Biometrics 62, 97–102
March 2006

DOI: 10.1111/j.1541-0420.2005.00439.x

Site Occupancy Models with Heterogeneous Detection Probabilities

J. Andrew Royle

USGS Patuxent Wildlife Research Center, 12100 Beech Forest Road, Laurel,
Maryland 20708, U.S.A.
email: aroyle@usgs.gov

Summary. Models for estimating the probability of occurrence of a species in the presence of imperfect detection are important in many ecological disciplines. In these "site occupancy" models, the possibility of heterogeneity in detection probabilities among sites must be considered because variation in abundance (and other factors) among sampled sites induces variation in detection probability (p). In this article, I develop occurrence probability models that allow for heterogeneous detection probabilities by considering several common classes of mixture distributions for p. For any mixing distribution, the likelihood has the general form of a zero-inflated binomial mixture for which inference based upon integrated likelihood is straightforward. A recent paper by Link (2003, *Biometrics* **59,** 1123–1130) demonstrates that in closed population models used for estimating population size, different classes of mixture distributions are indistinguishable from data, yet can produce very different inferences about population size. I demonstrate that this problem can also arise in models for estimating site occupancy in the presence of heterogeneous detection probabilities. The implications of this are discussed in the context of an application to avian survey data and the development of animal monitoring programs.

Key words: Animal sampling; Detection probability; Occurrence probability; Site occupancy.

1. Introduction

Modeling the occurrence of a species is an important objective of many ecological studies. So-called site occupancy (also probability of occurrence, and proportion of area occupied) is a quantity of some interest in ecology, and recent monitoring activities for some species have adopted a focus on occurrence-based summaries of species status (e.g., the Amphibian Research & Monitoring Initiative [ARMI]; Hall and Langtimm, 2001). This practical interest is motivated, at least in part, by the generality and simplicity of sampling designs that yield information about occurrence. Models of site occupancy require only observations of presence/absence and, consequently, sampling is efficient in many field situations, and across taxa.

When the probability of detecting a species is <1, nondetection of that species at a site is not indicative of nonoccurrence. Much recent work has focused on estimating occurrence-based summaries in the presence of imperfect detection (e.g., Bayley and Peterson, 2001; Kéry, 2002; MacKenzie et al., 2002; Royle and Nichols, 2003; Dorazio and Royle, 2005). Conventional models for estimating site occupancy regard detection probability, p, as constant, or as a function of measurable covariates. One important source of variation in detection probabilities that has been considered extensively in conventional capture–recapture problems is that due to heterogeneity among individuals. In the context of estimating site occupancy, analogous models would allow for variation in detection probability among sites. It is necessary to consider such models for estimating site occupancy because factors that influence detectability are many and varied, and it may not be possible to control for all important sources of variation in detectability through a careful study design in many field situations.

The literature on modeling heterogeneity in the context of estimating the size of a closed population is vast (e.g., Burnham and Overton, 1978; Norris and Pollock, 1996; Coull and Agresti, 1999; Fienberg, Johnson, and Junker, 1999; Dorazio and Royle, 2003; Link, 2003). These models suppose that p varies by individual and consider various mixture distributions on p including beta, logit-normal, or finite mixtures of point support. In this article, I develop analogous classes of models for estimating site occupancy, ψ, that allow for heterogeneous detection probabilities. In Section 2, I present a general formulation of these models and consider several commonly used mixture distributions for p including continuous and discrete mixtures. In Section 4, I present an application of these models to avian survey data for a number of species.

The second objective of this article, considered in Section 5, is to address the issue raised by Link (2003) concerning the identifiability of model parameters in the presence of heterogeneous detection probabilities. Link (2003) demonstrated that in closed population models for estimating abundance, abundance is *not* identifiable in the presence of heterogeneity in the sense that different mixture distributions may give rise to identically distributed data (or nearly so), yet produce substantially different inferences about abundance. It is natural to consider whether this phenomenon is important in analogous models for estimating site occupancy. In closed population models for estimating abundance, the inability to observe the "zero frequency" (number of uncaptured individuals) is the genesis of the nonidentifiability problem reported by Link (2003). However, in occupancy models, the zero frequency *is* observed. Unfortunately, site occupancy models introduce uncertainty about nondetection and nonoccurrence (by introduction of the site occupancy parameter) that may largely mitigate any ability to partition the observed zeros into those due to nondetection and nonoccurrence in the presence of heterogeneity.

2. Estimating Site Occupancy with Imperfect Detection

MacKenzie et al. (2002) consider the problem of estimating the probability of occurrence based on observations of presence/absence subject to imperfect detection. The sampling design requires that R sites are sampled (for presence/absence) on T occasions. Repeated sampling at a site yields a "detection history," e.g., (0, 0, 1, 0) for a site sampled four times and detection occurred only at the third sample. Samples may yield no detections (e.g., (0, 0, 0, 0)) either because the species was present but not detected, or because the species was not present. The statistical problem is essentially to partition the "zero sites" into those where the species occurs but was not detected (sampling zeros) and those where it does not occur (structural zeros).

For a model in which there is no temporal variation in detection probability, the data are the number of detections for each of the R sites: y_i; $i = 1, 2, \ldots, R$. For sites where the species occurs, y may be viewed as a binomial random variable. To accommodate nonoccurrence, we introduce the zero-inflation parameter, $1 - \psi$. Then, the sampling distribution of y_i is the zero-inflated binomial distribution

$$[y_i \mid p, \psi] = \text{Bin}(y_i; T, p)\psi + (1 - \psi)I(y_i = 0), \tag{1}$$

where $I(y_i = 0)$ is an indicator function that evaluates to 1 for $y_i = 0$ and 0 otherwise. The joint likelihood based on data from R sites is the product of R versions (one for each site) of equation (1). Zero-inflated probability distributions are commonly used in many biological applications (e.g., see Ridout, Demetrio, and Hinde, 1998).

The extension of this model considered subsequently is to the case where p varies among sample locations according to some class of probability distributions.

3. Site Occupancy Models with Heterogeneous Detection

For the case where there is no additional structure on p or ψ (e.g., time variation or covariate effects), the sufficient statistic is the vector of count frequencies n_k; $k = 0, 1, \ldots, T$ where n_k is the number of sites at which k detections were observed. Then, inference is based on the $(T + 1)$cell multinomial describing this frequency distribution.

Let $g(p; \theta)$ be some class of mixture distributions (e.g., beta) for p with parameter θ. Conditional on a site being "occupied," the probability of obtaining k detections at a site may be computed by integrating the binomial likelihood over g

$$\pi_g^{(c)}(k; \theta) = \Pr(y = k \mid \theta) = \int_0^1 \text{Bin}(k; T, p) g(p; \theta)\, dp. \tag{2}$$

The superscript (c) here indicates that this is conditioned on occurrence. The marginal (zero-inflated) cell probabilities for the data are of the form

$$\pi_g(k; \theta, \psi) = \psi \pi_g^{(c)}(k; \theta) + I(k = 0)(1 - \psi). \tag{3}$$

Then, the likelihood based on the observed data is the zero-inflated g-binomial mixture given by

$$L(\psi, \theta \mid \mathbf{n}) = \prod_{k=0}^{T} \pi_g(k; \theta, \psi)^{n_k}.$$

This general form holds regardless of the choice of mixture distribution g. For any choice of g, evaluation of the likelihood requires only that the $T + 1$ univariate integrals in (2) be computed. Several common choices of g are considered subsequently.

3.1 *Beta Mixtures*

A natural choice for g is the beta distribution with parameters α and β. In this case, the integrated likelihood is a zero-inflated beta-binomial with

$$\pi_{bb}^{(c)}(k; \theta) = \frac{\Gamma(T+1)}{\Gamma(k+1)\Gamma(T-k+1)} \frac{\Gamma(\alpha+k)\Gamma(T+\beta-k)}{\Gamma(\alpha+\beta+T)} \times \frac{\Gamma(\alpha+\beta)}{\Gamma(\alpha)\Gamma(\beta)}. \tag{4}$$

This is commonly parameterized in terms of the mean $\mu = \alpha/(\alpha + \beta)$ and precision $\tau = \alpha + \beta$, or the location-scale parameters μ and $\sigma = (\mu(1 - \mu)/(\tau + 1))^{1/2}$.

3.2 *Finite Mixtures*

Another common class of mixture distributions are the finite mixtures (Norris and Pollock, 1996; Pledger, 2000) which suppose that g consists of a mixture of m components with support points $p_1, p_2, \ldots, p_m$ and masses $g_1, g_2, \ldots, 1 - \sum_{j=1}^{m-1} g_j$. For example, consider g to be a two-component finite mixture with support points p_1 and p_2 and masses g_1 and $1 - g_1$. Then,

$$\pi_{fm}^{(c)}(k; \theta) = \text{Bin}(k; T, p_1)g_1 + \text{Bin}(k; T, p_2)(1 - g_1),$$

with parameter $\theta = (p_1, p_2, g_1)$.

3.3 *Normal Mixtures*

Another natural class of models for describing variation in p is the logit-normal class of models (Coull and Agresti, 1999) in which $\text{logit}(p_i)$ is assumed to have a normal distribution with mean μ and standard deviation σ. In this case, the integral in equation (2) must be evaluated numerically.

3.4 *Abundance Models*

Royle and Nichols (2003) noted that heterogeneity in p can be induced by variation in abundance among sites and they developed a model of heterogeneity based on this consideration. Specifically, let N_i be the abundance at site i, then binomial sampling yields that the "net" probability of detection (of at least one individual) is

$$p_i = 1 - (1 - r)^{N_i}, \tag{5}$$

where r is the *individual* capture probability. Then, they place the mixture distribution on N_i, and construct the integrated likelihood by integrating over possible abundance states. For example, suppose that N_i has a Poisson distribution with mean λ. Then, the integrated likelihood is

$$\pi_{rn}(k;\theta) = \sum_{N_i=0}^{\infty} \text{Bin}(k; T, 1-(1-r)^{N_i}) \frac{e^{-\lambda}\lambda^{N_i}}{N_i!}.$$

The appeal of this model is that in some cases it may be reasonable to view λ as density. However, there is no reason that N_i must be interpreted as abundance per se. It could be viewed as a generic random effect that yields variation in p_i and thus we might view this model merely as an alternative mixing distribution that accommodates heterogeneity in detection probability. Indeed, a referee of an earlier version of this article noted a relationship between this model and the logit-normal model. Specifically, note that the complementary log-log link of p_i, under the structure specified by equation (5), implies that

$$\log(\log(1-p_i)) = \log(\log(1-r)) + \log(N_i),$$

thus yielding a model with an additive random effect, say $u_i \equiv \log(N_i)$.

4. Application

Here we demonstrate the application of models allowing for heterogeneous detection probabilities to avian survey data. The specific data used here consist of observations of detection/nondetection from a single route (50 "stops"), sampled 11 times during an approximately 1-month sample period. We consider data on five species: blue jay, common yellowthroat, song sparrow, catbird, and tree swallow (JAY, CYT, SOSP, CATB, and TREE, respectively). The data for each species are given in Table 1. In addition to the constant-p model, we consider heterogeneity models based on the logit-normal (LN), beta-binomial (BB), two-component finite mixture (FM2), and the Royle and Nichols abundance (RN) models.

To evaluate model adequacy, we consider the usual deviance statistic based on the expected cell frequencies under the model in question

$$D_g = 2\sum_k n_k(\log(\pi_{\text{sat}}(k)) - \log(\pi_g(k;\hat{\theta}))).$$

Here, $\pi_{\text{sat}}(k)$ is the observed proportion of sites with k detections and $\pi_g(k)$ are the model-based estimates obtained by plugging the maximum likelihood estimate (MLE) of model parameters into equation (3). Asymptotically, we can expect this statistic to have a chi-square distribution on $T - dim(g) - 1$ degrees of freedom where $dim(g)$ is the number of parameters in the heterogeneity model g. In practice, we often fail to achieve large samples, which renders the null distribution of this statistic invalid. In such cases we would probably rely on the usual strategies for assessing goodness of fit. That is, in small samples we might pool cells where the expected cell count is less than five. Alternatively, we might estimate the small-sample null distribution using common resampling strategies such as the parametric bootstrap. This would be preferred when the number of cell frequencies is small so as to impair pooling of cells. Rather than undertaking a bootstrap characterization of the proper null distribution for each model we will present deviance statistics merely as indices of relative fit which seems to be a generally accepted use in such instances.

Table 1

Number of detections (out of $T = 11$ samples) of each species at each of the 50 BBS sample locations. Sample occupancy is the observed fraction of occupied sites.

	Number of detections												Sample
Species	0	1	2	3	4	5	6	7	8	9	10	11	occupancy
JAY	17	9	11	6	5	2	0	0	0	0	0	0	0.66
CYT	14	6	7	5	3	1	4	5	3	2	0	0	0.72
SOSP	24	5	1	5	1	5	3	1	3	1	0	1	0.52
CATB	31	6	4	5	2	0	2	0	0	0	0	0	0.38
TREE	21	7	8	1	3	3	1	2	3	0	1	0	0.58

Parameter estimates, asymptotic standard errors for $\hat{\psi}$, and summary statistics obtained from fitting each of the candidate models to the avian survey data are presented in Table 2. Note that the estimated mean and standard deviation of p under each model are given. Except for the FM2 and RN models there is a direct correspondence between the model parameters and the mean and variance of p. For the FM2 model the parameter estimates were (p_1, p_2, g_1): (0.077, 0.210, 0.142), (0.177, 0.621, 0.588), (0.182, 0.595, 0.478), (0.156 0.389, 0.786), and (0.139, 0.592, 0.661) for the JAY, CYT, SOSP, CATB, and TREE, respectively. Under the RN model, the parameter estimates for these species (in the same order) were (r, λ): (0.094, 0.806), (0.233, 0.765), (0.310, 0.567), (0.167, 0.428), and (0.222, 0.641) for the five species.

For the JAY, we should probably feel confident concluding that there does not exist heterogeneity in p among sites. None of the heterogeneity models yields more than a marginal improvement in negative log likelihood ($-2ll$ in Table 2) and the constant-p model fits the data as well as the heterogeneity models. Estimates of site occupancy are similar between the constant-p and most of the heterogeneity models (excepting the RN model, which was worst fitting).

Heterogeneity in detection probability is indicated for the remaining species. For CYT, the two-component finite mixture appears favored over all other models (by at least 6 $-2ll$ units), with $\hat{\psi} = 0.773$, and has a much smaller deviance. For SOSP, the beta or LN models appear to be equally favored with $\hat{\psi} = 0.597$ under the beta model and $\hat{\psi} = 0.578$ under the LN model. Although we are unable to make a decisive conclusion regarding the preferred model in this case, the fact that the estimates are similar among the two best models might render the model selection issue irrelevant. Heterogeneity is also indicated for the remaining two species (CATB, TREE), but it is less clear which heterogeneity model is to be preferred. All heterogeneity models have similar negative log likelihoods and fit the data about equally well. While this ambiguity may be innocuous for the CATB data because the estimates of ψ are similar across the four models (ranging from 0.428 under RN to 0.466 under the beta model), the differences are very pronounced for the TREE data. We consider this problem further in the following section.

Table 2

Parameter estimates and summary statistics from each site occupancy model fit to avian survey data

Species	Model	$\widehat{E[p]}$	$\hat{\sigma}_p$	$\hat{\psi}$	$SE(\hat{\psi})$	$-2ll$	np	DEV
JAY	const	0.199	0	0.723	0.054	164.08	2	1.88
	LN	0.197	0.023	0.728	0.062	164.07	3	1.87
	BB	0.197	0.024	0.728	0.063	164.07	3	1.87
	FM2	0.191	0.046	0.753	0.143	164.01	4	1.81
	RN	0.178	0.085	0.806	0.077	165.45	2	3.25
CYT	const	0.385	0	0.723	0.045	250.55	2	41.22
	LN	0.341	0.218	0.809	0.065	218.26	3	8.93
	BB	0.330	0.221	0.835	0.082	218.26	3	8.93
	FM2	0.360	0.218	0.773	0.052	212.40	4	3.07
	RN	0.374	0.147	0.765	0.042	222.77	2	13.44
SOSP	const	0.419	0	0.521	0.050	215.08	2	37.89
	LN	0.377	0.237	0.578	0.065	188.43	3	11.24
	BB	0.365	0.243	0.597	0.077	188.18	3	10.99
	FM2	0.398	0.206	0.549	0.056	192.18	4	14.99
	RN	0.403	0.132	0.567	0.049	196.25	2	19.06
CATB	const	0.219	0	0.407	0.053	132.11	2	8.05
	LN	0.196	0.102	0.454	0.073	130.08	3	6.02
	BB	0.191	0.106	0.466	0.087	130.04	3	5.98
	FM2	0.206	0.095	0.433	0.063	129.99	4	5.92
	RN	0.209	0.075	0.428	0.058	130.33	2	6.27
TREE	const	0.331	0	0.587	0.050	223.96	2	43.69
	LN	0.249	0.225	0.777	0.127	189.97	3	9.70
	BB	0.211	0.225	0.916	0.286	190.03	3	9.78
	FM2	0.293	0.215	0.665	0.070	189.11	4	8.84
	RN	0.317	0.123	0.641	0.049	199.91	2	19.64

5. Identifiability of ψ

In closed population models for estimating population size different classes of mixture distributions can yield very similar, or even identical, expected cell frequencies (and hence deviance), yet suggest very different population sizes (Link, 2003). Here I consider the relevance of Link's result to the classes of models described previously.

Note that it is possible to construct two distributions, say g and f, that yield the same marginal cell probabilities for the data, but different ψ's. If the first T moments of f are proportional to the first T moments of g (Link, 2003, gives an example), then $\pi_f(k) = c\pi_g(k)$ for some c and for $k = 1, 2, \ldots, T$. Consequently, $1 - \pi_f(0) = c(1 - \pi_g(0))$ and $\pi_f(0) = 1 - c(1 - \pi_g(0))$. Thus, the zero-inflated f mixture has cell probabilities

$$\pi_f(k; \theta, \psi) = \psi\pi_f(k; \theta) + I(k = 0)(1 - \psi),$$

whereas the equivalent zero-inflated g-mixture has cell probabilities

$$\pi_g(k; \theta, \psi) = c\psi\pi_g(k) + I(k = 0)(1 - c\psi).$$

Clearly, ψ is not identifiable across mixtures in this case. If truth is that p is distributed according to f with site occupancy ψ, then an equivalent representation for the distribution of the data can be obtained by considering g with site occupancy $c\psi$. Thus, unless $c = 1$ (in which case $f = g$), a unique estimate of site occupancy will not exist when there is uncertainty about the mixture distribution.

In general, common families of mixture distributions are not related in this fashion. However, Link (2003) established empirically that in certain instances one could not reasonably expect to distinguish between alternative mixture distributions from data. To demonstrate this, he chose g (e.g., LN) and f (e.g., beta) to minimize twice the Kullback–Liebler distance between π_g and π_f

$$KL_f = 2\sum_{k=0}^{T} \pi_g(k) \log\left(\frac{\pi_g(k)}{\pi_f(k)}\right).$$

Link (2003) noted that R (number of sampled sites) times KL_f is the noncentrality parameter that can be used to assess the power of a goodness-of-fit test of f against the g alternative.

For example, suppose that data are collected on $T = 5$ sampling occasions, that the logits of p_i have a normal distribution with $\mu = -2$ and $\sigma = 1$ (corresponding to Table 2 in Link, 2003), and that $\psi = 0.75$. The closest (in the Kullback–Liebler sense) models in each of the other classes are given in Table 3.

The marginal cell probabilities are very similar across the five models (indeed, nearly identical in some cases) and the noncentrality parameter is small in all cases, indicating low power to choose among them. This is consistent with the results reported by Link (2003) and elsewhere (Williams, 1982; Dorazio and Royle, 2003). For example, if one obtained a sample of size $R = 200$, the power to correctly reject the constant-p model is 0.274. Importantly, the models imply very different values of ψ.

Considering the results of Section 4, it is natural to question whether the results summarized in Table 3 are overly influenced by the imposed design, e.g., low mean p, or that

Table 3

Marginal cell probabilities and ψ for different models of p. The mixture models are logit-normal (LN), beta (BB), two-component finite mixture (FM2), and the Poisson abundance model of Royle and Nichols (RN). p-value is the power to reject that model under the specified logit-normal with $R = 200$ sample sites.

		Cell probabilities							
Model	ψ	$k = 0$	$k = 1$	$k = 2$	$k = 3$	$k = 4$	$k = 5$	$2 \times KL$	p-value
LN ($\mu = -2$, $\sigma = 1$)	0.75	0.630	0.222	0.097	0.037	0.011	0.002	0	0.05
const. $p = 0.23$	0.51	0.630	0.205	0.122	0.037	0.005	0.000	0.017	0.27
BB ($\mu = 0.13$, $\tau = 5.87$)	0.91	0.630	0.223	0.096	0.037	0.011	0.002	2.8e-05	0.05
FM2 $p = (0.15, 0.48)$, $f_1 = 0.87$	0.62	0.630	0.222	0.097	0.036	0.012	0.002	6.6e-05	0.05
RN ($r = 0.15$, $\lambda = 0.84$)	0.57	0.629	0.213	0.111	0.037	0.008	0.001	0.0048	0.10

there are only five sample occasions (yielding six observed cell frequencies). One might hope that the models become more distinguishable (or that ψ becomes more similar) by either increasing T or increasing the mean detection probability (which may be controllable to some extent by design). Empirical studies (not reported here) indicated that higher T (e.g., $T = 10, 20$), generally led to a marginal improvement in the ability to distinguish some models, but generally did not resolve the problem as two or more models would have nearly the same expected cell frequencies, but very different values of ψ. On the other hand, when the mass of $g(p)$ is moved away from zero, by increasing μ, the expected frequency distributions remain very similar but the differences in ψ become much less pronounced regardless of T, rendering the cost of model misspecification less important. Similarly, in some applications it might be possible to reduce heterogeneity by careful design, or identification of covariates that influence detection probability. Repeating the previous exercises with $\sigma = 0.5$ (with $T = 5, 10$, $\mu = -2, -1$, and $\psi = 0.75$) yields the results shown in Table 4. We see that again power to choose among alternatives is low, near the nominal 0.05 level in some cases, but the effect of misspecification is less severe.

6. Discussion

In this article, I developed models for estimating site occupancy that allow for heterogeneity in detection probability among sample units. I also considered the issue of identifiability of ψ across mixture distributions raised by Link (2003) in a related class of models.

The effect of misspecification of heterogeneity models is related to the degree of heterogeneity and the mean detection probability. As Link (2004) noted, differences among mixtures are more pronounced as the mass of $g(p)$ is concentrated near zero. This effect was apparent for two of the species to which several heterogeneity models were fit (Section 4). For these species (catbird and tree swallow), low indicated detection probability in combination with heterogeneity yielded ambiguous conclusions about which mixture model should be favored. Importantly, substantial differences in estimated site occupancy across mixtures resulted for the tree swallow. Less substantial differences were apparent for the catbird data, in part because less heterogeneity in p was indicated. Indeed, for the catbird, heterogeneity models improved $-2ll$ only by about two units whereas, for the tree swallow, the heterogeneity models yielded substantially greater improvements.

Table 4

Site occupancy parameter (ψ) and Kullback–Liebler distance from LN model under alternative mixture distributions for p with $T = 5$ and $T = 10$. LN model has $\mu = (-2, -1)$, $\sigma = 0.5$, and $\psi = 0.75$. p-value is the power to reject that model under the specified LN with $R = 200$ sample sites.

	$T = 5$			$T = 10$		
Model	ψ	$2 \times KL$	p-value	ψ	$2 \times KL$	p-value
			$\mu = -2$			
LN	0.750	0.0	0.05	0.750	0.0	0.05
const	0.648	9.0e-4	0.06	0.672	8.4e-3	0.11
BB	0.768	9.2e-7	0.05	0.764	2.3e-5	0.05
FM2	0.720	3.3e-7	0.05	0.726	4.7e-5	0.05
RN	0.782	2.9e-4	0.05	0.774	3.7e-4	0.05
			$\mu = -1$			
LN	0.750	0.0	0.05	0.750	0.0	0.05
const	0.694	4.9e-3	0.10	0.721	3.8e-2	0.43
BB	0.755	1.8e-6	0.05	0.753	4.0e-5	0.05
FM2	0.737	3.9e-6	0.05	0.741	4.4e-4	0.05
RN	0.785	1.3e-3	0.06	0.762	8.1e-3	0.11

For the remaining three species, there was little or no ambiguity among classes of mixture models. These results suggest that while identifiability of ψ is an important problem to be aware of, it may not inhibit inference.

However, there are cases where the mass of $g(p)$ is such that substantially different values of ψ are indicated under different models. When faced with discrepant estimates of ψ, obtained from models that all appear to fit the data, it is not clear what conclusion (if any) can be drawn. Consequently, monitoring programs that emphasize site occupancy as a metric of population status should consider the possibility that the existence of heterogeneity may diminish the utility of such metrics. One might view low mean detection or high levels of heterogeneity as suggesting that the species of interest cannot be reliably, or effectively, sampled. This should be viewed as a biological sampling issue to be considered in survey design prior to data collection, and not a statistical issue to rectify after the fact by considering complex models of the detection process. The results of Section 5 suggest that the latter is not a viable option. Thus, monitoring efforts that focus on site occupancy should take steps to minimize heterogeneity or to increase detection probability by, for example, establishment of rigorous sampling protocols, or identification of covariates that affect detection probability. While this may be difficult in the context of estimating population size, in surveys for estimating site occupancy it is often possible to measure a number of covariates about the site being sampled, and the conditions under which sampling occurs, and sampling methods are more flexible in many cases. In closed population models, high p may be unrealistic or, if achieved, suggests that most of the population was captured, hence reducing the need for complex estimators of abundance in the first place. This is not the case in site occupancy models wherein there will always be ambiguous zeros, at least for species for which site occupancy is a useful metric of a demographic state.

Acknowledgements

I would like to thank the associate editor and an anonymous referee for many constructive comments, and Dr William A. Link for helpful discussions on the identifiability problem.

References

Bayley, P. B. and Peterson, J. T. (2001). An approach to estimate probability of presence and richness of fish species. *Transactions of the American Fisheries Society* **130,** 620–633.

Burnham, K. P. and Overton, W. S. (1978). Estimation of the size of a closed population when capture probabilities vary among animals. *Biometrika* **65,** 625–633.

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

Dorazio, R. M. and Royle, J. A. (2003). Mixture models for estimating the size of a closed population when capture rates vary among individuals. *Biometrics* **59,** 351–364.

Dorazio, R. M. and Royle, J. A. (2005). Estimating size and composition of biological communities by modeling the occurrence of species. *Journal of the American Statistical Association* **100,** 389–398.

Fienberg, S. E., Johnson, M. S., and Junker, B. W. (1999). Classical multilevel and Bayesian approaches to population size estimation using multiple lists. *Journal of the Royal Statistical Society of London, Series A* **163,** 383–405.

Kéry, M. (2002). Inferring the absence of a species—A case study of snakes. *Journal of Wildlife Management* **66,** 330–338.

Hall, R. J. and Langtimm, C. A. (2001). The U.S. National Amphibian Research and Monitoring Initiative and the role of protected areas. *George Wright Forum* **18,** 14–25.

Link, W. A. (2003). Nonidentifiability of population size from capture–recapture data with heterogeneous detection probabilities. *Biometrics* **59,** 1123–1130.

MacKenzie, D. I., Nichols, J. D., Lachman, G. B., Droege, S., Royle, J. A., and Langtimm, C. A. (2002). Estimating site occupancy rates when detection probabilities are less than one. *Ecology* **83,** 2248–2255.

Norris, J. L. and Pollock, K. H. (1996). Nonparametric MLE under two closed-capture models with heterogeneity. *Biometrics* **52,** 639–649.

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

Ridout, M. S., Demetrio, C. G. B., and Hinde, J. P. (1998). Models for count data with many zeros. *Proceedings of XIXth International Biometric Society Conference* (*IBC*98), 179–192.

Royle, J. A. and Nichols, J. D. (2003). Estimating abundance from repeated presence-absence data or point counts. *Ecology* **84,** 777–790.

Williams, D. A. (1982). Extra-binomial variation in logistic-linear models. *Applied Statistics* **31,** 144–148.

Received September 2004. *Revised June* 2005.
Accepted July 2005.