of continuous-time stochastic processes including the Wiener process is given by Cox and Miller (1965, p. 203).

To describe the distribution of the covariate for the ith individual at the sth capture occasion we restrict the model to the discrete capture occasions and define a new stochastic process $Z_{is} = Z_{ia_i} + X(t_s), s = 1, \ldots, k$. To find the distribution of Z_{is} conditional on Z_{ia_i} for $s > a_i$, the term Z_{is} can be written as the sum of the initial value and the differences between successive capture occasions,

$$\begin{aligned} Z_{i,s+1} &= Z_{ia_i} + \sum_{r=a_i}^{s} (Z_{i,r+1} - Z_{ir}) \\ &= Z_{ia_i} + \sum_{r=a_i}^{s} (X(t_{r+1}) - X(t_r)). \end{aligned}$$

The properties above then imply that the $Z_{i,s+1}$ given Z_{ia_i} is the sum of independent normal terms, and further that

$Z_{i,s+1}|Z_{ia_i},\ldots,Z_{i1} \sim Z_{i,s+1}|Z_{is}$. Thus, $Z_{i,a_i+1},\ldots,Z_{ik}$ form a Markov chain with transition kernel:

$$Z_{i,s+1} \,|\, Z_{is} = z_{is} \sim N\big(z_{is} + \mu_s\Delta_s, \sigma^2\Delta_s\big). \tag{1}$$

In essence, given one value of the covariate, the next value is normally distributed with mean proportional to the rate of drift between the capture occasions and the difference in time, and variance proportional to the difference in time. The density of the distribution in equation (1) will be denoted by $f(\cdot\,|\,\cdot)$.

The final concerns in the model are the links between the covariate values and the capture and survival probabilities. While many standard link functions may be used, we have chosen to use logistic functions. To accommodate variations in Δ_s, we model the survival probability per unit time assuming that survival between capture occasions s and $s+1$ depends only on the value z_{is}. Survival and capture probabilities as functions of the covariate are

$$\phi_{is} = \left[\frac{e^{\beta_0+\beta_1 z_{is}}}{1+e^{\beta_0+\beta_1 z_{is}}}\right]^{\Delta_s}$$

$$p_{is} = \frac{e^{\gamma_0+\gamma_1 z_{is}}}{1+e^{\gamma_0+\gamma_1 z_{is}}}. \tag{2}$$

The parameters $\beta = (\beta_0,\ \beta_1)$ and $\gamma = (\gamma_0,\ \gamma_1)$ govern the behavior of the survival and capture curves and are of primary interest in understanding the population dynamics. In particular, if $\beta_1 \neq 0$ then the survival rate is dependent on the covariate and if $\gamma_1 \neq 0$ then the capture rate is dependent on the covariate.

2.4 *Parameter Estimation*

Parameter estimation for this model is complicated by both missing predictors (the unobserved covariates) and missing response values (the unknown survival information). This distinguishes the problem from similar work like that of Ibrahim, Chen, and Lipsitz (1999) who describe a Monte Carlo EM algorithm for estimating parameters in general regression models with missing covariates. To obtain parameter estimates in spite of the missing data, we perform Bayesian inference based on the componentwise MH algorithm.

For simplicity, we assume equally spaced capture occasions such that $\Delta_s = 1$ for all s. In this case, the distribution in equation (1) simplifies

$$Z_{i,s+1} \,|\, Z_{is} = z_s \sim N\big(z_{is} + \mu_s, \sigma^2\big),$$

where μ_s and σ^2 now represent the mean and variance of the difference in the covariate per capture occasion. The survival probability, equation (2), becomes

$$\phi_{is} = \frac{e^{\beta_0+\beta_1 z_{is}}}{1+e^{\beta_0+\beta_1 z_{is}}},$$

and can be interpreted as the probability of survival from one occasion to the next. In the general case where Δ_s varies the mechanics of parameter estimation are the same, but the calculations required in the MH algorithm are slightly more complex. Also, when the Δ_s is not constant, the drift means, drift variance, and survival probabilities should be interpreted as effects per unit time and not per capture occasion.

We begin parameter estimation by defining the complete data likelihood (CDL) for the model. The CDL is formed as if all z_{is} and d_i were observed, thus removing the integrals needed to account for all possible values of the unobserved covariates in the usual likelihood. Letting $\mathbf{x}$ denote the completed data and Θ the entire set of parameters, the CDL is

$$L(\Theta\,|\,\mathbf{x}) \propto \prod_{i=1}^{n}\left[\left(\prod_{s=a_i+1}^{k} f(z_{is}\,|\,z_{i,s-1})\right)\left(\prod_{s=a_i}^{d_i-1}\phi_{is}\right)\cdot(1-\phi_{id_i})^{I[d_i<k]}\left(\prod_{s=a_i+1}^{d_i} p_{is}^{\omega_{is}}\big(1-p_{is}^{\omega_{is}}\big)\right)\right]. \tag{3}$$

Note that this expression is essentially the same as the complete data likelihood used in the EM algorithm described by Van Deusen (2002). The sole difference is that the first product in (3) includes the conditional density of z_{is} for all $s > a_i$, while the corresponding product in Van Deusen (2002) includes only terms up to d_i. It may seem unusual to include covariates at times after d_i, but extending the sum simplifies the generation of d_i on each MH iteration and does not affect the final estimates.

To construct the posterior distribution, one must also specify the prior distribution of the parameters. In the example that follows, the prior distributions for each parameter were assumed to be independent and were chosen to simplify the calculations. More descriptive priors should be used when information is available. For the drift means and variance, we used conjugate prior distributions that are the normal and the inverse gamma, respectively. Improper flat priors were used for γ and β. Assuming constant mean and variance for the priors of each μ_s, denoted by μ_0 and σ_μ^2, and denoting the shape and scale parameters of the prior for σ^2 as α_σ, β_σ, the prior density for Θ is

$$\pi(\Theta) = \exp\left\{-\frac{\sum_{s=1}^{k-1}(\mu_s-\mu_0)^2}{2\sigma_\mu^2} - \frac{1}{\beta_\sigma\sigma^2}\right\}\cdot\left(\frac{1}{\sigma^2}\right)^{\alpha_\sigma-1}. \tag{4}$$

The posterior density is proportional to the product of the expressions in (3) and (4).

Because of the complexity of the posterior distribution, estimates of the parameters were generated through the componentwise MH algorithm (Chib and Greenberg, 1995). Details of the algorithm used in the example are provided in the Appendix.

3. Example

The meadow vole or field mouse, *Microtus pennsylvanicus*, is a small rodent that forms colonies in grasslands, wet meadows, fields, and swamps throughout Canada and the northern United States. Adult voles measure between 14.0 and 19.5 cm and weigh between 20 and 70 g (Whitaker, 1997, p. 640), though there is extensive variation in size across the voles' range (Banfield, 1974, p. 209). The meadow vole breeds from early spring to late fall, occasionally even throughout the winter, and is the most prolific North American mammal (Hamilton and Whitaker, 1979, p. 218). The maximum lifespan of the meadow vole is thought to be 16 months, but

juvenile mortality is high and the vast majority of individuals do not reach maturity (Banfield, 1974, p. 210).

Previous studies of meadow vole weight dynamics have reached different conclusions about the relationship between mass and survival. An early study by Hamilton (1941) showed that the mean mass of voles declined during the winter, which he attributed to selective mortality of larger voles during the fall. More recent studies of voles in Minnesota and Manitoba found similar seasonal trends in the distribution of body mass, but suggested that this was due to decreases in individual mass and immigration of smaller voles to the study area during the winter months (Brown, 1973; Iverson and Turner, 1974). Iverson and Turner (1974) further conclude that mortality may be higher among smaller voles at certain times of the year.

Data for our analysis were gathered at the Patuxent Wildlife Center in Laurel, Maryland, from fall 1981 through spring 1982. The voles were captured and weighed on four primary occasions consisting of approximately 5 days and separated by about 1 month (Nichols et al., 1992). If an individual was captured more than once during a primary period, only the first weight was recorded. Several analyses of these data using multistate models were described by Nichols et al. (1992). Here we study the effect of body mass on survival by fitting the model in Section 2 to a subset of the data and provide results of a new multistate model for comparison.

The data collected at Patuxent contained records for 515 voles, though less than half were used in our analysis. First, only 215 voles were observed prior to the fourth occasion and contributed to the model likelihood in equation (3). Further, one of the key assumptions in our model is that the changes in the covariate between successive capture occasions have the same distribution for all animals. Because this is unlikely to be satisfied when considering changes in body mass for both immature and mature animals, we removed all observations for immature voles, defined as those weighing 22 g or less (Nichols et al., 1992). It was possible to remove only the observations where an animal's mass was below 22 g, rather than excluding the entire record, because no vole was observed with a mass greater than 22 g and subsequently less than or equal to 22 g. One other individual with a reported mass of 0 g on the final capture occasion was removed entirely. The final data set contained a total of 450 captures for 199 voles.

To compute parameter estimates, the MH algorithm ran for 1,000,000 iterations and the final 200,000 were retained for calculating point estimates and credible intervals. Values of the hyper-parameters in (4) used in this application were $\mu_0 = 0$, $\sigma_\mu^2 = 100$, $\alpha_\sigma = 0.001$, and $\beta_\sigma = 1000$. Convergence of the algorithm was assessed by repeating this procedure ten times starting with a wide range of initial values and comparing the chains according to the burn-in period and final parameter estimates (Carlin and Louis, 1996, p. 196). Point estimates and credible intervals computed from all ten chains were similar for all parameters, and trace plots of the simulated parameter values along with the associated Gelman and Rubin diagnostics suggested that any effects of the initial values were lost within the first 5000 iterations.

Estimates computed from the first MH run are presented in Table 1. The primary conclusion is that the data showed no significant effect of body mass on either the capture or survival rates. However, the results did show a significant increase of body mass of 1.80 g on average between capture occasions 3 and 4. As described in Nichols et al. (1992), the final capture event combines several capture occasions through the late winter and spring of 1982. These results suggest that the body mass of adult meadow voles at Patuxent remained constant through the fall of 1981 and early winter of 1982, and then increased in the early spring. Note, however, the relatively large variance, $\hat{\sigma}^2 = 26.6$ g^2, which suggests that this behavior varied considerably between individual voles. Plots of the estimated capture and survival functions with 95% credible envelopes are shown in Figure 1. These suggest that values near $\phi_{it} = 0.8$ and $p_{it} = 0.9$ can be fit for all individuals.

Table 1
Parameter estimates and 95% credible intervals for the parameters of the model of meadow vole survival generated using continuous body mass as a covariate

Parameter	Estimate
μ_1	−0.56 (−1.90, 0.79)
μ_2	0.11 (−1.00, 1.22)
μ_3	1.80 (0.82, 2.77)
σ^2	26.61 (22.31, 31.84)
β_1	1.39 (0.30, 2.57)
β_2	0.00 (−0.03, 0.02)
γ_1	2.02 (−0.97, 5.63)
γ_2	0.01 (−0.07, 0.09)

For comparison, we also constructed a multistate model of the data. Following Nichols et al. (1992) the adult voles were divided into three distinct mass categories: (B) 22–33 g, (C) 34–45 g, and (D) >45 g (in Nichols et al., 1992, class A refers to the immature animals that were excluded from this

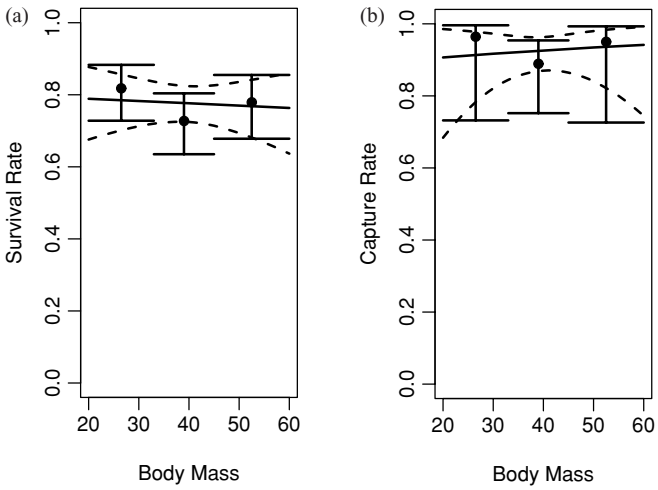

Figure 1. Estimated survival (a) and capture (b) rates as functions of body mass (grams) for the meadow vole, *Microtus pennsylvanicus*. For the model using continuous body mass the estimated rates are indicated by the solid lines, and the 95% credible intervals by the broken lines. Points represent the estimates from the multistate model with 95% confidence intervals for each of the three mass classes.

analysis). To match the assumptions of the proposed continuous covariate model, the parameters controlling transitions between the classes were allowed to vary over time while the capture and survival rates for each class were held constant. Several transitions were not observed in the data set and in these cases the associated parameter was fixed to 0 to generate an identifiable model. Estimates of the remaining 17 parameters (Table 2) were computed using program MARK (White and Burnham, 1999).

Estimates of the capture and survival rates for the multistate model match very well with the results of the previous model (Figure 1). Once again, there is no evidence of an effect of body mass on either the capture or survival rates. Further, for both the survival and capture rates the 95% confidence intervals produced from the multistate model overlap the 95% credible envelopes from the previous model over the entire range of body mass.

The results from the proposed model also agree with previously published studies of the body mass and survival of *Microtus pennsylvanicus*. Estimates of the survival probabilities for the original Patuxent data set are not provided in Nichols et al. (1992), but can be derived from results given. For the model assuming equal transition rates over time (Model II), the estimated survival probabilities are near 0.8 for every weight class and do not significantly differ from each other. This is consistent with our findings. Moreover, the body mass dynamics suggested by our model match well with the conclusions of Brown (1973) who describes a slight decrease in voles' weight in the fall, steady weight through the winter, and an increase in the spring. Plots from Iverson and Turner (1974) suggest a similar pattern, though the observed changes are much more dramatic.

Table 2

Maximum likelihood estimates and 95% confidence intervals of the parameters for the multistate model of meadow vole survival. Parameters ϕ_j and p_j denote the survival and capture rates for weight class j; r_{jlt} denotes the transition rate from state j to l at time t. Rates corresponding to transitions that were not observed in the data are not included.

Parameter	Estimate
ϕ_B	0.82 (0.73, 0.88)
ϕ_C	0.73 (0.64, 0.80)
ϕ_D	0.78 (0.68, 0.86)
p_B	0.96 (0.73, 1.00)
p_C	0.89 (0.75, 0.95)
p_D	0.95 (0.73, 0.99)
r_{BC1}	0.85 (0.40, 0.98)
r_{BC2}	0.21 (0.10, 0.40)
r_{BC3}	0.33 (0.22, 0.46)
r_{BD3}	0.04 (0.01, 0.13)
r_{CB3}	0.17 (0.08, 0.33)
r_{CD1}	0.13 (0.04, 0.37)
r_{CD2}	0.09 (0.03, 0.27)
r_{CD3}	0.10 (0.04, 0.26)
r_{DC1}	0.19 (0.08, 0.39)
r_{DC2}	0.37 (0.21, 0.57)
r_{DC3}	0.28 (0.11, 0.55)

However, the new method does have clear advantages over each of the other approaches. Though the studies of Brown (1973) and Iverson and Turner (1974) are both straightforward capture–recapture experiments, neither uses capture–recapture methodology and this may have systematically biased their results. Instead, Brown (1973) subjectively combines observations from different animals observed on different occasions to create a "composite picture of growth" while Iverson and Turner (1974) draw conclusions by comparing the mean body mass of animals captured on each occasion. Both studies ignore possible effects of the capture mechanism. The primary advantage of our model over the multistate model is the inclusion of time-dependent, continuous covariates. This is discussed further in Section 4. Another drawback of the multistate model is that it may be difficult to draw conclusions about the changes in a covariate based on the estimated transition rates for different states. For example, in the multistate model we fit to the meadow vole data only two of the transition rates show significant changes over the four occasions (Table 2). Both of these represent transitions that occur only between the third and fourth capture occasion, though one is an increase in mass (from state B to D) while the other is a decrease in mass (from state C to B). This makes it hard to derive conclusions about general trends in the voles' body mass. Moreover, the standard errors of both estimates are very high because very few individuals are observed making these specific transitions ($n = 6$ and $n = 2$, respectively). In contrast, our model provides an exact distribution for the difference in the covariate on each capture occasion.

4. Discussion

This article introduces a new method that extends the CJS model to incorporate the effects of covariates that are continuous, individual, and time dependent. This is accomplished using a Markov chain to model difference in the covariate over time, similar to the way the multistate model describes movements of animals between discrete states. However, the two models make very different assumptions regarding both the behavior of the covariate and its relationship to the capture and survival rates. We conclude by comparing these assumptions, and by commenting on the limitations of the proposed model and suggesting ways these might be avoided in future applications.

One of the main advantages of the multistate model is that it places no constraints on the model parameters. That is, the capture, survival, and transition rates for animals in each state vary independently of the rates for animals in the other states. For researchers who are studying systems about which little is known or who stress the need to remain objective, this may seem very attractive. The major disadvantage is that the multistate model makes the implicit assumption that animals behave exactly alike if they belong to the same state and not alike if they belong to different states. For example, the multistate model used in the analysis of Section 3 implies that if a vole weighs 44 g then (a) it behaves exactly like all others weighing between 35 and 44 g at that time, and (b) no information about the individual can be drawn from the voles weighing 45 g or more. Discretizing a continuous variable can also lead to serious problems with model fit. If the variable

is divided into too few discrete states then the assumption that animals in the same state behave alike may not be appropriate. This will increase the variability of the parameter estimates and may hide the true relationships in the system. On the other hand, the number of parameters grows exponentially as the number of states increases and can quickly exhaust the information in the data. A model with a large number of states may contain unidentifiable parameters or be too complex to interpret. In Section 3, the multistate model required 11 parameters to describe changes in the voles' body masses over time while the proposed model required only 4.

In contrast to the multistate model, the proposed model for continuous covariates imposes constraints on the distribution of the covariate and on the functional form of the relationships between the covariate and the capture and survival rates, but allows the capture and survival rates to change continuously. The first assumption of the model presented in Section 2 is that the change in the covariate between two capture occasions is normally distributed with constant variance and the same mean for all individuals. This is motivated by the idea that all animals in a population will react in a similar manner to environmental pressures. If there is reason to believe that different subpopulations behave differently, as in the expected difference between juveniles and adults in Section 3, the model might be extended by allowing different means and/or variances for different subpopulations. The simplest extension would model two distinct means, $\mu_s = \mu_{s0} + \mu_{s1}\delta$, where δ is an indicator variable that distinguishes two subgroups. More complicated models might use many different means, if the covariate depended on location for example, or even a continuous function, $\mu_s = g_s(\cdot)$, where $g_s(\cdot)$ may depend on the covariate itself or on other auxiliary variables. Different distributions could also be used in cases where normality cannot be assumed and biological considerations suggest another form.

The second assumption of our model is that the relationships between the survival and capture rates and the covariate can be described by simple link functions that are constant over time. We have used the logistic link in our work so far because of its convenience in modeling probabilities, but this choice further imposes that the relationships must be monotonic. In future applications, we hope to explore the possibility of using splines or other piecewise functions to allow more flexibility in the relationships. The assumption of time homogeneity is also likely to be restrictive in many applications. Indeed, in their analysis of the Patuxent meadow vole data, Nichols et al. (1992) show that a multistate model with time-specific parameters provides significant improvement over a model with fixed parameters. Complete flexibility over time could be achieved in the proposed model using separate capture and survival functions on each capture occasion, though this would introduce a large number of new parameters. Instead, we propose a solution based on the proportional hazards model used in survival analysis (Cox, 1972). This model allows survival rates to change over time under the assumption that the relative effect of different values of the covariate is constant and introduces only one extra parameter for each capture occasion. We anticipate that both survival and capture rates could be modeled in a similar manner.

An important practical difference between the multistate model and the proposed model is that estimation becomes much more difficult when considering a continuous covariate. The likelihood for the proposed model requires integrals to account for every missing covariate value, and maximum likelihood estimates cannot be found analytically. While several classical methods exist for estimating parameters in problems with large amounts of missing data, e.g., the EM algorithm and its derivatives, we recommend Bayesian estimation because of its natural view of missing data. Bayesian inference does not differentiate between missing data and parameters, both of which are considered unobserved random variables, and so both can be handled in the same framework. This is very appealing for problems with large amounts of missing data. Other researchers have already begun promoting the use of Bayesian estimation for ecological statistics and capture–recapture experiments (see e.g., Dupuis, 1995; Brooks, Catchpole, and Morgan, 2000; Poole, 2002).

We believe that the model presented here provides an important alternative to other models that impose constraints upon either the time dependence, continuity, or individuality of the covariate. We hope that other researchers will apply the proposed model in their own research and work to develop their own variations. To assist this, the source code from `C` programs implementing the MH algorithm and the complete data set for the meadow vole example are available at the *Biometrics* website: `http://www.tibs.org`.

Acknowledgements

This work was completed during Simon Bonner's M.Sc. which was partially funded by a Postgraduate Scholarship Type A from the National Science and Engineering Research Council, and by a Graduate Fellowship award from Simon Fraser University. Data for the study of the meadow vole were generously provided by Dr J. Nichols of the United States Fish and Wildlife Service. Our sincere thanks go to Laura Cowen, Alyssa MacLean, Dr Tim Swartz, and the two referees for their constructive comments on previous drafts and their help during the review process. Thanks also to Dr K. Larry Weldon who provided advice as part of Simon Bonner's M.Sc. committee.

References

Arnason, A. N. (1973). The estimation of population size, migration rates, and survival in a stratified population. *Research in Population Ecology* **15,** 1–8.

Banfield, A. W. F. (1974). *The Mammals of Canada*. Toronto: University of Toronto Press.

Brooks, S. P., Catchpole, E. A., and Morgan, B. J. T. (2000). Bayesian animal survival estimation. *Statistical Science* **15,** 357–376.

Brown, E. B. (1973). Changes in patterns of seasonal growth of *Microtus pennsylvanicus*. *Ecology* **54,** 1103–1110.

Carlin, B. P. and Louis, T. A. (1996). *Bayes and Empirical Bayes Methods for Data Analysis*. London: Chapman and Hall.

Chib, S. and Greenberg, E. (1995). Understanding the Metropolis-Hastings algorithm. *American Statistician* **49,** 327–335.

Cormack, R. M. (1964). Estimates of survival from the sighting of marked animals. *Biometrika* **51,** 429–438.

Cox, D. R. (1972). Regression models and life-tables. *Journal of the Royal Statistical Society* **34,** 187–220.

Cox, D. R. and Miller, H. D. (1965). *The Theory of Stochastic Processes*. London: Spottiswoode Ballantyne.

Dupuis, J. A. (1995). Bayesian estimation of movement and survival probabilities from capture–recapture data. *Biometrika* **82,** 761–772.

Gilks, W. R., Richardson, S., and Spiegelhalter, D. J. (1996). Introducing Markov chain Monte Carlo. In *Markov Chain Monte Carlo in Practice*, W. R. Gilks, S. Richardson, and D. J. Spiegelhalter (eds), 1–19. New York: Chapman and Hall.

Hamilton, W. J., Jr. (1941). Reproduction of the field mouse *Microtus pennsylvanicus*. Memorandum, Cornell University Agricultural Experiment Station.

Hamilton, W. J., Jr. and Whitaker, J. O., Jr. (1979). *Mammals of the Eastern United States*, 2nd edition. New York: Cornell University Press.

Ibrahim, J. G., Chen, M.-H., and Lipsitz, S. R. (1999). Monte Carlo EM for missing covariates in parametric regression models. *Biometrics* **55,** 591–596.

Iverson, S. L. and Turner, B. N. (1974). Winter weight dynamics in *Microtus pennsylvanicus*. *Ecology* **55,** 1030–1041.

Jolly, G. M. (1965). Explicit estimates from capture–recapture data with both death and immigration-stochastic model. *Biometrika* **52,** 225–247.

Lebreton, J.-D., Burnham, K. P., Clobert, J., and Anderson, D. R. (1992). Modelling survival and testing biological hypotheses using marked animals: A unified approach with case studies. *Ecological Monographs* **62,** 67–118.

Nichols, J. D., Sauer, J. R., Pollock, K. H., and Hestbeck, J. B. (1992). Estimating transition probabilities for stage-based population projection matrices using capture-recapture data. *Ecology* **73,** 306–312.

Pollock, K. H. (2002). The use of auxiliary variables in capture-recapture modelling: An overview. *Journal of Applied Statistics* **29,** 85–102.

Poole, D. (2002). Bayesian estimation of survival from mark-recapture data. *Journal of Agricultural, Biological, and Environmental Statistics* **7,** 264–276.

Schwarz, C. J., Schweigert, J. F., and Arnason, A. N. (1993). Estimating migration rates using tag-recovery data. *Biometrics* **49,** 177–193.

Seber, G. A. F. (1965). A note on the multiple recapture census. *Biometrika* **52,** 249–259.

Van Deusen, P. C. (2002). An EM-algorithm for capture-recapture estimation. *Environmental and Ecological Statistics* **9,** 151–165.

Whitaker,, J. O., Jr. (1997). *National Audobon Society Field Guide to North American Mammals*. New York: Alfred A. Knopf.

White, G. C. and Burnham, K. P. (1999). Program MARK-survival estimation from populations of marked animals. *Bird Study* **46**(suppl.), 120–138.

Williams, B. K., Nichols, J. D., and Conroy, M. J. (2002). *Analysis and Management of Animal Populations*. San Diego, California: Academic Press.

Received September 2003. *Revised February* 2005. *Accepted March* 2005.

Appendix

Details of the Metropolis–Hastings Algorithm

The MH algorithm generates a Markov chain by iteratively simulating values for the unknown quantities from a prespecified proposal distribution. These values are accepted with a certain probability and become the next state in the chain, or else they are rejected and the state remains the same. The acceptance probability for a simulated value depends on the likelihood function, the proposal distribution, and the chain's current state. Under general conditions, the distribution of the simulated values converges to the joint posterior distribution of the parameters and missing data, such that values from the tail of the chain can be treated as a sample from the posterior (Chib and Greenberg, 1995). The componentwise (or single-component) MH algorithm is a variation in which the unknown quantities are divided into low dimensional sets, and each set is updated in sequence at every iteration of the chain. This method greatly simplifies the computation when there are a large number of unknown parameters or missing data points. Values from the resulting chain still converge in distribution to the posterior (Gilks, Richardson, and Spiegelhalter, 1996).

The full conditional distributions for the parameters in the model of Section 2 can be derived from the product of equations (3) and (4). The proposal distributions used in the example are given below. In the proposal distribution for the unknown quantity θ, the quantity's current value is denoted as θ and the proposed value as θ'.

1. Missing covariates (z_{is})

$$z'_{is} \,|\, \mathbf{x}, \Theta \sim \begin{cases} \mathrm{N}\left(\dfrac{(z_{i,s-1}+\mu_{s-1})+(z_{i,s+1}-\mu_s)}{2}, \dfrac{\sigma^2}{2}\right) & s < k \\ \mathrm{N}\left(z_{i,s-1}+\mu_{s-1}, \sigma^2\right) & s = k \end{cases}$$

2. Missing survival information (d_i)

$$P(d'_i = s \,|\, \mathbf{x}, \Theta) \propto \begin{cases} 0 & s < b \\ 1-\phi_{ib} & s = b \\ \displaystyle\prod_{r=b+1}^{s} [\phi_{i,r-1}(1-p_{ir})] \cdot (1-\phi_{is})^{I[s<k]} & b < s \le k \end{cases}$$

3. Drift means (μ_s)

$$\mu'_s \,|\, \mathbf{x}, \Theta \sim \mathrm{N}\left(\frac{\displaystyle\sum_{i=1}^{n} I[a_i \ge s](z_{i,s+1}-z_{is}) + \frac{\sigma^2}{\sigma_\mu^2}\mu_{s0}}{\displaystyle\sum_{i=1}^{n} I[a_i \ge s] + \frac{\sigma^2}{\sigma_\mu^2}}, \frac{\sigma^2}{\displaystyle\sum_{i=1}^{n} I[a_i \ge s] + \frac{\sigma^2}{\sigma_\mu^2}} \right)$$

4. Drift variance (σ^2)

$$\sigma^{2\prime} \,|\, \mathbf{x}, \Theta \sim \mathrm{IG}\left(\frac{\sum_{i=1}^{n}(k-a_i)}{2} + \alpha_\sigma,\right.$$

$$\left.\left(\frac{1}{\beta_\sigma} + \frac{\sum_{i=1}^{n}\sum_{r=a_i}^{k-1}(z_{i,r+1} - z_{ir} - \mu_r)^2}{2}\right)^{-1}\right)$$

5. Coefficients of the survival function (β)
   Both elements sampled simultaneously from a bivariate normal distribution

$$\beta' \,|\, \beta \sim \mathrm{N}\big(\beta, I_\beta(\beta, \mathbf{x})^{-1}\big),$$

   where $I_\beta(\beta,\ \mathbf{x})$ is the observed information matrix for β computed given the current parameters and complete data.
6. Coefficients of the capture function (γ)
   Both elements sampled simultaneously from a bivariate normal distribution

$$\gamma' \,|\, \gamma \sim \mathrm{N}\big(\gamma, I_\gamma(\gamma, \mathbf{x})^{-1}\big),$$

   where $I_\gamma(\gamma,\ \mathbf{x})$ is the observed information matrix for γ computed given the current parameters and complete data.