J Ornithol (2005) 146: 151–159
DOI 10.1007/s10336-005-0074-9

ORIGINAL ARTICLE

Matthew G. Betts · Neal P. P. Simon
Joseph J. Nocera

Point count summary statistics differentially predict reproductive activity in bird-habitat relationship studies

Received: 22 September 2004 / Revised: 20 December 2004 / Accepted: 23 December 2004 / Published online: 9 March 2005

Abstract Point count summary statistics (e.g. *mean abundance*, *maximum abundance*, *frequency and presence/absence*) reflect different assumptions about behavioral and population processes. In this paper we (1) determine the frequency and usage trends of different point count summary statistics in recent ornithological literature, and (2) assess how well point count data, summarized using five common statistics, predict an alternate measure of habitat quality–reproductive activity. For the 100 journal years we reviewed (10 journals over 10 years), 148 papers used point counts to evaluate bird habitat relationships. The number of papers using point counts has increased over the decade. *Mean abundance*, the most common summary statistic, was used more than twice as frequently as the next most common summary statistics. Only 25.7% (38 papers) provided a justification for use of a particular summary technique. We conducted point counts in three Canadian study regions (New Brunswick, Nova Scotia, and Newfoundland and Labrador) comprised of two ecosystem types (forest and grassland). While there was a statistically significant positive correlation between point count data and reproductive activity data for most species, we found that point counts were often unsuccessful at predicting reproductive activity in forest birds. For species where point counts adequately predicted reproductive activity *mean abundance* and *frequency* were consistently the best predictors. Our results indicate that statistics using information on intra-season (multiple-visit) occupancy tend to be better estimators of reproductive activity.

Keywords Habitat relationships · Intra-season occupancy · Point count data · Reproductive activity

Communicated by F. Bairlein

M. G. Betts (✉)
Greater Fundy Ecosystem Research Group,
Hugh John Flemming Forestry Centre,
1350 Regent St, Fredericton, New Brunswick,
Canada, E3C 2G6
E-mail: m.betts@unb.ca

N. P. P. Simon
Newfoundland and Labrador Department of Natural Resources,
3014 Station B, Happy Valley-Goose Bay,
Newfoundland and Labrador, Canada, A0P 1E0

J. J. Nocera
New Brunswick Cooperative Fish and Wildlife Research Unit,
Biology Department, University of New Brunswick,
Bag Service#45111, Fredericton,
New Brunswick, Canada, E3B 6E1

In habitat-relationship research, point counting, which involves the visual and auditory detection of birds within fixed or variable-radius plots, is one of the most commonly used methods to survey birds (Ralph et al. 1995; Thompson et al. 2002). Considerable effort has been directed towards improving the accuracy and efficiency of this technique. Suggestions have been provided for the optimal number of visits (Siegel et al. 2001), frequency of visits (Brooks et al. 2001), length of count (Lynch 1995; Dawson et al. 1995; Drapeau et al. 1999), sample size (Smith et al. 1995; Smith and Twedt 1999), and point count radius and detection probabilities (Thompson and Schwalbach 1995; Nichols et al. 2000; Farnsworth et al. 2002; Thompson et al. 2002).

There is a diversity of point count data summary statistics including: *total abundance* (summed for each point across many visits), *mean abundance* (total per point/number of visits), *maximum abundance* (the greatest number of individual birds recorded at a point out of all visits), *presence/absence*, and *frequency* (presence/absence per point summed over all visits). If distance to observed birds is estimated during counts, these data may also be used to estimate true density (Bart and Earnst 2002). However, such estimates of true density are still uncommon in the avian literature. Justifications for selection of point count summary statistics are infrequently provided. In the determination of bird-habitat relationships, this is problematic given that each data summary statistic may emphasize different (unstated) assumptions about behavioral and population processes.

For instance, Bibby et al. (2000) suggested the use of maximum abundance, which is based on the assumption that abundance is generally underestimated by point counts (see Miller et al. 2004). Therefore, the maximum number of birds detected in any visit represents the minimum number at that location (Bibby et al. 2000). This may be troublesome because the maximum abundance is the statistic most likely to be biased by the inclusion of non-breeders, e.g. migrants or floaters, and thus inflates estimates of habitat quality. In landscapes with strong source-sink dynamics, maximum abundance could be a poor indicator of habitat quality (Van Horne 1983; Pulliam 1988).

Likewise, frequency, mean abundance and total abundance may reflect seasonal singing frequency in addition to actual density. Studies have shown that unmated males sing more frequently later in the season than mated males (Gibbs and Wenny 1993; Mcshea and Rappole 1997; Amrhein et al. 2002) suggesting these statistics could overestimate habitat quality. Mean abundance and frequency are the most likely of the summary statistics to underestimate true abundance because birds present may not be detected in particular visits resulting in decreased abundance estimates (Bibby et al. 2000).

The use of presence/absence is based on the assumption that the count of an individual in one visit is equivalent to counts of multiple individuals in one location or among visits. This may mask more subtle differences in habitat quality, particularly for more abundant species.

Research that relates various point count summary statistics to different measures of habitat quality such as reproductive success, nest density, or survival is rare (but see Jones et al. 2000), likely due the fact that such data are difficult and expensive to collect. Recent approaches to assessing 'reproductive activity' at broad scales (Vickery et al. 1992; Gunn et al. 2000) provide the opportunity to compare relative merits of point count data summary statistics to an alternate indicator of habitat quality.

Our objectives were to: (1) determine the frequency and usage trends of different point count summary statistics in recent ornithological literature, and (2) compare the efficacy of various point count data summarization methods at predicting a different index of habitat quality–'reproductive activity'. We determined reproductive activity indirectly from observations of reproductive behavior including pairing, food and nest material carrying, and the presence of fledglings (sensu Vickery et al. 1992; Gunn et al. 2000; Doran et al. 2005). Determining the relationship between point count data and reproductive indices is useful in light of the increased use of these indices in bird habitat studies (Vickery et al. 1992; Dale et al. 1997; Powell and Collier 1998; Martin and Morrison 1999; Rangen et al. 2000; Cornell Lab of Ornithology 2003; Christoferson and Morrison 2001; Harris and Reed 2002; Stevens et al. 2003). If point count data can effectively predict reproductive activity, the additional effort expended to collect reproductive activity data may be unnecessary.

Methods

Literature review

We searched for all articles using point count statistics in bird habitat studies in ten journals for the period 1992–2002: *Auk*, *Condor*, *Conservation Biology*, *Ecological Applications*, *Journal of Avian Biology* (formerly *Ornis Scandinavica*), *Journal of Field Ornithology*, *Journal of Ornithology*, *Journal of Wildlife Management*, *Wildlife Society Bulletin*, and *Wilson Bulletin*. We selected journals from both Europe and North America that most commonly publish bird-habitat relationship studies. We felt that 10 years was an adequate timeframe to investigate trends in reporting, and offered an opportunity to assess differences between the period preceeding, and following, the publication of the point count methodology book edited by Ralph et al. (1995).

For all titles of articles and short communications that referred to research on bird-habitat associations, we read abstracts to determine whether point counts were used. If the abstract indicated possible use of point counts, we read the entire methodology section of the article. We quantified the number of studies that summarized their count data according to one or more of the following: mean abundance/total abundance, maximum abundance, presence/absence, and frequency. Mean and total abundance were grouped as these statistics result in identical data structure. Secondly, we quantified the number of those reports that provided justification for use of the selected statistic, beyond implicit assumptions.

We used logistic regression in S-Plus (Mathsoft 2000) to determine whether the choice of point count summary statistics are increasingly justified in the literature ($\alpha = 0.05$). 'Year' was the predictor variable and 'presence or absence of justification' was the binomial response.

Point counts and reproductive activity

We conducted point counts and reproductive activity assessments for grassland and forest passerines (Table 1) at three sites in eastern Canada in 2002. In all cases, our observers were trained by an experienced observer who conducted simultaneous point counts to establish a consistent protocol before proceeding to collect data independently.

We counted forest birds at 320 points in a 4,000-km^2 area in southeastern New Brunswick (NB: 46°04′N, 64°56′W) and at 225 points in a 770-km^2 area around Goose Bay, Newfoundland and Labrador (NL: 53°20′N, 60°25′W), Canada. Point counts were conducted using 50- and 100-m radii. In NB, we conducted three counts

Table 1 Focal species surveyed by point counts at study sites in New Brunswick (NB), Newfoundland and Labrador (NL), and Nova Scotia (NS), Canada in 2002

Common name	Scientific name	NB	NL	NS
Blue-headed vireo	*Vireo solitarius*	X		
Red-eyed vireo	*Vireo olivaceus*	X		
Gray jay	*Perisoreus canadensis*		X	
Boreal chickadee	*Poecile hudsonica*		X	
Golden-crowned kinglet	*Regulus satrapa*	X		
Ruby-crowned kinglet	*Regulus calendula*	X	X	
Hermit thrush	*Catharus guttatus*	X		
American robin	*Turdus migratorius*	X		
Magnolia warbler	*Dendroica magnolia*	X		
Black-throated blue warb1ler	*Dendroica caerulescens*	X		
Yellow-rumped warbler	*Dendroica coronata*	X	X	
Black-throated green warbler	*Dendroica virens*	X		
Blackburnian warbler	*Dendroica fusca*	X		
Bay-breasted warbler	*Dendroica castanea*	X		
American redstart	*Setophaga ruticilla*	X		
Ovenbird	*Seiurus aurocapilla*	X		
Savannah sparrow	*Passerculus sandwichensis*			X
Nelson's sharp-tailed sparrow	*Ammodramus nelsoni*			X
Song sparrow	*Melospiza melodia*			X
Dark-eyed junco	*Junco hyemalis*	X	X	
Bobolink	*Dolichonyx oryzivorus*			X

of 5-min duration between 0530–1030 hours AST from 5 June and 11 July. Whereas two counts of 5-min duration were conducted between 0400–1100 hours AST from 10 June to 25 July at the NL site.

We surveyed three focal grassland bird species (Table 1) at 40 point count plots (50 m radii only) in hayfields of the Annapolis Valley, Nova Scotia (NS: centered around 44°45′N, 65°31′W) from 29 May to 2 August. A fourth species, song sparrow (*Melospiza melodia*) is more associated with old fields and farmland edges, but was adequately surveyed with our methodology and will hereafter be inclusively discussed with "grassland birds". We conducted 5-min counts 10–12 times (mean = 10.55 counts per plot; i.e. once every 5.3 days) between 30 min post-sunrise and 1000 hours AST (weather permitting: wind <25 kph, no precipitation).

We summarized all of our point count data, by species, with the five statistics of interest: mean, maximum abundance, presence/absence, presence/absence in the first round (using only observations from the first visit), and frequency. We used presence/absence in the first round to approximate data collected in studies with only one point count visit. We summarized data in these ways at both 50- and 100-m extents for points in the forest bird studies, but for only the 50-m extent in the grassland bird study (100-m was not available).

We used the procedure of Gunn et al. (2000) to index breeding activity of five focal species at the NL site and 15 focal species at the NB site (see Table 1) over the course of four visits to each point count station. Observers were alternated between visits to prevent observer bias. We played 5-min playbacks of black-capped chickadee (*Poecile atricapillus*) mobbing calls to elicit an interspecific mobbing response from birds holding territories surrounding sample points (Betts et al. 2005). We recorded all evidence of reproductive activity during the 5-min sample period at each of the four visits. We considered reproductive activity to include: male and female present in close proximity (suspected pair), adults carrying nest material or food, and adults feeding fledglings (after Gunn et al. 2000). Similarly, we recorded evidence of reproductive activity at the grassland bird study site (NS) except rather than using a playback tape, the data were collected through passive observation (Vickery et al. 1992; Christoferson and Morrison 2001) of birds within 50-m point count radii during the 10–12 standard point counts. As per the method detailed by Vickery et al. (1992), point count plots were considered reproductively active if at least one pair was consistently observed at a plot over a period of at least 3 weeks, or when evidence of breeding (e.g. delivery of food or nest material) was observed at least once.

We used playback observations (forest birds) and passive observations (grassland birds) as a binary index of reproductive activity in which any of the above observations constituted 'evidence of reproductive activity' and either the presence of a male or no observation of the species constituted 'no evidence of reproductive activity'. Such binary reproductive activity data are correlated with actual reproductive output at coarse scales of resolution (Doran et al. 2005).

We caution that the above index is only a coarse measure of reproductive success that is not as accurate as data from intensive nest monitoring. However, the sample size and spatial scale required to test the efficacy of point count statistics makes the collection of nest data practically infeasible. Tests of the mobbing playback method to date have shown that reproductive activity observations reflect actual reproductive success (defined as actual reproductive output) for forest birds in both fragmented (Gunn et al. 2000) and contiguous forested landscapes (Doran et al. 2005). The passively-derived index for grassland birds has also been shown to adequately reflect true reproductive activity (Vickery et al. 1992).

To test how well the point count summary statistics of interest predicted our different measure of habitat quality–reproductive activity, we modeled this variable as a function of each summary statistic using logistic regression (Mathsoft 2000). Each summary statistic was used as a univariate predictor and reproductive activity was the binomial response. Results of logistic regression models are often judged to be considerably more reliable if predicted values >0.5 correspond with observed occurrences and values <0.5 with absences. However, the 0.5 threshold is arbitrary and lacks ecological and statistical basis (Osborne et al. 2001). For this reason, we used receiver operating characteristic (ROC) curves as a measure of model accuracy. The ROC curve describes the relationship between the sensitivity (number of positive observations correctly predicted as positive) and number of false positive predictions (Hanley and McNeil 1982). The area under the ROC curve is a single index of classification accuracy that is independent of species prevalence and arbitrary threshold effects (Hanley and McNeil 1982; Fielding 2002). The area under the curve (AUC) can be interpreted as the probability that a model will correctly distinguish between positive and negative observations (in this case, evidence of reproductive activity versus no such evidence). As a result of these strengths, AUC is increasingly used in ecology literature as a measure of prediction success (Pearce and Ferrier 2000; Manel et al. 2001; Store and Jokimaki 2003; Vernier et al. 2004). The AUC values of 0.5 represent no discrimination (a poor model). As a general rule, if AUC2000). Statistical significance for all models was determined by using Chi-square to test for differences between each model and the intercept-only model.

We also sought to determine whether detection frequency (the proportion of sample points across all visits in which a species was present) influenced which point count statistic best explained reproductive activity across all species. To test for this relationship, we used Spearman rank correlation with a robust point estimate of ρ (correlation coefficient) based on a 0.8 fraction of the data (using *R* (Version 1.8.1) open-source software; R Development Core Team 2003).

Results

Literature review

For the 100 journal years we reviewed, 148 papers used point counts to evaluate bird-habitat relationships. The number of papers using point count methods increased over the decade (Fig. 1). Mean abundance, the most common summary technique, was used more than twice as frequently as the next most common summary statistics, i.e., presence/absence and maximum abundance (Fig. 2). The remaining techniques, including estimates of density using distance-sampling (Buckland et al. 1993) were uncommon. In two studies the summary statistic used was not stated. Overall, 25.7% (38 papers) justified use of a summary technique. There was no temporal pattern in justification of summary statistics over the 10-year period ($\chi^2 = 0.02$, $df = 147$, $P = 0.88$).

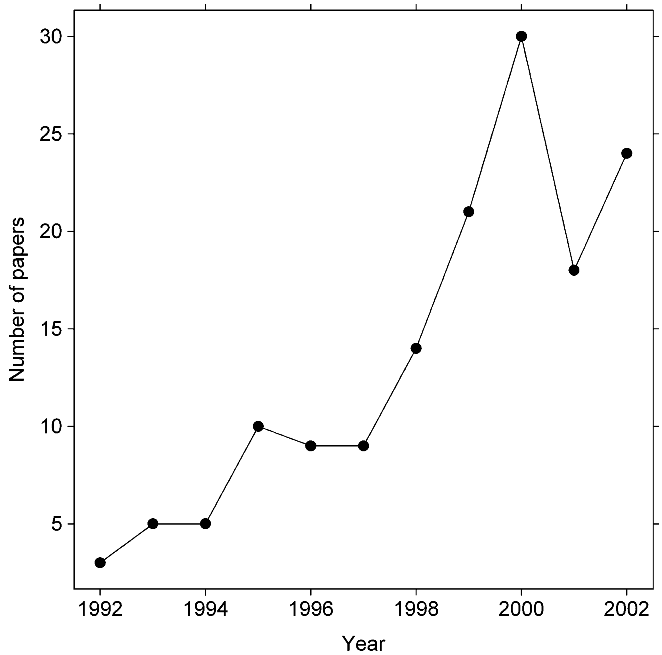

Fig. 1 Absolute number of papers using point count methodology from ten major ecological/ornithological journals (1992-2002)

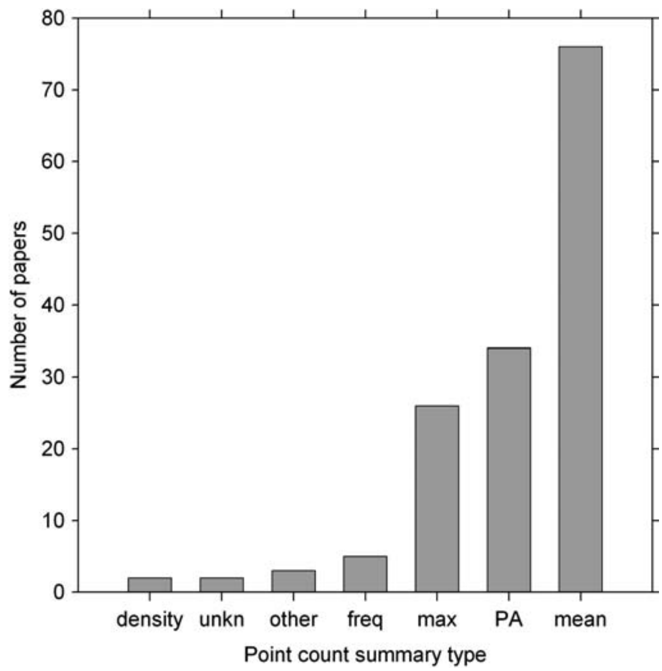

Fig. 2 Frequency of point count statistic use in the ornithological literature from 1992–2002. Summary technique abbreviations are listed in Table 2

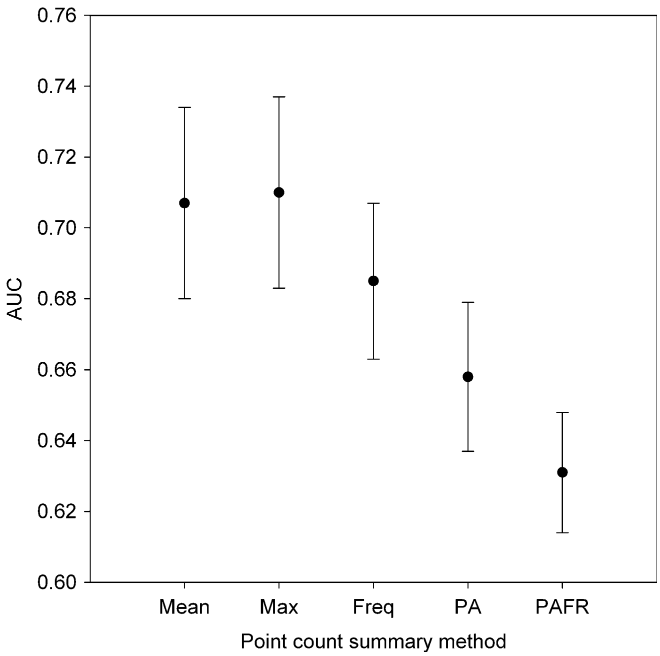

Fig. 3 Mean area under the curve (AUC) (±SE) for models using five different point count summary statistics to predict reproductive activity. Data presented are for 21 species from three study regions (50-m radius counts). AUC is a measure of model prediction success (see Methods)

Point counts and reproductive activity

Across our three regions (at both scales), we determined that summary statistics containing information about intra-season site occupancy (mean abundance and frequency) were usually the most accurate at predicting reproductive activity. However, maximum abundance, on average, was slightly more accurate (mean difference in AUC = 0.025). We found presence/absence and presence/absence in the first round to be the least effective predictors of reproductive activity (Fig. 3). Across summary statistics, there was a high degree of variability in predictive accuracy for each species (Fig. 3, Table 2).

For species where point counts accurately predicted reproductive activity, we determined that mean abundance and frequency were consistently the best predictors at both scales (Table 2). For species where point count data was a relatively poor predictor of reproductive activity (i.e. AUC < 0.7), there was little pattern in the performance of summary statistics.

The ability of summary techniques to predict reproductive activity between areas was quite variable. The grassland study had the strongest relationship between point count data and reproductive activity (mean AUC = 0.81). The average predictive accuracy (AUC) was 22 and 32% greater in the grassland sites than the New Brunswick and Newfoundland and Labrador sites respectively (mean AUC = 0.66 and 0.61).

Although maximum abundance approached statistical significance ($P = 0.06$, Table 3), detection frequency of bird species did not appear to influence which point count statistic best predicted reproductive activity (Table 3).

Discussion

Our literature review showed the number of papers using point count methods increased dramatically between 1992 and 2002. This increase is likely due, in part, to the recent growth of landscape ecology research that often requires extensive, time-efficient survey techniques. The most common statistic has been mean abundance, followed by presence/absence. The relative rate of statistics usage has not significantly changed over the 10-year period.

We tested how well different methods used to summarize point count data predicted an alternate assessment of habitat quality–reproductive activity. For species for which point count data were effective at predicting reproductive activity (AUC > 0.7), mean abundance and frequency were the best statistics. Despite our expectation that presence/absence and presence/absence in the first round would perform better for uncommon species, prevalence (commonness) within a study area did not influence which statistic best predicted reproductive activity. In fact, we found presence/absence and presence/absence in the first round to consistently be the worst predictors.

Both mean abundance and frequency contain information about intra-season (multiple-visit) site occupancy. If a species is observed in point counts at a sample station repeatedly within a season, it is more likely to be associated with a reproductive activity observation. Conversely, presence/absence and presence/absence in the first round contain no information about intra-season (multiple-visit) occupancy. Birds present only once are weighted equally to birds present in two or three rounds. The consistently poorer predictions of reproductive activity by presence/absence, presence/absence in the first round, and maximum abundance could have several causes: (1) although all summary statistics may be biased by the presence of mobile, non-breeding flocks or floaters, this bias would be greater for presence/absence in the first round as in the early season, a greater proportion of observations may be migrants or prospecting males that have not settled to defend a territory; or (2) biases are also likely to result from single-visit observer error in distance estimation or species recognition. Statistics that include information on intra-season occupancy are less prone to these biases. If a species is observed at a station more than once it is unlikely to be a floater. Further, it is less probable that different observers would be prone to the same error in the same location in consecutive visits. It appears that advantages gained by statistics using intra-season information outweigh biases that may result from

Table 2 AUC values used to assess the accuracy of summary statistics at predicting reproductive activity by species, scale and region. Prevalence is the percentage of stations at which birds were detected (see Methods for *n* for each region). Results are in order of decreasing predictive accuracy (mean AUC across statistics), divided into models with good (AUC > 0.7) and poor (AUC < 0.7) predictive accuracy. For each species, highest AUC values are in bold

Species	Scale	Area	Prevalence	Freq	*P*	Mean	*P*	Max	*P*	PA	*P*	PAFR	*P*
AUC > 0.7													
Song sparrow	50	NS	23	**1.00**	0.0001	0.995	0.0001	0.932	0.0001	0.919	0.0001	0.806	0.033
Bobolink	50	NS	95	0.936	0.0001	**0.967**	0.0001	0.904	0.0001	0.559	0.072	0.797	0.0001
Red-eyed vireo	100	NB	49	0.850	0.0001	**0.863**	0.0001	0.824	0.0001	0.777	0.0001	0.735	0.0001
Ovenbird	50	NB	38	0.845	0.0001	**0.847**	0.0001	0.834	0.0001	0.819	0.0001	0.669	0.0001
Ovenbird	100	NB	54	0.823	0.0001	**0.842**	0.0001	0.819	0.0001	0.728	0.0001	0.749	0.0001
Red-eyed vireo	50	NB	33	0.817	0.0001	**0.820**	0.0001	0.809	0.0001	0.799	0.0001	0.711	0.0001
Savannah sparrow	50	NS	100	**0.986**	0.0001	0.910	0.0001	0.736	0.051	0.500	NA	0.777	0.016
Black-throated blue warbler	100	NB	18	0.771	0.0001	**0.788**	0.0001	0.776	0.0001	0.771	0.0001	0.749	0.0001
Gray jay	100	NL	27	0.771	0.0001	**0.780**	0.0001	0.770	0.0001	0.760	0.0001	0.660	0.0001
Gray jay	50	NL	19	0.770	0.0001	**0.780**	0.0001	**0.780**	0.0001	0.770	0.0001	0.640	0.0001
American redstart	50	NB	24	0.765	0.0001	**0.768**	0.0001	0.757	0.0001	0.750	0.0001	0.691	0.0001
American redstart	100	NB	26	0.765	0.0001	**0.765**	0.0001	0.751	0.0001	0.740	0.0001	0.693	0.0001
Magnolia warbler	100	NB	60	0.728	0.0001	**0.732**	0.0001	0.710	0.0001	0.684	0.0001	0.695	0.0001
Black-throated blue warbler	50	NB	13	**0.752**	0.0001	0.732	0.0001	0.730	0.0001	0.725	0.0001	0.616	0.0001
Magnolia warbler	50	NB	56	0.732	0.0001	**0.741**	0.0001	0.696	0.0001	0.683	0.0001	0.677	0.0001
Golden-crowned kinglet	100	NB	38	**0.728**	0.0001	0.726	0.0001	0.701	0.0001	0.694	0.0001	0.668	0.0001
Golden-crowned kinglet	50	NB	36	**0.719**	0.0001	0.718	0.0001	0.695	0.0001	0.695	0.0001	0.659	0.0001
Nelson's sharp-tailed sparrow	50	NS	75	**0.751**	0.009	0.737	0.029	0.694	0.031	0.643	0.0001	0.657	0.104
Blackburnian warbler	50	NB	24	0.703	0.0001	**0.716**	0.0001	0.706	0.0001	0.702	0.0001	0.616	0.0001
Bay-breasted warbler	50	NB	27	**0.714**	0.0001	0.712	0.0001	0.694	0.0001	0.691	0.0001	0.620	0.0001
Bay-breasted warbler	100	NB	29	**0.714**	0.0001	0.713	0.0001	0.691	0.0001	0.688	0.0001	0.614	0.0001
AUC < 0.7													
Northern parula	50	NB	9	0.673	0.012	**0.681**	0.001	0.670	0.013	0.670	0.008	0.610	0.005
Blue-headed vireo	100	NB	25	**0.646**	0.0001	0.644	0.0001	0.640	0.0004	0.637	0.0005	0.566	0.016
Black-throated green warbler	100	NB	88	**0.642**	0.0007	0.632	0.002	0.614	0.0007	0.589	0.0001	0.615	0.0009
Blue-headed vireo	50	NB	16	0.63	0.0001	**0.631**	0.0001	0.620	0.0003	0.624	0.0003	0.558	0.0008
Yellow-rumped warbler	100	NB	55	**0.646**	0.0001	0.640	0.0003	0.625	0.002	0.622	0.0002	0.511	0.713
Boreal chickadee	50	NL	7	**0.620**	0.0001	**0.620**	0.0001	**0.620**	0.0001	**0.620**	0.0001	**0.620**	0.0008
Black-throated green warbler	50	NB	78	0.610	0.0007	0.609	0.0024	0.612	0.0007	0.612	0.0001	**0.593**	0.0009
American robin	100	NB	41	0.613	0.214	0.617	0.175	**0.637**	0.040	0.622	0.080	0.530	0.184
Yellow-rumped warbler	50	NL	51	0.600	0.007	0.600	0.012	0.590	0.024	0.590	0.013	**0.630**	0.0002
Dark-eyed Junco	100	NB	36	0.605	0.0001	0.604	0.0001	0.607	0.0001	**0.612**	0.0001	0.565	0.0001
Boreal chickadee	100	NL	8	**0.610**	0.0002	**0.610**	0.0001	**0.610**	0.0001	**0.610**	0.0002	0.550	0.008
Ruby-crowned kinglet	100	NL	80	**0.610**	0.007	0.600	0.017	0.560	0.082	0.570	0.013	**0.610**	0.002
Dark-eyed junco	50	NB	20	**0.620**	0.003	**0.620**	0.003	0.526	0.544	**0.620**	0.003	0.527	0.213
Yellow-rumped warbler	100	NL	67	0.580	0.047	0.570	0.036	0.560	0.085	0.550	0.127	**0.630**	0.0004
Yellow-rumped warbler	50	NB	40	0.587	0.012	**0.592**	0.009	0.579	0.017	0.577	0.016	0.517	0.013
Dark-eyed junco	50	NL	38	0.570	0.019	0.570	0.253	**0.580**	0.030	**0.580**	0.014	0.510	0.570
Hermit thrush	50	NB	13	0.564	0.250	0.560	0.460	0.563	0.304	**0.565**	0.195	0.552	0.169
Dark-eyed junco	100	NL	60	0.550	0.266	0.560	0.263	**0.570**	0.071	**0.570**	0.047	0.510	0.869
Ruby-crowned kinglet	50	NL	51	0.550	0.205	0.50	0.142	**0.560**	0.124	0.550	0.198	0.540	0.269
American robin	50	NB	9	0.534	0.250	0.535	0.260	0.531	0.450	0.535	0.149	**0.554**	0.100
Hermit thrush	100	NB	45	0.500	0.900	0.515	0.915	0.512	0.890	0.517	0.810	**0.587**	0.184

[a] Study regions: *NS* Nova Scotia, *NB* New Brunswick, *NL* Newfoundland and Labador

[b] Summary technique: *Mean* mean abundance; *Freq* frequency; *Max* maximum abundance; *PA* presence/absence; *PAFR* presence/absence in the first round

Table 3 Results of Spearman rank correlation used to determine whether detection frequency (the proportion of sample points across all visits in which a species was present) influenced which point count statistic best explained reproductive activity for all species

Model	Robust *rho*	*P*
Mean	0.27	0.27
Frequency	0.03	0.68
Maximum	0.19	0.06
Presence/absence in the first round	−0.03	0.89
Presence/absence	−0.30	0.88

an inverse relationship between singing frequency and reproductive success (e.g., Mcshea and Rappole 1997; Amrhein et al. 2002).

Our findings that statistics containing information on intra-season occupancy tend to be the best predictors of reproductive activity coincide with point count statistic usage in the ornithological literature over the past decade, where mean is most commonly used. However, the next most used statistic, presence/absence, less adequately predicted reproductive activity than mean (mean difference in AUC = 0.052). The prevalence of this statistic in the literature is perhaps due to the reliance on logistic regression, which requires binomial response data. Thus, while the use of this statistic is often statistically necessary, our results indicate that use of presence/absence will result in some loss of information about reproductive activity.

The difference between frequency, mean and maximum in predictive ability tended to be small for most species (Table 2, Fig. 3). Given these similarities, it is important to note that both Jones et al. (2000) and Brooks et al. (2001) found that maximum abundance consistently overestimated abundance in comparison to densities derived from spot-mapping.

While a positive correlation between point count data and reproductive activity data for most species was statistically significant, point counts were often unsuccessful at predicting reproductive activity in forest birds (range in best model AUC = 0.5–0.86). Ten species had AUC values lower than 0.7, indicating relatively poor predictive accuracy of point count data, however summarized. This indicates the importance of collecting reproductive data for many species, particularly in forest environments. The poor ability of point count data to predict reproductive activity for some species of forest birds could be due to a variety of factors including: (1) highly variable site and landscape-level characteristics in the New Brunswick and Newfoundland and Labrador study sites resulting in a broad range of habitat quality for most species; (2) some low quality habitats may be acting as attractors for unpaired males (Hagan et al. 1996); and (3) the number of visits to these sites (five) may have been too few to detect all sample points where reproductive activity occurred.

The grassland site (Nova Scotia) had three of four species for which reproductive activity was best predicted by point count data. This may be due to increased visibility in open areas of hayfields, or to the fact that the grassland sites had the highest number of visits per point count station (10–12 over the season). This discrepancy may lend support to the use of a passive index of reproductive activity developed in grasslands (Vickery et al. 1992) and use of a mobbing-playback index of reproductive activity (Gunn et al. 2000) for areas of reduced visibility (e.g., forests). In the NS study area, reproductive activity indices for song sparrow provided no additional information over point count data [AUC (frequency) = 1.0].

It is important to note that our estimates of reproductive activity (Vickery et al. 1992; Gunn et al. 2000) are reflective of key parameters associated with habitat quality such as annual reproductive output (Thompson et al. 2001) but are likely not as informative as true nest data. However, these data are expensive to collect and are rarely available on a scale sufficient to test the adequacy of point count statistics. Clearly, where resources are available, habitat research should rely on these direct measures of habitat quality rather than indices. Similarly, if the objective is to estimate density, distance sampling may be appropriate (Buckland et al. 1993; Rosenstock et al. 2002) or any other of a rapidly growing number of capture-recapture and similar methods available to estimate true density (for a review see, e.g., Williams et al. 2002).

Zusammenfassung

Statistische Kennzahlen bei Punkt-Stopp-Zählungen reflektieren das Brutgeschehen von Vögeln in einem Lebensraum unterschiedlich

Zusammenfassende statistische Kennzahlen bei Punkt-Stopp-Zählungen (z. B. mittlere Dichte, maximale Dichte, Stetigkeit, An-/Abwesenheit) spiegeln unterschiedliche Annahmen zum zugrunde liegenden Verhalten und der populationsbiologischer Prozesse wieder. In diesem Beitrag wollen wir (a) analysieren, mit welchen Verfahren und in welcher Häufigkeit derzeit gearbeitet wird, und (b) prüfen, wie gut Punkt-Stopp-Zählungen die reproduktive Aktivität in einem Lebensraum wiedergeben, als alternatives Maß für Habitatqualität. In den 10 über jeweils10 Jahre gesichteten Zeitschriften fanden sich 148 Arbeiten, die Punkt-Stopp-Zählungen zur Beurteilung der Habitatqualität benutzen. Dabei nahm die Häufigkeit der Anwendung dieser Methode über die 10 Jahre zu. Am häufigsten wurde die mittlere Dichte bestimmt, doch lediglich in 25.7% der Fälle (38 Arbeiten) wurde begründet, warum die gewählte Statistik überhaupt benutzt wurde. Wir selbst verwendeten Punkt-Stopp-Zählungen in drei Untersuchungsgebieten in Kanada (New Brunswick, Nova Scotia, Newfoundland und Labrador) und jeweils im Wald und in der Prärie. Obwohl es durchaus statistisch

signifikante Beziehungen zwischen Punkt-Stopp-Zähldaten und der Brutaktivität bei den meisten Arten gab, fanden wir, dass Punkt-Stopp-Zählungen meist ungeeignet waren, die Brutaktivität von Vögeln im Wald vorherzusagen. Für Arten, bei denen Punkt-Stopp-Zählungen adäquate Vorhersagen zur Brutaktivität ergaben, waren mittlere Dichte und Stetigkeit die für Vorhersagen durchweg besten Parameter. Unsere Ergebnisse verdeutlichen weiterhin, dass Statistiken, die das Vorkommen innerhalb einer Saison berücksichtigen, die für eine Schätzung der Fortpflanzungsaktivität besseren Größen sind.

Acknowledgements We thank A. Breton, M. Kéry, C.J. Ralph, P. Taylor, and an anonymous reviewer for constructive comments on early drafts of this manuscript. Thanks are also due to M. Krawchuk and T. Warren who provided feedback in the conceptual stage of this research. We extend our greatest appreciation to our many field assistants who persevered to help collect these data at all three study sites: J. Colbert, L. Elson,T. Fitzgerald, W. Fitzgerald, S. Glinz, R. Emerson, A. Hadley, J. Haselmayer, T. Jarvis, M. Michelin, R. Neville, M. Peckford, F. Phillips, B. Rodrigues and B. Zitske. Research funding for M.G.B. was provided by NSERC postgraduate scholarship, the Fundy Model Forest, Fundy National Park and the New Brunswick Environmental Trust Fund. N.P.P.S. received support from Newfoundland and Labrador Department of Natural Resources. J.J.N. was supported by a doctoral scholarship (#2.37A.2S) from Wildlife Habitat Canada, and also gratefully acknowledges financial and logistical assistance from the NS Habitat Conservation Fund, Canadian Wildlife Federation, NS Dept. Natural Resources, and Canadian Wildlife Service. All field methods used in this research complied with the current laws of Canada.

References

Amrhein V, Korner P, Naguib M (2002) Nocturnal and diurnal singing activity in the nightingale: correlations with mating status and breeding cycle. Anim Behav 64:939–944

Bart J, Earnst S (2002) Double sampling to estimate density and population trends in birds. Auk 119:36–45

Betts MG, Hadley AS, Doran PJ (2005) Avian mobbing response is restricted by territory boundaries: experimental evidence from two species of forest warblers. Ethology (in press)

Bibby CJ, Burgess ND, Hill DA, Mustoe SH (2000) Bird census techniques, 2nd edn. Academic Press, London

Brooks MA, Harrigan BC, Johnson KM, Lowe DE, Lowery JP, McGlothlin JW, Sasso MM, Smith SA, Cristol DA (2001) Revisit schedule does not affect results of point counts. J Field Ornithol 72:404–411

Buckland ST, Anderson D, Burnham K, Laake J (1993) Distance sampling: estimating abundance of biological populations. Chapman & Hall, London

Christoferson LL, Morrison ML (2001) Integrating methods to determine breeding and nesting status of 3 western songbirds. Wildlife Soc Bull 29:688–696

Cornell Lab of Ornithology (2003) Birds in forested landscapes. http://www.birds.cornell.edu/bfl/survey instructs/visit protocols.htm

Dale BC, Martin PA, Taylor PS (1997) Effects of hay management on grassland songbirds in Saskatchewan. Wildlife Soc Bull 25:616–626

Dawson D, Smith D, Robbins C (1995) Point count length and detection of forest neotropical migrant birds. In: Ralph CJ, Sauer J, Droege S (eds) Monitoring bird populations by point counts. Gen. Tech. Rep. PSW-GTR-149. Albany, CA: Pacific Southwest Research Station, Forest Service, U.S. Department of Agriculture, pp 35–43

Doran PJ, Gulezian PZ, Betts MG (2005) A test of the mobbing playback method as a means to estimate bird reproductive success. J Field Ornithol (in press)

Drapeau P, Leduc A, McNeil R (1999) Refining the use of point counts at the scale of individual points in studies of bird-habitat relationships. J Avian Biol 30:367–382

Farnsworth GL, Pollock KH, Nichols JD, Simons TR, Hines JE, Sauer JR (2002) A removal model for estimating detection probabilities from point-count surveys. Auk 119:414–425

Fielding A (2002) What are appropriate characteristics of an accuracy measure?. In: Scott JM, Heglund PJ, Morrison ML, Haufler JB, Raphael MG, Wall WA, Sampson FB (eds) Predicting species occurrences, issues of accuracy and scale. Island Press, Washington, pp 271–280

Gibbs JP, Wenny D (1993) Song output as a population estimator–effect of male pairing status. J Field Ornithol 64:316–322

Gunn JS, Desrochers A, Villard M-A, Bourque J, Ibarzabal J (2000) Playbacks of mobbing calls of Black-capped Chickadees as a method to estimate reproductive activity of forest birds. J Field Ornithol 71:472–483

Hagan JM, Vanderhaegen W, Mckinley PS (1996) The early development of forest fragmentation effects on birds. Conserv Biol 10:188–202

Hanley J, McNeil B (1982) The meaning and use of the area under a receiver operating characteristic (ROC) curve. Radiology 143:29–36

Harris R, Reed J (2002) Effects of forest-clearcut edges on a forest-breeding songbird. Can J Zool 80:1026–1037

Hosmer D, Lemeshow S (2000) Applied logistic regression. Wiley, Hoboken, N.J.

Jones J, McLeish WJ, Robertson RJ (2000) Density influences census technique accuracy for cerulean warblers in eastern Ontario. J Field Ornithol 71:46–56

Lynch J (1995) Effects of point count duration, time-of-day, and aural stimuli on detectability of migratory and resident bird species in Quintana Roo, Mexico. In: Ralph CJ, Sauer J, Droege S (eds) Monitoring bird populations by point counts. Gen. Tech. Rep. PSW-GTR-149. Albany, CA, Pacific Southwest Research Station, Forest Service, U.S. Department of Agriculture, pp 1–6

Manel S, Williams H, Omerod S (2001) Evaluating presence-absence models in ecology: the need to account for prevalence. J Appl Ecol 38:921–931

Martin JA, Morrison ML (1999) Distribution, abundance, and habitat characteristics of the buff-breasted flycatcher in Arizona. Condor 101:272–281

Mathsoft (2000) S-PLUS 2000. Seattle, Washington

Mcshea W, Rappole J (1997) Variable song rates in three species of passerines and implications for estimating bird populations. J Field Ornithol 68:367–375

Miller, JR, Dixon, MD, Turner, MG (2004) Response of avian communities in large-river floodplains to environmental variation at multiple scales. Ecol Appl 14:1394–1410

Nichols JD, Hines JE, Sauer JR Fallon FW, Fallon JE, Heglund PJ (2000) A double-oberver approach for estimating detection probability and abundance from point counts. Auk 117:393–408

Osborne P, Alonso JC, Bryant RG (2001) Modelling landscape-scale habitat use using GIS and remote sensing: a case study with great bustards. J Appl Ecol 38:458–471

Pearce J, Ferrier S (2000) Evaluating the predictive performance of habitat models developed using logistic regression. Ecol Model 133:225–245

Powell AN, Collier CL (1998) Reproductive success of Belding's savannah sparrows in a highly fragmented landscape. Auk 115:508–513

Pulliam HR (1988) Sources, sinks, and population regulation. Am Nat 132:652–661

Ralph CJ, Sauer J, Droege S (eds) (1995) Monitoring bird populations by point counts. Gen. Tech. Rep. PSW-GTR-149. Pacific southwest research station forest service U.S. Department of Agriculture, Albany, Calif.

Rangen SA, Hobson KA, Clark RG (2000) A comparison of density and reproductive indices of songbirds in young and old boreal forest. Wildlife Soc Bull 28:110–118

R Development Core Team (2003) R: a language environment for statistical computing. R foundation for statistical computing. Vienna, Austria

Rosenstock S, Anderson D, Giesen K, Leukering T, Carter M (2002) Landbird counting techniques: current practices and an alternative. Auk 119:46–53

Siegel, RB, Desante DF, Nott MP (2001) Using point counts to establish conservation priorities: how many visits are optimal? J Field Ornithol 72:228–235

Smith WP, Twedt DJ (1999) Temporal differences in point counts of bottomland forest birds. Wilson Bull 111:139–143

Smith W, Twedt D, Cooper R, Wiedenfeld D, Hamel P, Ford R (1995) Sample size and allocation of effort in point count sampling of birds in bottomland hardwood forests. In: Ralph CJ, Sauer J, Droege S (eds) Monitoring bird populations by point counts. Gen. Tech. Rep. PSW-GTR-149. Pacific Southwest research station forest service U.S. Department of Agriculture Albany CA, pp 7–17

Stevens C, Gabor T, Diamond A (2003) Use of restored small wetlands by breeding waterfowl in Prince Edward Island, Canada. Restor Ecol 11:3–12

Store R, Jokimaki J (2003) A GIS-based multi-scale approach to habitat suitability modeling. Ecol Model 169:1–15

Thompson F, Schwalbach M (1995) Analysis of sample size, counting time, and plot size from an avian point count survey on Hoosier National Forest, Indiana. In: Ralph CJ, Sauer J, Droege S (eds) Monitoring bird populations by point counts. Gen. Tech. Rep. PSW-GTR-149. Pacific Southwest research station forest service U.S. Department of Agriculture, Albany CA, pp 45–48

Thompson B, Knadle G, Brubaker D, Brubaker K (2001) Nest success is not an adequate comparative estimate of avian reproduction. J Field Ornithol 72:27–536

Thompson, FR, Burhans DE, Root B (2002) Effects of point count protocol on bird abundance and variability estimates and power to detect population trends. J Field Ornithol 73:141–150

Van Horne B (1983) Density as a misleading indicator of habitat quality. J Wildl Manag 47:893–901

Vernier L, Pearce J, McKee J, McKenney DW, Niemi G (2004) Climate and satellite-derived land cover for predicting breeding bird distribution in the Great Lakes Basin. J Biogeogr 31:315–331

Vickery PD, Hunter ML, Wells JV (1992) Use of a new reproductive index to evaluate relationships between habitat quality and breeding success. Auk 109:697–705

Williams BK, Nichols JD, Conroy MJ (2002) Analysis and management of animal populations. Academic Press, San Diego