Appendix A: Details of Bayesian analysis for bilateral photo-identification data

Here we provide details necessary to implement Markov chain Monte Carlo (MCMC) sampling for our integrated mark-recapture model utilizing bilateral photo-ID records. Specifically, we provide details for the closed population model M_t (which permits time-dependent capture probabilities; e.g., Otis et al. 1978). We give details for three different model formulations: 1) one that accounts for bilateral photographs via a latent multinomial; 2) a "traditional" approach limiting records to right-sided or left-sided encounters only; and 3) the approach suggested by Corkrey et al. (2008). As all analyses share a number of features, we have organized this appendix in the following manner. First, we specify prior distributions that will be common to all models. Second, we provide posterior distributions (up to a proportionality constant) and full conditional distributions needed to perform MCMC sampling, including the complete sampling algorithm used for our bilateral latent multinomial model.

1. Prior distributions

Similar to Link et al. (2010), the following priors were assumed for all analyses:

$$p_t \sim \text{Beta}(a_0^t, b_0^t),$$

$$\boldsymbol{\delta} = \left(\delta^L, \delta^R, \delta^B\right) \sim \text{Dirichlet}(c_0^L, c_0^R, c_0^B),$$

$$N \sim \text{Uniform}(1, N^*).$$

We set $a_0^t, b_0^t, c_0^L, c_0^R,$ and c_0^B equal to 1 to yield prior distributions that are uniform over all points of support. We note that a logit link function (for $\mathbf{p}$) or a multinomial logit link function (for $\boldsymbol{\delta}$) could provide additional flexibility for increasing the complexity of the model (e.g., when modeling variation in these parameters as a function of multiple predictors in an analogous fashion to logistic regression).

The prior for N is uniform with a large upper limit $N^* >> N$. Note that in some cases we actually work with the parameter f_0, the number of all zero encounter histories,

rather than N. In such cases, the number of uniquely identified animals, n, is also used in posterior calculations.

2. Posterior and full conditional distributions

Case 1: Model M_t, latent multinomial model

In this case, the posterior distribution of model parameters is specified up to a proportionality constant:

$$[\mathbf{x}, \mathbf{p}, \boldsymbol{\delta} \mid \mathbf{f}] \propto [\mathbf{x} \mid \mathbf{p}, \boldsymbol{\delta}, N] \mathrm{I}(\mathbf{A}'\mathbf{x} = \mathbf{f})[\mathbf{p}][\boldsymbol{\delta}]. \tag{A.1}$$

Here, $\mathrm{I}(\mathbf{A}'\mathbf{x} = \mathbf{f})$ is an indicator function that takes on a value of 1 if $\mathbf{A}'\mathbf{x} = \mathbf{f}$ and 0 otherwise (that is, only combinations of latent histories which match the observed data receive positive support). Also note that abundance is calculated as $N = \sum_i x_i$ but does not appear explicitly in the posterior. Since sampling is with replacement, a multinomial distribution is often used for modeling $[\mathbf{x} \mid \mathbf{p}, \boldsymbol{\delta}, N]$:

$$[\mathbf{x} \mid \mathbf{p}, \boldsymbol{\delta}, N] = \frac{N!}{\prod_i x_i!} \prod_i \pi_i^{x_i}.$$

Here, π_i gives the probability of latent encounter history i, and is calculated as

$$\pi_i = \prod_{t=1}^{T} p_t^{\mathrm{I}(\omega_{it} \neq 0)} (1 - p_t)^{\mathrm{I}(\omega_{it}=0)} (\delta^L)^{\mathrm{I}(\omega_{it}=L)} (\delta^R)^{\mathrm{I}(\omega_{it}=R)} (\delta^B)^{\mathrm{I}(\omega_{it}=B)}$$

(see Eq. 6 of Link et al. (2010) for a similar calculation for misidentification errors). We update using the full conditional distributions for each parameter θ, represented here by $[\theta \mid \cdot]$:

$$[\mathbf{x} \mid \cdot] \propto \mathrm{I}(\mathbf{A}'\mathbf{x} = \mathbf{f}) \frac{N!}{\prod_{i'} x_{i'}!} \prod_{i'} \pi_{i'}^{x_{i'}},$$

$$[p_t \mid \cdot] \sim \mathrm{Beta}(a^t, b^t),$$

and

$$[\boldsymbol{\delta} \mid \cdot] \sim \mathrm{Dirichlet}(c^L, c^R, c^B),$$

where

$$a^t = a_0^t + \sum_i x_i \mathrm{I}(\omega_{it} \neq 0),$$

$$b^t = b_0^t + \sum_i x_i \mathrm{I}(\omega_{it} = 0),$$

$$c^L = c_0^L + \sum_i x_i \sum_t \mathrm{I}(\omega_{it} = L),$$

$$c^R = c_0^R + \sum_i x_i \sum_t \mathrm{I}(\omega_{it} = R),$$

$$c^B = c_0^B + \sum_i x_i \sum_t \mathrm{I}(\omega_{it} = B),$$

and i' references all non-zero entries of the particular basis vector that is being used to update $\mathbf{x}$.

We now describe the MCMC sampling algorithm used for drawing from the posterior of our bilateral latent multinomial model M_t (Eq. A.1). Similar to Link et al. (2010), the Gibbs sampling procedure goes as follows:

1. Initialize all parameters, including an initial $\mathbf{x}$ from the feasible set of latent history frequencies. Start the chain at iteration $g = 1$.

2. Calculate a^t, b^t, c^L, c^R, and c^B for $t = 1, \ldots, T$ based on the current value $\mathbf{x}^{(g)}$.

3. Update $\mathbf{p}^{(g+1)}$ and $\boldsymbol{\delta}^{(g+1)}$ from their full conditional distributions.

4. Randomly draw j without replacement from the integer set $\{1, \ldots, r\}$, where r is the number of basis vectors spanning the null space of $\mathbf{A}'$. Then draw q_j from a discrete uniform distribution over the integers $\{-D_j, \ldots, -1, 1, \ldots, D_j\}$ and propose a candidate latent frequency

$$\mathbf{x}^* = \mathbf{x}^{(g)} + q_j \mathbf{v}_j,$$

where $\mathbf{v}_j$ is the jth basis vector of the null space of $\mathbf{A}'$. The tuning parameters D_j are chosen by the user.

5. Accept $\mathbf{x}^*$ (i.e., set $\mathbf{x}^{(g+1)} = \mathbf{x}^*$) with probability $R = \min\left(1, \frac{[\mathbf{x}^*|\cdot]}{[\mathbf{x}^{(g)}|\cdot]}\right)$.

6. Repeat steps 4 and 5 until a candidate latent frequency has been proposed from all of the r basis vectors.

7. Increment g by 1 and return to step 2.

For all analyses, we initialized the tuning parameters $D_1 = 10$ and $D_j = 1$ for $j = 2, \ldots, r$. During tuning periods, we adapted D_j at every 10th iteration of the chain by multiplying D_j by 0.95 if the acceptance rate was below 0.3 and dividing by 0.95 if the acceptance rate was above 0.4.

Case 2: Model M_t, 1-sided photographs

When data from one side of the animal are modeled, the problem reduces to a straightforward analysis of model M_t. Although generalized formulations for Bayesian implementation of this model are available (e.g., Smith 1991), here we adopt a similar formulation to Case 1 (above) to promote comparability of results. In particular, let $\mathbf{H}_i$ give the encounter history for animal i, whose entries h_{it} take on a value of 1 if individual i is detected at time t and 0 otherwise. In this case, the posterior is proportional to

$$[f_0, \mathbf{p} \mid \mathbf{H}, n] \propto [\mathbf{H} \mid f_0, \mathbf{p}, n][\mathbf{p}][f_0]. \tag{A.2}$$

As with N, the prior $[f_0]$ is uniform and thus may be omitted. The likelihood component $[\mathbf{H} \mid f_0, \mathbf{p}, n]$ is proportional to

$$\frac{(n + f_0)!}{f_0!}\left\{\prod_{i=1}^{n} \pi_i\right\} \pi_0^{f_0}.$$

Here, π_i gives the probability of observing encounter history i and is specified as

$$\pi_i = \prod_t p_t^{h_{it}}(1 - p_t)^{1-h_{it}}.$$

Likewise, π_0 gives the probability of never being encountered:

$$\pi_0 = \prod_t (1 - p_t).$$

Full conditional distributions are then obtained via factorization of Eq. A.2 to only include elements that change as a function of the parameter of interest, and are given as

$$[f_0 \mid \cdot] \propto \frac{(n + f_0)!}{f_0!} \pi_0^{f_0}$$

and

$$[p_t \mid \cdot] \sim \mathrm{Beta}(a^t, b^t),$$

where

$$a^t = a_0^t + \sum_i \mathrm{I}(h_{it} = 1)$$

and

$$b^t = b_0^t + f_0 + \sum_i \mathrm{I}(h_{it} = 0).$$

Case 3: Model M_t, approach of Corkrey et al. (2008)

To implement M_t for bilateral photographs using the formulation based on Corkrey et al. (2008), the posterior is first specified as

$$[f_0, \mathbf{p}_t^L, \mathbf{p}_t^R \mid \mathbf{H}, n'] \propto [\mathbf{H} \mid f_0, \mathbf{p}_t^L, \mathbf{p}_t^R, n'][\mathbf{p}_t^L][\mathbf{p}_t^R][f_0].$$

Here, n' gives the total number of recorded histories, which are treated as if they reference a unique animal. The priors for left- (p_t^L) and right-side (p_t^R) detection probabilities are assigned similarly for p_t in cases 1 and 2, with $p_t^L \sim \mathrm{Beta}(a_0^L, b_0^L)$ and $p_t^R \sim \mathrm{Beta}(a_0^R, b_0^R)$ where a_0^L, b_0^L, a_0^R, and b_0^R are set equal to 1. The likelihood component $[\mathbf{H} \mid f_0, \mathbf{p}_t^L, \mathbf{p}_t^R, n']$ is then proportional to

$$\frac{(n' + f_0)!}{f_0!} \left\{ \prod_{i=1}^{n'} \pi_i \right\} \pi_0^{f_0},$$

where π_i is calculated as

$$\prod_t \left\{ p_t^R p_t^L \right\}^{I(h_{it}=B)} \left\{ p_t^R (1 - p_t^L) \right\}^{I(h_{it}=R)} \left\{ p_t^L (1 - p_t^R) \right\}^{I(h_{it}=L)} \left\{ (1 - p_t^R)(1 - p_t^L) \right\}^{I(h_{it}=0)}$$

and

$$\pi_0 = \prod_t (1 - p_t^L)(1 - p_t^R).$$

Full conditional distributions are then

$$[f_0 \mid \cdot] \propto \frac{(n' + f_0)!}{f_0!} \pi_0^{f_0},$$

$$[p_t^L \mid \cdot] \sim \text{Beta}(a_t^L, b_t^L),$$

and

$$[p_t^R \mid \cdot] \sim \text{Beta}(a_t^R, b_t^R),$$

where

$$a_t^L = a_0^L + \sum_i \text{I}(h_{it} = L) + \text{I}(h_{it} = B),$$

$$b_t^L = b_0^L + f_0 + \sum_i \text{I}(h_{it} = 0) + \text{I}(h_{it} = R),$$

$$a_t^R = a_0^R + \sum_i \text{I}(h_{it} = R) + \text{I}(h_{it} = B),$$

and

$$b_t^R = b_0^R + f_0 + \sum_i \text{I}(h_{it} = 0) + \text{I}(h_{it} = L).$$

LITERATURE CITED

Corkrey, R., S. Brooks, D. Lusseau, K. Parsons, J. W. Durban, P. S. Hammond, and P. M. Thompson. 2008. A Bayesian capture-recapture population model with simultaneous estimation of heterogeneity. Journal of the American Statistical Association **103**:948–960.

Link, W. A., J. Yoshizaki, L. L. Bailey, and K. H. Pollock. 2010. Uncovering a latent multinomial: analysis of mark-recapture data with misidentification. Biometrics **66**:178–185.

Otis, D. L., K. P. Burnham, G. C. White, and D. R. Anderson. 1978. Statistical-inference from capture data on closed animal populations. Wildlife Monographs **62**:7–135.

Smith, P. J. 1991. Bayesian analyses for a multiple capture-recapture model. Biometrika **78**:399–407.