Biometrics 60, 108–115
March 2004

N-Mixture Models for Estimating Population Size from Spatially Replicated Counts

J. Andrew Royle

Division of Migratory Bird Management, U.S. Fish and Wildlife Service,
11510 American Holly Drive, Laurel, Maryland 20708, U.S.A.
email: Andy_Royle@fws.gov

Summary. Spatial replication is a common theme in count surveys of animals. Such surveys often generate sparse count data from which it is difficult to estimate population size while formally accounting for detection probability. In this article, I describe a class of models (*N*-mixture models) which allow for estimation of population size from such data. The key idea is to view site-specific population sizes, N, as independent random variables distributed according to some mixing distribution (e.g., Poisson). Prior parameters are estimated from the marginal likelihood of the data, having integrated over the prior distribution for N. Carroll and Lombard (1985, *Journal of American Statistical Association* **80,** 423–426) proposed a class of estimators based on mixing over a prior distribution for detection probability. Their estimator can be applied in limited settings, but is sensitive to prior parameter values that are fixed a priori. Spatial replication provides additional information regarding the parameters of the prior distribution on N that is exploited by the *N*-mixture models and which leads to reasonable estimates of abundance from sparse data. A simulation study demonstrates superior operating characteristics (bias, confidence interval coverage) of the *N*-mixture estimator compared to the Caroll and Lombard estimator. Both estimators are applied to point count data on six species of birds illustrating the sensitivity to choice of prior on p and substantially different estimates of abundance as a consequence.

Key words: Avian point counts; Binomial population size estimation; North American Breeding Bird Survey.

1. Introduction

Estimation of the size of a binomial population using capture–recapture methods has a long history in studies involving animal populations (Seber, 1982; Pollock, 2000; Williams, Nichols, and Conroy, 2002). The essence of such methods is accounting for capture probability using recapture, resighting, or similar information on marked individuals in the population.

Capture–recapture studies can be difficult to implement. These studies typically require intense effort, and may not even be practical in instances where marking is difficult, or when the population experiences intense movement, mortality, and other forces. Even where practical, if the population size is small or the capture probability is low, it may be difficult to acquire enough data to ensure reasonable estimation of N. On the other hand, counting organisms without explicit regard to their identity is convenient and forms the basis of many large-scale animal monitoring programs (e.g., North American Breeding Bird Survey [BBS], North American Amphibian Monitoring Program [NAAMP], Christmas Bird Count [CBC]) and many more focused ecological studies. Monitoring programs such as these yield a large number of spatially indexed counts that can be characterized by their general sparsity for most species that are encountered, i.e., a large number of zeros and small counts. This is due in part to small populations and that many species have low detection probabilities due to their secretive habits and are usually detected only by their song or call.

As an illustration of the type of data that typically arise from avian point counts, Table 1 contains point count data from the American redstart (*Setophaga ruticilla*) for five locations sampled on 10 days. These data are described in more detail in Section 5. The data of Table 1 are typical of those collected in many ecological studies that involve counting organisms in that sampling is replicated in space and time, and that the resulting counts are relatively sparse.

For data such as these, estimation of site-specific abundance is not usually of direct interest. Data limitations aside, there is little ecological motivation for interest in such small, localized populations. However, an estimate of average abundance over a number of locations within some region can form the basis of assessment of population change over time. Also, although covariate information is not available for the analyses presented in Section 5, an understanding of the relationship between abundance and explicit descriptions of habitat or landscape is an underlying theme in many ecological investigations.

Abundance, or population size, is fundamental in the statement of such objectives. However, most analyses rely on simple counts such as those given in Table 1 without explicit

Downloaded from https://academic.oup.com/biometrics/article/60/1/108/7289296 by James Cummings user on 16 September 2024

Table 1
Point count data for the American redstart from five sample locations

	Sampling occasion									
	1	2	3	4	5	6	7	8	9	10
Site 1	0	0	0	1	0	0	0	0	0	0
Site 2	0	0	0	0	0	0	0	0	0	0
Site 3	1	0	0	0	0	0	0	0	0	0
Site 4	0	1	1	3	1	2	2	1	0	1
Site 5	2	0	1	1	0	0	1	0	0	0

regard to the underlying detection process which generated those counts. No effective approach for the estimation of abundance from such data while formally accommodating the detection process has been proposed. The focus of this article is the development of such methods.

For closed populations, Haldane (1942) and Olkin, Petkau, and Zidek (1981) have shown that it is possible to estimate population size from temporally repeated counts of organisms without regard to unique identification of individuals. That is, given observed counts $n_1, n_2, \ldots, n_T$ from a Binomial(N, p) distribution, then a simple method of moments estimator (MME), or the maximum likelihood estimator (MLE), may be used. However, the resulting estimators can be very unstable when the data suggest that p is low, which led Olkin et al. (1981) to consider refinements for these unstable situations. Carroll and Lombard (1985) suggest an alternative estimator (henceforth the "CL estimator") which involves integrating the nuisance parameter p from the likelihood under a beta prior distribution with parameters a and b which must be fixed a priori. Conceivably, the CL estimator can be applied to the type of data given in Table 1, either by application to the data at each site, or by pooling the data across sites. While such methods are interesting from a theoretical perspective, and would seem ideally suited for avian point count data, they have not been adopted in practice for several reasons which are elaborated on in Section 3.

In this article I develop a model for estimating abundance under a sampling design wherein temporally replicated counts are collected at a number of locations in space. Under this model, site-specific abundance is viewed as a random effect, and the marginal likelihood of the counts is obtained by integrating the binomial likelihood for the observed counts over possible values of abundance for each site. This notion is thus similar to other integrated likelihood approaches which address estimating abundance (N) from replicated binomial data (e.g., Carroll and Lombard, 1985) except that here the site-specific *abundance* parameters are removed from the likelihood by integration, instead of the detection probability parameter (p). Importantly, the parameters of the prior distribution are retained in the model and estimated from the data, thus facilitating a variety of useful modeling extensions.

The proposed "N-mixture" model for estimating abundance is described in Section 2. In Section 3, I briefly describe the Carroll and Lombard (1985) estimator as it might be applied to spatially replicated data, and its potential limitations. A simulation study to evaluate the proposed estimator, and a comparison with the CL estimator, are presented in Section 4. An application involving avian point count data is presented in Section 5.

2. A Model for Spatially Replicated Counts

The study design considered here is that in which animals are counted during $t = 1, 2, \ldots, T$ sampling occasions and at $i = 1, 2, \ldots, R$ locations (sites). Let n_{it} be the number of distinct individuals counted at location i, in time t. It is assumed that the population being sampled is closed with respect to mortality, recruitment, and movement so that the counts may be viewed as i.i.d. binomial random variables,

$$n_{it} \sim \text{Binomial}(N_i, p),$$

where N_i is the number of individuals available for sampling (i.e., the population size at location i) and p is the detection (or capture) probability. The likelihood for data from site i is

$$L(N_i, p \mid \{n_{i1}, \ldots, n_{iT}\}) = \prod_{t=1}^{T} \binom{N_i}{n_{it}} p^{n_{it}} (1-p)^{N_i - n_{it}}, \quad (1)$$

which is that considered by Olkin et al. (1981) and Carroll and Lombard (1985).

Replicate samples at R locations in space yield R site-specific likelihoods conditional on $\{N_1, N_2, \ldots, N_R\}$, and p. The joint likelihood is

$$L(\{N_i\}, p \mid \{n_{it}\}) = \prod_{i=1}^{R} \left\{ \prod_{t=1}^{T} \binom{N_i}{n_{it}} p^{n_{it}} (1-p)^{N_i - n_{it}} \right\}. \quad (2)$$

This likelihood contains $R + 1$ unknowns, the R abundance parameters, $\{N_i : i = 1, 2, \ldots, R\}$, and a single detection probability parameter, p. Notwithstanding numerical difficulty in carrying out this maximizing, the general instability problems motivating Olkin et al. (1981) and Carroll and Lombard (1985) are important here because of the nature of the data being considered (sparse counts) and because of the large number of abundance parameters (the N_i), some of which may in fact be zero.

2.1 *The Poisson Mixture Estimator*

Carroll and Lombard's (1985) solution to dealing with the instability problem was to integrate the conditional likelihood for p and N against a beta prior distribution for p. Clearly the rationale for integrating p from (2) is tenuous because it does little to simplify the problem, and the estimation of p is aided by the addition of the R replicates regardless of the sparsity of the actual counts.

The solution proposed here is to view the N_i as nuisance parameters, assumed to be independent random variables with density $f(N; \theta)$, and then integrate (1) over this prior distribution on abundance. One sensible choice of f is the Poisson density,

$$f(N; \lambda) = \frac{e^{-\lambda} \lambda^N}{N!}.$$

The Poisson assumption on N is appealing because it arises under the assumption that the distribution of individuals within some region is a homogeneous Poisson point process, which is a natural (and common) model for the distribution of organisms (see Section 2.2).

Downloaded from https://academic.oup.com/biometrics/article/60/1/108/7289296 by James Cummings user on 16 September 2024

Estimation can now focus on the integrated likelihood:

$$L(p,\theta \mid \{n_{it}\}) = \prod_{i=1}^{R}\left\{\sum_{N_i=\max_t n_{it}}^{\infty}\left(\prod_{t=1}^{T}\mathrm{Bin}(n_{it};N_i,p)\right)f(N_i;\theta)\right\}, \quad (3)$$

where θ is the (possibly vector-valued) parameter of $f(N;\theta)$ ($\theta = \lambda$ for the Poisson model) and $\mathrm{Bin}(n_{it};\, N_i,\, p)$ is the binomial likelihood.

It is a simple matter to maximize (3) numerically using conventional methods. In practice, the summation over N_i is restricted to a finite, but large, bound, say K. The impact of this decision can be evaluated by inspecting the likelihood for various values of K. In the examples of Section 5, $K = 200$ was used, but the resulting MLEs did not change for K greater than about 20.

2.2 *Prior Distributions on N*

The Poisson prior distribution on abundance is a natural choice because it arises under the assumption that animals are distributed randomly in space. While this is probably best justified within a homogeneous landscape, potential departures from randomness that arise as a result of landscape heterogeneity could be explained explicitly (using covariates) or by modifying the prior distribution to accommodate extra-Poisson variation.

In this regard, a natural prior distribution for abundance is the negative binomial, which results from a Gamma prior distribution on site-specific intensity parameters λ_i. That is,

$$f(N;\alpha,r) = \frac{\Gamma(N+\alpha)}{\Gamma(\alpha)N!}r^{\alpha}(1-r)^{N},$$

which is commonly parameterized in terms of the mean, $\mu = \alpha(1-r)/r$, or the Gamma parameters $(\alpha,\ \beta)$ with $r = \beta/(1+\beta)$. In the analyses of Section 5, I use the $(\alpha,\ \mu)$ parameterization of the negative binomial.

In practice, both models (and perhaps others) can be fit to the data at hand and one could engage in various model selection exercises in an attempt to choose the most appropriate one (see Section 5).

2.3 *Estimating Abundance*

Estimation based on (3) does not directly yield estimates of abundance. Instead, maximization of (3) yields an estimate of θ, the parameter(s) of the prior distribution on N, and p. From these, there are several approaches to generating estimates of actual abundance.

First, in cases where site-specific abundance is not of direct interest (which is probably in most situations), an estimate of *total abundance* can be generated by area expansion. This can proceed in several ways depending on the context. If sample units are of known area (e.g., a fixed-radius point count) then it seems reasonable to estimate the total abundance by a simple area expansion based on $\hat{\lambda}$ (under the Poisson assumption) since λ is the density *per sample unit*. In cases where the sample area is unknown, obtaining an estimate of total population size is difficult. However, $\hat{\lambda}$ is still an estimate of *mean abundance* per sample unit and it may thus serve as a useful measure of abundance, accounting for detection probability. Therefore, if there were R sample units then $\hat{N} = R\hat{\lambda}$ is an estimate of the total abundance of the *sampled* area. This would be useful in studies involving temporal changes, impacts of climate or weather, and other factors. Finally, if covariates are thought to impact abundance (Section 2.4), then an estimate of total abundance can be constructed by summing site-specific estimates of each λ_i assuming that the covariates are known (i.e., mapped) over the region of interest.

Second, an estimate of θ can be used in conjunction with Bayes Theorem to yield an estimate of the conditional (on θ) posterior distribution of any particular N_i:

$$\Pr(N=k \mid n_1,n_2,\ldots,n_T,\theta,p) = \frac{\Pr(n_1,n_2,\ldots,n_T \mid N=k,p)\Pr(N=k;\theta)}{\sum_{k=0}^{\infty}\Pr(n_1,n_2,\ldots,n_T \mid N=k,p)\Pr(N=k;\theta)}, \quad (4)$$

using estimates $\hat{\theta}$ and $\hat{p}$ in place of θ and p. Thus, given the observed vector of counts $n_{i1},\ n_{i2},\ldots,n_{iT}$, one may characterize the posterior distribution of N_i, which can be used to make inferences about particular values of N_i.

This approach to estimating N_i can be described as a "plug-in" empirical Bayes procedure; prior parameters are estimated from the marginal likelihood and these estimates are then plugged into the conditional (on θ) posterior distribution. In some instances, it may be prudent to adopt a fully Bayesian perspective here and integrate over θ to yield the (marginal) posterior distribution of N_i. Although this is not much more difficult, it would require suitable prior distributions dependent on the choice of $f(N;\ \theta)$, and additional computation. A fully Bayesian approach might be preferred in small sample situations to more properly characterize the uncertainty in the estimates due to estimation of the prior parameters.

2.4 *Modeling Covariate Effects*

One of the principle benefits of the proposed modeling framework is that the effect of factors which lead to spatial variation in abundance or detection probability can be estimated directly. This is important because many field studies are not interested in abundance per se, but rather how abundance changes in response to habitat and other landscape characteristics. In the context of the Poisson model, the obvious modification is to adopt a log-linear model on the prior mean according to:

$$\log(\lambda_i) = \sum_{j=1}^{r} x_{ij}\beta_j,$$

where $x_{ij};\ j = 1,\ 2,\ldots,r$ are the r measurable covariates for site i.

A similar idea can be employed for modeling covariate effects on detection probability. For this, a linear model for $\mathrm{logit}(p_{it})$ seems reasonable.

3. Carroll and Lombard's Estimator

Carroll and Lombard's (1985) proposal was to integrate the product binomial likelihood (1) over a prior distribution for p, with fixed parameters. In principle, the CL estimator may be applied to data from each site, $(n_{i1},\ n_{i2},\ldots,n_{iT})$, in order to estimate N_i. For sparse data, this simple application of the basic CL estimator often yields estimates on the boundary of

Downloaded from https://academic.oup.com/biometrics/article/60/1/108/7289296 by James Cummings user on 16 September 2024

the parameter space (the number of unique individuals observed). Also, all-zero count histories pose obvious difficulty (Table 1).

One solution in sparse data situations involving spatial replication is to focus on estimating the total abundance (across all sites) by simply summing the counts for each sample period across sites, yielding the totals: $n_{.t} = \sum_{i=1}^{R} n_{it}$. Then, the CL estimator can be applied to these pooled counts to get an estimate of total abundance that is analogous to that of Section 2.3 (that is, $R\hat{\lambda}$ under the Poisson model). I will refer to this as the "pooled CL estimator." Use of this estimator does generally yield stable solutions in most of the examples given in Section 5. Note that, under the likelihood given by (2), the summed counts ($n_{.t}$) are not sufficient for estimating total abundance. Thus, some information about model parameters is disregarded. In particular, summing counts across sites ignores the structure implied by the two-dimensional array (counts made at the same site are exchangeable when conditioned on N_i).

There are several important limitations with the integrated likelihood approach suggested by Carroll and Lombard (1985). First, that approach requires specification of the beta parameters a and b, which is a difficult decision regardless of whether or not the prior is ostensibly "noninformative." In practice, the detection probability of individual animals is an important consideration and, intuitively, any estimator of N will be sensitive to a priori structure imposed on p (this is demonstrated in Section 5). Indeed, the CL estimator is biased for any particular value of p, unless a judicious (or fortunate) choice of a and b can be made. For many species of birds p is more likely small (closer to 0) than large, but the CL estimator requires prior decision making about what constitutes reasonable values of p. Thus the problem of how to choose a and b in a reasonably objective manner is fundamental and cannot be avoided by apparently innocuous assumptions.

Second, although instability of the CL estimator in sparse data problems can be overcome by pooling data across sites, there is often fundamental interest in the spatial aspect of the problem, such as modeling covariate effects on abundance. Spatial aggregation of the data precludes any possibility of investigating these effects.

Finally, in most field studies, information which influences detectability is collected (e.g., observer, time of day, weather, etc.). The CL approach does not provide a framework for modeling covariates which influence p because of its removal from the problem as a "nuisance" parameter. The implications of this are important. For example, estimating population change across years, or regional differences, are common goals in ecological studies. Often the data yielding information about different (yearly, regional) values of N are collected by different observers. Use of the CL estimator with the same prior parameters is a de facto statement that all observers are equal, or that all data are the same with respect to the detection process.

4. Simulation Study

A simulation study was conducted to evaluate the proposed N-mixture estimator in relation to the CL estimator. For these simulations, $R = (20, 50)$ sample sites were used with Poisson site-specific abundances having $\lambda = (2, 5)$. Thus, the *actual* population sizes vary around 40, 100, 100, and 250 individuals according to a Poisson(Rλ) distribution. For these four scenarios, binomial data were generated with $p = 0.25$ and $p = 0.50$. To be concise, only simulations with $T = 5$ are summarized here. Results did not vary dramatically for most cases when T was varied between 3 and 10.

The CL estimator with Beta(1,1) and the improper Beta(0,0) prior distributions on p were considered. The three estimators will be indicated by $\hat{N}_{\text{pois}}$, $\hat{N}_{cl00}$, and $\hat{N}_{cl11}$ where $\hat{N}_{\text{pois}}$ is the estimator $R\hat{\lambda}$ suggested in Section 2.3. Coverage of the nominal 95% asymptotic confidence interval was computed as well as various summaries of the sampling distribution of each estimator. The models were fit using the nlm procedure in the software package **R** (Ihaka and Gentleman, 1996). The numerical solution appeared stable in all but a small number of the $R = 20$, $r = 0.25$ cases, as diagnosed by a poorly conditioned Hessian. The reason for this being that in those cases, most of the observed counts were zero.

4.1 *Simulation Results*

Simulation results are summarized in Table 2. Results for $\hat{N}_{cl00}$ are omitted because they were similar to those for $\hat{N}_{cl11}$. For a single case, ($\lambda = 2$), the centered sampling distributions for the three estimators are shown in Figure 1. Density plots are of the difference $\hat{N} - N$ for the three estimators so that zero on the x-axis represents no discrepancy between $\hat{N}$ and the actual value.

These results clearly illustrate several points regarding the various estimators. First, $\hat{N}_{\text{pois}}$ is usually biased, owing to its general skewness, but its median and mode are typically about right. For example, in the $p = 0.25$ cases, its distribution is strongly skewed but its mass is centered around zero. Conversely, there is a dramatic bias in $\hat{N}_{cl}$ in all cases regardless of which prior is used. Even when $p = 0.50$ $\hat{N}_{cl}$ remains negatively biased under either prior specification although its mass is distributed more closely around zero than in the $p = 0.25$ case. Thus, while the CL estimator may be superior in terms of variance in some instances (e.g., when $p = 0.25$), its excessive bias is troublesome. This bias is reflected in the confidence interval coverage; The coverage (for a 95% confidence interval) for the CL estimators is very low. For the Poisson mixture estimator the confidence interval coverage is much closer to the nominal level, although conservative in the $\lambda = 2$, $p = 0.50$ cases. Simulation results for $\hat{p}$ are omitted, but in all cases examined, $\hat{p}$ was nearly unbiased and symmetric about p.

5. Application to Avian Point Count Data

The data analyzed here consist of repeated point counts of birds from a North American BBS (Robbins, Bystrak, and Geissler, 1986) route. The BBS is a large-scale survey which has been conducted since the 1960s, and consists of >4000 routes. Each route has 50 sample sites located one-half mile apart. From each sample location, an observer records the number of each species detected (by sight or song). Historically, BBS data are used as count indices, and to measure spatial and temporal variation in relative abundance (e.g., trends; Link and Sauer, 1997), because no rigorous methodological framework for direct modeling of detection probability from such data has been suggested. In 1991 a study was conducted

Downloaded from https://academic.oup.com/biometrics/article/60/1/108/7289296 by James Cummings user on 16 September 2024

Table 2

Summaries of the sampling distribution of Poisson mixture (Pois) and Carroll and Lombard (CL; 1985*) estimators of total abundance based on* 10,000 *simulated data sets. CL estimator uses beta*$(1,1)$ *prior distribution for* p. $\bar{N}$ *is the average abundance of the simulated realizations,* q_1, q_3, *and RMSE are the first and third quartiles and root mean-squared error of the estimator. Coverage is the fraction of* 95% *"asymptotic" confidence intervals which contained the true value of* N.

R	λ	p	$\bar{N}$	Estimator	q_1	Mean	Median	q_3	Coverage	RMSE
20	2	0.25	40	Pois	32.33	48.64	40.77	54.17	0.89	30.02
				CL	14.64	19.82	18.51	23.55	0.18	21.43
20	2	0.50	40	Pois	35.14	40.87	40.15	45.70	0.99	5.43
				CL	26.45	33.58	32.00	39.02	0.65	10.62
50	2	0.25	100	Pois	86.48	106.85	100.32	118.38	0.91	31.34
				CL	37.28	48.93	45.96	57.23	0.17	53.30
50	2	0.50	100	Pois	92.13	100.63	99.77	108.45	0.99	7.60
				CL	67.35	83.76	79.73	95.97	0.64	26.37
20	5	0.25	100	Pois	80.01	118.24	100.95	135.28	0.86	60.92
				CL	37.51	49.04	46.30	57.72	0.19	53.48
20	5	0.50	100	Pois	89.90	102.64	99.75	112.21	0.93	16.39
				CL	67.52	83.67	79.98	96.01	0.65	26.48
50	5	0.25	250	Pois	217.35	263.71	251.21	297.32	0.88	64.93
				CL	94.59	122.42	115.28	142.22	0.18	132.95
50	5	0.50	250	Pois	233.41	252.16	249.61	268.59	0.93	22.34
				CL	170.51	209.47	200.24	238.91	0.64	65.94

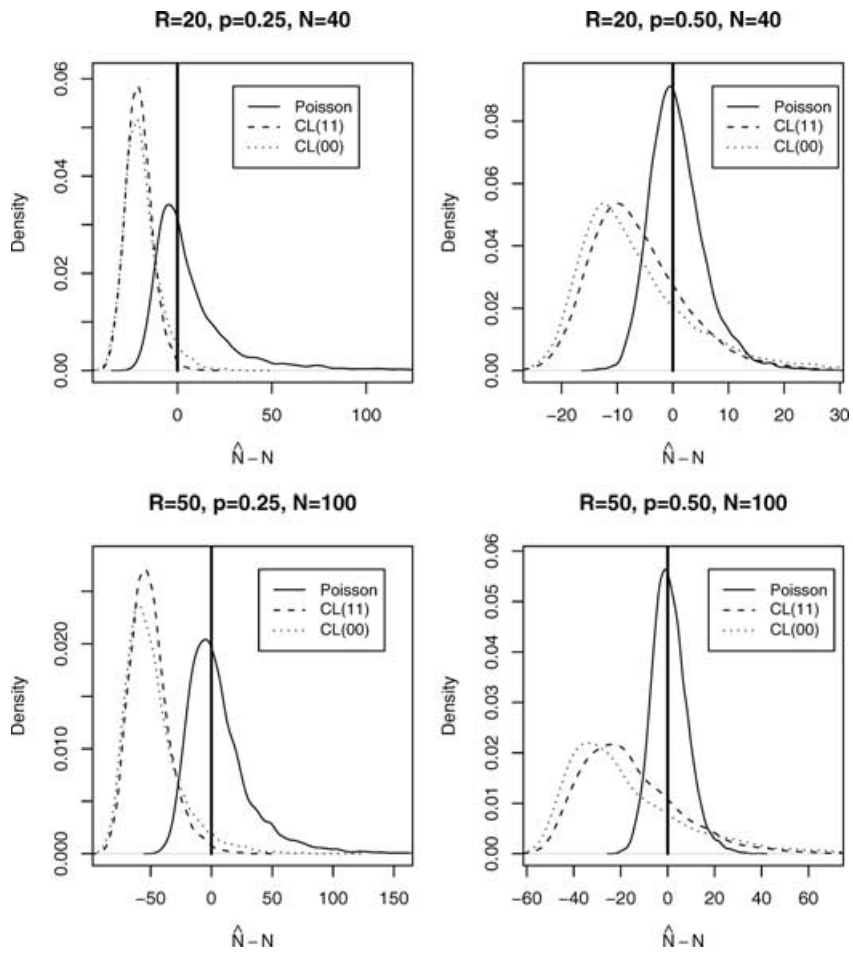

Figure 1. Centered sampling distributions ($\hat{N} - N$) for three estimators of abundance when $\lambda = 2$; Poisson mixture estimator (Poisson), Carroll and Lombard (1985) estimator using beta(1,1) (CL(1,1)) and beta(0,0) (CL(0,0)) prior distributions for p. R is the number of spatial replicates, p is detection probability.

Downloaded from https://academic.oup.com/biometrics/article/60/1/108/7289296 by James Cummings user on 16 September 2024

using BBS sampling protocol to examine variation in counts within a breeding season (Link et al., 1994). The data considered here are counts of six species from this study: American redstart (*Setophaga ruticilla*), ovenbird (*Seiurus aurocapillus*), gray catbird (*Dumetella carolinensis*), hermit thrush (*Catharus guttatus*), brown thrasher (*Toxostoma rufum*), and wood thrush (*Hylocichla mustelina*). Counts were carried out on 10 days during June 1991 by the same observer. Due to sampling protocol considerations, it is reasonable to assume that the population is closed in the sense that breeding birds have established territories so that observed birds are primarily local breeders. Consequently, the assumption that n_{it} for sample location $i = 1, 2, \ldots, 50$ and occasion $t = 1, 2, \ldots, 10$ are Binomial(N_i, p) seems plausible.

The data from the first five stops along the route containing 50 stops for the redstart were given in Table 1. These data are representative of that for the other species in the sense that they are sparse; counts are low, and there are many zeros.

5.1 *Carroll and Lombard Estimation*

Due to the sparse nature of the data (in particular, the all-zero count histories), the CL estimator does not perform well when applied to site-specific data, as one might expect. The majority of site-specific estimates of N were on the boundary (maximum number of unique individuals observed). However, it is possible to make progress using the CL estimator by pooling the data over the 50 stops as described in Section 3. In this way, an estimate of "total abundance" can be obtained. This was done for each of the six species and the CL estimator was used to estimate N using various values of a and b. The results are presented in Table 3.

These results illustrate two important problems. First, in those instances where the fewest number of individuals were seen (hermit thrush and thrasher), the CL estimator falls on the boundary of the parameter space unless substantial a priori mass on p is given to small values. Boundary estimates are suggestive of $p = 1$ which is highly suspicious given the secretive nature of these two species in particular. Sensitivity of $\hat{N}$ to the choice of prior parameters is also evident. The estimates behave roughly in accordance to the mean value of the imposed prior distribution and vary considerably (in relative terms) for even these minor changes in a and b. Experience suggests that many species of birds have low detection probabilities, which would favor use of something other than $a = b = 1$. But precisely which values of a and b should one choose? There is simply no basis for choosing an estimate of N without some information regarding detection probability. Thus, use of the CL estimator substitutes one problem (poor behavior of the MLE of N when the detection probability is low) for another problem—choosing a and b.

Table 4

Results of fitting Poisson and negative binomial mixture models to count data on six species of birds collected on 10 sampling occasions and at 50 spatial locations along a BBS route

	Poisson		Negative binomial				
Species	$\hat{\lambda}$	$\hat{p}$	$-2ll$	$\hat{\mu}$	$\hat{p}$	$-2ll$	Δll
Redstart	2.81	0.108	679.61	7.21	0.042	677.49	2.1
Ovenbird	2.12	0.317	900.70	2.63	0.254	893.91	6.8
Catbird	0.63	0.158	299.81	0.84	0.118	298.66	1.2
Hermit thrush	0.61	0.075	175.77	7.06	0.007	174.63	1.1
Thrasher	0.09	0.228	68.42	0.12	0.169	65.18	3.2
Wood thrush	1.76	0.262	762.13	1.90	0.243	761.48	0.7

5.2 *Estimation Using N-Mixture Models*

Results of fitting the Poisson and negative binomial mixture to the bird count data are summarized in Table 4. Because the Poisson model represents a special case of the negative binomial model, it must yield a better fit (the minimized value of minus twice the log likelihood for both models is shown in Table 4). Evaluation of the utility of the additional negative binomial parameter may be carried out by comparing the log-likelihood difference to a χ^2 distribution on one degree-of-freedom (Δll in Table 4). We see only a single instance (ovenbird) where the negative binomial is favored. In this instance, the results are not substantially different; The Poisson mean is 2.12, whereas the estimated negative binomial mean is slightly larger at 2.63. For some of the other cases (notably the redstart and hermit thrush), the negative binomial mean is unrealistically large, suggesting very high local abundance. However, in these instances, the overdispersed negative binomial model is not favored.

To generate abundance estimates analogous to those of Table 3 computed using the pooled Caroll and Lombard estimator, I used the "area expansion" estimator described in Section 2.3. Because data were collected at 50 sites and the model fits provide estimates of the mean number per site, an estimate of the total abundance is $\hat{N} = 50\hat{\lambda}$ (under the Poisson model). An estimate of the standard error

Table 3

Total abundance estimates for six bird species using different Carroll and Lombard estimators. max $n_{\cdot t}$ is the maximum count of birds over the 10 sample periods. $B(\cdot,\cdot)$ indicates the beta prior used in computing the estimate.

Species	$\max n_{\cdot t}$	B(0, 0)	B(1, 1)	B(1, 2)	B(2, 1)	B(1, 3)	B(1, 4)
Redstart	26	161.4	103.0	109.46	78.61	115.97	122.60
Ovenbird	41	51.8	53.5	58.62	51.49	65.78	75.24
Catbird	9	24.2	18.8	20.65	15.50	22.62	24.68
Hermit thrush	4	4.0	4.0	4.27	4.00	4.90	5.69
Thrasher	2	2.0	2.0	2.00	2.00	2.14	2.51
Wood thrush	30	64.4	58.9	66.03	51.67	74.02	82.75

Downloaded from https://academic.oup.com/biometrics/article/60/1/108/7289296 by James Cummings user on 16 September 2024

is $\mathrm{SE}(\hat{N}) = 50\mathrm{SE}(\hat{\lambda})$. Using the asymptotic variance (the inverse of the Hessian evaluated at the MLEs) for $\hat{\lambda}$ in each case, the total abundance estimates for the six species (and standard error) are $\hat{N}_{\mathrm{red}} = 140.5(26.4), \hat{N}_{\mathrm{oven}} = 106.0(9.2), \hat{N}_{\mathrm{cat}} = 31.5(5.7)$, $\hat{N}_{\mathrm{herm}} = 30.5(9.5)$, $\hat{N}_{\mathrm{thra}} = 4.5(1.6)$, and $\hat{N}_{\mathrm{wood}} = 88.0(8.9)$. All of these estimates are considerably greater than those based on $\hat{N}_{cl11}$, consistent with the apparent detection probabilities tending to be <0.5. As expected, as the beta prior parameters are chosen to favor lower values of p, the estimates become more consistent with those based on the Poisson mixture model.

6. Discussion

Sparse data from spatially replicated count surveys can be utilized to effectively estimate population sizes while properly accounting for the detection process when local (site-specific) abundance, N_i, can be modeled as exchangeable random variables. Under this approach, the characteristics of the latent distribution of N_i can be estimated from the integrated likelihood. Because the site-specific N_i are regarded as nuisance parameters and integrated from the likelihood, the proposed method does not yield direct estimates of abundance. Instead, estimates of quantities which are relevant to the *distribution of abundances* across sites (i.e., the parameters of the prior distribution of N_i), such as average abundance in the case of the Poisson prior, may be estimated. Although this is probably sufficient for many objectives (for example, in evaluating temporal change or geographic differences), it is possible to obtain an estimate of total abundance (e.g., area expansion), or even to estimate site-specific abundance from the estimated posterior distribution of N_i.

Imposing a prior distribution on abundance should not be perceived as a drawback of the proposed approach as there is flexibility in choice of prior distribution and decision making in this regard can be viewed as a model selection exercise. I considered the Poisson and negative binomial here because they are sensible within the context of the distribution of organisms in space, though other prior distributions may be considered.

The Carroll and Lombard (1985) approach requires specification of a prior distribution on detection probability, where prior parameters are not estimated from the data. One important disadvantage of this is that the estimator is sensitive to choice of prior parameters (e.g., see Table 3). Essentially, that approach resolves one problem (sensitivity to data) while inducing another (sensitivity to prior) which may be equally problematic. Another consequence of having fixed prior parameters for detection probability is that the estimator is badly biased. Although bias should not be the only consideration in evaluating the merits of a procedure, extreme bias would seem to mitigate the value of any gain in precision that biased estimators might yield.

6.1 *Discussion of Simulation Results*

Results of the simulation study suggest that the proposed N-mixture estimator performs well under a range of situations that are relevant to, for example, point counts of birds. The situations evaluated constitute what would generally be viewed as low to moderate abundance: 20 or 50 sites with an average of two or five individuals. Although the RMSE of the estimator was not always better than the CL estimators considered, its sampling distribution is more or less centered properly and its performance in terms of RMSE and coverage is compelling when viewed in light of the dramatic bias and poor coverage of the CL estimators. Interestingly, the bias of the CL estimator was substantial even when the prior used in construction of the CL estimator was consistent with the true value of p. Simulations (not reported here) indicate that the bias does diminish to tolerable levels as T increases, or as abundance increases.

The simulation studies could be criticized on the grounds that the data construction (i.e., spatial replication) favors the Poisson mixture model for estimating abundance as that model explicitly acknowledges the spatial replication. This is undeniable; the simulations were constructed to *evaluate* the performance of the Poisson mixture model in such problems. The pooled version of the CL estimator was used because the CL estimator cannot be applied effectively to site-specific sparse data of the type considered in this article. I attempted to resolve this issue by developing a modified CL estimator that properly accounts for the spatial replication. This estimator is that resulting from integration of both p and N from the conditional likelihood, while still retaining the fixed parameters in the prior distribution on p, but estimating the prior parameters of N. Generally, its performance was intermediate between the pooled CL estimator and the Poisson mixture estimator obtained by only mixing over the prior on N, but bias problems were still persistent as a consequence of retaining the fixed beta prior parameters.

6.2 *The Application*

Estimates of abundance were computed for avian point count data on six species of birds. These data typify those collected in many studies of avian biology and by many monitoring programs. Estimates of total abundance for the six species were sensitive to the choice of prior parameters in the CL estimator. This induces some arbitrariness in the choice of *which* estimate of abundance should be used and would likely be an important consideration in many analyses. For the N-mixture models, the negative binomial model was favored over the Poisson for only a single species (based on log likelihood). For that species, the estimates of mean abundance differed only slightly. However, for two of the remaining species the estimated mean abundance was substantially different under the two models. For example, the estimated mean abundance for the Hermit thrush was 0.61 under the Poisson model and 7.06 under the negative binomial, suggesting very high abundance at some sites. Thus, even though Poisson was favored, the implausibility of such high levels of abundance at the spatial scale of a point count suggests the possibility that neither model provides an adequate description of the data. For example, it could be the lurking covariates are inducing some heterogeneity in the data, but in a manner that is inconsistent with that implied by the negative binomial model. Hermit thrush is a forest bird, and so forest coverage at the stop level would be an interesting covariate to consider, if it could be obtained.

Downloaded from https://academic.oup.com/biometrics/article/60/1/108/7289296 by James Cummings user on 16 September 2024

Acknowledgements

The author thanks the associate editor and two anonymous referees for many helpful comments and suggestions on drafts of this manuscript.

Résumé

La réplication spatiale est un thème usuel dans les recensements animaux. De telles enquêtes fournissent souvent des données de comptage éparses, à partir desquelles il est difficile d'estimer la taille de la population, bien que prenant en compte formellement la probabilité de détection. Dans ce papier, je décris une classe de modèles (modèles de *N*-mélange) qui vise à estimer la taille de la population à partir de telles données. L'idée centrale est de considérer les tailles de population par site, N, comme des variables aléatoires indépendantes distribuées selon une certaine loi de mélange (par exemple de Poisson). Les paramètres *a priori* sont estimés à partir de la vraisemblance marginale des données, après intégration sur la distribution *a priori* de N. Carroll et Lombard (1985) ont proposé une classe d'estimateurs basés sur un mélange par une distribution *a priori* de la probabilité de détection. Leur estimation s'applique dans un cadre limité, et est sensible aux valeurs des paramètres fixés en *a priori*. La réplication spatiale fournit une information additionnelle concernant les paramètres de la distribution *a priori* de N, exploitée par les modèles de *N*-mélange, et qui conduit à des estimations raisonnables de l'abondance pour des données éparses. Une étude de simulation montre l'avantage des caractéristiques (biais, couverture de l'intervalle de confiance) de l'estimateur du *N*-mélange sur l'estimateur de Carroll et Lombard. Les deux estimateurs sont appliqués à des données de comptage pour six espèces d'oiseaux, illustrant la sensibilité au choix de l'*a priori* sur p, avec en conséquence des estimations substantiellement différentes pour l'abondance.

References

Carroll, R. J. and Lombard, F. (1985). A note on N estimators for the binomial distribution. *Journal of the American Statistical Association* **80,** 423–426.

Haldane, J. B. S. (1942). The fitting of binomial distributions. *Annals of Eugenics* **11,** 181–187.

Ihaka, R. and Gentleman, R. (1996). R: A language for data analysis and graphics. *Journal of Computational and Graphical Statistics* **5,** 299–314.

Link, W. A. and Sauer, J. R. (1997). Estimation of population trajectories from count data. *Biometrics* **53,** 63–72.

Link, W. A., Barker, R. J., Sauer, J. R., and Droege, S. (1994). Within-site variability in surveys of wildlife populations. *Ecology* **74,** 1097–1108.

Olkin, I., Petkau, A. J., and Zidek, J. V. (1981). A comparison of N estimators for the binomial distribution. *Journal of the American Statistical Association* **76,** 637–642.

Pollock, K. H. (2000). Capture-recapture models. *Journal of the American Statistical Association* **95,** 293–296.

Robbins, C. S., Bystrak, D. A., and Geissler, P. H. (1986). *The Breeding Bird Survey: Its First Fifteen Years* (1965–1979), Volume 157. Washington, D.C.: USDOI, Fish and Wildlife Service Resource Publication.

Seber, G. A. F. (1982). *The Estimation of Animal Abundance and Related Parameters,* 2nd edition. London: Charles Griffin.

Williams, B. K., Nichols, J. D., and Conroy, M. J. (2002). *Analysis and Management of Animal Populations.* San Diego: Academic Press.

Received April 2002. *Revised October* 2002.
Accepted July 2003.

Downloaded from https://academic.oup.com/biometrics/article/60/1/108/7289296 by James Cummings user on 16 September 2024