The Annals of Applied Statistics
2014, Vol. 8, No. 4, 2461–2484
DOI: 10.1214/14-AOAS783
In the Public Domain

AND MISIDENTIFICATION[1]

BY BRETT T. MCCLINTOCK*, LARISSA L. BAILEY†, BRIAN P. DREHER‡ AND WILLIAM A. LINK§

NOAA National Marine Mammal Laboratory, Colorado State University†, Colorado Parks and Wildlife‡, and USGS Patuxent Wildlife Research Center§*

As noninvasive sampling techniques for animal populations have become more popular, there has been increasing interest in the development of capture–recapture models that can accommodate both imperfect detection and misidentification of individuals (e.g., due to genotyping error). However, current methods do not allow for individual variation in parameters, such as detection or survival probability. Here we develop misidentification models for capture–recapture data that can simultaneously account for temporal variation, behavioral effects and individual heterogeneity in parameters. To facilitate Bayesian inference using our approach, we extend standard probit regression techniques to latent multinomial models where the dimension and zeros of the response cannot be observed. We also present a novel Metropolis–Hastings within Gibbs algorithm for fitting these models using Markov chain Monte Carlo. Using closed population abundance models for illustration, we re-visit a DNA capture–recapture population study of black bears in Michigan, USA and find evidence of misidentification due to genotyping error, as well as temporal, behavioral and individual variation in detection probability. We also estimate a salamander population of known size from laboratory experiments evaluating the effectiveness of a marking technique commonly used for amphibians and fish. Our model was able to reliably estimate the size of this population and provided evidence of individual heterogeneity in misidentification probability that is attributable to variable mark quality. Our approach is more computationally demanding than previously proposed methods, but it provides the flexibility necessary for a much broader suite of models to be explored while properly accounting for uncertainty introduced by misidentification and imperfect detection. In the absence of misidentification, our probit formulation also provides a convenient and efficient Gibbs sampler for Bayesian analysis of traditional closed population capture–recapture data.

Received December 2013; revised May 2014.

[1]Funding for the bear example was provided in part by Federal Aid in Wildlife Restoration Project W-147-R-2, Michigan Department of Natural Resources, United States Fish and Wildlife Service, and Michigan State University.

Key words and phrases. Data augmentation, individual heterogeneity, latent multinomial, mark-recapture, missing data, population size, probit regression, record linkage.

1. Introduction. Capture–recapture methods are commonly used to estimate demographic parameters for wildlife [e.g., Williams, Nichols and Conroy (2002)] and human [e.g., Yip et al. (1995a, 1995b)] populations. Passive (or "noninvasive") sampling techniques are becoming more common in capture–recapture studies, largely because these techniques can be less expensive and less invasive than the physical capture of animals [e.g., Karanth and Nichols (1998), Mackey et al. (2008), Ruell et al. (2009)]. Passive sampling techniques in capture–recapture studies include the use of photographs [Karanth and Nichols (1998), Langtimm et al. (1998), Mackey et al. (2008)], visual sightings [e.g., Hall, McConnell and Barker (2001), Kauffman, Frick and Linthicum (2003)] or genetic material [Dreher et al. (2007), Ruell et al. (2009)] to individually identify animals. When individual animals are identifiable by natural or artificial marks, these techniques can provide information about key demographic parameters such as abundance, survival and recruitment. They are therefore very useful for informing management decisions, as well as for testing ecological or evolutionary hypotheses.

Unfortunately, use of passive sampling techniques in capture–recapture studies is not entirely without problems. For example, matching photographs to individuals can be prone to identification error due to variable image quality [e.g., Bonner and Holmberg (2013), Hastings, Hiby and Small (2008), Link et al. (2010), McClintock et al. (2013a), Morrison et al. (2011)], and genetic samples (e.g., scat or hair) are susceptible to genotyping error [e.g., Dreher et al. (2007), Lukacs and Burnham (2005), Wright et al. (2009)]. Individual identifications from photographs, visual sightings or genetic samples are all susceptible to observer recording error. Sampling designs can also result in differential exposures of individuals to sampling (e.g., due to home range behavior or opportunistic sampling). Such individual heterogeneity in detection probabilities can severely bias estimators and is a common culprit in the underestimation of abundance in capture–recapture studies.

Link et al. (2010) recently developed a novel approach for the analysis of capture–recapture data when individual identification errors occur. This pioneering contribution focused on the closed population abundance model allowing for temporal variation in parameters [Darroch (1958), Otis et al. (1978)], and therefore does not accommodate individual-level variation in parameters, such as detection [e.g., Basu and Ebrahimi (2001), Coull and Agresti (1999), Fienberg, Johnson and Junker (1999), King and Brooks (2008), Manrique-Vallier and Fienberg (2008), Pledger (2000)] or survival [e.g., Gimenez and Choquet (2010), Royle (2008)] probability. Here, we develop models to simultaneously account for temporal variation, behavioral response (e.g., trap "happy" or "shy" effects), individual heterogeneity and misidentification in capture–recapture analyses. To facilitate Bayesian inference using our approach, we also extend standard probit regression data augmentation techniques [e.g., Albert and Chib (1993)] to latent multinomial models where the dimension and zeros of the response cannot be observed.

2. Methods.

2.1. *Detailed problem description.* Consider a "classic" capture–recapture study, where sampling is conducted over T sampling occasions and the identity of each animal is known with certainty when it is observed (i.e., there is no misidentification). When encounters are simple binary responses and $T = 2$, there are three possible recorded encounter histories for each individual: "11" (encountered on both occasions), "10" (encountered on the first occasion but not the second) and "01" (encountered on the second occasion but not the first). If the encounter history for animal i is denoted $\mathbf{h}_i$, a classic approach is to assume that $\mathbf{h}_i$ is a realization from a multinomial process, where the probability of observing $\mathbf{h}_i$ is a function of unknown demographic parameters ($\boldsymbol{\theta}$) and (usually nuisance) parameters related to the observation process ($\boldsymbol{\rho}$). For example, $\boldsymbol{\theta}$ might consist of survival probabilities and $\boldsymbol{\rho}$ of detection probabilities. In this case, the number of unique animals encountered (n) is known with certainty, and when conditioning on first capture, a standard likelihood for capture–recapture data is proportional to

$$[\mathbf{h}|\boldsymbol{\theta}, \boldsymbol{\rho}] = \prod_{i=1}^{n} \Pr(\mathbf{h}_i|\boldsymbol{\theta}, \boldsymbol{\rho}), \tag{1}$$

where $[\mathbf{h}|\boldsymbol{\theta}, \boldsymbol{\rho}]$ denotes the conditional distribution for $\mathbf{h}$ given $\boldsymbol{\theta}$ and $\boldsymbol{\rho}$. We note that "00" encounter histories are not observed, hence, additional modifications to equation (1) are needed to make inferences about individuals that are never encountered.

In contrast to the preceding scenario, now consider the situation where individuals may be misidentified. When such errors can occur, three types of encounters for any of the T sampling occasions are possible. These include a nonencounter (denoted by "0"), a correctly identified encounter (denoted by "1") or a misidentified encounter (denoted by "2"). Misidentified encounters result in "ghost" encounter histories [Link et al. (2010), Yoshizaki (2007)], and an individual encountered in >1 sampling occasion could therefore yield a number of possible recorded histories. For example, when presented with the recorded histories "10" and "01," we do not know whether these observations arose from the same animal seen on both occasions (latent histories "12," "21" or "22") or whether it was indeed two different animals each seen on one occasion (latent histories "10" and "01," "20" and "01," "10" and "02," or "20" and "02"). Under misidentification, encounter histories are not uniquely associated with animals, so equation (1) is no longer valid for making inferences about $\boldsymbol{\theta}$ and $\boldsymbol{\rho}$.

Assuming the same misidentification cannot occur more than once (i.e., a ghost cannot be detected more than once) and an encounter cannot be misidentified as a legitimate marked individual, Link et al. (2010) proposed a closed population abundance model allowing for temporal variation in detection probability under this misidentification scenario. In the next section, we generalize their approach to a much broader suite of misidentification models that can simultaneously accommodate temporal, behavioral and individual effects on $\boldsymbol{\theta}$ and $\boldsymbol{\rho}$.

2.2. *Accounting for individual heterogeneity and misidentification.* Consider the marginal likelihood obtained by summing the "complete data likelihood" over all possible values of the latent encounter histories:

$$[\mathbf{f}|\boldsymbol{\theta}, \boldsymbol{\rho}] = \sum_{\mathbf{h}} [\mathbf{h}|\boldsymbol{\theta}, \boldsymbol{\rho}][\mathbf{f}|\mathbf{h}, \boldsymbol{\theta}, \boldsymbol{\rho}], \tag{2}$$

where $\mathbf{f}$ is a vector of recorded history frequencies indicating the number of times each of the possible recorded histories was observed (see Table 1 for notation definitions). The complete data likelihood therefore derives from distributions $[\mathbf{h}|\boldsymbol{\theta}, \boldsymbol{\rho}]$ for latent capture–recapture data and distributions $[\mathbf{f}|\mathbf{h}, \boldsymbol{\theta}, \boldsymbol{\rho}]$ describing their conversion to observed (potentially misidentified) data (see Table 2). We note that this extension is applicable to all sorts of capture–recapture models [e.g., those reviewed by Williams, Nichols and Conroy (2002)] and could apply to data subject to errors other than misidentification [e.g., incomplete mark observations sensu McClintock et al. (2013b)]. Evaluating equation (2) involves a multidimensional summation, thus making maximum likelihood estimation difficult. Link et al. (2010) averted this problem by adopting a Bayesian perspective and sampling from the posterior distribution using Markov chain Monte Carlo (MCMC), but their approach requires the assumption of no individual variation in $\boldsymbol{\theta}$ and $\boldsymbol{\rho}$.

We will for convenience refer to the latent and recorded histories using indices. With three possible latent encounter types (0, 1 and 2), the latent history for individual i, $\mathbf{h}_i = (h_{i1}, h_{i2}, \dots, h_{iT})$, is identified by

$$j = 1 + \sum_{t=1}^{T} h_{it} 3^{t-1},$$

such that $H_i = j$ indicates individual i has latent encounter history j. For example, $H_i = 16$ for $T = 3$ indicates individual i has latent history $\mathbf{h}_i = 021$, $H_i > 1$ indicates individual i was encountered at least once, and $\Pr(H_i = j)$ is the probability that individual i has latent history j. Similarly, a binary recorded history $\boldsymbol{\omega} = (\omega_1, \omega_2, \dots, \omega_T)$ is identified by

$$k = \sum_{t=1}^{T} \omega_t 2^{t-1},$$

such that f_k is the observed frequency of recorded history k.

To implement our method, it is necessary to construct a matrix $\mathbf{A}$, such that $\mathbf{f} = \mathbf{A}'\mathbf{x}$, where the latent history frequency vector $\mathbf{x}$ has elements $x_j = \sum_i \mathrm{I}(H_i = j)$ indicating the number of individuals with latent history j, and $\mathrm{I}(H_i = j)$ is an indicator function having the value 1 when $H_i = j$ and 0 otherwise. The matrix $\mathbf{A}$ formally describes the relationship between the recorded and latent histories, and intuition about how $\mathbf{A}$ is constructed is best provided through a simple example. Suppose $T = 3$ for binary (i.e., detection, nondetection) recorded histories as in Table 2. The $3^T \times (2^T - 1)$ matrix $\mathbf{A}$ for this example can be constructed from the

TABLE 1

Definitions of parameters, latent variables, data and modeling constructs used in the latent multinomial model allowing misidentification with temporal, behavioral and individual-level variation in parameters. Note that bold symbols represent collections (vectors) of parameters

Parameters	**Definition**
$\boldsymbol{\theta}$	Vector of demographic process parameters (e.g., abundance or survival probability).
$\boldsymbol{\rho}$	Vector of observation process parameters (e.g., encounter or misidentification probability).
p_{it}	Probability that individual i is encountered at time t.
α	Probability that an individual, encountered at time t, is correctly identified.
Latent variables	
$\mathbf{h}_i$	The latent encounter history for individual i, $(h_{i1}, h_{i2}, \ldots, h_{iT})$.
h_{it}	Encounter type for the latent encounter history of individual i at time t; $h_{it} = 0$ represents no encounter, $h_{it} = 1$ a correctly identified encounter, and $h_{it} = 2$ a misidentified encounter.
H_i	Latent encounter history index for individual i, such that $H_i = j$ indicates individual i has latent history j. For $h_{it} \in \{0, 1, 2\}$ the 3^T possible latent histories are identified by $j = 1 + \sum_{t=1}^{T} h_{it} 3^{t-1}$ (see Table 2).
x_j	Latent frequency of encounter history j, where $x_j = \sum_i \mathrm{I}(H_i = j)$. Note that $\mathbf{x}$ denotes a column vector of such frequencies, for example, $\mathbf{x} = (x_1, x_2, \ldots, x_{3^T})'$ for $h_{it} \in \{0, 1, 2\}$.
Data	
T	Number of sampling occasions.
f_k	Frequency for recorded (observed) encounter history k. Note that $\mathbf{f}$ denotes a column vector of such frequencies, for example, $\mathbf{f} = (f_1, f_2, \ldots, f_{2^T-1})'$ for $\omega_t \in \{0, 1\}$.
Modeling constructs	
$\boldsymbol{\omega}$	Recorded encounter history, $(\omega_1, \omega_2, \ldots, \omega_T)$.
ω_t	Observation type for a recorded history at time t; $\omega_t = 0$ represents no detection and $\omega_t = 1$ a detection. For $\omega_t \in \{0, 1\}$ the $2^T - 1$ possible recorded histories are identified by $k = \sum_{t=1}^{T} \omega_t 2^{t-1}$ (see Table 2).
C_i	Occasion of first capture for individual i. For example, $C_i = 3$ if individual i has latent encounter history $\mathbf{h}_i = 0021$ ($H_i = 46$).

corresponding contributed records column in Table 2 by simply replacing each dot (.) with a 0 and any other entry with a 1. Thus, the rows of $\mathbf{A}$ correspond to the 3^T possible latent encounter histories and the columns correspond to the $2^T - 1$ possible recorded histories. For example, the sixth row of $\mathbf{A}$ indicates that latent

TABLE 2

Latent and recorded histories from marked individual encounters with $T = 3$ sampling occasions subject to misidentification. The probability of each latent history for individual i, $\Pr(H_i = j)$, is for a closed population abundance model, where p_{it} is the probability that individual i is encountered at time t, and α is the probability that an individual, encountered at time t, is correctly identified. Contributed records column shows the recorded histories (k) arising from specific latent histories (j). For example, latent history 25, "022," gives rise to recorded histories "010" and "001" (for which $k = 2$ and 4)

j	Latent history ($\mathbf{h}_i$)	$\Pr(H_i = j)$	Contributed records (k from j)	k	Recorded history (ω)
1	000	$(1-p_{i1})(1-p_{i2})(1-p_{i3})$	.....	1	100
2	100	$p_{i1}\alpha(1-p_{i2})(1-p_{i3})$	1.....	2	010
3	200	$p_{i1}(1-\alpha)(1-p_{i2})(1-p_{i3})$	1.....	3	110
4	010	$(1-p_{i1})p_{i2}\alpha(1-p_{i3})$	.2.....	4	001
5	110	$p_{i1}\alpha p_{i2}\alpha(1-p_{i3})$	..3....	5	101
6	210	$p_{i1}(1-\alpha)p_{i2}\alpha(1-p_{i3})$	12.....	6	011
7	020	$(1-p_{i1})p_{i2}(1-\alpha)(1-p_{i3})$	.2.....	7	111
8	120	$p_{i1}\alpha p_{i2}(1-\alpha)(1-p_{i3})$	12.....		
9	220	$p_{i1}(1-\alpha)p_{i2}(1-\alpha)(1-p_{i3})$	12.....		
10	001	$(1-p_{i1})(1-p_{i2})p_{i3}\alpha$	...4...		
11	101	$p_{i1}\alpha(1-p_{i2})p_{i3}\alpha$	....5..		
12	201	$p_{i1}(1-\alpha)(1-p_{i2})p_{i3}\alpha$	1..4...		
13	011	$(1-p_{i1})p_{i2}\alpha p_{i3}\alpha$	.....6.		
14	111	$p_{i1}\alpha p_{i2}\alpha p_{i3}\alpha$	7		
15	211	$p_{i1}(1-\alpha)p_{i2}\alpha p_{i3}\alpha$	1....6.		
16	021	$(1-p_{i1})p_{i2}(1-\alpha)p_{i3}\alpha$	.2.4...		
17	121	$p_{i1}\alpha p_{i2}(1-\alpha)p_{i3}\alpha$	.2..5..		
18	221	$p_{i1}(1-\alpha)p_{i2}(1-\alpha)p_{i3}\alpha$	12.4...		
19	002	$(1-p_{i1})(1-p_{i2})p_{i3}(1-\alpha)$	...4...		
20	102	$p_{i1}\alpha(1-p_{i2})p_{i3}(1-\alpha)$	1..4...		
21	202	$p_{i1}(1-\alpha)(1-p_{i2})p_{i3}(1-\alpha)$	1..4...		
22	012	$(1-p_{i1})p_{i2}\alpha p_{i3}(1-\alpha)$	.2.4...		
23	112	$p_{i1}\alpha p_{i2}\alpha p_{i3}(1-\alpha)$	..34...		
24	212	$p_{i1}(1-\alpha)p_{i2}\alpha p_{i3}(1-\alpha)$	12.4...		
25	022	$(1-p_{i1})p_{i2}(1-\alpha)p_{i3}(1-\alpha)$	.2.4...		
26	122	$p_{i1}\alpha p_{i2}(1-\alpha)p_{i3}(1-\alpha)$	12.4...		
27	222	$p_{i1}(1-\alpha)p_{i2}(1-\alpha)p_{i3}(1-\alpha)$	12.4...		

history 210 ($j = 6$) gives rise to the recorded histories 100 ($k = 1$) and 010 ($k = 2$) for binary recorded histories when $T = 3$.

We treat the latent individual encounter histories as unobserved quantities (just like $\boldsymbol{\theta}$ and $\boldsymbol{\rho}$) and use Bayesian analysis methods to evaluate the joint posterior distribution

$$[\mathbf{h}, \boldsymbol{\theta}, \boldsymbol{\rho}|\mathbf{f}] \propto [\mathbf{h}|\boldsymbol{\theta}, \boldsymbol{\rho}][\mathbf{f}|\mathbf{h}, \boldsymbol{\theta}, \boldsymbol{\rho}][\boldsymbol{\theta}, \boldsymbol{\rho}], \tag{3}$$

where $[\mathbf{f}|\mathbf{h}, \boldsymbol{\theta}, \boldsymbol{\rho}] = \mathrm{I}(\mathbf{A}'\mathbf{x} = \mathbf{f})$. We note that $[\mathbf{f}|\mathbf{h}, \boldsymbol{\theta}, \boldsymbol{\rho}]$ does not depend on $\boldsymbol{\theta}$ or $\boldsymbol{\rho}$; the relation is deterministic rather than stochastic in the cases we consider here. One of the keys to sampling from equation (3) using MCMC is proposing latent history frequencies $\mathbf{x}$ that satisfy $\mathbf{A}'\mathbf{x} = \mathbf{f}$. This is accomplished by utilizing basis vectors for the null space of $\mathbf{A}'$. Once the $\mathbf{A}$ matrix is defined, a basis for the null space of $\mathbf{A}'$ can be determined by solving the system of equations $\mathbf{A}'\mathbf{x} = \mathbf{0}$. For binary recorded histories with $T = 2$, one such basis is the set of $3^T - 2^T + 1 = 6$ column vectors $\{\mathbf{v}\}$, where $\mathbf{v}_1 = (1, 0, 0, 0, 0, 0, 0, 0, 0)'$, $\mathbf{v}_2 = (0, -1, 1, 0, 0, 0, 0, 0, 0)'$, $\mathbf{v}_3 = (0, -1, 0, -1, 0, 1, 0, 0, 0)'$, $\mathbf{v}_4 = (0, 0, 0, -1, 0, 0, 1, 0, 0)'$, $\mathbf{v}_5 = (0, -1, 0, -1, 0, 0, 0, 1, 0)'$ and $\mathbf{v}_6 = (0, -1, 0, -1, 0, 0, 0, 0, 1)'$.

When there is no individual heterogeneity in parameters, one may propose and update $\mathbf{x}$ from the set of basis vectors without explicit consideration of $\mathbf{h}_i$ [Link et al. (2010)]. However, when allowing for individual heterogeneity, one must explicitly consider $\mathbf{h}_i$ for each individual in the population. An efficient MCMC algorithm therefore needs to regularly propose reasonable $\mathbf{h}_i$ in combinations that satisfy $\mathbf{A}'\mathbf{x} = \mathbf{f}$. As illustrated in Sections 2.2.1 and 2.2.2 for closed population abundance models, we accomplish this by apportioning each latent history frequency x_j to individuals with probabilities proportional to $\Pr(H_i = j)$.

2.2.1. *Model $M_{t,b,h,\alpha}$.* For illustration, we now focus our efforts on extending the closed population capture–recapture model $M_{t,b,h}$ [King and Brooks (2008), Otis et al. (1978)], which estimates abundance (N) assuming temporal variation, behavioral effects and individual heterogeneity in detection probabilities. Our extension includes all of these effects while accounting for misidentification; we denote this model as $M_{t,b,h,\alpha}$. Before proceeding, we again note that our proposed approach may be used for other capture–recapture models [e.g., Williams, Nichols and Conroy (2002)] by modifying them accordingly for misidentification; the mathematical form for the $M_{t,b,h,\alpha}$ likelihood is simply substituted directly for $[\mathbf{h}|\boldsymbol{\theta}, \boldsymbol{\rho}]$ in equation (3).

We adopt a Bayesian perspective and utilize data augmentation both to account for individuals that were never detected [e.g., Royle, Dorazio and Link (2007)] and to formulate a probit model for detection probability [e.g., Albert and Chib (1993)]. The data augmentation framework is useful because of computational efficiencies it produces, and our procedure treats N as a binomial random variable with known index M (typically $M \gg N$) and parameter ψ. In this context, M is often described as a "superpopulation" size of indicators $q_i \sim \mathrm{Bernoulli}(\psi)$, where individuals with $q_i = 1$ are considered "real individuals" or "individuals available for capture," and $N = \sum_{i=1}^{M} q_i$. For the $\sum_{j=2}^{3^T} x_j$ individuals with $H_i > 1$, we know $q_i = 1$. For the remaining $M - \sum_{j=2}^{3^T} x_j$ individuals that were never detected, $H_i = 1$ and q_i is unknown. A closed population misidentification model allowing temporal and

individual variation in detection probability may then be represented as

$$q_i|\psi \sim \text{Bernoulli}(\psi),$$
$$h_{it}|q_i, p_{it} \sim \text{Categorical}(1 - q_i p_{it}, \alpha q_i p_{it}, (1-\alpha) q_i p_{it})$$

for $h_{it} \in \{0, 1, 2\}$, where p_{it} is the probability of detection for individual i at time t, and α is the probability that an individual is correctly identified, given detection. Because we assume $N|\psi \sim \text{Binomial}(M, \psi)$, a judicious choice of prior can yield the desired prior for N when marginalized over ψ. For example, $\psi \sim \text{Beta}(1, 1)$ produces a discrete uniform prior on N.

As a more computationally efficient alternative to the ubiquitous logit link function for heterogeneous detection probabilities in Bayesian capture–recapture analyses [e.g., Castledine (1981), Fienberg, Johnson and Junker (1999), George and Robert (1992), King and Brooks (2008), Link (2013), Royle, Dorazio and Link (2007)], we use data augmentation to formulate a probit model, $p_{it} = \Phi(\mathbf{w}_{it}'\boldsymbol{\beta} + \gamma_i)$, where Φ is the standard normal cumulative distribution function, $\mathbf{w}_{it}$ is a vector of covariates for individual i at time t, $\boldsymbol{\beta}$ is a vector of regression coefficients, and γ_i is an individual-level effect. Let $y_{it} = \mathrm{I}(h_{it} > 0)$ be an indicator for the binary detection process, and let $\tilde{y}_{it}$ be a continuous latent version of this process, where $\tilde{y}_{it}|\boldsymbol{\beta}, \gamma_i \sim \mathcal{N}(\mathbf{w}_{it}'\boldsymbol{\beta} + \gamma_i, 1)$. Assuming $y_{it} = 1$ if $\tilde{y}_{it} > 0$ and $q_i = 1$, and assuming $y_{it} = 0$ if $\tilde{y}_{it} < 0$ and $q_i = 1$ or $q_i = 0$, then it follows that $y_{it}|q_i, \tilde{y}_{it} \sim \text{Bernoulli}(q_i \mathrm{I}(\tilde{y}_{it} > 0))$. This approach shares some similarities with recent extensions of the probit regression model of Albert and Chib (1993) to imperfectly-detected species occurrence data [Dorazio and Rodriguez (2012), Johnson et al. (2013)], but our extension allows for individual-level effects and a response variable of unknown dimension.

For our probit model allowing temporal, behavioral and individual effects in detection probability, we define $\mathbf{w}_{it} = (\mathrm{I}(t=1), \mathrm{I}(t=2), \dots, \mathrm{I}(t=T), \mathrm{I}(t > C_i))$ and $\boldsymbol{\beta} = (\beta_1, \beta_2, \dots, \beta_{T+1})$, where C_i denotes the first capture occasion for individual i (with $C_i = \infty$ for individuals with $H_i = 1$). Given the recorded history frequencies $\mathbf{f} = (f_1, f_2, \dots, f_{2^T-1})$, the joint posterior distribution for model $M_{t,b,h,\alpha}$ is then

$$\begin{aligned}[\mathbf{h}, \mathbf{q}, \tilde{\mathbf{y}}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \psi, \alpha, \sigma_\gamma^2|\mathbf{f}] \propto{}& [\mathbf{h}|\mathbf{q}, \tilde{\mathbf{y}}, \alpha]\mathrm{I}(\mathbf{A}'\mathbf{x} = \mathbf{f}) \\ &\times [\mathbf{q}|\psi][\tilde{\mathbf{y}}|\boldsymbol{\beta}, \boldsymbol{\gamma}][\boldsymbol{\beta}|\boldsymbol{\mu}_\beta, \boldsymbol{\Sigma}_\beta][\boldsymbol{\gamma}|\sigma_\gamma^2] \\ &\times [\psi][\alpha][\sigma_\gamma^2],\end{aligned} \tag{4}$$

where

$$\begin{aligned}[\mathbf{h}|\mathbf{q}, \tilde{\mathbf{y}}, \alpha] \propto{}& \prod_{i=1}^{M}\prod_{t=1}^{T}\{q_i \mathrm{I}(\tilde{y}_{it} > 0)\}^{\mathrm{I}(h_{it}>0)}\{1 - q_i \mathrm{I}(\tilde{y}_{it} > 0)\}^{\mathrm{I}(h_{it}=0)} \\ &\times \alpha^{\mathrm{I}(h_{it}=1)}(1-\alpha)^{\mathrm{I}(h_{it}=2)}\end{aligned}$$

and

$$\mathbf{x} = \left(\sum_{i=1}^{M} \mathrm{I}(H_i = 1), \sum_{i=1}^{M} \mathrm{I}(H_i = 2), \ldots, \sum_{i=1}^{M} \mathrm{I}(H_i = 3^T)\right) = (x_1, x_2, \ldots, x_{3^T}).$$

We complete our Bayesian formulation by assigning the priors

$$\begin{aligned} \boldsymbol{\beta} | \boldsymbol{\mu}_\beta, \boldsymbol{\Sigma}_\beta &\sim \mathcal{N}(\boldsymbol{\mu}_\beta, \boldsymbol{\Sigma}_\beta), \\ \gamma_i | \sigma_\gamma^2 &\sim \mathcal{N}(0, \sigma_\gamma^2), \\ \alpha &\sim \mathrm{Beta}(a_\alpha, b_\alpha), \\ \psi &\sim \mathrm{Beta}(a_\psi, b_\psi), \end{aligned}$$

and $\sigma_\gamma^2 \sim \Gamma^{-1}(a_{\sigma_\gamma}, b_{\sigma_\alpha})$, where $\boldsymbol{\mu}_\beta$ and $\boldsymbol{\Sigma}_\beta$ are the prior mean and covariance matrix for $\boldsymbol{\beta}$. By choosing $b_\psi = 1$ and a very small positive value for a_ψ, one can approximate the scale prior $[N] \propto 1/N$ [Link (2013)]. We note that simpler closed population abundance models may be specified by modifying model $M_{t,b,h,\alpha}$ accordingly. For example, set $\beta_{T+1} = 0$ to remove behavior effects, set $\beta_1 = \beta_2 = \cdots = \beta_T$ to remove temporal variation, or set $\gamma_i = 0$ for $i = 1, \ldots, M$ to remove individual effects.

Given the $3^T \times (2^T - 1)$ matrix $\mathbf{A}$ for binary recorded histories and a set of basis vectors $\{\mathbf{v}\} = \{\mathbf{v}_1, \mathbf{v}_2, \ldots, \mathbf{v}_{3^T - 2^T + 1}\}$ for the null space of $\mathbf{A}'$ (where $\mathbf{v}_1$ is the basis vector corresponding to the all-zero latent history frequency), we propose the following MCMC algorithm for sampling from the posterior distribution of model $M_{t,b,h,\alpha}$ [equation (4)]. We utilize Metropolis–Hastings updates for the latent encounter histories, but our judicious choice of priors enables Gibbs updates for $\mathbf{q}$, $\tilde{\mathbf{y}}$ and all parameters:

1. Initialize all parameters and latent variables, including an initial feasible set of M latent individual histories ($\mathbf{h}$) with corresponding frequencies $\mathbf{x}$ satisfying $\mathbf{A}'\mathbf{x} = \mathbf{f}$. One such initial vector $\mathbf{x}$ is readily available by assuming $\alpha = 1$, such that latent frequencies corresponding to histories with 2's are zeros, with a one-to-one matching of the remaining latent frequencies with the recorded history frequencies ($\mathbf{f}$). This creates $\sum_{k=1}^{2^T-1} f_k$ individual histories (with corresponding $H_i > 1$), none of which is the all-zero history. To complete the initialization, assign $x_1 = M - \sum_{k=1}^{2^T-1} f_k$ individuals to the all-zero history (with corresponding $H_i = 1$).
2. Update $\tilde{y}_{it}$ for $i = 1, \ldots, M$ and $t = 1, \ldots, T$ from the full conditional distribution:

$$\tilde{y}_{it} | \cdot \sim \begin{cases} \mathcal{TN}_{(0,\infty)}(\mathbf{w}_{it}'\boldsymbol{\beta} + \gamma_i, 1), & \text{if } h_{it} > 0 \text{ and } q_i = 1, \\ \mathcal{TN}_{(-\infty,0)}(\mathbf{w}_{it}'\boldsymbol{\beta} + \gamma_i, 1), & \text{if } h_{it} = 0 \text{ and } q_i = 1, \\ \mathcal{N}(\mathbf{w}_{it}'\boldsymbol{\beta} + \gamma_i, 1), & \text{otherwise,} \end{cases}$$

where $\mathcal{TN}_{(L,U)}$ is a normal distribution truncated at L and U.

3. Update $\boldsymbol{\beta}$ from the full conditional distribution:

$$\boldsymbol{\beta}|\cdot \sim \mathcal{N}((\boldsymbol{\Sigma}_\beta^{-1} + \mathbf{W}'\mathbf{W})^{-1}(\boldsymbol{\Sigma}_\beta^{-1}\boldsymbol{\mu}_\beta + \mathbf{W}'(\tilde{\mathbf{y}} - \boldsymbol{\gamma} \otimes \mathbf{1}_T)), (\boldsymbol{\Sigma}_\beta^{-1} + \mathbf{W}'\mathbf{W})^{-1}),$$

where $\mathbf{W}$ is the $MT \times (T+1)$ design matrix with rows $\mathbf{w}_{it}'$ and $\mathbf{1}_T$ is the all-ones vector of length T.

4. Update γ_i for $i = 1, \ldots, M$ from the full conditional distribution:

$$\gamma_i|\cdot \sim \mathcal{N}\left(\frac{\sigma_\gamma^2 \sum_{t=1}^T (\tilde{y}_{it} - \mathbf{w}_{it}'\boldsymbol{\beta})}{1 + T\sigma_\gamma^2}, \frac{\sigma_\gamma^2}{1 + T\sigma_\gamma^2}\right).$$

5. Update σ_γ^2 from the full conditional distribution:

$$\sigma_\gamma^2|\cdot \sim \Gamma^{-1}\left(a_{\sigma_\gamma} + \frac{M}{2}, b_{\sigma_\gamma} + \frac{\boldsymbol{\gamma}'\boldsymbol{\gamma}}{2}\right).$$

6. Update α from the full conditional distribution:

$$\alpha|\cdot \sim \text{Beta}\left(a_\alpha + \sum_{i=1}^M \sum_{t=1}^T \mathrm{I}(h_{it} = 1), b_\alpha + \sum_{i=1}^M \sum_{t=1}^T \mathrm{I}(h_{it} = 2)\right).$$

7. Update q_i for the x_1 individuals with $H_i = 1$ from the full conditional distribution by drawing from a Bernoulli distribution with probability

$$\Pr(q_i = 1 | H_i = 1) = \frac{\psi \prod_{t=1}^T \{1 - \Phi(\mathbf{w}_{it}'\boldsymbol{\beta} + \gamma_i)\}}{\psi \prod_{t=1}^T \{1 - \Phi(\mathbf{w}_{it}'\boldsymbol{\beta} + \gamma_i)\} + (1 - \psi)}.$$

8. Update ψ from the full conditional distribution:

$$\psi|\cdot \sim \text{Beta}\left(a_\psi + \sum_{i=1}^M q_i, b_\psi + M - \sum_{i=1}^M q_i\right).$$

9. Update the set of M latent encounter histories ($\mathbf{h}$) using a Metropolis–Hastings step.

(a) Set $H_i^* = H_i$ for $i = 1, \ldots, M$ and $x_1^* = x_1$. Randomly draw r from the integer set $\{2, \ldots, 3^T - 2^T + 1\}$ corresponding to basis vectors $\{\mathbf{v}_2, \ldots, \mathbf{v}_{3^T - 2^T + 1}\}$. Next draw k_r from a discrete uniform distribution over the integers $\{-D_r, \ldots, -1, 1, \ldots, D_r\}$, where D_r is a tuning parameter. Propose a latent history frequency vector

$$\mathbf{x}^* = \mathbf{x} + k_r \mathbf{v}_r.$$

If any $x_j^* < 0$ for $j = 2, \ldots, 3^T$ or $M - \sum_{j=2}^{3^T} x_j^* < 0$, go to step 10.

(b) Apportion $\mathbf{x}^*$ to individuals with probabilities proportional to $\Pr(H_i = j)$. With probability 0.5, continue to step 9(b)(i) followed by step 9(b)(ii); otherwise proceed with step 9(b)(ii) followed by step 9(b)(i).

(i) For each $x_j^* < x_j$ $(j = 2, \ldots, 3^T)$, draw a set $\{O_r^{j-}\} = \{o_1^{j-}, o_2^{j-}, \ldots, o_{k_r}^{j-}\}$ of individuals (of size k_r) without replacement from the x_j individuals with capture history j (i.e., $H_i^* = j$) with respective probabilities

$$\Pr(H_i^{**} = 1) = \prod_{t=1}^{T} \{1 - \Phi(\mathbf{w}_{it}'\boldsymbol{\beta} + \gamma_i)\},$$

and set $H_i^* = 1$ for individuals $i \in \{O_r^{j-}\}$. After cycling through each j for which $x_j^* < x_j$ $(j = 2, \ldots, 3^T)$, set $x_1^* = \sum_{i=1}^{M} \mathrm{I}(H_i^* = 1)$.

(ii) For each $x_j^* > x_j$ $(j = 2, \ldots, 3^T)$, draw a set $\{O_r^{j+}\} = \{o_1^{j+}, o_2^{j+}, \ldots, o_{k_r}^{j+}\}$ of individuals (of size k_r) without replacement from the x_1^* individuals that were never captured with respective probabilities

$$\Pr(H_i^{**} = j) = \prod_{t=1}^{T} \Phi(\mathbf{w}_{it}'\boldsymbol{\beta} + \gamma_i)^{\mathrm{I}(h_{it}^{**} > 0)} \{1 - \Phi(\mathbf{w}_{it}'\boldsymbol{\beta} + \gamma_i)\}^{\mathrm{I}(h_{it}^{**} = 0)} \times \alpha^{\mathrm{I}(h_{it}^{**} = 1)} (1 - \alpha)^{\mathrm{I}(h_{it}^{**} = 2)}.$$

Set $H_i^* = j$ and $q_i^* = 1$ for individuals $i \in \{O_r^{j+}\}$, and set $x_1^* = \sum_{i=1}^{M} \mathrm{I}(H_i^* = 1)$. Cycle through each j for which $x_j^* > x_j$ $(j = 2, \ldots, 3^T)$.

(c) Propose q_i^* for the x_1^* individuals with $H_i^* = 1$ as in step 7. Accept the proposed latent histories (i.e., set $\mathbf{x} = \mathbf{x}^*$, $H_i = H_i^*$ and $q_i = q_i^*$) with probability $\min(1, R_r)$, where

$$R_r = \left(\left[\prod_{i : q_i^* = 1} \left\{ \sum_{j=1}^{3^T} \Pr(H_i^{**} = j) \mathrm{I}(H_i^* = j) \right\} \right] \times [\mathbf{q}^* | \psi][\mathbf{h} | \mathbf{h}^*, \boldsymbol{\beta}, \boldsymbol{\gamma}, \alpha][\mathbf{q} | \psi, \boldsymbol{\beta}, \boldsymbol{\gamma}] \right) \Big/ \left(\left[\prod_{i : q_i = 1} \left\{ \sum_{j=1}^{3^T} \Pr(H_i^{**} = j) \mathrm{I}(H_i = j) \right\} \right] \times [\mathbf{q} | \psi][\mathbf{h}^* | \mathbf{h}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \alpha][\mathbf{q}^* | \psi, \boldsymbol{\beta}, \boldsymbol{\gamma}] \right),$$

$[\mathbf{h}^* | \mathbf{h}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \alpha]$ is the proposal density for $\mathbf{h}^*$, and $[\mathbf{q}^* | \psi, \boldsymbol{\beta}, \boldsymbol{\gamma}]$ is the proposal density for $\mathbf{q}^*$. Here, $[\mathbf{h}^* | \mathbf{h}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \alpha]$ is the product of the (ordered) conditional inclusion probabilities, $\Pr(H_i^{**} = 1)$ for $i \in \{O_r^{j-}\}$ and $\Pr(H_i^{**} = j)$ for $i \in \{O_r^{j+}\}$, that were, respectively, selected in steps 9(b)(i) and 9(b)(ii) under

unequal probability sampling without replacement [e.g., Thompson (1992), page 53]:

$$[\mathbf{h}^*|\mathbf{h}, \boldsymbol{\beta}, \boldsymbol{\gamma}, \alpha]$$
$$= \left[\prod_{\substack{j: x_j^* < x_j;\\ j>1}} \prod_{s=1}^{k_r} \frac{\Pr(H^{**}_{o_s^{j-}} = 1)}{\sum_{i: H_i = j} \Pr(H_i^{**} = 1) - \sum_{m=1}^{s-1} \Pr(H^{**}_{o_m^{j-}} = 1)}\right]$$
$$\times \left[\prod_{\substack{j: x_j^* > x_j;\\ j>1}} \prod_{s=1}^{k_r} \frac{\Pr(H^{**}_{o_s^{j+}} = j)}{\sum_{i: H_i^* = 1} \Pr(H_i^{**} = j) - \sum_{m=1}^{s-1} \Pr(H^{**}_{o_m^{j+}} = j)}\right]$$

and

$$[\mathbf{h}|\mathbf{h}^*, \boldsymbol{\beta}, \boldsymbol{\gamma}, \alpha]$$
$$= \left[\prod_{\substack{j: x_j^* < x_j;\\ j>1}} \prod_{s=1}^{k_r} \frac{\Pr(H^{**}_{o_s^{j-}} = j)}{\sum_{i: H_i^* = 1} \Pr(H_i^{**} = j) - \sum_{m=1}^{s-1} \Pr(H^{**}_{o_m^{j-}} = j)}\right]$$
$$\times \left[\prod_{\substack{j: x_j^* > x_j;\\ j>1}} \prod_{s=1}^{k_r} \frac{\Pr(H^{**}_{o_s^{j+}} = 1)}{\sum_{i: H_i^* = j} \Pr(H_i^{**} = 1) - \sum_{m=1}^{s-1} \Pr(H^{**}_{o_m^{j+}} = 1)}\right].$$

10. Return to step 2 and repeat as needed.

Note that N is obtained by calculating $N = \sum_{i=1}^{M} q_i$ at each iteration of the algorithm.

2.2.2. *Model M_{t,b,α_h}.* In some applications, one may be more concerned about individual heterogeneity in misidentification than detection probability. For example, the quality of visual identifiers (e.g., artificial marks, naturally occurring pelt or scar patterns) or genetic material (e.g., hair or fecal samples) may vary by individual [Lukacs and Burnham (2005)], and some individuals may therefore be more or less likely to spawn ghost histories (see *Blue Ridge two-lined salamander*). We can modify model $M_{t,b}$ [King and Brooks (2008), Otis et al. (1978)] to accommodate temporal variation, behavioral effects and individual heterogeneity in correct identification probability, obtaining a model we call M_{t,b,α_h}.

Similar to Section 2.2.1, we specify a probit model for the probability of correctly identifying an individual, given detection, $\alpha_i = \Phi(\mu_\alpha + \varepsilon_i)$, where μ_α is an intercept term and ε_i is an individual-level effect. Let u_{it} be an indicator for the binary correct identification process, and let $\tilde{u}_{it}$ be a continuous latent version of this process, where $\tilde{u}_{it}|\mu_\alpha, \varepsilon_i \sim \mathcal{N}(\mu_\alpha + \varepsilon_i, 1)$. Assuming $u_{it} = 1$ if $\tilde{u}_{it} > 0$ and $h_{it} = 1$, and assuming $u_{it} = 0$ if $\tilde{u}_{it} < 0$ and $h_{it} = 2$ or $h_{it} = 0$, then it follows that

$u_{it}|h_{it}, \tilde{u}_{it} \sim \text{Bernoulli}(\mathrm{I}(h_{it} > 0)\mathrm{I}(\tilde{u}_{it} > 0))$. The joint posterior distribution for model M_{t,b,α_h} is then

$$
\begin{aligned}
[\mathbf{h}, \mathbf{q}, \tilde{\mathbf{y}}, \tilde{\mathbf{u}}, \boldsymbol{\beta}, \mu_\alpha, \boldsymbol{\varepsilon}, \psi, \sigma_\varepsilon^2|\mathbf{f}] \propto{}& [\mathbf{h}|\mathbf{q}, \tilde{\mathbf{y}}, \tilde{\mathbf{u}}]\mathrm{I}(\mathbf{A}'\mathbf{x} = \mathbf{f}) \\
&\times [\mathbf{q}|\psi][\tilde{\mathbf{y}}|\boldsymbol{\beta}][\tilde{\mathbf{u}}|\mu_\alpha, \boldsymbol{\varepsilon}][\boldsymbol{\beta}|\boldsymbol{\mu}_\beta, \boldsymbol{\Sigma}_\beta][\boldsymbol{\varepsilon}|\sigma_\varepsilon^2] \qquad (5)\\
&\times [\psi][\mu_\alpha][\sigma_\varepsilon^2],
\end{aligned}
$$

where

$$
\begin{aligned}
[\mathbf{h}|\mathbf{q}, \tilde{\mathbf{y}}, \tilde{\mathbf{u}}] \propto \prod_{i=1}^{M}\prod_{t=1}^{T}&\{q_i\mathrm{I}(\tilde{y}_{it} > 0)\}^{\mathrm{I}(h_{it}>0)}\{1 - q_i\mathrm{I}(\tilde{y}_{it} > 0)\}^{\mathrm{I}(h_{it}=0)} \\
&\times \{\mathrm{I}(h_{it} > 0)\mathrm{I}(\tilde{u}_{it} > 0)\}^{\mathrm{I}(h_{it}=1)} \\
&\times \{1 - \mathrm{I}(h_{it} > 0)\mathrm{I}(\tilde{u}_{it} > 0)\}^{\mathrm{I}(h_{it}=2)},
\end{aligned}
$$

$\tilde{y}_{it}|\boldsymbol{\beta} \sim \mathcal{N}(\mathbf{w}_{it}'\boldsymbol{\beta}, 1)$, and all other components of the model are specified as in Section 2.2.1 for model $M_{t,b,h,\alpha}$. We assign the additional priors $\mu_\alpha \sim \mathcal{N}(\mu_{\mu_\alpha}, \sigma_{\mu_\alpha}^2)$, $\varepsilon_i|\sigma_\varepsilon^2 \sim \mathcal{N}(0, \sigma_\varepsilon^2)$, and $\sigma_\varepsilon^2 \sim \Gamma^{-1}(a_{\sigma_\varepsilon}, b_{\sigma_\varepsilon})$.

It is straightforward to modify the MCMC algorithm described in Section 2.2.1 for sampling from the posterior distribution of model M_{t,b,α_h} [equation (5)]. The additional parameters and $\tilde{\mathbf{u}}$ are simply updated from their full conditional distributions:

$$
\tilde{u}_{it}|\cdot \sim \begin{cases} \mathcal{T}N_{(0,\infty)}(\mu_\alpha + \varepsilon_i, 1), & \text{if } h_{it} = 1, \\ \mathcal{T}N_{(-\infty,0)}(\mu_\alpha + \varepsilon_i, 1), & \text{if } h_{it} = 2, \\ \mathcal{N}(\mu_\alpha + \varepsilon_i, 1), & \text{otherwise,} \end{cases}
$$

$$
\mu_\alpha|\cdot \sim \mathcal{N}\left(\left(\frac{1}{\sigma_{\mu_\alpha}^2} + MT\right)^{-1}\left(\frac{\mu_{\mu_\alpha}}{\sigma_{\mu_\alpha}^2} + \sum_{i=1}^{M}\sum_{t=1}^{T}\{\tilde{u}_{it} - \varepsilon_i\}\right), \left(\frac{1}{\sigma_{\mu_\alpha}^2} + MT\right)^{-1}\right),
$$

$$
\varepsilon_i|\cdot \sim \mathcal{N}\left(\frac{\sigma_\varepsilon^2 \sum_{t=1}^{T}\{\tilde{u}_{it} - \mu_\alpha\}}{1 + T\sigma_\varepsilon^2}, \frac{\sigma_\varepsilon^2}{1 + T\sigma_\varepsilon^2}\right)
$$

and

$$
\sigma_\varepsilon^2|\cdot \sim \Gamma^{-1}\left(a_{\sigma_\varepsilon} + \frac{M}{2}, b_{\sigma_\varepsilon} + \frac{\boldsymbol{\varepsilon}'\boldsymbol{\varepsilon}}{2}\right).
$$

The only other notable difference from our algorithm for model $M_{t,b,h,\alpha}$ is that we instead use $\Pr(H_i^{**} = 1) = 1 - \Pr(H_i^{**} = j)$ to propose individuals that were never detected in the step corresponding to 9(b)(i) above. This is because under model M_{t,b,α_h}, all individuals have the same probability of never being detected.

3. Example applications.

3.1. *Black bears of the Northern Lower Peninsula, Michigan, USA.* In an impressive field and analytical effort, Dreher et al. (2007) applied a closed popula-

tion model that incorporates individual misidentification due to genotyping error [Lukacs and Burnham (2005)] to estimate black bear (*Ursus americanus*) abundance in the Northern Lower Peninsula of Michigan, USA. DNA samples were collected from baited barbed wire hair snares on five occasions from 22 June–26 July 2003. A sixth DNA sampling occasion occurred through the extraction of teeth and muscle tissue from bears registered during the recreational harvest in the autumn (hence, $T = 6$). In addition, a random sample of hand-pulled hair samples collected from harvested bears provided auxiliary information about the probability of a genotyping error using hair-snare samples. Complete details of the data collection, genetic analysis and statistical analysis can be found in Dreher et al. (2007).

Here we re-visit the DNA capture–recapture data of Dreher et al. (2007) using our closed population abundance model allowing for temporal variation, behavioral effects, individual heterogeneity and misidentification (Section 2.2.1). Our motivation is twofold: (1) individual heterogeneity in detection from hair-snare samples was suspected by Dreher et al. (2007), but not incorporated into their misidentification model; and (2) the misidentification model proposed by Lukacs and Burnham (2005) relies on several assumptions that are unlikely to be met in practice and does not properly account for ghost capture histories that result from misidentification [Link et al. (2010), Yoshizaki (2007), Yoshizaki et al. (2011)].

Based on the best-supported model from Dreher et al. (2007), we fit model $M_{\text{hunt},b,h,\alpha}$, which allows for different detection probabilities for the two methods of capture (i.e., hair snare or harvest; indicated by "hunt"), a behavioral response to the baited hair snares, individual heterogeneity in detection probability from hair-snare sampling, and misidentification of hair snare samples due to genotyping error. Allowing misidentification to occur only for the hair-snare sampling occasions ($t = 1, \dots, 5$), we have

$$[\mathbf{h}|\mathbf{q}, \tilde{\mathbf{y}}, \alpha] \propto \prod_{i=1}^{M} \left[\prod_{t=1}^{5} \{q_i \mathrm{I}(\tilde{y}_{it} > 0)\}^{\mathrm{I}(h_{it}>0)} \{1 - q_i \mathrm{I}(\tilde{y}_{it} > 0)\}^{\mathrm{I}(h_{it}=0)} \times \alpha^{\mathrm{I}(h_{it}=1)} (1-\alpha)^{\mathrm{I}(h_{it}=2)} \right] \times (q_i p_{\text{hunt}})^{h_{i6}} (1 - q_i p_{\text{hunt}})^{1-h_{i6}},$$

$p_{it} = \Phi(\mathbf{w}_{it}'\boldsymbol{\beta} + \gamma_i)$ for $t = 1, \dots, 5$, $\mathbf{w}_{it} = (1, \mathrm{I}(t > C_i))$, $\boldsymbol{\beta} = (\beta_1, \beta_2)$, and $p_{\text{hunt}}|a_p, b_p \sim \text{Beta}(a_p, b_p)$.

The $\mathbf{A}$ matrix, posterior and MCMC algorithm described in Section 2.2.1 are modified accordingly, where the reduced $[2(3^{T-1})] \times (2^T - 1)$ $\mathbf{A}$ matrix does not include misidentification for the harvest sampling occasion ($t = 6$), and p_{hunt} is updated from the full conditional distribution: $p_{\text{hunt}}|\cdot \sim \text{Beta}(a_p + \sum_{i=1}^{M} q_i h_{i6}, b_p + \sum_{i=1}^{M} q_i(1 - h_{i6}))$. We used weakly informative priors by setting

$\boldsymbol{\mu}_\beta = \mathbf{0}$ and $\boldsymbol{\Sigma}_\beta = \text{diag}(10, 10)$, $a_\psi = 10^{-6}$ and $a_{\sigma_\gamma} = b_{\sigma_\gamma} = a_p = b_p = b_\psi = 1$. Based on the auxiliary data about genotyping error of hair samples collected from harvested bears (where 91 out of 95 samples were correctly assigned to genotype), we used an informative prior with $a_\alpha = 91$ and $b_\alpha = 4$. To investigate prior sensitivity, we conducted an additional analysis using an uninformative prior on α by specifying $a_\alpha = b_\alpha = 1$.

Our MCMC algorithm was written in the C programming language [Kernighan and Ritchie (1988)] with data pre- and post-processing performed in R via the .C interface [R Core Team (2012)]. Starting with $D_r = 1$ and rounding to the nearest integer, we tuned the MH sampler every 5000 iterations by multiplying or dividing D_r by 0.95 if the acceptance rate for basis vector r was ≤ 0.44 or > 0.44, respectively, where acceptance rates were calculated as the number of accepted moves divided by the number of attempted moves. After pilot tuning and burn-in of 500,000 iterations from overdispersed starting values, we obtained three chains of 10 million iterations for both analyses. With $M = 5000$, our analyses required about 48 hrs on a computer running 64-bit Windows 7 (3.4 GHz Intel Core i7 processor, 16 Gb RAM). Slow mixing necessitated long runs, likely due to correlated parameters and low movement rates for the MH sampler. Similar to Link et al. (2010), low movement rates for the MH sampler resulted from many of the 486 possible latent histories having very low probability. Chain convergence was assessed by visual inspection and the Gelman–Rubin–Brooks (GRB) diagnostic in the R package CODA [Plummer et al. (2006)]. For both analyses, all univariate GRB diagnostics were < 1.002 and the multivariate GRB diagnostic was 1.001 for monitored parameters $(N, \beta_1, \beta_2, \sigma_\gamma^2, p_{\text{hunt}}, \alpha)$. Based on sample autocorrelations, mixing was somewhat slower using the uninformative prior on α, but effective sample sizes exceeded 5000 for all parameters.

Using the informative prior on α, we estimated posterior median $N = 1945$ with a 95% credible interval (CI) of 1470–2681 (Figure 1). Similar to Dreher et al. (2007), we found evidence of a trap "happy" behavioral response to the baited hair snares, with posterior mean $\beta_2 = 0.50$ (95% CI = 0.08–0.89). Estimates for p_{hunt} suggest about 21% of this population was harvested and reported to officials (Table 3). As suspected by Dreher et al. (2007), we found evidence of individual heterogeneity in detection from hair snares, with posterior median $\sigma_\gamma = 0.63$ (95% CI = 0.42–0.92). Because unmodeled individual heterogeneity tends to cause underestimation of abundance, this likely explains our posterior distribution for N having support at higher values than the original estimates using models that did not account for individual heterogeneity. For example, Dreher et al. (2007) estimated $N = 1882$ with a 95% confidence interval of 1389–2551 for model $M_{\text{hunt},b}$ using the misidentification model proposed by Lukacs and Burnham (2005).

We estimated posterior mean $\alpha = 0.95$ (95% CI = 0.90–0.99), with slight evidence of higher misidentification probabilities from the hair-snare samples than

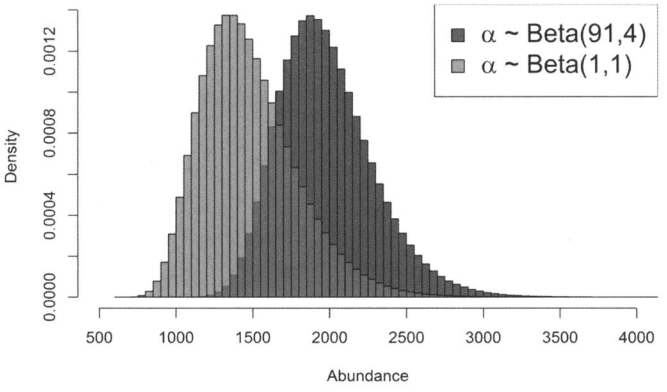

FIG. 1. *Posterior distributions for abundance of black bears in the Northern Lower Peninsula of Michigan, USA, from DNA capture–recapture surveys conducted in summer and autumn* 2003. *Results are for analyses using an uninformative prior (light) and an informative prior (dark) on the probability of correctly identifying an individual, given detection* (α).

from the auxiliary hair samples collected from harvested bears (Figure 2). The auxiliary genotyping error data proved quite informative; the analogous analysis using an uninformative Beta(1, 1) prior on α yielded posterior median $N = 1436$ (95% CI = 988–2203; Figure 1) and posterior mean $\alpha = 0.77$ (95% CI = 0.61–0.95; Figure 2). In the absence of prior information, the recorded histories may only provide minimal information about misidentification, such as the range of α for which there is very little support. Put another way, the frequencies of potential ghost histories alone suggest $\alpha > 0.55$. If α were in fact $\ll 0.55$, we would expect many more ghost histories to have been observed relative to the observed nonghost

TABLE 3

Posterior summaries and effective sample sizes (ESS) for model $M_{\text{hunt},b,h,\alpha}$ *using black bear DNA capture–recapture data collected in the Northern Lower Peninsula of Michigan, USA in* 2003. *Mean capture and recapture probabilities were derived as* $\bar{p} = \int_{-\infty}^{\infty} \Phi(\beta_1 + \gamma)[\gamma|\sigma_\gamma^2]\,d\gamma$ *and* $\bar{c} = \int_{-\infty}^{\infty} \Phi(\beta_1 + \beta_2 + \gamma)[\gamma|\sigma_\gamma^2]\,d\gamma$, *respectively*

					95%		
Parm.	**Mean**	**Median**	**Mode**	**SD**	**LCI**	**UCI**	**ESS**
N	1978.9	1945	1875	310.5	1470	2681	69,711
$\bar{p}$	0.02	0.02	0.02	0.00	0.01	0.03	89,171
$\bar{c}$	0.05	0.05	0.04	0.02	0.02	0.10	31,380
p_{hunt}	0.21	0.21	0.21	0.03	0.15	0.28	80,241
α	0.95	0.96	0.96	0.02	0.90	0.99	11,107
β_1	−2.48	−2.46	−2.44	0.16	−2.84	−2.21	16,568
β_2	0.50	0.51	0.53	0.21	0.08	0.89	55,550
σ_γ	0.64	0.63	0.61	0.13	0.42	0.92	16,371

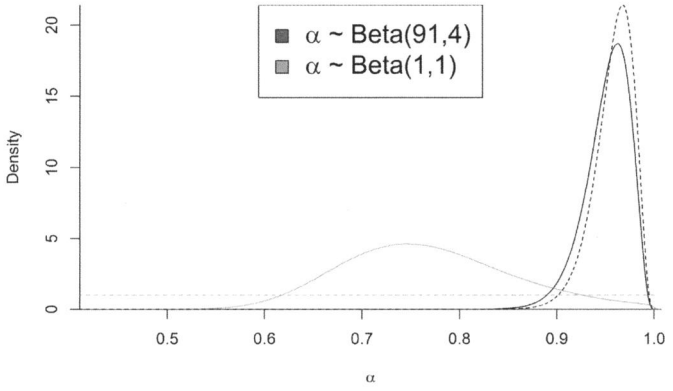

FIG. 2. *Posterior (solid lines) and prior (dashed lines) densities for the probability of correctly genotyping DNA hair-snare samples (α) collected from black bears in the Northern Lower Peninsula of Michigan, USA. Results are for analyses using an uninformative prior (light) and an informative prior (dark) on α.*

histories [see Section 4.1 in Link et al. (2010) for further discussion]. Nevertheless, when using the informative prior for α, we found relatively little contrary information about misidentification from the recorded histories. Given this prior sensitivity, care should be taken in specifying informative priors for α. For example, there could be reason to suspect that hair samples collected from harvested bears are of higher quality than hair-snare samples (e.g., due to degradation by environmental factors), in which case hair-snare misidentification could potentially be underestimated (and abundance overestimated) from this prior.

3.2. *Blue Ridge two-lined salamanders.* Bailey (2004) conducted a laboratory experiment evaluating the ability of observers to individually identify Blue Ridge two-lined salamanders (*Eurycea bislineata wilderae*) marked with a subcutaneous injection of elastomer (a silicone-based material manufactured by Northwest Marine Technology, Inc., Shaw Island, Washington, USA). Out of a pool of 20 marked salamanders, each of 14 observers viewed 10 randomly chosen individuals. Two different lights were used for viewing the marks: a dive light with blue filter lens (hereafter blue light) and a deep blue 7-LED flashlight (hereafter black light). Observers first viewed each salamander with one light (randomly assigned), and then the 10 individuals were re-randomized and presented to the observers for identification using the other light. Bailey (2004) found no difference in observer ability to correctly identify individuals based on the light used, but found mark quality strongly influenced observers' ability to correctly identify individuals. For example, one individual salamander accounted for 10 of 15 misidentifications resulting from missed marks (because one of its marks was quite small).

This laboratory experiment affords an opportunity to apply model M_{t,α_h} on a population of known size ($N = 20$) with no individual variation in detection

and suspected individual variation in misidentification probability. Although some ghosts were identified by multiple observers, here we analyze a subset of $T = 8$ observers for which all ghost encounter histories contain a single detection and no identification errors matched a legitimate individual (thus satisfying these assumptions of the model). Because the true encounter history for each marked individual was known, recorded history data were simply generated from the true histories. For example, suppose an individual was presented to 4 of the 8 observers and, using the blue light, they recorded the true encounter history .12...21 (where a dot indicates this individual was not presented to the corresponding observer), then the blue light recorded histories spawned from this true encounter history would be 01000001, 00100000 and 00000010. We performed separate analyses for the blue and black light recorded histories to examine potential differences in misidentification probabilities, as well as our model's ability to accurately estimate the number of salamanders used in the experiment.

Allowing for temporal variation in detection and individual variation in misidentification probability, we modify the posterior and MCMC algorithm described in Section 2.2.2 accordingly. Setting $p_{it} = p_t$ and assuming $p_t|a_p, b_p \sim \text{Beta}(a_p, b_p)$, p_t can be updated from the full conditional distribution: $p_t|\cdot \sim \text{Beta}(a_p + \sum_{i=1}^{M} q_i \text{I}(h_{it} > 0), b_p + \sum_{i=1}^{M} q_i \text{I}(h_{it} = 0))$. We used weakly informative priors by setting $\mu_{\mu_\alpha} = 0$, $\sigma^2_{\mu_\alpha} = 10$, $a_\psi = 10^{-6}$, and $a_{\sigma_\varepsilon} = b_{\sigma_\varepsilon} = a_p = b_p = b_\psi = 1$.

For both the blue and black light analyses, we obtained three chains of 10 million iterations after initial pilot tuning and a burn-in of 500,000 iterations from overdispersed starting values. With $M = 200$, our analyses required about 2 hrs to complete. As in the black bear example, relatively slow mixing necessitated long runs, likely due to correlated parameters and low movement rates for the MH sampler. For both analyses, all univariate GRB diagnostics were <1.05 and the multivariate GRB diagnostic was <1.008 for monitored parameters $(N, p_t, \mu_\alpha, \sigma^2_\varepsilon, \alpha)$. Based on sample autocorrelations, mixing was somewhat slower for several parameters in the blue light analyses, but all effective sample sizes exceeded 8000 for both analyses.

For the blue light analysis, we found posterior median $N = 20$ with a 95% credible interval of 18–23. For the black light analysis, we found posterior median $N = 21$ (95% CI = 19–25) (Table 4). Hence, our model was able to reliably estimate N using either light source. As in Bailey (2004), we found misidentification probabilities were similar for the blue and black lights, with posterior mean $\bar{\alpha} = \int_{-\infty}^{\infty} \Phi(\mu_\alpha + \varepsilon)[\varepsilon|\sigma^2_\varepsilon]\, d\varepsilon = 0.88$ (95% CI = 0.74–0.97) and 0.88 (95% CI = 0.76–0.97), respectively. We found some evidence of individual heterogeneity in misidentification probabilities attributable to variable mark quality, with posterior median $\sigma_\varepsilon = 1.34$ (95% CI = 0.54–4.51) and $\sigma_\varepsilon = 1.07$ (95% CI = 0.50–3.45) for the blue and black lights, respectively.

TABLE 4

Posterior summaries and effective sample sizes (ESS) for model M_{t,α_h} using salamander data generated from laboratory experiments evaluating the effectiveness of a subcutaneously injected marking material with two light sources (blue and black). The population was of known size ($N = 20$) with 8 total misidentifications (0–3 per individual) using the blue light and 9 total misidentifications (0–2 per individual) using the black light

	Blue light							Black light						
				95%							95%			
Parm.	**Mean**	**Median**	**Mode**	**SD**	**LCI**	**UCI**	**ESS**	**Mean**	**Median**	**Mode**	**SD**	**LCI**	**UCI**	**ESS**
N	19.8	20	20	1.3	18	23	11,935	21.3	21	21	1.7	19	25	34,562
p_1	0.51	0.51	0.51	0.11	0.30	0.72	1,378,542	0.47	0.47	0.46	0.11	0.27	0.68	924,064
p_2	0.51	0.51	0.50	0.11	0.30	0.72	1,383,346	0.47	0.47	0.47	0.11	0.27	0.68	925,735
p_3	0.51	0.51	0.51	0.11	0.30	0.72	1,379,963	0.47	0.47	0.47	0.11	0.27	0.68	922,474
p_4	0.51	0.51	0.51	0.11	0.30	0.72	1,381,844	0.43	0.43	0.42	0.10	0.24	0.64	1,175,373
p_5	0.51	0.51	0.51	0.11	0.30	0.72	1,389,232	0.47	0.47	0.47	0.11	0.27	0.68	924,540
p_6	0.51	0.51	0.50	0.11	0.30	0.72	1,379,287	0.47	0.47	0.47	0.11	0.27	0.68	924,655
p_7	0.51	0.51	0.51	0.11	0.30	0.72	1,384,098	0.47	0.47	0.47	0.11	0.27	0.68	920,696
p_8	0.51	0.51	0.50	0.11	0.30	0.72	1,381,843	0.47	0.47	0.47	0.11	0.27	0.68	927,481
μ_α	2.37	2.06	1.62	1.11	1.05	5.37	8667	2.06	1.83	1.59	0.88	1.02	4.44	11,134
σ_ε	1.65	1.34	0.88	1.06	0.54	4.51	8264	1.30	1.07	0.79	0.80	0.50	3.45	10,593
α	0.88	0.89	0.90	0.06	0.74	0.97	58,385	0.88	0.89	0.90	0.05	0.76	0.97	71,694

4. Discussion. We have presented a general model formulation and MCMC model-fitting algorithm for capture–recapture models allowing for misidentification and individual heterogeneity in parameters. Our approach is computationally more demanding than the closed population misidentification model proposed by Lukacs and Burnham (2005), implemented in Program MARK [White and Burnham (1999)], that allows for individual heterogeneity in detection probability using a finite mixture distribution. However, Lukacs and Burnham (2005) do not properly account for misidentification [Link et al. (2010), Yoshizaki (2007), Yoshizaki et al. (2011)], and their approach performs particularly poorly when detection probabilities are too low (<0.1) or too high (>0.3), as well as when $\alpha < 0.95$ [Lukacs and Burnham (2005)]. The computational cost of our approach may therefore be worth the additional effort, but similar to Link et al. (2010), the computational demands of using basis vectors to propose $\mathbf{x}$ (and allocate $\mathbf{h}$ accordingly) can be impractical for large T. These computational demands can be somewhat reduced by eliminating basis vectors that will always produce negative latent history frequencies for a given $\mathbf{f}$, but in the absence of gains in computing power, more efficient methods for evaluating equation (2) will likely be needed for $T > 10$.

Owing to the complexity of the model, we found mixing to be relatively slow and recommend long runs when implementing our proposed MCMC algorithm. Other capture–recapture models of somewhat similar complexity have also exhibited slow mixing that is likely due to correlated parameters [e.g., Bonner and Schofield (2013), Fienberg, Johnson and Junker (1999), Link (2013)] and low movement rates for the MH sampler [e.g., Link et al. (2010)]. Computational efficiency could potentially be improved by accounting for individual heterogeneity using observed or "semi-complete" data likelihoods in place of complete data likelihoods [Bonner and Schofield (2013), Fienberg, Johnson and Junker (1999), R. King, B. T. McClintock, D. Kidney and D. L. Borchers, *unpublished manuscript*]. Capture–recapture data tend to be somewhat sparse, and in application many of the possible latent histories could have very low probability. Instead of drawing basis vectors with equal probability in step (a) of the algorithm described in Section 2.2.1, movement rates of the MH sampler could potentially be improved by drawing r with probabilities proportional to those of the corresponding latent histories proposed by each basis vector.

Alternative models have been proposed for handling matching uncertainty in wildlife populations [e.g., Tancredi et al. (2013), Wright et al. (2009)] or "record linkage" in human populations [e.g., Tancredi and Liseo (2011)]. These approaches rely on auxiliary information, such as genotype or family name, to match recorded histories, but they do not account for individual heterogeneity in parameters. By integrating a simpler form of record linkage and individual heterogeneity into a unified missing data framework, our work constitutes a step toward the "grand synthesis" identified by Fienberg and Manrique-Vallier (2009) in the context of multiple recapture estimation, but further development is needed to integrate auxiliary information into the matching process.

To facilitate Bayesian inference using our approach, we extended standard probit regression techniques to latent multinomial models where both the dimension and zeros of the response are unobserved due to imperfect detection and misidentification. We note that in the absence of misidentification (i.e., $\alpha = 1$), our probit model provides a convenient Gibbs sampler for Bayesian analysis of traditional closed population capture–recapture data with heterogeneous detection probabilities. By avoiding the need to tune proposal distributions, our probit formulation is potentially a more efficient alternative to traditional capture–capture models that rely on the logit link function to account for variability in detection probability or other parameters. However, we note that the logit link is sometimes desirable due to its ease of interpretation of the resulting odds-ratio; recent work by Polson, Scott and Windle (2013) could potentially be adapted to yield a Gibbs sampler for capture–recapture models using the logit link.

Capture–recapture models are more robust to individual capture heterogeneity when absolute abundance is not the focal parameter [e.g., Williams, Nichols and Conroy (2002)]. In this case, it may be more sensible to focus on individual heterogeneity in demographic parameters, such as survival probability [e.g., Gimenez and Choquet (2010), Royle (2008)]. A similarly-structured MCMC algorithm to those described in Sections 2.2.1 and 2.2.2 can be employed for other capture–recapture models extended for misidentification, including open population models, such as the Cormack–Jolly–Seber (CJS) and more recent multi-state formulations [e.g., Morrison et al. (2011), Pradel (2005)]. This is accomplished by substituting the desired form for the likelihood $[\mathbf{h}|\boldsymbol{\theta}, \boldsymbol{\rho}]$ in equation (3) and assigning corresponding priors for $\boldsymbol{\theta}$ and $\boldsymbol{\rho}$. The proposal density $[\mathbf{h}^*|\mathbf{h}, \boldsymbol{\theta}, \boldsymbol{\rho}]$ and the set of basis vectors $\{\mathbf{v}\}$ used to update $\mathbf{h}$ and $\mathbf{x}$, respectively, will depend on the particular model and the relationship between recorded and latent histories (formally described by $\mathbf{A}$).

Bonner and Holmberg (2013) and McClintock et al. (2013a) recently developed methods for integrated analyses of multiple sources of capture–recapture data, such as those arising from photo and DNA records. The methods developed in this paper for incorporating parameter heterogeneity could be extended to these latent multinomial models as well. Covariates explaining individual heterogeneity in parameters [e.g., King, Brooks and Coulson (2008)] could also be accommodated. While we have generalized the approach of Link et al. (2010) to a broader suite of misidentification models, we have maintained several key assumptions that may not be reasonable for many passive sampling data sets (e.g., those based on visual sightings). Some challenging (but needed) extensions include the evolving mark problem examined by Yoshizaki et al. (2009), allowing for ghost histories to consist of multiple encounters, and allowing identification errors to match legitimate individuals.

Acknowledgments. D. Etter and P. Lukacs for helpful discussions and assistance with the bear data. K. Scribner and S. Winterstein for pivotal roles in the

bear study. D. Johnson for helpful discussions. The findings and conclusions in the paper are those of the author(s) and do not necessarily represent the views of the National Marine Fisheries Service, NOAA. Any use of trade, product or firm names does not imply an endorsement by the US Government.

REFERENCES

ALBERT, J. H. and CHIB, S. (1993). Bayesian analysis of binary and polychotomous response data. *J. Amer. Statist. Assoc.* **88** 669–679. MR1224394

BAILEY, L. L. (2004). Evaluating elastomer marking and photo-identification methods for terrestrial salamanders: Marking effects and observer bias. *Herpetological Review* **35** 38–41.

BASU, S. and EBRAHIMI, N. (2001). Bayesian capture–recapture methods for error detection and estimation of population size: Heterogeneity and dependence. *Biometrika* **88** 269–279. MR1841274

BONNER, S. J. and HOLMBERG, J. (2013). Mark-recapture with multiple, non-invasive marks. *Biometrics* **69** 766–775. MR3106605

BONNER, S. J. and SCHOFIELD, M. R. (2013). MC(MC)MC: Exploring Monte Carlo integration with MCMC for mark-recapture models with individual covariates. *Methods in Ecology and Evolution*. DOI:10.1111/2041-210X.12095.

CASTLEDINE, B. J. (1981). A Bayesian analysis of multiple-recapture sampling for a closed population. *Biometrika* **68** 197–210. MR0614956

COULL, B. A. and AGRESTI, A. (1999). The use of mixed logit models to reflect heterogeneity in capture–recapture studies. *Biometrics* **55** 294–301.

DARROCH, J. N. (1958). The multiple-recapture census. I. Estimation of a closed population. *Biometrika* **45** 343–359. MR0119360

DORAZIO, R. M. and RODRIGUEZ, D. T. (2012). A Gibbs sampler for Bayesian analysis of site-occupancy data. *Methods in Ecology and Evolution* **3** 1093–1098.

DREHER, B. P., WINTERSTEIN, S. R., SCRIBNER, K. T., LUKACS, P. M., ETTER, D. R., ROSA, G. J. M., LOPEZ, V. A., LIBANTS, S. and FILCEK, K. B. (2007). Noninvasive estimation of black bear abundance incorporating genotyping errors and harvested bears. *Journal of Wildlife Management* **71** 2684–2693.

FIENBERG, S. E., JOHNSON, M. S. and JUNKER, B. W. (1999). Classical multilevel and Bayesian approaches to population size estimation using multiple lists. *J. Roy. Statist. Soc. Ser. A* **162** 383–406.

FIENBERG, S. E. and MANRIQUE-VALLIER, D. (2009). Integrated methodology for multiple systems estimation and record linkage using a missing data formulation. *AStA Adv. Stat. Anal.* **93** 49–60. MR2476299

GEORGE, E. I. and ROBERT, C. P. (1992). Capture–recapture estimation via Gibbs sampling. *Biometrika* **79** 677–683. MR1209469

GIMENEZ, O. and CHOQUET, R. (2010). Individual heterogeneity in studies on marked animals using numerical integration: Capture–recapture mixed models. *Ecology* **91** 951–957.

HALL, A. J., MCCONNELL, B. J. and BARKER, R. J. (2001). Factors affecting first-year survival in grey seals and their implications for life history strategy. *Journal of Animal Ecology* **70** 138–149.

HASTINGS, K. K., HIBY, L. A. and SMALL, R. J. (2008). Evaluation of a computer-assisted photograph-matching system to monitor naturally marked harbor seals at Tugidak Island, Alaska. *Journal of Mammalogy* **89** 1201–1211.

JOHNSON, D. S., CONN, P. B., HOOTEN, M. B., RAY, J. C. and PONE, B. A. (2013). Spatial occupancy models for large data sets. *Ecology* **94** 801–808.

KARANTH, K. U. and NICHOLS, J. D. (1998). Estimation of tiger densities in India using photographic captures and recaptures. *Ecology* **79** 2852–2862.

KAUFFMAN, M. J., FRICK, W. F. and LINTHICUM, J. (2003). Estimation of habitat-specific demography and population growth for peregrine falcans in California. *Ecological Applications* **13** 1802–1816.

KERNIGHAN, B. W. and RITCHIE, D. M. (1988). *The C Programming Language*, 2nd ed. Prentice Hall, Englewood Cliffs, NJ.

KING, R. and BROOKS, S. P. (2008). On the Bayesian estimation of a closed population size in the presence of heterogeneity and model uncertainty. *Biometrics* **64** 816–824. MR2526632

KING, R., BROOKS, S. P. and COULSON, T. (2008). Analyzing complex capture–recapture data in the presence of individual and temporal covariates and model uncertainty. *Biometrics* **64** 1187–1195. MR2522267

LANGTIMM, C. A., O'SHEA, T. J., PRADEL, R. and BECK, C. A. (1998). Estimates of annual survival probabilities for adult Florida manatees *Trichechus manatus latirostris*. *Ecology* **79** 981–997.

LINK, W. A. (2013). A cautionary note on the discrete uniform prior for the binomial N. *Ecology* **94** 2173–2179.

LINK, W. A., YOSHIZAKI, J., BAILEY, L. L. and POLLOCK, K. H. (2010). Uncovering a latent multinomial: Analysis of mark-recapture data with misidentification. *Biometrics* **66** 178–185. MR2756704

LUKACS, P. M. and BURNHAM, K. P. (2005). Estimating population size from DNA-based closed capture–recapture data incorporating genotyping error. *Journal of Wildlife Management* **69** 396–403.

MACKEY, B. L., DURBAN, J. W., MIDDLEMAS, S. J. and THOMPSON, P. M. (2008). A Bayesian estimate of harbour seal survival using sparse photo-identification data. *Journal of Zoology* **274** 18–27.

MANRIQUE-VALLIER, D. and FIENBERG, S. E. (2008). Population size estimation using individual level mixture models. *Biom. J.* **50** 1051–1063. MR2649394

MCCLINTOCK, B. T., CONN, P. B., ALONSO, R. S. and CROOKS, K. R. (2013a). Integrated modeling of bilateral photo-identification data in mark-recapture analyses. *Ecology* **94** 1464–1471.

MCCLINTOCK, B. T., HILL, J. M., FRITZ, L., CHUMBLEY, K., LUXA, K. and DIEFENBACH, D. R. (2013b). Mark-resight abundance estimation under incomplete identification of marked individuals. *Methods in Ecology and Evolution*. DOI:10.1111/2041-210X.12140.

MORRISON, T. A., YOSHIZAKI, J., NICHOLS, J. D. and BOLGER, D. T. (2011). Estimating survival in photographic capture–recapture studies: Overcoming misidentification error. *Methods in Ecology and Evolution* **2** 454–463.

OTIS, D. L., BURNHAM, K. P., WHITE, G. C. and ANDERSON, D. R. (1978). Statistical-inference from capture data on closed animal populations. *Wildlife Monographs* **62** 7–135.

PLEDGER, S. (2000). Unified maximum likelihood estimates for closed capture–recapture models using mixtures. *Biometrics* **56** 434–442.

PLUMMER, M., BEST, N., COWLES, K. and VINES, K. (2006). CODA: Convergence diagnosis and output analysis for MCMC. *R News* **6** 7–11.

POLSON, N. G., SCOTT, J. G. and WINDLE, J. (2013). Bayesian inference for logistic models using Pólya–Gamma latent variables. *J. Amer. Statist. Assoc.* **108** 1339–1349. MR3174712

PRADEL, R. (2005). Multievent: An extension of multistate capture–recapture models to uncertain states. *Biometrics* **61** 442–447. MR2140915

R CORE TEAM (2012). *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0.

ROYLE, J. A. (2008). Modeling individual effects in the Cormack–Jolly–Seber model: A state-space formulation. *Biometrics* **64** 364–370, 664. MR2432405

ROYLE, J. A., DORAZIO, R. M. and LINK, W. A. (2007). Analysis of multinomial models with unknown index using data augmentation. *J. Comput. Graph. Statist.* **16** 67–85. MR2345748

RUELL, E. W., RILEY, S. P. D., DOUGLAS, M. R., POLLINGER, J. P. and CROOKS, K. R. (2009). Estimating bobcat population sizes and densities in a fragmented urban landscape using noninvasive capture–recapture sampling. *Journal of Mammalogy* **90** 129–135.

TANCREDI, A. and LISEO, B. (2011). A hierarchical Bayesian approach to record linkage and population size problems. *Ann. Appl. Stat.* **5** 1553–1585. MR2849786

TANCREDI, A., AUGER-MÉTHÉ, M., MARCOUX, M. and LISEO, B. (2013). Accounting for matching uncertainty in two stage capture–recapture experiments using photographic measurements of natural marks. *Environ. Ecol. Stat.* **20** 647–665. MR3128764

THOMPSON, S. K. (1992). *Sampling*. Wiley, New York. MR1193031

WHITE, G. C. and BURNHAM, K. P. (1999). Program MARK: Survival estimation from populations of marked animals. *Bird Study* **46 Supplement** 120–138.

WILLIAMS, B. K., NICHOLS, J. D. and CONROY, M. J. (2002). *Analysis and Management of Animal Populations*. Academic Press, San Diego, CA.

WRIGHT, J. A., BARKER, R. J., SCHOFIELD, M. R., FRANTZ, A. C., BYROM, A. E. and GLEESON, D. M. (2009). Incorporating genotype uncertainty into mark-recapture-type models for estimating abundance using DNA samples. *Biometrics* **65** 833–840. MR2649856

YIP, P. S. F., BRUNO, G., TAJIMA, N., SEBER, G. A. F., BUCKLAND, S. T., CORMACK, R. M., UNWIN, N., CHANG, Y. F., FIENBERG, S. E., JUNKER, B. W., LAPORTE, R. E., LIBMAN, I. M. and MCCARTY, D. J. (1995a). Capture–recapture and multiple-record systems estimation I: History and theoretical development. *American Journal of Epidemiology* **142** 1047–1058.

YIP, P. S. F., BRUNO, G., TAJIMA, N., SEBER, G. A. F., BUCKLAND, S. T., CORMACK, R. M., UNWIN, N., CHANG, Y. F., FIENBERG, S. E., JUNKER, B. W., LAPORTE, R. E., LIBMAN, I. M. and MCCARTY, D. J. (1995b). Capture–recapture and multiple-record systems estimation II: Applications in human diseases. *American Journal of Epidemiology* **142** 1059–1068.

YOSHIZAKI, J. (2007). Use of natural tags in closed population capture–recapture studies: Modeling misidentification. Ph.D. thesis, North Carolina State Univ., Raleigh, NC.

YOSHIZAKI, J., POLLOCK, K. H., BROWNIE, C. and WEBSTER, R. A. (2009). Modeling misidentification errors in capture–recapture studies using photographic identification of evolving marks. *Ecology* **90** 3–9.

YOSHIZAKI, J., BROWNIE, C., POLLOCK, K. H. and LINK, W. A. (2011). Modeling misidentification errors that result from use of genetic tags in capture–recapture studies. *Environ. Ecol. Stat.* **18** 27–55. MR2783681

B. T. MCCLINTOCK
NATIONAL MARINE MAMMAL LABORATORY
ALASKA FISHERIES SCIENCE CENTER, NOAA-NMFS
7600 SAND POINT WAY NE
SEATTLE, WASHINGTON 98115
USA
E-MAIL: brett.mcclintock@noaa.gov

B. P. DREHER
COLORADO PARKS AND WILDLIFE
4255 SINTON ROAD
COLORADO SPRINGS, COLORADO 80907
USA
E-MAIL: brian.dreher@state.co.us

L. L. BAILEY
DEPARTMENT OF FISH, WILDLIFE, AND CONSERVATION BIOLOGY
COLORADO STATE UNIVERSITY
FORT COLLINS, COLORADO 80523
USA
E-MAIL: larissa.bailey@colostate.edu

W. A. LINK
USGS PATUXENT WILDLIFE RESEARCH CENTER
12100 BEECH FOREST ROAD
LAUREL, MARYLAND 20708
USA
E-MAIL: wlink@usgs.gov