

Research Article

Comparison of Methods for Estimating Density of Forest Songbirds From Point Counts

JENNIFER L. REIDY,¹ *Department of Fisheries and Wildlife Sciences, University of Missouri, 302 Anheuser-Busch Natural Resources Building, Columbia, MO 65211, USA*

FRANK R. THOMPSON III, *United States Department of Agriculture, Forest Service, Northern Research Station, University of Missouri, 202 Anheuser-Busch Natural Resources Building, Columbia, MO 65211, USA*

J. WESLEY BAILEY,² *Department of Fisheries and Wildlife Sciences, University of Missouri, 302 Anheuser-Busch Natural Resources Building, Columbia, MO 65211, USA*

ABSTRACT New analytical methods have been promoted for estimating the probability of detection and density of birds from count data but few studies have compared these methods using real data. We compared estimates of detection probability and density from distance and time-removal models and survey protocols based on 5- or 10-min counts and outer radii of 50 or 100 m. We surveyed singing male Acadian flycatchers (*Empidonax virens*), cerulean warblers (*Dendroica cerulea*), Kentucky warblers (*Oporornis formosus*), Louisiana waterthrushes (*Parkesia motacilla*), wood thrushes (*Hylocichla mustelina*), and worm-eating warblers (*Helminthophila vermivorum*) in bottomland and upland forest across 5 states in the Central Hardwoods Bird Conservation Region during the breeding season in 2007 and 2008. Detection probabilities differed between distance and time-removal models and species detectabilities were affected differently by year, forest type, and state. Density estimates from distance models were generally higher than from time-removal models, resulting from lower detection probabilities estimated by distance models. We found support for individual heterogeneity (modeled as a finite mixture model) in the time-removal models and that 50-m radius counts generated density estimates approximately twice as high as 100-m radius counts. Users should be aware that in addition to estimating different components of detectability, density estimates derived from distance and time-removal models can be affected by survey protocol because some count durations and plot radii may better meet model assumptions than others. The choice of a method may not affect the use of estimates for relative comparisons (e.g., when comparing habitats) but could affect conclusions when used to estimate population size. We recommend careful consideration of assumptions when deciding on point-count protocol and selection of analysis methods. © 2011 The Wildlife Society.[†]

KEY WORDS bottomland forest, central hardwoods, detectability, distance sampling, survey protocol, time-removal sampling, upland forest.

Point-count surveys are a traditional and popular technique used in avian ecology; often, data collected from these surveys are used to estimate abundance or density of bird species (Ralph et al. 1995, Rosenstock et al. 2002). Typically, abundance is reported as an uncorrected count of individuals detected per point and is treated as an index of abundance. The use of counts of birds as an index has received considerable criticism because of the inherent assumption of constant detection probability across time and space (Burnham 1981, Thompson 2002). Substantial evidence exists that many factors can affect detection probability, including time of day and season, weather, breeding status, distance to detected individuals, habitat type, and observer ability (reviewed in

Johnson 2008). Most monitoring efforts have relied on design-based approaches to control factors that affect detection probabilities (Verner 1985, Ralph et al. 1995). Researchers relying on design-based methods attempt to minimize variation in detectability or randomize differences in detectability to minimize bias. Despite well-standardized methodology, most surveys are not well designed and will not be able to account for all variation in detection probabilities, and many factors remain beyond control such that substantial variation can still exist in detection probability (Johnson 1995, Pendleton 1995). Important sources of variation often overlooked and not usually controlled for in study designs include observer ability (particularly problematic for long-term studies) and physical and behavioral attributes of the study species (which can affect likelihood of detection; reviewed in Rosenstock et al. 2002).

Recent advances have made modeling detection probability of point-count data more feasible. Several approaches have been developed to account for detection information, including double-observer sampling (Nichols et al. 2000), double sampling (Bart and Earnst 2002), time-removal sampling (Farnsworth et al. 2002), time-of-detection sampling

Received: 7 March 2010; Accepted: 8 October 2010

[†]This article is a US Government work and, as such, is in the public domain in the United States of America.

¹E-mail: jennifer.reidy@gmail.com

²Present Address: Forest Wildlife and Populations Research Group, Minnesota Department of Natural Resources, 1201 East Highway 2, Grand Rapids, MN 55744, USA.

(2007c), and distance sampling (Buckland et al. 2001), each method having its own assumptions. Despite general agreement that detection probability is rarely perfect or constant, many field studies may not be able to meet assumptions to estimate detection probability and, even if they do, models may not perform well for certain species or survey protocols (Allredge et al. 2008, Johnson 2008, Nichols et al. 2009). Beyond model assumptions, selection of a method has implications to the count or survey protocol used in the field. The former 2 techniques require substantially more field work than do typical point counts; whereas the latter 3 approaches can more easily be accommodated within the framework of standard point counts by simply collecting time and distance data. However, the additional collection of ancillary data can affect an observer's ability to detect multiple or hard-to-detect species and is highly subject to error (Allredge et al. 2008, Johnson 2008, Nichols et al. 2009).

Additionally, there is confusion over which of these newer methods to use for analysis of count data to estimate detection probabilities. Each method models detection probability based on different information associated with observations and is best chosen by a thorough examination of species biology, study design, and examining how well the data meets model assumptions. Detection probability comprises 2 broad components. Availability (p_a) is the probability a bird is present and gives the appropriate cue, whether visual or aural (Marsh and Sinclair 1989, McCallum 2005, Laake et al. 2008). The other component of detection probability, detectability (p_d), has been termed perception bias (Marsh and Sinclair 1989, Laake et al. 2008) and is the probability the observer detects the bird given that the bird is available, which can be divided into components of conspicuousness and abundance of cues (McCallum 2005). Double sampling and time-removal sampling are better at estimating p_a , whereas double-observer and distance sampling are better suited to estimating p_d . No current method incorporates all components of detection probability easily (McCallum 2005, Nichols et al. 2009). Work is underway using hybrids of these methods (Nichols et al. 2009), such as including distance as a covariate in time-removal or mark-recapture models (Alpizar-Jara and Pollock 1996), combining time-of-detection and double observer sampling (Stanislav et al. 2010) or mark-recapture and distance modeling (Laake et al. 2008), or incorporating availability as a divisor in distance models (Diefenbach et al. 2007, Gale et al. 2009).

Because estimation of detection probability is new, there is little information about how survey protocol affects detectability. Several studies have explored the effect of survey protocol on unadjusted estimates (Jones et al. 2000, Norvell et al. 2003). Few have evaluated the effect of survey protocol on detection probabilities, which in turn affect density estimates. The little research that has been done has focused on distance sampling, and conclusions showed that count duration and maximum radius used affected estimates (Kissling and Garton 2006, Cimprich 2009). To our knowledge, no study has evaluated the effect of count

duration or outer radius on detection probability using time-removal models.

We worked with partners of the Central Hardwoods Joint Venture to implement a point-count survey targeting Acadian flycatchers (*Empidonax vireescens*), cerulean warblers (*Dendroica cerulea*), Kentucky warblers (*Oporornis formosus*), Louisiana waterthrushes (*Parkesia motacilla*), wood thrushes (*Hylocichla mustelina*), and worm-eating warblers (*Helminthos vermivorum*) in the Central Hardwoods Bird Conservation Region (CHBCR). These birds are priority species for the joint venture (due to declining populations or having a large percentage of their population within CHBCR, Fitzgerald et al. 2003) but also represent a suite of forest-breeding songbirds that cover a gradient of ecological niches and likely vary in their song rate and volume and height of singing perches. Our objectives were to: 1) survey density of the 6 priority species on selected sites in the CHBCR in such a way that we could compare multiple point-count protocols and analysis methods; 2) identify the best distance and time-removal models for estimating density considering forest type, year, state (as a surrogate for observers), distance, and heterogeneity effects; and 3) compare estimates from the best models for each method using different count durations and radii around points.

STUDY AREA

We conducted our study in forests on public lands in Arkansas, Illinois, Kentucky, Missouri, and Tennessee within the CHBCR (Fig. 1). The CHBCR comprised >3 million ha of rolling hills covered primarily with hardwood forests interspersed with glades and woodlands and dissected by deep river valleys (Fitzgerald et al. 2003). Other landforms within the CHBCR included steep-sided ridges and hills, karst terrain, gently rolling lowland plains, and bottomlands along major rivers, with associated terraces and meander scars (McNab and Avers 1994). The Mississippi River floodplain bisected the CHBCR between Illinois and Missouri into 2 regions: the Ozark Highlands and Boston Mountains to the west and the Interior Low Plateaus to the east (U.S. North American Bird Conservation Initiative Committee 2000).

METHODS

To address objectives ancillary to this study, we considered forested lands that represented potential habitat for cerulean warblers. We used a Geographic Information System (GIS) to intersect a map of forest derived from National Landcover Data (Homer et al. 2004) with public lands in the CHBCR. Because of the extensive forested lands in Missouri, we further constrained sampling to within forest breaks landtype associations within the Meramec River Hills and Current River Hills ecological subsections (Nigh and Schroeder 2002), because this is believed to represent the best potential cerulean warbler habitat in the state (Rosenberg et al. 2000). We then classified forest as upland or bottomland using a digital elevation model. Wooded bottomland forest consisted of mixed hardwoods stands often with American sycamore (*Plantanus occidentalis*) dominating the canopy

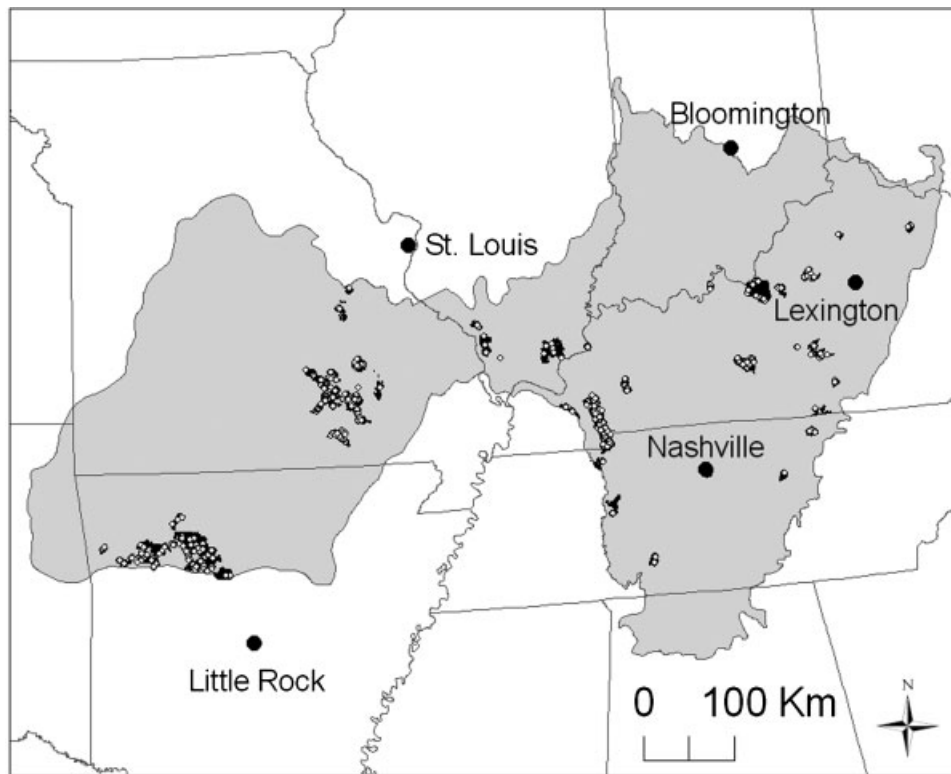


Figure 1. Survey area (black areas) and point counts (white circles) in public forested land we used in a survey of abundance of 6 priority species in the Central Hardwoods Bird Conservation Region (gray area), May to June 2007–2008.

along riparian edges. Oak (*Quercus* spp.)-hickory (*Carya* spp.) forest dominated upland habitats as well as occasional stands of mixed oaks and shortleaf pine (*Pinus echinata*).

Point Counts

We conducted counts at points along transects starting from random locations along secondary and unimproved roads in upland habitats and rivers accessible by canoes in bottomland habitats. We located points in upland forest at 250-m intervals on U-shaped transects that departed perpendicularly and started 30–50 m from a road. Roads had a dirt surface and very little traffic and were narrow with tree canopy overhead. We oriented transects in bottomland forest perpendicular to the river but transects were only 1–2 points long because bottomlands were narrow.

We conducted counts from 23 May to 30 June 2007–2008; from sunrise to 1000 hours; and during periods of no or low wind, no or light precipitation, and temperatures $>10^{\circ}\text{C}$. At each point we recorded the state, site, transect, point, observer, date, start time, Universal Transverse Mercator (UTM) coordinates, and weather condition (temperature, wind speed, cloud cover, and precipitation). For each singing male detected, we recorded the time of initial detection and the distance and direction to the bird from the observer. We measured distances using a Bushnell Yardage Pro laser range-finder (Bushnell, Overland Park, KS). However, when topography or vegetation density made use of the range-finder difficult, we estimated distances ($<10\%$ of detections). Primary observers were trained in distance measurements

prior to surveying by a biologist from each state (such that observers from each state would corroborate each other).

Distance Modeling

We estimated density based on the distance to detected individuals at points, assuming detectability decreases with increasing distance between the observer and the detected individual (Buckland et al. 2001). Assumptions for distance sampling are: 1) objects at the point are always detected, 2) objects are detected at their initial location (no movement in response to observer or otherwise), and 3) distances are measured accurately (Buckland et al. 2001). Distance sampling addresses one component of detection probability, detectability (p_d). We evaluated the effect of covariates on detectability for each species in multiple covariate distance sampling (MCDS) engine within Program DISTANCE (Version 6.0, www.ruwpa.st-and.ac.uk/distance/, accessed 26 Oct 2010) based on data collected at each point. We truncated the data at the greatest 10% of recorded distances as recommended by Buckland et al. (2001). We fit models with half-normal key function and cosine series expansion and hazard-rate key function with simple polynomial series expansion, and we used model selection criteria (defined below) to determine the best fit to the data for each species. We hypothesized that detectability could be affected by year, forest type, and observer. However, because we had many observers, many of whom conducted few surveys, we evaluated observer effects using state as a surrogate because observers were trained and supervised by each state.

We compared support for models incorporating year, forest type, and state as covariates evaluated singly and additively, as well as a null model with a global detection function. We assessed model fit with chi-square goodness-of-fit and evaluated model support using Akaike's Information Criterion adjusted for small sample sizes (AIC_c); the most supported model was the model with the lowest AIC_c ($\Delta AIC_c = 0$). We report ΔAIC_c for each species for each model and detection probability for the most supported model. Because Program DISTANCE does not allow multiple levels of stratification, we partitioned points by forest type to enable estimation of density by forest type and year. If year was supported in the covariate analysis above, we estimated density with year-level detection probabilities; otherwise, we estimated density using a global detection function. We compared model fit using the half-normal, hazard-rate, and uniform key functions with the cosine, simple polynomial, and hermite polynomial series expansions, respectively. We estimated density (singing males/ha) for each species by year and forest type.

Time-Removal Models

We estimated detection probability and density using a removal model framework (Farnsworth et al. 2002) in Program MARK (Version 5.1, www.cnr.colostate.edu/~gwhite/mark/mark.htm, accessed 29 Jan 2010) with Huggins closed-capture and Huggins full-heterogeneity models (Huggins 1989). Huggins models are based on conditional likelihood theory, where individuals not detected (i.e., encounter history of 000 for 3 time intervals) are not included in analysis. Because individuals not detected are conditioned out of the likelihood, these models are able to analyze data incorporating covariates. Time-removal models estimate density based on the time interval in which an individual is detected. Detection probability is modeled as the decline in the number of new detections over sequential time intervals (Farnsworth et al. 2002). Models incorporating heterogeneity allow detection to vary across groups, typically two: those easy to detect and those difficult to detect. Model assumptions are: 1) population of interest is closed, 2) there is no double-counting, 3) all members of group 1 are detected in the first interval, 4) all members of group 2 not detected in the first interval have a constant detection probability, and 5) for limited-radius counts, observers accurately assign birds to the appropriate radius (Farnsworth et al. 2002). Time-removal models address another component of detection probability, availability (p_a). We created encounter histories by placing detections of each species at each point in five 2-min intervals and truncated the data to a maximum observation distance of 100 m. We first evaluated support for intercept-only models with and without heterogeneity, then used the most supported model to build models incorporating the covariates year, forest type, state, and distance singly and additively, as well as a null model with no covariates. We report ΔAIC_c (model support) and detection probability for each species for each model.

We created encounter histories for each species based on different survey protocols to evaluate the effect of count duration and plot radius on density estimation. We partitioned the full dataset by 2 count durations (5 min and 10 min) and 2 plot radii (50 m and 100 m). Whereas we created the full model above by binning count data into five 2-min intervals, we binned the 5-min count (based on the first 5 min of the survey) into five 1-min intervals. For each of these count durations, we used a truncation radius of 100 m or 50 m (both commonly used outer radii). For each survey protocol, we evaluated the same model set as above for the full model (100 m and 10 min). We report density (singing males/ha) estimated using the most supported model for each survey protocol for each species in both years and forest types. Model selection results are reported for the full dataset (100-m radius 10-min counts) (Fig. S1, available online at www.onlinelibrary.wiley.com).

RESULTS

During 2007 and 2008, 56 observers completed 310 point transects comprising 2,771 point counts (Table 1; Fig. 1). The dataset included 1,246 and 1,132 point counts in upland habitat and 97 and 296 point counts in bottomland habitat in 2007 and 2008, respectively.

Data truncation for distance models resulted in a maximum distance of 100 m, 90 m, 100 m, 109 m, 150 m, and 100 m for Acadian flycatchers, cerulean warblers, Kentucky warblers, Louisiana waterthrushes, wood thrushes, and worm-eating warblers, respectively. Distance models fit the data well (Fig. 2); all top models had coefficients of variation $< 6\%$ for the 4 most abundant species and $< 15\%$ for the 2 least abundant. Cerulean warbler detections appeared to be subject to heaping errors (Fig. 2), either due to distance measurement error or movement away from the point. Wood thrushes had a broad detection shoulder, representing high detection probabilities to larger distances, whereas Louisiana waterthrushes displayed a sharp decline or narrow shoulder in detections with increasing distance (Fig. 2). We were not able to evaluate state as a covariate for any species using distance models because the shape function differed or there were too few observations in some states for models to converge. There was strong support for effects of year, forest type, or year + forest type for all species except cerulean warbler. There was substantial model selection uncertainty for cerulean warblers, probably due to a small sample size (Table 2). However because for most species there were few or no competing models with $\Delta AIC_c < 2$, and if there was, density estimates were very similar, we only present estimates from the most-supported model.

For time-removal models, 67–94% of detections occurred within the first 5 min and 36–75% of detections were within 50 m for all species except wood thrush, for which 21–37% of detections occurred within 50 m. Intercept models with heterogeneity (M_h) were supported over models without heterogeneity for 5 species ($\Delta AIC_c > 3.5$ for models without heterogeneity) and both models received almost equal support for 1 species (Louisiana waterthrush), so heterogen-

Table 1. Number of points and individuals detected during point counts of 6 priority songbirds in 2007–2008 in the Central Hardwoods Bird Conservation Region by forest type (upland and bottomland), state, and year. We did not conduct bottomland forest surveys in Tennessee.

	Upland						Bottomland						Total							
	Arkansas		Illinois		Kentucky		Missouri		Tennessee		Arkansas		Illinois		Kentucky		Missouri			
	2007	2008	2007	2008	2007	2008	2007	2008	2007	2008	2007	2008	2007	2008	2007	2008	2007	2008		
No. of points	439	46	41	0	256	324	460	568	50	194	0	51	0	24	18	90	79	131	1,343	1,428
No. of detections	92	20	29	0	97	230	230	331	67	317	15	15	32	32	11	72	31	29	557	1,046
Acadian flycatcher	18	3	0	0	8	26	32	42	4	3	0	0	0	0	1	9	42	51	105	134
Cerulean warbler	24	8	5	0	88	138	74	87	19	43	9	9	24	7	27	34	41	251	377	377
Kentucky warbler	6	4	14	0	10	33	16	17	0	12	12	12	4	1	1	21	25	72	72	175
Louisiana waterthrush	41	4	9	0	207	197	52	64	52	60	5	5	15	12	12	43	21	43	394	431
Wood thrush	99	5	27	0	15	131	210	184	11	130	3	3	22	4	4	26	9	4	375	505
Worm-eating warbler																				

city was included in all models with covariates. Detection probabilities for all models were generally high, resulting in estimates almost the same as the number of detections and narrow confidence intervals. State and forest type were included in the top model for 5 species; year and distance were included in the top model for 3 species (Table 2). Because for most species there were few or no competing models with $\Delta AIC_c < 2$, and if there was, density estimates were similar, we only present estimates from the most supported model and estimates are conditional on the best model.

Detection probabilities varied widely between species and methods, ranging from 0.08 for Louisiana waterthrush to approximately 1.00 for several species (Fig. 3). Distance models produced lower detection probabilities for all species. Louisiana waterthrushes, based on the sharp decline of detections with increasing distance, had a very low detection probability particularly for bottomland forest detections.

On average, distance models produced the highest density estimates, but they were similar to estimates from time-removal models based on a 50-m radius. Time-removal models based on 50-m radius estimated densities about twice as great as models based on 100-m radius for 5 of 6 species (Fig. 4). Density estimates from time-removal models differed more between 50-m and 100-m plot radii than between 5-min and 10-min counts. We generally detected about twice as many individuals within 100 m than within 50 m for all species (1.56–2.74, except Louisiana waterthrush [0.3–2.1] and wood thrush [2.8–4.8]), but density was often twice as high for the smaller plot. Additionally, density estimates averaged 27% (range –9.0–97%) higher for 10-min counts than for 5-min counts. Differences in density between plot radii were mostly the result of differences in the number of detections and area sampled because detection probabilities were similarly high for most survey protocols. There was a tendency for confidence limits for density estimates from time-removal models to be wider for surveys based on 50-m radii than 100-m radii (Fig. 4).

DISCUSSION

We compared density estimates derived from the most supported distance and time-removal models using a standardized design. Comparison of estimates from different models can be of limited value when the true density is unknown. However, given limited application of time-removal models in large-scale survey efforts, we believe our application and comparison of alternative survey protocols is informative. We used a few target birds, as recommended by Alldredge et al. (2008) and Nichols et al. (2009), where we could use collected data in both distance and time-removal models. Based on measures such as model fit or support and the precision of estimates, we successfully applied these models to estimate bird densities. However, the estimates varied, sometimes greatly, when we used different subsets of the data representing different survey protocols in the models.

Our hypothesis that detection probability varied by year and forest type was supported in both modeling approaches. These results are similar to Norvell et al. (2003), who found

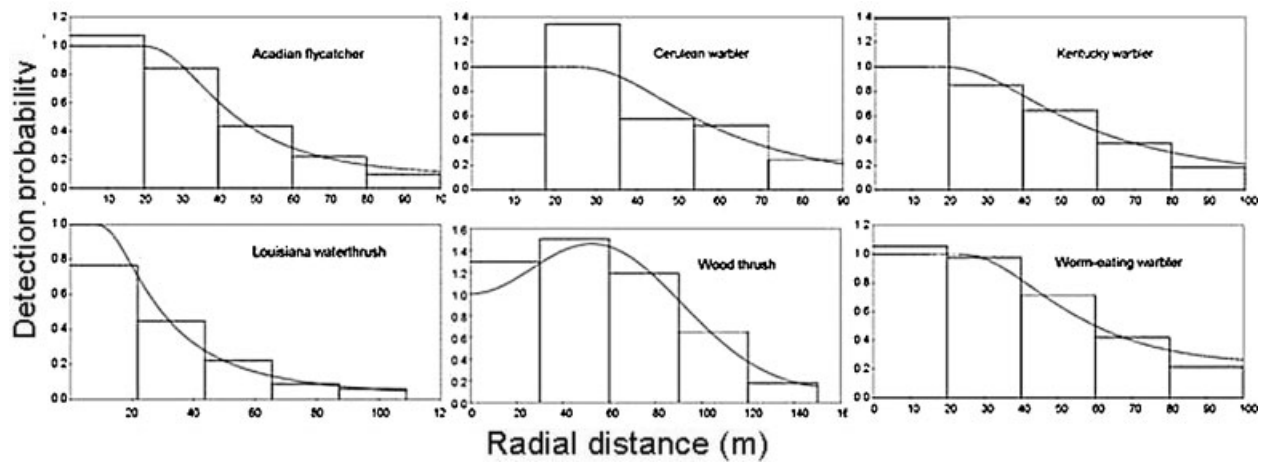


Figure 2. Detection-probability histograms from top distance models for Acadian flycatcher, cerulean warbler, Kentucky warbler, Louisiana waterthrush, wood thrush, and worm-eating warbler detected in the Central Hardwoods Bird Conservation Region during early summer 2007–2008. We truncated detections at the largest 10% of distances for each species.

detection probabilities varied widely across years. Whereas Norvell et al. (2003) did not include a habitat type as a variable, they noted that their sampling may have failed to address habitat-specific detection probabilities. Year is not a very informative variable and is often a proxy for unexplained variation. Possible explanations for a year effect include hiring more skilled observers 1 yr than the other, number of observers used each year, changes in weather patterns, and differences in species' detectability based on mating and nesting status (birds may have experienced higher nest pre-

paration 1 yr leading to differences in nesting phenology and song rates). We think that detectability varied by forest type due to terrain and topography and by species-specific habitat preferences and use of space.

Our observer surrogate was supported in time-removal models for 5 of 6 species. Although states represented groups of observers trained together, an observer effect may have been obscured by mixing it with another potential source of variation (states may have acted as groups for more than just observers). However, we have no a priori reason to speculate

Table 2. Number of parameters (K) and model support (difference in Akaike's Information Criteria from the top model [ΔAIC_c]) using distance and time-removal models for 6 priority species detected during point counts in the Central Hardwoods Bird Conservation Region from May to June 2007 and 2008. We evaluated the effect of year and forest type (singly and additively) in distance models, and year, forest type, state, and distance (singly and additively) in time-removal models. We based models on 10-min counts with the largest 10% of distances truncated for distance models and with a 100-m fixed radius for time-removal models. The null model used a global detection function.

	K	ΔAIC_c					
		Acadian flycatcher	Cerulean warbler	Kentucky warbler	Louisiana waterthrush	Wood thrush	Worm-eating warbler
Distance							
Null	2	13.53	0.00	10.61	4.31	27.09	25.88
Year	4	6.75	1.21	10.30	6.78	1.73	0.00
Forest type	4	0.00	1.14	12.74	0.00	19.96	18.10
Year + forest type	5	1.45	3.09	0.00	2.22	0.00	1.24
Time-removal ^a							
Null	3	64.02	5.46	30.10	19.01	24.81	54.64
Year	4	24.70	7.43	22.48	18.55	15.54	11.74
Forest type	4	61.05	6.89	31.95	12.18	23.05	47.45
State	7	20.69	4.00	2.93	16.90	21.06	18.00
Distance	4	52.30	4.88	31.99	7.61	23.32	55.25
Year + forest type	5	23.08	8.90	20.79	12.83	17.20	9.79
Year + state	8	19.95	5.57	0.00	18.92	10.63	5.19
Year + distance	5	17.09	6.87	24.34	7.22	15.58	13.34
Forest type + state	8	18.47	3.78	6.11	14.25	22.43	0.00
Forest type + distance	5	49.32	6.38	32.29	0.52	21.81	48.15
State + distance	8	16.50	1.75	2.49	5.77	20.47	16.94
Year + forest type + state	9	0.00	5.77	4.24	14.37	4.53	1.46
Year + forest type + distance	6	14.68	8.40	21.91	0.00	17.46	10.42
Year + state + distance	9	12.65	2.72	0.01	7.34	11.62	5.56
Forest type + state + distance	9	1.71	0.00	6.69	7.48	0.00	0.99
Year + forest type + state + distance	10	14.50	2.01	4.58	4.50	1.43	3.36

^a All models include heterogeneity.

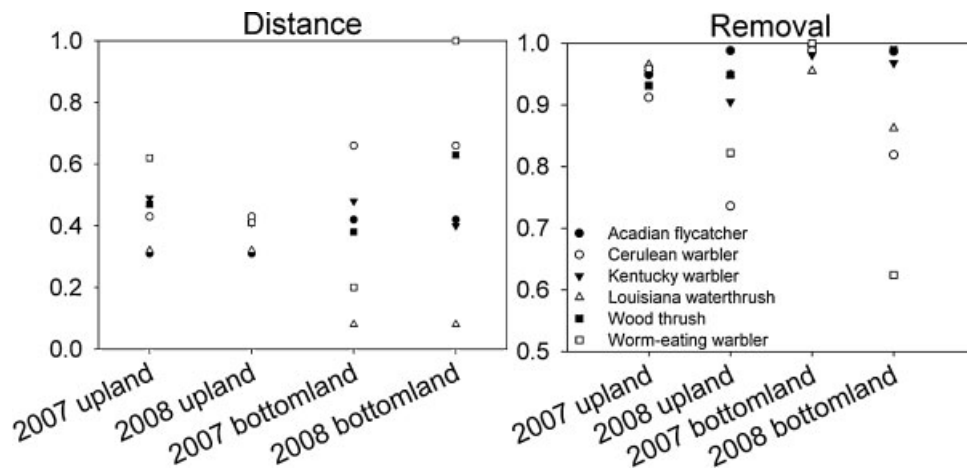


Figure 3. Estimated detection probabilities based on most supported covariates in distance and time-removal models by forest type, state, and year from point-count surveys in the Central Hardwoods Bird Conservation Region, 2007–2008; scales vary between graphs. We truncated distance models at 10% of the largest distances for each species. We based detection probabilities for time-removal models on 10-min counts and 100-m fixed radius. We did not conduct bottomland surveys in Tennessee.

that detectability differed by state for any other reason (forest type, terrain, weather). Other studies have provided more direct evidence of observer effects (Farnsworth et al. 2002, Norvell et al. 2003, Simons et al. 2007, Gale et al. 2009) and observer bias is an issue that will continue to plague count-based surveys, especially large-scale monitoring projects. Small-scale studies can more easily control for observer bias. Large-scale studies should attempt to implement a more rigorous system to control for and identify sources of observer bias, including longer training periods and testing. Design-based methods such as randomizing observers can minimize bias but do not eliminate variability due to observers. Surveys should control for as many sources of variation in detection probability as possible, but it is unlikely all sources of variation can be accounted for in the design.

Distance models fit the data well and detection probabilities ranged from 0.08 to 0.66. Thompson and La Sorte (2008) and Norvell et al. (2003) found similar detection probabilities for songbirds from distance models. Simulations show distance models perform well when all assumptions are met (Efford and Dawson 2009). However, meeting model assumptions in the field is difficult and errors are pervasive in distance measurements (Allredge et al. 2007a, 2008; Simons et al. 2007; Johnson 2008). Despite this, ignoring the effects of distance on detectability certainly leads to biased estimates of detection probability and therefore to density (Efford and Dawson 2009). Conclusions from real data comparing count data to known populations are conflicting. Cimprich (2009) concluded black-capped vireo (*Vireo atricapilla*) density was overestimated using distance sampling compared to a known, color-marked population. Gale et al. (2009) concluded density from count data was biased low for 4 species and high for 1 compared with known abundance of 8 tropical forest birds estimated with distance sampling and found incorporating a measure of availability did not improve accuracy of estimates.

Time-removal models with heterogeneity were supported over those without, similar to Farnsworth et al. (2002).

Including covariates removed some of the heterogeneity, but not all, as evidenced by the higher density within 50 m than within 100 m. Unmodeled heterogeneity in mark-recapture data, such as time-removal models, is a well-documented problem in the line-transect literature and may result in negative bias greater than that resulting from violation of failing to detect all individuals on the line or point using distance models (Laake et al. 2008). Controlling for and modeling as many sources of variation can reduce heterogeneity-related bias, but all sources are unlikely to be accounted for (Laake 1999, Borchers et al. 2006).

Time-removal models produced estimates of detection probabilities (0.62–1.00) that were greater on average than distance models and similar to Farnsworth et al. (2002) and Thompson and La Sorte (2008). These results suggested availability is high, particularly at closer distances. Dawson et al. (1995) estimated the probability of detecting a bird given it was known to be present at a point ranged from 47% to 85% within the first 5 min and 66% to 92% within the first 10 min of a 20-min count. Diefenbach et al. (2007) estimated availability separately for 2 grassland songbirds and found availability to be substantially lower than 1 (0.50 and 0.21, respectively, for a 10-min aural count).

Time-removal methods model the decline in new detections across time, so it is unsurprising that detection probabilities were high given that our target birds have high singing frequencies and small territories or loud vocalizations. We expect that detection probabilities would be lower for species with lower singing frequencies or larger territories. Using time to estimate detection probability has its own shortcomings, as experimental counts have shown observers often double count individuals, record detections in the wrong time interval, and count detections from outside the outer radius as in Allredge et al. (2007b). Bias due to spatial heterogeneity in detection probability can be reduced by incorporating distance to detected birds in time-removal models (Efford and Dawson 2009). In our study, distance received some support as a covariate in time-removal models

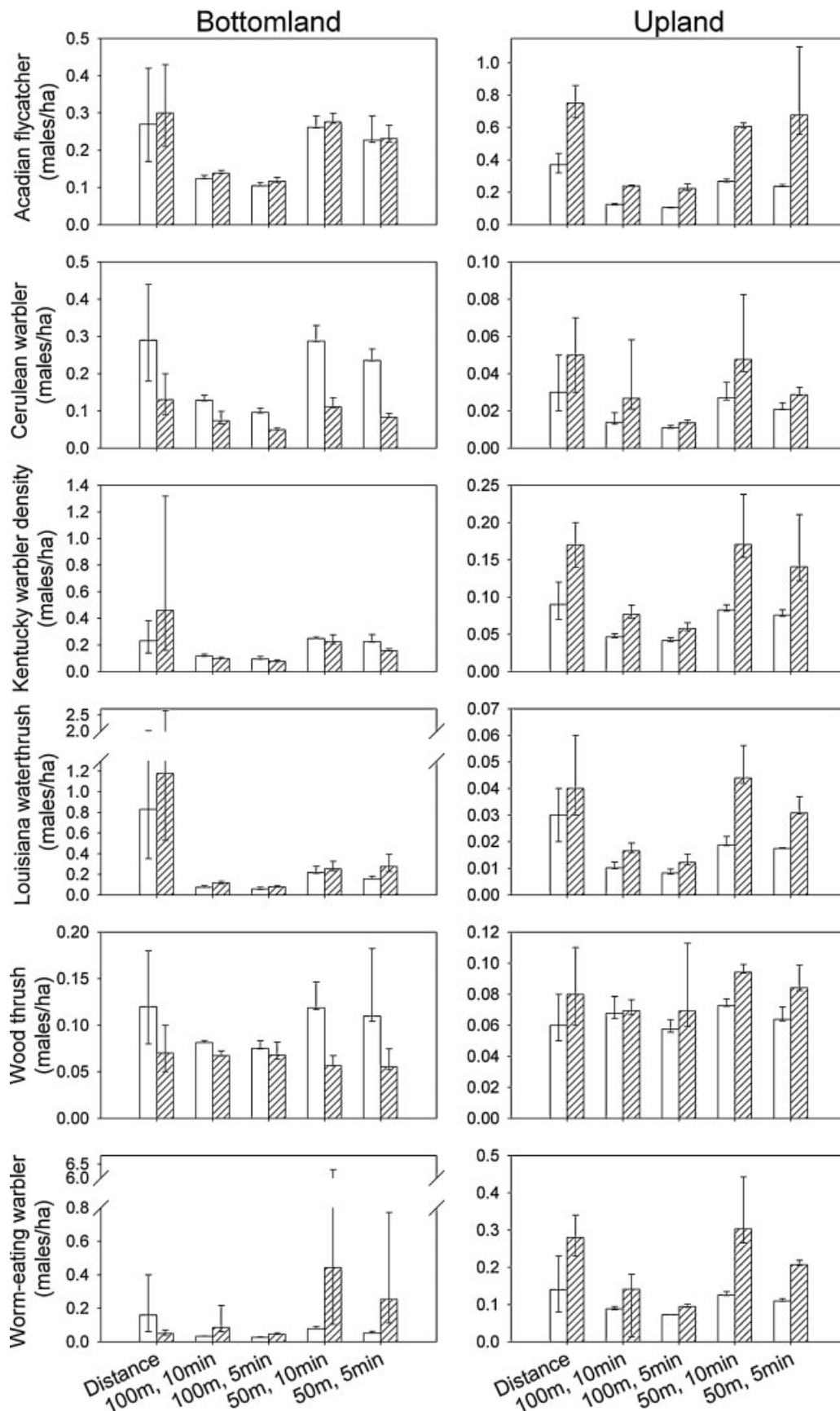


Figure 4. Density estimates and 95% confidence intervals from top distance model and time-removal models based on different survey protocols (100 m and 10 min, 100 m and 5 min, 50 m and 10 min, and 50 m and 5 min) in bottomland (left) and upland (right) forests in 2007 (white bars) and 2008 (hatched bars). Scales vary among figures.

for all species. Examination of simulated counts demonstrate that if detection probabilities are high, then population estimates show little bias using numerous methods to estimate detection probability, but large bias exists when heterogeneity related to distance exists (Efford and Dawson 2009).

Our evaluation of survey protocols highlights the importance of count duration and plot radius to density estimation because they define the population of interest (more detections are expected over a longer duration and an increased area). Cimprich (2009) concluded a 3-min count was more accurate at estimating density of black-capped vireos than were 5-min or 6-min counts and found that longer counts produced higher density estimates. Lee and Marsden (2008) also showed a positive association between density estimates and count duration. Such results suggest individuals may have been double-counted or that movement of birds occurred during the survey, both violations of the model assumptions. We found a similar trend with higher density estimates for 10-min counts than for 5-min counts. Jones et al. (2000) found observers overestimated density of cerulean warblers within a 50-m radius and underestimated density at 100-m or unlimited radius (on unadjusted counts) compared to a known population. Similarly, Simons et al. (2007) demonstrated observers overestimated density within 50 m and underestimated density in unlimited radius counts using experimental surveys. Our density estimates were often twice as high for the 50-m plot as for the 100-m plot. Differences in our estimates between the 2 plot radii suggests either errors in distance estimation that were dependent on distance or that our models did not adequately capture spatial heterogeneity, even though many included distance as a covariate.

The appropriate design to accurately estimate density will likely require preliminary investigation to establish protocol criteria such as count duration and maximum radius to use (Barker et al. 1993, Jones et al. 2000, Cimprich 2009). We included only detections of singing males in our analysis because we conducted surveys in forested terrain where we were unlikely to detect individuals by sight except at very close distances. We also focused on singing males to minimize heterogeneity in detectability resulting from differences in detectability among male and female birds. Studies that combine visual and auditory detections may arrive at different conclusions. Our investigation highlights the discrepancy between density estimates using different survey protocols and the potential problems of comparing estimates based on different protocols or models. Our study represents a realistic dataset collected from multiple observers across a large geographic area with varying levels of training in species identification, distance estimation, and general field techniques. Even if we could better standardize all the above, there still exist differences in observers' hearing ability and terrain, as well as individual species' biology (song rate, singing volume).

We agree with concerns raised by Johnson (2008) and Efford and Dawson (2009) concerning the reliability of methods used to estimate absolute abundance. Nevertheless we could fit distance and time-removal models

to typical point count data, and these methods represent one way to address effects of covariates on detection probability. Based on our findings, we recommend using time-removal models for species with a constant singing rate, and do not recommend using this method for birds that sing in bouts. Distance sampling may be best applied to studies conducted in more open habitats, where observations are composed of both visual and auditory detections and distance estimates are more accurate, and for species that use space randomly. Using a combination of mark-recapture and distance sampling, as developed by Laake (1999) and Borchers et al. (2006), to estimate each component of the detection probability shows promise but needs to be evaluated with known abundance.

MANAGEMENT IMPLICATIONS

We recommend that detection probability be addressed in some manner, whether by design-based methods (Johnson 2008), statistical or model-based methods that consider factors affecting detection probability (Link and Sauer 1998, Johnson 2008, Thompson and La Sorte 2008), or modeling approaches that attempt to estimate detection probability. Given that density estimates can vary based on the survey protocol and method used to estimate density, investigators and managers should be careful when comparing estimates. Relative relationships of species' density among forest types and years were mostly consistent across methods and survey protocols, so these models may still provide an option for addressing factors affecting detectability when interested in relative comparisons among treatments or habitats. When choosing a modeling approach and survey protocol we suggest investigators and managers consider species biology (e.g., territory size, vocal range, singing rates) and model assumptions.

Surveys that collect exact time and distance data offer maximum flexibility for analysis, allowing data to be partitioned to evaluate different approaches based on species biology, habitat characteristics, and additional covariate data such as observers. Finally, we believe that targeting a few species of interest helps eliminate errors and allows time to collect data required for each detection during the count. Given the large effect that the modeling method and survey protocol had on density estimates, we suggest investigators collect data that allow evaluation of multiple approaches and their assumptions. Additionally, we recommend investigators report both adjusted and unadjusted count data.

ACKNOWLEDGMENTS

We thank J. Fitzgerald for her efforts to promote and facilitate this project through the Central Hardwoods Joint Venture and for reviewing this manuscript. K. Cornell-Duerr provided assistance with analyses early in the study. We thank the following joint venture partners for conducting bird surveys for this project: Arkansas Game and Fish Commission, Southern Illinois University, Kentucky Department of Fish and Wildlife Resources, Missouri Department of Conservation, and Tennessee Wildlife Resources Agency. This project was funded by

the Central Hardwoods Joint Venture through the American Bird Conservancy and the U.S. Department of Agriculture Forest Service.

LITERATURE CITED

- Allredge, M. W., T. R. Simons, and K. H. Pollock. 2007a. A field evaluation of distance measurement error in auditory avian point count surveys. *Journal of Wildlife Management* 71:2759–2766.
- Allredge, M. W., T. R. Simons, K. H. Pollock, and K. Pacifici. 2007b. A field evaluation of the time-of-detection method to estimate population size and density for aural avian point counts. *Avian Conservation and Ecology* 2:13–24.
- Allredge, M. W., K. H. Pollack, T. R. Simons, J. A. Collazo, and S. A. Shriner. 2007c. Time-of-detection method for estimating abundance from point count surveys. *Auk* 124:653–664.
- Allredge, M. W., K. Pacifici, T. R. Simons, and K. H. Pollock. 2008. A novel field evaluation of the effectiveness of distance and independent observer sampling to estimate aural avian detection probabilities. *Journal of Applied Ecology* 45:1349–1356.
- Alpizar-Jara, R., and K. H. Pollock. 1996. A combination line transect and capture-recapture sampling model for multiple observers in aerial surveys. *Environmental and Ecological Statistics* 3:311–327.
- Barker, R. J., J. R. Sauer, and W. A. Link. 1993. Optimal allocation of point-count sampling effort. *Auk* 110:752–758.
- Bart, J., and S. Earnst. 2002. Double sampling to estimate density and population trends in birds. *Auk* 119:36–45.
- Borchers, D. L., J. L. Laake, C. Southwell, and C. G. M. Paxton. 2006. Accommodating unmodelled heterogeneity in double-observer distance sampling surveys. *Biometrics* 62:372–378.
- Buckland, S. T., D. R. Anderson, K. P. Burnham, J. L. Laake, D. L. Borchers, and L. Thomas. 2001. Introduction to distance sampling: estimating abundance of biological populations. Oxford University Press, Oxford, United Kingdom.
- Burnham, K. P. 1981. Summarizing remarks: environmental influences. *Studies in Avian Biology* 6:324–325.
- Cimprich, D. A. 2009. Effect of count duration on abundance estimates of black-capped vireos. *Journal of Field Ornithology* 80:94–100.
- Dawson, D. K., D. R. Smith, and C. S. Robbins. 1995. Point count length and detection of forest Neotropical migrant birds. Pages 35–43 in C. J. Ralph, J. R. Sauer, and S. Droege, editors. *Monitoring bird populations by point counts*. U.S. Department of Agriculture, Forest Service General, Technical Report PSW-GTR-149, Pacific Southwest Research Station, Albany, California, USA.
- Diefenbach, D. R., M. R. Marshall, J. A. Mattice, and D. W. Brauning. 2007. Incorporating availability for detection in estimates of bird abundance. *Auk* 124:96–106.
- Efford, M. G., and D. K. Dawson. 2009. Effect of distance-related heterogeneity on population size estimates from point counts. *Auk* 126:100–111.
- Farnsworth, G. L., K. H. Pollock, J. D. Nichols, T. R. Simons, J. E. Hines, and J. R. Sauer. 2002. A removal model for estimating detection probabilities from point-count surveys. *Auk* 119:414–425.
- Fitzgerald, J. A., G. Wathen, M. D. Howery, W. P. Lisowsky, D. F. McKenzie, and D. N. Pashley. 2003. The Central Hardwoods Joint Venture Concept Plan. Tennessee Wildlife Resources Agency, USA.
- Gale, G. A., P. D. Round, A. J. Pierce, S. Nimnuan, A. Pattanavibool, and W. Y. Brockelman. 2009. A field test of distance sampling methods for a tropical forest bird community. *Auk* 126:439–448.
- Homer, C., C. Haug, L. Yang, B. Wylie, and M. Coan. 2004. Development of a 2001 national landcover database for the United States. *Photogrammetric Engineering and Remote Sensing* 70:829–840.
- Huggins, R. M. 1989. On the statistical analysis of capture experiments. *Biometrika* 76:133–140.
- Johnson, D. H. 1995. Point counts of birds: what are we estimating? Pages 117–123 in C. J. Ralph, J. R. Sauer, and S. Droege, editors. *Monitoring bird populations by point counts*. U.S. Department of Agriculture, Forest Service General, Technical Report PSW-GTR-149, Pacific Southwest Research Station, Albany, California, USA.
- Johnson, D. H. 2008. In defense of indices: the case of bird surveys. *Journal of Wildlife Management* 72:857–868.
- Jones, J., W. J. McLeish, and R. J. Robertson. 2000. Density influences census technique accuracy for cerulean warblers in eastern Ontario. *Journal of Field Ornithology* 71:46–56.
- Kissling, M. L., and E. O. Garton. 2006. Estimating detection probability and density from point count surveys: a combination of distance and double-observer sampling. *Auk* 123:735–752.
- Laake, J. L. 1999. Distance sampling with independent observers: reducing bias from heterogeneity by weakening the conditional independence assumption. Pages 137–148 in G. W. Garner, S. C. Amstrup, J. L. Laake, B. F. J. Manly, L. L. McDonald, and D. G. Robertson, editors. *Marine mammal survey and assessment methods*. Balkema, Rotterdam, The Netherlands.
- Laake, J., M. J. Dawson, and J. Hone. 2008. Visibility bias in aerial survey: mark-recapture, line-transect or both? *Wildlife Research* 35:299–309.
- Lee, D. C., and S. J. Marsden. 2008. Adjusting count period strategies to improve the accuracy of forest bird abundance estimates from point transect distance sampling surveys. *Ibis* 150:315–325.
- Link, W. A., and J. R. Sauer. 1998. Estimating population change from count data: applications to the North American Breeding Bird Survey. *Ecological Applications* 8:258–268.
- 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.
- McCallum, D. A. 2005. A conceptual guide to detection probability for point counts and other count-based survey methods. Pages 754–761 in C. J. Ralph and T. D. Rich, editors. *Bird conservation implementation and integration in the Americas: proceedings of the Third International Partners in Flight Conference*. U.S. Department of Agriculture, Forest Service General, Technical Report PSW-GTR-191, Pacific Southwest Research Station, Albany, California, USA.
- McNab W. H., and P. E. Avers. editors. 1994. *Ecological subregions of the United States: section descriptions*. U.S. Department of Agriculture, Forest Service Publication WO-WSA-5, Washington, D.C., USA.
- 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. *Auk* 117:393–408.
- Nichols, J. D., L. Thomas, and P. B. Conn. 2009. Inferences about landbird abundance from count data: recent advances and future directions. Pages 201–235 in D. L. Thomson, E. G. Cooch, and M. J. Conroy, editors. *Modeling demographic processes in marked populations*. Springer, New York, New York, USA.
- Nigh, T. A., and W. A. Schroeder. 2002. *Atlas of Missouri ecoregions*. Missouri Department of Conservation, Jefferson City, Missouri, USA.
- Norvell, R. E., F. P. Howe, and J. R. Parrish. 2003. A seven-year comparison of relative-abundance and distance-sampling methods. *Auk* 120:1013–1028.
- Pendleton, G. W. 1995. Effects of sampling strategy, detection probability, and independence of counts on the use of point counts. Pages 131–133 in C. J. Ralph, J. R. Sauer, and S. Droege, editors. *Monitoring bird populations by point counts*. U.S. Department of Agriculture, Forest Service General, Technical Report PSW-GTR-149, Pacific Southwest Research Station, Albany, California, USA.
- Ralph, C. J., J. R. Sauer, and S. Droege. 1995. *Monitoring bird populations by point counts*. U.S. Department of Agriculture, Forest Service General, Technical Report PSW-GTR-149, Pacific Southwest Research Station, Albany, California, USA.
- Rosenberg, K. V., S. E. Barker, and R. W. Rohrbach. 2000. An atlas of cerulean warbler populations. <<http://www.birds.cornell.edu/cwap/cwapresultsdec18.pdf>>. Accessed 16 Dec 2009.
- Rosenstock, S. S., D. R. Anderson, K. M. Giesen, T. Leukering, and M. F. Carter. 2002. Landbird counting techniques: current practices and an alternative. *Auk* 119:46–53.
- Simons, T. R., M. W. Allredge, K. H. Pollock, and J. M. Wettroth. 2007. Experimental analysis of the auditory detection process on avian point counts. *Auk* 124:986–999.
- Stanislav, S. J., K. H. Pollock, T. R. Simons, and M. W. Allredge. 2010. 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* 5:(1). <<http://www.ace-eco.org/vol5/iss1/art3/>>. Accessed 9 Aug 2010.

Thompson, W. L. 2002. Towards reliable bird surveys: accounting for individuals present but not detected. *Auk* 119:18–25.

Thompson, F. R. III, and F. A. La Sorte. 2008. Comparison of methods for estimating bird abundance and trends from historical count data. *Journal of Wildlife Management* 72:1674–1682.

U.S. NABCI Committee. 2000. North American Bird Conservation Initiative Bird Conservation Region descriptions: a supplement to the

North American Bird Conservation Initiative Bird Conservation Regions map. U.S. Fish and Wildlife Service, Arlington, Virginia: <<http://www.nabci-us.org/aboutnabci/bcrdescrip.pdf>>. Accessed 14 Oct 2009.

Verner, J. 1985. Assessment of counting techniques. *Current Ornithology* 2:247–302.

Associate Editor: Steven S. Rosenstock.