A hierarchical Bayesian approach to multi-state mark–recapture: simulations and applications

Anna M. Calvert[1,2]*, Simon J. Bonner[3]†, Ian D. Jonsen[4]†, Joanna Mills Flemming[5]†, Sandra J. Walde[1] and Philip D. Taylor[2]

[1]*Department of Biology, Dalhousie University, Halifax NS, Canada B3H 4J1;* [2]*Atlantic Cooperative Wildlife Ecology Research Network & Biology Department, Acadia University, Wolfville NS, Canada B4P 2R6;* [3]*Department of Statistics and Actuarial Science, Simon Fraser University, Burnaby BC, Canada V5A 1S6;* [4]*Bedford Institute of Oceanography, Fisheries and Oceans Canada, Dartmouth NS, Canada B2Y 4A2; and* [5]*Department of Mathematics and Statistics, Dalhousie University, Halifax NS, Canada B3H 3J5*

Summary

1. Mark–recapture models are valuable for assessing diverse demographic and behavioural parameters, yet the precision of traditional estimates is often constrained by sparse empirical data. Bayesian inference explicitly recognizes estimation uncertainty, and hierarchical Bayes has proven particularly useful for dealing with sparseness by combining information across data sets.

2. We developed a general hierarchical Bayesian multi-state mark–recapture model, tested its performance on simulated data sets and applied it to real ecological data on stopovers by migratory birds.

3. Our hierarchical model performed well in terms of both precision and accuracy of parameters when tested with simulated data of varying quality (sample size, capture and survivorship probabilities). It also provided more precise and accurate parameter estimates than a non-hierarchical model when data were sparse.

4. A specific version of the model, designed for estimation of daily transience and departure of migratory birds at a mid-route stopover, was applied to 11 years of autumn migration data from Atlantic Canada. Hierarchical estimates of departure and transience were more precise than those derived from parallel non-hierarchical and frequentist methods, and indicated that inter-annual variability in parameters suggested by these other methods was largely due to sampling error.

5. *Synthesis and applications.* Estimates of demographic parameters, often derived from mark–recapture studies, provide the basis for evaluating the status of species at risk, for developing conservation and management strategies and for evaluating the results of current protocols. The hierarchical Bayesian multi-state mark–recapture model presented here permits partitioning of complex parameter variation across space or time, and the simultaneous analysis of multiple data sets results in a marked increase in the precision of estimates derived from sparse capture data. Its structural flexibility should make it a valuable tool for conservation ecologists and wildlife managers.

Key-words: departure, MARK, migration, movement, random effects, songbird, state-space, stopover, survival, WinBUGS

Introduction

Demographic trait estimation has been greatly facilitated by the monitoring of individually marked animals over time, via live recaptures/observations and dead recoveries (hereafter all termed 'recapture'), providing quantitative information for wildlife conservation and management. Mathematical mark–recapture models based upon these encounters were originally devised for the study of abundance and survivorship (Cormack 1964; Jolly 1965; Seber 1965) and subsequently modified to address a broader variety of ecological questions (e.g. Pollock *et al.* 1990; Lebreton *et al.* 1992). The multi-state model structure (Arnason 1972, 1973; Schwarz *et al.* 1993), in particular, permits the estimation of additional parameters including recruitment (Pradel 1996), breeding probability (Schwarz & Arnason 2000), harvest mortality (Calvert & Gauthier 2005), and population growth rate (Caswell & Fujiwara 2004). The leading software package for implementation of these models is the program MARK (White &

*Correspondence: E-mail: anna.calvert@dal.ca
†These three authors made equal contributions to this work.

Burnham 1999), and related methods continue to be developed (e.g. Williams, Nichols & Conroy 2002; Bonner & Schwarz 2006).

Traditional frequentist methods of estimating parameters from mark–recapture data are sensitive to low survival and capture probabilities, small sample sizes, and the number of sampling intervals, such that sparse data greatly limit the precision of parameter estimates (Pollock *et al.* 1990; O'Brien, Robert & Tiandry 2005; Morris *et al.* 2006). However, greater analytical power is available through a Bayesian approach (Harwood & Stokes 2003; Gelman *et al.* 2004), where (i) parameters are considered random variables rather than fixed unknown values, (ii) prior knowledge about parameter distributions can be directly incorporated into estimation, and (iii) data are used to estimate the probability of a given hypothesis. Bayesian methods permit greater precision of parameter estimates due to incorporation of prior information (McCarthy & Masters 2005), explicit recognition of uncertainty (Harwood & Stokes 2003), and enhanced evaluation of complex variation from sparse data (Clark *et al.* 2005; Clark & Gelfand 2006). These advantages, in combination with computational advances, have driven a recent expansion of Bayesian methods in ecology (Ellison 2004; Clark 2005), including applications to mark–recapture modelling (e.g. Poole 2002; Gimenez *et al.* 2007; Dupuis & Schwarz 2007).

Nevertheless, practical constraints in traditional mark–recapture modelling frameworks may limit the analysis of complex data structures. For instance, the 'robust design' family of models (Pollock 1982; Kendall, Pollock & Brownie 1995) stratify time intervals into primary and secondary sampling periods (e.g. year and day, respectively), but require that the population be closed to immigration and emigration within primary sampling periods. Recent extensions allow estimation of short-term survivorship (e.g. across days within a year; Schwarz & Stobo 1997), but the multi-state robust design model does not permit state-transitions within primary sampling periods. Modelling of complex multi-state mark–recapture data could therefore benefit from additional structural flexibility, such as that available through a hierarchical Bayesian framework (Link *et al.* 2002; Clark & Gelfand 2006; Jonsen, Myers & James 2006).

Hierarchical Bayes accommodates stochasticity at multiple levels in a manner similar to frequentist random-effects models (Clark 2005; Zheng *et al.* 2007), and has been applied to studies of animal movement (Jonsen, Myers & Flemming 2003), species richness (Kéry & Royle 2008), and metapopulation dynamics (Royle & Kéry 2007). It is especially advantageous over non-hierarchical models when parameter variation requires partitioning across several related spatial or temporal replicates (Clark *et al.* 2005; Link & Barker 2005; Royle & Dorazio 2008).

Our need for such a flexible multi-scale approach arose during a study based upon daily monitoring of songbirds at an autumn migration stopover site over consecutive years. In a previous study (Calvert, Taylor & Walde 2009), we used a multi-state mark–recapture model formulation derived from Schaub *et al.* (2004) to estimate annual values of two stopover parameters: departure (analogous to mortality), and transi-

ence (analogous to movement). However, small within-year sample sizes limited the precision of the estimates. We were interested in using a hierarchical Bayesian analysis to reduce the impact of such sparseness by 'borrowing' information across different years, thus providing a balance between an approach that ignores variance among years and one where years are analysed independently. Additionally, we sought a model that could evaluate stopover decisions across two temporal scales (daily and annual variation), with the potential to further partition variance across other scales (e.g. among stopover sites, between taxonomic groups). Beyond this specific application, we perceived a demand for an accessible multi-state mark–recapture framework allowing flexible hierarchical structuring and able to deal with sparse ecological data.

The objectives of this study were therefore to: (i) build a hierarchical Bayesian multi-state mark–recapture model framework flexible enough for application to any system with a nested sampling scheme (e.g. daily monitoring within a season, conducted across years) or other hierarchical structuring; (ii) assess the accuracy and precision of estimates obtained with this model at varying levels of data quality and sample size; (iii) verify that the hierarchical structuring in this model permits improved estimation from poor-quality data relative to non-hierarchical methods; and (iv) apply this model to the assessment of stopover decisions, using 11 years of daily migration monitoring data.

Model structure and notation

MULTI-STATE MARK–RECAPTURE MODELS

Multi-state mark–recapture analysis was first suggested by Arnason (1972, 1973), and extended by Hestbeck, Nichols & Malecki (1991), Brownie *et al.* (1993), and Schwarz, Schweigert & Arnason (1993). Based upon the Cormack–Jolly–Seber model that estimates survivorship (φ) while accounting for probability of encounter (p), the multi-state formulation additionally permits individuals to move stochastically among 'states' with probability $\psi_{k,j}$ (movement from state k to j), where each state may be characterized by different φ or p probabilities. If we also allow for temporal variation in parameters, animals alive at time t in state k will be captured with probability $p_{k,t}$, survive to time $t + 1$ with probability $\varphi_{k,t}$ and will shift to state j by time $t + 1$ with the probability $\psi_{k,j,t}$. The movement parameter ψ was originally devised to distinguish mortality from emigration out of the study area (i.e. where states are geographical locations), but it can equally be used to represent other state transitions (Lebreton & Pradel 2002; White *et al.* 2006). Similarly, the definition of the φ parameter is not restricted to true demographic survivorship, and hence, the combination of non-standard φ and ψ interpretations allows for the modelling of diverse ecological scenarios beyond estimation of survival and emigration (e.g. Schaub, Liechti & Jenni 2004). As with single-state models, multi-state mark–recapture models rely upon some critical assumptions, such as: (i) marks are not lost, and all marks and states are correctly identified; (ii) the marking of individuals

does not affect their probabilities of capture, survival or movement; (iii) every individual in state k present in the population at time t is subject to the same capture probability $p_{k,t}$, the same survival probability $\varphi_{k,t}$, and the same vector of movement probabilities $\psi_{k,j,t}$; and (iv) the fate of each individual is independent of the fates of others. Pollock *et al.* (1990) and Williams *et al.* (2002) further discuss model assumptions and likelihoods, and Clark *et al.* (2005) consider how some assumptions may be relaxed under hierarchical Bayes; our model did not relax these assumptions.

HIERARCHICAL BAYESIAN APPROACH

Multi-state mark–recapture estimation using Bayesian methods has been described by Dupuis (1995), King & Brooks (2002), Clark *et al.* (2005), and Gimenez *et al.* (2007). Our approach builds on their work. We describe below the components of our hierarchical multi-state mark–recapture model as implemented within the software packages WinBUGS 1.4 (Lunn *et al.* 2000) and OpenBUGS (Thomas *et al.* 2006), hereafter collectively referred to as 'BUGS' (i.e. the same model notation is used in both platforms).

DATA

Y = total number of 'years' y sampled
T_y = total number of 'days' t sampled in year y
K = total number of possible states k (or j)
N_y = total number of individuals i marked in year y
$c_{y,i}$ = first capture (i.e. marking) date of individual i in year y

We define 'years (y)' and 'days (t)' as the primary and secondary sampling periods, respectively, where Y and T_y represent the total number of sampling intervals at each level, but these could be generalized to any other data structure as long as t is nested within y. The number of individuals marked annually can vary across years (N_y), and the first encounter date of each individual within each year is also specified ($c_{y,i}$).

Note that unlike the robust design models (Pollock 1982; Kendall *et al.* 1995), our model considers each year's data to be conditionally independent (i.e. with shared hyperprior distributions). No assumption is made about the behaviour across years for individuals that may be sampled in multiple data sets.

STATE AND OBSERVATION VECTORS

$x_{y,i,t}$ = observed state of individual i at time t within year y
$z_{y,i,t}$ = true state of individual i at time t within year y
$w_{y,i,t}$ = observation indicator for individual i at time t within year y

The two state vectors defined in this model are x, the time-series of observed states (1: K) for each individual (i.e. the encounter histories); and z, a latent variable representing the time series of true states for each individual [irrespective of whether it was observed, but known (i.e. equal to x) when

$w = 1$]. Both x and z consist of values ranging between 1 and $K + 1$; we used a non-0 value ($K + 1$) for the unobserved state in x and the dead state in z, for notational convenience. The observation vector w is a binary indicator of whether an individual was observed [irrespective of state, i.e. $w = 1$ if $x \in (1, \dots, K)$; $w = 0$ otherwise].

PARAMETERS

$\varphi_{y,k}$ = probability, within year y, that an individual that is alive and in state k survives a 1-day interval, independent of whether it changes state in that interval (i.e. assumed equal across all 1-day intervals)
$p_{y,k}$ = probability, within year y, that an individual that is alive and in state k will be observed on a given day
$\psi_{y,k,j}$ = probability, within year y, that an individual that is alive and in state k will move to state j over a 1-day interval (i.e. assumed equal across all one-day intervals)

Parameter definitions apply when $k \in (1, \dots, K)$; see below for transitions to and from $K + 1$. Extensions of the model could accommodate parameter variation among secondary sampling periods (t), but are not shown.

MODEL

Likelihood

The general form of the likelihood function for this hierarchical multi-state mark–recapture model is:

$$P\{x \mid \varphi,\Psi,p\} = P\{z,w \mid \varphi,\psi,p\} =$$
$$\sum_{y=1}^{Y}\sum_{i=1}^{N_y}\sum_{t=c_{y,i}+1}^{T} P\{z_{y,i,t} \mid z_{y,i,t-1}, \varphi_{y,k}, \Psi_{y,k,j}\} \times P\{w_{y,i,t} \mid z_{y,i,t}, p_{y,k}\} \quad \text{eqn 1}$$

where $k = z_{y,i,t-1}$ and $j = z_{y,i,t}$.

Note that (i) the model conditions on time of first capture (i.e. starting at $t = c_{y,i} + 1$), and it is assumed that the true state is known at the time of first capture (i.e. $z_{y,i,c} = x_{y,i,c}$); (ii) the probability of encounter history data x is equal to the joint probability of observation vector w and state vector z; and (iii) the likelihood separates process from observations, in parallel to state-space models (e.g. Jonsen *et al.* 2003, Gimenez *et al.* 2007).

Process model

The first likelihood component represents the process, i.e. the true state z in one time period conditional upon the true state in the previous time period ($z_{y,i,t} \mid z_{y,i,t-1}$). This is defined by a categorical distribution with probabilities given by:

$$P(z_{y,i,t} = k \mid z_{y,i,t-1} = j) = \begin{cases} \varphi_{y,k}\psi_{y,k,j} & k = 1, \dots, K, \ j = 1, \dots, K \\ 1 - \varphi_{y,k} & k = 1, \dots, K, \ j = K+1 \\ 1 & k = K+1, \ j = K+1 \\ 0 & k = K+1, \ j = 1, \dots, K \end{cases}$$

$$\text{eqn 2}$$

Thus, the true state z of an individual i in year y at time t depends on its previous state, as well as the combined probability that it survived in that state ($z_{y,i,t-1}$) and moved from that state to the current state ($z_{y,i,t,1}$). The second category in equation 2 reflects individuals that do not survive from $t-1$ to t, while the final two categories simply state that dead animals (i.e. those no longer in the study population) must remain dead.

Observation model

The second likelihood component represents the observation model, that is the probability of observation indicator time series $w_{y,i,t}$ conditional upon the latent process time series $z_{y,i,t}$ (i.e. $w_{y,i,t} | z_{y,i,t}$). This probability was defined by a Bernoulli distribution, given by:

$$P(w_{y,i,t} | z_{y,i,t} = k) \sim \text{Bernoulli } [p_{y,k}]$$

for $k = 1, \dots, K$; note that $p_{y,K+1} = 0$ eqn 3

Priors

Bayesian inference requires the definition of prior distributions for all parameters to be estimated (in this case φ, p and ψ). When possible, prior distributions should describe the researchers' previous knowledge about parameter values, but uninformative (or weakly informative) prior distributions may be used if little is initially known about the parameters, or to reduce the influence of priors on estimation. Daily sampling within years naturally leads to a hierarchical model with two levels, where prior distributions are specified in the form of 'hyperpriors' (Carlin & Louis 1996). At the first level, we assume that the parameters in each primary period (e.g. daily capture probabilities in each year) form a random sample from a distribution with unknown mean and variance (the hyperparameters). At the second level, we model these means and variances as random samples from a further distribution with unknown mean and variance (the hyperpriors); these values reflect the prior parameter knowledge (if any).

Prior distributions for all parameters were defined on the logit scale so that the parameters would be naturally bounded between 0 and 1. Across all years, parameters of the same type were assigned normal priors on the logit scale with equal, but unknown, mean and variance; for example, the capture probability of individuals in state k in year y was assigned the prior distribution: $\text{logit}(p_{y,k,t}) \sim N(\mu_{pk}, 1/\tau_{pk})$, with mean and variance state-dependent but equal in all years (precision τ, and not variance $1/\tau$, is specified in BUGS notation). These parameters (the hyperparameters: μ_{pk}, $\mu_{\varphi k}$, $\mu_{\psi k}$, τ_{pk}, $\tau_{\varphi k}$ and $\tau_{\psi k}$) were then assigned further distributions (hyperpriors) dependent on state alone. We used the conjugate distributions $\mu_k \sim N(0,1000)$ and $\tau_k \sim \text{Gamma}(0·001,0·001)$. As all probabilities of movement must sum to 1 for each starting state, priors were specified for $\psi_{1,1}$ and $\psi_{2,2}$, with $\psi_{1,2} = 1 - \psi_{1,1}$ and $\psi_{2,1} = 1 - \psi_{2,2}$; in cases with more than two states, ψ would represent a multinomial (instead of binomial) process and priors would have to be accordingly defined.

ALTERNATIVE AND NON-HIERARCHICAL MODEL STRUCTURES

The BUGS code for the hierarchical multi-state mark–recapture model described above, where parameters vary with k and y but not with t (giving the equivalent of an 'annual average' estimate for each daily parameter) is outlined in Supporting Information, Appendix S1. In order to test specific biological hypotheses, alternative versions of this general model can be constructed to reflect differences in the structure of parameter variation. For instance, the model could be specified to reflect daily changes in parameters by simply adding another index to the parameters (e.g. $\varphi_{y,k,t}$). Similarly, reducing the sources of variation can be accomplished by removing an index from a particular parameter (e.g. no state or daily variation on survivorship: φ_y), or fixing parameters at a known value (see case study below). Effects of external covariates can also be modelled: an effect of daily wind speed s_t upon daily variation in survivorship, for example, could be defined as $\varphi_t = B_0 + B_1 * s_t$ (where B_0 is the intercept and B_1 is the slope, each defined by a probability distribution). Finally, a non-hierarchical equivalent of the model can be described for analysis of a single year's data, using prior distribution parameters $\mu = 0$ and $\tau = 0·667$ instead of the hyperparameters defined in Supporting Information, Appendix S1, as follows:

$$P\{z,w \mid \varphi,\psi,p\} =$$
$$\sum_{i=1}^{N}\sum_{t=c_i+1}^{T} P\{z_{i,t} \mid z_{i,t-1},\varphi_k,\psi_{k,j}\} \times P\{w_{i,t} \mid z_{i,t},p_{y,k}\}$$ eqn 4

where $k = z_{i,t-1}$ and $j = z_{i,t}$

Assessment of model performance

SIMULATED DATA

A main objective of this hierarchical model was to reduce the problems of parameter bias and imprecision posed by limited data (e.g. O'Brien *et al.* 2005), and specifically by small sample size, low capture probability, or low survivorship. We therefore designed a three-step simulation study to assess model performance with varying data quality, based upon a generalized model where $K = 2$, $Y = 10$, $T = 10$, and all $c_{y,i} = 1$, with model structure [$\varphi(y,k)$ $p(y)$ $\psi(y,k,j)$]. All simulations were built around a hypothetical situation where state $k = 1$ represented a 'sink' state (i.e. lower survivorship and more immigration than emigration) and $k = 2$, a 'source' state ($\varphi_1 < \varphi_2$ and $\psi_{2,1} > \psi_{1,2}$). True parameter values (see Supporting Information, Appendix S2) for each of the 10 years were independently drawn from logit-normal distributions with known means (described below) and variance (between 0·005 and 0·01), so that true and estimated values had the same hierarchical structuring.

First, to verify that model performance was satisfactory when capture probability, survivorship and sample size were high, we created 25 sets of simulated hierarchical data based on 'good' parameter values. Each data set was characterized by: $N_y = 200$ per state at $t = 1$; mean annual survivorship probabilities $\varphi_{y,1} = 0·70$ and $\varphi_{y,2} = 0·90$; mean annual capture

probabilities $p_y = 0·80$; and mean annual movement probabilities $\psi_{y,1,2} = 0·15$ and $\psi_{y,2,1} = 0·35$.

Secondly, to test model performance with 'poor' data quality, we created 25 sets of simulated hierarchical data for each of 12 different scenarios representing combinations of low (mean $p_y = 0·20$) or very low ($p_y = 0·05$) capture probability, high (mean $\varphi_{y,1} = 0·70$, $\varphi_{y,2} = 0·90$), moderate ($\varphi_{y,1} = 0·40$, $\varphi_{y,2} = 0·60$) or low ($\varphi_{y,1} = 0·10$, $\varphi_{y,2} = 0·30$) survivorship, and small ($N_y = 50$ individuals marked in each state at $t = 1$) or large ($N_y = 200$ per state at $t = 1$) sample size. The 12 scenarios resulted from all combinations of each of these three factors ($3\varphi \times 2p \times 2N$), where all true mean movement probabilities were $\psi_{y,1,2} = 0·15$ and $\psi_{y,2,1} = 0·35$.

Thirdly, to determine whether the hierarchical model structure would improve estimation in years of sparse data, we created 25 sets of simulated hierarchical data for a scenario with 'mixed' data quality (as measured by capture probability) among years. Capture probability was taken from the low (mean $p_y = 0·20$) distribution in seven years ($y = 1–4$, $8–10$), and from the very low (mean $p_y = 0·05$) distribution in three years ($y = 5–7$); survivorship was moderate (means $\varphi_{y,1} = 0·40$, $\varphi_{y,2} = 0·60$), sample sizes small ($N_y = 50$/state at $t = 1$), and mean movement probabilities were $\psi_{y,1,2} = 0·15$ and $\psi_{y,2,1} = 0·35$. We compared parameter estimates for these data from both the hierarchical model described above and the non-hierarchical equivalent (eqn 4).

CONVERGENCE AND MODEL FIT

Model convergence occurred most slowly with the poorest data quality, and therefore, we based our convergence tests on our weakest data sets: very low capture probability (mean $p_y = 0·05$), low survivorship (means $\varphi_{y,1} = 0·10$, $\varphi_{y,2} = 0·30$), and small sample size ($N_y = 50$ per state). We monitored the scale reduction factor (the factor by which posterior intervals are expected to shrink if simulation continued; Gelman & Rubin 1992) based on 3 chains (each with different initial values), and found that it closely approached the desired value of 1 (< 1·05 for all parameters) within the first 10 000 iterations. Consequently, for all simulated data sets we set a burn-in run of 10 000 iterations, and estimated parameters based on 50 000 subsequent iterations.

We do not address the issue of model fit and selection in this study. General methods for assessing model fit and for determining the 'best' model have been developed in detail (e.g. Burnham & Anderson 2002), and specific Bayesian model selection tools are available (e.g. Spiegelhalter *et al.* 2002). In particular, Reversible Jump MCMC methods involve the estimation of posterior model probabilities in addition to posterior parameter distributions (King & Brooks 2002; Gimenez *et al.* 2009), and represent a natural Bayesian approach to model selection and uncertainty; they would be a useful next step with this model.

ESTIMATION ACCURACY AND PRECISION

The accuracy and precision of parameter estimation based on the good simulated data ($N_y = 200$ per state, mean $p_y = 0·80$,

mean $\varphi_{y,1} = 0·70$ and $\varphi_{y,2} = 0·90$) are illustrated in Fig. 1. For all three parameters, across both states, the posterior medians of annual estimates and of the hierarchical means (Hm) were very close to the true mean values used to simulate the data. Additionally, the 95% credible intervals always included the true mean values and were narrow around the posterior median, though slightly wider for capture probability p than for the other parameters. Finally, the estimates appeared to capture interannual variation in parameter values very well, with little 'shrinkage' of the annual medians toward the hierarchical mean.

For the 12 scenarios using poor quality data, Fig. 2 illustrates the estimation bias (posterior median – true value) for each parameter resulting from each of the 25 replicate data sets across the varying levels of survivorship, capture probability and sample size. Bias was greatest with the most sparse data (small N, very low p and low φ) and decreased with improving data quality; mean bias was < 0·10 in 32/36 parameter-scenario combinations, with the following four exceptions: estimates of movement (ψ) when $N = 50$, mean $p_y = 0·05$ and mean $\varphi_y = 0·10$ and $0·30$; estimates of capture probability (p) when $N = 50$, mean $p_y = 0·05$ or $0·20$, and mean $\varphi_y = 0·10$ and $0·30$; and estimates of p also when $N = 200$, mean $p = 0·05$ and mean $\varphi_y = 0·10$ and $0·30$. Additionally, shrinkage of annual estimates toward the hierarchical mean was stronger with sparse data and declined with increasing data quality (Supporting Information, Appendix S2). Bias of the estimates of hierarchical mean (Hm – mean of true annual values), and the width of their 95% credible intervals, are shown in Supporting Information, Appendix S3. As with the annual parameters, Hm estimates were close to true values (bias < 0·10) for most of the parameter-scenario combinations (29/36) and bias decreased as data quality improved (higher N, p, φ), though again bias was greatest for capture probability p. Precision of Hm estimates also improved with data quality (Supporting Information, Appendix S3).

Finally, the hierarchical and non-hierarchical estimates based on the data of mixed quality (mean $p_y = 0·20$ for most years but $0·05$ for years 5–7) are illustrated in Fig. 3 (estimates shown for $\phi_{y,2}$ and $\psi_{y,1,2}$; see Supporting Information, Appendix S2 for all parameter estimates). Relative to non-hierarchical estimates, the hierarchical estimates were much less variable among years, closer to true values, and had narrower credible intervals. For the three years with very low p, estimates from the non-hierarchical model were less accurate, and with larger 95% credible intervals, than in years with higher p; however, the hierarchical estimates from these three years were also further shrunk toward the hierarchical mean. Posterior medians, 95% credible interval size and true parameter values for all three components of the simulation study are presented in full detail in Supporting Information, Appendix S2.

Specific ecological application

MIGRATORY STOPOVER DECISIONS

We applied this model to an analysis of migratory stopover decisions by songbirds as described in Calvert *et al.* (2009).
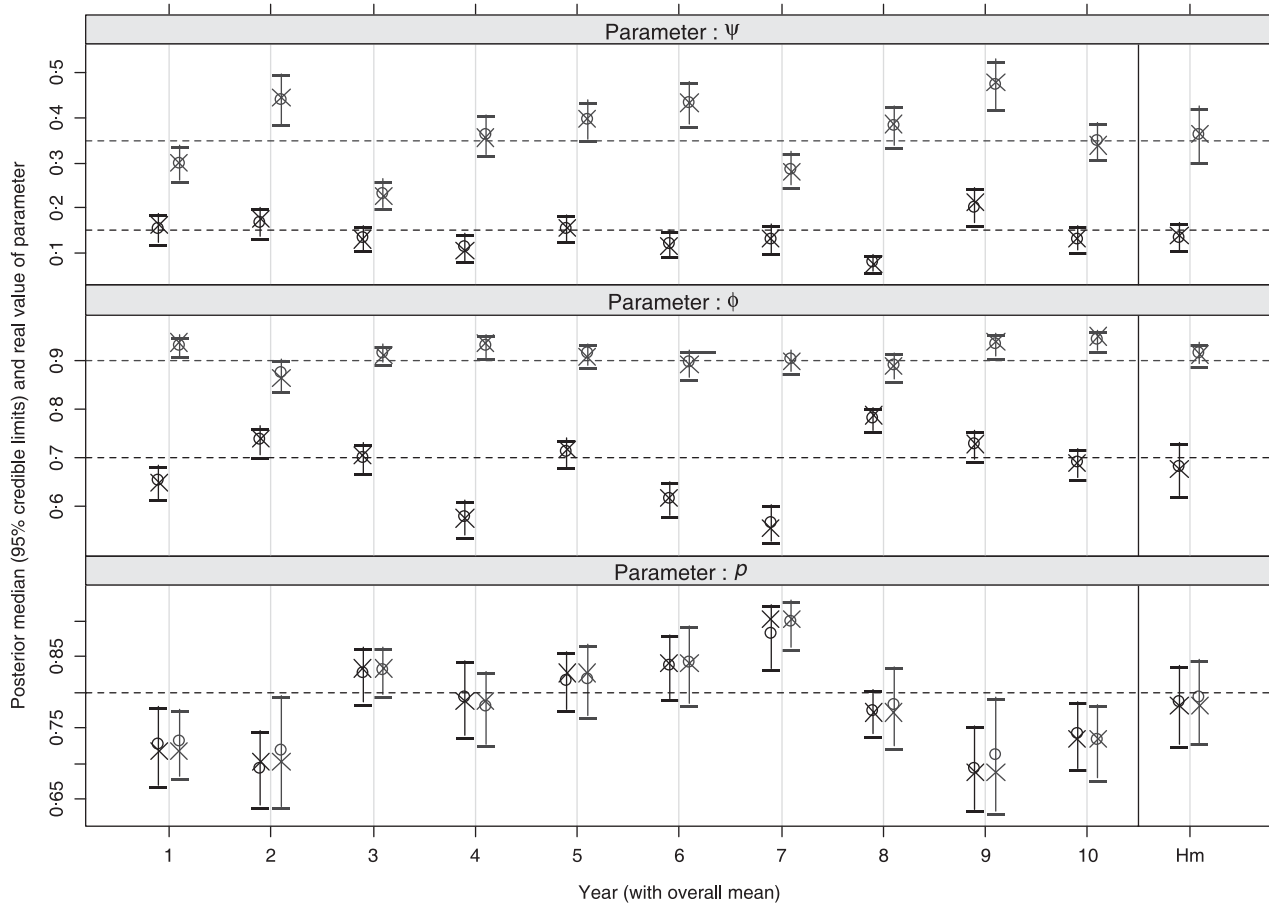
Fig. 1. Posterior medians (○) and 95% credible intervals (bars) of movement (ψ), survivorship (φ) and capture probability (p), for 10 years and hierarchical mean *Hm*, based on 25 replicate sets of 'good' simulated data ($N_y = 200$ per state at $t1$, mean $p_1 = p_2 = 0.80$, mean $\varphi_1 = 0.70$ and $\varphi_2 = 0.90$; mean $\psi_{y,1,2} = 0.15$, $\psi_{2,1} = 0.35$); black symbols are for p_1, φ_1, ψ_{12}, grey for p_2, φ_2, ψ_{21}. True parameter values used to simulate data (x) and the mean of the hierarchical distribution used to simulate them (dotted lines) are shown.

The data derive from autumn migration monitoring of temperate-breeding songbirds at an island stopover site in Atlantic Canada (Bon Portage Island, NS, 43°28′N, 65°44′W), where daily mark–recapture data on stopped individuals were collected from 1996 to 2006. Each year, between 52 and 562 (median 207) hatch-year yellow-rumped 'Myrtle' warblers (*Dendroica coronata*, hereafter MYWA) were leg-ringed during peak autumn migration (23 September–31 October) as part of the constant-effort mist-netting programme of the Canadian Migration Monitoring Network, and all subsequent days' recaptures were noted.

Derived from Schaub *et al.* (2004), mark–recapture analysis of stopover decisions involves a three-state model, with an 'initial' state (first capture only), a 'non-transient' state (potential for subsequent recapture), and a 'transient' state (departure within 24 h preventing recapture). By considering stopover and departure analogous to 'survival' and 'mortality' respectively, we constrain survivorship in the initial state to be $\varphi_1 = 1$ and transient survivorship to be $\varphi_3 = 0$; we thus only estimate survivorship (i.e. 1-departure) for non-transient individuals, φ_2. Similarly, we only estimate capture probability for non-transient birds (p_2); capture probability is inestimable for birds in the initial state, and is 0 by definition for transient

birds, and therefore, both p_1 and p_3 were set to 0 in the model notation. Finally, movement probabilities are also restricted: initial individuals cannot remain as initials ($\psi_{1,1} = 0$), and instead move to the non-transient ($\psi_{1,2}$) or transient ($\psi_{1,3}$, 'transience') states. The move to be transient or not is a permanent one, such that once in state 2 or 3, individuals remain in that state ($\psi_{2,2} = 1$ and $\psi_{3,3} = 1$); see Schaub *et al.* (2004) for parameter vectors and transition matrices.

For the study of stopover decisions, therefore, only three parameters are estimated: stopover (i.e. one-departure probability) for non-transient birds, φ_2; capture probability of stopping birds, p_2; and probability of transience upon arrival, $\psi_{1,3}$ (or of non-transience, $\psi_{1,2} = 1 - \psi_{1,3}$). In BUGS, specific constraints are reflected within the model code (see Supporting Information, Appendix S1): all constrained parameters are fixed to 0 or 1 as described above and not given prior distributions. As we were interested in inter-annual variability in transience and departure probabilities of migratory passerines (Calvert *et al.* 2009), we built a model constraining daily ψ and φ to be equal across days *t* within each year *y*. We were also interested in the influence of individuals' fuel load upon transience and departure, as migrants' decisions may be affected by the quantity of fuel (in the form of fat) they are
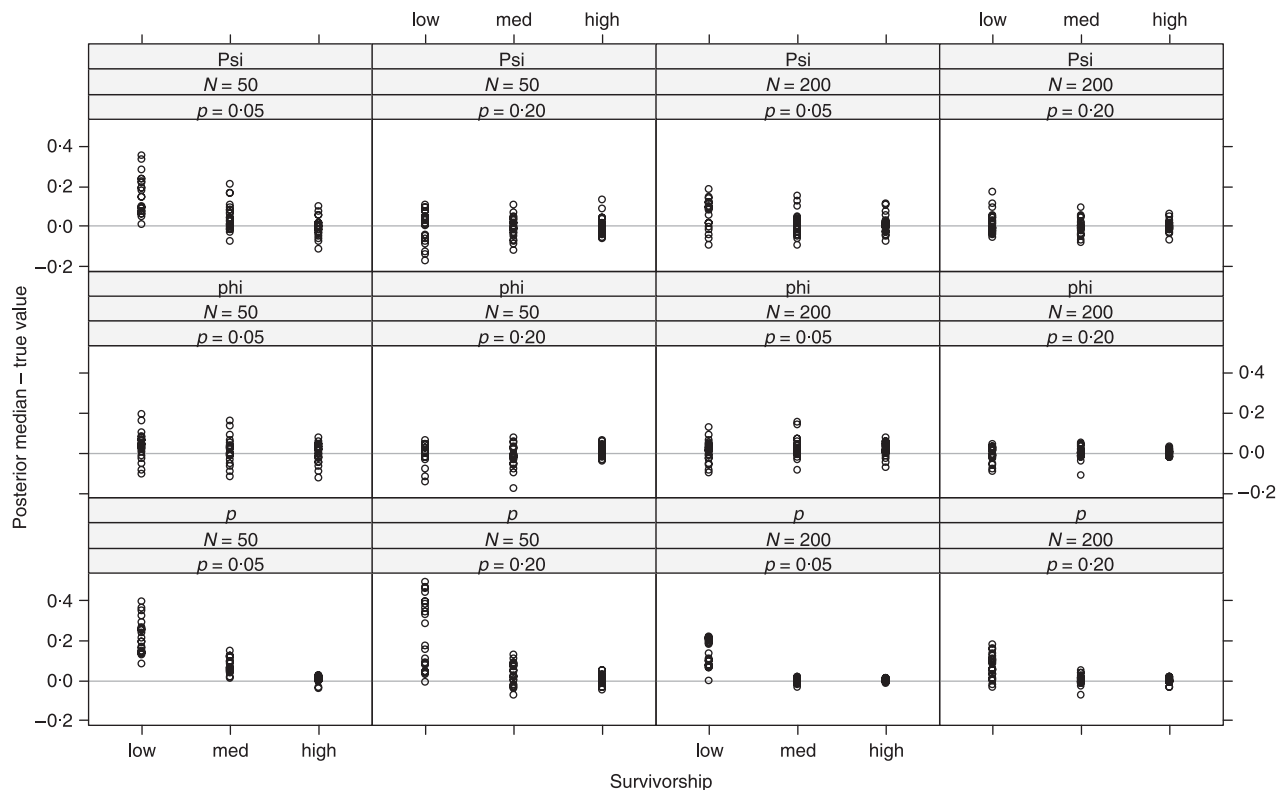
Fig. 2. Estimation bias (posterior median – true value) of ψ (top), φ (middle) and p (bottom) for 25 replicate data sets across each level of the cross-design 'poor data' simulation study; horizontal grey lines indicate accurate estimation, i.e. bias = 0. The design involved three levels of mean survivorship [high ($\varphi_{y,1} = 0.70$, $\varphi_{y,2} = 0.90$), med ($\varphi_{y,1} = 0.40$, $\varphi_{y,2} = 0.60$) or low ($\varphi_{y,1} = 0.10$, $\varphi_{y,2} = 0.30$)], two levels of mean capture probability [low $p_y = (0.20)$ or very low ($p_y = 0.05$)], and two levels of sample size [small ($N_y = 50$ individuals marked in each state at $t = 1$) or large ($N_y = 200$ per state at $t = 1$)]; all mean movement values were $\psi_{y,1,2} = 0.15$, $\psi_{y,2,1} = 0.35$.

carrying, particularly prior to flight across barriers such as open ocean (Alerstam & Lindstrom 1990). We divided the ringed sample between birds that had a 'high' fat-load (furculum at least half-full of fat, some with additional fat on breast, abdomen or below wings) or 'low' fat-load (furculum less than half-full of fat or empty) at the time of marking (see Calvert *et al.* 2009).

MODEL STRUCTURE

These sources of variability were incorporated into a general MYWA stopover model with structure [$\varphi(y*k) p(y*k) \psi(y*k)$] (see Supporting Information, Appendix S1). Structural features of the data set included $Y = 11$, $K = 3$, T_y ranging from 28 to 39 (in some years ringing efforts were halted early, but parameters were only estimated for the duration of the ringing season), and N_y ranging from 15–272 for high-fat birds, and 34–290 for low-fat birds. For each fat-group, we ran a 20 000-iteration burn-in period, and then obtained parameter estimates based upon a further 50 000 iterations, with both a hierarchical (Supporting Information, Appendix S1) and a non-hierarchical version of this same model (as outlined above). For comparison with BUGS estimates, we also estimated the parameters with an equivalent frequentist multi-state model in MARK ([$\varphi(f*k) p(f*k) \psi(f*k)$]), i.e. fat-load-('high' vs. 'low') and state-specific parameters,

with each year run separately; estimates from Calvert *et al.* 2009).

PARAMETER ESTIMATES

Figures 4 and 5 illustrate the posterior median and 95% credible intervals for annual MYWA departure (φ) and transience (ψ) probabilities for each of the fat groups from both the hierarchical and non-hierarchical BUGS models, and the hierarchical mean of each parameter across the 11 years of data (*Hm*). Estimates of φ and ψ from MARK models are also presented, although boundary estimates (0 or 1, typically in years of small sample size) are not shown. Median departure probabilities were generally higher for high-fat than low-fat birds, as previously seen (Calvert *et al.* 2009), but the fat-load difference was more evident in the hierarchical model estimates (where all years' values were derived from a common distribution) than in the non-hierarchical estimates (where each year's value was estimated independently). Wide credible intervals, however, provided little evidence for a fat effect on transience. Estimates of capture probability, p, were low and overlapping between the two fat groups, with hierarchical means of 0.040 (posterior median; 95% credible intervals 0.021–0.092) and 0.053 (0.033–0.085) for high- and low-fat, respectively.

Credible intervals on Bayesian parameter estimates and confidence limits on MARK estimates all showed a high
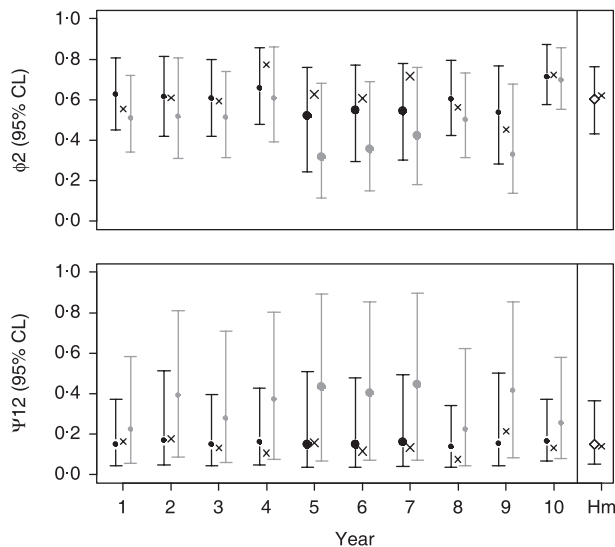
Fig. 3. Survivorship (φ_2, top) and movement (ψ_{12}, bottom) probabilities (median and 95% credible intervals) estimated from hierarchical (black, with hierarchical mean *Hm* shown by ◇) and non-hierarchical (grey) Bayesian multi-state mark–recapture models, relative to true parameter values (x), for 'mixed' data. Estimates are means from 25 replicate sets of simulated data with $N = 50$ individuals marked/state at day 1, moderate true survivorship (mean $\varphi_{y,1} = 0.40$, $\varphi_{y,2} = 0.60$), and mean true movement probabilities of $\psi_{y,1,2} = 0.15$, $\psi_{y,2,1} = 0.35$. Capture probabilities were drawn from distributions with mean $p = 0.05$ for years 5,6,7 (large symbols), but mean $p = 0.20$ for years 1–4, 8–10 (small symbols), to compare estimates between years of variable data quality.

degree of overlap for a given parameter within a particular year. Relative to non-hierarchical and MARK estimates, median hierarchical parameter estimates were 'shrunk' toward the hierarchical mean *Hm*, suggesting that much of the apparent inter-annual variability in estimates from the other methods could be due to sampling error (see Discussion). The 95% credible intervals around non-hierarchical annual estimates were generally larger than those from the hierarchical model, especially for departure and capture probabilities, indicating that the hierarchical modelling of several years' data produced more precise parameter estimates than year-specific models; this difference was particularly evident in years where few individual birds were marked. Annual capture probability estimates (not shown) from the hierarchical model similarly showed narrower credible intervals and more shrinkage toward *Hm* than estimates from the non-hierarchical model, further suggesting important sampling error among years (see Discussion).

Discussion

ACCURACY, PRECISION, AND HIERARCHICAL MODELLING

Our hierarchical Bayes multi-state mark–recapture model expanded upon previously developed models (e.g. Dupuis 1995; King & Brooks 2002; Clark *et al*. 2005). Using simulated

data, we first tested our model under an idealized scenario where: (i) many individuals are marked (large *N*), (ii) individuals remain in the population a long time (high φ), and (iii) they are likely to be observed if present (high *p*). Under these conditions, the hierarchical model performed well, providing annual estimates and hierarchical means that were very close to true values, had narrow credible intervals, and accurately detected inter-annual variation. Estimates were particularly precise for the two parameters of usual biological significance (survivorship φ and movement probability ψ), and showed no bias for any of the parameters. Supplementary tests with alternate hyperprior distributions on parameter means suggested little sensitivity of posterior estimates to choice of priors (see Supporting Information, Appendix S4). We concluded that the model structure was properly defined, and that both accurate and precise estimation is possible with this approach if sufficient data are available.

The second component of the simulation study was designed to test how well our hierarchical model would perform under more realistic biological conditions. For instance, in previous work (Calvert *et al*. 2009), even the most abundant migratory passerines at Bon Portage Island were captured rarely in some years, resulting in annual sample sizes ranging between 8 and 579 individuals ringed during peak migration. Additionally, daily capture probabilities of migrating songbirds are typically low, often on the order of 2·3–10·0% (Bachler & Schaub 2007; Salewski, Thoma & Schaub 2007; Calvert *et al*. 2009). The poor-quality simulated data (with $N = 50$ or 200 per state, variable survivorship, and capture probabilities of $p = 0.05$ or 0·20) were thus chosen to better represent the sparse data common in real multi-state mark–recapture studies.

As expected, poor data reduced the precision of parameter estimates. There was relatively little bias in estimates of hierarchical means, however, except for a tendency for a positive bias when true values were very close to zero (i.e. for *p* and ψ). Annual estimates tended to underestimate inter-year variation, but correctly captured the trends (i.e. the direction of deviation away from the hierarchical mean). Bias and precision of capture probability estimates appeared particularly sensitive to survivorship, especially when low survivorship was combined with either small sample size or very low capture probability; the little information in the data resulted in a posterior distribution of *p* that was highly influenced by the prior.

An important benefit of a hierarchical approach, as seen in our simulations and as explained elsewhere (e.g. Kéry & Royle 2008; Royle & Dorazio 2008), is that estimates are less biased and more precise. This can come at a slight cost of under-estimating inter-annual variation when data are sparse, because hierarchical models interpret deviation of individual parameter estimates that have low precision largely as sampling error (Efron & Morris 1975); deviation of more precise estimates is interpreted as process stochasticity. A non-hierarchical approach, however, can overestimate inter-annual variation by not accounting for sampling error. In the analysis of the Myrtle warbler data, the direction of inter-annual variation in hierarchical estimates was consistent with those
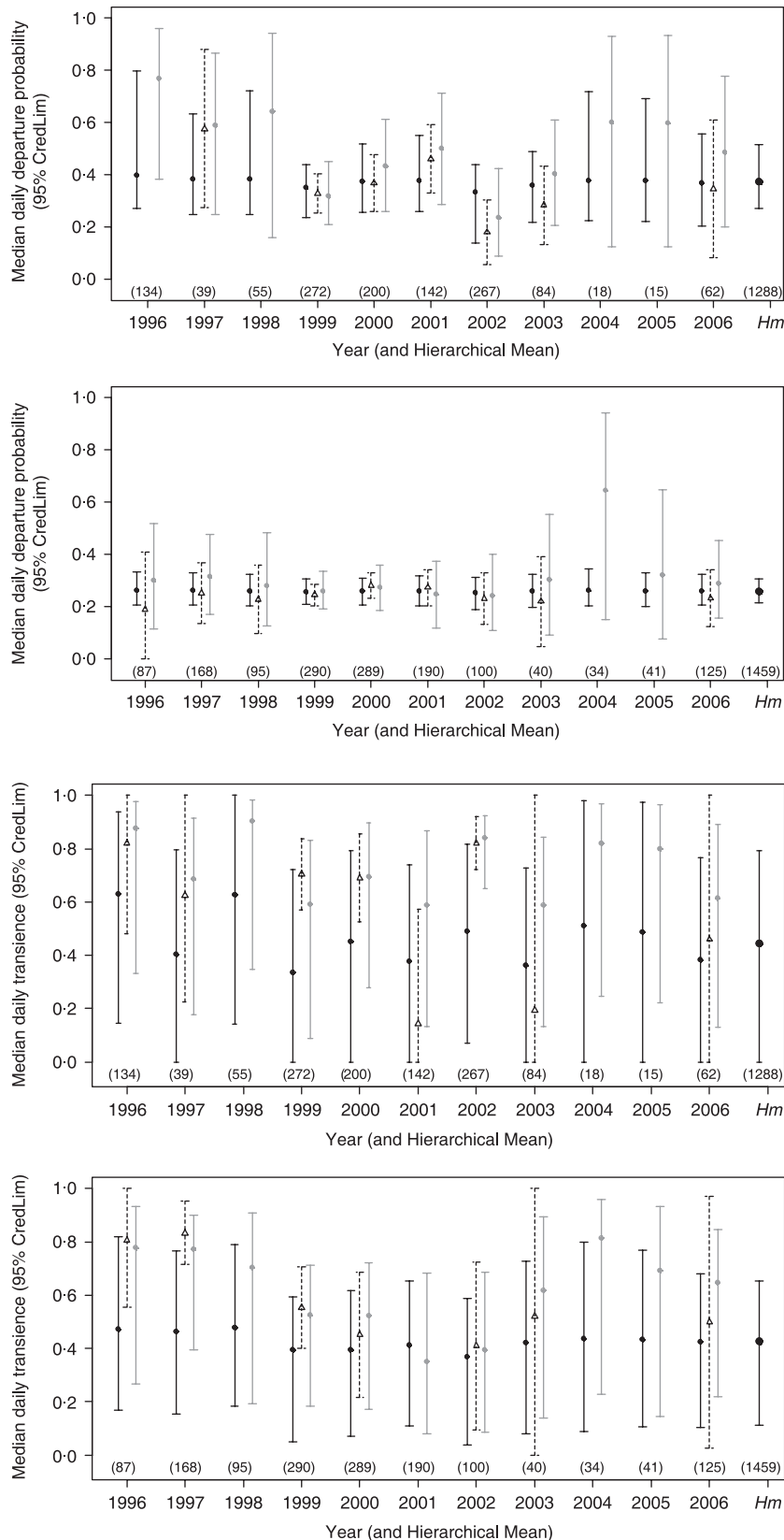
Fig. 4. Annual departure probabilities $(1 - \varphi_{y,2}$; median and 95% credible intervals) for high-fat (top) and low-fat (bottom) Myrtle warblers (*Dendroica coronata*) monitored at Bon Portage Island, Canada, from 1996–2006, estimated from hierarchical (black) and non-hierarchical (grey) Bayesian models; '*Hm*' indicates the hierarchical mean. Also shown are annual number of individuals ringed (at bottom of plot), and mean estimates (± SE) from frequentist MARK models (△), when estimable (boundary estimates, i.e. 0 or 1, are not shown; from Calvert *et al.* 2009).

Fig. 5. Annual transience probabilities $(\psi_{y,1,3}$; median and 95% credible intervals) for high-fat (top) and low-fat (bottom) Myrtle warblers (*Dendroica coronata*) monitored at Bon Portage Island, Canada, from 1996–2006, estimated from hierarchical (black) and non-hierarchical (grey) Bayesian models; '*Hm*' indicates the hierarchical mean. Also shown are annual number of individuals ringed (at bottom of plot), and mean estimates (± SE) from frequentist MARK models (△), when estimable (boundary estimates, i.e. 0 or 1, are not shown; from Calvert *et al.* 2009).

from the non-hierarchical Bayes model and the frequentist MARK estimates but the median values were closer to *Hm*, reflecting the interpretation of small sample sizes and poor capture rate as parameter uncertainty rather than process variance (Efron & Morris 1975; Clark 2005). However, due to smaller credible intervals on annual estimates, the hierarchical model still provided stronger evidence that high-fat birds depart this coastal stopover sooner than low-fat birds.

MODEL FLEXIBILITY AND APPLICATIONS TO COMPLEX ECOLOGICAL QUESTIONS

Our hierarchical Bayes model offers several advantages to the analysis of ecological mark–recapture data. Most obviously, the borrowing of information across hierarchical levels (in this case, years; Clark 2005; Jonsen *et al.* 2006) enables analysis of data sets too sparse to be analysed individually (e.g. Askey *et al.* 2007). Thus, while other studies have analysed passerine stopover behaviour separately for each year (e.g. Salewski & Schaub 2007), hierarchical Bayes permits the estimation of both a hierarchical mean (i.e. an average across years) and annual parameter values that might otherwise be inestimable or imprecise (Morris *et al.* 2006; Salewski *et al.* 2007). More broadly, this modelling approach may improve estimates of demographic parameters for endangered or elusive animals where sparse data (few individuals marked and low capture probabilities) are inevitable. Predictive probability distributions of future values, representing total variability in each parameter, can also be derived from hierarchical Bayes estimates.

The development of this hierarchical Bayesian model was driven by our need for modelling flexibility in multi-state analyses of songbird stopover behaviours, but its generality makes it available to many more diverse applications. The hierarchical time-structure (sampling days across study years) could be replaced with an alternate time-structure (e.g. weeks within seasons), or a spatial-structure where data are analysed hierarchically among study regions or populations. As with other multi-state models (reviewed in White *et al.* 2006), group structure could allow age or sex variation in parameters, states could represent geographical locations (e.g. Hestbeck *et al.* 1991) or reproductive stages (e.g. Nichols *et al.* 1994), and parameters could vary with temporal or individual covariates. Future analyses of stopover behaviour, for instance, may include additional hierarchical structuring among species or between adjacent stopover sites, or even alternative state definitions (e.g. fat-loads as states; Schaub 2006). Finally, although it was designed for a complex situation with multiple states, this model could easily be adapted to simpler study systems, such as single-state models or other mark–recapture variations, offering a valuable tool for diverse demographic studies.

Conclusions

Hierarchical Bayes accommodates complex variation and covariates (e.g. Zheng *et al.* 2007; Kéry & Royle 2008), and can improve parameter estimation relative to frequentist non-hierarchical methods (Poole 2002; Dupuis & Schwarz 2007). As a result, its value is increasing for both mark–recapture studies and other types of movement analyses (e.g. state-space modelling: Jonsen *et al.* 2003, 2006; Gimenez *et al.* 2007; Royle & Kéry 2007). The Bayesian multi-state mark–recapture model described here (in both its hierarchical and non-hierarchical variations) performed well across a range of data quality, yet estimation bias was high when low survivorship was combined with low sample size or low capture probability. Consequently, if survivorship is unknown (it is often a

parameter of interest to be estimated) or expected to be low, effort should be made to increase both the number of individuals marked and the probability of capture (Salewski *et al.* 2007). At stopover, for instance, where capture probabilities are typically very low (Morris *et al.* 2006; Calvert *et al.* 2009), this might be accomplished by improving the placement or number of nets within the available habitat (e.g. Bachler & Schaub 2007) or by employing additional markers such as colour rings to increase the probability of re-observation.

Our application of hierarchical Bayes to multi-state mark–recapture studies is structurally flexible, amenable to the analyses of complex and unusual data at various spatial scales, and deals with sparseness by borrowing strength across all available data sets. Its structuring of parameter variation across time or space will therefore complement existing tools for the modelling of real ecological mark–recapture data that are critical to demographic quantifications in wildlife conservation and management.

Acknowledgements

References

Alerstam, T. & Lindstrom, A. (1990) Optimal bird migration: the relative importance of time, energy and safety. *Bird Migration: Physiology and Ecophysiology* (ed. E. Gwinner), pp. 331–351. Springer-Verlag, Berlin.

Arnason, A.N. (1972) Parameter estimates from mark–recapture experiments on two populations subject to migration and death. *Researches on Population Ecology*, **13**, 97–113.

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

Askey, P.J., Post, J.R., Parkinson, E.A., Rivot, E., Paul, A.J & Biro, P.A. (2007) Estimation of gillnet efficiency and selectivity across multiple sampling units: a hierarchical Bayesian analysis using mark–recapture data. *Fisheries Research*, **83**, 162–174.

Bachler, E. & Schaub, M. (2007) The effects of permanent local emigration and encounter technique on stopover duration estimates as revealed by telemetry and mark–recapture. *Condor*, **109**, 142–152.

Bonner, S.J. & Schwarz, C.J. (2006) An extension of the Cormack–Jolly–Seber model for continuous covariates with application to *Microtus pennsylvanicus*. *Biometrics*, **62**, 142–149.

Brownie, C., Hines, J.E., Nichols, J.D., Pollock, K.H. & Hestbeck, J.B. (1993) Capture–recapture studies for multiple strata including non-Markovian transitions. *Biometrics*, **49**, 1173–1187.

Burnham, K.P. & Anderson, D.R. (2002) *Model Selection and Multi-model Inference: A Practical Information–Theoretic Approach*. Springer-Verlag, New York.

Calvert, A.M. & Gauthier, G. (2005). Effects of exceptional conservation measures on survival and seasonal hunting mortality in greater snow geese. *Journal of Applied Ecology*, **42**, 442–452.

Calvert, A.M., Taylor, P.D. & Walde, S. (2009) Cross-scale environmental influences on migratory stopover behaviour. *Global Change Biology*, **15**, 744–759.

Carlin, B.P. & Louis, T.A. (1996) *Bayesian Methods for Data Analysis*. Chapman & Hall, London.

Caswell, H. & Fujiwara, M. (2004) Beyond survival estimation: mark–recapture, matrix population models, and population dynamics. *Animal Biodiversity and Conservation*, **27**, 471–488.

Clark, J.S. & Gelfand, A.E. (2006) A future for models and data in environmental science. *Trends in Ecology & Evolution*, **21**, 375–380.

Clark, J.S. (2005) Why environmental scientists are becoming Bayesians. *Ecology Letters*, **8**, 2–14.

Clark, J.S., Ferraz, G., Oguge, N., Hays, H. & DiCostanzo, J. (2005) Hierarchical Bayes for structured, variable populations: from recapture data to life-history prediction. *Ecology*, **86**, 2232–2244.

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

Dupuis, J.A. & Schwarz, C.J. (2007) A Bayesian approach to the multistate Jolly–Seber capture–recapture model. *Biometrics*, **63**, 1015–1022.

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

Efron, B. & Morris, C. (1975) Data analysis using Stein's estimator and its generalizations. *Journal of the American Statistical Association*, **70**, 311–319.

Ellison, A.M. (2004) Bayesian inference in ecology. *Ecology Letters*, **7**, 509–520.

Gelman, A. & Rubin, D.B. (1992) Inference from iterative simulation using multiple sequences. *Statistical Science*, **7**, 457–472.

Gelman, A., Carlin, J.B., Stern, H.S. & Rubin, D.B. (2004) *Bayesian Data Analysis*, 2nd edn. CRC Press, London.

Gimenez, O., Bonner, S., King, R., Parker, R.A., Brooks, S.P., Jamieson, L.E., Grosbois, V., Morgan, B.J.T. & Thomas, L. (2009) WinBUGS for population ecologists: Bayesian modeling using Markov Chain Monte Carlo methods. *Environmental and Ecological Statistics*, **3**, 883–915.

Gimenez, O., Rossi, V., Choquet, R., Dehais, C., Doris, B., Varella, H., Vila, J.-P. & Pradel, R. (2007) State-space modelling of data on marked individuals. *Ecological Modelling*, **206**, 431–438.

Harwood, J. & Stokes, K. (2003) Coping with uncertainty in ecological advice: lessons from fisheries. *Trends in Ecology & Evolution*, **18**, 617–622.

Hestbeck, J.B., Nichols, J.D. & Malecki, R.A. (1991) Estimates of movement and site fidelity using mark–resight data of wintering Canada geese. *Ecology*, **72**, 523–533.

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

Jonsen, I.D., Myers, R.A. & Flemming, J.M. (2003) Meta-analysis of animal movement using state-space models. *Ecology*, **84**, 3055–3063.

Jonsen, I.D., Myers, R.A. & James, M.C. (2006) Robust hierarchical state-space models reveal diel variation in travel rates of migrating leatherback turtles. *Journal of Animal Ecology*, **75**, 1046–1057.

Kendall, W.L., Pollock, K.H. & Brownie, C. (1995) A likelihood-based approach to capture–recapture estimation of demographic parameters under the robust design. *Biometrics*, **51**, 293–308.

Kéry, M. & Royle, J.A. (2008) Hierarchical Bayes estimation of species richness and occupancy in spatially replicated surveys. *Journal of Applied Ecology*, **45**, 589–598.

King, R. & Brooks, S.P. (2002) Bayesian model discrimination for multiple strata capture–recapture data. *Biometrika*, **89**, 785–806.

Lebreton, J.D. & Pradel, R. (2002) Multistate recapture models: modelling incomplete individual histories. *Journal of Applied Statistics*, **29**, 353–369.

Lebreton, J.-D., Burnham, K.P., Clobert, J. & 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. & Barker, R.J. (2005). Modeling association among demographic parameters in analysis of open population capture–recapture data. *Biometrics*, **61**, 46–54.

Link, W.A., Cam, E., Nichols, J.D. & Cooch, E.G. (2002) Of BUGS and birds: Markov chain Monte Carlo for hierarchical modelling in wildlife research. *Journal of Wildlife Management*, **66**, 277–291.

Lunn, D.J., Thomas, A., Best, N. & Spiegelhalter, D. (2000) WinBUGS – a Bayesian modelling framework: concepts, structure, and extensibility. *Statistics and Computing*, **10**, 325–337.

McCarthy, M.A. & Masters, P. (2005) Profiting from prior information in Bayesian analyses of ecological data. *Journal of Applied Ecology*, **42**, 1012–1019.

Morris, S.R., Larracuente, A.M., Covino, K.M., Mustillo, M.S., Mattern, K.E., Liebner, D.A. & Sheets, H.D. (2006) Utility of open population models: limitations posed by parameter estimability in the study of migratory stopover. *Wilson Journal of Ornithology*, **118**, 513–526.

Nichols, J.D., Hines, J.E., Pollock, K.H., Hinz, R.L. & Link, W.A. (1994) A methodological framework for estimating breeding proportions and testing biological hypotheses about the cost of reproduction with capture–recapture data. *Ecology*, **75**, 2052–2065.

O'Brien, S., Robert, B. & Tiandry, H. (2005) Consequences of violating the recapture duration assumption of mark–recapture models: a test using simulated and empirical data from an endangered tortoise population. *Journal of Applied Ecology*, **42**, 1096–1104.

Pollock, K.H. (1982) A capture–recapture design robust to unequal probability of capture. *Journal of Wildlife Management*, **46**, 752–757.

Pollock, K.H., Nichols, J.D., Brownie, C. & Hines, J.E. (1990) Statistical inference for capture–recapture experiments. *Wildlife Monographs*, **107**, 1–97.

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

Pradel, R. (1996) Utilization of capture–mark–recapture for the study of recruitment and population growth rate. *Biometrics*, **52**, 703–709.

Royle, J.A. & Dorazio, R.M. (2008) *Hierarchical Modeling and Inference in Ecology*. Academic Press, San Diego, CA, USA.

Royle, J.A. & Kéry, M. (2007) A Bayesian state-space formulation of dynamic occupancy models. *Ecology*, **88**, 1813–1823.

Salewski, V. & Schaub, M. (2007) Stopover duration of Palearctic passerine migrants in the western Sahara – independent of fat stores? *Ibis*, **149**, 223–236.

Salewski, V., Thoma, M. & Schaub, M. (2007) Stopover of migrating birds: simultaneous analysis of different marking methods enhances the power of capture–recapture analyses. *Journal of Ornithology*, **148**, 29–37.

Schaub, M. (2006) How to study departure decisions of migrants from stopover sites using capture–recapture data. *Acta Zoologica Sinica*, **52**, 602–605.

Schaub, M., Liechti, F. & Jenni, L. (2004) Departure of migrating European robins, *Erithacus rubecula*, from a stopover site in relation to wind and rain. *Animal Behaviour* **67**, 229–237.

Schwarz, C.J. & Arnason, A.N. (2000) Estimation of age-specific breeding probabilities from capture–recapture data. *Biometrics*, **56**, 59–64.

Schwarz, C.J. & Stobo, W.T. (1997) Estimating temporary migration using the robust design. *Biometrics*, **53**, 178–194.

Schwarz, C.J., Schweigert, J.F. & 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.

Spiegelhalter, D.J., Best, N.G., Carlin, B.P. & van der Linde, A. (2002) Bayesian measures of model complexity and fit. *Journal of the Royal Statistical Society, Series B*, **64**, 1–34.

Thomas, A., O'Hara, B., Ligges, U. & Stutz, S. (2006) Making BUGS Open. *R News*, **6**, 12–17.

White, G.C. & Burnham, K.P. (1999) Program MARK: survival estimation from populations of marked animals. *Bird Study*, **46**, S120–S138.

White, G.C., Kendall, W.L. & Barker, R.J. (2006) Multistate survival models and their extensions in program MARK. *Journal of Wildlife Management*, **70**, 1521–1529.

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

Zheng, C., Ovaskainen, O., Saastamoinen, M. & Hanski, I. (2007) Age-dependent survival analyzed with Bayesian models of mark–recapture data. *Ecology*, **88**, 1970–1976.

Supporting Information

Additional Supporting Information may be found in the online version of this article:

Appendix S1. BUGS code for a general 2-state hierarchical model (S1-A) and the specific hierarchical migratory stopover model (S1-B).

Appendix S2. True values used to simulate data, and parameters estimated in BUGS, for all three components of the simulation study.

Appendix S3. Estimation bias of movement, survivorship and capture probabilities from the 'poor data' simulation study, as estimated with the hierarchical BUGS model.

Appendix S4. Discussion of alternate hyperprior distributions for parameter mean μ and precision τ.

Please note: Wiley-Blackwell are not responsible for the content or functionality of any supporting materials supplied by the authors. Any queries (other than missing material) should be directed to the corresponding author for the article.