Ecology, 84(4), 2003, pp. 1058–1066
© 2003 by the Ecological Society of America

ADJUSTING MULTISTATE CAPTURE–RECAPTURE MODELS FOR MISCLASSIFICATION BIAS: MANATEE BREEDING PROPORTIONS

WILLIAM L. KENDALL,[1] JAMES E. HINES, AND JAMES D. NICHOLS

USGS Patuxent Wildlife Research Center, 11510 American Holly Drive, Laurel, Maryland 20708-4017 USA

Abstract. Matrix population models are important tools for research and management of populations. Estimating the parameters of these models is an important step in applying them to real populations. Multistate capture–recapture methods have provided a useful means for estimating survival and parameters of transition between locations or life history states but have mostly relied on the assumption that the state occupied by each detected animal is known with certainty. Nevertheless, in some cases animals can be misclassified. Using multiple capture sessions within each period of interest, we developed a method that adjusts estimates of transition probabilities for bias due to misclassification. We applied this method to 10 years of sighting data for a population of Florida manatees (*Trichechus manatus latirostris*) in order to estimate the annual probability of transition from nonbreeding to breeding status. Some sighted females were unequivocally classified as breeders because they were clearly accompanied by a first-year calf. The remainder were classified, sometimes erroneously, as nonbreeders because an attendant first-year calf was not observed or was classified as more than one year old. We estimated a conditional breeding probability of 0.31 ± 0.04 (estimate ± 1 SE) when we ignored misclassification bias, and 0.61 ± 0.09 when we accounted for misclassification.

Key words: breeding probabilities; capture–resighting models; Florida; manatee; Markovian transition probabilities; multistate models; photo identification; Pollock's robust design; survival probabilities; Trichechus manatus.

INTRODUCTION

Matrix population models can often be useful to population and evolutionary ecologists, conservation biologists, and wildlife managers, for describing or predicting the dynamics of a population over time (e.g., Tuljapurkar 1990, Caswell 2001). Within these models, animals are assigned a state at each point in time. In some cases, transitions between states occur deterministically, as with age-based models such as Leslie matrices (Bernardelli 1941, Lewis 1942, Leslie 1945, 1948). We refer to age transitions as deterministic, because once an animal's age is known for any time period, its age at all previous and subsequent time periods is known with certainty. In contrast, with stage-based models (e.g., Lefkovitch 1965, Goodman 1969, Crouse et al. 1987), transitions are not obligatory at each time step, and transition parameters can be treated as expected proportions (fixed proportions of the animals in a given state that survive the subsequent time step make prescribed transitions) or probabilities (each animal ''flips a coin'' or series of coins at each point in time to determine whether or not it makes a transition to a different state).

If the intent is to apply a model to a real population, then values for the parameters of the model need to be determined or estimated from field data (Nichols et al. 1992). Multistate capture–recapture statistical models are a relatively recent development for estimating these parameters when detection of individuals is less than certain. Arnason (1972, 1973) pioneered this development for animals that can move among different sampled locations, and Nichols et al. (1992, 1994) noted that this class of models could also be useful for animals making transitions among a finite number of physiological or behavioral states. The early models of Arnason (1972, 1973) were generalized by Brownie et al. (1993) and Schwarz et al. (1993), and appropriate software was developed (Hines 1994, White and Burnham 1999). These advances have resulted in many applications of multistate capture–recapture models, as reviewed recently by Lebreton and Pradel (2002). These models and subsequent applications all assume that, in each case, the true state of an encountered animal is known with certainty. The only exception to this general statement involves estimation methods developed for the case in which a member of the population might occupy an unobservable state (e.g., it is in a location that is not exposed to capture efforts; see review in Kendall and Nichols [2002]).

Despite the appeal of multistate models for estimation of state transition probabilities, there are many cases in which an animal can be observed but its current state is unknown, or is not known with certainty. For example, the young of many species cannot be sexed without internal examination or genetic analysis of blood or tissue. For some species, such as Sandhill

Manuscript received 15 May 2002; revised 9 September 2002; accepted 17 September 2002. Corresponding Editor: F. R. Adler.
[1] E-mail: William_Kendall@usgs.gov

Cranes (*Grus canadensis*), this sexual monomorphism continues throughout life, and the only external way to assign sex is either indirectly through morphometrics, or directly through behavioral cues such as courtship displays (S. Nesbitt, *personal communication*). Often state is defined by reproductive activity, and individuals observed in certain activities (e.g., birds incubating on nests) can be unambiguously assigned to the ''breeder'' state, but nonbreeders are difficult to identify with certainty.

Despite the fact that uncertainty in state assignment is a common phenomenon in capture–recapture studies, there has been relatively little statistical work on this issue. An ad hoc approach to this problem is to define an additional state and label it as ''unknown'' or ''uncertain'' (e.g., Wood et al. 1998, Conroy et al. 1999). An animal would be assigned to this state unless its true state could be assigned with certainty. However, this uncertain state would consist of a mixture of true states (e.g., males and females). The resulting heterogeneity in demographic or sampling parameters could cause bias problems in estimates of interest. Lebreton and Pradel (2002) considered the problem of uncertain state assignment and outlined a general approach to the problem. Fujiwara and Caswell (2002) developed an approach that uses assignment probabilities that are assumed to be known from other sources of information. In this paper, we also use extra information available by trapping at two different temporal scales, under the robust design of Pollock (1982), and we develop a joint likelihood that incorporates the uncertainty associated with estimates of assignment or classification probabilities obtained from this additional information.

Specifically, we focus on estimation problems associated with developing a stage-based population projection model of Florida manatees (*Trichechus manatus latirostris*), although we believe that the generalities and some of the specifics of this case could be applied to many other taxa. Here the state of interest is female breeding status, and the principal question is ''given that an adult female did or did not produce a calf last year, what is the probability that she produces a calf this year?'' In this photo-identification study of naturally marked individuals, an adult female's breeding status is determined by whether or not she has a first-year calf with her when she is sighted. The crux of the problem is this: there is a chance that the calf will not be visible, even when its mother is, thus causing the female to be classified as a nonbreeder when, in fact, she is a breeder. We will first outline the design of an ongoing sight–resighting study of manatees. We will then present an analysis of these data using the Arnason-Schwarz multistate estimation method, ignoring the misclassification problem previously described. Finally, we will outline our robust design approach for simultaneously estimating and adjusting for this probability of misclassification, using multiple sighting periods within each year. We will compare estimates of survival rates and conditional breeding probabilities with and without this adjustment.

Manatee Mark–Resighting Study

Field methods and initial data selection

The manatee study population occurs along the north Gulf coast of Florida, USA. In winter (November–February) when waters turn cold, individuals in the region converge on two artesian-spring, warm-water refuges, the Crystal and Homosassa rivers. In warm months (March–September), individuals disperse south toward Tampa Bay and west toward Alabama to feed near shore in estuarine or freshwater habitats (Powell and Rathbun 1984, Rathbun et al. 1995). Mating generally is observed from February to July. Births generally occur from May to September and only rarely in winter. Gestation is ~12 mo. Litter size is usually one, with calf dependency of 1–2 yr. Site fidelity to the winter refuges is high among observed individuals (Rathbun et al. 1995).

Since 1978, the USGS Sirenia Project has annually photographed and documented sightings at the winter aggregation sites of known individuals, identified by distinctive scar patterns and features. Manatees routinely acquire marks in the wild from nonfatal boat injuries, fishing line entanglements, and lesions from fungal infections and cold damage. Slide transparencies were taken of marks, using underwater cameras fitted with a wide-angle lens as photographers swam near the manatees. Photographs were screened yearly for matches with known individuals entered in a photo catalog, the Manatee Individual Photo-identification System (MIPS). Digitized images and PC-based search technologies assisted researchers in matching photographs to cataloged individuals. Data on sex and reproductive status (with a first-year, second-year, or unknown-age calf) were collected at each sighting. Beck and Reid (1995) described the MIPS system and protocols in detail.

Data used to construct female capture histories were restricted to sightings only after the individual met strict MIPS criteria for cataloging and adult status (see Langtimm et al. [1998] for a complete discussion). The 90-d sampling period each year was defined as 15 November through 12 February, when females are most easily photographed at the sites, births are rare, and the likelihood of weaning is low. A sighting history was constructed for each female, consisting of its nonsighting (0) during the winter sample or its sighting with (C) or without (N) a first-year calf for each year of the study. Females were coded as (C) if they were seen at least once during the winter sample period with an attendant first-year calf. Great effort was expended to age calves and match them with the mother. However, if reasonable doubt remained about age or association, a female was assigned to state N. The final data set consisted of the sighting histories of 120 females from

TABLE 1. Number of adult female manatees from the northwest Florida population that are sighted each year with (R_i^{C}), without (R_i^{N}), or apparently without ($R_i^{N\prime}$) a first-year calf, that are next sighted t years later with ($m_{i,i+t}^{C}$) or without ($m_{i,i+t}^{N\prime}$), a first-year calf, respectively.

Year and initial status	Initial no. adult females	No. adult female manatees with or without first-year calf, year $i + t$									
		$m_{i,i+1}^{C}$	$m_{i,i+1}^{N\prime}$	$m_{i,i+2}^{C}$	$m_{i,i+2}^{N\prime}$	$m_{i,i+3}^{C}$	$m_{i,i+3}^{N\prime}$	$m_{i,i+4}^{C}$	$m_{i,i+4}^{N\prime}$	$m_{i,i+5}^{C}$	$m_{i,i+5}^{N\prime}$
1984											
R_i^{C}	10	0	6†	2	1	0	1	0	0	0	0
R_i^{N}†	6	0	0	0	5	0	0	0	0	0	0
$R_i^{N\prime}$	9	1	4	1	2	0	1	0	0	0	0
1985											
R_i^{C}	7	0	5†	1	0	0	0	0	1	0	0
R_i^{N}†	6	4	2	0	0	0	0	0	0	0	0
$R_i^{N\prime}$	19	2	14	0	2	0	0	0	0	0	0
1986											
R_i^{C}	13	0	12†	1	0	0	0	0	0	0	0
R_i^{N}†	5	0	4	0	0	0	0	0	0	0	0
$R_i^{N\prime}$	39	7	24	0	4	0	1	0	0	1	0
1987											
R_i^{C}	11	0	9†	0	1	1	0	0	0	0	0
R_i^{N}†	12	5	2	1	2	1	0	0	0	0	0
$R_i^{N\prime}$	41	9	22	1	3	0	0	0	1	0	0
1988											
R_i^{C}	15	0	10†	2	2	0	1	0	0	0	0
R_i^{N}†	9	6	1	0	2	0	0	0	0	0	0
$R_i^{N\prime}$	36	12	15	0	6	0	1	0	0	0	1
1989											
R_i^{C}	22	0	12†	3	4	0	1	0	1	0	0
R_i^{N}†	10	1	6	0	2	1	0	0	0	0	0
$R_i^{N\prime}$	33	7	12	4	2	1	5	0	0	0	0
1990											
R_i^{C}	13	0	6†	2	3	0	0	0	1	0	0
R_i^{N}†	12	3	5	0	1	0	0	0	1	0	0
$R_i^{N\prime}$	34	0	14	9	6	0	2	0	0	0	2
1991											
R_i^{C}	12	0	11†	0	0	0	0	0	0	0	0
R_i^{N}†	6	2	3	0	1	0	0	0	0	0	0
$R_i^{N\prime}$	33	4	21	0	4	0	1	0	0	0	1
1992											
R_i^{C}	22	0	8†	1	5	0	0	0	3	1	0
R_i^{N}†	11	2	2	1	5	0	0	0	0	0	0
$R_i^{N\prime}$	47	4	19	1	7	0	0	0	3	0	1
1993											
R_i^{C}	6	0	4†	0	1	1	0	0	0	0	0
R_i^{N}†	8	1	6	0	1	0	0	0	0	0	0
$R_i^{N\prime}$	34	3	16	0	4	1	1	0	2	0	0

† Known to be without first-year calf, based on being sighted with first-year calf in the previous year.

winter 1984–1985 through winter 1993–1994. Sighting data across years are summarized in Table 1.

Estimation ignoring misclassification

Given two states (C, with first-year calf; and N, no first-year calf), S_i^{C}, S_i^{N} is the probability that a female with or without, respectively, a first-year calf in primary sampling period i survives to period $i + 1$ and remains part of the population; and ψ_i^{CC}, ψ_i^{NC} is the probability that an animal that is with or without, respectively, a first-year calf in primary sampling period i, and survives from period i to $i + 1$, will be with a first-year calf in period $i + 1$.

We applied the Arnason-Schwarz model (see Brownie et al. 1993) to manatee resighting data, using program MARK (White and Burnham 1999). The most general model that we fit was $\{S(b, T), \psi(b, T), p(b, T)\}$, indicating variation across breeding states (b) and years (T) in survival, breeding, and detection probabilities (p). We fixed $\psi_i^{CC} \equiv 0$, because the minimal interbirth interval is 30 mo for manatees (Lefebvre and O'Shea 1995). We then evaluated goodness of fit using

the Pearson test in program MSSURVIV (White 1983, Hines 1994), pooling for cells with expected values <2. This yielded a moderate lack of fit and a variance adjustment factor of 1.87 (χ^2 = 48, df = 26), which we used to compute the Akaike Information Criterion, adjusted for small sample size and lack of fit, $QAIC_c$ (see Burnham and Anderson 1998:53). Based on $QAIC_c$, the best-fitting models by far were $\{S(.,.), \psi(b,.), p(.,T)\}$ with $\Delta QAIC_c = 0$, and $\{S(b,.), \psi(b,.), p(.,T)\}$ with $\Delta QAIC_c = 0.27$. A "." indicates no variation across breeding state or year; e.g., $S(b,.)$ indicates that survival probability is different for breeders and nonbreeders, but is constant across years. Survival and breeding probability estimates from the former model were $\hat{S}_{.}^{C} = \hat{S}_{.}^{N} = 0.97 \pm 0.012$ (estimate ± 1 SE) and $\hat{\psi}_{.}^{NC} = 0.31 \pm 0.037$, respectively. For the latter model, $\hat{S}_{.}^{C} = 1.0 \pm <0.001$, but SE is unreliable when the estimate is at the extreme of its permissible range), and $\hat{S}_{.}^{N} = 0.96 \pm 0.016$. Under this model, therefore, female manatees that do not reproduce in a given year have a 31% chance of reproducing in the following year. However, these estimates are based on the incorrect assumption that all sighted adult females are assigned to the correct breeding state. In reality, some females with calves were misclassified as females without calves (see subsequent analyses).

The bias in estimated breeding probability induced by misclassification of breeders as nonbreeders can extend beyond reproductive parameters. This misclassification also implies that estimated survival and sighting probabilities for nonbreeders actually apply to a mixture of breeders and nonbreeders. This would tend to bias estimated differences in these parameters between breeders and nonbreeders. The superior fit of a model that ignores state-specific differences in survival rate and sighting probability, as previously noted, could indicate that the differences are really negligible, or could instead be an artifact of misclassification.

Adjusting for misclassification bias

To account for misclassification, we use multiple encounters with the same female within a year, and examine the sighting history of that female and her first-year calf. Then we estimate the probability that a calf is detected given that it is present, and thereby adjust breeding probabilities for this detection probability. Specifically, we need to split the sampling effort (in this case 90 d of searching) for a given primary period i (in this case a year) into two or more secondary sampling intervals (see Table 2 for a summary of within-year sighting histories for the manatee data set). We assume that the population of adult females and calves is closed (to additions or deletions) for the duration of the entire 90-d sampling effort for primary period i. This use of within-period encounter information to better estimate detection probability is the essence of Pollock's robust design (Pollock 1982). We also assume that there is no chance that a calf is assigned to the wrong female.

TABLE 2. Number of adult female manatees from the northwest Florida population of manatees that exhibit the indicated sighting history within each winter, 1984–1985 to 1993–1994.

	Sighting histories†							
	Known with calf					Known without first-year calf‡		
Year	CC	CN	NC	C0	0C	NN	N0	0N
1984	2	0	0	4	4	0	3	3
1985	2	0	2	1	2	1	3	2
1986	3	2	2	3	3	3	2	0
1987	4	2	0	3	2	6	5	1
1988	2	3	3	2	5	3	3	3
1989	5	1	1	10	5	6	2	2
1990	4	0	1	5	3	4	7	1
1991	2	2	0	2	6	2	0	4
1992	10	2	2	6	2	5	5	1
1993	3	1	0	1	1	3	1	4

† C and N denote being sighted with or without a first-year calf, respectively.

‡ Those adult females also sighted in the previous winter with a first-year calf.

For the manatee problem, we broke sightings for a given year into two sampling intervals. Using the notation in Table 3, we illustrate how to adjust for misclassification bias in the Appendix, by listing representative expected cell frequencies for the between-years portion of the model for years 1 to 3. As presented, the model is general, with no restrictions on transition probabilities. Even for such few periods, the model is complex. We provide an explanation for two of the listed cells:

$$E(m_{12}^{CN\prime}) = R_1^C(\phi_1^{CC} p_2^{C(1-\delta)} + \phi_1^{CN} p_2^N).$$

This cell represents those seen as breeders in year 1 and seen as apparent nonbreeders in year 2. The first term in parentheses on the right-hand side represents the expected number that breed in the following year but are mistaken for nonbreeders, whereas the second term represents the expected number that really are nonbreeders. For the manatee case, the first term drops out because we set $\psi_1^{CC} = 0$. The second cell represents those seen as apparent nonbreeders in years 1 and 2, and constitutes a mixture of breeders and nonbreeders in each year:

$$E(m_{12}^{N\prime N\prime}) = R_1^{N\prime}[\pi_1(\phi_1^{NC} p_2^{C(1-\delta)} + \phi_1^{NN} p_2^N) + (1 - \pi_1)(\phi_1^{CC} p_2^{C(1-\delta)} + \phi_1^{CN} p_2^N)].$$

The first term in brackets represents the number of these that really are nonbreeders in year 1 (their expected proportion of the total released as apparent nonbreeders is denoted by the mixture parameter π_1), and the second term represents the number of these that actually are breeders in year 1.

One sighting session per year is insufficient to estimate the parameters of this model. Therefore, we use

TABLE 3. Definitions of parameters and statistics for multistate models under the robust design with misclassification.

Notation	Definition
R_i^{C}	Number of adult females seen with first-year calf in at least one secondary sampling period in year i.
R_i^{N}	Number of adult females seen in year i known to have no first-year calf (females cannot produce a new calf two years in a row, so this represents the subset of those seen in year i and also seen with a first-year calf in year $i - 1$).
$R_i^{\mathrm{N}'}$	Number of adult females seen in primary period i that apparently have no first-year calf (i.e., they could have one, but none was seen in either secondary sampling period).
m_{hi}^{rs}	Number of adult females seen in year h in state r next seen in year i in state s (e.g., r, s = C for with first-year calf, N for without first-year calf, N′ for apparently without first-year calf).
$X_i^{\mathrm{C}\delta(\omega)}$, $X_i^{\mathrm{N}(\omega)}$	Number of adult females (seen with first-year calf at least once in year i or known to have no first-year calf in year i, respectively) with sighting history $\omega_i \in \Omega_i$ across two samples within year i, where ω is series of 0's (when adult female not seen), C's (when first-year calf seen with adult female), and N's (when adult female seen but no first-year calf seen with her).
S_i^{C}, S_i^{N}	Probability that adult female with or without, respectively, first-year calf in primary sampling period i survives to period $i + 1$ and remains part of the population.
ψ_i^{CC}, ψ_i^{NC}	Probability that adult female with or without, respectively, a first-year calf in primary sampling period i, and that survives from period i to $i + 1$, will be with a first-year calf in period $i + 1$.
ϕ_i^{CC}, ϕ_i^{CN}	$S_i^{\mathrm{C}}\psi_i^{\mathrm{CC}}$, $S_i^{\mathrm{C}}(1 - \psi_i^{\mathrm{CC}})$, respectively.
ϕ_i^{NC}, ϕ_i^{NN}	$S_i^{\mathrm{N}}\psi_i^{\mathrm{NC}}$, $S_i^{\mathrm{N}}(1 - \psi_i^{\mathrm{NC}})$, respectively.
p_{ij}^{C}	Probability that adult female alive, in the population, and with a first-year calf is sighted (does not assume calf is sighted) in secondary sample j of year i.
δ_{ij}^{C}	Conditional on an adult female with a first-year calf being seen in sample j of year i, the probability that her calf is seen with her at that time.
$p_i^{\mathrm{C}\delta}$	Probability that adult female alive, in the population, and with a calf is sighted with her calf in year i. For l_i samples per year, $p_i^{\mathrm{C}\delta} = 1 - \Pi_{j=1}^{l_i}(1 - p_{ij}^{\mathrm{C}}\,\delta_{ij}^{\mathrm{C}})$.
$p_i^{\mathrm{C}(1-\delta)}$	Probability that adult female alive, in the population, and with a first-year calf is sighted in year i, but her calf is not sighted, $p_i^{\mathrm{C}(1-\delta)} = \Pi_{j=1}^{l_i}(1 - p_{ij}^{\mathrm{C}}\,\delta_{ij}^{\mathrm{C}}) - \Pi_{j=1}^{l_i}(1 - p_{ij}^{\mathrm{C}})$.
p_{ij}^{N}	Probability that adult female alive, in the population, and without a calf is sighted in secondary sample j of year i.
p_i^{N}	Probability that adult female alive, in the population, and without a calf is sighted in year i: $p_i^{\mathrm{N}} = 1 - \Pi_{j=1}^{l_i}(1 - p_{ij}^{\mathrm{N}})$.
π_i	Probability that adult female seen and "released" in year i, apparently without a first-year calf, indeed has no first-year calf.

within-year information to estimate, $p_i^{\mathrm{C}\delta}$, $p_i^{\mathrm{C}(1-\delta)}$, and p_i^{N}, thereby permitting the estimation of π_i, ψ_i^{CC}, ψ_i^{NC}, S_i^{C}, and S_i^{N}. For two secondary samples in a given year, conditioning on the number of adult females seen at least once with a calf in year i (denoted as $m_{\cdot i}^{\cdot\mathrm{C}}$, specifically the number of animals observed as with-calf in year i, regardless of the year and state of last observation), and assuming that adult sighting probability and conditional calf sighting probability vary only by time,

$$E(X_i^{\mathrm{C}\delta(\mathrm{CC})}) = m_{\cdot i}^{\cdot\mathrm{C}}\frac{p_{i1}^{\mathrm{C}}\delta_{i1}^{\mathrm{C}}p_{i2}^{\mathrm{C}}\delta_{i2}^{\mathrm{C}}}{[1 - (1 - p_{i1}^{\mathrm{C}}\delta_{i1}^{\mathrm{C}})(1 - p_{i2}^{\mathrm{C}}\delta_{i2}^{\mathrm{C}})]} \quad (1a)$$

$$E(X_i^{\mathrm{C}\delta(\mathrm{CN})}) = m_{\cdot i}^{\cdot\mathrm{C}}\frac{p_{i1}^{\mathrm{C}}\delta_{i1}^{\mathrm{C}}p_{i2}^{\mathrm{C}}(1 - \delta_{i2}^{\mathrm{C}})}{[1 - (1 - p_{i1}^{\mathrm{C}}\delta_{i1}^{\mathrm{C}})(1 - p_{i2}^{\mathrm{C}}\delta_{i2}^{\mathrm{C}})]} \quad (1b)$$

$$E(X_i^{\mathrm{C}\delta(\mathrm{NC})}) = m_{\cdot i}^{\cdot\mathrm{C}}\frac{p_{i1}^{\mathrm{C}}(1 - \delta_{i1}^{\mathrm{C}})p_{i2}^{\mathrm{C}}\delta_{i2}^{\mathrm{C}}}{[1 - (1 - p_{i1}^{\mathrm{C}}\delta_{i1}^{\mathrm{C}})(1 - p_{i2}^{\mathrm{C}}\delta_{i2}^{\mathrm{C}})]} \quad (2a)$$

$$E(X_i^{\mathrm{C}\delta(\mathrm{C0})}) = m_{\cdot i}^{\cdot\mathrm{C}}\frac{p_{i1}^{\mathrm{C}}\delta_{i1}^{\mathrm{C}}(1 - p_{i2}^{\mathrm{C}})}{[1 - (1 - p_{i1}^{\mathrm{C}}\delta_{i1}^{\mathrm{C}})(1 - p_{i2}^{\mathrm{C}}\delta_{i2}^{\mathrm{C}})]} \quad (2b)$$

$$E(X_i^{\mathrm{C}\delta(0\mathrm{C})}) = m_{\cdot i}^{\cdot\mathrm{C}}\frac{(1 - p_{i1}^{\mathrm{C}})p_{i2}^{\mathrm{C}}\delta_{i2}^{\mathrm{C}}}{[1 - (1 - p_{i1}^{\mathrm{C}}\delta_{i1}^{\mathrm{C}})(1 - p_{i2}^{\mathrm{C}}\delta_{i2}^{\mathrm{C}})]} \quad (3)$$

Next, conditioning on the number of adult females that are seen at least once and are known to not have a first-year calf in year i (denote these as $m_{i-1,i}^{\mathrm{CN}}$, females observed in year i that were observed with calf in the previous year),

$$E(X_i^{\mathrm{N}(\mathrm{NN})}) = m_{i-1,i}^{\mathrm{CN}}\frac{p_{i1}^{\mathrm{N}}p_{i2}^{\mathrm{N}}}{[1 - (1 - p_{i1}^{\mathrm{N}})(1 - p_{i2}^{\mathrm{N}})]} \quad (4a)$$

$$E(X_i^{\mathrm{N}(\mathrm{N0})}) = m_{i-1,i}^{\mathrm{CN}}\frac{p_{i1}^{\mathrm{N}}(1 - p_{i2}^{\mathrm{N}})}{[1 - (1 - p_{i1}^{\mathrm{N}})(1 - p_{i2}^{\mathrm{N}})]} \quad (4b)$$

$$E(X_i^{\mathrm{N}(0\mathrm{N})}) = m_{i-1,i}^{\mathrm{CN}}\frac{(1 - p_{i1}^{\mathrm{N}})p_{i2}^{\mathrm{N}}}{[1 - (1 - p_{i1}^{\mathrm{N}})(1 - p_{i2}^{\mathrm{N}})]}. \quad (5)$$

Expressions (4) and (5) essentially constitute a conditional Lincoln-Petersen model, as in Kendall et al. (1997), and for observed breeders, expressions (1) through (3) constitute an extension of this model. Under this formulation, detection and misclassification probabilities are estimated from within-year information from known breeders and known nonbreeders. The contribution from apparent nonbreeders is limited to estimation of survival and breeding probabilities through between-year detection information, as described in the Appendix.

Refined data selection, analysis, and results

We implemented the models for misclassification just described by developing a modification of program

TABLE 4. Comparison of fit for models that account for misclassification of adult female manatees with first-year calves, for the northwest Florida population, 1984–1985 to 1993–1994. The model is listed if numerical convergence is achieved.

Model	No. parameters	$\Delta QAIC_c$†	Akaike weight‡
$S(.,.)$, $\psi(b,.)$, $p(.,T,.)$, $\delta(.,.)$, $\pi(.)$	14	0	0.755
$S(.,.)$, $\psi(b,.)$, $p(.,T,.)$, $\delta(.,.)$, $\pi(T)$	22	2.3	0.239
$S(b,.)$, $\psi(b,.)$, $p(b, T,.)$, $\delta(.,.)$, $\pi(T)$	33	10.0	0.005
$S(.,T)$, $\psi(b,.)$, $p(b, T,.)$, $\delta(.,.)$, $\pi(.)$	32	13.7§	
$S(.,T)$, $\psi(b,.)$, $p(b, T,.)$, $\delta(.,.)$, $\pi(T)$	40	14.9	<0.001
$S(.,T)$, $\psi(b,.)$, $p(b, T,.)$, $\delta(.,t)$, $\pi(T)$	41	17.0§	
$S(.,.)$, $\psi(b,.)$, $p(.,T,.)$, $\delta(T,.)$, $\pi(T)$	31	17.4	<0.001
$S(b, T)$, $\psi(b,.)$, $p(b, T,.)$, $\delta(.,.)$, $\pi(T)$	49	31.3	<0.001

† Difference in Akaike Information Criterion, adjusted for small sample size and lack of fit based on a sample size of 708 releases and a variance inflation factor of 1.5.

‡ See Buckland et al. (1997) and Burnham and Anderson (1998).

§ Rounding error limited numerical convergence to one decimal place. Akaike weights were not computed.

SURVIV (White 1983), and more specifically MSSURVIV (Hines 1994), and call it MSSRVmis.[2] We validated our programming and confirmed the estimability of all parameters under this new model by analyzing, as real data, expected cell frequencies under a set of parameter values. This analytical-numerical approach is reviewed in Burnham et al. (1987:215).

We considered two secondary sighting sessions per year, so the resighting histories included a series of C's (adult female seen with first-year calf), N's (adult female seen without first-year calf), and 0's (adult female not seen), two for each year of the study. Equal detection probabilities for each sighting session within a season, although not necessary, produce more precise estimates. Therefore we delineated the time periods differently for each year to reflect a comparable number of sightings in each session. We began with a general model $\{S(b, T), \psi(b, T), p(b, T, t), \delta(T, t), \pi(T)\}$, where T refers to time variation across years and t denotes time variation across sampling periods within a year. We then considered versions of this model in which parameters are constrained equal over breeding state or time. Unfortunately, we could not fit many of the more general models (i.e., numerical convergence was not achieved to at least two decimal places), due to sparseness of data (Tables 1 and 2). This included all models in which conditional breeding probability was time dependent.

We have listed the models that we considered and could fit, along with a comparison of $QAIC_c$ values for each model, in Table 4. $QAIC_c$ values were based on a variance inflation factor of 1.5 (computed from a Pearson goodness-of-fit test as previously described: $\chi^2 = 71.1$, df = 41). Based on $QAIC_c$, the best-fitting model was $\{S(.,.),\psi(b,.),p(.,T,.),\delta(.,.),\pi(.)\}$. Survival and conditional breeding probability estimates under this model were 0.96 ± 0.01 (estimate ± 1 SE) and 0.61 ± 0.09, respectively. The survival rate estimate and its standard error are consistent with our multistate approach ignoring misclassification, and with Langtimm et al. (1998).

The estimate of 0.61 for ψ^{NC} (breeding probability, given nonbreeding status in the previous year) from the best-fitting model contrasts sharply with the estimate of 0.31 ± 0.037 when misclassification is ignored. Calf detection probability estimates for the best-fitting model were $\hat{\delta}_{i1} = \hat{\delta}_{i2} = 0.72 \pm 0.07$, and average sighting probability for adults per half-season session was 0.48 ± 0.04. This result implies that, on average, there was a 22% chance that a calf was missed, given that its mother was sighted in the winter of year i (i.e., on average $\hat{p}_i^{C(1-\delta)}/p_i^C = 0.22 \pm 0.05$, based on Taylor series approximation, where $p_i^C = [1 - (1 - p_{i1}^C)(1 - p_{i2}^C)]$, the probability that a female with a first-year calf is sighted at least once in year i).

Extent of bias when misclassification is ignored

These manatee results indicate at least one scenario in which misclassification of state can cause severe bias in breeding probabilities. To give more perspective on this problem, for several sets of parameter values we generated expected data under the misclassification model previously described and then analyzed the data using the Arnason-Schwarz model that ignores misclassification, to estimate resulting bias (see Burnham et al. 1987:215). Table 5 contains parameter values and expected values for estimators of survival and conditional breeding probability (ψ_i^{NC}). For simplicity, both the model that generated the data and the model used to analyze the data assumed that $\psi_i^{CC} = 0$ and that all parameters were equal over time.

In each case, bias in breeding probability was substantial. Even for the case in which there was a 90% chance that a calf was seen each time its mother was seen, there was still a 15% bias in conditional breeding probability. These scenarios are not exhaustive and therefore we do not claim generality, but in each case, the expected value of the $\hat{\psi}_i^{NC}$ was $<\psi_i^{NC}(p_{i+1}^{C\delta}/p_{i+1}^C)$,

[2] URL: ⟨www.pwrc-mbr.usgs.gov/software⟩

TABLE 5. Effect of ignoring misclassification of breeding state in the estimation of demographic parameters (defined in Table 3) using multistate capture–recapture models, based on a hypothetical five-year study.

Parameter			Expected value†		
p_{ij}^{C}	p_{ij}^{N}	δ_{ij}^{C}	$\hat{\psi}_i^{NC}$	$\hat{S}_i^{C}$	$\hat{S}_i^{N}$
A) Survival probability, $S_i^N = 0.80$					
0.60	0.70	0.50, 0.75	0.406, 0.536	0.798, 0.799	0.801, 0.800
0.70	0.60	0.50, 0.75	0.393, 0.538	0.797, 0.799	0.803, 0.801
0.50	0.50	0.50, 0.75	0.385, 0.523	0.795, 0.798	0.804, 0.802
B) Survival probability, $S_i^N = 0.65$					
0.60	0.70	0.50, 0.75	0.399, 0.526	0.797, 0.798	0.668, 0.666
0.70	0.60	0.50, 0.75	0.387, 0.528	0.794, 0.797	0.671, 0.667
0.50	0.50	0.50, 0.75	0.381, 0.520	0.790, 0.796	0.674, 0.661
C) Survival probability, $S_i^N = 0.80$					
0.60	0.60	0.90	0.598	0.800	0.800

Notes: Indicated parameter values are based on a sampling period partitioned into two halves, but we assume that this is ignored by the estimation method. Breeders survive with probability $S_i^C = 0.8$ but do not breed in two consecutive years; nonbreeders breed with probability $\hat{\psi}_i^{NC} = 0.7$ in the following year. All parameters are assumed constant across time.

† Based on analyzing expected data under model (see Burnham et al. 1987:214).

where the factor in parentheses is the probability that the calf is seen at some point in year $i + 1$, given that its mother is seen in that year. As expected, estimates of survival probability were virtually unbiased when survival was independent of state. When survival probability was different for breeders and nonbreeders, the bias was in the predicted direction, but was small for these cases. We anticipate that all biases indicated here would be more pronounced if $\psi_i^{CC} \neq 0$ or parameters were allowed to vary across time. An exhaustive evaluation of scenarios is impractical here. We suggest that the practitioner should evaluate scenarios pertinent to a particular case.

DISCUSSION

We have demonstrated how quite substantial bias in estimates of population parameters, due to misclassification of the state of an observed animal, can be corrected using an additional source of information. In this case, we exploited multiple sampling periods within a season, as in Pollock's robust design. This is extremely logical, given that the probability of misclassification is a direct function of the detection probability of, in this case a first-year calf, in other cases perhaps a behavior. In addition to the usual assumptions associated with multistate capture–recapture methods (see Brownie et al. 1993), we required three others.

First, we assumed that all adult females in the north Gulf coast population were equally exposed to sighting effort in each of the two sampling periods per season. Second, we assumed the population of adult females and calves was closed for each season. That is, if an adult female was in the study area for one of the two sampling periods within a season, then she was there for both sampling periods within that season. Similarly, if a female had a first-year calf at all in a given sampling period, then we assume that the calf was also present in the other sampling period in that season. Third, we assumed that a breeder can be misclassified as a nonbreeder, but not the reverse (i.e., a first-year calf that is seen is not assigned to the wrong mother).

To satisfy these assumptions, we selected a sampling period when both birth and weaning of a first-year calf is unlikely, and ignored observations of calves when age (first-year or older) or parentage was in doubt. In addition, calf weaning or mortality between the first and second sampling periods within a season should not cause bias as long as the probability of weaning or death is equal for calves seen or not seen in the first sampling period. As with the Lincoln-Petersen method (see Seber 1982:70), the probability of weaning or death is confounded with $1 - \delta_{i2}$, thereby simply lowering the effective sighting probability for the calf in the second period.

This method assumes that some female manatees are known to be nonbreeders. Due to the long gestation period, female manatees cannot produce a new calf in two consecutive years, so this is a reasonable approach. Although this will not be the case for many species, the method can be modified to remove this requirement, perhaps at the cost of precision.

We anticipate that this technique will be useful for other applications in which misclassification occurs. For example, some behaviors, such as incubating eggs on a nest, identify animals as breeders. Cam et al. (2002) acknowledged that pre-breeding kittiwakes (*Rissa tridactyla*) are difficult to partition into squatters and non-squatters, because squatting behavior is not consistently displayed. In epidemiological studies of marked animals, the clinical signs of a disease might be easy to identify when seen, but in a resighting study, such signs could be missed by an observer due to viewing conditions.

In their review of multistate models, Lebreton and Pradel (2002) described a general structure for considering uncertainty in assigning an animal to a state. They pointed out that ''Although potentially important, those multi-event models are likely to suffer from identifiability problems.'' Here we have solved those identifiability problems by exploiting the assumption that nonbreeders cannot be misclassified as breeders (i.e., for one state, each time an animal occupies that state, there is a non-negligible chance it will be unambiguously assigned), and by using the robust design to directly estimate the misclassification probability.

Acknowledgments

C. A. Langtimm and C. A. Beck, of the USGS Sirenia Project, provided the manatee photo-identification data. C. A. Langtimm provided very helpful comments and the context and design details of the manatee study. Comments by two anonymous referees improved the manuscript.

Literature Cited

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 in Population Ecology **15**:1–8.

Beck, C. A., and J. P. Reid. 1995. An automated photo-identification catalog for studies of the life history of the Florida manatee. Pages 120–134 *in* T. J. O'Shea, B. B. Ackerman, and H. F. Percival, editors. Population biology of the Florida manatee. U.S. Department of the Interior. National Biological Service. Information and Technology Report **1**.

Bernardelli, H. 1941. Population waves. Journal of the Burma Research Society **31**:1–18.

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

Buckland, S. T., K. P. Burnham, and N. H. Augustin. 1997. Model selection: an integral part of inference. Biometrics **38**:469–477.

Burnham, K. P., and D. R. Anderson. 1998. Model selection and inference. Springer-Verlag, New York, New York, USA.

Burnham, K. P., D. R. Anderson, G. C. White, C. Brownie, and K. H. Pollock. 1987. Design and analysis methods for fish survival experiments based on release–recapture. American Fisheries Society Monograph 5. Bethesda, Maryland, USA.

Cam, E., B. Cadiou, J. E. Hines, and J. Y. Monnat. 2002. Influence of behavioural tactics on recruitment and reproductive trajectory in the kittiwake. Pages 163–187 *in* B. J. T. Morgan and D. L. Thomson, editors. Statistical analysis of data from marked bird populations. Journal of Applied Statistics **29**(1–4).

Caswell, H. 2001. Matrix population models. Second edition. Sinauer, Sunderland, Massachusetts, USA.

Conroy, M. J., J. C. Senar, J. E. Hines, and J. Domenech. 1999. Development and application of a mark–recapture model incorporating predicted sex and transitory behavior. Bird Study **46**(supplement):S62–S73.

Crouse, D. T., L. B. Crowder, and H. Caswell. 1987. A stage-based population model for loggerhead sea turtles and implications for conservation. Ecology **68**:1412–1423.

Fujiwara, M., and H. Caswell. 2002. Estimating population projection matrices from multi-stage mark–recapture data. Ecology **83**:3257–3265.

Goodman, L. A. 1969. The analysis of population growth when the birth and death rates depend upon several factors. Biometrics **25**:659–681.

Hines, J. E. 1994. MSSURVIV user's manual. USGS Patuxent Wildlife Research Center, Laurel, Maryland, USA.

Kendall, W. L., and J. D. Nichols. 2002. Estimating state-transition probabilities for unobservable states using capture–recapture/resighting data. Ecology **83**:3276–3284.

Kendall, W. L., J. D. Nichols, and J. E. Hines. 1997. Estimating temporary emigration using capture–recapture data with Pollock's robust design. Ecology **78**:563–578.

Langtimm, C. A., T. J. O'Shea, R. Pradel, and C. A. Beck. 1998. Estimates of annual survival probabilities for adult Florida manatees (*Trichechus manatus latirostris*). Ecology **79**:981–997.

Lebreton, J.-D., and R. Pradel. 2002. Multistate recapture models: modelling incomplete individual histories. Pages 353–370 *in* B. J. T. Morgan and D. L. Thomson, editors. Statistical analysis of data from marked bird populations. Journal of Applied Statistics **29**(1–4).

Lefebvre, L. L., and T. J. O'Shea. 1995. Florida manatees. Pages 267–269 *in* E. T. Laroe, G. S. Farris, C. E. Puckett, P. D. Doran, and M. J. Mac, editors. Our living resources: coastal and marine ecosystems. National Biological Service, U.S. Government Printing Office, Washington, D.C., USA.

Leftkovich, L. P. 1965. The study of population growth in organisms grouped by stages. Biometrics **21**:1–18.

Leslie, P. H. 1945. On the use of matrices in certain population mathematics. Biometrika **33**:183–212.

Leslie, P. H. 1948. Some further notes on the use of matrices in population mathematics. Biometrika **35**:183–212.

Lewis, E. G. 1942. On the generation and growth of a population. Sankhya: The Indian Journal of Statistics **6**:93–96.

Nichols, J. D., J. E. Hines, K. H. Pollock, R. L. Hinz, and W. A. Link. 1994. Estimating breeding proportions and testing hypotheses about costs of reproduction with capture–recapture data. Ecology **75**:2052–2065.

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

Pollock, K. H. 1982. A capture–recapture design robust to unequal probability of capture. Journal of Wildlife Management **46**:757–760.

Powell, J. A., and G. B. Rathbun. 1984. Distribution and abundance of manatees along the northern coast of the Gulf of Mexico. Northeast Gulf Science **7**:1–28.

Rathbun, G. B., J. P. Reid, R. K. Bonde, and J. A. Powell. 1995. Reproduction in free-ranging Florida manatees. Pages 135–156 *in* T. J. O'Shea, B. B. Ackerman, and H. F. Percival, editors. Population biology of the Florida manatee. U.S. Department of the Interior, National Biological Service. Information and Technology Report **1**.

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

Seber, G. A. F. 1982. The estimation of animal abundance and related parameters. Second edition. Macmillan, New York, New York, USA.

Tuljapurkar, S. 1990. Population dynamics in variable environments. Springer-Verlag, New York, New York, USA.

White, G. C. 1983. Numerical estimation of survival rates from band-recovery and biotelemetry data. Journal of Wildlife Management **47**:716–728.

White, G. C., and K. P. Burnham. 1999. Program MARK for survival estimation. Bird Study **46**(supplement):120–139.

Wood, K. V., J. D. Nichols, H. F. Percival, and J. E. Hines. 1998. Size–sex variation in survival rates and abundance of pig frogs, *Rana grylio*, in northern Florida wetlands. Journal of Herpetology **32**:527–535.

APPENDIX

Examples of expected cell frequencies for a general model, primary periods 1 to 3. Specifics for manatees, where females cannot produce a new calf in two consecutive years, can be obtained by dropping terms that contain ϕ_i^{CC}.

$$E(m_{12}^{CC}) = R_1^C \phi_1^{CC} p_2^{C\delta}$$

$$E(m_{12}^{CN'}) = R_1^C(\phi_1^{CC} p_2^{C(1-\delta)} + \phi_1^{CN} p_2^N)$$

$$E(m_{13}^{CC}) = R_1^C[\phi_1^{CC}(1 - p_2^{C\delta} - p_2^{C(1-\delta)})\phi_2^{CC} p_3^{C\delta} + \phi_1^{CN}(1 - p_2^N)\phi_2^{NC} p_3^{C\delta}]$$

$$E(m_{13}^{CN'}) = R_1^C[\phi_1^{CC}(1 - p_2^{C\delta} - p_2^{C(1-\delta)})\phi_2^{CC} p_3^{C(1-\delta)} + \phi_1^{CC}(1 - p_2^{C\delta} - p_2^{C(1-\delta)})\phi_2^{CN} p_3^N + \phi_1^{CN}(1 - p_2^N)\phi_2^{NC} p_3^{C(1-\delta)} + \phi_1^{CN}(1 - p_2^N)\phi_2^{NN} p_3^N]$$

$$E(m_{12}^{NC}) = R_1^N \phi_1^{NC} p_2^{C\delta}$$

$$E(m_{12}^{N'C}) = R_1^{N'}[\pi_1 \phi_1^{NC} + (1 - \pi_1)\phi_1^{CC}] p_2^{C\delta}$$

$$E(m_{12}^{NN'}) = R_1^N[\phi_1^{NN} p_2^N + \phi_1^{NC} p_2^{C(1-\delta)}]$$

$$E(m_{12}^{N'N'}) = R_1^{N'}[\pi_1(\phi_1^{NC} p_2^{C(1-\delta)} + \phi_1^{NN} p_2^N) + (1 - \pi_1)(\phi_1^{CC} p_2^{C(1-\delta)} + \phi_1^{CN} p_2^N)]$$

$$E(m_{13}^{NC}) = R_1^N[\phi_1^{NC}(1 - p_2^{C\delta} - p_2^{C(1-\delta)})\phi_2^{CC} + \phi_1^{NN}(1 - p_2^N)\phi_2^{NC}] p_3^{C\delta}$$

$$E(m_{13}^{NN'}) = R_1^N[\phi_1^{NC}(1 - p_2^{C\delta} - p_2^{C(1-\delta)})\phi_2^{CC} p_3^{C(1-\delta)} + \phi_1^{NC}(1 - p_2^{C\delta} - p_2^{C(1-\delta)})\phi_2^{CN} p_3^N + \phi_1^{NN}(1 - p_2^N)\phi_2^{NC} p_3^{C(1-\delta)} + \phi_1^{NN}(1 - p_2^N)\phi_2^{NN} p_3^N]$$

$$E(m_{13}^{N'N'}) = R_1^N[\pi_1\{\phi_1^{NC}(1 - p_2^{C\delta} - p_2^{C(1-\delta)})\phi_2^{CC} p_3^{C(1-\delta)} + \phi_1^{NC}(1 - p_2^{C\delta} - p_2^{C(1-\delta)})\phi_2^{CN} p_3^N + \phi_1^{NN}(1 - p_2^N)\phi_2^{NC} p_3^{C(1-\delta)} + \phi_1^{NN}(1 - p_2^N)\phi_2^{NN} p_3^N\} + (1 - \pi_1) \times \{\phi_1^{CC}(1 - p_2^{C\delta} - p_2^{C(1-\delta)})\phi_2^{CC} p_3^{C(1-\delta)} + \phi_1^{CC}(1 - p_2^{C\delta} - p_2^{C(1-\delta)})\phi_2^{CN} p_3^N + \phi_1^{CN}(1 - p_2^N)\phi_2^{NC} p_3^{C(1-\delta)} + \phi_1^{CN}(1 - p_2^N)\phi_2^{NN} p_3^N\}].$$