

Bird population density estimated from acoustic signals

Deanna K. Dawson*¹ and Murray G. Efford²

¹USGS Patuxent Wildlife Research Center, 12100 Beech Forest Road, Laurel, MD 20708, USA and ²Zoology Department, University of Otago, PO Box 56, Dunedin 9010, New Zealand

Summary

1. Many animal species are detected primarily by sound. Although songs, calls and other sounds are often used for population assessment, as in bird point counts and hydrophone surveys of cetaceans, there are few rigorous methods for estimating population density from acoustic data.

2. The problem has several parts – distinguishing individuals, adjusting for individuals that are missed, and adjusting for the area sampled. Spatially explicit capture–recapture (SECR) is a statistical methodology that addresses jointly the second and third parts of the problem. We have extended SECR to use uncalibrated information from acoustic signals on the distance to each source.

3. We applied this extension of SECR to data from an acoustic survey of ovenbird *Seiurus aurocapilla* density in an eastern US deciduous forest with multiple four-microphone arrays. We modelled average power from spectrograms of ovenbird songs measured within a window of 0.7 s duration and frequencies between 4200 and 5200 Hz.

4. The resulting estimates of the density of singing males (0.19 ha^{-1} SE 0.03 ha^{-1}) were consistent with estimates of the adult male population density from mist-netting (0.36 ha^{-1} SE 0.12 ha^{-1}). The fitted model predicts sound attenuation of 0.11 dB m^{-1} (SE 0.01 dB m^{-1}) in excess of losses from spherical spreading.

5. *Synthesis and applications.* Our method for estimating animal population density from acoustic signals fills a gap in the census methods available for visually cryptic but vocal taxa, including many species of bird and cetacean. The necessary equipment is simple and readily available; as few as two microphones may provide adequate estimates, given spatial replication. The method requires that individuals detected at the same place are acoustically distinguishable and all individuals vocalize during the recording interval, or that the *per capita* rate of vocalization is known. We believe these requirements can be met, with suitable field methods, for a significant number of songbird species.

Key-words: acoustic localization, bird counting, density estimation, microphone array, passive acoustic methods, sound attenuation, spatially explicit capture–recapture

Introduction

Animals of many species sing or call in a repetitive and species-specific fashion for intraspecific communication. Vocalizations are an important source of evidence on the abundance of birds, amphibians and cetaceans. In surveys of land birds, an observer usually hears more birds than are seen, and counts based largely on aural cues are routinely used to index bird abundance. Where bird species richness is high or when skilled observers are not available, standardized sound recordings have been suggested as an alternative method for surveying bird populations (Haselmayer & Quinn 2000; Hobson *et al.* 2002; Rempel *et al.* 2005; Alberta Biodiversity Monitoring

Institute 2007; Brandes 2008; Celis-Murillo, Deppe & Allen 2009), allowing field data to be collected by staff unskilled in identifying birds by sound and providing an archival record that can later be interpreted or re-sampled by experts. As generally applied, recordings are made at points to yield species lists for assessments of species richness or occurrence patterns, or counts of birds that index abundance. While an index is sufficient for some purposes (Caughley 1977; Dawson 1981; Johnson 2008), estimates of absolute density are sometimes needed (Kendeigh 1944; Nichols, Thomas & Conn 2009). This requires that counts be adjusted for incomplete detection and the area sampled. In this study, we present methods to estimate bird population density with data extracted from sound recordings. Although we focus on singing birds, the methods have wider application.

*Correspondence author. E-mail: ddawson@usgs.gov

Simultaneous sound recordings from multiple microphones may be used to localize singing birds (Magyar, Schleidt & Miller 1978; McGregor *et al.* 1997; Bower & Clark 2005; Mennill *et al.* 2006). However, localization can be imprecise (Spiesberger & Fristrup 1990) and does not lead directly to an estimate of density. We bypass the localization of individual birds by fitting a statistical model to the signals received at each microphone. The model includes a parameter for the density of sound sources in the surrounding area, and thus provides a direct estimate of this density. An array of microphones resembles a trapping grid in that a sound (animal) may be detected at some microphones (traps) and not others. Unlike traps, microphones have the characteristics of passive 'proximity' detectors (Efford, Borchers & Byrom 2009) – detection of one sound usually does not preclude detection of another at the same microphone, and a sound may be detected at multiple microphones. The probability of being detected by at least one detector in an array may be modelled by combining detector-wise probabilities, each of which we assume to be a decreasing function of distance from the source. This is the essence of spatially explicit capture–recapture (SECR) methods (Borchers & Efford 2008; Efford *et al.* 2009a; Efford, Dawson & Borchers 2009). Density may be estimated from an array of proximity detectors by maximizing the appropriate likelihood (Efford *et al.* 2009a). Sounds are ephemeral and cannot strictly be recaptured, but with multiple microphones there are 'recaptures' in space that may be analysed by SECR methods (Efford *et al.* 2009b).

Having introduced spatially explicit acoustic methods, we need now to clarify the relationship between individual birds and the sounds they produce. A bird may sing repeatedly in one recording interval; each repeated unit is a song burst, or 'cue' (e.g. Buckland *et al.* 2001). A cue is transient and distinguished from other cues by its time and location. Cue density (e.g. song bursts per hectare per minute) is a measurable parameter distinct from population density (individuals per hectare). To estimate population density the observer or analyst must either distinguish cues from different individuals or convert cue density to population density. Converting cue density to population density requires a separate estimate of the rate of cue production (e.g. song bursts per individual per minute), usually obtained by continuous observation of selected individuals (Gates & Smith 1972; Burnham *et al.* 2004; Buckland 2006). The calibration must be repeated for each study because singing rates can vary both within and among individuals, depending on their pairing status (e.g. Gibbs & Wenny 1993) or breeding stage (e.g. Lein 1981; Wilson & Bart 1985), and other factors.

We do not consider in this study methods that rely on estimating cue production rate, although the models we describe may be used to estimate cue density. Rather, we assume that cues of individual birds can be distinguished. An observer conducting a bird point count implicitly ascribes aural cues to individuals 'on the fly' using a combination of loudness, direction, temporal overlap with other cues and possibly other attributes. These same criteria have been used to ascribe cues to individuals from recordings (Rempel *et al.* 2005; Celis-

Murillo *et al.* 2009). For some species, the vocal cues of individuals can be distinguished by their internal characteristics (e.g. Galeotti & Pavan 1991; McGregor & Byle 1992; Gilbert, McGregor & Tyler 1994; Jones & Smith 1997; Puglisi & Adamo 2004; Vögeli *et al.* 2008).

The preceding method for binary acoustic data (sounds either detected or not detected at each microphone) may be extended by modelling variation between microphones in the strength of the received signal (Efford *et al.* 2009b). Sound energy declines with distance from the source because of spherical spreading (in proportion to the inverse-square of distance), scattering and absorption by vegetation, atmospheric absorption and other effects (Wiley & Richards 1982). We first describe the method in general terms and then demonstrate its application in a field study of ovenbirds *Seiurus aurocapilla* (Linnaeus, 1766).

Materials and methods

DATA AND MODEL

We consider an array of K microphones, each at a known unique location. The number and spacing of microphones are design issues that we consider later. If a small array is used (e.g. $K \leq 4$), then it should be deployed at a probability sample of points (e.g. stratified random sample or a systematic grid with random origin) within a region of interest. Simultaneous recordings of short duration (e.g. 5 min) are reviewed aurally, and by scanning the spectrogram, to distinguish vocalizing individuals of the species of interest. In our experience, automatic algorithms for recognizing vocalizations may assist but not replace a human interpreter.

The data comprise n observations where each observation is a vector \mathbf{o}_i of length K describing the reception of one sound at the different microphones. Partial observations are included (i.e. some elements of \mathbf{o}_i may be missing because a sound is detected at some microphones but not others); null observations (sound not detected at any microphone) are excluded. We describe in simple terms the models for the distribution of sound sources and the received signal strength. Each of these is capable of elaboration (e.g. inclusion of covariates), but we defer this for clarity.

DISTRIBUTION OF SOUND SOURCES

We assume that both the microphones and the sound sources lie in a plane. The location of each sound source is unknown and is not estimated, but the locations follow a random 2-dimensional distribution and remain fixed for the duration of the recording. For specificity, we assume a homogeneous Poisson distribution with a single parameter D (density, ha^{-1}). Equivalent inhomogeneous Poisson models allow for spatial trend or habitat effects (Borchers & Efford 2008; Efford *et al.* 2009b). Because in our model sounds are detected independently of each other, we can pool data from different arrays with the same configuration. D is then equal to the product of the average density and the number of arrays.

RECEIVED SIGNAL STRENGTH

We implemented the general signal-strength model of Efford *et al.* (2009b) for the specific case of relative signal power measured in decibels (i.e. on a logarithmic scale). We use as a reference value the power

at 1 m from the source (β_0). The expected signal strength at d metres from the source ($d \geq 1$) is then

$$\mu_S = \beta_0 - 10 \log_{10}(d^2) + \beta_1(d-1). \quad \text{eqn 1}$$

The second term describes the spherical spread of sound energy, and the final term represents log-linear attenuation from all other causes ($\beta_1 < 0$). We set $\mu_S = \beta_0$ for $d < 1$, when spherical spreading generally does not apply (Wiley & Richards 1982), thus avoiding computational problems at $d = 0$. We ignore ground effects, which are important for sounds below 1000 Hz (Wiley & Richards 1982; Tarrero *et al.* 2008). Received signal strength is modelled as a random variable $S = \mu_S + \varepsilon$, where ε is normally distributed with zero mean and variance σ_s^2 . The parameters β_0 , β_1 , and σ_s are estimated by fitting the model.

DETECTION

Sounds are assumed to be detected independently at different microphones, conditional on their source location. Detection is defined as occurring when S exceeds a threshold c . This criterion, combined with the preceding model for signal strength, gives a model for the probability of detection as a function of distance (a 'detection function'). The choice of threshold is somewhat arbitrary, and within a certain range may have little effect on the estimates. We address this later.

LIKELIHOOD

The general SECR likelihood for Poisson-distributed animals and an array of detectors is presented elsewhere (Borchers & Efford 2008; Efford *et al.* 2009a, 2009b). In this study, we specify the probability model for data from a signal-strength detector that may be plugged into that likelihood. Let ω_{ik} represent the signal from animal i at detector k and \mathbf{x} represent the x - y coordinates of a potential source location. Put $y = \omega_{ik} - c$ and $\gamma(\mathbf{x}, k) = (c - \mu_s(\mathbf{x}, k))/\sigma_s$, where μ_s is the expected value from eqn 1. The likelihood component associated with the signal from animal i at detector k is then

$$\Pr(\omega_{ik}|\mathbf{x}) = \Phi(\gamma(\mathbf{x}, k))^{1-\delta_{ik}} \left\{ (2\pi\sigma_s^2)^{-1/2} \exp\left(-(\gamma(\mathbf{x}, k)\sigma_s)^2/(2\sigma_s^2)\right) \right\}^{\delta_{ik}}, \quad \text{eqn 2}$$

where Φ is the standard normal distribution function and δ_{ik} is an indicator variable for whether the observed signal strength exceeds c (i.e. $y > 0$). The probability that signal strength exceeds c at one or more detectors is

$$P(\mathbf{x}) = 1 - \prod_{k=1}^K (1 - \Phi(\gamma(\mathbf{x}, k))). \quad \text{eqn 3}$$

Specification of $\Pr(\omega_{ik})$ and $P(\mathbf{x})$ adapts the general SECR likelihood for a proximity detector (Efford *et al.* 2009a) to model

signal strength (see Efford *et al.* 2009b for further detail). Estimation is by numerically maximizing the likelihood. Care must be taken to avoid local maxima in models with spherical spreading.

FIELD STUDY

We conducted an acoustic survey of breeding birds in deciduous forest at the Patuxent Research Refuge (39°3' N, 76°48' W) near Laurel, Maryland, USA. The forest was described by Stamm, Davis & Robbins (1960), and has changed little since. Four microphones (PA3 mini-microphone; Supercircuits, Austin, TX, USA), mounted 0.4 m above-ground, were arranged in a square with 21-m sides and connected by 15-m cables to a central digital recorder (R-4; Edirol Ltd., London, UK). The built-in preamplifier of the PA3 microphone supplied a line-level signal to the recorder that was sampled digitally at 44.1 kHz and saved uncompressed as a 4-channel 16-bit pulse-code-modulated (.wav) file. Over 5 days during 12–19 June 2007, the microphone array was shifted to each of 75 points on a 50-m square grid (Fig. 1) across a 14-ha study area that was also sampled by mist-netting (see below). Sound was recorded at each array position for 5 min. All recordings were made between 0600 and 1200 Eastern Daylight Time.

SOUND MEASUREMENT

Although the sounds of all vocalizing birds were recorded in our survey, we selected ovenbird as the study species because its song is concise and distinctive, and because males sing from the lower stratum of the forest (i.e. either near the bottom of the canopy or on the ground; D. K. Dawson, pers. obs., Lein 1981). Female ovenbirds do not sing. One interpreter listened to each recording, while also viewing the spectrogram, to distinguish individual singing ovenbirds. Within (but not across) recordings, song bursts were ascribed to individuals by contrasting their relative intensity across the four channels (i.e. microphone-specific recordings) and temporal pattern with that of other song bursts. The first song burst attributed to an individual was selected for measurement, unless it was partly obscured by another bird's song, in which case its next song burst was selected. We used the software Raven Pro 1.3 (Charif, Waack & Strickman 2008) to measure the average power of each selected song burst for each channel. Power is reported by Raven on a logarithmic (decibel) scale. The reference value for this scale is arbitrary, and depends on the microphone and recording system. Thus, it is critical that the same equipment and settings be used for all recordings and that automatic gain control, if present, be switched off. We manually positioned a time-frequency window (0.7 s \times 1000 Hz) on the spectrogram of each song burst (spectrogram record length 256 samples) to sample the portion of the song burst with maximum power (Fig. 2). It was placed at either 2780–3780 or 4200–5200 Hz to avoid

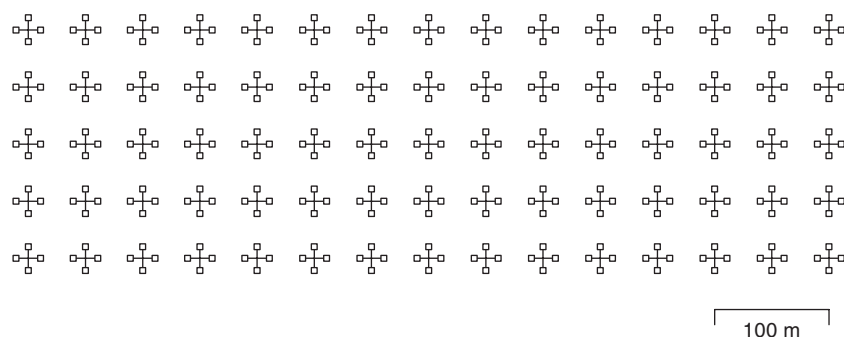


Fig. 1. Microphone arrays for study of ovenbird density at Patuxent Research Refuge, Maryland, USA. Recordings were made simultaneously on four microphones (squares) at each of the 75 array positions over five sampling days in June 2007. On each day, positions 100 m apart were sampled sequentially.

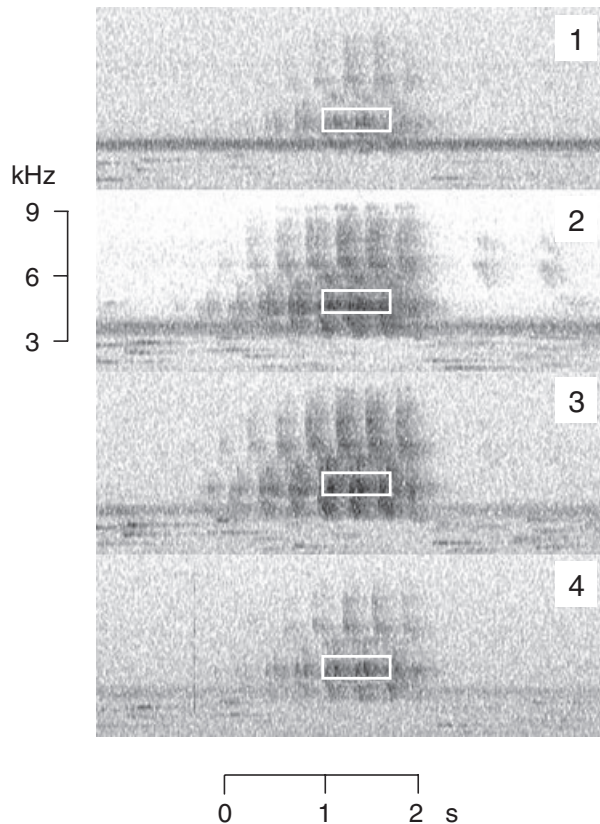


Fig. 2. Ovenbird song burst recorded at four microphones at corners of a 21-m square. White rectangle indicates a 0.7 s × 1000 Hz window used for power measurements. Dark horizontal band prominent on microphones 1 and 2 is insect sound.

overlap with a continuous band of insect sound that partly obscured the spectrogram of most ovenbird songs. When insects were silent ($n = 26$), measurements from both frequency bands were taken and a linear regression was constructed to adjust observations made in the lower band only; later analyses used only average power in the 4200–5200 Hz band. The measurement windows were aligned across channels (i.e. no adjustment was made for differences in the time that a sound arrived at each microphone).

The power we measured includes a component of ambient noise. After each measured song burst, a corresponding measurement was made of the background noise at each microphone, within a 0.7 s window in the same frequency range. We then modelled signal data S' from which the noise component had been removed using the formula

$$S' = 10 \log_{10} \left(10^{S/10} - 10^{N/10} \right),$$

where S was the measured signal and N was the measured noise, both in decibels. This adjustment has little effect on S unless $S - N$ is small ($S - S' < 1$ dB when $S - N > 7$ dB).

CHOICE OF SIGNAL THRESHOLD

Our model requires the user to specify a threshold of relative power below which signals are ignored. As the signal to noise ratio decreases, a signal becomes less detectable. N does not appear in the present model; we deal with the effects of noise by selecting a thresh-

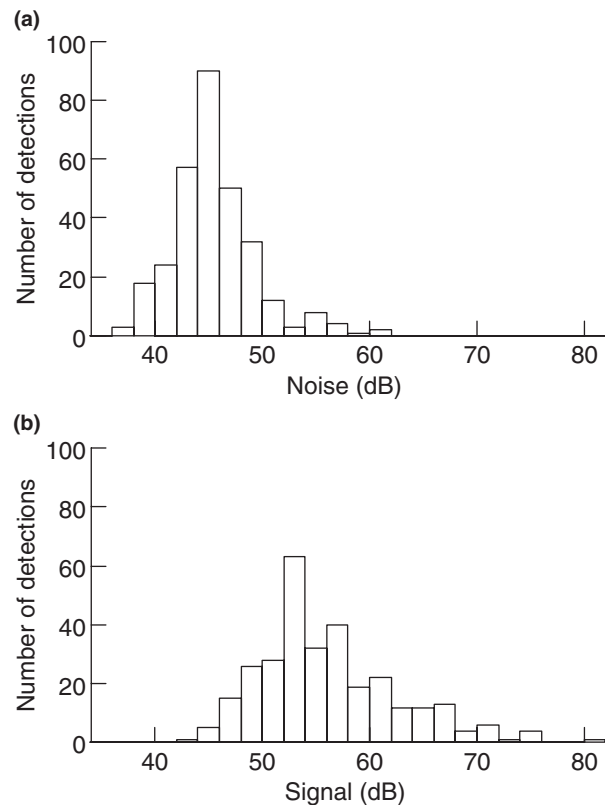


Fig. 3. Distribution of average power in frequency band 4200–5200 Hz measured at four microphones for 76 ovenbird song bursts (304 detections). (a) Background noise. (b) Song (before adjustment for noise).

old of S that is high enough to ensure detection regardless of noise, over the likely range of values for noise. It is important that this is set so that all signals above the threshold are detected, but not so high as to sacrifice precision. We chose a threshold of 52.5 dB, a value that exceeded the ambient noise in 95% of observations from our study area (Fig. 3a).

Simulations in which all signals are detected (measured) show that results are robust to the choice of threshold (Efford *et al.* 2009b), but in field data some signals above the nominated S threshold are likely to be immeasurable because of noise (including simultaneous singing by other birds). We checked this effect empirically by varying the threshold and examining the estimates for any trend that might indicate bias.

ACOUSTIC ESTIMATES OF OVENBIRD DENSITY

We fitted a uniform (Poisson) model for the distribution of singing male ovenbirds in two dimensions by maximizing the likelihood for the signal-strength model in eqns 1–3 (see also Efford *et al.* 2009b). Each likelihood computation used a grid of 4096 points extending 200 m beyond the sampled points. Calculations were performed with function 'optim' in R 2.7.0 (R Development Core Team 2008); external C code was written for the computationally intensive likelihood calculation. For maximization, density (D), $-\beta_1$, and σ_s were varied on a log scale to constrain the corresponding parameters to positive values (D , σ_s) or negative values (β_1) as appropriate. Starting values were varied to find the global maximum. Asymptotic sampling variances (including Poisson spatial variance) were obtained from

the information matrix. Confidence limits for D , $-\beta_1$ and σ_s were back-transformed from the log scale, and standard errors were obtained by the delta method (Lebreton *et al.* 1992).

We compared a full model including spherical spreading to one with only log-linear attenuation, using Akaike's Information Criterion (AIC), although we note that these models have the same number of parameters.

CAPTURE-RECAPTURE ESTIMATES OF OVENBIRD DENSITY

For comparison with acoustic estimates, we estimated the density of the adult male ovenbird population by SECR analysis of mist-netting data collected on the study area. Forty-four nets spaced 30 m apart on the perimeter of a 600 m \times 100 m rectangle were operated for *c.* 9 h on each of 10 non-consecutive days during late May and June in each year 2005–2008. Birds received individually numbered bands, and both newly banded and previously banded birds were released at the net where captured. Mist nets were treated as binary proximity detectors (Efford *et al.* 2009a), and density was estimated by fitting a 'hazard rate' detection function (cf. Hayes & Buckland 1983): $g(d) = g_0 (1 - \exp(-(d/\sigma)^{-z}))$, where d is the distance between a net and the bird's home range centre. The three parameters of this function (g_0 , σ and z) respectively vary the intercept (detection probability when $d = 0$), the spatial scale, and the shape of the curve. Variances were calculated as for the acoustic analysis. Pooling of detection parameters over years potentially provided a means to reduce the sampling variance of density estimates for the year of interest (2007). AIC was used to compare alternative models, specifically those with and without pooled detection parameters, and to assess possible net avoidance after first capture.

Results

AURAL INTERPRETATION OF RECORDINGS

Ovenbird song was detected on 65 of the 75 different 5-min recordings. More than one individual was distinguished in 37 recordings (21 recordings with two birds, 12 with three birds and four with four birds). The mean number of individuals per recording was 1.88 (SE 0.11).

MEASUREMENTS OF SIGNAL AND NOISE

Measurements of power were made for a subset ($n = 76$) of the ovenbird song bursts; the spectrograms of others were too indistinct for measurement on any channel. Background noise was tightly distributed around the median power of 45.3 dB (interquartile range 43.2–47.5 dB), with some skewness because of occasional values up to 61.3 dB (Fig. 3a). Unadjusted signal power was more variable (median 55.0 dB, interquartile range 52.1–59.7 dB; Fig. 3b). Adjusting signal power for concurrent background noise (median 54.1 dB, interquartile range 50.7–59.4 dB) had little effect except in the left tail of the distribution, as expected. The observed ratio of unadjusted signal power to noise had a median of 9.0 dB (minimum 0.9 dB, interquartile range 6.0–13.8 dB, maximum 36.3 dB).

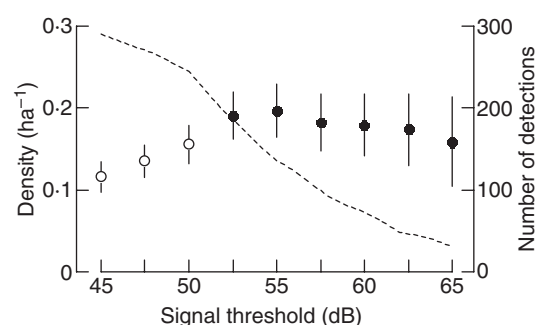


Fig. 4. Acoustic estimates of density of adult male ovenbirds (± 1 SE) and number of detections (dashed line) for varying thresholds of adjusted signal strength. Estimates were robust to the choice of threshold above 50 dB (filled circles); when the threshold was ≤ 50 dB (open circles), estimates were biased because some signals above the nominal threshold were scarcely detectable above the background noise and hence immeasurable.

ACOUSTIC ESTIMATES OF DENSITY

The signal-strength model was fitted to the adjusted signal strengths across a range of threshold values (Fig. 4). Low values for the threshold (below the 52.5 dB value justified by the distribution of noise associated with measured song bursts) yielded low estimates of density and we infer that these were negatively biased. Estimates were robust to the choice of threshold for thresholds ≥ 52.5 dB. Increasing the threshold reduced the number of microphones at which song bursts were detected and hence reduced precision; the apparent decline in the estimate with the highest threshold (65 dB) probably reflects sampling error. We therefore accepted the model with a threshold of 52.5 dB; this gave a density estimate of 0.19 ha^{-1} [95% confidence interval (CI) 0.14 – 0.26 ha^{-1}] (other parameter estimates in Table 1). A model without spherical spreading did not fit the data so well ($\Delta\text{AIC} = 10.5$), but gave

Table 1. Acoustic estimate of ovenbird density: parameter estimates from signal-strength model (threshold 52.5 dB; 180 detections of 60 song bursts) and asymptotic standard errors

Parameter	Full model		Log-linear	
	Estimate (95% CI)	SE	Estimate (95% CI)	SE
Density ha^{-1}	0.19 (0.14, 0.26)	0.03	0.19 (0.13, 0.26)	0.03
β_0 dB	103.8 (102.3, 105.3)	0.8	78.2 (75.5, 80.8)	1.4
β_1 dB m^{-1}	−0.11 (−0.10, −0.14)	0.01	−0.25 (−0.22, −0.28)	0.01
σ_s dB	1.68 (1.43, 1.97)	0.14	1.89 (1.61, 2.23)	0.16

β_0 is the estimated average power 1 m from the source; β_1 is the estimated sound attenuation in excess of spherical spreading; σ_s is the standard deviation of measured power, conditional on distance from the source. The scale for power has an arbitrary origin. Results are also shown for a log-linear model without explicit allowance for spherical spreading; in this case, β_0 is the intercept at 0 m and β_1 is the estimated sound attenuation, inclusive of spherical spreading. Attenuation near the source is almost certainly not log-linear, so the intercept is not a reliable estimate of intensity at the source, and cannot be compared with the value of β_0 from the model with spherical spreading.

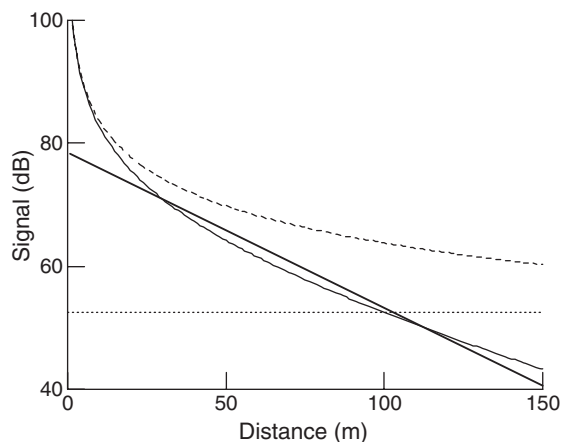


Fig. 5. Attenuation of ovenbird song predicted by the full model ($\beta_0 = 103.8$ dB, $\beta_1 = -0.114$ dB m^{-1} ; continuous curve) and the log-linear model ($\beta_0 = 78.2$ dB, $\beta_1 = -0.25$ dB m^{-1} ; heavy line), compared with that from spherical spreading alone (dashed curve). Horizontal dotted line indicates the signal threshold (52.5 dB).

essentially the same estimate of density (Table 1). The fitted models for signal attenuation are shown in Fig. 5. The resulting acoustic detection model (Fig. 6) has $g(0) \approx 1.0$ because the difference between the power at source ($\hat{\beta}_0 = 103.8$ dB) and the threshold ($c = 52.5$ dB) greatly exceeded the signal error ($\hat{\sigma}_s = 1.7$ dB).

The two subsets of diagonally opposed microphones each yielded 90 detections, from 54 and 56 song bursts respectively. Density estimates from these data (0.16 ha $^{-1}$, CI 0.07 – 0.37 ha $^{-1}$; 0.28 ha $^{-1}$, CI 0.21 – 0.37 ha $^{-1}$) had substantially lower precision, but were otherwise consistent with the four-microphone estimates (Table 1).

CAPTURE–RECAPTURE ESTIMATES OF DENSITY

Eleven to 13 adult male ovenbirds were captured each year. Models with detection parameters pooled over years were preferred by AIC to models with annual parameters, so we report only pooled models. The daily number of captures declined by

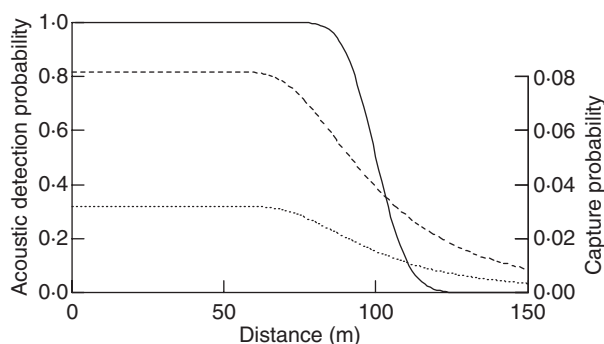


Fig. 6. Detection functions fitted to acoustic and capture–recapture ovenbird data. Solid line – probability of a song burst exceeding the signal threshold (52.5 dB) as a function of distance between bird and microphone. Probability of capture during one day as a function of the distance between home range centre and mist net for naïve birds (dashed line) and birds captured previously (dotted line).

about 75% between the first and last days of mist-netting, and a model including a learned net response (decrease in capture probability after first capture) fitted substantially better than a null model (constant capture probability) ($\Delta AIC = 8.0$). The estimated density of male ovenbirds in 2007 was 0.36 ha $^{-1}$ (CI 0.19 – 0.68 ha $^{-1}$) (other parameter estimates in Table 2).

Discussion

Acoustic methods have been identified as a priority for research on bird population assessment (Nichols *et al.* 2009). We provide the first demonstration of how acoustic data collected with a microphone array can be analysed to estimate avian population density. Mellinger *et al.* (2007) identified the development of statistical methods for estimating populations acoustically as critical for understanding cetacean populations; the method we describe may be applied without modification to data from hydrophone arrays.

Acoustic estimates of density with four microphones were more precise than those obtained by mist-netting. They also required less field effort, and did not expose birds to the stress of capture. The method is suitable for use on a wide range of species, especially those for which individuals can be distinguished by song characteristics. It deserves consideration for any study that requires estimates of bird density, because even experienced observers have difficulty collecting the field data required for other methods to estimate density from aural cues (e.g. Alldredge *et al.* 2008). However, care must be taken to collect data consistent with the assumptions of the method, which we now consider in more detail.

ASSUMPTIONS

Sound sources are assumed to follow a uniform random (Poisson) distribution, although this may easily be generalized to an inhomogeneous Poisson distribution if density varies systematically (Borchers & Efford 2008; Efford *et al.* 2009b). Other studies suggest density estimates by SECR are robust to mis-specification of the distribution (Efford 2004; Efford *et al.* 2009a), and we anticipate this will apply also to acoustic estimates. Overdispersion (clumping) of sound sources will cause bias in Poisson estimates of sampling variance. Where a small array is deployed at multiple sites, as in our study, overdispersion might be addressed by estimating separately the asymptotic variance of the signal attenuation parameters and the variance among sites (Buckland *et al.* 2001).

Sound sources are assumed to be distributed in the plane of the microphones. This will not apply for ground-level microphones when birds sing from the forest canopy. The principal effect will be to invalidate the ‘hard-wired’ spherical spreading term of the attenuation model, as even a sound from directly overhead will have dissipated some of its energy before reaching the ground. We suggest in this case dropping the inverse-square term for spherical spreading and modelling the joint effect of spherical spreading and other attenuation as a single log-linear term.

Table 2. Density of adult male ovenbirds estimated by mist-netting in the breeding season at Patuxent Research Refuge, Maryland, USA

Parameter	Estimate (95% confidence interval)	SE
Density 2005 (ha ⁻¹)	0.37 (0.20, 0.70)	0.12
Density 2006 (ha ⁻¹)	0.33 (0.17, 0.64)	0.11
Density 2007 (ha ⁻¹)	0.36 (0.19, 0.68)	0.12
Density 2008 (ha ⁻¹)	0.30 (0.15, 0.60)	0.10
g_0 (day ⁻¹)	0.082 (0.043, 0.150)	0.026
$\Delta\text{logit}(g_0)$	-0.99 (-1.53, -0.44)	0.28
σ (m)	90 (69, 118)	12
z	4.3 (3.2, 5.8)	0.7

Spatially explicit model with capture rate parameters pooled over years. The model includes a net avoidance effect ' $\Delta\text{logit}(g_0)$ ' that is an additive change in the logit of distance-specific capture probability between the first and subsequent captures.

The model assumes that at least some sounds are detected at more than one microphone, and that average sound energy declines uniformly in all directions. Directional vocalization has been studied for rather few bird species (see references in Patricelli, Dantzker & Bradbury 2008), and more work is needed to understand its significance for density estimation. If vocalizations are directional, but oriented randomly with respect to the microphones, the effect may be to increase the variance of reception (parameter σ_s) without introducing significant bias in density estimates. For highly directional sounds (e.g. the sonar clicks of some cetaceans Au *et al.* 1999; Møhl *et al.* 2003; Zimmer *et al.* 2005), it may be worthwhile to include source orientation in the model to allow for the autocorrelation of reception at adjacent microphones (cf. Mungamuru & Aarabi 2004). The effect of directionality at source may be modified by habitat structure. For example, sounds in forest are diffused by reflection off vegetation, and we predict the effect will be small.

Our probability model assumes independence of detection at different microphones; this is nearly the case when microphones are well-separated. When microphones are closely spaced, stochastic components of environmental sound attenuation and degradation impinge nearly equally on all microphones, resulting in spatial autocorrelation of residuals from the signal attenuation model across the 'sound shadow' of a source. Future models may address non-independence in sound detection by breaking the detection process into systematic components that are common to all microphones (e.g. sound pressure at source and source orientation), and those that are more nearly random and independent between microphones (e.g. sound path and proximity of competing sound sources).

Sound sources are assumed to be immobile within a recording. Movement is likely to cause bias similar to that in distance analysis, because detection locations of moving birds are shifted towards the microphones (cf. Buckland *et al.* 2001; Buckland 2006); movement also makes attribution of sounds to individuals less certain, unless individuals can be distinguished by song characteristics. A sequence of song bursts by a

moving bird should be traced back to the first detectable instance, and the power of this signal used in analyses, even if it is the faintest. This is analogous to the 'snapshot' distance method (Buckland 2006).

The risk of bird movement increases in longer recordings, suggesting that recording duration should be minimized. This conflicts with the further assumption that the duration should be sufficient for all birds to vocalize, so a compromise is needed. Intermittent vocalization, which leads to incomplete availability of birds for detection, is a general problem for passive field methods of density estimation (Diefenbach *et al.* 2007; Nichols *et al.* 2009). As researchers learn more about bird singing behaviour and movement, it may become possible to design spatial models for density estimation that realistically incorporate these behavioural effects. These models might also be used to automate the process of attributing cues to individual birds.

Although the basic model assumes uniformity among microphones and ambient conditions, variation may be modelled by including covariates. If microphones differ in their sensitivity, then it is feasible to include microphone as a categorical covariate in the model, or to use some measure of sensitivity as a continuous covariate (analogous to a trap-level covariate in SECR; Borchers & Efford 2008). Failure to allow for variation among microphones inflates the estimate of σ_s , and appears to cause positive bias in estimates of density when the variation is large (i.e. between-microphone SD > 2 dB in simulations with other parameters matching the ovenbird study; M. G. Efford, unpublished results). Ambient conditions (e.g. background noise) at the time of each recording may be treated as an individual-level covariate for sounds detected during each recording (cf. Huggins 1989; Borchers & Efford 2008).

Our model employs a study-specific threshold of signal strength (power). Signals below the threshold are treated as 'undetected' even if they appear on the spectrogram. It is critical that all signals over the threshold are detected and measured. This will not be the case if the threshold is set so low relative to the background noise that some song bursts with power above the threshold are not measurable on any microphone. In our study, a threshold of 52.5 dB, justified by the distribution of background noise (Fig. 3a), appeared to be effective (Fig. 4). The choice of threshold entails a trade-off between bias and precision, but the model as formulated does not allow us to use information-theoretic criteria (e.g. AIC; Burnham & Anderson 2002) to select an optimal model because varying the threshold varies the data used. The solution probably lies in a larger model that is fitted by conditioning on the measurements of background noise.

COMMENTS ON STUDY DESIGN

This method may be used with any number of microphones ($K \geq 2$) in almost any configuration, but these and other design decisions (spacing of microphones and duration of recording) will strongly affect both the feasibility of the study and the precision of the results. Hardware costs are of the same order as a

pair of good binoculars. Signal processing for one species takes only about 25% of total field time. For surveying large areas, it is desirable to disperse the sampling effort to allow for spatial variation in density. This is feasible with a small array of two or four microphones moved to multiple sites as in our study. Sites should be selected with a probability-based sampling design (Thompson 1992; Buckland *et al.* 2001). Sites would ideally be spaced widely enough that birds are not re-sampled to ensure independence, but there are indications from distance sampling that variance estimates may be almost unbiased even when samples overlap (Buckland 2006). Two people could quickly lay out cables and set up the microphones and recorder on our study area, which is flat, with little ground or shrub cover; it was a slower, but manageable, task for one person to do repeatedly. After set up, the start of recording should be delayed briefly, allowing bird activity to resume as before. Wireless microphones would remove the need to carry and lay out cables at each site, but they are expensive and demand more power. An alternative that is attractive for large arrays is to use an autonomous recorder and microphone at each array position; autonomous recordings need to be time-aligned for analysis, but the required precision (about 0.05 s) is very much less than would be needed for localization (about 0.001 s).

COMPARISON WITH OTHER METHODS FOR DENSITY ESTIMATION

Mist-netting data, analysed by SECR methods, provided alternative estimates of ovenbird density. The population of males sampled by netting presumably included birds that did not sing during the recordings, which may explain the somewhat higher estimate of density by this method. We found evidence for net avoidance, but were able to incorporate this in the model and have no reason to believe the estimates are biased. However, mist-netting has severe limitations as a general method. Even when detection parameters were estimated from four seasons of pooled data, the CI on adult male ovenbird density in 2007 was very large (0.19–0.68 ha⁻¹). The method is not readily scaled for large regions, as nets must be somewhat clustered to produce recaptures.

We are aware of four other approaches to estimating density with data from a microphone array. Marques *et al.* (2009) treat detectors as replicate point samples, rather than analysing patterns of signal reception among detectors. The other methods have not been published in any detail or applied in the field, to our knowledge, but it is useful to outline them here, and to identify their limitations or advantages. It has been suggested that sound sources be localized from the relative arrival times of sounds at different microphones to generate a sample of observed distances for analysis by conventional distance methods (Burnham *et al.* 2004; Buckland 2006; Mellinger *et al.* 2007; Marques *et al.* 2009). With this approach, data from sounds detected at only one or two microphones must be discarded because, for these, localization is impossible. Spatial ambiguity can remain even when sounds are detected at three microphones (Spiesberger 2001); localization errors will bias estimates of density by distance analysis. Also, there is no natu-

ral origin for the measurement of distances, and it may be difficult to select an origin such that $g(0) \approx 1.0$ as required for conventional distance analysis. An alternative approach is to fit a likelihood-based SECR model directly to the observed arrival times of sounds, with an additional parameter for normal errors in time measurement. Our simulations show that this method consistently matches or out-performs distance analysis (M.G. Efford & D.K. Dawson, unpublished results). We do not advocate it because accurate measurement of sound arrival times is difficult; we failed to extract usable data from the recordings we made in forest, although it may be possible from recordings of bird song in open, less reverberant habitats. Another approach is to model the direction of arrival of sounds at two or more clusters of microphones, determined by beamforming (steerable array) methods in software (Chen, Yao & Hudson 2002). Maximum likelihood estimation is feasible, given suitable data (M.G. Efford, unpublished results). This approach has the potential advantage that beamforming improves the separation of signal from noise in the chosen direction (Spiesberger & Fristrup 1990), but data collection is technically demanding. In summary, distance methods are difficult to apply and not needed for microphone array data; methods based on time delays pose technical challenges and are unproven.

Finally, we find the signal-strength method attractive because it relates detection to measurable acoustic processes. This should motivate investigation of the acoustic properties of habitats in relation to transmission of vocal cues, an area that has received rather little attention since the 1970s (Padgham 2004; Barker 2008). We measured excess attenuation of 11 dB/100 m, a value consistent with measurements of sound attenuation in forests for frequencies above 1000 Hz (Wiley & Richards 1982). The acoustic properties of forests can be approximated with physical models (e.g. Tarrero *et al.* 2008), and these may ultimately allow attenuation (the β_1 parameter) to be predicted from habitat characteristics and atmospheric conditions. Loudness at source (β_0) may be relatively constant within species. It is not clear whether externally derived values of β_0 and β_1 can replace estimates in particular studies, but at least they will provide strong priors for Bayesian versions of these analyses.

Acknowledgements

The authors thank Larissa Bailey, Dave Dawson, Charles Francis, Tiago Marques, Jim Nichols and an anonymous reviewer for comments on a draft, and David Borchers and Andy Royle for general discussion. Use of trade, product, or firm names does not imply endorsement by the U.S. Government.

References

- Alberta Biodiversity Monitoring Institute (2007) Terrestrial field data collection protocols, Version 2007-12-13. Available at: <http://www.abmi.ca> (last accessed on 26 June 2009).
- Aldredge, M.W., Pacifici, K., Simons, T.R. & Pollock, K.H. (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.
- Au, W.W.L., Kastelein, R.A., Rippe, T. & Schooneman, N.M. (1999) Transmission beam pattern and echolocation signals of a harbor porpoise (*Phocoena phocoena*). *Journal of the Acoustical Society of America*, **106**, 3699–3709.

- Barker, N.K. (2008) Bird song structure and transmission in the Neotropics: trends, methods and future directions. *Ornitologia Neotropical*, **19**, 175–199.
- Borchers, D.L. & Efford, M.G. (2008) Spatially explicit maximum likelihood methods for capture–recapture studies. *Biometrics*, **64**, 377–385.
- Bower, J.L. & Clark, C.W. (2005) A field test of the accuracy of a passive acoustic location system. *Bioacoustics*, **15**, 1–14.
- Brandes, T.S. (2008) Automated sound recording and analysis techniques for bird surveys and conservation. *Bird Conservation International*, **18**, S163–S173.
- Buckland, S.T. (2006) Point-transect surveys for songbirds: robust methodologies. *Auk*, **123**, 345–357.
- Buckland, S.T., Anderson, D.R., Burnham, K.P., Laake, J.L., Borchers, D.L. & Thomas, L. (2001) *Introduction to Distance Sampling: Estimating Abundance of Biological Populations*. Oxford University Press, Oxford.
- Burnham, K.P. & Anderson, D.R. (2002) *Model Selection and Multimodel Inference: A Practical Information-Theoretic Approach*, 2nd edn. Springer-Verlag, New York.
- Burnham, K.P., Buckland, S.T., Laake, J.L., Borchers, D.L., Marques, T.A., Bishop, J.R.B. & Thomas, L. (2004) Further topics in distance sampling. *Advanced Distance Sampling* (eds S.T. Buckland, D.R. Anderson, K.P. Burnham, J.L. Laake, D.L. Borchers & L. Thomas), pp. 307–392. Oxford University Press, Oxford.
- Caughley, G. (1977) *Analysis of Vertebrate Populations*. John Wiley & Sons, London.
- Celis-Murillo, A., Deppe, J.L. & Allen, M.F. (2009) Using soundscape recordings to estimate bird species abundance, richness, and composition. *Journal of Field Ornithology*, **80**, 64–78.
- Charif, R.A., Waack, A.M. & Strickman, L.M. (2008) *Raven Pro 1.3 User's Manual*. Cornell Laboratory of Ornithology, Ithaca, New York.
- Chen, J.C., Yao, K. & Hudson, R.E. (2002) Source localization and beamforming. *IEEE Signal Processing Magazine*, **19**, 30–39.
- Dawson, D.G. (1981) Counting birds for a relative measure (index) of density. *Estimating Numbers of Terrestrial Birds* (eds C.J. Ralph & J.M. Scott). *Studies in Avian Biology*, **6**, 12–16. Allen Press, Inc., Lawrence, Kansas.
- Diefenbach, D.R., Marshall, M.R., Mattice, J.A. & Brauning, D.W. (2007) Incorporating availability for detection in estimates of bird abundance. *Auk*, **124**, 96–106.
- Efford, M.G. (2004) Density estimation in live-trapping studies. *Oikos*, **106**, 598–610.
- Efford, M.G., Borchers, D.L. & Byrom, A.E. (2009a) Density estimation by spatially explicit capture–recapture: likelihood-based methods. *Modeling Demographic Processes in Marked Populations* (eds D.L. Thomson, E.G. Cooch & M.J. Conroy), pp. 255–269. Springer, New York.
- Efford, M.G., Dawson, D.K. & Borchers, D.L. (2009b) Population density estimated from locations of individuals on a passive detector array. *Ecology*, **90**, 2676–2682.
- Galeotti, P. & Pavan, G. (1991) Individual recognition of male Tawny Owls (*Strix aluco*) using spectrograms of their territorial calls. *Ethology, Ecology & Evolution*, **3**, 113–126.
- Gates, C.E. & Smith, W.B. (1972) Estimation of density of mourning doves from aural information. *Biometrics*, **28**, 345–359.
- Gibbs, J.P. & Wenny, D.G. (1993) Song output as a population estimator: effect of male pairing status. *Journal of Field Ornithology*, **64**, 316–322.
- Gilbert, G., McGregor, P.K. & Tyler, G. (1994) Vocal individuality as a census tool: practical considerations illustrated by a study of two rare species. *Journal of Field Ornithology*, **65**, 335–348.
- Haselmayer, J. & Quinn, J.S. (2000) A comparison of point counts and sound recording as bird survey methods in Amazonian Southeast Peru. *Condor*, **102**, 887–893.
- Hayes, R.J. & Buckland, S.T. (1983) Radial-distance models for the line-transect method. *Biometrics*, **39**, 29–42.
- Hobson, K.A., Rempel, R.S., Greenwood, H., Turnbull, B. & Van Wilgenburg, S.L. (2002) Acoustic surveys of birds using electronic recordings: new potential from an omnidirectional microphone system. *Wildlife Society Bulletin*, **30**, 709–720.
- Huggins, R.M. (1989) On the statistical analysis of capture experiments. *Biometrika*, **76**, 133–140.
- Johnson, D.H. (2008) In defense of indices: the case of bird surveys. *Journal of Wildlife Management*, **72**, 857–868.
- Jones, D.N. & Smith, G.C. (1997) Vocalisations of the Marbled Frogmouth II: an assessment of vocal individuality as a potential census technique. *Emu*, **97**, 296–304.
- Kendeigh, S.C. (1944) Measurement of bird populations. *Ecological Monographs*, **14**, 67–106.
- Lebreton, J.-D., Burnham, K.P., Clobert, J. & Anderson, D.R. (1992) Modeling survival and testing biological hypotheses using marked animals: a unified approach with case studies. *Ecological Monographs*, **62**, 67–118.
- Lein, M.R. (1981) Display behavior of Ovenbirds (*Seiurus aurocapillus*) II. Song variation and singing behavior. *Wilson Bulletin*, **93**, 21–41.
- Magyar, I., Schleidt, W.M. & Miller, R. (1978) Localization of sound-producing animals using the arrival time differences of their signals at an array of microphones. *Experientia*, **34**, 676–677.
- Marques, T.A., Thomas, L., Ward, J., DiMarzio, N. & Tyack, P.L. (2009) Estimating cetacean population density using fixed passive acoustic sensors: an example with Blainville's beaked whales. *Journal of the Acoustical Society of America*, **125**, 1982–1994.
- McGregor, P.K. & Byle, P. (1992) Individually distinctive bittern booms: potential as a census tool. *Bioacoustics*, **4**, 93–109.
- McGregor, P.K., Dabelsteen, T., Clark, C.W., Bower, J.L., Tavares, J.P. & Holland, J. (1997) Accuracy of a passive acoustic location system: empirical studies in terrestrial habitats. *Ethology, Ecology & Evolution*, **9**, 269–286.
- Mellinger, D.K., Stafford, K.M., Moore, S.E., Dziak, R.P. & Matsumoto, H. (2007) An overview of fixed passive acoustic observation methods for cetaceans. *Oceanography*, **20**, 36–45.
- Mennill, D.J., Burt, J.M., Fristrup, K.M. & Vehrencamp, S.L. (2006) Accuracy of an acoustic location system for monitoring the position of duetting songbirds in tropical forest. *Journal of the Acoustical Society of America*, **119**, 2832–2839.
- Möhl, B., Wahlberg, M., Madsen, P.T., Heerfordt, A. & Lund, A. (2003) The monopulsed nature of sperm whale clicks. *Journal of the Acoustical Society of America*, **114**, 1143–1154.
- Mungamuru, B. & Aarabi, P. (2004) Enhanced sound localization. *IEEE Transactions on Systems, Man, and Cybernetics. Part B: Cybernetics*, **34**, 1526–1540.
- Nichols, J.D., Thomas, L. & Conn, P.B. (2009) Inferences about landbird abundance from count data: recent advances and future directions. *Modeling Demographic Processes in Marked Populations* (eds D.L. Thomson, E.G. Cooch & M.J. Conroy), pp. 201–235. Springer, New York.
- Padgham, M. (2004) Reverberation and frequency attenuation in forests – implications for acoustic communication in animals. *Journal of the Acoustical Society of America*, **115**, 402–410.
- Patricelli, G.L., Dantzker, M.S. & Bradbury, J.W. (2008) Acoustic directionality of red-winged blackbird (*Agelaius phoeniceus*) song relates to amplitude and singing behaviours. *Animal Behaviour*, **76**, 1389–1401.
- Puglisi, L. & Adamo, C. (2004) Discrimination of individual voices in male Great Bitterns (*Botaurus stellaris*) in Italy. *Auk*, **121**, 541–547.
- R Development Core Team (2008) *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria. Available at: <http://www.R-project.org>, accessed on 27 May 2008.
- Rempel, R.S., Hobson, K.A., Holborn, G., Van Wilgenburg, S.L. & Elliott, J. (2005) Bioacoustic monitoring of forest songbirds: interpreter variability and effects of configuration and digital processing methods in the laboratory. *Journal of Field Ornithology*, **76**, 1–11.
- Spiesberger, J.L. (2001) Hyperbolic location errors due to insufficient numbers of receivers. *Journal of the Acoustical Society of America*, **109**, 3076–3079.
- Spiesberger, J.L. & Fristrup, K.M. (1990) Passive localization of calling animals and sensing of their acoustic environment using acoustic tomography. *American Naturalist*, **135**, 107–153.
- Stamm, D.D., Davis, D.E. & Robbins, C.S. (1960) A method of studying wild bird populations by mist-netting and banding. *Bird-Banding*, **31**, 115–130.
- Tarrero, A.I., Martín, M.A., González, J., Machimbarrena, M. & Jacobsen, F. (2008) Sound propagation in forests: a comparison of experimental results and values predicted by the Nord2000 model. *Applied Acoustics*, **69**, 662–671.
- Thompson, S.K. (1992) *Sampling*. John Wiley & Sons, New York.
- Vögeli, M., Laiolo, P., Serrano, D. & Tella, J.L. (2008) Who are we sampling? Apparent survival differs between methods in a secretive species. *Oikos*, **117**, 1816–1823.
- Wiley, R.H. & Richards, D.G. (1982) Adaptations for acoustic communication in birds: sound transmission and signal detection. *Acoustic Communication in Birds*, Vol. 1, *Production, Perception, and Design Features of Sounds* (eds D.E. Kroodsma, E.H. Miller & H. Ouellet), pp. 131–181. Academic Press, New York.
- Wilson, D.M. & Bart, J. (1985) Reliability of singing bird surveys: effects of song phenology during the breeding season. *Condor*, **87**, 69–73.
- Zimmer, W.M.X., Johnson, M.P., Madsen, P.T. & Tyack, P.L. (2005) Echolocation clicks of free-ranging Cuvier's beaked whales (*Ziphius cavirostris*). *Journal of the Acoustical Society of America*, **117**, 3919–3927.

Received 19 December 2008; accepted 9 October 2009

Handling Editor: Esteban Fernandez-Juricic