

Modeling Individual Effects in the Cormack–Jolly–Seber Model: A State–Space Formulation

J. Andrew Royle

USGS Patuxent Wildlife Research Center, 12100 Beech Forest Road,
Laurel, Maryland 20708, U.S.A.
email: aroyle@usgs.gov

SUMMARY. In population and evolutionary biology, there exists considerable interest in individual heterogeneity in parameters of demographic models for open populations. However, flexible and practical solutions to the development of such models have proven to be elusive. In this article, I provide a state-space formulation of open population capture–recapture models with individual effects. The state-space formulation provides a generic and flexible framework for modeling and inference in models with individual effects, and it yields a practical means of estimation in these complex problems via contemporary methods of Markov chain Monte Carlo. A straightforward implementation can be achieved in the software package **WinBUGS**. I provide an analysis of a simple model with constant parameter detection and survival probability parameters. A second example is based on data from a 7-year study of European dippers, in which a model with year and individual effects is fitted.

KEY WORDS: Animal demography; Capture–recapture; Cormack–Jolly–Seber; CJS; Survival estimation; WinBUGS.

1. Introduction

The use of data from marked individuals in studies of animal populations is widespread in contemporary ecology. The so-called “Cormack–Jolly–Seber” (CJS) models (Lebreton et al., 1992; Brooks, Catchpole, and Morgan, 2000) constitute one of the most popular classes of models in which primary interest is focused on obtaining information about animal survival in the presence of imperfect detection of individuals.

Models for animal demography that contain individual effects are important in many situations. Two broad classes of such models exist: those in which heterogeneity is modeled as unstructured variation, due to an individual “random effect” (Norris and Pollock, 1996; Coull and Agresti, 1999; Cam et al., 2002; Dorazio and Royle, 2003; Royle, 2006) or models that seek to describe heterogeneity explicitly, by the use of individual covariates (Pollock, Hines, and Nichols, 1984; Huggins, 1989; Alho, 1990). Methods of inference for both classes of problems have, individually, received considerable attention, most of this focused on the case of closed populations. The development of general and practical solutions for modeling individual heterogeneity in open populations has proved challenging, and there is a dearth of published research on methods of modeling individual effects in such cases. Cam et al. (2002) discuss evolutionary interest in individual heterogeneity, and provide analysis of the special case in which individuals are detected with probability 1 (see also Link et al., 2002). Nichols et al. (1982) and Pollock and Raveling (1982) motivate interest in the problem of heterogeneity in survival in studies of North American waterfowl populations. In this context, Burnham and Rexstad (1993) developed a model of

heterogeneous survival rates by recognizing that heterogeneity induces a particular functional decline in band recovery probability, yielding a model that they fitted by nonlinear least squares. Pledger, Pollock, and Norris (2003) developed a CJS model containing heterogeneity in both detection probability and survival probability using finite-mixture models, and they adopt a framework for inference based on integrated likelihood.

In this article, I provide a state-space parameterization of CJS models with individual heterogeneity in which a model for the data is described explicitly by two distinct components: (1) a model for the unobserved, or partially observed, individual state process (the “process model”), and (2) a model for the observations, conditional on the state process (the “observation model”). Fundamental to this formulation is the definition of a set of individual state variables, say $z(i, t)$, which describe whether individual i is alive and available for capture at time t . In the classical CJS situation, the process model, that governing the dynamics of $z(i, t)$, is a simple survival process, whereas the observation model consists of a collection of independent Bernoulli trials. The hierarchical, state-space formulation has been used effectively in other animal demographic modeling contexts. Newman et al. (2006) provide a conceptually clear formulation of the state-space framework applied to models of animal populations. Dupuis (1995) gave a hierarchical formulation of so-called “multistate” models. Dorazio et al. (2006) used a state-space formulation to develop models of animal community structure (i.e., for estimating species richness and other summaries of biological communities). The state-space formulation of CJS

models is closely related to metapopulation occupancy models (Royle, Dorazio, and Link, 2007) for which it can be exploited to yield important extensions to the modeling of individual effects (Royle and Kery, 2007).

The state-space representation of the CJS model is described in Section 2. Estimation and inference using Markov chain Monte Carlo (MCMC) methods is discussed briefly in Section 2.1 for the case where individual heterogeneity exists in both p and ϕ . An application of the constant- p , constant- ϕ model is given in Section 3 and, in Section 4, the European dipper data from Lebreton et al. (1992) is analyzed using a model with individual heterogeneity in both p and ϕ . In both examples, the freely available software package WinBUGS is used for analysis.

2. State-Space Parameterization of the CJS Model

We suppose that a population is sampled on T occasions yielding capture histories on n unique individuals. The capture history for individual i , first captured at time f_i , is the vector $\{y(i, t)\}_{t=f_i}^T$. The strategy that I adopt here for developing a model is analogous to that found in classical state-space time-series models (Jones, 1993; Berliner, 1996) in which the model is specified by distinct models for the (typically unobserved or partially observed) process of interest and then for the observations conditional on the latent process. I will denote the state process here by the binary state variable $z(i, t)$ indicating whether individual i is alive at time t . In the present case, the $z(i, t)$ state variables are governed by a simple survival process (described subsequently). The second component model is a model for the binary observations $y(i, t)$ conditional on the latent state variables $z(i, t)$.

The State Model: Let $\{z(i, t); i = 1, \dots, n, t = 1, \dots, T\}$ be Bernoulli random variables describing whether individual i is alive ($z(i, t) = 1$) or dead ($z(i, t) = 0$) at time t . The CJS model is developed conditional on the time of first capture of each individual, f_i . That is, $z(i, f_i) = 1$ with probability 1. The survival process is given by the conditional model

$$z(i, t) | z(i, t-1) \sim \text{Bernoulli}(z(i, t-1)\phi_{t-1}) \quad (1)$$

for $t = f_i + 1, \dots, T$. In other words, if an individual is alive at time $t - 1$, its survival outcome is a Bernoulli random variable with parameter ϕ_t . If an individual is not alive at $t - 1$, then $z(i, t)$ is Bernoulli with success probability 0, that is, $z(i, t) = 0$ with probability 1.

The Observation Model: Conditional on the state process, $y(i, t)$ are independent Bernoulli random variables,

$$y(i, t) | z(i, t) \sim \text{Bernoulli}(p_t z(i, t)) \quad (2)$$

Thus, if $z(i, t) = 0$, then $y(i, t) = 0$ with probability 1, otherwise $y(i, t)$ is a Bernoulli trial with parameter p_t .

The models have been described here with time-varying parameters. Although this is a common model, so too are age-structured models, or models in which either p_t or ϕ_t vary in response to known covariates. These and other modifications are straightforward to construct and, using the MCMC methods described subsequently, equally practical to analyze.

2.1 Posterior Inference by Markov chain Monte Carlo

It is straightforward to develop a Gibbs sampling MCMC algorithm for the state-space representation of the model, either by explicit construction of the full conditional distributions (see Web Appendix A) or by exploiting the freely available software WinBUGS (Gilks, Thomas, and Spiegelhalter, 1994; or its variants). Indeed, one of the most practical benefits of the state-space representation of the CJS model is that it does have a fairly simple implementation in WinBUGS. Thus, explicit attention to the development of numerical methods for carrying out the analysis by MCMC in such complex models can be avoided. In the examples that follow (Sections 4 and 5), I use WinBUGS to conduct the analysis.

3. Example 1: A Small Data Set

Here I provide an analysis of a small data set consisting of eight individuals, captured on five sampling occasions. The data are given in Table 1.

These data were made up arbitrarily, using no formal rule, and there are no values of p and ϕ that correspond to “truth.” The data are used here only to provide an example of fitting the constant- p , constant- ϕ model using the state-space representation. The WinBUGS model specification for this model with uniform(0,1) priors for both p and ϕ is given in Figure 1. Implementation requires learning enough of WinBUGS to input

Table 1
A sample data set consisting of encounter histories for 8 individuals

$t = 1$	$t = 2$	$t = 3$	$t = 4$	$t = 5$
1	1	0	0	1
—	1	1	1	0
—	1	0	0	1
—	—	1	1	1
—	—	1	0	0
—	—	1	0	1
—	—	—	1	0
—	—	—	1	0

```
model {
  # Prior distributions for model parameters
  phi ~ dunif(0,1)
  p ~ dunif(0,1)

  for(i in 1:nind){
    z[i,first[i]]~dbern(1)
    for(j in (first[i]+1):nyear){
      mu1[i,j]<-p*z[i,j]
      x[i,j]~dbern(mu1[i,j])
      mu2[i,j]<-phi*z[i,j-1]
      z[i,j]~dbern(mu2[i,j])
    }
  }
}
```

Figure 1. WinBUGS model specification for the CJS model with constant p and ϕ .

Table 2

Posterior summary statistics from analysis of the small data set under a constant- p , constant- ϕ model

Parameter	Mean	SD	0.025	0.25	0.50	0.75	0.975
ϕ	0.903	0.082	0.69	0.861	0.924	0.968	0.997
p	0.481	0.118	0.26	0.398	0.478	0.560	0.718

the data array x , the vector “first,” and the fixed constants “nind” and “nyear”. The input for WinBUGS is given in Web Appendix B. Alternatively, users familiar with the free software R (Ihaka and Gentleman, 1996) can exploit the add-on library R2WinBUGS (Sturtz, Ligges, and Gelman, 2005), bypassing the WinBUGS GUI. In Figure 1, the line `z[i,first[i]] ~ dbern(1)` enforces that the first state of each individual is fixed, whereas the meaning of the remaining model specification should be fairly self-evident. In particular, note that uniform(0,1) priors are assumed for both ϕ and p , and the distributions of $z(i, t) | z(i, t-1)$ and $y(i, t) | z(i, t)$ are those described by equations (1) and (2), respectively.

Posterior summary statistics for these data are given in Table 2. These were based on 10,000 Monte Carlo draws from the posterior distribution after 2000 samples were discarded as “burn-in.”

4. Example 2: The European Dipper Data

In this example, I consider data from a 7-year study of European dippers (*Cinclus cinclus*).¹ The data are originally from Marzolin (1988), as described by Lebreton et al. (1992), and were also analyzed by Brooks et al. (2000). Here I will use the state-space formulation of the individual effects model to fit a model in which both survival and capture probability vary by year and individual. Using conventional shorthand for describing models, this model would be indicated as $\{p(t+h), \phi(t+h)\}$. The model assumed here is

$$u_{it} \equiv \text{logit}(p_{it}) = a_t + \alpha_i$$

and

$$v_{it} \equiv \text{logit}(\phi_{it}) = b_t + \beta_i,$$

where p_{it} is the probability of capture for individual i for year t and ϕ_{it} is the probability of survival of individual i over the interval $(t, t+1)$, a_t and b_t are the fixed yearly effects, and α_i and β_i are latent individual effects. The individual effects α_i and β_i are assumed to be mean zero random effects with variances σ_p^2 and σ_ϕ^2 , respectively. Note that with $T = 7$ samples, not all of the fixed effects are identifiable. In particular, in the classical formulation of the fixed effects model, only p_2, \dots, p_6 , ϕ_1, \dots, ϕ_5 and the product $\theta = p_7\phi_6$ are identifiable (typically, p_7 is set to 1 and the interpretation of ϕ_6 is so qualified, or vice versa). Thus, the corresponding logit-scale fixed effects are identifiable in the present construction.

Specification of prior distributions for the model parameters merits some discussion. Priors with compact support typically yield better-mixing Markov chains. Thus, for all parameters, I imposed proper uniform priors. I assumed uniform(0,1) priors on the inverse logit of the annual fixed effects:

$\text{logit}^{-1}(a_t) \sim \text{uniform}(0, 1)$, $\text{logit}^{-1}(b_t) \sim \text{uniform}(0, 1)$. Gelman (2006) recommends uniform(0,B) priors for the standard deviation parameters σ_p and σ_ϕ , and his discussion of the general issue of selecting priors for variance components is highly relevant to the present model. This is brought into context shortly with regard to the issue of prior sensitivity.

The WinBUGS model specification for the heterogeneity model is given in Web Appendix C. I also provide a modification for the case where individual covariates are thought to influence survival probability (although such covariates are not available in the present example). Posterior summaries were based on three Markov chains of length 60,000 after a 5000 sample burn-in. The estimated posterior distributions of model parameters are summarized in Table 3. For comparison, the maximum likelihood estimates (MLE) of the year effects in a model without heterogeneity are also given.² The results indicate considerable heterogeneity in p , and a lesser degree of heterogeneity in ϕ . The degree of heterogeneity is naturally by the densities of $p_i = \text{logit}^{-1}(a_0 + \alpha_i)$ and $\phi_i = \text{logit}^{-1}(b_0 + \beta_i)$ for some baseline a_0 and b_0 . Setting these at the posterior mean of a_1 and b_1 , these densities are shown in Figure 2. Ideally, these figures could be regarded as applying to the population of individuals that enter the adult age class. If p and ϕ are independent, then this should be the case for the description of heterogeneity in p_i . However, better-surviving individuals should have a higher probability of appearing in the sample as compared to the population of adults at large, and so the heterogeneity in ϕ_i may be biased. This is discussed in Section 5 (see also Web Appendix E).

The posterior means are broadly consistent with the MLEs computed under a model without heterogeneity. There are slight differences owing both to the fact that the posteriors are not symmetric (and hence posterior means should deviate from MLEs) and also that the effect of heterogeneity is to bias estimates of detection probability low. The effect on estimated detection probabilities is not so pronounced here, in part because the mean survival and detection probabilities are reasonably high, and thus mass of both densities is concentrated away from 0 (Figure 2). Posterior uncertainty is only moderately larger than classical asymptotic estimates under the model without individual effects.

4.1 Sensitivity to Prior Distributions

The use of ostensibly vague inverse gamma prior distributions for σ^2 (i.e., $\text{IG}(\epsilon, \epsilon)$ for small ϵ) is fairly conventional in Bayesian analysis of models containing random effects or other latent structure. Gelman (2006) considered the appropriateness of this prior distribution, noting that “...for datasets in which low values of σ are possible, inferences become very sensitive to ϵ in this model, and the prior distribution hardly looks noninformative...” (Gelman, 2006, p. 522). I carried out a similar analysis for the CJS model using various $\text{IG}(\epsilon, \epsilon)$ priors for σ_p^2 (the posterior of σ_ϕ appeared not to be so sensitive to the prior specification).

¹ These data can be found on E. G. Cooch’s website: <http://www.phidot.org/software/mark/docs/book/>.

² I thank Jim Hines for computing these in the software package MARK.

Table 3

Maximum likelihood estimates of year-effects model and posterior summaries of the model with year effects and individual heterogeneity fit to the European dipper data. q_m is the m th percentile of the posterior distribution.

Parameter	Bayesian				MLEs			
	Mean	SD	$q_{0.025}$	$q_{0.50}$	$q_{0.975}$	MLE	SE	95% CI
ϕ_1	0.748	0.134	0.476	0.754	0.977	0.718	0.156	(0.361,0.920)
ϕ_2	0.464	0.080	0.320	0.460	0.632	0.435	0.069	(0.308,0.571)
ϕ_3	0.494	0.070	0.362	0.491	0.639	0.478	0.060	(0.364,0.594)
ϕ_4	0.624	0.066	0.495	0.623	0.753	0.626	0.059	(0.505,0.733)
ϕ_5	0.599	0.064	0.473	0.599	0.725	0.598	0.056	(0.486,0.702)
p_1	0.671	0.155	0.344	0.684	0.929	0.696	0.166	(0.330,0.914)
p_2	0.879	0.087	0.659	0.899	0.987	0.923	0.073	(0.616,0.989)
p_3	0.878	0.076	0.691	0.893	0.977	0.913	0.058	(0.714,0.978)
p_4	0.900	0.058	0.761	0.911	0.980	0.901	0.054	(0.736,0.967)
p_5	0.922	0.052	0.791	0.933	0.989	0.932	0.046	(0.768,0.983)
σ_p	1.395	0.973	0.124	1.233	3.773			
σ_ϕ	0.352	0.232	0.036	0.314	0.887			

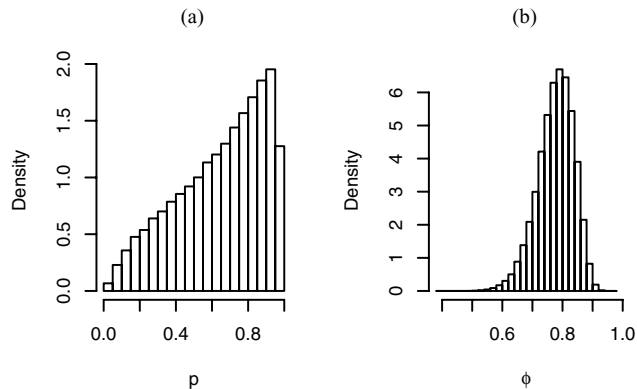


Figure 2. Estimated individual heterogeneity distributions of p (panel a) and ϕ (panel b) for the European dipper data.

In Figure 3, the three panels correspond to the estimated posterior under a uniform(0,8) prior (left panel) for σ_p , the estimated posterior under an IG(1,1) prior for σ_p^2 (middle panel), and under an IG(0.01, 0.01) prior for σ_p^2 (right panel). In all cases, the posterior is restricted to $[0,5]$, where most of the posterior mass occurs (the histograms were restandardized to sum to 1 over that range). These results are strikingly similar to those reported by Gelman (2006). In particular, note that the uniform prior does not affect the posterior so much, whereas for the conventional vague conditionally conjugate prior, the posterior does appear to be influenced by the choice of ϵ . As summarized by Gelman (2006), “... the inverse-gamma(ϵ , ϵ) prior is not at all ‘noninformative’ for this problem since the resulting posterior distribution remains highly sensitive to the choice of ϵ .” In the present case, support for the proper uniform prior seems justified.

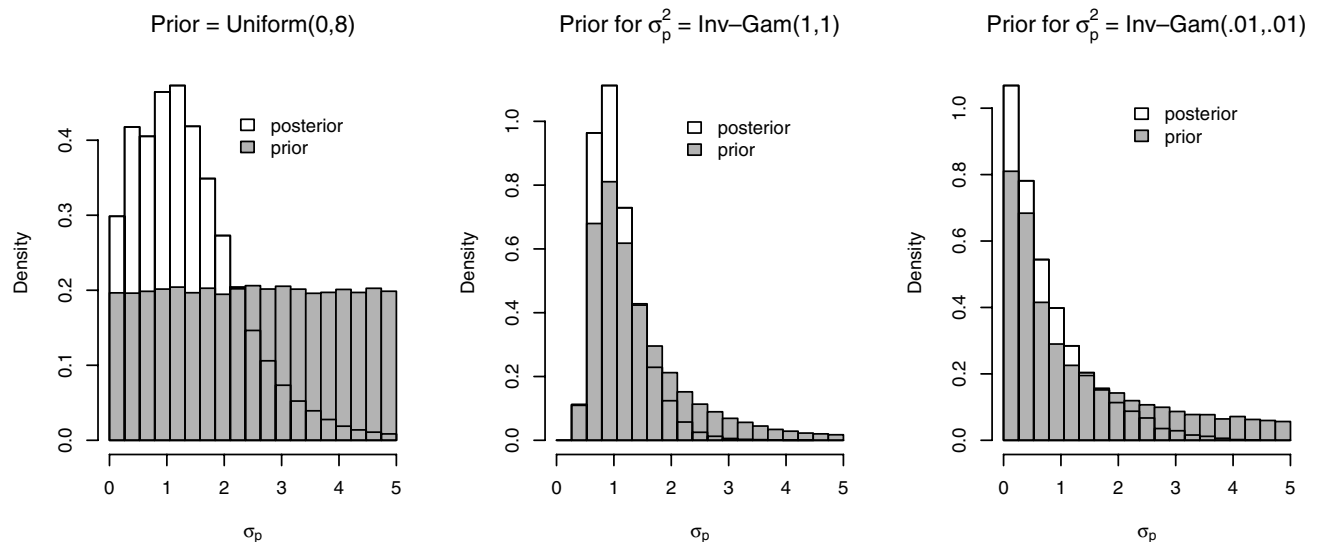


Figure 3. Estimated posterior distributions of σ_p and corresponding histograms obtained under several prior distributions. This figure appears in color in the electronic version of this article.

Table 4

Posterior model probabilities for the four models considered for the dipper data

Model	Posterior probability
Model 1: $p(t) + \phi(t)$	0.711
Model 2: $p(t + h) + \phi(t)$	0.229
Model 3: $p(t) + \phi(t + h)$	0.044
Model 4: $p(t + h) + \phi(t + h)$	0.016

4.2 Model Selection

The estimated variance components are very imprecise (and the mass of the posterior for σ_ϕ is close to zero) raising the question as to whether the heterogeneity components should be included in the model at all. To evaluate this, I computed the posterior model probabilities for each of the four models: model 1 (M1), year effects on p and ϕ (i.e., no heterogeneity); model 2, M1 plus heterogeneity in p ; model 3, M1 plus heterogeneity in ϕ ; and model 4, M1 plus heterogeneity in both parameters. This was implemented in **WinBUGS** by defining two indicator variables, say w_1 and w_2 , both having Bernoulli (0.5) prior distributions and premultiplying the additive random effect (see Web Appendix D). This idea was borrowed from Congdon (2005, section 3.2) for computing posterior model probabilities in variable selection problems. Model 1 is indicated if both $w_1 = 0$ and $w_2 = 0$, model 2 is indicated if $w_1 = 1$ but $w_2 = 0$, etc. The posterior probabilities of the four possible events were computed from the MCMC histories (Table 4). The results suggest that heterogeneity in ϕ is probably unimportant, whereas heterogeneity in p may not be negligible. In practice, we would obtain parameter estimates by model averaging. For the survival probability parameters, the model averaged estimates (posterior SD) were ($\phi_1 = 0.729$ (0.133), $\phi_2 = 0.455$ (0.073), $\phi_3 = 0.485$ (0.063), $\phi_4 = 0.628$ (0.061), and $\phi_5 = 0.609$ (0.060)). These are similar to those reported in Table 3, but slightly more precise in the absence of heterogeneity.

5. Discussion

Models of animal survival that allow for the parameterization of individual effects, or heterogeneity, are necessary in many studies of evolutionary biology and conservation and management (Nichols et al., 1982; Service, 2000; Cam et al., 2002; Cooch, Cam, and Link, 2002; Cam, Monnat, and Royle, 2004). Yet, accessible methods for conducting inference under such models with random effects have proven elusive. Pledger et al. (2003) proposed a model in which heterogeneity in individual parameters (both p and ϕ) was modeled according to a finite mixture,³ but implementation of continuous mixture analogs of such models has not been demonstrated.

The state-space parameterization provides a flexible framework for developing models of individual variation, and a practical solution of conducting inference in these complex problems via contemporary methods of MCMC. A straightforward implementation of heterogeneity models can be achieved

in the software package **WinBUGS**. In my view, this is the primary benefit of the state-space formulation of such models, as it allows attention to be focused on the conceptual formulation of models, while deemphasizing the technical aspects of developing MCMC algorithms. Such details render the application of Bayesian models well beyond the purview of most practitioners, which has been the major impediment to the widespread adoption of random effects type models. That such models are easily implemented in **WinBUGS** is not a panacea as **WinBUGS** does not effectively update parameters in all cases. In some problems containing individual effects, I have noted poor updating of random effects, which can be diagnosed by Markov chains for σ^2 that do not mix, when started from diffuse values. For such cases, it is not too difficult to develop one's own MCMC algorithms (e.g., see Web Appendix A), or modifying the **WinBUGS** implementation can improve performance (see Web Appendix D).

In the present context of capture–recapture models, the state-space formulation yields a solution to important model extensions such as the inclusion of individual heterogeneity in demographic parameters. I demonstrated the fitting of individual effects in a capture–recapture study on the European dipper, data that have been analyzed extensively in many contexts, but absent consideration of individual heterogeneity in parameters.

In the formulation of the model for capture–recapture data adopted here, also in previous efforts by others (Burnham and Rexstad, 1993; Pledger et al., 2003; Bonner and Schwarz, 2006), the model is developed conditional on the first capture of each individual. In general, it is not clear whether, or under what conditions, the estimated heterogeneity distribution (or covariate effects as the case may be) is an estimate of the *population* quantity that is most relevant to ecological and evolutionary theory. Intuitively, individuals with higher intrinsic survival rates will appear in the sample in higher proportions than individuals with lower intrinsic survival rates, because they have more net exposure to sampling through time. In the face of this potential bias of the sample toward “better surviving” individuals, two strategies might justify the use of capture–recapture models containing heterogeneity. First is to consider adults only and assert interest in heterogeneity in adult survival, and that the resulting estimates are approximately unbiased for that representing the population of breeders at large. Defining $g(\phi)$ for adult breeders is basically the approach adopted by Cam et al. (2002, 2004) and Link et al. (2002). Burnham and Rexstad (1993) also considered adults only, and the dipper data analyzed in Section 4 are of adults. In both Burnham and Rexstad (1993) and the present analysis, $p < 1$, and it stands to reason that some bias will arise in equating the estimated heterogeneity distribution to that of the breeding population at large. (i.e., longer surviving individuals, with larger individual survival effects, have a higher probability of entering the sample). The second strategy is to deemphasize the population distribution as the object of interest and, instead, use the individual effects model as a means of accounting for nuisance variation in the data. That heterogeneity masks other effects that might be the focus of investigation is widely known (e.g., senescent decline in survival rates; Service, 2000; Cam et al., 2002), and thus to elucidate these effects it is necessary

³ Software can be found at www.mcs.vuw.ac.nz/~shirley/pubs.

to accommodate heterogeneity among individuals. Finally, a preferred approach, where possible, is to develop the model based only on cohort 0 individuals. In this case, the distribution of individual effects is the population quantity of interest, provided that p and ϕ are independent for individuals.

I evaluated the potential bias due to sampling by Monte Carlo simulation. Some of the results are presented in Web Appendix E. Two cases were evaluated, that in which entrants into the population had a year to survive (or not) before they were exposed to sampling. This was meant to evaluate the effect of sampling adults only on potential bias in natal heterogeneity. Second, entrants into the population were exposed to sampling immediately, and the sample should more properly match the group of entrants. Viewing the entrants as “new” adults, this second comparison provides an assessment of potential bias in adult heterogeneity due to sampling. The basic conclusion is that potential bias in μ_ϕ and σ_ϕ increases dramatically (positive bias and negative bias, respectively) as the mass of $g(\phi)$ concentrates near 1.0 (as μ_ϕ and σ_ϕ increase), whereas the potential bias of μ_ϕ and σ_ϕ is relatively unaffected by heterogeneity in p . The effect of shifting the mass of $g(\phi)$ toward 1 is, not surprisingly, more pronounced if individuals have more exposure to mortality before being sampled. The results suggest that perhaps in the analysis of the Dipper data, the effect of sampling is that the estimator of μ_ϕ is biased slightly high, but that there is not much bias in σ_ϕ .

5. Supplementary Materials

Web Appendices referenced in Sections 2.1, 3, 4, and 5 are available under the Paper Information link at the *Biometrics* website <http://www.biometrics.tibs.org>.

ACKNOWLEDGEMENTS

I am grateful to Emmanuelle Cam, University Paul Sabatier, Toulouse, France and Marc Kéry at the Swiss Ornithological Institute for their collaboration on applications that motivated this work, the associate editor, K. B. Newman, and S. Pledger for many helpful comments and suggestions.

REFERENCES

- Alho, J. M. (1990). Logistic regression in capture-recapture models. *Biometrics* **46**, 623–635.
- Berliner, L. M. (1996). Hierarchical Bayesian time series models. In *Maximum Entropy and Bayesian Methods*, K. M. Hanson and R. N. Silver (eds), 15–22. Dordrecht: Kluwer Academic Publishers.
- Bonner, S. J. and Schwarz, C. J. (2006). An extension of the Cormack–Jolly–Seber model for continuous covariates with application to *Microtus pennsylvanicus*. *Biometrics* **62**, 142–149.
- Brooks, S. P., Catchpole, E. A., and Morgan, B. J. T. (2000). Bayesian animal survival estimation. *Statistical Science* **15**, 357–376.
- Burnham, K. P. and Rexstad, E. A. (1993). Modeling heterogeneity in survival rates of banded waterfowl. *Biometrics* **49**, 1194–1208.
- Cam, E., Link, W. A., Cooch, E. G., Monnat, J.-Y., and Danchin, E. (2002). Individual covariation in life-history traits: Seeing the trees despite the forest. *American Naturalist* **159**, 96–105.
- Cam, E., Monnat, J.-Y., and Royle, J. A. (2004). Dispersal and individual quality in a long lived species. *Oikos* **106**, 386–398.
- Congdon, P. 2005. *Bayesian Models for Categorical Data*. New York: Wiley.
- Cooch, E. G., Cam, E., and Link, W. A. (2002). Occam’s shadow: Levels of analysis in evolutionary ecology: Where to next? *Journal of Applied Statistics* **23**, 19–48.
- Coull, B. A. and Agresti, A. (1999). The use of mixed logit models to reflect heterogeneity in capture-recapture studies. *Biometrics* **55**, 294–301.
- Dorazio, R. M. and Royle, J. A. (2003). Mixture models for estimating the size of a closed population when capture rates vary among individuals. *Biometrics* **59**, 351–364.
- Dorazio, R. M., Royle, J. A., Söderström, B., and Glimskär, A. (2006). Estimating species richness and accumulation by modeling species occurrence and detectability. *Ecology* **84**, 842–854.
- Dupuis, J. A. (1995). Bayesian estimation of movement and survival probabilities from capture-recapture data. *Biometrika* **82**, 761–772.
- Gelman, A. (2006). Prior distributions for variance parameters in hierarchical models (comment on article by Browne and Draper). *Bayesian Analysis* **1**, 515–534.
- Gilks, W. R., Thomas, A., and Spiegelhalter, D. J. (1994). A language and program for complex Bayesian modelling. *The Statistician* **43**, 169–178.
- Huggins, R. M. (1989). On the statistical analysis of capture experiments. *Biometrika* **76**, 133–140.
- Ihaka, R. and Gentleman, R. (1996). R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics* **5**, 299–314.
- Jones, R. H. (1993). *Longitudinal Data with Serial Correlation: A State-Space Approach*. London: Chapman & Hall.
- Lebreton, J.-D., Burnham, K. P., Clobert, J., and Anderson, D. R. (1992). Modeling survival and testing biological hypotheses using marked animals: A unified approach with case studies. *Ecological Monographs* **62**, 67–118.
- Link, W. A., Cam, E., Nichols, J. D., and Cooch, E. G. (2002). Of BUGS and birds: Markov chain Monte Carlo for hierarchical modeling in wildlife research. *Journal of Wildlife Management* **66**, 277–291.
- Marzolin, G. (1988). Polygynie du Cincle plongeur (*Cinclus cinclus*) dans les côtes de Lorraine. *L’Oiseau et la Revue Française d’Ornithologie* **58**, 277–286.
- Newman, K. B., Buckland, S. T., Lindley, S. T., Thomas, L., and Fernández, C. (2006). Hidden process models for animal population dynamics. *Ecological Applications* **16**, 74–86.
- Nichols, J. D., Stokes, S. L., Hines, J. E., and Conroy, M. J. (1982). Additional comments on the assumption of homogeneous survival rates in modern bird banding estimation models. *Journal of Wildlife Management* **46**, 953–962.

- Norris III, J. L. and Pollock, K. H. (1996). Nonparametric MLE under two closed capture-recapture models with heterogeneity. *Biometrics* **52**, 639–649.
- Pledger, S., Pollock, K. H., and Norris, J. L. (2003). Open capture-recapture models with heterogeneity: I. Cormack-Jolly-Seber model. *Biometrics* **59**, 786–794.
- Pollock, K. H. and Raveling, D. G. (1982). Assumptions of modern band-recovery models, with emphasis on heterogeneous survival rates. *Journal of Wildlife Management* **46**, 88–98.
- Pollock, K. H., Hines, J. E., and Nichols, J. D. (1984). The use of auxiliary variables in capture-recapture and removal experiments. *Biometrics* **40**, 329–340.
- Royle, J. A. (2006). Site occupancy models with heterogeneous detection probability. *Biometrics* **62**, 97–102.
- Royle, J. A. and Kéry, M. (2007). A Bayesian state-space formulation of dynamic occupancy models. *Ecology* **88** (7), 1813–1823.
- Royle, J. A., Dorazio, R. M., and Link, W. A. (2007). Analysis of multinomial models with unknown index using data augmentation. *Journal of Computational and Graphical Statistics* **16**, 67–85.
- Service, P. M. (2000). Heterogeneity in individual mortality risk and its importance for the evolutionary studies of senescence. *American Naturalist* **156**, 1–13.
- Sturtz, S., Ligges, U., and Gelman, A. (2005). R2WinBUGS: A package for running WinBUGS from R. *Journal of Statistical Software* **12**, 1–17.

Received February 2006. Revised May 2007.

Accepted June 2007.