

D. R. Breininger, J. D. Nichols, B. W. Duncan, E. D. Stolen, G. M. Carter, D. K. Hunt, and J. H. Drese. Multistate modeling of habitat dynamics: factors affecting Florida scrub transition probabilities. *Ecology* 91:3354-3364.

Appendix B. Approaches for developing multistate capture-recapture models to estimate habitat and ecosystem transition probabilities.

This appendix introduces a broader capture-recapture context of multistate models to the estimation of habitat or ecosystem transition probabilities, including transitions to an absorbing state. We expand on how to use this inference framework to estimate average annual transition probabilities when there are not data for all years and when there are misclassification problems. We also show how to test for process stationarity.

Multistate modeling is not new (e.g., Darroch 1961), and data classification problems in multistate modeling are not unique to ecology (e.g., Commenges 1999, Jackson et al. 2003). Multistate modeling has been made more accessible by free software that evolves by incorporating new modeling techniques (e.g., White and Burnham 1999, White et al. 2006) supported by evolving online textbooks with contributions by international scientists (e.g., Cooch and White 2009).

The basic Arnason-Schwarz model (Arnason 1972, Brownie et al. 1993, Schwarz et al. 1993) assumes that state transitions from one sampling occasion to the next represent a first-order Markov process, in that the probability of making a specified transition between occasions t and $t + 1$ depends only on the state at time t , and not on states in previous periods. The basic parameters for this model defined in the capture-recapture context are:

ϕ^{rs}_t = the probability that an organism alive in state r at time t is alive and in state s at time $t + 1$,

p^s_t = the probability that a marked organism alive in state s at time t is recaptured or resighted on that sampling occasion.

Because ϕ^{rs}_t reflects the probability of both surviving and making a specified state transition, it is often of interest to compute the following derived parameters (Hestbeck et al., 1991; Brownie et al., 1993; Schwarz et al., 1993) using:

$S^r_t = \sum \phi^{rs}_t$ = the probability that an organism alive in state r at time t survives (and does not permanently emigrate from the study locations) until time $t + 1$.

$\Psi^{rs}_t = \phi^{rs}_t / S^r_t$ = the probability of being in state s at time $t + 1$, given that the organism was in state r at time t and survived between times t and $t + 1$.

If survival between t and $t + 1$ depends only on state at time t (and not on state at time $t + 1$), then the ψ^{rs}_t can be viewed as conditional (on survival) state transition probabilities and used to provide inferences about the state transition process. In such cases, it is reasonable to parameterize the multistate models in terms of S^r_t and ψ^{rs}_t based on the following equality:

$$\phi^{rs}_t = S^r_t \psi^{rs}_t$$

This parameterization permits direct estimation and modeling of S^r_t and ψ^{rs}_t .

For our habitat dynamics application, we constrained detection probabilities to be 1, reflecting our ability to classify each site for each of the three years of the study:

$$p_{94} = p_{99} = p_{04} = 1.$$

Habitat state dynamics are potentially associated with two parameters, S and ψ . The “survival” parameter, S , actually refers to a site or grid cell. In the case of the 924 sites included in our study, all sites remained in one of the four habitat states for the duration of the study. However, if development (e.g., shopping center construction) essentially removed a site from the set of sites at which natural habitat change was possible; one might be interested in estimating S , the probability that a site did not enter this absorbing state. Because none of our sites was removed from consideration in this manner, we constrained site survival to be:

$$S_t = 1, \text{ for all years, } t.$$

Inferences about habitat dynamics were based primarily on the conditional transition parameters: ψ^{rs}_t . For all sites in each possible state r at time t , the number of sites in each of the 4 states (s) at time $t+1$ followed a multinomial distribution governed by the transition probabilities ψ^{rs}_t .

The estimation approach described in the text was focused on transition probabilities corresponding to 5-year time intervals. For some modeling purposes, it may be more useful to use the same data to estimate “average” annual transition probabilities (e.g., see Breininger et al. 2009). Consider the same data used in the analyses above consisting of site-specific habitat states of each of the 924 10-ha grid cells for each of three years, 1994, 1999, and 2004. If our interest was instead on annual transition probabilities, then the relevant data for each cell could be written as a detection history including all years of the study. For example, consider the following detection history: *Sh* 0 0 0 0 *Op* 0 0 0 0 *Tm*. In 1994 this site was classified at Short (*Sh*), in 1999 as Optimal (*Op*) and in 2004 as Tall mix (*Tm*). The intervening 0’s reflect years at which the habitat state was unknown. We note that detection and habitat classification are perfect during years 1994, 1999, and 2004 of our sample history (i.e., we know true habitat state for these periods) and detection probability is 0 for the intervening periods (i.e., we have no observations about habitat state during these years). In order to specify the observation process (terminology of state space modeling e.g., Royle and Dorazio 2008) within program MARK, we would constrain detection probabilities equal to 1 for the observation years and 0 for the intervening years:

$$p_{1994} = p_{1999} = p_{2004} = 1;$$

$$P_{1995} = P_{1996} = P_{1997} = P_{1998} = P_{2000} = P_{2001} = P_{2002} = P_{2003} = 0.$$

The “survival” parameter, S , still refers the probability that a site present in one sample period will be present in the next sample period, one year later in this case. Because all 924 sites remained in one of the four habitat states, we would again constrain site survival to be 1:

$$S_t = 1, \text{ for all years, } t.$$

Inferences about habitat dynamics are based primarily on the conditional transition parameters, ψ_t^{rs} . Because our data on habitat state of sites were available for only 3 of the 11 years of the study, our inferences about time-specificity of transition probabilities would still be limited. For example, we cannot draw specific inferences about the transition probabilities between years 2002 and 2003 if these differed from those of other years. However, we can draw limited inference about time constancy of transition parameters, as we can test whether “average” transition probabilities differ between two periods, 1994-1999 and 1999-2004. Under the general time model, we would constrain the transition probabilities as follows:

$$\begin{aligned} \psi_{1994}^{rs} &= \psi_{1995}^{rs} = \psi_{1996}^{rs} = \psi_{1997}^{rs} = \psi_{1998}^{rs} = \psi_1^{rs}; \\ \psi_{1999}^{rs} &= \psi_{2000}^{rs} = \psi_{2001}^{rs} = \psi_{2002}^{rs} = \psi_{2003}^{rs} = \psi_2^{rs}; \end{aligned}$$

for all habitat states r and s . The above parameterization, when combined with the specified constraints on p and S , provides estimates of “average” transition probabilities for the two time periods for which such inference is possible. In the matrix notation of equations 1 and 2, the estimates of transition probability for the periods 1994-1999, $\hat{\psi}_1^{rs}$ and 1999-2004, $\hat{\psi}_2^{rs}$ represent averages in the sense that they should satisfy:

$$\begin{aligned} \mathbf{n}_{1999} &= (\Psi_1)^5 \mathbf{n}_{1994}, \\ \mathbf{n}_{2004} &= (\Psi_2)^5 \mathbf{n}_{1999}. \end{aligned}$$

The hypothesis of time-specificity for these two periods, 1994-1999 and 1999-2004, can be tested by comparing this model with two sets of annual transition probabilities against a model in which annual transitions were constrained to be equal for the 2 time periods, $\psi_1^{rs} = \psi_2^{rs}$. This comparison should yield exactly the same results as that described in the main text for 5-year intervals. The model described here is simply a reparameterization of the model described in the text. The advantage of the reparameterization is simply that program MARK will provide direct estimates of annual transition probabilities and their variances rather than estimates corresponding to 5-year intervals. Such an approach is also useful if time periods between observations differ, and if covariates are measured at different temporal frequencies.

We estimated the probability of remaining in the same state by subtraction, which is the default option in Program MARK. This was reasonable because we were most interested in factors that produced change in landscapes. There are, however, situations when it is sensible to estimate another transition parameter by subtraction which is easily performed in MARK (Cooch and White 2009).

Our goals were to draw inferences about habitat dynamics of all managed areas within a large conservation area; hence our models were developed for all potential Florida Scrub-Jay territories within the area of interest. We identified those potential covariates most responsible for spatial variation in habitat state and state dynamics.

However, spatial autocorrelation is of interest in many studies and could be investigated using neighborhood effects and autologistic models (e.g., Royle and Dorazio 2008).

Another potential problem faced in many studies of habitat using remotely sensed data involves misclassification of states based on the misinterpretation of image data. In this study, we believe that misclassification was uncommon. If any misclassification did occur, it most likely involved distinguishing short and optimal states where there was no strong threshold between the amount of medium-height scrub needed for adequate Scrub-Jay recruitment and survival (Breininger and Oddy 2004). Although we believe that misclassification was not an important issue in our study, we recognize that it is a common problem in many remote sensing studies. The usual approach to misclassification is to hope that errors are offsetting (Scanlan and Archer 1991).

Multistate models provide a useful framework for dealing with misclassification. There are a number of approaches to directly incorporate uncertainty in state determination using multistate models that do not require the elimination of data of questionable classification (e.g., Fujiwara and Caswell 2002; Kendall et al. 2003; Nichols et al. 2004, 2007; Royle 2004, Royle and Link 2005, Runge et al. 2007, Choquet et al. 2009, Conn and Cooch 2009, MacKenzie et al. 2009). A key element of most of these approaches is to have some subset of data for which truth is known, either because one state permits unambiguous state determination or because there are some sample units in which a second survey is conducted, permitting unambiguous state determination (“ground truthing”) for all states within the subset. Thus, direct approaches for incorporating classification uncertainty into multistate analyses exist, and the specific approach selected will depend on the sampling situation.

Many have noted the importance of developing innovative techniques to study complex interactions in spatial and temporal variability that can separate noise and causes of variability (Levin 1992, Fuhlendorf et al. 2006). Approaches to separate sampling variation from true process variation were developed for capture-recapture models some time ago (e.g., Burnham et al. 1987, Link and Nichols 1994, Gould and Nichols 1998), and multiple approaches based on random effects modeling are now available for capture-recapture models in general, and multistate models in particular (Cooch and White 2009, Lebreton et al. 2009).

Model support can be assessed using model weights (w_i), which provide relative measures of model support given the data (Burnham and Anderson 2002, Anderson 2008). The hypothesis of time-specificity for two periods (e.g., 1994-1999 and 1999-2004), versus no differences between periods (stationarity) could be tested by comparing a model with two sets of time-specific transition probabilities against a model in which transitions were constrained to be equal for the 2 time periods, $\psi_{94}^{rs} = \psi_{99}^{rs}$, using evidence ratios (w_i / w_j) that give the relative likelihood of model i to model j . In our study there was no support for a model without fire and cutting effects on transition probabilities and no support for a model that excluded time (Appendix A), indicating that dynamics were likely not stationary over time.

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.

- Breininger, D.R., and D. M. Oddy. 2004. Do habitat potential, population density, and fires influence Florida Scrub-Jay source-sink dynamics? *Ecological Applications* 14: 1079-1089.
- Breininger, D. R., J. D. Nichols, G. M. Carter, and D. M. Oddy. 2009. Habitat-specific breeder survival of Florida Scrub-Jays: inferences using multistate models. *Ecology* 90:3180-3189.
- 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. 2002. *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*. 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 Number 5*.
- Commenges, D. 1999. Multi-state models in epidemiology. *Lifetime Data Analysis* 5:315-327.
- Choquet, R., Rouan, L., Pradel, R. 2009. Program E-SURGE: a software application for fitting Multievent models. *Environmental and Ecological Statistics* 3:845-865.
- Conn, P. B., and E. G. Cooch. 2009. Multistate capture-recapture analysis under imperfect state observation an application to disease models. *Journal of Applied Ecology* 46:486-492.
- Cooch, E., and G. White. 2009. Program MARK: a gentle introduction. <http://www.phidot.org/software/mark> (accessed November 2009).
- Darroch, J. N. 1961. The two-sample capture-recapture census when tagging and sampling are stratified. *Biometrika* 48:241-260.
- Fuhlendorf, S. D. W. C. Harrel, D. M. Engle, R. G. Hamilton, C. A. Davis, and D. M. Leslie, Jr. 2006. Should heterogeneity be the basis for conservation? Grassland birds response to fire and grazing. *Ecological Applications* 16:1706-1716.
- Gould, W. R., and J. D. Nichols. 1998. Estimation of temporal variability of survival in animal populations. *Ecology* 79:2531-2538.
- 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.
- Kendall, W. L., J. E. Hines, and J. D. Nichols. 2003. Adjusting multistate capture-recapture models for misclassification bias manatee breeding proportions. *Ecology* 84:1058-1066.
- Jackson, C. H., L. D. Sharples, S. G. Thompson, S. W. Duffy, and E. Couto. 2003. Multistate Markov models for disease progression with classification error. *The Statistician* 52:193-209.
- Lebreton, J. D., J. D. Nichols, R. Barker, R. Pradel, and J. Spendelov. 2009. Modeling individual animal histories with multistate capture-recapture models. *Advances in Ecological Research* 41:87-173.
- Levin, S. A. 1992. The problem of pattern and scale in Ecology: The Robert H. MacArthur Lecture. *Ecology* 73:1943-1967.
- Link, W. A., and J. D. Nichols. 1994. On the importance of sampling variance to investigations of temporal variation in animal population size. *Oikos* 67:165-171.

- Mackenzie, D. I., J. D. Nichols, J. E. Hines, M. G. Knutson, and A. B. Franklin. 2003. Estimating site occupancy, colonization, and local extinction when a species is detected imperfectly. *Ecology* 84:2200-2207.
- Nichols, J. D., W. L. Kendall, J. E. Hines, and J. A. Spendelov. 2004. Estimation of sex-specific survival from capture-recapture data when sex is not always known. *Ecology* 85:3192-3201.
- Royle, J. A., and R. M. Dorazio. 2008. *Hierarchical Modeling and Inference in ecology: the analysis of data from populations, metapopulations, and communities*. Academic Press, San Diego, California, USA.
- Runge, J. P., J. E. Hines, and J. D. Nichols. 2007. Estimating species-specific survival and movement when species identification is uncertain. *Ecology* 80:282-288.
- Scanlan, J.C., and S. Archer. 1991. Simulated dynamics of succession in a North American subtropical *Prosopis* savanna. *Journal of Vegetation Science* 2:625-634.
- Schwarz, C. J., J. Schweigert, and A. N. Arnason. 1993. Using tag-recovery data to estimate migration rates. *Biometrics* 49:177-194.
- Tanner, J. E., T. P. Hughes, and J. H. Connell. 1994. The role of history in community dynamics: a modeling approach. *Ecology* 77:108-117.
- White, G. C., and K. P. Burnham. 1999. Program MARK for survival estimation. *Bird Study* 46:S120-S139.
- White, G. C., W. L. Kendall, and R. J. Baker. 2006. Multistate survival models and their extensions in Program MARK. *Journal of Wildlife Management* 70:1521-1529.
- Wootton, J. T. 2001. Causes of species diversity differences: a comparative analysis of Markov models. *Ecology Letters* 4:46-56.