



Research Papers

A Field Evaluation of the Time-of-Detection Method to Estimate Population Size and Density for Aural Avian Point Counts

Évaluation sur le terrain de la méthode fondée sur le temps de détection pour estimer l'effectif et la densité des populations d'oiseaux à partir de points d'écoute

*Mathew W. Alldredge*¹, *Theodore R. Simons*², *Kenneth H. Pollock*³, and *Krishna Pacifici*³

ABSTRACT. The time-of-detection method for aural avian point counts is a new method of estimating abundance, allowing for uncertain probability of detection. The method has been specifically designed to allow for variation in singing rates of birds. It involves dividing the time interval of the point count into several subintervals and recording the detection history of the subintervals when each bird sings. The method can be viewed as generating data equivalent to closed capture–recapture information. The method is different from the distance and multiple-observer methods in that it is not required that all the birds sing during the point count. As this method is new and there is some concern as to how well individual birds can be followed, we carried out a field test of the method using simulated known populations of singing birds, using a laptop computer to send signals to audio stations distributed around a point. The system mimics actual aural avian point counts, but also allows us to know the size and spatial distribution of the populations we are sampling. Fifty 8-min point counts (broken into four 2-min intervals) using eight species of birds were simulated. Singing rate of an individual bird of a species was simulated following a Markovian process (singing bouts followed by periods of silence), which we felt was more realistic than a truly random process. The main emphasis of our paper is to compare results from species singing at (high and low) homogenous rates per interval with those singing at (high and low) heterogeneous rates. Population size was estimated accurately for the species simulated, with a high homogeneous probability of singing. Populations of simulated species with lower but homogeneous singing probabilities were somewhat underestimated. Populations of species simulated with heterogeneous singing probabilities were substantially underestimated. Underestimation was caused by both the very low detection probabilities of all distant individuals and by individuals with low singing rates also having very low detection probabilities.

RÉSUMÉ. La méthode fondée sur le temps de détection utilisée dans le contexte des points d'écoute représente une nouvelle approche pour estimer l'abondance des oiseaux en tenant compte de la probabilité incertaine de détection. Cette méthode a été spécialement conçue pour tenir compte du taux variable de chant observé chez les oiseaux. Elle consiste à diviser l'intervalle de temps passé à un point d'écoute en sous-intervalles et de noter l'historique de détection des sous-intervalles où chaque individu est détecté. Les données obtenues par cette méthode peuvent être considérées comme étant équivalentes à celles obtenues par la capture-recapture dans une population fermée. La méthode diffère des approches fondées sur la distance et les observateurs multiples du fait qu'elle n'exige pas que tous les individus chantent durant le point d'écoute. Puisque cette méthode est nouvelle et qu'il existe une incertitude quant à la capacité de suivre les individus, nous avons effectué une évaluation sur le terrain de sa précision en utilisant des simulations sur des populations connues d'oiseaux chanteurs à l'aide d'un ordinateur portable qui envoyait des signaux à des stations audio distribuées autour du point d'écoute. Ce système imite un véritable point

¹Colorado Division of Wildlife, ²Cooperative Fish and Wildlife Research Unit, ³Dept of Zoology, North Carolina State University



Sponsored by the Society of
Canadian Ornithologists and
Bird Studies Canada

Parrainée par la Société des
ornithologistes du Canada et
Études d'oiseaux Canada



BIRD STUDIES
ÉTUDES D'OISEAUX CANADA

d'écoute tout en permettant de connaître l'effectif et la distribution des populations échantillonnées. Cinquante points d'écoute de 8 min (séparés en intervalles de 2 min) ont été simulés pour huit espèces d'oiseaux. Le taux de vocalisation de chaque individu a été simulé à l'aide d'une chaîne de Markov (périodes de chant suivies de périodes de silence), ce qui nous semblait plus réaliste qu'un processus purement aléatoire. L'objectif principal de notre article était de comparer les résultats pour des espèces chantant à une fréquence homogène (élevée ou faible) par intervalle avec d'autres espèces chantant à des fréquences hétérogènes (élevées ou faibles). L'effectif de la population a été estimé précisément pour les espèces simulées ayant une fréquence de vocalisation homogène et élevée. L'effectif des populations des espèces simulées présentant une fréquence de vocalisation faible mais homogène a été légèrement sous-estimé. Dans les cas des espèces chantant à des fréquences hétérogènes, les populations étaient fortement sous-estimées. La sous-estimation était causée à la fois par la faible probabilité de détection de tous les individus éloignés et des individus chantant à des fréquences faibles.

Key Words: *Aural detections; availability process; avian point counts; detection probability; field tests; perception process; time-of-detection method*

INTRODUCTION

Bart (2005) estimated that there are at least 1000 independent programs that gather long-term data on bird abundance in the USA and Canada, and many are based on point counts. Point counts are widely used to monitor spatial and temporal patterns of bird abundance, to assess species–habitat relationships, to evaluate the response of populations to environmental change or management, and to estimate species diversity. Surveys of breeding birds rely heavily on auditory detections, which can range from 70% of observations in suburban landscapes to 94% of observations in closed-canopy deciduous forest (Simons et al. 2007). Hundreds of thousands of point counts are conducted annually in North America across a spectrum of scales, from short-term site-specific studies, to long-term continental-scale surveys, such as the Breeding Bird Survey (Sauer et al. 2005).

In fixed radius point transects, where all the birds are detected,

$$\hat{D} = N/a = n/k\pi w^2 \quad (1)$$

where D denotes density, N is the number of birds in the sampled area, and n is the number of birds detected. Note that w is the radius of the circle around the point, and the area surveyed is, therefore, $a = k\pi w^2$ if there are k points surveyed. Of course this is totally unrealistic in practice, as it will

virtually never be possible to detect all birds within the sampled area. When not all the animals are detected, the result generalizes to:

$$\hat{D} = n/a\hat{p} = n/k\pi w^2 \hat{p} \quad (2)$$

where p is the probability of an animal being detected in the circle and has to be estimated in some manner.

There are several approaches to the estimation of p , the probability of detection, in the literature, and they all make different assumptions about the detection process. Distance sampling (Buckland et al. 2001) requires that: points are chosen randomly, detection is certain at the point, detection is a decreasing function of distance, and that there is no movement before detection. A second method, the multiple independent observer approach, requires that several observers detect birds simultaneously on the same sample area and map their locations. Estimation then is based on a closed capture–recapture modeling approach for the detection histories; Alldredge et al. (2006) describes the assumptions in detail. Key issues are no matching errors and no movement before detection. Nichols et al. (2000) suggested a dependent double-observer variation of this method. Another method is the repeated count method (Royle and Nichols 2003, Kery et al. 2005). The focus of our paper is a new approach, the time-of-detection method of

Farnsworth et al. (2002) and Alldredge et al. (2007a).

The overall probability of detection of an individual animal is made up of an availability process (the probability of an animal being available for detection), and a detection process (the probability of an animal being detected, given that the animal is available for detection) (see, e.g., Marsh and Sinclair 1989, Pollock et al. 2004). Here, we focus on aural point counts, and symbolically, we represent this detection process by the following equation: $p = p_a p_d$, where p is the overall probability of detecting a bird that is present in the sampled circle during the sampling period, p_a is the probability that such a bird is available for detection (this is the probability that the bird sang if the detections are aural), and p_d is the conditional probability of detection given that the bird sang. For the distance and multiple observer methods, p_a is assumed to be 1 (or in other words we condition our inference on the birds that have sung).

The time-of-detection method, by contrast, estimates the overall detection probability and allows for birds to be unavailable because they do not sing. Farnsworth et al. (2002) were the first to realize that there is information on detection probability available from the times when birds are detected in a point count. They developed a “removal” method using only the time interval when individual birds are first detected. The approach is related to the removal method for trapped animals (Zippin 1958, Seber 1982, Williams et al. 2002). Later Alldredge et al. (2007a) suggested a more efficient approach that applies a k sample closed capture–recapture model to the detection histories. With this approach, that is based on tracking individual birds accurately over the whole point count, the time intervals in which a particular bird was detected and the time intervals in which that same bird was not detected are recorded for the set of k time intervals. For example, if there were $k = 4$ time periods then a detection history of {1101} would indicate that a bird was detected in periods 1, 2, and 4, but not in period 3. Capture–recapture approaches are much better than removal methods at modeling heterogeneity of detection probabilities. Alldredge et al. (2007a) implemented this capture–recapture approach using field data to illustrate the method’s strengths and weaknesses. In the rest of the article, we will be considering this approach when we refer to the time-of-detection method.

Following Alldredge et al. (2007a), the basic model assumptions of the time-of-detection method are as follows:

1. There is no change in the population of birds within the detection radius during the point count (i.e., the population is closed and birds do not move in or out);
2. There is no double counting of individuals (i.e., the observer keeps track of individual birds without error);
3. Observers accurately assign birds to within or beyond the radius used for a fixed-radius circle if one is used;
4. All individual birds of a species have a constant per minute probability of being detected in each interval.

These assumptions are not trivial because there may be movement in or out of the area, which violates assumption 1. Movement of individuals within the area during the count may cause violations of assumption 2, and difficulty in assigning distances to birds located aurally (Alldredge et al. 2007c) may cause violations of assumption 3. In addition, we know very little about the variability of singing rates in individual birds (assumption 4), but it appears to be influenced by a variety of factors (Gibbs and Wenny 1993).

Extensions to allow trap response and heterogeneity of capture probabilities in closed capture–recapture models can be advantageously adapted to this setting. If the probability of detection changes after the first detection (analogous to trap response in a capture–recapture setting), then assumption 4 can be weakened. “Trap response” models may be useful, and in this application, recapture probabilities are likely to be greater than first-capture probabilities. An observer may anticipate that an individual bird may call again, and thus it would be more likely to be detected if it does call (Alldredge et al. 2007a). If the probability of detection varies among individual birds, then heterogeneity models may also be useful, and Alldredge et al. (2007a) used heterogeneity in both their examples. Heterogeneity is likely because of variation in singing rates among individuals (Brewster 2007), as well as distance from the observer (Buckland et al. 2001) and many other factors. Much has been written about heterogeneity

models in the capture–recapture literature (Burnham and Overton 1978, Otis et al. 1978, Pollock et al. 1990, Williams et al. 2002). Link (2003) has noted problems with identifiability when these models are used. Modeling heterogeneity in detection probability using covariates can reduce issues associated with identifiability (Huggins 1989, 1991, Alho 1990). One covariate of particular importance is distance from the observer to the bird.

Birds with detection probabilities near 0 may be of special importance in time-of-detection applications. Alldredge (2004) and Brewster (2007) emphasize that avian singing rates may vary widely among individuals of the same species due to pairing status and other factors related to nesting phenology. Brewster (2007) estimated singing rates by following individual Ovenbirds (*Seiurus aurocapillus*) and Black-throated Blue Warblers (*Dendroica caerulescens*) for long periods. Some individuals had very low singing rates, which made them almost impossible to detect.

Our field evaluation uses the methods developed by Simons et al. (2007) consisting of an experimental system for simulating avian point count conditions when birds are detected aurally. The system uses a laptop computer to control a set of audio stations (amplified MP3 players) placed at known locations around a census point. The system can realistically simulate a known population of songbirds under a range of factors that affect aural detection probabilities. Earlier field tests using this system focused on factors influencing detection probability (Alldredge et al. 2007b), and measurement error in detection distance (Alldredge et al. 2007c).

The objective of this paper is to report on a field evaluation of the time-of-detection method for estimating aural detection probabilities and population sizes using our experimental system. We focus on comparing species with homogeneous singing rates vs. those with varying degrees and types of heterogeneous singing rates. After a detailed description of our research methods and results, we focus on the implications and importance of our work to field ornithologists designing point-count studies and suggest possibilities for future research.

METHODS

Field Methods

A bird song simulation system (Simons et al. 2007) was used to simulate point counts in a field setting in order to evaluate the time-of-detection point-count method. Thirty-five players were uniformly distributed with respect to area surrounding a single point in a mixed pine–hardwood forest at Howell Woods Environmental Science Center in the Piedmont Region of North Carolina. The forest has a dense understory that limits visibility to 30 m or less in almost all directions. All players were set 1 m above ground because, at the same site, previous experiments showed little effect of player height on detection probability (Alldredge et al. 2007b, Alldredge et al. 2007c); therefore, for simplicity, we chose to eliminate this variable from our experiments although it may be important in other forested habitats. Players were set at radial distances between 0 and 120 m, again based on the results of earlier experiments. For simplicity, single examples of each species' song were used for all our field tests including this one. Typical songs from Walton and Lawson (1999) were converted from the audio CD to 128 KPS, 44, 100 HZ, MP3 Format. Songs for all species were played at a volume of approximately 90 dB at a distance of 1 m.

Fifty point counts, each of 8 min duration were simulated over 2 days in early March 2006. Therefore, the conditions were those of no leaves on the deciduous trees on the plot. This time of year was chosen to minimize interference from resident bird calls. The same individual point was used for all counts, but the play lists of songs varied as described below. Each count was broken into four time intervals each of 2 min duration, and four experienced observers recorded birds using multi-colored pens to distinguish intervals of detection. Detection of a previously recorded bird in subsequent intervals was recorded by circling the previous detection in the appropriate color of the interval. To simplify their task of individually tracking each bird detected, observers were not required to estimate distance to the sound source of each bird, but they were asked to map birds to their approximate location.

There were a total of 18 singing birds per point count, and these included two individuals of eight species of primary interest (Acadian Flycatcher (ACFL), *Empidonax virescens*), Black and White

Warbler (BAWW, *Mniotilta varia*), Black-throated Blue Warbler (BTBW, *Dendroica caerulescens*), Black-throated Green Warbler (BTNW, *D. virens*), Hooded Warbler (HOWA, *Wilsonia citrina*), Scarlet Tanager (SCTA, *Piranga olivacea*), Ovenbird (OVEN, *Seiurus aurocapilla*), and Yellow-throated Warbler (YTWA, *D. dominica*). Four additional species were used to diversify the species list. None of the simulated species were found locally at the study area during the time of our experiments. For each of the eight primary species, there were exactly two individuals on each count, for a total of 100 individual singing birds on the 50 counts. We maintained a minimum 45° separation between the two individuals of the same species to minimize matching errors when recording observer results. For each individual, all calls came from only one speaker. That is, once assigned to play from a particular speaker, then all subsequent calls of that individual came from the same speaker so that “movement” of particular birds was not allowed.

A crucial point (described in detail below) is that true population sizes for some species were greater than 100 birds, because not all birds in the simulated population actually sang during the count, and hence were unavailable during the 8-min count. All birds had a positive probability of singing in each interval, although for some individuals of some species these probabilities were small.

Previous studies have focused on how various factors affect detection probabilities on auditory point counts (Simons et al. 2007, Alldredge et al. 2007b, c). Our main focus in this field test was to examine the singing process and potential biases that may exist with the time-of-detection method. The singing rate of individual birds was simulated following a Markovian process. We felt that singing bouts followed by periods of silence were more realistic than a truly random process (Collins 2004). The process is represented by two parameters; the probability that a bird does not sing in interval i , given that it did not sing in interval $i-1$ (γ'), and the probability that a bird does not sing in interval i , given that it did sing in interval $i-1$ (γ'').

Six of the eight species of interest were simulated under this Markovian process, and the true population size was manipulated so that the number of birds actually played during the experiment for each species was 100. The first three species considered (HOWA, BTNW, and YTWA) were all

simulated with a constant or homogeneous probability of singing at least once in an interval. Their probabilities of singing in at least one of the four 2-min intervals were 1.0, 0.8, and 0.6, respectively (Table 1). Total simulated populations of these three species were 100, 125, and 167, respectively. Singing rates during intervals in which these birds sang were six to nine songs per minute based on Robbins et al. (1983). The fourth species, BAWW, had an overall homogeneous probability of singing at least once in the total count of 0.8, but singing rates varied within the intervals in which the birds sang. Half of the simulated BAWW sang two songs per minute and the other half sang eight songs per minute. The total population of simulated BAWW was 125. The fifth and sixth species ACFL and SCTA, were simulated with heterogeneous probability of singing at least once in an interval. The overall populations of ACFL and SCTA were 133 birds. Fifty individuals were simulated with a probability of 1.0 of singing during the count, and 83 individuals were simulated with a probability of 0.6 of singing during the count. Singing rates for intervals in which a bird sang were six to nine songs per minute for these two species.

The probability of singing per interval for the two other species of interest, BTBW and OVEN, was based on empirical field data, where singing patterns of individual birds were recorded over 10- to 30-min intervals (Brewster 2007). Eight-minute long segments of these data were randomly selected to create simulated populations of birds (Table 1). The total potential population of BTBW was 133 birds, and the total potential population of OVEN was 127. Singing rates in intervals where birds sang were six to nine songs per minute. We emphasize that for each of the eight species, 100 birds actually sang during the 8-min count.

Statistical Methods

We calculated double-counting errors, errors from calls being assigned to the wrong time interval, and errors due to species being misidentified.

For each individual bird detected, a detection history over the four intervals was constructed using the standard capture–recapture format used by program MARK (White and Burnham 1999). For example, if a bird was only detected in the first interval, the detection history was {1000}, whereas if a bird was detected in intervals 1, 2, and 4 but not 3, its

Table 1. Experimental conditions for each of the eight species of interest. The first four species were homogeneous in population structure, in that all birds had the same probability of calling in each interval, whereas the next four species were heterogeneous. All species except the BAWW had a high singing rate per interval. Total and Singing Population sizes are presented as the probability of singing during the 8-min count. The last two columns summarize the Markovian nature of the singing process and are explained in detail in the text.

Species	Homogeneous population	Singing rate, if sings	Total population	Singing population	Probability sings in 8- min count	γ	γ'
HOWA	Yes	High	100	100	1.0	0.3	0.1
BTNW	Yes	High	125	100	0.8	0.83	0.45
YTWA	Yes	High	167	100	0.6	0.9	0.5
BAWW	Yes	Low/ High	125	100	0.8	0.83	0.45
ACFL	No	High	50 83	100	1.0 0.6	0.3 0.9	0.1 0.5
SCTA	No	High	50 83	100	1.0 0.6	0.3 0.9	0.1 0.5
BTBW	No	High	133	100	empirical	na	na
OVEN	No	High	127	100	empirical	na	na

detection history would be {1101}. Analyses were on a species-by-species basis for each observer and followed the methods outlined by Alldredge et al. (2007a).

We used the “Huggins Closed Captures” estimation method (Huggins 1989, 1991) as implemented in program MARK (White and Burnham 1999). Models included constant detection probability for all individuals (M_0), time effects on detection probability (M_T), and differences in detection probability due to previous detection (M_p). Unobservable heterogeneity (M_h) models were also fitted. We used the two-point mixture models (Pledger 2000). This is a very simple way of introducing heterogeneity where there is a proportion of the animals with one capture probability and the remainder of the animals in the population have another capture probability. (For example, 40% of the animals might have a capture probability of 0.3, and the remaining 60% of the

animals might have a capture probability of 0.8). Akaike's information criterion (AIC_c) model selection, corrected for small sample size, was used to determine the most likely model given the data (Burnham and Anderson 2002). Maximum likelihood estimates were then computed for selected models.

We carried out additional analyses for ACFL and SCTA where we modeled heterogeneity in detection probabilities as an observable group effect (M_g) in combination with other time or behavioral effects. However, this analysis would not be possible with actual point-count data because these groups are not observable in wild birds.

RESULTS

Double Counting and Interval Recording Errors

Double counting, counting a single bird as more than one bird, was a significant source of error among the four experienced observers. This error is produced when observers are unable to accurately localize the source of a song. Double-counting rates ranged from 0.9% to 3.4% (SE = 0.6%) of total observations among observers. Double counting occurred in a variety of forms. In most cases, single birds were recorded as two birds throughout the count interval. This created two or more capture histories that clearly indicated a single individual tracked as two birds throughout the count. Occasionally, observers mapped a bird in one location at the beginning of the count, then mapped the same bird in a new location and continued to track it at the new location for the remainder of the count. These cases produced two or more capture histories of the form xx00 for the original bird and 00xx for the double count, where x could be either a one or a zero.

Overall, 2.0%–4.1% (SE = 0.43%) of observations were recorded in the wrong time interval among the four observers. This occurred when two or more individuals of the same species sang during a count, and observers attributed a song to the wrong individual. Misidentification errors were rare, ranging from 0.1% to 0.6% (SE = 0.09%) of total observations among the four experienced observers.

Species with Simulated Homogeneous Probabilities of Singing

Model M_h was selected for all four HOWA data sets, which were simulated with a homogeneous singing probability with all birds singing in at least one interval. The AIC_c importance weights were high (Table 2) for heterogeneity models, indicating that some unobservable source of heterogeneity was affecting the detection probabilities for this species. Estimates of HOWA population size were accurate and precise for all four observers.

The other two species with homogeneous singing probabilities (0.8 and 0.6, respectively) and constant singing rates for singing birds were BTNW and YTWA. Model M_0 was selected or was a competitive model ($\Delta AIC_c < 1.0$) for BTNW and

YTWA (Table 2). Population size estimates for BTNW for the four observers (95, 97, 87, and 98) were substantial underestimates of the simulated population of 125 birds. Model M_b was selected as the top YTWA model for three of the four observers. Models accounting for heterogeneity in detection probabilities had little support for these data. Model M_b gave reasonably accurate estimates of YTWA for two of the observers and overestimates for the other, but standard errors were very large. Model M_0 was a competitive model ($\Delta AIC_c \leq 0.46$) for these three data sets, and although it gave precise estimates, it substantially underestimated the actual population size of 167 birds (Observer 1, population estimate 111, SE = 9.35; Observer 3, population estimate 121, SE = 12.92; Observer 4, population estimate 102, SE = 8.91). M_0 was chosen as the top model for Observer 2, where the population estimate is 123 with SE = 12.13.

Species with Simulated Variable Song Production

Although BAWW was simulated with heterogeneity in detection probability due to differences in singing rates, we found very little support for heterogeneity models. Model M_0 was selected for all four data sets (importance weights ≤ 0.24). The population size of 125 was underestimated by all four observers with estimates of 106, 118, 103, and 95 birds.

Species with Empirical Heterogeneous Probabilities of Singing

Heterogeneity models M_{th} , M_{th} or M_{bh} were selected for all BTBW and OVEN data sets, which were simulated based on song data recorded from breeding birds in the field. The variable importance weights for heterogeneity were 1 for BTBW and ≥ 0.88 for OVEN, indicating that models accounting for individual variation in detection probabilities were strongly supported by the data. An increase in detection probabilities (model M_{bh}) following initial detection was also supported by the data from observers 1 and 4. Models incorporating differences in detection probability associated with time intervals were supported for BTBW and OVEN, but they were only selected as the best model for observer 4 and the BTBW. This was expected because no data sets were simulated with variability in detection probability associated with time. True population sizes (133 for BTBW and 127 for

Table 2. Top models chosen using AICc and estimated population sizes and standard errors for four experienced observers. Singing rates of the first four species were homogeneous (all birds had the same probability of calling in each interval). The next four species were heterogeneous. ACFL and SCTA are presented first when groups are modeled with unidentifiable heterogeneity (h) and second when modeled with identifiable heterogeneity (g). All species except the BAWW had a high homogeneous singing rate per interval.

Species	N	Observer 1			Observer 2			Observer 3			Observer 4		
		Model	Est	SE	Model	Est	SE	Model	Est	SE	Model	Est	SE
HOWA	100	M _h	99	0.41	M _h	102	1.69	M _h	99	1.37	M _h	101	1.62
BTNW	125	M ₀	95	4.91	M ₀	97	5.29	M ₀	87	4.61	M ₀	98	5.56
YTWA	167	M _b	162	63.94	M ₀	123	12.13	M _b	224	163.59	M _b	161	76.54
BAWW	125	M ₀	106	14.79	M ₀	118	14.74	M ₀	103	14.4	M ₀	95	11.04
BTBW	133	M _h	93	5.3	M _h	89	2.49	M _h	95	4.58	M _{th}	100	5.54
OVEN	127	M _{bh}	89	1.36	M _h	84	1.26	M _h	99	5.29	M _{bh}	87	0.46
ACFL	133	M _h	96	4.62	M _h	87	4.60	M _h	91	6.12	M _h	89	4.75
SCTA	133	M _h	108	5.33	M _h	105	5.54	M _h	106	5.81	M _h	108	5.84
ACFL	133	M _g	103	8.04	M _{gb}	126	29.03	M _{gb}	98	10.77	M _g	93	5.47
SCTA	133	M _g	116	9.13	M _{gb}	104	10.57	M _g	114	10.18	M _g	121	12.04

OVEN) were substantially underestimated. Observers never estimated populations greater than the 100 birds that actually sang in at least one interval. Some individuals of these two species had very low detection probabilities.

Species with Two-Group Heterogeneous Singing Probabilities

Finite mixture heterogeneity models (Pledger 2000) performed best for the ACFL and SCTA. The ACFL population estimates were substantially negatively biased, with estimates of 96, 87, 91, and 89 individuals. The SCTA population estimates were similar with estimates of 108, 105, 106, and 108 individuals.

Estimates improved when we repeated these analyses to include two groups with different probabilities of singing in at least one interval. It is not surprising that the group effect (individuals matched to high or low availability) was very important for both ACFL and SCTA. The AIC_c importance weights of 1 for the group effect indicate that all models included this effect (Table 2). The selected models for observer 2 for both species, and observer 3 for ACFL also included a behavioral effect, indicating that the probability of subsequent detection of an individual increased after first detection. The total populations for these two species were 133. ACFL population estimates were substantially negatively biased (population size estimates 103, 98, and 93, respectively) although less so than when the group effect was not used. The exception was observer 2 (population size estimate

126). Populations were slightly underestimated for SCTA (population size estimates 116, 104, 114, and 121), although less so than when the group effect was not used.

Portions of the ACFL and SCTA populations were simulated with an availability of 1, and we found that the population size estimate for ACFL was 46, SE = 0.47 and the population size estimate for SCTA was 51, SE = 0.43, where the true population size was 50. These estimates have small SEs and means close to the truth so that they are accurate and precise, similar to the results we obtained for HOWA described earlier.

DISCUSSION

We believe that this is a unique and important field test of the performance of the time-of-detection method of estimating population size on aural point counts. Knowing the true number of birds of each species singing allowed us to directly assess the performance of the method. We were able to compare the true population density parameter with its corresponding estimate under different experimental conditions. In contrast, many field tests reported in the literature only compare different population estimates with no knowledge of the true population parameters.

We consider our test as a “quasi-experiment” rather than a true experiment because of important practical limitations. We used one single point in one habitat under leaf-off conditions, therefore eliminating the effects of spatial or temporal variation (environmental variation in topography, background noise, foliage, etc.). We were unable to choose our observers randomly and did not model observer effects in our analyses. We used one example of each species song in the interests of simplicity and control. We also were unable to allow for individual birds to move during the count. However, we were able to control for other important factors including song orientation, song height, and song volume, and we were able to randomize the order in which species called within different fixed-distance bands.

Our system does not allow us to evaluate the effect of mixed-cue detections where, e.g., a bird is first detected by a sound cue and then sighted and localized from a mix of sound and visual cues. This is an important issue that we leave to other

researchers to study. However, we believe that our results are very important and relevant to many field ornithologists because such a high percentage of birds are detected aurally in forested habitats (Simons et al. 2007). An important strength of our study was that we could assess in detail the effects of variable detection probability caused by variability in the singing rates of individual birds. We found that the method estimated population size accurately for species simulated with high homogeneous singing rates. Populations of species with lower but homogeneous singing rates were somewhat underestimated, but in general, the method performed fairly well. When probabilities of singing were heterogeneous among individuals of the same species, the method often substantially underestimated populations. Heterogeneity is known to cause underestimation in capture–recapture models, even in models allowing heterogeneity in capture probabilities. Otis et al. (1978) carried out detailed simulation studies and showed the problem occurs whenever animals have detection probabilities close to zero. If heterogeneity exists, but detection probabilities are not near zero, heterogeneity models perform reasonably well.

We believe the problem of aural sampling of avian species with very low singing rates, and hence low overall detection probabilities, is a very difficult one. Our evaluation showed serious negative bias when using the time-of-detection method, but unfortunately, two other methods of estimating detection probability using distance and multiple-observer sampling will perform even worse than the time-of-detection methods because they assume that all birds sing during the count interval. A possible solution that should be investigated further involves the use of the repeated counts method (Kery et al. 2005). However, the repeated counts method is costly and may have its own limitations, and so we do not discuss it further in this paper.

We attempted to make this evaluation as realistic as possible by including species with very heterogeneous singing rates and a variety of song characteristics, including non-random singing bouts based on empirical field data. In fact, our evaluation may have included more situations with lower singing rates than would be expected. Nevertheless, in other respects, our field evaluation was somewhat oversimplified compared with real point counts where more individuals and species are present, where the individuals of each species may be closer

than a 45° angle of separation, where observers attempt to estimate the distance to singing birds, and where birds constantly move and change their orientation with respect to observers. In addition, the species we simulate (HOWA, SCTA, ACFL, OVEN) had very distinct and conspicuous calls, and our observers were very familiar with the simulated songs because of their participation in previous experiments. Thus, in most cases, levels of accuracy and precision reported from these experiments are probably higher than those expected in actual field studies where only aural detections are used.

The time-of-detection method is quite demanding on observers in that they have to keep track of individual birds to obtain an accurate detection history. In our study, the observers had to track 18 total individual birds and focus on getting detection history data for 16 birds (two individuals of each of the eight species). Unfortunately, we were unable to evaluate this component of the method, and in future work, we would like to vary the number of birds at a point and see how that changes the quality of the data collected by the observers.

This paper has focused on the use of the time-of-detection method for aural detections because of the nature of our song simulation system and because such a high proportion of avian detections in forested systems are aural. We believe that the time-of-detection method has value for this situation. We are less certain of the value of the method for visual detections and believe this question requires further research. Humans process auditory and visual information differently. We believe that an observer's field of view limits the proportion of a plot that can be sampled visually in each time interval, whereas hearing provides a much better snapshot of the whole plot and is, therefore, better suited to this method. An important related question concerns the analysis of data comprised of detections from different types of cues (aural only, visual only, and a mix of aural and visual). For example, visual detections may include both males and females of some species, whereas auditory detections are presumed to indicate breeding males in most surveys. We are concerned about mixing visual and auditory observations because the detection processes are so different for different types of cues. We believe this is still an open question that should be examined in future research.

When using the time-of-detection method, we recommend the capture–recapture version (Allredge

et al. 2007a), which uses the full detection history rather than the removal version (Farnsworth et al. 2002), which just uses the time of first detection. At least four time intervals of equal length with 8–10 min total time might be reasonable, but we did not study this question in these experiments. A key assumption of the method is that observers localize and track individual birds without error. Therefore, the longer the count interval, and the greater the rate of bird movement, the greater the likelihood that this assumption is violated. We believe the method has promise when localization errors can be kept to a minimum, e.g., for species such as Whip-poor-wills (*Caprimulgus vociferous*) with large territories, high singing rates, and low movement rates. We encourage other researchers to try the method on a variety of species so that our knowledge of when it is most useful can be increased. The quality of the data will also be influenced by the number of species and individuals being tracked. Therefore, we suggest that researchers consider applying the method to a few key species in habitats where many individuals of many species are present.

Finally, we believe that research on combined methods, such as combined multiple-observer and time-of-detection methods (Pollock et al., submitted) could prove quite useful. This approach can provide estimates of both components of the detection process (the probability of singing and the probability of detection, given that the bird sings). Repeated-count methods could also be combined with methods such as distance, multiple observers, and time-of-detection methods. These combined approaches, although more complex and expensive to implement could provide very useful insights into the overall detection process.

Responses to this article can be read online at:
<http://www.ace-eco.org/vol2/iss2/art13/responses/>

Acknowledgments:

We are very grateful to the many volunteers who assisted with this research: Jerome Brewster, Adam Efird, Mark Johns, Ryan Myers, Shiloh Schulte, and Clyde Sorenson. Electrical engineering students at NCSU: John Marsh, Marc Williams, Michael Foster, and Wendy Moore provided valuable technical assistance. Funding for this research was provided

by the USGS Status and Trends Program, the US Forest Service, the US National Park Service, and the North Carolina Wildlife Resources Commission. We also thank Dr. Barbara Brunhuber for substantial editorial assistance.

LITERATURE CITED

- Alho, J. M.** 1990. Logistic regression in capture–recapture models. *Biometrics* **46**:623–635.
- Allredge, M. W.** 2004. Avian point-count surveys: estimating components of the detection process. Dissertation, North Carolina State University, Raleigh, North Carolina, USA.
- Allredge, M. W., K. H. Pollock, and T. R. Simons.** 2006. Estimating detection probabilities from multiple observer point counts. *The Auk* **123**:1172–1182.
- Allredge, M. W., K. H. Pollock, T. R. Simons, J. A. Collazo, and S. A. Shriner.** 2007a. Time of detection method for estimating abundance from point count surveys. *The Auk* **124**:653–664.
- Allredge, M. W., T. R. Simons, and K. H. Pollock.** 2007b. Factors affecting aural detections of songbirds. *Ecological Applications* **17**:948–955.
- Allredge, M. W., T. R. Simons, and K. H. Pollock.** 2007c. An experimental evaluation of distance measurement error in avian point count surveys. *Journal of Wildlife Management* (in press).
- Bart, J.** 2005. Monitoring the abundance of bird populations. *The Auk* **122**:15–25.
- Brewster, J. P.** 2007. *Spatial and temporal variation in the singing rates of two forest songbirds, the ovenbird and the Black-throated Blue Warbler: implications for aural counts of songbirds*. Thesis, Department of Zoology, North Carolina State University, Raleigh, North Carolina, USA.
- Buckland, S. T., D. R. Anderson, K. P. Burnham, J. L. Laake, D. L. Borchers, and L. Thomas.** 2001. *Introduction to distance sampling*. Oxford University Press, Oxford, UK.
- Burnham, K. P., and D. R. Anderson.** 2002. *Model selection and inference: a practical information theoretic approach*. Springer-Verlag, New York, New York, USA.
- Burnham, K. P., and W. S. Overton.** 1978. Estimation of the size of a closed population when capture probabilities vary among animals. *Biometrika* **65**:625–633.
- Collins, S.** 2004. Vocal fighting and flirting: the functions of birdsong. Pages 69–72 in P. Marler and H. Slabbekoorn, editors. *Nature's music: the science of birdsong*. Elsevier Academic Press, San Diego, California, USA.
- Farnsworth, G., K. H. Pollock, J. D. Nichols, T. R. Simons, J. E. Hines, and J. R. Sauer.** 2002. A removal model for estimating the detection probability during point counts divided into time intervals. *The Auk* **119**:414–425.
- Gibbs, J. P., and D. G. Wenny.** 1993. Song output as a population estimator: effect of male pairing status. *Journal of Field Ornithology* **64**:316–322.
- Huggins, R. M.** 1989. On the statistical analysis of capture experiments. *Biometrika* **76**:133–140.
- Huggins, R. M.** 1991. Some practical aspects of a conditional likelihood approach to capture experiments. *Biometrics* **47**:725–732.
- Kery, M., J. A. Royle, and H. Schmid.** 2005. Modeling avian abundance from replicated counts using binomial mixture models. *Ecological Applications* **15**:1450–1461.
- Link, W. A.** 2003. Nonidentifiability of population size from capture–recapture data with heterogeneous detection probabilities. *Biometrics* **59**:1123–1130.
- Marsh, H., and D. F. Sinclair.** 1989. Correcting for visibility bias in strip transect aerial surveys of aquatic fauna. *Journal of Wildlife Management* **53**:1017–1024.
- Nichols, J. D., J. E. Hines, J. R., Sauer, F. W. Fallon, J. E. Fallon, and P. J. Heglund.** 2000. A double observer approach for estimating detection probability and abundance from point counts. *The Auk* **117**:393–408.
- Otis, D. L., K. P. Burnham, G. C. White, and D. L. Anderson.** 1978. Statistical inference from

capture data on closed animal populations. *Wildlife Monographs* **62**. The Wildlife Society, Bethesda, Maryland, USA.

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

Pollock, K. H., M. W. Alldredge, and T. R. Simons. Submitted. Separation of availability and perception processes for aural detection in avian point counts: a combined multiple observer and time-of-detection approach. *Avian Conservation and Ecology - Écologie et conservation des oiseaux*.

Pollock, K. H., H. Marsh, L. L. Bailey, G. L. Farnsworth, T. R. Simons, and M. W. Alldredge. 2004. Separating components of detection probability in abundance estimation: an overview with diverse examples. Pages 43–58 in W. L. Thompson, editor. *Sampling rare and elusive species: concepts, designs and techniques for estimating population parameters*. Island Press, Washington D.C., USA.

Pollock, K. H., J. D. Nichols, C. Brownie, and J. E. Hines. 1990. Statistical inference for capture–recapture experiments. *Wildlife Monographs* **107**:1–97.

Robbins, C. S., B. Brun, and H. S. Zin. 1983. *Birds of North America: a guide to field identification*. Golden Press, New York, New York, USA.

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

Sauer, J. R., J. E. Hines, and J. Fallon. 2005. The North American breeding bird survey, results and analysis 1966–2004. Version 2005.2, US Geological Survey, Patuxent Wildlife Research Center, Laurel, Maryland, USA. (online) URL: <http://www.mbr-pwrc.usgs.gov/bbs/bbs.html>.

Seber, G. A. F. 1982. *The estimation of animal abundance and related parameters*. Second edition, Charles W. Griffin, London, UK.

Simons, T. R., M. W. Alldredge, K. H. Pollock, and J. M. Wettröth. 2007. Experimental analysis of the auditory detection process on avian point counts. *The Auk* **124**:986–999.

Walton, R. K., and R. W. Lawson. 1999. *Birding by ear: eastern/central*. Petersen Field Guides (Audio CD). Houghton Mifflin, Boston, Massachusetts, USA.

White, G. C., and K. P. Burnham. 1999. Program MARK: survival estimation from populations of marked animals. *Bird Study* **46**:S120–S139.

Williams, B. K., J. D. Nichols, and M. J. Conroy. 2002. Analysis and management of animal populations: modeling, estimation, and decision making. Academic Press, San Diego, California, USA.

Zippin, C. 1958. The removal method of population estimation. *Journal of Wildlife Management* **22**:82–90.