Russell Alpizar-Jara · James D. Nichols ·
James E. Hines · John R. Sauer · Kenneth H. Pollock ·
Christopher S. Rosenberry

The relationship between species detection probability and local extinction probability

Abstract In community-level ecological studies, generally not all species present in sampled areas are detected. Many authors have proposed the use of estimation methods that allow detection probabilities that are <1 and that are heterogeneous among species. These methods can also be used to estimate community-dynamic parameters such as species local extinction probability and turnover rates (Nichols et al. Ecol Appl 8:1213–1225; Conserv Biol 12:1390–1398). Here, we present an ad hoc approach to estimating community-level vital rates in the presence of joint heterogeneity of detection probabilities and vital rates. The method consists of partitioning the number of species into two groups using the detection frequencies and then estimating vital rates (e.g., local extinction probabilities) for each group. Estimators from each group are combined in a weighted estimator of vital rates that accounts for the effect of heterogeneity. Using data from the North American Breeding Bird Survey, we computed such estimates and tested the hypothesis that detection probabilities and local extinction probabilities were negatively related. Our analyses support the hypothesis that species detection probability covaries negatively with local probability of extinction and turnover rates. A simulation study was conducted to assess the performance of vital parameter estimators as well as other estimators relevant to questions about heterogeneity, such as coefficient of variation of detection probabilities and proportion of species in each group. Both the weighted estimator suggested in this paper and the original unweighted estimator for local extinction probability performed fairly well and provided no basis for preferring one to the other.

Keywords Animal community ecology · Capture–recapture models · Community dynamics · North American Breeding Bird Survey · Species turnover

R. Alpizar-Jara (✉)
Departamento de Matemática, Universidade de Évora,
Rua Romão Ramalho 59,
7000-671 Évora, Portugal
e-mail: alpizar@uevora.pt

J. D. Nichols · J. E. Hines · J. R. Sauer
US Geological Survey, Patuxent Wildlife Research Center,
12100 Beech Forest Road,
Laurel, MD, 20708-4030, USA

K. H. Pollock
Zoology Department, North Carolina State University,
Box 7617 Raleigh, NC, 27695-7617, USA

C. S. Rosenberry
Fisheries and Wildlife Program, Department of Zoology, North
Carolina State University,
Raleigh, NC, 27695, USA

Introduction

Studies of animal communities must deal with species detection probabilities that are <1. Whether such studies are based on direct counts or captures of animals, it is likely that not all species in a sampled community are detected. Some ecologists have used estimators based on capture–recapture modeling to estimate species richness (Burnham and Overton 1979; Derleth et al. 1989; Karr et al. 1990; Dawson et al. 1995; Thiollay 1995; Bennetts et al. 1996; Boulinier et al. 1998a, b).

The capture–recapture approach to estimating species richness has been extended to estimation of other community-dynamic parameters such as local rates of extinction, turnover, and spatial overlap (Boulinier et al. 1998b; Nichols et al. 1998a, b). Many of these community-dynamic estimators are based on the robust design of Pollock (1982). These estimators may be written as functions of the more frequently used species richness estimators developed for the case of heterogeneous detection probabilities among species (Burnham and Overton 1979; Smith and van Belle 1984; Chao et al. 1992; Bunge and Fitzpatrick 1993; Lee and Chao 1994; Norris and Pollock 1996). Thus, heterogeneity of detection probability presents no problems with these estimators of parameters. However, Nichols et al. (1998a) speculated that frequently there will be a negative sampling covariance between species detection probability and many of

these parameters, and that this covariation could produce bias.

The objectives of this work are to:

1. Present the ecological reasoning underlying the prediction that species detection probabilities and parameters such as local extinction probability should covary negatively.
2. Test this prediction using data from the North American Breeding Bird Survey (BBS).
3. Suggest a class of ad hoc estimators that may provide a reasonable alternative to the standard robust design estimators of Nichols et al. (1998a, b) in the situation where the heterogeneity among species includes negative covariation between detection probability and the parameter of interest.
4. Evaluate the performance of these estimators, as well as selected others relevant to questions about heterogeneity, via simulation.

In the Discussion, we note that existence of covariation between species detection probability and community-level parameters has important implications for other areas of investigation in community ecology such as nested subset patterns of species composition (e.g., Patterson 1984) and the relationship between abundance and extinction probability (Gaston 1994; McKinney 1997).

Background

Community sampling

Several approaches can be used to sample a community. A very common method is based on some sort of quadrat sampling where an area of interest is subdivided into quadrats or other sampling units. Each species detected in the survey will have been detected on some quadrats and not on others, and these data are encoded as a row of 1's and 0's indicating the exact quadrats on which each species is detected (e.g., 01101 indicates a species detected on quadrats 2, 3, and 5, but not on 1 and 4). The detection history data are then used for estimation of species richness (Bunge and Fitzpatrick 1993; Nichols and Conroy 1996; Boulinier et al. 1998a). If such data are obtained from the same area for multiple seasons or years, then it is possible to estimate parameters associated with community dynamics, such as local species turnover and extinction probability (Boulinier et al. 1998b; Nichols et al. 1998a).

Species detection probabilities

In the standard population-dynamic use of capture–recapture models, the capture probability (p_{ik}) denotes the conditional probability that animal k is caught during sample period i, given that the animal is alive and present in the sampled area at period i. Variation among individuals in this probability results from variation in

such factors as movement patterns relative to traps or nets and wariness with respect to capture devices. The community-dynamic use of capture–recapture models requires a different view, as we define detection probability for a particular species (p_{ij}; where j denotes species and i denotes sample period), as the conditional probability that at least one individual of the species is detected at a sampling location, given the presence of at least one individual of the species in the area sampled at period i. If p_{ijk} denotes the probability of detecting individual k of species j at sampling period i, then we can write the species detection probability as:

$$p_{ij} = 1 - \prod_{k=1}^{n_j} \left(1 - p_{ijk}\right), \tag{1}$$

where n_j is the number of individuals of species j exposed to sampling efforts. If all individuals of the species have the same detection probability (denoted as $p_{ij\cdot}$), then we can write the species detection probability as:

$$p_{ij} = 1 - \left(1 - p_{ij\cdot}\right)^{n_j}. \tag{2}$$

Thus, variation in species-specific detection probabilities arises from variation in the number of individuals in each species as well as from the factors that cause individuals of different species to differ in detection probability (e.g., movement patterns, size, secretiveness, vocalizations, wariness, etc.). For the sampling techniques likely to be used in most community surveys, we expect variation in species detection probabilities to be very large (see Sauer et al. 1994) and speculate that such variation will typically exceed that variation among individuals within a species.

Expected relationship between detection probability and community vital rates

Some community-level parameters are likely a function of the number of individuals in the species in the sampled time and location. For example, Nichols et al. (1998a) noted that extinction probability is negatively related to population size, with small populations experiencing relatively high extinction probabilities as a result of demographic and/or environmental stochasticity (Bailey 1964; MacArthur and Wilson 1967; Goel and Richter-Dyn 1974; Gilpin and Soulé 1986; Boyce 1992; Burgman et al. 1993). Turnover is a function of both extinction and colonization probabilities (Nichols et al. 1998a) and should also vary inversely with population size and, therefore, species detection probability. Regarding spatial dynamics, the probability that a species in one location is also present at another location is likely to be a function of local and regional abundance, such that an abundant species in one location is expected to have a higher probability than a rare species of being found at a nearby location (Hanski et al. 1993; Gaston 1994). The number of

individuals in a species is thus a determinant of both species detection probability and certain community-dynamic parameters. Thus, we expect a relationship between these parameters and species detection probability.

The estimation problem

We focus our discussion on an estimator of local probability of extinction but would like to emphasize that the analyses can easily be extended to other community-dynamic parameters such as turnover rates (Nichols et al. 1998a). The approach assumes that data are available for two sampling periods (for instance, an initial year i and a later year j).

Nichols et al. (1998a) proposed the following estimator for local extinction probability:

$$1 - \hat{\varphi}_{ij} = 1 - \frac{\hat{M}_j}{R_i}, \tag{3}$$

where $\hat{\varphi}_{ij}$ the estimated probability that a species present in the sampled area at time i is present in the sampled area at time j ($j > i$), R_i is the number of species actually encountered in the sampling at time i, and $\hat{M}_j$ is the estimated number of the R_i species that are also present at time j. The $\hat{M}_j$ estimate is obtained using a richness estimator with species occurrence data (the detection histories) at time j of those species that were also members of R_i (denote this number of species that were members of R_i and detected at time j as m_j). This estimator of species richness can explicitly deal with heterogeneous detection probabilities (e.g., by using the jackknife estimator of Burnham and Overton 1978, 1979). The potential for bias in community parameters such as local extinction probability occurs because of the conditioning on the R_i species detected at time i and the possibility that these species will likely not include an equal proportion of species with high and low detection probabilities. As local extinction probability may be a function of the number of individuals, the R_i species detected may have a higher probability of being present at time j than the entire group of species present at time i. Thus, although we may be able to obtain an unbiased estimate of the local extinction probability of the R_i species detected at time i, this estimate may be positively biased for all species present at time i, as it is based on a subset of species with relatively low extinction probabilities.

Nichols et al. (1998b) have proposed the use of an estimator similar to that of Eq. 3 to estimate spatial turnover and similarity. In this case, i and j present two different areas sampled perhaps in the same year. The parameter ϕ_{ij} now represents the probability that a species present in community i is also present in community j, so its complement $(1-\phi_{ij})$ reflects spatial turnover or dissimilarities. If the species detected from community i tend to be the more locally abundant species (conferring higher detection probabilities), then if community j is in the same general area or region as community i, we might expect to find members of R_i in community j with higher average probability than for the entire community of species at i, because of a possible relationship between abundance and range (Hanski et al. 1993; Gaston 1994). The possibility of R_i being non-representative with respect to probability of occurrence in community j represents the same kind of problem noted with respect to the estimator (Eq. 3).

Materials and methods

Analyses of avian data

Data from the North American BBS (see Robbins et al. 1986 for details on the sampling scheme) have been analyzed to test our hypotheses about relationships between detection probabilities and community-dynamic parameters. The BBS consists of >4,000 roadside routes 39.4 km in length, surveyed once each year during the peak of the breeding season. A competent observer conducts fifty 3-min point counts at 0.8-km intervals along the survey route, recording all birds seen and heard at each stop. For the present analysis, we focused on the years 1976 and 1996 for three selected states representing differences in habitat and bird species composition: North Dakota, Maryland and Wisconsin. Previous work with BBS data from these states indicated the utility of a specific capture–recapture model that accounts for heterogeneous detection probabilities (known as the heterogeneity model) for estimating species richness and related quantities (Boulinier et al. 1998a).

Estimation methods

We were interested in developing estimators for local extinction probabilities (and other parameters) that account, at least to some degree, for heterogeneous detection probabilities and likely heterogeneous extinction probabilities. The proposed ad hoc approach categorizes species detected on a BBS route in 1 year (i.e. 1976) into two classes based on differences in species-specific detection frequencies (f_i = number of species detected on exactly i stops on the BBS route, $i=1, \cdots, 50$), as these frequencies will be closely related to underlying detection probabilities. Specifically, we designated the species detected on only one or two stops ($f_1 + f_2$) as belonging to the "low-p" group, and the species detected on more than two stops ($f_3 + \cdots + f_{50}$) as members of the "high-p" group. We expected this partitioning of species would result in two groups of species with different average detection probabilities. Species in the low-p group were predicted to have lower detection probabilities than species in the high-p group.

Equation 3 was used to estimate local survival probability (complement of local extinction probability) for the interval 1976–1996 for species on each BBS route designated as high-p $(\hat{\varphi}_{ij}^{\mathrm{h}})$ and low-p $(\hat{\varphi}_{ij}^{\mathrm{l}})$ based on detection frequencies in 1976 (year i). Using the group-specific estimates of extinction probabilities, a weighted estimator for the complement of local extinction probability is given by

$$\hat{\varphi}_{ij}^{\mathrm{w}} = \hat{\theta}\hat{\varphi}_{ij}^{\mathrm{l}} + (1 - \hat{\theta})\hat{\varphi}_{ij}^{\mathrm{h}}, \tag{4}$$

where $\hat{\theta} = \left(\hat{f}_0 + f_1 + f_2\right)/\hat{N},$ is the estimated proportion of species in the low-p group and $\hat{N}$ is an estimator of the total number of species in 1976. Note that the numerator of $\hat{\theta}$ includes the estimated number of species not detected (seen on zero stops) on the route, $\hat{f}_0 = \hat{N} - \sum_{k=1}^{t=50} f_k$. For comparison purposes, we computed extinction probabilities based on the weighted (Eq. 4) and the unweighted (Eq. 3) estimators.

The use of two groups of species with different detection probabilities is arbitrary, as we could have subdivided observed species into a greater number of groups and estimated local extinction for each group. We had several reasons for using only two groups. First, we wanted to retain an adequate number of species in each group to still obtain reasonable estimates of local extinction probability. Second, Carothers (1973) provided a justification for using a two-group distribution to investigate the effects of heterogeneous capture probabilities on capture–recapture estimators, noting that the first two moments of the detection probability distribution were sufficient to explain most of the heterogeneity and the effects of skewness and kurtosis were almost negligible. Carothers (1973, 1979) then successfully used a two-group distribution in investigation of bias in open-model capture–recapture estimators. Finally, during the course of this investigations we became aware of the work of Pledger (2000), who studied use of mixture models with fixed numbers of groups for estimation under closed population capture–recapture models and concluded that "a simple dichotomy of the animals is enough to substantially correct for heterogeneity-induced bias in the estimation of population size".

We thus used two groups of species defined by the different detection frequencies in 1976, and estimated the average route-level detection probabilities for each group based on jackknife estimates of species richness (Burnham and Overton 1979; Boulinier et al. 1998a). The estimator of route-level detection probability is given by

$$\hat{p}_{j,\mathrm{r}}^{g} = \frac{m_j^g}{\hat{M}_j^g}, \quad g = l, \ h,$$

where m_j^g is the number of species assigned to group g in year i that were also detected in year j, and $\hat{M}_j^g$ is the number of species assigned to group g in year i that were estimated to be present in j.

As a measure of the degree of heterogeneity in species detection probabilities, we used the estimator for the coefficient of variation of capture probabilities under the heterogeneity model (M_h) proposed by Lee and Chao (1994). This estimator is given by

$$\hat{\kappa} = \sqrt{\max\left(\frac{t \sum_{k=1}^{t} f_k \sum_{k=1}^{t} k(k-1)f_k}{(t-1)\sum_{k=2}^{t} kf_k \sum_{k=1}^{t} kf_k} - 1, 0\right)}, \tag{5}$$

where all the summations go up to t=50 stops. This is an estimator of the coefficient of variation for stop-level rather than route-level species detection probabilities.

Predictions and associated tests

As discussed and justified in the Background section, the following predictions were tested based on 1976 and 1996 data for the three states: local probability of extinction for the high-p groups should be smaller than for the low-p groups $(\hat{\varphi}_{ij}^{\mathrm{l}} - \hat{\varphi}_{ij}^{\mathrm{h}} < 0)$, the proposed weighted estimates for overall survival (based on estimates from the two groups) will be smaller than the unweighted estimates $(\hat{\varphi}_{ij}^{\mathrm{w}} - \hat{\varphi}_{ij} < 0)$, and the precision of the weighted estimator may be poorer as it is based on more parameters to be estimated. We used Wilcoxon signed-rank tests for paired samples based on the estimated differences of the estimators for the corresponding parameters.

We also expected that species characterized by low detection probabilities in 1976 will likely have low detection probabilities in 1996. This prediction was tested using route-level detection probabilities for each group estimated from the same data used to test for differences in local extinction $(\hat{p}_{j,\mathrm{r}}^{\mathrm{h}} - \hat{p}_{j,\mathrm{r}}^{\mathrm{l}} > 0)$. We view this test as a check on the efficacy of our approach to classifying species by detection probability using detection frequencies.

Computer simulation analyses

We used computer simulation to evaluate estimator performance. Our primary interest was in the performance of the proposed original unweighted (Eq. 3) and weighted (Eq. 4) community parameter estimators. We also evaluated the performance of other estimators relevant to the issue of heterogeneous detection probabilities such as our ad hoc "estimator" for the proportion of species in the low-p group, $\hat{\theta}$, and the coefficient of variation of

detection probability, $\hat{\kappa}$, given by (Eq. 5) and taken from Eqs. 3.2, 3.20 and 3.23 of Lee and Chao (1994).

All simulations were based on two-group distributions for the reasons discussed above (Carothers 1973, 1979; Pledger 2000). Species detections were simulated for two points in time, and the resulting data were used to estimate the quantities noted above. The number of species in the community at each time period was set at 100. Solving a system of equations that accounts for estimates of average route-level detection probabilities and stop-level coefficient of variation, we were able to define plausible parameter sets (based on BBS data) for our simulations. If species detection probabilities are constant for all stops, then route- and stop-level detection probabilities are related by: $p_r = 1 - (1-p)^t$, where p_r and p denote route- and stop-level detection probabilities, respectively, and t is the number of stops on the route. Fourteen different scenarios were chosen for our investigations (Table 1).

The proportion of species in the low-p group was permitted to vary from 0.3 to 0.9, and the average species detection probability at the route level was varied from 0.6 to 0.8 (Table 1) in order to reflect our experience with BBS data sets. These average route-level detection probabilities produced average stop-level detection probabilities ranging from about 0.02 to 0.07 (Table 1), and we investigated coefficients of variation for the stop-level detection probabilities (κ) of 0.5, 1.0, and 1.5. For these two-group distributions, the coefficient of variation is given by

$$\kappa = \frac{\sqrt{\theta(1-\theta)}|p^h - p^l|}{\theta p^l + (1-\theta)p^h}, \tag{6}$$

where p^l and p^h are the stop-level detection probabilities

for the low-p and high-p groups, respectively (see Carothers 1973).

The complement of species extinction probability was set equal to 0.70 for the low-p group and 0.95 for the high-p group. The different proportions of species in these two groups (θ) led to average local survival probabilities ranging between 0.725 and 0.875 (Table 1). Simulation procedures generally followed the bootstrap methodology presented in Nichols et al. (1998a), and all results are based on 200 simulations. We used software developed by Hines et al. (1999), based on the jackknife estimator of Burnham and Overton (1978, 1979) for the heterogeneity model.

Each estimator was evaluated by computing percent relative bias (% bias) and root mean squared error (RMSE). Let γ denote the parameter of interest and $\hat{\bar{\gamma}}$ the mean of the associated estimates over all 200 simulations (i.e., $\hat{\bar{\gamma}} = \left(\sum_{i=1}^{200} \hat{\gamma}_i\right)/200$). Bias was simply computed as $\text{bias}(\hat{\gamma}) = \hat{\bar{\gamma}} - \gamma$, and percent relative bias was computed as

$$\% \text{bias}(\hat{\gamma}) = \frac{100\left(\hat{\bar{\gamma}} - \gamma\right)}{\gamma}. \tag{7}$$

The variance of the estimator, $\hat{\gamma}$, was computed as $\widehat{\text{var}}(\hat{\gamma}) = \frac{\sum_{i=1}^{200}\left(\hat{\gamma}_i - \hat{\bar{\gamma}}\right)^2}{199}$, and the RMSE was then computed as

$$\text{RMSE}(\hat{\gamma}) = \sqrt{[\text{bias}(\hat{\gamma})]^2 + \widehat{\text{var}}(\hat{\gamma})}. \tag{8}$$

Table 1 Simulation scenarios to assess the performance of community dynamic estimators. $N=100$, $\phi^l=0.70$, and $\phi^h=0.95$ for each set of parameters; where N is total species richness, ϕ^l is the complement of local extinction probability in the low-probability (low-p) group, and ϕ^h is the complement of local extinction probability in the high-p group, θ proportion of species in the low-p group, ϕ^w complement of overall (weighted) local extinction probability, p^l stop-level species detection probability in the low-p group, p^h stop-level species detection probability in the high-p group, $\bar{p}$ average stop-level species detection probability, κ stop-level coefficient of variation of species detection probabilities, p_r^l route-level species detection probability in the low-p group, p_r^h route-level species detection probability in the high-p group, $\bar{p}_r$ average route-level species detection probability

No.	θ	ϕ^w	Stop level				Route level		
			p^l	p^h	$\bar{p}$	κ	p_r^l	p_r^h	$\bar{p}_r$
1	0.3	0.875	0.00503	0.02827	0.02130	0.5	0.22291	0.76161	0.6
2	0.5	0.825	0.01037	0.03110	0.02073	0.5	0.40607	0.79393	0.6
3	0.7	0.775	0.01133	0.08290	0.03280	1.0	0.43423	0.98679	0.6
4	0.7	0.775	0.01683	0.12318	0.04874	1.0	0.57203	0.99860	0.7
5	0.7	0.775	0.02475	0.18111	0.07165	1.0	0.71431	0.99995	0.8
6	0.9	0.725	0.01648	0.04944	0.01978	0.5	0.56436	0.92077	0.6
7	0.9	0.725	0.02195	0.06585	0.02634	0.5	0.67035	0.96682	0.7
8	0.9	0.725	0.02972	0.08917	0.03567	0.5	0.77882	0.99063	0.8
9	0.9	0.725	0.01612	0.09671	0.02418	1.0	0.55624	0.99381	0.6
10	0.9	0.725	0.02174	0.13043	0.03261	1.0	0.66677	0.99908	0.7
11	0.9	0.725	0.02963	0.17780	0.04445	1.0	0.77778	0.99994	0.8
12	0.9	0.725	0.01609	0.17697	0.03218	1.5	0.55556	0.99994	0.6
13	0.9	0.725	0.02173	0.23906	0.04347	1.5	0.66667	1.00000	0.7
14	0.9	0.725	0.02963	0.32597	0.05927	1.5	0.77778	1.00000	0.8

Table 2 Community dynamic estimates based on avian species from Breeding Bird Survey (BBS) routes surveyed in 1976 and 1996. Numbers of routes on which the estimates are based are given for each state. Estimated SEs are given *in parentheses*. $\hat{N}$ estimate of total species richness, $\hat{\bar{p}}_r$ estimate of average route-level species detection probability, $\hat{\bar{\kappa}}$ estimate of the coefficient of variation for species detection probabilities, $\hat{\bar{\theta}}$ estimate of the proportion of species in the low-p group, $\hat{\bar{R}}^l$ average number of species detected in the low-p group, $\hat{\bar{R}}^h$ average number of species detected in the high-p group

| State | Maryland (39 routes) | | North Dakota (21 routes) | | Wisconsin (49 routes) | |
Year	1976	1996	1976	1996	1976	1996
$\hat{\bar{N}}$	72.9 (2.05)	78.4 (1.99)	66.0 (4.24)	80.8 (4.61)	81.1 (2.69)	85.2 (2.57)
$\hat{\bar{p}}_r$	0.92 (0.008)	0.92 (0.009)	0.90 (0.025)	0.88 (0.02)	0.92 (0.01)	0.89 (0.01)
$\hat{\bar{\kappa}}$	0.997 (0.013)	1.022 (0.014)	1.344 (0.035)	1.240 (0.041)	1.097 (0.020)	1.122 (0.014)
$\hat{\bar{\theta}}$	0.43 (0.01)	0.44 (0.01)	0.56 (0.02)	0.56 (0.02)	0.52 (0.01)	0.50 (0.01)
$\hat{\bar{R}}^l$	19.2 (0.59)	22.0 (0.81)	22.0 (1.11)	24.8 (1.34)	24.5 (0.92)	25.8 (0.84)
$\hat{\bar{R}}^h$	41.2 (1.02)	43.8 (1.18)	28.0 (1.51)	34.8 (2.18)	38.0 (1.33)	41.7 (1.01)

We view RMSE as the best overall metric reflecting estimator performance.

Results

Analysis of BBS data

Results are based on BBS routes monitored in 1976 and 1996 in Maryland (39 routes), North Dakota (21) and Wisconsin (49). Table 2 shows that the mean avian species richness estimates on these routes for all three states were smaller in 1976 than in 1996. The mean estimated κ of Lee and Chao (1994), Eq. 5, were largest in North Dakota and smallest in Maryland for both years, but detection probabilities appear to be quite variable among species in all three states. The state averages of our ad hoc estimates for the proportion of species in the low-p group ranged form 0.43 to 0.56, yielding adequate numbers of species detected in each species group.

In Table 3, we present other community parameter estimates and results from the Wilcoxon signed rank test statistics associated with our predictions. Our ad hoc method of partitioning the observed detection frequencies into two groups in 1976 yields estimates of route-level species detection probabilities for 1996 that are very consistent among the states. The state averages for route-level detection probability ranged between 0.80 and 0.83 for the low-p groups and between 0.90 and 0.93 for the high-p groups. We found strong evidence of different detection probabilities for the two groups in the three states. The state means for the complement of local extinction probability ranged from 0.70 to 0.75 for the low-p group and were >0.95 for the high-p group. The results provided strong evidence of lower extinction probabilities for species in the high-p group. The weighted estimates for the complements of local extinction (Eq. 4) were indeed generally lower than the estimates based on the unweighted estimators, but differences were smaller than anticipated.

Simulation results

Our ad hoc approach to estimating the complement of local extinction probability typically provided negatively biased estimates for the high-p group and positively biased estimates for the low-p group (Table 4). Both weighted and unweighted estimates of overall survival tended to be positively biased for most scenarios, and the percent relative bias for the weighted estimates was smaller, as expected (Table 5). However, the unweighted estimator

Table 3 Estimates of community-dynamic rate parameters based on avian species from BBS routes surveyed in 1976 and 1996 and significant values for the Wilcoxon signed-rank statistic for the associated comparison tests. Estimated SEs are given *in parentheses*. $\hat{\bar{p}}_r^l$ state average route-level detection probabilities for the low-p groups in 1996, $\hat{\bar{p}}_r^h$ state average route-level detection probabilities for the high-p groups in 1996, $\hat{\bar{\varphi}}^l$ estimate of the complement of local extinction probabilities for the low-p group, $\hat{\bar{\varphi}}^h$ estimate of the complement of local extinction probabilities for the high-p group, $\hat{\bar{\varphi}}^w$ estimate of the weighted complement of local extinction probabilities, $\hat{\bar{\varphi}}$ estimate of the unweighted complement of local extinction probabilities

Estimator	Maryland	North Dakota	Wisconsin
$\hat{\bar{p}}_r^l$	0.83 (0.019)	0.83 (0.020)	0.81 (0.020)
$\hat{\bar{p}}_r^h$	0.92 (0.008)	0.91 (0.018)	0.93 (0.007)
p-value[a]	≤0.0001	0.0024	≤0.0001
$\hat{\bar{\varphi}}^l$	0.75 (0.026)	0.73 (0.036)	0.70 (0.025)
$\hat{\bar{\varphi}}^h$	0.98 (0.005)	0.97 (0.011)	0.96 (0.007)
p-value[b]	≤0.0001	≤0.0001	≤0.0001
$\hat{\bar{\varphi}}^w$	0.88 (0.013)	0.84 (0.021)	0.83 (0.015)
$\hat{\bar{\varphi}}$	0.90 (0.012)	0.87 (0.016)	0.85 (0.014)
p-value[c]	0.0201	0.2262	0.0039

[a] For differences between route-level detection probabilities for the low- and high-p groups
[b] For differences between local extinction probabilities for the low- and high-p groups
[c] For differences between weighted and unweighted estimators of local extinction probabilities

Table 4 Simulation results for the complement of local extinction probability estimator in the low-p and high-p groups; the true parameter values were set to ϕ^l=0.70, and ϕ^h=0.95. $\hat{\varphi}^l$ estimates of the complement of local extinction probability for the low-p group, $\hat{\varphi}^h$ estimates of the complement of local extinction probability for the high-p group, *%BIAS* relative bias, *RMSE* root mean squared error; for other terms, see Table 1

Scenario	Estimates		%BIAS		RMSE	
	$\hat{\varphi}^l$	$\hat{\varphi}^h$	$\hat{\varphi}^l$	$\hat{\varphi}^h$	$\hat{\varphi}^l$	$\hat{\varphi}^h$
1	0.872	0.878	24.62	−7.57	0.206	0.165
2	0.823	0.873	17.64	−8.13	0.184	0.172
3	0.657	0.961	−6.11	1.11	0.213	0.060
4	0.687	0.941	−1.82	−0.98	0.213	0.063
5	0.729	0.91	4.16	−4.25	0.165	0.081
6	0.714	0.785	1.96	−17.34	0.172	0.249
7	0.747	0.800	6.69	−15.76	0.161	0.217
8	0.732	0.784	4.53	−17.46	0.128	0.207
9	0.723	0.834	3.28	−12.21	0.202	0.182
10	0.724	0.834	3.38	−12.23	0.161	0.181
11	0.725	0.801	3.50	−15.73	0.134	0.185
12	0.718	0.836	2.51	−12.03	0.205	0.171
13	0.724	0.838	3.48	−11.78	0.164	0.175
14	0.724	0.802	3.45	−15.58	0.134	0.185

Table 5 Simulation results for the complement of local extinction probability estimator where ϕ^w corresponds to the true parameter values based on the weighted approach. $\hat{\varphi}^w$ weighted estimates of the complement of local extinction probability, $\hat{\varphi}$ unweighted estimates of the complement of local extinction probability; for abbreviations, see Table 4

Scenario	ϕ^w	Estimates		%BIAS		RMSE	
		$\hat{\varphi}^w$	$\hat{\varphi}$	$\hat{\varphi}^w$	$\hat{\varphi}$	$\hat{\varphi}^w$	$\hat{\varphi}$
1	0.875	0.873	0.902	−0.18	3.13	0.101	0.102
2	0.825	0.829	0.853	0.44	3.39	0.123	0.122
3	0.775	0.742	0.743	−4.28	−4.08	0.156	0.133
4	0.775	0.768	0.789	−0.93	1.84	0.149	0.131
5	0.775	0.796	0.811	2.71	4.66	0.113	0.109
6	0.725	0.719	0.721	−0.81	−0.57	0.159	0.152
7	0.725	0.754	0.770	4.05	6.23	0.138	0.142
8	0.725	0.744	0.768	2.57	5.96	0.102	0.114
9	0.725	0.736	0.740	1.48	2.00	0.177	0.164
10	0.725	0.742	0.766	2.32	5.59	0.137	0.146
11	0.725	0.743	0.767	2.45	5.86	0.104	0.114
12	0.725	0.732	0.745	1.03	2.81	0.177	0.165
13	0.725	0.744	0.767	2.57	5.77	0.139	0.145
14	0.725	0.743	0.767	2.47	5.85	0.103	0.113

did not perform nearly as badly as we had expected, with %bias never exceeding 6.5% (Table 5). The RMSE for the weighted and unweighted estimators were very similar, leaving little basis for preference of one estimator over the other.

The estimator for the coefficient of variation of stop-level detection probability among species exhibited negative bias in all 14 scenarios (Table 6). Examination of

Table 6 Simulation results for proportion of species in the low-p group and coefficient of variation estimator. $\hat{\theta}$ estimates of the proportion of species in the low-p group, $\hat{\kappa}$ estimates of the stop level coefficient of variation of detection probabilities. For abbreviations, see Table 4

Scenario	True values		$\hat{\theta}$			$\hat{\kappa}$		
	θ	κ	Estimate	% BIAS	RMSE	Estimate	% BIAS	RMSE
1	0.3	0.5	0.871	190.4	0.57	0.205	−58.9	0.35
2	0.5	0.5	0.888	77.6	0.39	0.261	−47.8	0.30
3	0.7	1.0	0.723	3.3	0.07	0.621	−37.9	0.39
4	0.7	1.0	0.686	−1.9	0.07	0.754	−24.6	0.26
5	0.7	1.0	0.634	−9.5	0.09	0.849	−15.1	0.17
6	0.9	0.5	0.913	1.4	0.03	0.330	−34.0	0.27
7	0.9	0.5	0.857	−4.7	0.06	0.383	−23.3	0.21
8	0.9	0.5	0.767	−14.7	0.14	0.422	−15.7	0.18
9	0.9	1.0	0.879	−2.3	0.04	0.777	−22.3	0.27
10	0.9	1.0	0.832	−7.5	0.08	0.857	−14.3	0.18
11	0.9	1.0	0.753	−16.4	0.16	0.895	−10.5	0.16
12	0.9	1.5	0.868	−3.5	0.05	1.182	−21.2	0.34
13	0.9	1.5	0.830	−7.8	0.08	1.285	−14.3	0.24
14	0.9	1.5	0.752	−16.4	0.16	1.36	−9.3	0.17

RMSE suggests that the Lee and Chao (1994) estimator for coefficient of variation performs better when heterogeneity in detection probabilities is higher. We also examined the performance of $\hat{\theta}$ using our ad hoc approach of simply grouping species based on capture frequencies. The RMSE was reasonable when θ was relatively large (e.g., θ≥ 0.7), but was quite high for smaller θ (Table 6).

Discussion

Our analysis of BBS data provided strong evidence of a negative relationship between detection probability and local extinction probability. We have noted that species-level detection probabilities can be written as functions of detection probabilities for individual animals and the number of individuals in a species (e.g., Eqs. 1, 2). If the sole reasons for heterogeneous detection probabilities among species were differences among individuals of different species in behavior and activity patterns, then we would have expected to see no difference between local extinction probabilities of species detected frequently vs. infrequently. We believe that the substantial difference between local extinction probability estimates for our low-p and high-p groups can be attributed directly to differences in abundance, with the low-p species being much less abundant.

The results provide empirical support for the model-based prediction of a negative relationship between abundance and extinction probability (e.g. Bailey 1964; MacArthur and Wilson 1967; Goel and Richter-Dyn 1974; Gilpin and Soulé 1986; Boyce 1992; Burgman et al.

1993). Although other investigators have claimed to provide such evidence (e.g., see review of McKinney 1997), results have typically been ambiguous because species detection probabilities have not been considered as providing a possible alternative explanation. For example, assume a situation in which there is a relationship between detection probability and abundance, but no relationship between abundance and extinction probability. Of all species detected at time t, those with larger population size will be more likely detected at some later time, $t+\Delta$, and less abundant species will be less likely detected, lending the appearance of a negative relationship between extinction probability and abundance. In the work reported here, detection probabilities were explicitly included in our efforts to estimate numbers of surviving species from the low-p and high-p groups. We view the differences in local extinction probabilities for species in the two groups as an interesting and important ecological result.

This covariation analysis between species detection probabilities and community-level parameters was motivated by our interest in estimating local species extinction probability, local species turnover, and probability of species sharing two areas. The existence of detection probabilities <1 and their covariation with community parameters are also relevant to other ecological relationships involving communities. Consider nested subset analysis, in which communities in archipelagos, for example, are investigated to determine the degree to which species in smaller biotas are also found in larger ones (e.g., Patterson and Atmar 1986). The existence of detection probabilities <1 is typically not considered in nested subset investigations and could affect analyses in various ways. For example, consider a situation of three perfectly nested communities, in the sense that all species found in the smallest community are found in both larger communities, and all species found in the second-largest community are also found in the largest community. Species missed (not detected) in the second-largest community, yet detected in the largest and smallest community, would appear to deviate from the actual pattern of perfect nesting. On the other hand, the covariation between detection probability and both extinction probability and species sharing probability (the probability that a species present on one area is also present on another) could result in the species missed (not detected) on one area being those species least likely to be present on another area in the archipelago. This problem of species least likely to be shared by two communities also being least likely to be detected when present, would cause communities to appear nested to a greater degree than they actually were. Thus, we strongly recommend the use of species co-occurrence estimators (Nichols et al. 1998b) for nested subset analyses (see Cam et al. 2000). We also believe that our results are relevant to various other types of investigations in community ecology and emphasize the importance of considering detection probability explicitly in analyses.

Our empirical analyses were also directed at questions regarding the estimation of local extinction probability in

the face of the negative relationship of this quantity and detection probability. Our ad hoc weighted estimates of local species survival were higher than the estimates based on the original estimators proposed by Nichols et al. (1998a). However, the difference was much smaller than we had expected, suggesting that bias associated with the original estimators is not large. There was a slight tendency for the precision of estimates to be larger for our ad hoc weighted estimator, although the difference was again smaller than we had expected.

The simulation results permitted us to address estimator performance directly. The ad hoc weighted estimator and the original unweighted estimator performed similarly, as both tended to be slightly positively biased. RMSE for the two estimators were very similar, leaving no real basis for selecting one estimator over the other. Ease of computation leads us to slightly prefer the original unweighted estimator, which should be a reasonable selection for future analyses of BBS data.

The estimator for κ from Lee and Chao (1994) showed negative bias for all scenarios. Comparison of the estimators assessed here with results obtained using the estimators of Pledger (2000) deserves a good deal of attention. Pledger's approach seems promising because it allows estimation of species detection probability in each group as well as the probability of group membership (i.e., analogous to our ad hoc θ).

All our simulation results are obviously scenario-dependent as we tried to use sampling schemes that reflected results from the BBS data. Analyses of different kinds of community-level data may merit additional simulations tailored to other sampling situations.

In conclusion, we provided strong empirical evidence that local extinction probability covaries negatively with species detection probability and probably abundance of individuals within species. This empirical result is based on methods that permit heterogeneity in species detection probabilities and thus differs from virtually all previous work on this and related topics. Community-dynamic estimators (e.g., Nichols et al. 1998a, b) based on the jackknife estimator of Burnham and Overton (1978, 1979) appear to perform reasonably well for community analyses of data obtained under sampling situations similar to that of the BBS. When community-level work is conducted using data collected under different sampling schemes, then it may be wise to revisit questions about estimator performance.

660

References

Bailey NTJ (1964) The elements of stochastic processes with applications to the natural sciences. Wiley, New York

Bennetts RE, White GC, Hawksworth FG, Severs SE (1996) The influence of dwarf mistletoe on bird communities in Colorado ponderosa pine forests. Ecol Appl 6:899-909

Boulinier T, Nichols JD, Sauer JR, Hines JE, Pollock KH (1998a) Estimating species richness to make inferences in community ecology: the importance of heterogeneity in species detectability. Ecology 79:1018–1028

Boulinier T, Nichols JD, Hines JE, Sauer JR, Flather CH, Pollock KH (1998b) Higher temporal variability of forest breeding bird communities in fragmented landscapes. Proc Natl Acad Sci U S A 95:7497–7501

Boyce MS (1992) Population viability analysis. Ann Rev Ecol Syst 23:481–506

Bunge J, Fitzpatrick M (1993) Estimating the number of species: a review. J Am State Assoc 88:364–373

Burgman MA, Ferson S, Akçakaya HR (1993) Risk assessment in conservation biology. Chapman Hall, London

Burnham KP, Overton WS (1978) Estimation of the size of a closed population when capture probabilities vary among animals. Biometrika 65:625–633

Burnham KP, Overton WS (1979) Robust estimation of population size when capture probabilities vary among animals. Ecology 60:927–936

Cam E, Nichols JD, Hines JE, Sauer JR (2000) Inferences about nested subset structure when not all species are detected. Oikos 91:428–434

Carothers AD (1973) The effects of unequal catchability on Jolly-Seber estimates. Biometrics 29:79–100

Carothers AD (1979) Quantifying unequal catchability and its effect on survival estimates in an actual population. J Anim Ecol 48:863–869

Chao A, Lee SM, Jeng SL (1992) Estimation of population size for capture—recapture data when capture probabilities vary by time and individual animal. Biometrics 48:201–216

Dawson DK, Sauer JR, Wood PA, Berlanga M, Wilson MH, Robbins CS (1995) Estimating bird species richness from capture and count data. J Appl Stat 22:1063–1068

Derleth EL, McAuley DG, Dwyer TJ (1989) Avian community response to small-scale habitat disturbance in Maine. Can J Zool 67:385–390

Gaston KJ (1994) Rarity. Chapman and Hall, London

Gilpin ME, Soulé ME (1986) Minimum viable populations: processes of species extinctions. In: Soulé ME (ed) Conservation biology: the science of scarcity and diversity. Sinauer, Sunderland, Mass., pp 19–34

Goel NS, Richter-Dyn N (1974) Stochastic models in biology. Academic Press, New York

Hanski I, Kouki J, Halkka A (1993) Three explanations for the positive relationship between distribution and abundance of species. In: Ricklefs RE, Schluter D (eds) Species diversity in ecological communities. University of Chicago Press, Chicago, Ill., pp 108–116

Hines JE, Boulinier T, Nichols JD, Sauer JR, Pollock KH (1999) COMDYN: software to study the dynamics of animal communities using a capture–recapture approach. Bird Study 46:209–217

Karr JR, Robinson SK, Blake JG, Bierregaard RO Jr (1990) Birds of four neotropical forests. In: Gentry AH (ed) Four neotropical rainforests. Yale University Press, New Haven, Conn., pp 237–269

Lee SM, Chao A (1994) Estimating population size via sample coverage for closed capture–recapture models. Biometrics 50:88–97

MacArthur RH, Wilson EO (1967) The theory of island biogeography. Princeton University Press, Princeton, N.J.

McKinney ML (1997) Extinction vulnerability and selectivity: combining ecological and paleontological views. Annu Rev Ecol Syst 28:495–516

Nichols JD, Conroy MJ (1996) Estimation of species richness. In: Wilson DE, Cole FR, Nichols JD, Rudran R, Foster M (eds) Measuring and monitoring biological diversity. Standard methods for mammals. Smithsonian Institute Press, Washington, D.C., pp 226–234

Nichols JD, Boulinier T, Hines JE, Pollock KH, Sauer JR (1998a) Estimating rates of local extinction, colonization and turnover in animal communities. Ecol Appl 8:1213–1225

Nichols JD, Boulinier T, Hines JE, Pollock KH, Sauer JR (1998b) Inference methods for spatial variation in species richness and community composition when not all species are detected. Conserv Biol 12:1390–1398

Norris JL III, Pollock KH (1996) Nonparametric MLE under two closed capture–recapture models with heterogeneity. Biometrics 52:639–649

Patterson BD (1984) Mammalian extinction and biogeography in the southern rocky mountains. In: Nitecki MH (ed) Extinctions. University of Chicago Press, Chicago, Ill., pp 247–293

Patterson BD, Atmar W (1986) Nested subsets and the structure of insular mammalian faunas and archipelagos. Biol J Linn Soc 28:65–82

Pledger S (2000) Unified maximum likelihood estimates for closed capture–recapture models using mixtures. Biometrics 56:434-442

Pollock KH (1982) A capture–recapture sampling design robust to unequal catchability. J Wildl Manage 46:752–757

Robbins CS, Bystrack D, Geissler PH (1986) The breeding bird survey: its first fifteen years, 1965–1979. Resources publication no. 157. US Fish and Wildlife Services, , Washington, D.C.

Sauer JR, Barker RJ, Geissler PH (1994) Statistical aspects of modeling population change from population size data. In: Kendall RJ, Lacher TE Jr (eds) Wildlife toxicology and population modeling. Integrated studies of agroecosystems. CRC, Boca Raton, Fla., pp 451–466

Smith EP, van Belle G (1984) Nonparametric estimation of species richness. Biometrics 40:119–129

Thiollay JM (1995) The role of traditional agroforests in the conservation of rain forest bird diversity in Sumatra. Conserv Biol 9:335–353