Estimating Breeding Proportions and Testing Hypotheses about Costs of Reproduction with Capture-Recapture Data

James D. Nichols; James E. Hines; Kenneth H. Pollock; Robert L. Hinz; William A. Link

Ecology, Volume 75, Issue 7 (Oct., 1994), 2052-2065.

Stable URL:
http://links.jstor.org/sici?sici=0012-9658%28199410%2975%3A7%3C2052%3AEBPATH%3E2.0.CO%3B2-5

Your use of the JSTOR archive indicates your acceptance of JSTOR's Terms and Conditions of Use, available at http://www.jstor.org/about/terms.html. JSTOR's Terms and Conditions of Use provides, in part, that unless you have obtained prior permission, you may not download an entire issue of a journal or multiple copies of articles, and you may use content in the JSTOR archive only for your personal, non-commercial use.

Each copy of any part of a JSTOR transmission must contain the same copyright notice that appears on the screen or printed page of such transmission.

Ecology is published by The Ecological Society of America. Please contact the publisher for further permissions regarding the use of this work. Publisher contact information may be obtained at http://www.jstor.org/journals/esa.html.

Ecology
©1994 The Ecological Society of America

JSTOR and the JSTOR logo are trademarks of JSTOR, and are Registered in the U.S. Patent and Trademark Office. For more information on JSTOR contact jstor-info@umich.edu.

©2003 JSTOR

http://www.jstor.org/
Tue Apr 15 17:11:08 2003

Ecology, 75(7), 1994, pp. 2052–2065
© 1994 by the Ecological Society of America

ESTIMATING BREEDING PROPORTIONS AND TESTING HYPOTHESES ABOUT COSTS OF REPRODUCTION WITH CAPTURE–RECAPTURE DATA[1]

JAMES D. NICHOLS AND JAMES E. HINES
National Biological Survey, Patuxent Wildlife Research Center, Laurel, Maryland 20708 USA

KENNETH H. POLLOCK
Institute of Statistics, North Carolina State University, Box 8203, Raleigh, North Carolina 27695-8203 USA

ROBERT L. HINZ AND WILLIAM A. LINK
National Biological Survey, Patuxent Wildlife Research Center, Laurel, Maryland 20708 USA

Abstract. The proportion of animals in a population that breeds is an important determinant of population growth rate. Usual estimates of this quantity from field sampling data assume that the probability of appearing in the capture or count statistic is the same for animals that do and do not breed. A similar assumption is required by most existing methods used to test ecologically interesting hypotheses about reproductive costs using field sampling data. However, in many field sampling situations breeding and nonbreeding animals are likely to exhibit different probabilities of being seen or caught. In this paper, we propose the use of multistate capture–recapture models for these estimation and testing problems. This methodology permits a formal test of the hypothesis of equal capture/sighting probabilities for breeding and nonbreeding individuals. Two estimators of breeding proportion (and associated standard errors) are presented, one for the case of equal capture probabilities and one for the case of unequal capture probabilities. The multistate modeling framework also yields formal tests of hypotheses about reproductive costs to future reproduction or survival or both fitness components. The general methodology is illustrated using capture–recapture data on female meadow voles, *Microtus pennsylvanicus*. Resulting estimates of the proportion of reproductively active females showed strong seasonal variation, as expected, with low breeding proportions in midwinter. We found no evidence of reproductive costs extracted in subsequent survival or reproduction. We believe that this methodological framework has wide application to problems in animal ecology concerning breeding proportions and phenotypic reproductive costs.

Key words: capture/sighting probability; Microtus pennsylvanicus; *multistate capture–recapture models; proportion of animals breeding; reproductive costs; survival rate.*

INTRODUCTION

Variation among individual organisms simultaneously forms the basis for most of the interesting questions in animal ecology and evolutionary biology and creates difficult methodological problems for the sampling of animal populations in the field (Johnson et al. 1986). One source of individual variation that carries important consequences for fitness and population dynamics is reproductive effort. If, at some time i, some animals breed and some do not, then we would like to be able to estimate this breeding proportion in order to model population dynamics.

In order to estimate this proportion of animals that breed, we must consider possible variation in sampling probabilities of breeders and nonbreeders. For example, if we are using observational methods for sampling animals, then we must consider the possibility that different behaviors and activity patterns of breeders and nonbreeders may lead to different probabilities of seeing the two types of animals. If we are using capture–recapture methods, then we cannot ignore the possibility that the differing movement patterns and nutritional requirements of breeders and nonbreeders may produce different capture probabilities.

We may also be interested in whether breeding at time i carries fitness consequences as measured by changes in either the probability of surviving the interval between i and $i + 1$ or the probability of breeding at time $i + 1$. However, if individually marked breeders and nonbreeders exhibit different probabilities of appearing in our samples (e.g., of being resighted or recaptured at time $i + 1$), then we must deal with these differences in our procedures for estimating survival and breeding probabilities.

For example, if we are using standard capture–recapture/resighting models (Seber 1982, Pollock et al.

[1] Manuscript received 11 June 1993; revised 21 December 1993; accepted 18 January 1994; final version received 28 February 1994.

1990, Lebreton et al. 1992, Nichols 1992), then we might consider stratifying animals caught at time i into breeders and nonbreeders and then conducting separate capture–recapture/resighting analyses on the two groups. However, such stratification appropriate at time i would not necessarily be appropriate for any subsequent time period (e.g., $i + 1$). That is, it would likely be inappropriate to assume that all breeders at i will breed in each successive sampling period of the study and that all nonbreeders at i will likewise remain nonbreeders for the duration of the study. Instead, we would like a methodology that permits capture probability to vary as a function of breeding state and that permits animals to change breeding state from one time to the next.

In this paper we present a methodological framework for addressing questions about breeding proportions and costs of reproduction using capture–recapture/resighting data from marked individuals. The methods are flexible in that they permit the estimation of quantities of interest and the testing of associated hypotheses in the presence of time-specific sampling probabilities that may differ between individuals that do and do not breed.

Study Design and Data

We envision a standard capture–recapture/resighting study of an open animal population. We assume that an animal population is sampled for K sampling periods representing successive potential breeding periods. For example, animals that reproduce annually would be sampled during the breeding season with one sample period per year. For aseasonal or "continuous" breeders, the interval between sampling periods might correspond to the shortest time between two successive breeding attempts (the "recycle" time).

Upon initial capture, animals are given a mark permitting individual identification. Subsequent sampling of marked animals may involve either actual capture or observation but must include both individual identification and an assessment of breeding condition. Thus, a sampled animal's tag must be read and its breeding status must be determined (e.g., via morphological characteristics or behaviors associated with reproduction, or by presence of young animals).

To simplify the analysis, we will not consider age-specific variation in breeding or survival probabilities. Our analysis will exclude animals that have not reached potential reproductive age, but we will also exclude young potential breeders if they could exhibit different probabilities of breeding than adults. Similarly, if senescent decline in breeding or survival is suspected then we will exclude older animals. Our analysis will thus concern animals in prime reproductive age classes. The analysis can be modified to include age specificity in breeding and/or survival probabilities, but the single-age case presented here is most easily explained and should be useful for many data sets.

The data resulting from the study should be summarized in capture history form. A capture history is simply a row of numbers conveying information about an animal's capture status (whether or not it was caught or seen) and, if it was caught/seen, its reproductive state (breeding or nonbreeding), in each sample period of the study. A "0" denotes no capture/observation and indicates that an animal was not seen or caught during a sample period. A "1" indicates that the animal was seen/caught and was found to be in breeding condition. A "2" indicates that the animal was seen/caught and was found not to be in breeding condition. For example, consider the following capture history for a seven-period ($K = 7$) capture–recapture study: 0010120. The animal(s) to which this history pertains was first caught in period 3 and was in breeding condition at that time. The animal was not caught in period 4, but was caught in period 5, again as a breeder. In period 6 the animal was caught as a nonbreeder, and it was not caught in period 7. Every animal caught during the study has an associated capture history, and these data form the basis for all analyses.

Statistical Modeling and Estimation

Estimation

We rely on the multistate modeling approach of Brownie et al. (1993; also see Arnason 1973, Hestbeck et al. 1991, Nichols et al. 1992, 1993, Schwarz et al. 1993). Our state variable is breeding condition, and we define our two states simply as breeding (1) and nonbreeding (2). We will use the capture history data to estimate time- and state-specific capture probabilities and state transition probabilities.

Define the following model parameters:

p_i^s = probability of recapture or resighting at time i, for an animal in state s ($s = 1, 2$) at time i ($i = 1, \ldots, K$).

ϕ_i^{rs} = probability of being alive and in state s at time $i + 1$, for an animal alive and in state r ($r = 1, 2$) at time i ($i = 1, \ldots, K - 1$).

These parameters are used to model the events giving rise to the various possible capture histories. Fig. 1 shows a tree diagram of animal fates and associated probabilities for a two-sample example under the Arnason-Schwarz model (see Brownie et al. 1993).

For computational ease, capture history data are summarized in the following statistics:

n_i^r = the number of animals caught in state r ($r = 1, 2$) at time i ($i = 1, \ldots, K$).

R_i^r = the number of marked animals released in state r at time i (this number may be smaller than n_i^r if animals die on capture or are otherwise removed from the study population),

m_{ij}^{rs} = the number of marked animals recaptured/resighted in state s at time j that were last captured/sighted in state r at time i.

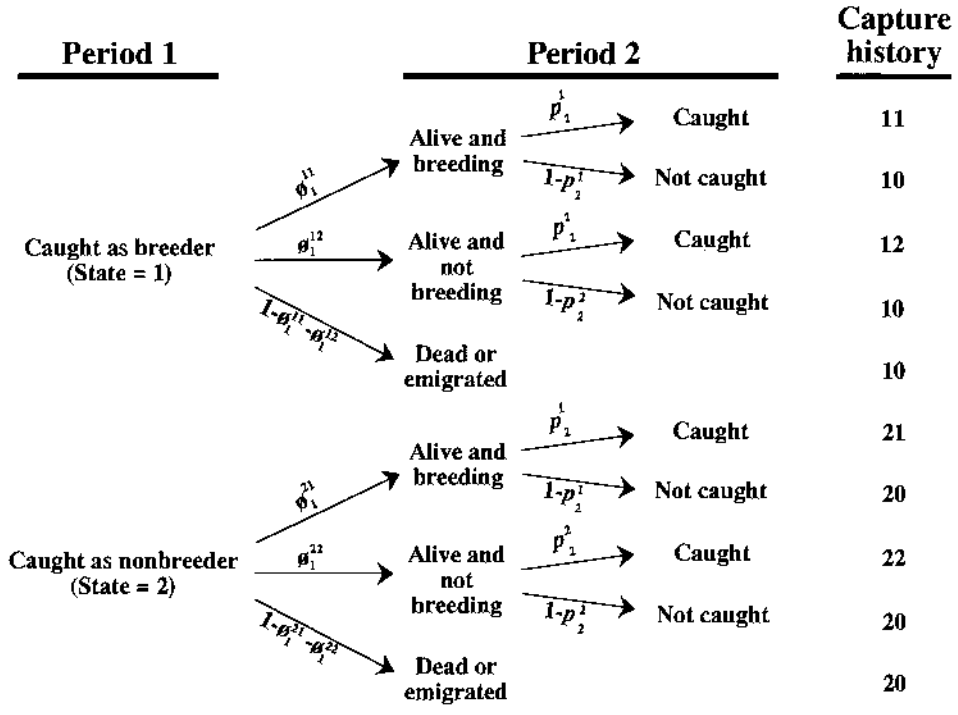

FIG. 1. Tree diagram of events and associated probabilities for animals released as breeders or nonbreeders in period 1 of a two-period study under the described model.

These summary statistics can be used in conjunction with program MSSURVIV (Brownie et al. 1993) to obtain maximum likelihood estimates of the model parameters under the Arnason-Schwarz and related reduced-parameter models.

Estimates of state-specific capture probabilities can be used to estimate the number of breeders (N_i^1) and nonbreeders (N_i^2) in the population at sample period i, as well as the proportion breeding, γ_i:

$$\hat{N}_i^r = n_i^r/\hat{p}_i^r, \quad r = 1, 2 \quad \text{and} \quad i = 2, \ldots, K - 1, \tag{1}$$

$$\hat{\gamma}_i = \hat{N}_i^1/(\hat{N}_i^1 + \hat{N}_i^2), \qquad r = 1, 2 \quad \text{and } i = 2, \ldots, K - 1. \tag{2}$$

Approximate variances of these estimators are:

$$\mathrm{var}(\hat{N}_i^r \mid n_i^r) = \frac{(n_i^r)^2 \mathrm{var}(\hat{p}_i^r)}{(p_i^r)^4}, \tag{3}$$

$$\mathrm{var}(\hat{N}_i^r) = \frac{\mathrm{var}(\hat{p}_i^r)}{(p_i^r)^4}[N_i^r p_i^r(1 - p_i^r + N_i^r p_i^r)] + \frac{N_i^r(1 - p_i^r)}{p_i^r}, \tag{4}$$

$$\mathrm{var}(\hat{\gamma}_i) = \frac{(1 - \gamma_i)^2 \mathrm{var}(\hat{N}_i^1) + \gamma_i^2 \mathrm{var}(\hat{N}_i^2) - 2\gamma_i(1 - \gamma_i)\mathrm{cov}(\hat{N}_i^1, \hat{N}_i^2)}{(N_i^1 + N_i^2)^2}. \tag{5}$$

When using Eqs. 3–5 to compute variance estimates, estimates of N_i^r and γ_i^r are computed using Eqs. 1 and 2. Estimates of p_i^r and $\mathrm{var}(\hat{p}_i^r)$ come directly from the output of program MSSURVIV. Because of the independence of n_i^1 and n_i^2 we can write the cov $(\hat{N}_i^1, \hat{N}_i^2)$ as:

$$\mathrm{cov}(\hat{N}_i^1, \hat{N}_i^2) = N_i^1 N_i^2 \left[\frac{\mathrm{cov}(\hat{p}_i^1, \hat{p}_i^2)}{p_i^1 p_i^2} - \frac{\mathrm{var}(\hat{p}_i^1)\mathrm{var}(\hat{p}_i^2)}{(p_i^1)^2(p_i^2)^2}\right], \tag{6}$$

where $\mathrm{cov}(\hat{p}_i^1, \hat{p}_i^2)$ is obtained from MSSURVIV.

The usual method of estimating breeding proportion from capture–recapture data is to compute the proportion of breeders in the sample:

$$\hat{\gamma}_i = \frac{n_i^1}{n_i^1 + n_i^2}. \tag{7}$$

It is clear that Eq. 7 will be biased if capture probabilities of breeders and nonbreeders are not equal (i.e., if $p_i^1 \neq p_i^2$). However, if capture probability does not depend on reproductive state, then we would prefer Eq. 7 to Eq. 2 because Eq. 7 will have a smaller variance. The conditional (on the total captured sample, $n_i^1 + n_i^2$) variance of $\hat{\gamma}_i$ in Eq. 7 is:

$$\mathrm{var}(\hat{\gamma}_i \mid n_i^1 + n_i^2) = \frac{\gamma_i(1 - \gamma_i)}{n_i^1 + n_i^2}\left[\frac{N_i^1 + N_i^2 - (n_i^1 + n_i^2)}{N_i^1 + N_i^2 - 1}\right], \tag{8}$$

and can be estimated by:

$$\widehat{\text{var}}(\hat{\gamma}_i \mid n_i^1 + n_i^2) = \frac{\hat{\gamma}_i(1 - \hat{\gamma}_i)(1 - \hat{p}_i)}{n_i^1 + n_i^2}. \quad (9)$$

The unconditional variance is:

$$\text{var}(\hat{\gamma}_i) = \frac{\gamma_i(1 - \gamma_i)(1 - p_i)}{p_i(N_i^1 + N_i^2)} \cdot \left[1 + \frac{1 - p_i}{(N_i^1 + N_i^2)p_i}\right], \quad (10)$$

where p_i is written without a superscript, indicating the equal capture probabilities of breeders and nonbreeders.

We can formally test the hypothesis of $p_i^1 = p_i^2$ with likelihood ratio tests between nested models using program MSSURVIV. Using notation similar to that of Lebreton et al. (1992), we can denote the Arnason–Schwarz model with time- and state-specific capture and transition probabilities as model (ϕ_i^{rs}, p_i^r). We can then develop a reduced-parameter model that incorporates our assumption of equal capture probabilities of breeders and nonbreeders, model (ϕ_i^{rs}, p_i). This model is the same as the Arnason–Schwarz model, except that capture probability no longer depends on reproductive state. The likelihood ratio test of the null hypothesis, model (ϕ_i^{rs}, p_i), vs. the alternative hypothesis, model (ϕ_i^{rs}, p_i^r), provides a formal test of the dependence of capture probability on reproductive state. If we have no evidence to reject the null hypothesis, $p_i^1 = p_i^2$, then we can estimate the breeding proportion using Eq. 7.

An alternative approach to selecting an estimator (Eq. 2 or 7) for γ_i is to compute and compare Akaike's Information Criterion (AIC, Akaike 1973, Burnham and Anderson 1992, Lebreton et al. 1992) for the two models, (ϕ_i^{rs}, p_i^r) and (ϕ_i^{rs}, p_i). AIC can be viewed as an objective function for the optimization problem of model selection, reflecting the dual objectives of model selection, adequate description of the data, and use of fewest model parameters. If the AIC is lower for model (ϕ_i^{rs}, p_i) than for model (ϕ_i^{rs}, p_i^r), then we would select Eq. 7 for estimating γ.

This approach of using count statistics (such as n_i^r) of different groups of animals directly in estimators and tests, when capture probabilities of the groups are equal, is advocated and discussed more fully by Skalski and Robson (1992). Of course, it is possible that failure to reject the null hypothesis of equal capture probabilities for breeding and nonbreeding animals could simply be a consequence of low test power, or similarly that low AIC for model (ϕ_i^{rs}, p_i) could simply be a consequence of sparse data. In such cases, use of Eq. 7 could lead to biased estimates of breeding proportion. However, the alternative approach of always using the more general estimator (Eq. 2), even when the more general model (ϕ_i^{rs}, p_i^r) is not supported by the data, is unattractive because it will frequently yield estimates with unnecessarily low precision. Our approach (also see Skalski and Robson 1992) of selecting an estimator based on results of a test of the equality of capture probabilities (or alternatively on a model selection procedure) is based on the Principle of Parsimony, the idea of basing estimation on a model that contains a sufficient number of parameters to avoid bias, yet not so many parameters that precision is lost unnecessarily (Burnham and Anderson 1992, Lebreton et al. 1992). It would be possible to evaluate the performance of our approach using a criterion such as mean squared error (Mood et al. 1974) or residual sum of squares (K.P. Burnham et al., *unpublished manuscript*) that incorporates both sampling variance and bias. K.P. Burnham et al. (*unpublished manuscript*) used computer simulation to study estimator performance for some capture–recapture models and concluded that model selection using AIC performed well and yielded estimates with minimum residual sums of squares.

Testing hypotheses about reproductive costs

The transition probabilities, ϕ_i^{rs}, are relevant to ideas about reproductive costs, and the proposed estimation framework can also be used to develop hypothesis tests about such costs. For example, ϕ_i^{r1} is the probability that an animal in breeding state r at time i survives until time $i + 1$ and breeds at that time. Under a null hypothesis of no reproductive cost, these transition probabilities should be similar for breeders and nonbreeders at time i (H_0: $\phi_i^{11} = \phi_i^{21}$). Under the alternative hypothesis of a reproductive cost, the transition probability associated with breeding at $i + 1$ would be smaller for an animal that was also a breeder at i (H_A: $\phi_i^{11} < \phi_i^{21}$). We can formally test the null hypothesis by building a model with the equality constraint of H_0 ($\phi_i^{11} = \phi_i^{21}$; ϕ_i^{12} and ϕ_i^{22} remain unconstrained). The likelihood ratio test of this null hypothesis model against the more general alternative hypothesis model without this parameter constraint is a formal test of the hypothesis of interest.

Reproduction at time i may extract costs in terms of future survival or future reproduction. The transition probabilities (ϕ_i^{rs}) and tests involving these probabilities include both survival and reproductive components. However, in some circumstances it is possible to decompose the transition probabilities into separate component probabilities and to test component-specific hypotheses. Define the following parameters:

S_i^r = probability that an animal in breeding state r at time i survives until period $i + 1$ ($i = 1, \ldots, K - 1$),

ψ_i^{rs} = probability that an animal in breeding state r at time i, is in breeding state s at time $i + 1$, given that the animal survived until $i + 1$ ($i = 1, \ldots, K - 1$).

If the survival probability, S_i^r, depends only on state at time i, and not on state at time $i + 1$, then we can write:

$$\phi_i^{rs} = S_i^r \psi_i^{rs}. \quad (11)$$

S_i^r and ψ_i^{rs} are identifiable under this parameterization because of the constraint that $\psi_i^{r1} + \psi_i^{r2} = 1$ (that is, an animal must be either a breeder or a nonbreeder at any time i), and the resultant ability to express ψ_i^{r2} as $(1 - \psi_i^{r1})$.

We can use the above parameterization to test more specific hypotheses about whether reproductive costs involve future survival, future reproduction, or both factors. For example, our general model permits different survival probabilities for breeders and nonbreeders (S_i^1, S_i^2). However, we can build a constrained model in which these survival parameters are set equal ($S_i^1 = S_i^2$), corresponding to the null hypothesis of no reproductive cost in terms of survival. Similarly, the general model permits different probabilities of being a breeder at time $i + 1$ for animals that did and did not breed at time i (ψ_i^{11}, ψ_i^{21}). However, a constrained model with $\psi_i^{11} = \psi_i^{21}$ would correspond to the null hypothesis that reproductive state at time i was not associated with different probabilities of breeding at time $i + 1$ (i.e., that there was no cost extracted in terms of future reproduction).

All of the model-based tests described above proceed under the assumption that costs of reproducing at time i are extracted during the subsequent interval (i, $i + 1$). Expressed differently, we are assuming that we are dealing with a first-order Markov process in which the transition probability, ϕ_i^{rs}, depends only on the state at time i. However, we can also develop models and associated tests for extended costs in which transition probabilities depend on reproductive state at times i and $i - 1$. The p_i^s parameters remain the same under this "memory" model, but new transition parameters are required (Hestbeck et al. 1991, Brownie et al. 1993):

$\phi_{i-1,i}^{rst}$ = probability that an animal in state r at time $i - 1$ and state s at time i is in state t at time $i + 1$.

We can formally test the hypothesis of no extended reproductive costs by testing this general memory model against the Arnason–Schwarz model described initially. The Arnason–Schwarz model can be viewed as a constrained model in which $\phi_{i-1,i}^{1st} = \phi_{i-1,i}^{2st}$.

Example

Here we illustrate some of the described methods using capture–recapture data from a livetrapping study of meadow voles, *Microtus pennsylvanicus*, at Patuxent Wildlife Research Center, Laurel, Maryland.

Field methods

The data come from the first 9 mo of an experimental study of metapopulation dynamics. The study design uses four pairs of grids, each grid a 7 × 15 rectangle of trapping stations with adjacent stations within each row or column separated by (7.6 m) (25 ft). Trapping was conducted for five consecutive days on each pair of grids every 8 wk from September 1991 through June 1992.

A single Sherman live-trap containing cotton and baited with rolled oats was placed at each station. Traps were prebaited (baited and left open) 3–5 d before initiating trapping at each sample period. The trapping schedule consisted of setting traps one evening, "running" the traps, and then closing them the following morning, resetting the traps that afternoon, running and closing the traps the following morning, etc., so that animals were captured on five consecutive mornings.

Newly captured animals were marked with individually coded monel fingerling tags placed in their ears. Animals were sexed and weighed on each occasion. For females we recorded nipple size (small, medium, large), whether or not the vagina was perforate, and whether the animal was obviously pregnant. We recorded testes position (scrotal or not) for males.

Reproductive costs are typically hypothesized to be greater for females than for males, so we focus on female meadow voles in this example analysis. In a recent evaluation of external features of small mammals used as indicators of reproductive status, McCravy and Rose (1992) concluded that nipple size was the most useful variable for indicating breeding status of females and estimating breeding proportions. Gestation period for meadow voles is ≈21 d, and duration of lactation is 14 d (Asdell 1964, Eisenberg 1981). Thus, our intersample period exceeded the "recycle time" of the meadow vole, making it likely that an animal classified as a breeder (medium or large nipples) in two successive trapping periods gave birth to two separate litters (i.e., it is unlikely that two successive classifications as a breeder would reflect a single breeding event; thus each classification as a breeder should represent a distinct breeding event). We restricted our analysis to subadult and adult females (≥22 g, see Krebs et al. 1969).

Statistical methods

We analyzed the resulting data with a variety of models developed using program MSSURVIV (Brownie et al. 1993). Our initial models were developed using the full transition probabilities, ϕ_i^{rs}, and capture probabilities, p_i^s. We used results of the initial modeling with the full transition probabilities to suggest appropriate models with the decomposed transition probability parameterization (using survival, S_i^r, and conditional breeding, ψ_i^{rs}, probabilities). The detailed modeling was then conducted using the latter parameterization.

Our model selection strategy followed that suggested by Lebreton et al. (1992) in many respects. We began with a very general model (p_i^s, S_i^r, ψ_i^{rs}), the basic Arnason–Schwarz model with all parameters varying by time and reproductive state, and then used likelihood ratio tests of this model vs. reduced-parameter models in which one parameter at a time was constrained with

respect to either time or reproductive state. If the null hypothesis (the reduced-parameter model) was rejected, then we concluded that the additional parameters of the alternative hypothesis (more general) model were needed to adequately describe the capture–recapture data. If the null hypothesis was not rejected, then we concluded that we had no reason to use the extra parameters of the more general model. In such cases, the reduced-parameter model was then used as the more general model in further likelihood ratio tests against even simpler models with fewer parameters. These likelihood ratio tests are valid only when the more general model fits the data adequately. Model fit was assessed using the goodness-of-fit G statistic computed by program SURVIV (White 1983). We also computed AIC values for each model.

We examined a relatively large number of models (>50) for each capture–recapture data set. In the course of this modeling, we computed a number of likelihood ratio tests that were relevant to each hypothesis of interest. For example, a likelihood ratio test of model (p_i^r, S_i, ψ_i^{rs}) vs. model (p_i^r, S_i^r, ψ_i^{rs}) tests the hypothesis of state-specific survival probabilities. However, a test of model (p_i, S, ψ_i^{rs}) vs. model (p_i, S^r, ψ_i^{rs}) tests the same hypothesis, provided that the alternative hypothesis (model [p_i, S^r, ψ_i^{rs}]) fits the data adequately. Whenever possible, we based our inferences for hypotheses of interest on tests between the model with the lowest AIC and the appropriate "neighboring" model (i.e., the model that was identical to the low-AIC model except for the parameter of interest). This approach should have yielded the most powerful tests. For each test involving parameter variation with reproductive state, we computed an overall or composite test statistic by summing the likelihood ratio χ^2 statistics and their associated degrees of freedom over all four grids.

In cases where the null hypothesis model was not rejected for hypotheses of primary interest (e.g., those concerning reproductive costs), we approximated test power. This was accomplished by conditioning on the observed R_i^r and then computing the expected values of the m_{ij}^{rs} statistics for the alternative hypothesis model. These expectations (we did not round numbers to integers, because we wanted to obtain good approximations) were input as "data" to program MSSURVIV. The resulting likelihood ratio chi-square statistic approximates the noncentrality parameter, λ, of a noncentral chi-square distribution from which power was directly obtained (Burnham et al. 1987:214–217, Lebreton et al. 1992).

Results

The meadow vole populations on the eight study grids were not large when the study began in the fall of 1991. Numbers did increase in the spring of 1992. Initially, we modeled the data from each of the eight grids separately. However, because of the relatively large number of parameters requiring estimation, the multistate models are "data-hungry." Because of data requirements and the relatively small meadow vole populations during most of the trapping periods, we combined data from each pair of grids to form a single data set (e.g., grids 1a and 1b were combined and referred to as grid 1). All of our general models fit the resulting four data sets well, indicating that capture probabilities and demographic characteristics of the voles on the paired grids were similar, as anticipated. The m_{ij}^{rs} arrays for the four grids are presented in the tables of Appendix A and were the data used in all of our modeling.

Over 50 models were fit to each of the four data sets, and roughly half of the models provided reasonable fits as indicated by the goodness-of-fit test statistics. The following models yielded the lowest AIC values and were judged the most appropriate for the respective data sets: grid 1 (p, S, ψ_i^{rs}), grid 2 (p_i, S, ψ_i), grid 3 (p, S, ψ_i), and grid 4 (p_i, S, ψ_i). Goodness-of-fit tests for these models provided no evidence that they were not appropriate descriptions of the respective data sets (all four goodness-of-fit P values > 0.50).

Tests for temporal variation were consistent among the four grids for both sets of demographic parameters, survival probabilities and conditional breeding probabilities. Tests provided no evidence of seasonal variation among the 2-mo survival probabilities on any of the grids, so survival was modeled as a constant. In contrast, there was strong evidence of temporal variation among the conditional breeding probabilities on each grid, requiring the use of time-specific parameters. Parameter estimates were relatively high in November and December 1991, declined to low values in January and February 1992, and then increased to high values in the spring months of 1992 (Table 1), providing a good picture of the seasonality of vole reproductive activity.

In contrast to the situation with demographic parameters, the different grids were not consistent with respect to temporal variation in capture probabilities. Capture probability was adequately modeled as a constant on grids 1 and 3 but required time-specific parameters on grids 2 and 4. This is not surprising, as the four grids were run on different days. Although trapping effort was constant over the entire study, environmental and other factors certainly varied over time, and it is reasonable that such variation produced variation in capture probabilities during some periods on some grids.

We needed to test for possible effects of reproductive state on capture probabilities in order to select the appropriate estimator (Eq. 2 or 7) for breeding proportion. Results of likelihood ratio tests of this hypothesis provided no evidence of different capture probabilities for voles that were and were not reproductively active (Table 2). The composite test statistic over all four grids was nonsignificant ($\chi^2_{10} = 9.5$, $P = 0.45$). We approximated the power of these tests (as-

TABLE 1. Conditional breeding transition probability estimates for female meadow voles at Patuxent Wildlife Research Center, Laurel, Maryland.

Sample period ($i + 1$)	Grid	Date	Conditional breeding transition probabilities*					
			$\hat{\psi}_i$	$\widehat{\text{SE}}(\hat{\psi}_i)$	$\hat{\psi}_i^{11}$	$\widehat{\text{SE}}(\hat{\psi}_i^{11})$	$\hat{\psi}_i^{21}$	$\widehat{\text{SE}}(\hat{\psi}_i^{21})$
2	1	18 Nov 1991	...	...	0.96†	0.038	0.30‡	0.252
	2	2 Dec 1991	1.00§	0.422	...	...	...	...
	3	16 Dec 1991	0.38	0.151	...	...	...	...
	4	30 Dec 1991	0.65	0.048	...	...	...	...
3	1	13 Jan 1992	...	...	0.18	0.083	0.00	0.000
	2	27 Jan 1992	0.31	0.113	...	...	...	...
	3	10 Feb 1992	0.00	0.294	...	...	...	...
	4	24 Feb 1992	0.00	0.605	...	...	...	...
4	1	9 Mar 1992	...	...	0.55	0.233	0.23	0.090
	2	23 Mar 1992	0.53	0.090	...	...	...	...
	3	6 Apr 1992	0.86	0.120	...	...	...	...
	4	20 Apr 1992	0.83	0.066	...	...	...	...
5	1	4 May 1992	...	...	1.00	0.000	0.96	0.034
	2	18 May 1992	0.95	0.049	...	...	...	...
	3	1 Jun 1992	1.00	0.273	...	...	...	...
	4	15 Jun 1992	0.97	0.029	...	...	...	...

* Estimates are based on the following models: grids 2 and 4 (p_i, S, ψ_i), grid 1 (p, S, ψ_i^{rs}), grid 3 (p, S, ψ_i).

† Estimate indicates that 0.96 of the grid 1 females caught as breeders in period 1 and surviving until period 2 were breeders in period 2.

‡ Estimate indicates that 0.30 of the grid 1 females caught as nonbreeders in period 1 and surviving until period 2 were breeders in period 2.

§ Estimate indicates that all of the grid 2 females that were caught in period 1 and survived until period 2 were breeders in period 2.

suming $\alpha = 0.05$) for the alternative hypothesis of $p_i^1 - \hat{p}_i = \hat{p}_i - p_i^2 = 0.10$, 0.15, and 0.20 (hence $\Delta = p_i^1 - p_i^2 = 0.20$, 0.30, and 0.40), where $\hat{p}_i$ is the estimated capture probability based on our data (Table 2). Thus we are investigating the alternative hypothesis of greater capture probabilities for reproductively active females, assuming that pregnancy and lactation are energetically expensive and that reproductive females will show greater attraction to the bait. The power approximations were 0.46 ($\Delta = 0.20$), 0.80 ($\Delta = 0.30$), and 0.96, ($\Delta = 0.40$) for grid 1 and 0.19 ($\Delta = 0.20$), 0.36 ($\Delta = 0.30$), and 0.56 ($\Delta = 0.40$) for grid 3, the two grids with no temporal variation in capture probability. Approximated power was very poor for the two grids requiring time-specific p_i and did not exceed 0.50 even for $\Delta = 0.40$. Approximate powers for the composite statistic were 0.26 ($\Delta = 0.20$), 0.60 ($\Delta = 0.30$), and 0.87 ($\Delta = 0.40$).

The extremely poor power for the two grids (2, 4) with time-specific capture probabilities was unanticipated and led us to consider an alternative testing approach. This approach used capture–recapture data over the 5 d (secondary periods) within each 2-mo primary sampling period as an alternative means of estimating and testing hypotheses about capture probabilities. This approach (see Appendix B) also led to a likelihood ratio test of the null hypothesis that capture probability did not differ for reproductivity active vs. inactive females. These tests provided evidence of generally higher capture probabilities of reproductively active females on grid 2 ($\chi^2_{10} = 26.2$, $P < 0.01$) but not on grid 4 ($\chi^2_{10} = 10.3$, $P = 0.42$). Approximate power for grid 4 was 0.39 for $\Delta = 0.20$, 0.80 for $\Delta = 0.30$, and 0.97 for $\Delta = 0.40$.

Because the multistate capture–recapture models described in this paper use data from primary sampling periods for estimation, we continued to rely on associated goodness-of-fit test statistics and AIC for testing hypotheses about survival and breeding transition probabilities (i.e., we did not elect to use models with

TABLE 2. AIC (Akaike's Information Criterion) values and likelihood ratio tests of the null hypothesis (H_0) that capture probability did not differ by reproductive state vs. the alternative hypothesis (H_a) of state-specific capture probability. Data are from female meadow voles trapped at Patuxent Wildlife Research Center, Laurel, Maryland, 1991–1992.

Grid	Models tested (H_0 vs. H_a)	Test statistic			Model AIC values	
		χ^2	df	P	H_0	H_a
1	(p, S, ψ_i^{rs}) vs. (p^r, S, ψ_i^{rs})	0.4	1	0.53	90.0	91.6
2	(p_i, S^r, ψ_i^{rs}) vs. (p_i^r, S^r, ψ_i^{rs})	2.9	4	0.57	85.3	90.4
3	(p, S, ψ_i) vs. (p^r, S, ψ_i)	1.4	1	0.24	64.4	65.0
4	(p_i, S^r, ψ_i^{rs}) vs. (p_i^r, S^r, ψ_i^{rs})	4.8	4	0.31	94.5	97.7

TABLE 3. Estimated breeding proportions for female meadow voles at Patuxent Wildlife Research Center, Laurel, Maryland.

Sample period (i)	Grid	Date	$p_i^1 = p_i^2$		$p_i^1 \neq p_i^2$	
			$\hat{\gamma}_i$*	$\widehat{\text{SE}}(\hat{\gamma}_i)$	$\hat{\gamma}_i$†	$\widehat{\text{SE}}(\hat{\gamma}_i)$
1	1	23 Sep 1991	0.89	0.025	...	...
	2	14 Oct 1991	0.86‡	...	0.85§	...
	3	28 Oct 1991	0.77	0.050	...	...
	4	11 Nov 1991	0.69‡	...	...	...
2	1	18 Nov 1991	0.86	0.032	...	...
	2	2 Dec 1991	0.66	0.044	0.41	0.205
	3	16 Dec 1991	0.38	0.055	...	...
	4	30 Dec 1991	0.29	0.041	...	...
3	1	13 Jan 1992	0.18	0.032	...	...
	2	27 Jan 1992	0.22	0.047	0.25	0.163
	3	10 Feb 1992	0.00	0.000	...	...
	4	24 Feb 1992	0.06	0.051	...	...
4	1	9 Mar 1992	0.29	0.029	...	...
	2	23 Mar 1992	0.52	0.022	0.46	0.064
	3	6 Apr 1992	0.77	0.043	...	...
	4	20 Apr 1992	0.87	0.012	...	...
5	1	4 May 1992	0.87	0.020	...	...
	2	18 May 1992	0.93	0.046	0.92§	...
	3	1 Jun 1992	0.94	0.020	...	...
	4	15 Jun 1992	0.96	0.016	...	...

* Computed using Eq. 7.
† Computed using Eq. 2.
‡ p_i was not estimable, so $\text{SE}(\hat{\gamma}_i)$ could not be adequately estimated.
§ $\text{var}(\hat{p}_i)$ was not estimable, so $\text{SE}(\hat{\gamma}_i)$ could not be adequately estimated.

state-specific p_i on grid 2, because these models were not parsimonious for the data summarized over primary periods [Table A2]; also most of the evidence of different capture probabilities for breeders and nonbreeders came from a single sample period, period 2). However, we did estimate breeding proportions, γ_i, for grid 2 using both Eq. 2 (assuming state-specific capture probabilities) and Eq. 7 (no state specificity).

The conclusion of no state-specific variation in capture probability for grids 1, 3, and 4 permitted us to estimate time-specific breeding proportions using Eq. 7. We estimated breeding proportions for grid 2 using the combined-state $\hat{p}_i$ from model (p_i, S, ψ_i) in conjunction with Eq. 7, and again using the state-specific $\hat{p}_i^r$ from the Appendix B approach together with Eq. 2. These estimates, $\hat{\gamma}_i$ (Table 3), apply to the entire population of adult and subadult female voles at period i, whereas the $\hat{\psi}_i^{rs}$ (Table 1) apply specifically to the survivors from period $i - 1$. The estimated proportions of Table 3 show the same basic pattern of seasonality as the estimates for survivors (Table 1). Note that when $p_i^1 \neq p_i^2$, resulting estimates of γ_i can differ substantially from $\hat{\gamma}_i$ estimated assuming equal state-specific capture probabilities (e.g., Table 3, grid 2, period 2).

The hypothesis of primary ecological interest involved possible effects of reproductive activity on the demographic parameters, S_i^r and ψ_i^{rs}. Likelihood ratio tests of the null hypothesis of no influence of reproductive state on survival vs. the alternative hypothesis of state-specific variation in survival probability provided no evidence to reject H_0 (Table 4). The composite statistic over all four grids ($\chi_4^2 = 3.8$, $P = 0.95$) provided no evidence that survival was influenced by reproductive state. Point estimates of 2-mo survival probability under H_a likewise showed no indication of a reproductive cost. In fact the largest difference between point estimates was for grid 2 where the survival estimate for reproductive females was larger than that for nonreproductive voles (Table 4). We approximated power of the tests of Table 4 for the situation where $S^2 - \hat{S} = \hat{S} - S^1 = 0.10$ and 0.15 (hence $\Delta = S^2 - S^1 = 0.20$ and 0.30), where $\hat{S}$ is the estimated survival probability under the null hypothesis models. Assuming $\alpha = 0.05$, power approximations were 0.75 ($\Delta = 0.20$) and 0.98 ($\Delta = 0.30$) for grid 1, 0.51 ($\Delta = 0.20$) and 0.86 ($\Delta = 0.30$) for grid 2, 0.37 ($\Delta = 0.20$) and 0.69 ($\Delta = 0.30$) for grid 3, and 0.50 ($\Delta = 0.20$) and 0.85 ($\Delta = 0.30$) for grid 4. Power of the composite statistic over all four grids was approximated as 0.93 for $\Delta = 0.20$.

The tests for an effect of reproductive state at time i on probability of reproducing at time $i + 1$ provided evidence of such an effect on grid 1 ($P < 0.03$), but not on the other three grids (Table 5). However, examination of the time-specific point estimates, $\hat{\psi}_i^{r1}$, for grid 1 showed $\hat{\psi}_i^{r1} > \hat{\psi}_i^{r1}$ for all four time periods (Table 1). The composite test statistic over all four grids was nonsignificant ($\chi_{16}^2 = 16.4$, $P = 0.43$). Thus, there was no evidence of a "cost" of reproduction. Instead, on grid 1 there was some evidence that females that were

TABLE 4. Likelihood ratio tests of the null hypothesis (H_0) that survival probability did not differ by reproductive state vs. the alternative hypothesis (H_a) of state-specific survival. Data are from female meadow voles trapped at Patuxent Wildlife Research Center, Laurel, Maryland, 1991–1992.

Grid	Models tested (H_0 vs. H_a)	Test statistic			2-mo survival estimates					
					H_0		H_a			
		χ^2	df	P	$\hat{S}$	$\widehat{\text{SE}}(\hat{S})$	$\hat{S}^1$	$\widehat{\text{SE}}(\hat{S}^1)$	$\hat{S}^2$	$\widehat{\text{SE}}(\hat{S}^2)$
1	(p, S, ψ_i^{rs}) vs. (p, S^r, ψ_i^{rs})	0.03	1	0.87	0.70	0.038	0.71	0.049	0.70	0.057
2	(p_i, S, ψ_i) vs. (p_i, S^r, ψ_i)	2.8	1	0.10	0.79	0.052	0.87	0.066	0.70	0.077
3	(p, S, ψ_i) vs. (p, S^r, ψ_i)	0.04	1	0.84	0.65	0.066	0.64	0.095	0.66	0.083
4	(p_i, S, ψ_i) vs. (p_i, S^r, ψ_i)	0.9	1	0.33	0.65	0.048	0.73	0.091	0.62	0.054

TABLE 5. AIC (Akaike's Information Criterion) values and likelihood ratio tests of the null hypothesis (H_0) that the probability of breeding in period $i + 1$ did not differ among surviving females that did and did not breed in period i, vs. the alternative hypothesis (H_a) that the probability of a surviving female breeding in period $i + 1$ was dependent on reproductive state at period i. Data are from meadow voles trapped at Patuxent Wildlife Research Center, Laurel, Maryland, 1991–1992.

Grid	Models tested (H_0 vs. H_a)	Test statistic			Model AIC values	
		χ^2	df	P	H_0	H_a
1	(p, S, ψ_i) vs. (p, S, ψ_i^{rs})	10.7	4	0.03	92.8	90.0
2	(p_i, S, ψ_i) vs. (p_i, S, ψ_i^{rs})	4.7	4	0.32	82.4	85.6
3	(p, S, ψ_i) vs. (p, S, ψ_i^{rs})	0.5	4	0.97	64.4	71.8
4	(p_i, S, ψ_i) vs. (p_i, S, ψ_i^{rs})	0.5	4	0.97	86.1	93.5

reproductively active at time i showed a greater probability of reproducing at time $i + 1$ than voles that were inactive at i. For the grids (2–4), where there was no evidence of an effect of reproductive state at i on breeding probability at $i + 1$, we approximated the power of our tests (Table 5). Under the alternative hypothesis, we assumed that $\psi_i^{21} - \hat{\psi}_i = \hat{\psi}_i - \psi_i^{11} = 0.10, 0.15$, and 0.20 (hence $\Delta = \psi_i^{21} - \psi_i^{11} = 0.20, 0.30$, and 0.40), where $\hat{\psi}_i$ was the time-specific estimate under the null hypothesis (Table 5). For $\alpha = 0.05$, power was approximated as 0.25 ($\Delta = 0.20$), 0.50 ($\Delta = 0.30$), and 0.76 ($\Delta = 0.40$) for grid 2, 0.18 ($\Delta = 0.20$), 0.33 ($\Delta = 0.30$), and 0.53 ($\Delta = 0.40$) for grid 3, and 0.27 ($\Delta = 0.20$), 0.47 ($\Delta = 0.30$), and 0.71 ($\Delta = 0.40$) for grid 4. Power of the composite statistic for all four grids was approximated as 0.54 ($\Delta = 0.20$), 0.88 ($\Delta = 0.30$), and 0.99 ($\Delta = 0.40$).

DISCUSSION

On three of the four grids used in our example, there was no evidence that capture probabilities were influenced by reproductive state of female voles. For these animals, we were able to use a simple estimator for breeding proportion, $\hat{\gamma}_i$ (Eq. 7). However, these estimates did not depend on an untested assumption, as is the case in all previous work of which we are aware. Instead, we framed the assumption as a hypothesis and tested it formally. On one grid, the robust design approach to testing provided some evidence of state-specific variation in capture probabilities. In this situation, we used the estimator of γ_i developed specifically for the situation of unequal capture probabilities for reproductively active and inactive animals (Eq. 2). In one sampling period, the difference in capture probabilities of breeding and nonbreeding females produced a substantial difference in the estimated breeding proportion, $\hat{\gamma}_i$.

The power of our tests for state-specific variation in capture probabilities using our multistate models was considered adequate when there was no temporal variation in p, but was very disappointing for the case where capture probabilities varied over time. This lack of power likely resulted from the fact that during most of the sample periods, females were not divided equally among the two states. Instead, one state contained most of the females. Most of the female voles were reproductively active in the late fall and early winter, inactive during midwinter, and active again in the spring. This seasonality resulted in poor sample sizes for one of the reproductive states during most periods. Models with capture probability constant over time have clear advantages in this situation. Power of tests for state specificity will tend to be greatest in cases where good sample sizes exist for animals in both of the possible states.

The low power of tests on two of the grids led to our development of a "robust design" approach to testing the hypothesis of state-specific capture probabilities. This provides yet another illustration of the utility of the robust design approach to capture–recapture studies (Pollock 1982, Pollock et al. 1990). This design permits a variety of alternatives to the traditional estimators and tests based on standard open models (Kendall and Pollock 1992). When the open-model estimators or tests are found to perform poorly in a particular analysis, these robust design alternatives can be used to great benefit. In the modeling described in Appendix B, we developed a Lincoln–Petersen type model to test hypotheses about capture probabilities. Other closed-model approaches were available (Otis et al. 1978, Pollock and Otto 1983, Chao 1987, Rexstad and Burnham 1991, Chao et al. 1992), but, based on the work of Menkins and Anderson (1988), we saw no need to go beyond the Lincoln–Petersen model for our purposes.

The power of tests for reproductive costs in survival (tests for state-specific S) was reasonable, primarily because we did not need to model survival as a time-specific parameter. We can thus be reasonably certain that there were no large (e.g., $\Delta > 0.20$) differences in survival between reproductively active and inactive females during this study. The power of tests for costs in future breeding probability was not good, because it was necessary to model the ψ_i as time-specific parameters. As noted above, numbers of animals in one of the reproductive states were low during most of the sampling periods. Power will be greatest when there are roughly equal numbers of reproductively active and inactive animals caught and when this proportion does

not change substantially from one sampling period to another.

Our estimates of time-specific breeding proportions, $\hat{\gamma}_i$, and conditional breeding probabilities of survivors, $\hat{\psi}_i$, indicated a strong seasonal component of variation, with reduced reproductive activity during winter. Specific hypotheses about sources of variation (e.g., specific weather effects) in these breeding proportions can be tested using program CONTRAST (Hines and Sauer 1989; also see Sauer and Williams 1989). These estimates can also be used in population modeling efforts for meadow voles.

Our formal tests of reproductive cost hypotheses used models that were developed specifically for this purpose and that properly incorporated sampling probabilities. Had we found evidence of reproductive costs, we would have been able to estimate their magnitude as illustrated in Table 4. Had we found evidence of costs, other types of hypotheses could have been tested with our modeling framework. For example, assume that we had found evidence that time-specific breeding probabilities depended on past reproductive state and that model (p, S, ψ_i^s) was an appropriate model for the data. We could have used this as an alternative hypothesis and constructed a null hypothesis model in which $\psi_i^{11} = c\psi_i^{21}$. The null hypothesis in this case is that breeding probability in period $i + 1$ for a female that bred in i is always a constant multiple (c) of the $i + 1$ breeding probability for females that did not breed in i. Thus, the null hypothesis is that the reproductive cost is constant (on a multiplicative scale), whereas the alternative hypothesis is that relative cost varies over time (e.g., perhaps cost is higher during winter, when resources are more scarce).

Because of our lack of evidence of reproductive costs and because of data limitations, we did not test hypotheses about costs that extended for 4 mo or more, rather than 2 mo. Such tests are possible using our modeling framework and can be implemented using program MSSURVIV (Brownie et al. 1993).

Our approach to modeling and estimation requires that all animals included in the analysis at a particular time and in a particular reproductive state have the same capture (p_i^s), survival (S_i^s), and breeding transition (ψ_i^s) probabilities (Brownie et al. 1993). This type of assumption is required by most capture–recapture models for open populations (Seber 1982, Burnham et al. 1987, Pollock et al. 1990, Lebreton et al. 1992). Certainly, such probabilities may differ by age, size, or some other characteristic of the animal. If such relevant characteristics can be identified upon capture/observation of the animal, then our modeling approach can be modified to deal with this additional variation. Age-specific extensions are straightforward and basically follow the same approach used by Pollock (1981) for single-state models.

If covariates such as size or condition are relevant, then we would increase the number of states to incorporate the additional source of variation. For example, assume a simple case where we categorize animals not only as breeders or nonbreeders but also as large or small (e.g., see Nichols et al. 1992). We would deal with this additional variable by defining four states (large breeders, large nonbreeders, small breeders, small nonbreeders) and then proceeding in the manner outlined in this paper. In our *Microtus* example, we dealt with the possible influence of size on probability of breeding by stratification (e.g., we restricted our analysis to voles $\geq$ 22 g) in an effort to provide the most simple illustration possible.

If capture, survival, and breeding transition probabilities are influenced by animal characteristics that cannot be discerned upon capture or observation (i.e., they are not visible to the investigator), then it may or may not be possible to deal with them in our modeling framework, depending on which quantities are influenced. For example, if the unmeasured variable is associated with capture probability, but not with survival or breeding transition probability, then the robust design approach of Nichols et al. (1992) could be adapted to deal with investigations of breeding probabilities and reproductive costs. Other situations will likely be more difficult to deal with, and in such cases the influence of unmeasured variation on model estimates and tests will depend partly on the covariation between the unmeasured variable, reproductive state, and model parameters.

The methods described here should be useful in addressing questions about reproductive costs for a variety of organisms. In fact, many field studies will yield data that are more suitable for these data-analytic methods than our example data. One undesirable aspect of our example was the relatively continuous nature of reproductive activity of meadow voles at our study latitude. This necessitated use of an intersample period that was sufficiently long to insure that an animal designated as a breeder in two successive periods bred twice. At the same time, we hoped that the intersample period was sufficiently short that it would be somewhat unlikely for a female to be designated as a nonbreeder in two successive periods, yet breed between samples. Such an occurrence would not have violated any assumptions of our method, but would have added extra variation to the analysis, reducing our ability to detect reproductive costs. The other undesirable aspect of our example was the inexact nature of our indicator of reproductive status. Although nipple size was the best external indicator available, it was still not perfect (McCravy and Rose 1992), and we undoubtedly designated some breeders as nonbreeders and vice versa. The consequence of this degree of uncertainty is again increased variation and resultant reduction in power of our tests for reproductive costs, as well as possible bias in our estimates of breeding proportion and model parameters.

We are aware of a number of field studies of repro-

ductive and population ecology in which breeding occurs at relatively discrete periods and for which assignment of reproductive status is fairly certain. For example, Gaillard et al. (1992) studied roe deer (*Capreolus capreolus*) using capture–recapture sampling each winter. Determination of pregnancy in females was based on plasma progesterone concentrations in blood samples. Petersen (1992) marked and resighted Emperor Geese (*Chen canagicus*) on the breeding grounds and observed whether returning females nested. Owen et al. (1988) and Black and Owen (1989) marked and resighted Barnacle Geese (*Branta leucopsis*) on the wintering grounds. They determined reproductive performance (during the previous breeding season) by presence of young during fall–early winter observations. Many intensive studies of marked vertebrates will yield data appropriate for use with the described methods (e.g., see Clutton-Brock 1988, Newton 1989), and we recommend these methods for future studies directed at breeding proportions and reproductive costs.

As a final note, we do not wish to overstate the utility of the methods presented here, especially with respect to investigations involving reproductive costs. As noted by Reznick (1985, 1992), studies claiming to investigate reproductive costs can be placed into four methodological categories: phenotypic correlations, experimental manipulations of life history, genetic correlations, and correlated responses to selection. The multistate estimation models permit proper incorporation of variable sampling probabilities into estimation and testing for studies directed at phenotypic correlations. The described *Microtus* study is an example of an investigation directed at phenotypic correlations. The models will also be useful in some field-experimental manipulation studies and may also be applicable to genetic studies conducted in the field, as a means of dealing with possible genotypic variation in capture/observation probabilities. In any event, we certainly do not advocate studies of phenotypic correlations over the other kinds of studies directed at reproductive costs. Reznick (1985, 1992) has argued effectively that genetic studies frequently are more relevant to evolutionary questions, although this perspective is not universally accepted (e.g., Partridge 1992). Instead, our intention is to present the multistate models as a means of properly dealing with variable sampling probabilities in field studies, and we believe that there are many possible applications to studies involving reproductive costs. The most obvious application is to studies of phenotypic correlation, although these methods should be useful in other types of cost studies as well. In fact, multistate models should be useful in any capture–recapture study where a state variable for individual animals may be associated with both demographic parameters (e.g., survival probability) and sampling probabilities (Nichols 1992, Nichols et al. 1992).

ACKNOWLEDGMENTS

We thank Janet M. Hinz, Nigel Noriega, and Heda Zweig for help with fieldwork and Bill Kendall, John Sauer, and two anonymous referees for helpful reviews of the initial manuscript.

LITERATURE CITED

Akaike, H. 1973. Information theory and an extension of the maximum likelihood principle. Pages 267–281 *in* B. N. Petran and F. Csaki, editors. International symposium on information theory. Second edition. Akademiai Kiado, Budapest, Hungary.

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

Asdell, S. A. 1964. Patterns of mammalian reproduction. Cornell University Press, Ithaca, New York, USA.

Black, J. M., and M. Owen. 1989. Agonistic behavior in barnacle goose flocks: assessment, investment and reproductive success. Animal Behavior **37**:199–209.

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.

Burnham, K.P., and D.R. Anderson. 1992. Data-based selection of an appropriate biological model: the key to modern data analysis. Pages 16–30 *in* D. R. McCullough and R. H. Barrett, editors. Wildlife 2001: populations. Elsevier, 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**.

Chao, A. 1987. Estimating the population size for capture-recapture data with unequal catchability. Biometrics **43**: 783–791.

Chao, A., S-M. Lee, and S-L. Jeng. 1992. Estimating population size for capture-recapture data when capture probabilities vary by time and individual animal. Biometrics **48**:201–216.

Clutton-Brock, T. H. 1988. Reproductive success. University of Chicago Press, Chicago, Illinois, USA.

Eisenberg, J. F. 1981. The mammalian radiations. University of Chicago Press, Chicago, Illinois, USA.

Gaillard, J.-M., A. J. Sempere, J.-M. Boutin, G. Van Laere, and B. Boisaubert. 1992. Effects of age and body weight on the proportion of females breeding in a population of roe deer (*Capreolus capreolus*). Canadian Journal of Zoology **70**:1541–1545.

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

Hines, J. E., and J. R. Sauer. 1989. Program CONTRAST—a general program for the analysis of several survival or recovery rate estimates. United States Fish and Wildlife Service, Fish and Wildlife Technical Report **24**.

Johnson, D. H., K. P. Burnham, and J. D. Nichols. 1986. The role of heterogeneity in animal population dynamics. Proceedings of the International Biometrics Conference **13**.

Kendall, W. L., and K. H. Pollock. 1992. The robust design in capture-recapture studies: review and evaluation by Monte Carlo simulation. Pages 31–43 *in* D. R. McCullough and R. H. Barrett, editors. Wildlife 2001: populations. Elsevier, New York, New York, USA.

Krebs, C. J., B. L. Keller, and R. H. Tamarin. 1969. *Microtus* population biology: demographic changes in fluctuating populations of *M. ochrogaster* and *M. pennsylvanicus* in southern Indiana. Ecology **50**:587–607.

Lebreton, J.-D., K. P. Burnham, J. Clobert, and D. R. Anderson. 1992. Modeling survival and testing biological

hypotheses using marked animals: a unified approach with case studies. Ecological Monographs **62**:67–118.

McCravy, K. W., and R. K. Rose. 1992. An analysis of external reproductive features as predictors of reproductive status in small mammals. Journal of Mammalogy **73**:151–159.

Menkins, G. E., Jr., and S. H. Anderson. 1988. Estimation of small-mammal population size. Ecology **69**:1952–1959.

Mood, A. M., F. A. Graybill, and D. C. Boes. 1974. Introduction to the theory of statistics. McGraw-Hill, New York, New York, USA.

Newton, I. 1989. Lifetime reproduction in birds. Academic Press, San Diego, California, USA.

Nichols, J. D. 1992. Capture-recapture models: using marked animals to study population dynamics. BioScience **42**:94–102.

Nichols, J. D., C. Brownie, J. E. Hines, K. H. Pollock, and J. B. Hestbeck. 1993. The estimation of exchanges among populations or subpopulations. Pages 265–279 *in* J.-D. Lebreton and P.M. North, editors. The use of marked animals in the study of bird population dynamics: models, methods, and software. Birkhauser Verlag, Berlin, Germany.

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.

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

Owen, M., J. M. Black, and H. Liber. 1988. Pair bond duration and timing of its formation in barnacle geese (*Branta leucopsis*). Pages 23–57 *in* M. W. Weller, editor. Waterfowl in winter. University of Minnesota Press, Minneapolis, Minnesota, USA.

Partridge, L. 1992. Measuring reproductive costs. Trends in Ecology and Evolution **7**:99.

Petersen, M. R. 1992. Reproductive ecology of emperor geese: annual and individual variation in nesting. Condor **94**:383–397.

Pollock, K. H. 1981. Capture-recapture models allowing for age-dependent survival and capture rates. Biometrics **37**: 521–529.

———. 1982. A capture-recapture design robust to unequal probability of capture. Journal of Wildlife Management **46**: 757–760.

Pollock, K. H., J. D. Nichols, C. Brownie, and J. E. Hines. 1990. Statistical inference for capture-recapture experiments. Wildlife Monographs **107**.

Pollock, K. H., and M. C. Otto. 1983. Robust estimation of population size in closed animal populations from capture-recapture experiments. Biometrics **39**:1035–1049.

Rexstad, E., and K. Burnham. 1991. User's guide for interactive program CAPTURE. Colorado Cooperative Fish and Wildlife Research Unit, Fort Collins, Colorado, USA.

Reznick, D. 1985. Costs of reproduction: an evaluation of the empirical evidence. Oikos **44**:257–267.

———. 1992. Measuring the costs of reproduction. Trends in Ecology and Evolution **7**:42–45.

Sauer, J. R., and B. K. Williams. 1989. Generalized procedures for testing hypotheses about survival and recovery rates. Journal of Wildlife Management **53**:137–142.

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. Estimation of animal abundance and related parameters. Macmillan, New York, New York, USA.

Skalski, J. R., and D. S. Robson. 1992. Techniques for wildlife investigations: design and analysis of capture data. Academic Press, San Diego, California, USA.

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

APPENDIX A

TABLE A1. Capture–recapture data for breeding and nonbreeding female meadow voles at Patuxent Wildlife Research Center summarized in m_{ij}^{g}-array format; grid 1.

Period of release	Number released	Recapture period							
		Period 2		Period 3		Period 4		Period 5	
		1*	2*	1	2	1	2	1	2
1	$R_1^1 = 47$	25	1	1	2	0	1	0	0
	$R_1^2 = 6$	1	2	0	0	0	0	0	0
2	$R_2^1 = 32$	...	...	3	15	1	4	0	0
	$R_2^2 = 5$	...	...	0	3	0	1	0	0
3	$R_3^1 = 8$	...	...	...	...	3	2	0	0
	$R_3^2 = 37$	...	...	...	...	6	17	3	0
4	$R_4^1 = 21$	...	...	...	...	...	...	10	0
	$R_4^2 = 54$	...	...	...	...	...	...	23	1

* Reproductive state 1 = breeding (medium-large nipples). Reproductive state 2 = nonbreeding (small nipples).

TABLE A2. Capture–recapture data for breeding and nonbreeding female meadow voles at Patuxent Wildlife Research Center summarized in m_{ij}^{rg}-array format; grid 2.

Period of release	Number released	Recapture period: Period 2		Period 3		Period 4		Period 5	
		1*	2*	1	2	1	2	1	2
1	$R_1^1 = 11$	8	0	0	2	0	0	0	0
	$R_1^2 = 1$	1	0	0	0	0	0	0	0
2	$R_2^1 = 19$	...	...	2	7	1	4	0	0
	$R_2^2 = 10$	...	...	3	2	0	0	0	0
3	$R_3^1 = 7$	...	...	...	...	4	1	0	0
	$R_3^2 = 28$	...	...	...	...	11	9	1	0
4	$R_4^1 = 27$	...	...	...	...	...	...	12	1
	$R_4^2 = 25$	...	...	...	...	...	...	6	0

* Reproductive state 1 = breeding (medium-large nipples). Reproductive state 2 = nonbreeding (small nipples).

TABLE A3. Capture–recapture data for breeding and nonbreeding female meadow voles at Patuxent Wildlife Research Center summarized in m_{ij}^{rg}-array format; grid 3.

Period of release	Number released	Recapture period: Period 2		Period 3		Period 4		Period 5	
		1*	2*	1	2	1	2	1	2
1	$R_1^1 = 16$	4	5	0	1	0	0	0	0
	$R_1^2 = 5$	1	3	0	0	1	0	0	0
2	$R_2^1 = 9$	...	...	0	2	1	0	0	0
	$R_2^2 = 15$	...	...	0	7	1	0	0	0
3	$R_3^1 = 0$	...	...	...	...	0	0	0	0
	$R_3^2 = 13$	...	...	...	...	3	1	1	0
4	$R_4^1 = 23$	...	...	...	...	...	...	13	0
	$R_4^2 = 7$	...	...	...	...	...	...	5	0

* Reproductive state 1 = breeding (medium-large nipples). Reproductive state 2 = nonbreeding (small nipples).

TABLE A4. Capture–recapture data for breeding and nonbreeding female meadow voles at Patuxent Wildlife Research Center summarized in m_{ij}^{rg}-array format; grid 4.

Period of release	Number released	Recapture period: Period 2		Period 3		Period 4		Period 5	
		1*	2*	1	2	1	2	1	2
1	$R_1^1 = 27$	5	10	0	3	0	1	0	0
	$R_1^2 = 12$	2	5	0	1	0	0	0	0
2	$R_2^1 = 12$	...	...	0	2	1	0	0	0
	$R_2^2 = 29$	...	...	0	5	1	0	0	0
3	$R_3^1 = 1$	...	...	...	...	1	0	0	0
	$R_3^2 = 34$	...	...	...	...	17	2	1	0
4	$R_4^1 = 65$	...	...	...	...	...	...	26	1
	$R_4^2 = 10$	...	...	...	...	...	...	6	0

* Reproductive state 1 = breeding (medium-large nipples). Reproductive state 2 = nonbreeding (small nipples).

APPENDIX B

The power approximations provided evidence of very low power for tests for dependence of capture probability on reproductive state, in cases where capture probability was modeled as a time-specific parameter (e.g., on grids 2 and 4). The robust design (Pollock 1982) nature of the meadow vole sampling program permitted construction of an alternative means of testing this hypothesis. Within each of the primary sampling periods, there were five consecutive days of trapping. We assumed population closure within these 5-d periods and used these data to develop Lincoln–Petersen models (e.g., see Seber 1982) for capture histories within each primary sampling period.

For this modeling, we denote the first 2 d as secondary period 1 and the final 3 d as secondary period 2. We denote the number of animals in reproductive state r with capture history h in primary period i as $X^r_{i,h}$. Because we have only two capture periods, we observe only three capture histories, 10, 01, and 11. If we condition on the animals that are captured, $n^r_i = X^r_{i,01} + X^r_{i,10} + X^r_{i,11}$, then we can write the expected values of the $X^r_{i,h}$ as:

$$E(X^r_{i,10} \mid n^r_i) = n^r_i p'^r_i(1 - p''^r_i)/(p'^r_i + p''^r_i - p'^r_i p''^r_i),$$

$$E(X^r_{i,01} \mid n^r_i) = n^r_i(1 - p'^r_i)p''^r_i/(p'^r_i + p''^r_i - p'^r_i p''^r_i),$$

and

$$E(X^r_{i,11} \mid n^r_i) = n^r_i p'^r_i p''^r_i/(p'^r_i + p''^r_i - p'^r_i p''^r_i),$$

where p'^r_i and p''^r_i denote the capture probabilities of animals in reproductive state r during primary period i at secondary periods 1 and 2, respectively. We denote the above model as model (p'^r_i).

A null hypothesis model of no state-specific variation in capture probability can be constructed by equating the state-specific capture probabilities in the above model ($p'^1_i = p'^2_i$; $p''^1_i = p''^2_i$). We denote this reduced-parameter model as model (p'_i). The likelihood ratio test of model (p'_i) vs. model (p'^r_i) provides a test of the hypothesis of equal secondary period capture probabilities for breeders and nonbreeders.

However, this is not the exact null hypothesis that we want, because our main interest is on the probability that an animal is caught at least once during the entire period i [i.e., our interest is in $p^r_i = 1 - (1 - p'^r_i)(1 - p''^r_i)$]. Thus, we reparameterized the Lincoln–Petersen model, writing p''^r_i in terms of p'^r_i and p^r_i:

$$p''^r_i = 1 - \left(\frac{1 - p^r_i}{1 - p'^r_i}\right).$$

We used a reduced-parameter model ($p^1_i = p^2_i$) as the null hypothesis (denoted model $[p_i]$) and the more general model (p^1_i, p^2_i) as the alternative (denoted model $[p^r_i]$). The likelihood ratio test of model (p_i) vs. model (p^r_i) tested the hypothesis of interest about state specificity of primary-period capture probabilities. We implemented these two models using program SURVIV (White 1983) and conducted the associated test for data from grids 2 and 4.