Inferences About Landbird Abundance from Count Data: Recent Advances and Future Directions

James D. Nichols, Len Thomas, and Paul B. Conn

Abstract We summarize results of a November 2006 workshop dealing with recent research on the estimation of landbird abundance from count data. Our conceptual framework includes a decomposition of the probability of detecting a bird potentially exposed to sampling efforts into four separate probabilities. Primary inference methods are described and include distance sampling, multiple observers, time of detection, and repeated counts. The detection parameters estimated by these different approaches differ, leading to different interpretations of resulting estimates of density and abundance. Simultaneous use of combinations of these different inference approaches can not only lead to increased precision but also provides the ability to decompose components of the detection process. Recent efforts to test the efficacy of these different approaches using natural systems and a new bird radio test system provide sobering conclusions about the ability of observers to detect and localize birds in auditory surveys. Recent research is reported on efforts to deal with such potential sources of error as bird misclassification, measurement error, and density gradients. Methods for inference about spatial and temporal variation in avian abundance are outlined. Discussion topics include opinions about the need to estimate detection probability when drawing inference about avian abundance, methodological recommendations based on the current state of knowledge and suggestions for future research.

1 Introduction

For decades, the majority of inferences about landbird abundance and density have been based on counts conducted by investigators, either stationed at points (e.g., Blondel et al. 1970) or walking along line transects (Emlen 1971; Jarvinen and Vaisanen 1975). Counts resulting from point and transect sampling have been treated frequently as indices to abundance, in the sense that the expected counts

J.D. Nichols (✉)
U.S. Geological Survey, Patuxent Wildlife Research Center, Laurel, MD, USA
e-mail: jnichols@usgs.gov

D.L. Thomson et al. (eds.), *Modeling Demographic Processes in Marked Populations*,
Environmental and Ecological Statistics 3, DOI 10.1007/978-0-387-78151-8_9,

have been assumed to represent an unknown, yet relatively constant, proportion of the sampled population. Early on, some investigators argued that the proportionality assumption is not likely to be widely met, or at least merits testing, and advocated the collection of supplemental data with counts that permit inference about the detection probabilities of individual birds and thus about true abundance and density (e.g., Ramsey and Scott 1979; Burnham et al. 1980, 1981). Debate about approaches for drawing inferences about population size and dynamics from avian count data has motivated symposia and workshops over the years (Ralph and Scott 1981; Ralph et al. 1995) and has persisted through the present time (e.g., Anderson 2001; Hutto and Young 2002, 2003; Rosenstock et al. 2002; Thompson 2002b; Ellingson and Lukacs 2003).

The past 5 years have been a period of especially active research on inference methods for avian count data. Such research has included development of new estimation methods, application of these and previous methods in investigations of relatively large scale, and serious testing of existing methods using novel experimental approaches. These developments are sufficiently recent that it has been difficult for investigators to keep up with progress that has been made. Thus, we hosted a small workshop at Patuxent Wildlife Research Center, Maryland USA, inviting 19 biometricians and avian population ecologists (see Acknowledgements) who have played large roles in the recent research. The purposes of the workshop were to obtain a synthesis of the current "state of the art" in methods for estimating landbird abundance from point count and related data, to highlight future research needs, and to determine how best to bridge the gap between statisticians and practitioners. This paper represents an effort to summarize some of the central conclusions and points of discussion from the workshop.

2 Conceptual Framework

2.1 Basic Framework

Discussions of use of count data as a basis for inference about animal populations frequently begin with 2 facts about sampling animal populations (e.g., Lancia et al. 1994, 2005; Borchers et al. 2002; Williams et al. 2002):

(1) Interest is frequently in areas that are sufficiently large that animals cannot be counted over the entire area for which inference is desired;
(2) At locations where investigators do obtain counts of animals by whatever means (counts of animals seen, heard, captured, etc.), these counts seldom include all animals at the sampled location.

Fact 1 is common to many areas of statistics, and traditional design-based sampling approaches are applicable (e.g., Cochran 1977; Thompson 2002a). These sampling approaches are designed to use data from locations at which counts are made to draw inferences about locations where counts are not made. In design-based sampling, the key to such extrapolation is to sample locations in such a manner that all locations about which inference is desired have some known, non-negligible probability of being included in the sample. These probabilities (sometimes called "coverage" probabilities) are determined by the type of design (e.g., random sampling, stratified random sampling, systematic sampling, adaptive cluster sampling), and can be computed based on a knowledge of the design (including desired sample size). Model-based sampling represents a somewhat different approach in which covariate relationships estimated from data on locations that are visited are assumed to apply to locations that are not visited. If this assumption holds, then as long as covariate information is available for all locations of interest, inferences about animal abundance can be made even from locations at which no counts are made. Although geographic variation and spatial sampling were not the primary foci of the workshop, discussion of these topics arose frequently, as their importance was clearly recognized.

Fact 2 involves detectability, and the workshop focused on approaches for dealing with this issue. Traditional discussions of detectability view counts of animals (C_i for location i) as random variables, the expectation of which can be written as the product of the true number of animals at the location at the time of the survey (N_i) and the detection probability (p_i), the probability that a member of N_i appears in C_i:

$$E(C_i) = N_i p_i \ . \tag{1}$$

In the context of the workshop, counts were usually the numbers of birds seen or heard, and discussion focused on how to translate these counts into inferences about true abundance or density.

For some purposes, estimates of true abundance or density are required, and can be obtained as:

$$\hat{N}_i = \frac{C_i}{\hat{p}_i} \ . \tag{2}$$

In many cases, we can view location i as corresponding to an individual sampling unit selected from a large area for which an abundance estimate, $\hat{N}$, is desired. In an effort to deal with facts 1 and 2 listed above, we define $p_{c,i}$ as the coverage probability of sample i within this large area and estimate N as:

$$\hat{N} = \sum_i \frac{C_i}{\hat{p}_i p_{c,i}} \ . \tag{3}$$

More frequently, inferences of interest involve not abundance itself, but ratios of abundance over space (often termed relative abundance) or time (often termed trend or rate of population change). One approach to inference about ratios of abundance is to standardize data collection procedures in hopes of obtaining similar detection probabilities for the different times or locations to be compared. If similar detection probabilities can be obtained, then ratios of the counts themselves provide reasonable estimates of ratios of abundances. Another approach is to hope that most of the relevant temporal and spatial variation in detection probability is associated with recorded covariates that have no possibility of also being associated with variation in true abundance. For example, observer identity is such a covariate and can be incorporated into analyses that use raw count data (e.g., Link and Sauer 1997). Those who do not believe it is safe to rely on standardization and covariate identification typically advocate collection of data needed to draw direct inference about detection probability and its variation. Given such data, it is possible to compare alternative models that express different hypotheses about how detection probability varies as a function of time, space, or recorded covariates. The workshop included discussion of the relative efficacy of these 2 general approaches: (1) use of raw counts with assumptions about relevant sources of variation in detection probability, versus (2) collection of data needed to draw inferences about variation in detection probabilities, using methods that also require assumptions about the detection process.

2.2 Decomposition of Coverage and Detection Probabilities

K.H. Pollock presented a conceptual framework for the workshop that extended the ideas presented above to include different components of detection (also see Pollock et al. 2002; Farnsworth et al. 2002, 2005). Specifically, he noted that detection can be broken into components associated with availability and detection given availability. The issue of availability has been discussed, mainly with respect to aquatic organisms that may be submerged at the time of the survey and thus not exposed to surface survey methods (e.g., Marsh and Sinclair 1989; Laake and Borchers 2004; Okamura et al. 2006). In an auditory survey, a bird that does not vocalize during the survey period is not available to be detected. A bird that does vocalize is available and may or may not be detected depending on the probability of detection given availability. During the course of the workshop, there was also discussion of temporary emigration and the possibility that a bird that sometimes uses a particular sampled site (i.e., the site is included in the bird's territory or home range) may not be present on the site at the time of the survey.

These ideas about geographic sampling and detection probability cause us to further subdivide the coverage and detection probabilities of Section 2.1 and view the probability that a bird in some large area of interest is actually detected during a survey within that area as the product of 4 conceptually distinct probabilities. We begin by considering all birds whose territories or home ranges lie at least partially within the large area about which inference is to be drawn. We refer to this number as superpopulation size, N^*, indicating that these birds have some probability of being exposed to sampling efforts at any given survey time. We are not necessarily interested in estimating N^*, but we simply identify members of N^* as the individuals that may appear in our sample counts. We then consider a randomly chosen bird from the superpopulation N^* and consider its probability of being counted during a survey.

One way to define coverage probability is the probability that the location of the bird is within a sampling unit (location at which a survey count is conducted) at the time of the survey. It is helpful to decompose coverage probability into two parts. The first is the probability that the bird's home range or territory at least partly overlaps a sampling unit. We denote this probability as p_s, and associate it with spatial sampling (Cochran 1977; Thompson 2002a). This probability depends on the spatial sampling design (how are sampling units selected, what are the sizes of the units) and on the size and shape of the bird's home range, or of the portion of it lying within the large area of interest. Conditional on a bird's home range overlapping at least one selected sample unit, the second probability of interest is the probability that a bird is present at a sample unit during the survey period (the time spent surveying at that unit). We refer to this as the probability of presence, p_p, as it indicates the probability that the bird is within the area exposed to sampling efforts for at least some of the survey period. For ease of extrapolation and discussion, we assume the simple case in which the range of a single bird can overlap at most a single selected sample unit, although more complicated situations are certainly possible. Decomposing coverage probability into these two components allows us to account for temporary emigration of birds from a sampling unit during the time of the survey; this occurs with probability $(1-p_p)$.

We now decompose detection probability into two parts. If a bird is located within the sample unit during the survey period, we still consider the possibility that the bird is not available for detection during that time. For example, in an auditory survey, a bird that fails to vocalize during the survey period is essentially unavailable for detection despite being present in the surveyed area. We use the term availability for this detection component and denote the associated probability as p_a. We further note that the methods we consider focus on birds for which $p_a > 0$. For example, if females of some species are simply invisible in the general period during which surveys are conducted, then none of our estimation approaches will be useful for them. Finally, we consider the probability of detection given presence and availability, p_d. In an auditory survey, this probability corresponds to the observer actually hearing a bird and being able to identify the species. This component of detection probability is likely to vary as a function of such factors as observer skill and sensory (e.g., hearing) abilities, vocalization characteristics of the bird species, habitat structure, and distance from the observer.

All of these probabilities are viewed as components of the process generating the count data, and their identification and specification during the workshop hopefully led to clear thinking and discussion of the various methods used to estimate bird numbers and density. In particular, we note that most of the estimation methods discussed at the workshop incorporate some sort of parameter representing detection probability. We will use the terms "detection probability" to describe the parameters estimated by different approaches, and we reserve the notation p_d for detection conditional on presence and availability.The nature of the detection probability parameter varies among the different methods, as it can reflect only p_d, the product $p_a p_d$, or even the product $p_p p_a p_d$. Combinations of methods provide opportunities for estimating these components separately and in some cases such decomposition may be useful. We believe that it is especially important that the investigator be aware of which detection parameter applies to a particular method, as this determines the population about which inferences are being made. We discuss this further when introducing each method, below.

Note that for ease of presentation and interpretation, the above components of the detection process are listed as single parameters, implying that every individual in the population of interest shares these common parameters. In reality, at least some of these components will vary by individual, such that these parameters are best viewed as averages. In some cases, these probabilities are best viewed at the level of the individual and modeled as a function of individual-level covariates.

2.3 Closed Populations and Open Fields

Frequently, we will express interest in inferences about a closed population of size N in a known area of size A. While this conceptualization is directly applicable in some circumstances, for example an island population, it is often the case that the survey boundaries are not closed to animal movement, so that population size is not a fixed quantity but varies even over short time intervals. The inclusion of p_p as a component of the detection process in the framework presented above reflects a recognition of this reality. Further, there are some circumstances when the area of inference A is not well defined, although these are often associated with circumstances where there is a poor survey design. M.G. Efford (collaborating with D.K. Dawson) made a presentation, part of which described an alternative perspective to the "closed population" paradigm, which he called an "open field". In this perspective, detectors are located within some area of interest and the focus is on estimating animal density, defined as the local intensity of a spatial point process.

2.4 Index Methods

In certain situations, it may be possible to diagnose population trends without explicitly estimating detection probabilities, treating counts as indices to abundance. In a recent survey of 224 papers using field counts of landbirds, Rosenstock et al. (2002) found that 95% viewed the counts as indices for inference about variation in abundance. Index methods traditionally assume constant detectability, so that changes in raw counts over time or space are viewed as representing changes in the population of interest. Use of indices can also involve covariates hypothesized to influence detection probability but not true abundance or density. Such analyses require that all systematic changes in detectability can be explained and modeled with covariates (e.g., Link and Sauer 1997). At the workshop, W.A. Link and D.H. Johnson gave provocative presentations that focused on the key points of consideration when making decisions about whether to use an index approach to point counts or to instead try to model and/or estimate detection probabilities. W.A. Link maintained that covariates influencing detectability can often be controlled for, and that it was indeed possible to make inferences about population trends without explicit estimation of detection probability. Both presenters argued that index approaches to monitoring, such as the BBS, have provided valuable information on population trends, and have helped to identify species of concern whose status warrants further investigation.

Subsequent discussion of index methods largely focused on the tenability of model assumptions. Workshop participants reached consensus that the assumption of constant detectability needed for index methods was sometimes overstated. Rather, it would suffice that expected detection probability, E(p), is constant over time (e.g., Nichols et al. 2000; Yoccoz et al. 2001; Conn et al. 2004). D. H. Johnson argued that even if there is a small trend in detectability, index methods may still be sufficient to diagnose large scale changes in abundance. The same general comments apply to covariate analyses of index data, although the interpretation of E(p) is in this case changed to the expected detection probability after available covariates have been used to control for variation in detectability.

Index-based methods are inherently attractive when their assumptions are met; they require fewer data to be collected and avoid potentially problematic assumptions about the functional form of individual heterogeneity in detection probability (Link 2003). On the other hand, when index assumptions are sufficiently violated, they may lead to erroneous inferences. When trends are estimated as ratios of raw counts or as regressions of counts on time, undetected trends in detection probability can either mask changes in abundance or cause one to observe a spurious trend in abundance. Use of appropriate covariates can correct for some of the factors influencing detection probability. For instance, Link and Sauer (1998) were able to detect, and correct for, changes in the competency of bird point count participants over time. However, these approaches cannot be used to correct for unrecorded covariates that influence detection, nor for covariates that may be associated with trends in abundance. For instance, decibel level has undoubtedly increased along roads in the United States over the past several decades, but historically has not been recorded at U. S. Breeding Bird Survey (BBS) point counts. Decibel level appears to be inexorably linked with the auditory detection process (Simons et al. 2007), and so may have led to unmodeled trends in detectability in BBS analyses. Even if decibel level had been recorded, this covariate is likely related to land use and proximity of human development, variables that may influence actual bird abundance. Global warming, succession, and introduction of invasive species are some of the examples proposed by workshop participants that might simultaneously influence both abundance and detectability. If these factors have large influences on expected detection probability, index methods will not be adequate for the analysis of population trends. In these cases, methods designed to explicitly estimate detection probability are needed.

3 Basic Inference Methods

3.1 Distance Sampling

L. Thomas presented to the workshop an overview of "conventional" distance sampling (CDS) methods (Burnham et al. 1980; Buckland et al. 1993, 2001). There are two main variations: line transects, where the survey is performed from a set of randomly located lines, and point transects, where it is from a set of randomly located points. The basic idea is the same for both. Observers record the distance from the line or point to all birds detected within some truncation distance, w (which in practice may be infinity, i.e., all detections are recorded, but some finite truncation distance is almost invariably specified at the analysis stage). The sample units are therefore a set of strips (line transects) or circles (point transects) of known size. Not all birds within the sample units are detected, but a fundamental assumption of the conventional methods is that all birds at zero distance are available and detected. Intuitively, one would expect that birds become harder to detect on average with increasing distance from the line or point. The key to distance sampling is to use the distribution of the observed distances to estimate the "detection function", denoted $g(y)$ – that is the probability of detecting a bird, given it is at distance y. This function can then be used to estimate the average probability of detecting a bird given that it is within a sample unit and available for detection – i.e., p_d. Note that the conventional methods are not designed to estimate availability, p_a – if this is <1 then additional data are required. Given an estimate of detection probability, p_d, it is straightforward to estimate density using design-based methods such as equation (3), since coverage probability is known by design. Let N_s represent the total number of birds whose home ranges overlap the set of sample units surveyed by the two observers. Because distance sampling estimates p_d, abundance estimates obtained using this approach are associated with the birds present at the sample locations during the sample period and available to be detected during that time. If ranges of individual bird do not overlap multiple surveyed sample units, then $E(\hat{N}_s) \approx N_s p_p p_a$.

Assumptions of CDS used in estimating p_d are (1) animals are distributed independently of the line or point locations; (2) all birds at zero distance are detected; (3) distances are measured without error; (4) observations at a line or point take place at an instant in time, so that animal movement is negligible. Assumption 1 (independent animal distribution) is true by design if a large number of sample units are located at random within the study area, and may be violated if there is non-random sample unit placement such as surveys along roads or trails. Substantial violation can lead to substantial bias in the estimator of abundance. Assumption 2 (g(0)=1) may be violated due to "availability bias" (i.e., non-availability of a component of the population, such as female birds not vocalizing in an aural survey) or "perception bias" (birds that are available being missed at zero distance). In both cases, additional information is required to estimate the proportion missed, either through a separate survey (e.g., observations of radiomarked or colormarked birds to directly estimate availability, Section 4.3) or more complex mark-recapture distance sampling methods (Section 3.7). Assumption 3 (no measurement error) may be more easily met in visual surveys, where laser rangefinders can be used to provide accurate distances, but is more problematic in aural surveys, as we discuss later. We also discuss methods designed to correct for bias caused by measurement error in cases where the measurement error process is relatively simple and well understood. Assumption 4 (no animal movement) can also be problematic in some field situations. Responsive movement of animals, such as attraction of birds to the observer before detection, can cause substantial bias, and even random movement can cause significant bias for point transects. One reason that bias tends to be worse in point transects is that the observer is (more or less) stationary; short survey periods (perhaps 3 min) can minimize the problems, but a preferable protocol is a "snapshot" method where only known locations of birds at a pre-defined instant are recorded – see Buckland (2006) for a more detailed description. Another potential solution is to use cue-counting methods, where distances to individual cues such as song bursts are recorded, rather than individual birds. Cue production rate is estimated in a separate survey, and used to convert an estimate of the abundance of cues back to abundance of animals. If possible, estimation of cue production rate is carried out at the same time/place of the actual survey. Detailed advice on survey design and field methods is given in Buckland et al. (2001, Chapter 7) and Strindberg et al. (2004), and recommendations specifically oriented to landbird studies are given in Buckland (2006). A fifth assumption, that each bird detection is an independent event, is not as important in practice.

While not strictly assumptions, there are some additional requirements for robust estimation. First, the detection function should have a "shoulder" – i.e., the probability of detection should remain at or close to 1 initially as distance from the line or point increases. This is often referred to as the "shape criterion". Second, the detection function should be smooth. Third, the models used for g(y) should be flexible, in the sense that they can take a wide variety of plausible shapes, so that they will be a good approximation to the true detection function given a large sample size. Such models are termed "model robust". Fourth, an adequate sample size of distances is required. "Adequate" is hard to define unambiguously, since more samples are required for "difficult" detection functions (e.g., small shoulder or steep fall-off in detectability), however Buckland et al. (2001, Section 7.2.2) recommend at least 60–80 observations for line transect studies and 75–100 for point transect studies. Under these conditions, the estimators of abundance are "pooling robust", meaning that even large variations among individuals in probability of detection due to observer, habitat, etc. cause little bias in the estimate of p_{d} and hence abundance. Thomas presented simulations that demonstrated this, but which also showed when heterogeneity in detectability is extreme (e.g., singing males and cryptic females), significant bias can arise (Thomas et al. in prep.). One potential solution to extreme heterogeneity is to include the factors causing the heterogeneity as additional covariates in the detection function model (Marques et al. 2007). This approach may also offer a partial solution to the sample size requirements if rarer species can be combined with more common ones and species used as a detection function covariate (Alldredge et al. 2007b). A fifth requirement is of an adequate sample of lines or points (minimum 10–20, Buckland et al. 2001, Section 7.2.1) for reliable estimation of the spatial component of variance of the abundance estimate.

3.2 Multiple Observers

M.W. Alldredge presented the basic ideas underlying models based on multiple observers and time of detection. The multiple-observer approach requires that 2 or more observers either sample a point together or traverse a line transect together, keeping track of observer-specific detections of individual birds. The approach is adapted from work by Cook and Jacobson (1979) on estimation approaches for aerial surveys and is closely related to capture–recapture modeling of closed population data (Otis et al. 1978; Seber 1982). Field sampling by multiple observers can be treated in either of two general ways, labeled dependent and independent. Both approaches have been used with avian point count data. Our descriptions will be of point counts, although we note that both approaches can be implemented along line transects as well.

Under a dependent double-observer approach, at each point one observer is designated as "primary" and the other as "secondary". The primary observer identifies all birds detected and communicates each detection to the secondary observer. The secondary observer records these detections of the primary observer, as well as additional birds that the primary observer does not detect (e.g., Nichols et al. 2000). Observers switch roles at different points such that each observer serves as primary observer for about half the sample points. The data for a series of point counts conducted in this manner by two observers can be summarized as four sufficient statistics for each species or group of species to be analyzed together: the number of birds detected by observer i (i=1,2) when that observer was the primary observer, and the number of extra birds detected by observer i when the other observer was primary observer. These data can then be used to estimate the number of birds exposed to sampling efforts at the group of surveyed points under a general model in which detection probabilities differ between observers and among species. Reduced-parameter models can then be developed to evaluate hypotheses about the similarity of detection probabilities for observers and bird species.

The general model for the dependent double-observer approach assumes that detection probability of an observer does not vary depending on the observer's role as primary or secondary. It is assumed that detection and recording of a bird by the secondary observer does not influence the probability that the primary observer detects the bird. The approach assumes that all birds of a species found within the sampled area (frequently defined by a specified fixed radius) at the time of the sample have the same probability of being detected by an observer (the probabilities may be different for the two observers). The number of birds exposed to sampling efforts is assumed to be fixed (population closure), and this assumption may limit the sample to a very short time period (e.g., 3 min). Nichols et al. (2000) provide a more detailed discussion of field sampling methods, analysis methods, underlying model assumptions and field approaches directed at meeting model assumptions.

The independent multiple-observer approach to point counts requires that two or more observers independently record detections from the same basic field sampling point for some specified short period of time (e.g., 3 min). The observers will typically have a schematic diagram of the surveyed area (concentric circles of different radii around the sample point) such that detections of a species are recorded on the diagram with a time of detection. Immediately following the count, observers confer and compare diagrams with the purpose of matching detections of the same birds and developing detection histories for every bird detected (Alldredge et al.

2006). For example, with two independent observers, three detection histories and associated sufficient statistics can be observed: x_{11} = number of birds detected by both observers, x_{10} = number of birds detected by observer 1 and not observer 2, and x_{01} = number of birds detected by observer 2 and not observer 1. These data are then analyzed as closed model capture–recapture data, where time-specific variation in the capture–recapture context is analogous to observer-specific variation in the point count context. Observer variation in detection probability can be incorporated into models, or detection probability can be modeled as a constant for all observers. If >2 observers conduct the sampling, then finite mixture heterogeneity models (e.g., Norris and Pollock 1996; Pledger 2000) can be fit that permit variation among individual birds in their probabilities of being detected (Alldredge et al. 2006).

Assumptions underlying the independent multiple observer approach include population closure and independence of detections among observers. It is further assumed that detection histories are correct (i.e., that there are no matching errors). The double observer models assume the same detection probabilities for the different individuals of the same species within the sample area for each observer. This homogeneity assumption can be relaxed with >2 observers using finite mixture heterogeneity models (Alldredge et al. 2006). If limited-radius counts are used, it is assumed that the observer correctly determines whether each detected bird is inside or outside of the specified radius.

The detection probabilities estimated using multiple-observer approaches pertain to the conditional probability of detection, p_d, given that the bird is present in the area exposed to sampling efforts (e.g., located inside the area defined by a fixed radius) at the time of the sample (probability associated with this event is p_p) and given that it vocalizes or is otherwise available during the sample period (associated probability is p_a). The detection probability is also conditional on the initial probability that the point count sample unit is overlapped by the home range of a particular bird (associated probability p_s), but as this probability is a component of all bird detections, we will omit it from our discussions of the different detection parameters estimated by the different methods. Let N_s represent the total number of birds whose home ranges overlap the set of sample units surveyed by the two observers. Because multiple-observer estimation focuses on p_d, the abundance estimated using this approach is associated with the birds present at the sample locations during the sample period and available to be detected during that time. If ranges of individual bird do not overlap multiple surveyed sample units, then $E(\hat{N}_s) \approx N_s p_p p_a$.

3.3 Time of Detection

The time of detection approach to abundance estimation from avian point counts requires only a single observer at each sampled point. The duration of the entire point count is divided into component time intervals (e.g., a 3-minute point count might be divided into three 1-minute time intervals). The initial development of this approach focused on the time interval of first detection for each bird detected in the count (Farnsworth et al. 2002). If we let K denote the total number of time intervals in a point count, then the sufficient statistics under this approach are the numbers of birds detected for the first time in each interval, $x_1, x_2, \ldots x_K$.

Modeling of these sufficient statistics requires an abundance parameter, N, and detection parameters, p. The modeling is identical to that of removal modeling in capture–recapture literature (Otis et al. 1978; Seber 1982). Estimation is not possible with interval-specific detection parameters, and these parameters are typically assumed to be constant over time when all intervals are of equal length. Farnsworth et al. (2002) also consider the situation where the intervals are of unequal length. Let t_i represent the length of interval i expressed in some relevant time unit (e.g., minutes). Then the probability of a bird being detected during interval i can be written as: $p_i = 1-(1-p)^{t_i}$, where p is the probability of detection for a single unit of time. Under an equivalent continuous time formulation, let ϕ_i be the instantaneous rate of detection during interval i, or "Poisson detectability coefficient" (Alldredge et al. 2007a). Then the probability of detection during interval i is: $p_i = 1 - e^{-\phi_i t_i}$. Heterogeneity among individual birds at a sample unit can be modeled using a finite mixture (e.g., Pledger 2000) or other approach (other estimators for M_{bh} of Otis et al. 1978).

It is also possible to treat time of detection data as standard capture–recapture data, rather than as simply removal data (e.g., Alldredge et al. 2007a). For example, instead of recording the time interval of first detection, the observer records all intervals of detection for each bird. For example, with two time intervals, three detection histories and associated sufficient statistics can be observed: x_{11}= number of birds detected in both time intervals, x_{10} = number of birds detected only in the first time interval, and x_{01} = number of birds detected only in the second time interval. These data are then analyzed using standard capture–recapture models for closed populations, with time interval of detection being equivalent to a sample period in closed capture–recapture. It seems likely that the initial detection of an individual might have a different (typically smaller) detection probability than subsequent detections, in which case analysis would be based on the time intervals of first detection.

Assumptions underlying the time of detection approach include population closure and independence of detections of an individual among the different sample intervals. If time of first detection only is modeled then the latter independence assumption is no longer relevant. The time of detection approach assumes the same detection probabilities for the different individuals of the same species within any sample interval. This homogeneity assumption can be relaxed with >2 intervals using finite mixture heterogeneity models. It is assumed that birds are not double-counted (1 bird mistakenly counted as 2). The time-of-detection approach is typically applied to sample plots defined by a fixed radius (fixed distance from the point). If fixed-radius counts are used, it is assumed that the observer correctly determines whether each detected bird is inside or outside of the specified radius. A final assumption concerns the circle defined by the modeling of the availability process and thus of p_a. Farnsworth et al. (2002) initially modeled availability as a random process in the sense that each bird had an equal probability of vocalizing in any time interval. However, if the process for individual birds is Markovian, in the sense that vocalization during one interval causes the probability of vocalization in subsequent intervals to be larger or smaller, then this process should be incorporated into the modeling and estimation (e.g., see analogous situation in capture–recapture with temporary emigration, Kendall et al. 1997).

The detection probabilities estimated using time of detection approaches pertain to the product of (1) the conditional probability of being available (p_a), given presence in the sample unit at the time of the survey (associated probability p_p), and (2) the conditional probability of bird detection, p_d, given presence and availability. The detection probability estimated by this approach is also conditional on the initial probability that the point count sample unit is overlapped by the home range of a particular bird (associated probability p_s), but this term is omitted in our development as in the example for multiple observers. Let N_s once again represent the total number of birds whose home ranges overlap the set of sampled points. Because estimation based on time of detection focuses on the product $p_a\, p_d$, the abundance estimated using this approach is associated with the birds present at the sample location during the sample period but is not conditioned on availability. Thus, if ranges of individual birds do not overlap multiple surveyed sample units, then $E(\hat{N}_s) \approx N_s p_p$.

3.4 Repeated Counts

J. A. Royle outlined approaches to the use of repeated count data from the same locations as a basis for inference. In doing so, he noted that the data arising from point counts can be viewed naturally in terms of hierarchical models with two basic components, an observation component and a process component. The observation component of such models deals with survey methods and avian detection, conditional on true abundance, whereas the process component deals with the distribution of true abundance over space or survey points. Royle noted that the likelihoods for the approaches described above (distance sampling, multiple observers, time of detection) can all be viewed as multinomial observation models in at least some instances (e.g., distance sampling with data grouped by intervals). He then specified two other data types resulting from repeated point counts at the same locations, the replicate counts themselves and the reduced presence–absence (detection–nondetection) data separating 0 and positive counts.

The sampling protocol involves simple point counts (no necessary collection of ancillary data on distance, time of detection, etc.) at the same locations at multiple times. The different sampling occasions are typically close together in time (e.g., 5 counts at each point during May or perhaps the breeding season) to achieve a kind of closure in the sense that the same group of breeding birds is potentially exposed to sampling efforts at each occasion (see more complete discussion below). The data arising from such sampling are the counts for each sampling occasion. Conditional on the true abundance of birds potentially exposed to sampling efforts at a point, the counts at one location can be viewed as binomial random variables with detection probability modeled as a constant, or perhaps as a function of site- or time-specific covariates (Royle 2004). If abundances are viewed as site-specific, then the resulting likelihood contains many abundance parameters and can be difficult to maximize (e.g., Carroll and Lombard 1985).

Under the approach outlined by Royle, the conditional (on detection probability and site-specific abundances) binomials can be viewed as the observation component of the likelihood. Royle (2004; see also Kery et al. 2005) then proposed use of a reasonable density (e.g., Poisson or negative binomial, with parameters possibly modeled as functions of covariates) for the distribution of true abundances over the spatial sampling locations as the process component of the model. Estimation under the resulting hierarchical model can be accomplished numerically, although a large number of sampling locations will typically be needed to achieve adequate performance.

When the repeat count data at each location are condensed into detections (at least 1 individual of the species counted at the sampling occasion) and nondetections (species not detected at that occasion), abundance can be estimated by relying on the relationship between detection probability at the level of the sample unit (p*) and the detection probability of individual birds (p). This relationship is a function of abundance at location i, as $p_i * = 1 - (1 - p)^{N_i}$ Presence–absence detection data provide information about detection probability at the level of the sample unit, and variation in this probability over space provides information about the distribution of N_i (Royle and Nichols 2003).

Under the basic approach using replicate counts, the binomial detection parameter is actually the product$p_p p_a p_d$. In order to appear in a count, an individual bird must be present in the sample unit at the time of the sampling, must be available for detection during the time of sampling and then must be detected. Similarly, the individual bird detection parameter of models for presence–absence data reflects this product. Thus, the detection parameters estimated using the approach of Royle (2004) and also Royle and Nichols (2003) will usually be smaller than detection probabilities using other approaches as they include the detection component associated with the complement of temporary emigration; the probability that a bird is in the portion of its range exposed to sampling efforts during the survey. Similarly, the abundance estimated using these approaches includes not just the animals that are present during a single survey, but all birds with some non-negligible probability of being in the sample unit during a survey (i.e., all birds whose ranges overlap the sample unit). In the language of capture–recapture modeling (e.g., Kendall et al. 1997; Williams et al. 2002), the abundance estimates produced by the replicate count approaches represent "superpopulation" sizes (N_s) of all of the birds whose home ranges overlap the sample units.

We note that the above inclusion of p_p as a component of the detection parameter estimated using a repeated counts approach is based on the usual field application in which the replicate counts are separated in time by a period during which substantial bird movement is expected (e.g., at least 24 hours). If repeat counts can be independently conducted on an area within a much shorter time frame (e.g., every third minute for 12 min), then this approach will also be conditional on the set of birds present during this period. Our point is that the different interpretation of estimates resulting from the repeated count approaches is not based on anything inherent in the approach itself, but is instead determined by the time intervals over which it is typically applied.

The population closure assumption underlying repeated count approaches applies to the superpopulation size rather than to abundance during any survey occasion. So the number of birds with ranges overlapping the sample unit is assumed not to change over the time period of the repeat counts. The probability of detecting an individual in a given survey (representing the product $p_p p_a p_d$) is generally assumed to remain constant over survey occasions, but this assumption can likely be relaxed in various ways. The repeat count approach generally assumes the same detection probabilities for the different individuals of the same species within any survey, but this assumption can likely be relaxed using various mixture distributions for the detection parameters. It is assumed that birds are not double-counted (1 bird mistakenly counted as 2) during a survey. Although not an assumption, the estimation of superpopulation size, rather than abundance of birds in a sample unit at a snapshot in time, has implications for interpretation and use of resulting estimates. In particular, use of estimates obtained from repeated counts to estimate abundance in some larger area of interest may be more difficult. In addition, the superpopulation of birds exposed to any particular sample unit depends on bird mobility and may change with bird density, such that comparative uses of resulting estimates must be evaluated carefully.

3.5 Double Sampling

The general term "double sampling" from general sample survey statistics (e.g., Cochran 1977; Thompson 2002a) has recently been used to refer to a specific approach to estimation of avian density from count data (Bart and Earnst 2002). The approach involves extensive rapid survey methods on a typically large number of sample units and intensive surveys on a subset of these sample units. If the intensive surveys yield unbiased estimates of true abundance, then the ratio of counts from rapid surveys to estimates based on intensive surveys provides estimates of detection probabilities (for avian applications see Smith 1995). The usual approach to double sampling in wildlife surveys (e.g., Pollock et al. 2002; MacKenzie and Royle 2005) uses methods such as those described above as the intensive surveys, as these methods require ancillary data (distances to detected birds, times of detection), multiple observers, or repeat counts. The rapid surveys often involve single counts with no ancillary data.

The intensive approach described by Bart and Earnst (2002) for Alaskan shorebirds required several hours per day for about 3 weeks, with much of this time reportedly spent searching for nests and counting territorial birds. No estimator was presented for abundance on the intensive survey plots, so we did not discuss or attempt to evaluate this approach. The definition of abundance on a plot as "number of territorial males whose first nest of the season, or territory centroid for non-nesters, was within the plot" (Bart and Earnst 2002) indicates interest in a subset of the superpopulation of all birds whose ranges overlap the sample unit. The basic assumption underlying this approach is that observers on intensive survey plots end the season with an unbiased estimate of abundance on the plot.

3.6 Interpreting Estimates

Both formal and informal discussion at the workshop was devoted to the issue of interpreting estimates from these basic approaches to abundance estimation. Interpretation is dependent on which components of the detection process are the bases for conditioning and which are considered as part of the process to be estimated. All of the approaches begin by conditioning on the birds whose home ranges overlap the selected sample units, with expected value for this set of birds $E(N_s) \approx N^* p_s$, where N^* again is the total number of birds in the entire area of interest from which samples are drawn. If ranges of individual birds do not overlap multiple selected sample units, then abundance estimates based on repeat sample approaches at the selected sample units estimate N_s. Estimates based on time of detection approaches estimate the number of birds present in the selected sample units during the survey period, $N_s p_p$. Estimates based on distance sampling and multiple observers estimate the number of birds that are present in the selected sample units during the survey period and available to be detected, $N_s p_p p_a$. From an "open field" perspective, the quantity $N_s p_p$ is roughly equivalent to the number of birds with home range centers lying within sample units (assuming that observers do not influence the probability of presence, etc.). Thus, one key distinction among abundance estimates based on these four approaches is that a repeated count estimate does not apply to a known fixed area (unless the time scale for repeat surveys is very short), whereas the other three approaches yield snapshot estimates that can be associated with sample units of known area. However, estimates or assumptions about p_a will typically be needed to make inferences about absolute abundance when distance sampling or multiple observer methods are employed. Finally, recall that the above expectations apply to birds for which $p_a > 0$. If a subset of birds is simply invisible to detection efforts, then abundance estimates will of course not include this subset.

The above considerations lead to three points that deserve emphasis. First, if abundance is estimated for a set of sample units at which point counts are conducted, different estimates are expected depending on which estimation approach is used. Second, approaches based on combinations of the above methods provide an opportunity to separate components of the detection process in cases where this might be useful. Third, the different interpretations of the closure assumption for the basic methods lead to the recognition that some approaches lend themselves more readily to estimation of abundance for the entire area from which samples are drawn, than do others. For example, investigators will frequently be interested in the number of birds whose range centers are located within the specified area of interest (define this number as N). Distance sampling, multiple observers, and time of detection all condition on the birds that are actually present in the sample units during the survey period (expectation $N_s p_p$). Given application of design-based sampling protocols, estimates based on these methods can be readily used to extrapolate (based on p_s) to an estimate of overall abundance, N, for the entire area of interest. However, repeated count approaches yield estimates of the number of birds whose ranges overlap the selected sample units such that extrapolation of these estimates to estimate N would require extra information (e.g., about the average number of sample units overlapped by each bird or the actual area sampled by a point count).

3.7 Combination Approaches

Several attendees noted the possibility of using both distance sampling and either multiple observer or time of detection sampling (e.g., D.L. Borchers, M.W. Alldredge, K.H. Pollock, M. Efford). The different kinds of information resulting from such combination-method point or transect counts can be viewed in various ways. For example, from the perspective of multiple-observer and time at detection approaches, distance may be viewed as a covariate that is used to deal with heterogeneous detection probabilities (Alldredge et al. 2006; 2007a, b). From the perspective of distance sampling, use of multiple observers at the same locations and times can be used to estimate the probability of detection on the transect line and/or to test the hypothesis that this probability is 1 (e.g., in aerial surveys). In addition, multiple observers at lagged times (e.g., while traversing an oceanic line transect) can be used to estimate p_a, where the complement, $1–p_a$, includes individuals that are submerged. The need for a time lag between multiple observer counts in order to estimate p_a is based on the same thinking as for time of detection approaches. Indeed, Farnsworth et al. (2005) combined time at detection and distance sampling in order to estimate both p_a and density, corrected for availability.

A well-developed literature now exists for combined applications of distance sampling and double-observer approaches (Alpizar-Jara and Pollock 1996, 1999; Manly et al. 1996; Hiby and Lovell 1998; Borchers et al. 1998a, b; Borchers 1999; Laake and Borchers 2004), although applications to avian point count data have been relatively recent (Alldredge et al. 2006; Kissling and Garton 2006). It is now widely recognized that such combination approaches can be used to relax assumptions required by single component approaches and sometimes to permit separation of different components of the detection process. For example, mark-recapture distance sampling (MRDS) allows estimation of $g(y)$ without the requirement that animals be distributed independently of the transect lines or points (Laake and Borchers 2004, Section 6.3.1.1). Intuitively this is done using the proportion of animals seen ("marked") by one observer at distance y that were also seen ("recaptured") by the other observer. A similar approach can be used to deal with responsive animal movement (Laake and Borchers 2004, Section 6.3.1.2).

When detection on a point or line is not certain, MRDS methods are typically not pooling robust in the sense of CDS. However, D. L. Borchers presented some recent work at the workshop indicating that if the distribution $g(y)$ is assumed to be known (e.g. uniform), one can use data on the distribution of observed distances to relax heterogeneity assumptions (Borchers et al. 2006).

K.H. Pollock discussed the possibility of using double observers and time of detection simultaneously. Pollock noted that this approach can be viewed as a robust design (Pollock 1982; Williams et al. 2002), with the point count divided into K

time intervals and detection–nondetection data for each observer at each interval. For example, in a 3-minute point count with 1-minute intervals and 2 observers, a detection history for a bird might be: 00 11 01. During the first interval the bird was not detected by either observer. Both observers detected it during the second, and only the second observer detected the bird during the final interval. This approach would permit separate estimation of 2 detection components, p_d and p_a. Pollock also noted that whereas the initial development of the time of detection approach assumed random availability (i.e., constant p_a, Farnsworth et al. 2002, or perhaps variation in p_a associated with time or distance, Alldredge et al. in 2007a), the combination approach would permit treatment of availability as a Markov process. Under such a model (similar to Markovian temporary emigration models of Kendall et al. 1997), the availability of a bird in a specific time interval could differ depending on whether the bird was available in the previous interval. Such Markovian modeling would seem to represent a reasonable hypothesis about avian singing behavior.

4 Field Tests, Field Trials, and a Simulation Experiment

4.1 Bird Radio System

One of the primary motivations for the workshop was the recent and ongoing work on field tests of estimation methods for avian point counts. In particular, the bird radio system developed by the group at N.C. State University has provided an excellent test system with which to test estimation methods themselves as well as specific assumptions that underlie those methods (e.g., Alldredge et al. 2007c; Simons et al. 2007). Workshop presentations by T.R. Simons, M. Alldredge and K. Pacifici noted that tests with their system showed that most observers had substantial difficulties estimating distances to birds that are detected aurally, particularly at large distances (beyond 60 m). Heaping of observations at certain distances was also a problem. Double counting of individual birds was a substantial problem, especially in experiments simulating relatively complex field situations (several bird species with heterogeneous singing rates and different speaker orientations). Density estimates based on distance sampling approaches were too high and estimated detection probabilities too small when observer-estimated distances were used. However, there was speculation during discussion that with more truncation better results might be obtained at the cost of decreased sample size of detections, and hence reduced precision. Distance sampling approaches based on true distances to detected bird vocalizations performed well.

Use of the radio bird system to test double-observer and time of detection approaches led to interesting insights as well. K. Pacifici noted that multiple observer approaches sometimes seemed to provide reasonable estimates and at other times yielded biased estimates. In the case of independent observers, there were substantial difficulties with matching, in that observers found it difficult to decide whether detected birds represented a bird detected by both observers or different birds detected by each observer separately. There was some experimentation with objective decision rules for determining matches. Despite matching difficulties for independent observers, estimation results were fairly similar for independent and dependent double-observer approaches. Double counting of birds occurred fairly frequently during experiments with time of detection methods. Placement of singing birds in the wrong time interval was another common error in these tests. In cases where observers were asked to restrict data to birds within a fixed radius, it was common to include birds outside the detection radius. Finally, many observers were unable to identify the species of all detected birds and thus had a number of "unknown" observations. Overall, estimates based on time of detection approaches were biased low and frequently appeared to better estimate $N_s p_p p_a$ than $N_s p_p$. Thus, the approach did not seem to account well for birds that were present but did not happen to vocalize during the survey period.

Although it was difficult to summarize results of all the various experimental tests conducted using the bird radio system, it was clear that none of the basic estimation methods performed as well as would be desired in the common situation of many individual birds from a diverse community. Although this work is still underway, workshop participants involved in the bird radio project offered several summary conclusions and recommendations. Auditory detection of singing birds was found to be much more problematic than previously thought. Localization of sound proved to be very difficult, leading to problems in matching birds when applying multiple-observer approaches, in double counting for single observers, and in estimating distances both for distance sampling and for the purpose of identifying birds lying in and out of fixed radius sample units. Observer performance was especially poor in more complicated experiments with larger numbers of species and individual birds. Performance was better in simple experiments with a small number of species, leading to the recommendation that many of the methods discussed above would be most useful in situations where observers concentrate on a very small number of focal species. In situations where interest must focus on an entire diverse community, it was proposed that alternative approaches such as occupancy estimation and modeling (e.g., MacKenzie et al. 2006) should be considered. The rationale for this recommendation is that occupancy estimation simply requires information on detection, or not, of each species, and does not require keeping track of individual birds over time and space. Another suggestion was that more effort be allocated to the field effort with different observers being responsible for different sub-taxa.

4.2 Robust Distance Sampling Methods Study

S.T. Buckland reported results of a comparison of various distance sampling methods against results from territory mapping for four species (common chaffinch, great tit, European robin and winter wren) in a Scottish woodland and parkland study area of approximately 40 ha (Buckland 2006). The methods compared were

(1) a conventional point transect of 5 min duration; (2) a point transect survey using the "snapshot" approach, with 3 min allowed before the snapshot moment and 2 min afterwards; (3) a point transect using cue counts, where the number of song bursts were recorded using 5 min surveys, and a separate note was kept of birds where the observer was confident all cues were heard so that cue production rate could be calculated; (4) a conventional line transect survey. The habitat was more open than in the bird radio study, resulting in more frequent visual detections, or at least identification of the tree or bush from which a bird was heard to sing. A laser binocular was used to accurately measure distance in these cases. The estimated 95% confidence limits contained the values obtained by territory mapping for all species and methods with one exception: the estimate for great tit from conventional point transects was too high, likely due to animal movement. The snapshot method was found to be the most efficient of the point transect methods, in terms of precision per hour field time, but line transect sampling was more efficient than all the point transect methods. This is likely a general finding, leading to a recommendation that line transect sampling be employed rather than point transect methods where they are possible. Where this is not feasible (for example in difficult terrain, or multi-species surveys where observers may get swamped) then snapshot methods are preferred over conventional point transects. Cue counting may be particularly useful for single-species surveys, but are unlikely to be of use when the goal is to survey many species simultaneously as the observer will quickly become overwhelmed. Insufficient resources were devoted to estimating cue rates in this study, but given more resources the method offers great potential to provide reliable results where animal movement is a problem. The study also included a computer simulation component, which showed that edge effects caused by undersampling near the edge of small study areas do not necessarily cause a serious problem with the reliability of results, nor does sampling at closely spaced points, so that the same birds are heard from multiple points.

4.3 Grassland Bird Availability Study

D.R. Diefenbach reported results of a study of grassland birds in central Pennsylvania, U.S.A. designed to estimate availability, p_a (Diefenbach et al. 2007). Color-marked and radio-marked grassland sparrows were followed for relatively long periods of time (e.g., 1 hour) while observers recorded their availability with respect to auditory detections at point counts, and auditory and sight-based detections at both point counts and line transects. Overall, availability of male sparrows was low ($\hat{p}_a < 0.5$) and quite variable for periods $\leq$ 10 min. Time of the breeding season was an important source of variation with availability being greatest in late May and early June but declining to very low levels late June and July. Diefenbach et al. (2007) concluded that substantial bias and heterogeneity in abundance estimates would be obtained if this source of detection probability were not incorporated in population estimates.

4.4 Other Field Trials

P.F. Doherty recounted experiences with organizing substantial field efforts based on avian point counts and occupancy surveys in Colorado and southern California. The California work was on an endangered species, and survey efforts included many points with no birds detected and only small numbers of detections at remaining points. His Colorado example included work on design-based sample allocation, and Doherty noted substantial problems in accessing a large number of selected sites. A relatively large proportion (10%) of landowners denied access to their lands, leading to a change in scope of inference. Doherty highlighted the frequent need to deal with unanticipated practical problems that may necessitate reevaluation of methods and approaches.

4.5 A Simulation Experiment

Using his "Open Field" view of abundance, M.G. Efford presented preliminary results from a simulation study, in which the probability of detection of an animal depended primarily on distance of the animal's home range centre from the detector (Efford and Dawson in prep.) He showed that if distance has a strong effect on probability of capture then many methods show substantial negative bias in estimated density, including multiple-observer, time of detection and repeated count methods (all using 4 "capture occasions"). This was the case even when using mixture models to try to account for distance induced heterogeneity. Using distance sampling methods or adding distance as a covariate in some of the above methods produced lower bias, although it appeared to be important which model was used for the relationship between detection probability and distance.

5 Methodological Advances and Extensions

5.1 Distance Sampling with Density Gradients

T. A. Marques presented recent work extending conventional distance sampling methods to deal with situations where animal distribution is not independent of distance from the transect line or point (Marques 2008). Such methods are useful in two contexts. The first is when there are few transects located at random: although random transect placement ensures independence on average, the actual distribution of animals for a single realization may be far from the average if there are few transects. The second is where transects are not located at random, but instead follow roads, trails, rivers, etc. In these circumstances, it is not possible with conventional distance sampling to distinguish between changes in animal density with respect to distance from the transect line and changes in detectability, and hence it is not possible to use the observed distances to estimate p_d without additional data. One solution is to use double-observer methods mentioned earlier. A second related approach, for line transects, is a “crossed design”, where two sets of transects are laid out perpendicular to one another; information in the along-transect distribution of observations is then used to estimate the distribution of animals perpendicular to the other transect (see Buckland et al. 2007a for an application to plant surveys). Indeed any additional survey that provides information about the distribution of observations perpendicular to the main line transect could be used. A third approach is to perform point transects along the linear feature (road, trail, etc.) and record both the distance and angle of observations. Then, under the assumption that the detection function is radially symmetric, it is possible to use differences in the distribution of observed distances in the along-road direction and perpendicular-to-road direction to simultaneously estimate detection probability and animal distribution. Marques presented a simulation study of this third method that included different effects of the linear feature (road) on animal density, and demonstrated marked reduction in bias relative to conventional methods, although the method did not appear very robust to misspecification of the model for animal density. It appears that the new method will not over-ride the advice that roadside (and other non-random) transects should be avoided whenever possible.

5.2 Measurement Error in Distance Sampling

Part of the presentation by Marques concerned the effects of measurement error on distance sampling estimators, and extensions to the conventional methods that account for measurement error (Marques 2004, 2008; Burnham et al. 2004: 11.9, Borchers et al. in prep.). Random, non-systematic measurement error can lead to bias in abundance estimates, with the magnitude of the effect depending on the type of survey (bias is generally much worse for point transects than line transects) and type of measurement error (e.g., constant variance with respect to distance from the transect vs. constant CV) and its magnitude. In general, measurement error close to the line or point has more effect than the same level of error far from the transect. Systematic errors cause larger bias than non-systematic errors. Given a model for the error process it is generally possible to estimate error process parameters jointly with detection function parameters, and so to reduce or eliminate the bias if the error process model is correct, at the expense of lower precision in the estimate of p_d. In some cases, the correction method is quite simple to implement (Marques 2004). However, the true process leading to measurement errors may be quite complex, as has been demonstrated by the bird radio studies, and the methods may not be robust to misspecification of the error model. Marques’ take-home message was that distance measurement error should be minimized wherever possible in the field through the use of, e.g., training, calibration exercises, appropriate field methods (e.g., allowing observers to move around during or after point transects to better locate birds) and technological aids (e.g., laser binoculars for visual surveys).

5.3 Bird Misclassification

R. Webster presented some joint work with K.H. Pollock and T.R. Simons on dealing with misidentification of individual birds in certain approaches to estimation using point count data. He focused on the time of detection approach, although this work has clear relevance to multiple observer approaches as well. The basic problem involves the misidentification of individual birds in the different intervals, such that detection histories contain errors. For example, assume that a bird is detected in both periods of a 2-period point count. Assume that the bird is misidentified as a new bird in period 2, so that the true detection history, 1 1, gives rise to two incorrect detection histories, 1 0 and 0 1. Webster modeled this problem by introducing a correct classification probability, α. For example, if p_t denotes detection probability for period t, then the underlying probability associated with detection history 1 0 would be written as: $p_1(1 - p_2) + p_1 p_2(1 - \alpha)$, with the second additive term corresponding to a misidentification of the original bird as a different individual. Webster noted that he was experimenting with approaches to fitting such models, including use of a χ^2 loss function. Subsequent discussion (e.g., by W.A. Link, L. Thomas) focused on the possibility of accounting for the dependencies among the detection histories resulting from a single study with misidentification.

5.4 Surveys of Cryptic Species

S.T. Buckland presented three recently-developed methods based on distance sampling that are appropriate for cryptic species not surveyed well by conventional methods. The first is a "lure point transect" (Buckland et al. 2006), where lures such as playbacks of territorial songs are used to elicit a response from animals of the target species. Conventional methods are not applicable because it is likely that the animal will respond by approaching the surveyors before detection – causing a positive bias in estimates of abundance using conventional estimators. Instead, the detection function is estimated in a separate survey, where trials are conducted by two observation teams, both of which search for animals without using the lure. When one team detects an animal, the other team deploys the lure and records whether they detect a response. A set of such trials allows a binary regression to be used to determine probability of detection given distance of the animal from the lure, and therefore average detection probability. If this detection probability can be assumed to apply also to birds detected in the main survey (for which their initial location before the lure was used is not known) then it is possible to calculate abundance. This method has been used in a survey of Scottish crossbills (Buckland et al. 2006).

The second approach is a "trapping point transect", which is essentially identical to the above approach, except that a trap is used to capture animals rather than a lure. Trials can be set up in much the same manner as for lure point transects: a marked animal is located (e.g., because it is radio-collared, or because it has been released from a different trap) at the same time as the trial trap is set a known distance away, and whether the animal is caught in the trial trap or not within some fixed time interval is recorded. The methods are described in Buckland et al. (2006) and are being tried on an endangered woodrat in Florida. Note that if repeated captures of marked animals in different traps are anticipated then standard mark-recapture methods may be used for estimating animal abundance, or the more recent methods of Efford (2004) and Efford et al. (2005).

The third approach is crossed line (or strip) transects (Buckland et al. 2007a), which were mentioned in Section 5.1. They are potentially suitable for sedentary cryptic objects (such as cryptic plants or bird nests) and allow for probability of detection on the transect line to be less than 1 and for non-independence of animal distribution with respect to the transect.

6 Space-time Variation in Abundance

Although much of the workshop emphasis was on approaches for dealing with the various components of detection, the inferences of most interest to biologists, and thus the ultimate objective of our efforts, involves variation of density or abundance over space and/or time. Several workshop attendees were involved in such work and gave brief presentations on their directions.

6.1 Spatial Variation

J.A. Royle presented examples of work on spatial modeling of avian abundance and occupancy based on avian count data. Specifically, he focused on replicate counts as a basis for inference about abundance (Royle 2004) and occupancy (Royle and Nichols 2003). In the terminology of his presentation on replicate count approaches, his focus in this presentation was on the process model component rather than the binomial component dealing with observation. He noted that the Poisson distribution provides a reasonable model for spatial distribution of abundance, N_i for location i. Modeling can be based on the Poisson mean, λ, and linear models can be developed for $\log\lambda_i$ as a function of covariates associated with site i. If counts exhibit excess variation with respect to the Poisson, then the negative binomial provides an alternative model for spatial distribution, and the negative binomial mean can again be modeled as a function of site-specific covariates. Royle presented several examples in which this basic approach was used for mapping avian distribution as a function of environmental covariates (Kery et al. 2005; Royle et al. 2005).

R. Webster was similarly interested in spatial modeling (with K.H. Pollock and T.R. Simons), but used the time of detection approach as a basis for modeling. Abundance was modeled as a Poisson–lognormal mixture. He explored the use of conditional autoregressive models in which abundance at a site was modeled as a function not only of selected site covariates (e.g., elevation), but also as a weighted (by distance) function of abundances at nearby sites. He noted that the approach provided a reasonable means of investigating spatial processes and that it could be used with repeated count, as well as time of detection, data.

S.T. Buckland described the general framework of Hedley (2000), in which the location of animals is seen as a realization of a random process with some spatially-indexed intensity (the density), and the number of animals in a given area can therefore be described by an inhomogeneous Poisson process (IPP). The detection process represents a thinning of the IPP, which yields another IPP. Given a parametric form for the intensity and detection processes, a likelihood can be derived, but it is often more convenient to treat the estimation of parameters for the two processes separately as they often operate on very different scales (the detection process over a small scale – e.g., 10's or 100's of meters, while we usually wish to model spatial variation in density at much larger scales). This notion has led to the two-stage methods for spatial modeling of line transect data of Hedley and Buckland (2004) and Hedley et al. (2004). Spatial smoothing is accomplished using generalized additive modeling, and improved methods are in development to deal with problems caused by irregular topography (Wood et al. in press) and local clustering that is not explained by the large-scale smooth (Bravington et al. in prep.).

6.2 Temporal Variation

A general overview of methods for temporal inferences when detection is uncertain (but focusing on distance sampling) is given by Thomas et al. (2004). One distinction made in that paper is between empirical modeling and process (or mechanistic) modeling. Aspects of the former approach were covered in a talk by W.A. Link (in joint work with J.R. Sauer) based on work with data from the North American Breeding Bird Survey (BBS) and the Christmas Bird Count (CBC) survey. Their hierarchical modeling approach to BBS data models the logarithm of the expected count at a survey site as a linear function with stratum-specific intercept, time slope, and year effect, with observer effects based on observer identity and year (reflecting temporal changes in observer abilities), and an overdispersion parameter (e.g., Link and Sauer 2002, 2007; Sauer and Link 2002). The year, observer and overdispersion are modeled as random effects. Their modeling of CBC data includes observer effort, rather than observer identity. Such models can be fit using BUGS (Spiegelhalter et al. 1995). Time trend and annual indices to abundance can be derived from estimates under such models.

Link and Sauer (2007) noted that these continental surveys have the potential to provide information about seasonality of population change. In particular, models for the logarithm of counts from each survey include components reflecting proportional change between approximately January and June and between July and December. Such modeling permits inference about the relative amount of temporal variation associated with different seasons of the year, an inference of potential use for managers.

Link and Sauer noted current interest in estimates not only of trend, but also of population size. They noted that BBS has two primary deficiencies that preclude direct inferences about abundance: the absence of information about detection probability on surveyed routes, and the restriction of the BBS to roads. Experimental work using distance and time of detection approaches on both on- and off-road routes provided data that can be incorporated into BBS analyses, again using a hierarchical modeling approach, to obtain abundance estimates.

Process modeling was discussed by L. Thomas, who pointed out that this approach is becoming increasingly popular as recent developments in computer intensive (largely Bayesian) statistical methods allow reasonably complex, high dimensional models to be fit to biological data. Overviews of recent developments in this area are given by Newman et al. (2006) and Buckland et al. (2007b), and an example using bird (lapwing) count data is given by Besbeas et al. (2005). Thomas described his work fitting biological models of grey seal population dynamics to data on seal pup numbers, using a Bayesian fitting algorithm called particle filtering or sequential importance sampling (early work is described in Thomas et al. 2005).

7 Discussion

The workshop format emphasized participant discussion, with the result that useful discussion followed virtually every presentation. Here we attempt to capture some of the ideas expressed, especially on topics that recurred throughout the workshop. One of these topics was the issue of Section 2.2 about whether to use raw counts as indices or to instead attempt to collect ancillary data in order to try to better deal with detection probabilities when drawing inferences about abundance and its variation over time and space. It seems clear that abundance estimation requires some sort of estimate of detection probability in order to translate counts into estimates of abundance (2). However, inferences about temporal or spatial variation in abundance do not require abundance estimates. Raw counts can be used for inferences about temporal or spatial variation in cases where (1) detection probabilities over the dimension of interest (space, time) are similar or (2) the primary sources of variation in detection probability over the dimension of interest are identified, measured, and used as covariates. This latter approach is only appropriate when the primary sources of variation in detection probability do not have the potential to also be associated with variation in abundance.

An important question about inferences for variation in abundance is then: when is it best to collect ancillary data needed to draw specific inference about detection probabilities and when is it better to forego the collection and subsequent modeling of ancillary data and base inference on counts themselves? Many participants held a belief similar to that expressed in books such as Skalski and Robson (1992), Borchers et al. (2002), and Williams et al. (2002), that collection and modeling of ancillary data was wise, even if model selection has the potential to result in inferences being based on models assuming constant detection probabilities. These participants argued that collection and modeling of data on detection probability guarded against incorrect inferences about abundance that can be caused by detection probabilities that differ over the dimension of comparison. Participants who did not necessarily agree with this approach suggested that the modeling associated with the detection process was also likely to produce incorrect inferences in some cases. All agreed that if detection probability is not formally incorporated into the estimation and modeling, then design issues become more important as the only means of dealing with variation in detection probability over dimensions of interest (e.g., via standardization).

In considering this difference in opinion, we recall the early days of capture-recapture modeling, in which some investigators advocated use of these models whereas others suggested that underlying model assumptions were too restrictive and that enumeration methods were preferable. Capture–recapture proponents investigated performance of model-based estimators in the face of likely assumption violations (e.g., Carothers 1973, 1979; Gilbert 1973) and in some cases compared performance with direct enumeration approaches (e.g., Jolly and Dickson 1983; Nichols and Pollock 1983; Conn et al. 2004). These exercises led many investigators to the conclusion that capture–recapture models incorporating detection probability parameters are typically preferable to raw captures as a basis for inference, even in the face of common assumption violations. We believe that such exercises investigating the robustness of the various point count approaches to likely assumption violations, such as those reported by M.G. Efford (with D.K. Dawson), would be useful here as well. In addition to computer simulation work, test systems such as the N.C. State University bird radio system could be very useful in such work.

A conclusion emerging from the bird radio system work about which there was general agreement was the inability to prescribe an omnibus estimation approach likely to be the wise choice in all situations. Instead, the need to tailor methods to the particular study was emphasized, with such features as structural habitat type, relative likelihood of visual versus auditory detections, species richness and study objectives (e.g., time trend versus habitat variation) all identified as important determinants of method selection.

Weaknesses associated with each of the basic estimation approaches were readily identified. Many of these specific weaknesses were associated with auditory detections and the general difficulty in localizing bird sound. For example, this difficulty led to problems with determining matches of individual birds between multiple independent observers. Inability to localize sound also produces problems in distinguishing multiple observations of the same individual within a point count (e.g., double counting and phantom individuals), and such problems can affect all of the basic methods. Difficulties in determining distances at which birds are heard cause problems not only with distance sampling but with all approaches using fixed radius plots. In addition to sound localization, undetected movement in or out of the sample plot has the potential to lead to bias under all of the considered methods, causing more difficulties with some approaches than others. Simulation results presented by M.G. Efford (with D.K. Dawson) further indicated that distance-induced heterogeneity in detection probabilities had substantial potential to cause negative bias in estimators of abundance that did not explicitly account for distance (Efford and Dawson in prep.). If these approaches are used, the suggestion was made to limit the effective area surveyed at each line or point to only include individuals that have naturally "high" detection probabilities, in the sense that they are close to the observer. Distance estimation is also typically easier for closer animals.

The bird radio system test results led to the conclusion that it will likely not be possible to use point count data to draw strong inferences about populations of all species in even a moderately rich bird community. This conclusion led to the recommendation that sampling be restricted to a small number of focal species, but this recommendation was countered by the claim that sometimes the entire bird community is of interest. Subsequent discussion included the observation that many avian studies and monitoring programs are not well-conceived in the sense of having concise objectives. All participants agreed on the importance of establishing clear study objectives and ensuring that sampling was consistent with these objectives. It was noted that in cases where community-level inferences really are of primary interest, then occupancy estimation and modeling (MacKenzie et al. 2002, 2006; Dorazio and Royle 2005) could provide a methodological alternative to abundance estimation that would yield achievable results consistent with many kinds of study objectives. Some participants suggested designs in which the observers focus on 1–3 species for abundance estimation and collect occupancy data on remaining species, whereas another recommendation was to devote one person to the focal species and an additional observer to collect occupancy data for the entire community.

Another aspect of study objectives that is relevant to method selection is whether one goal is to use abundance estimates on the selected sample units to estimate abundance for the entire area from which the samples were drawn. Such extrapolation seems reasonable for many of the methods, conditional on the ability to determine whether each detected bird is inside or outside of the prescribed sample unit. However, such extrapolation is expected to be more difficult for repeated count approaches (Royle and Nichols 2003; Royle 2004). The fact that detection probability estimates under these approaches include the probability of a bird whose range overlaps the sample unit actually being present in the unit at the time of sampling (p_p), means that resulting abundance estimates should include all birds with ranges that overlap sample units (Section 3.3). Extrapolation based on such estimates should estimate the product of overall abundance and the average number of sample units covered by an individual bird's home range (Section 3.6). If sample unit size is large relative to individual home range, this sample-units-per-home-range multiplier is likely to be small, so overall abundance estimates should be fairly good. If sample units and home ranges are of similar size, then it will be more difficult to estimate overall abundance from repeat count estimates without separately estimating the average number of units overlapped by the range of an individual bird. Note that such extrapolation to an overall abundance estimate will not be required by many objectives.

Participants discussed possible changes to existing methods that might be worthy of future consideration. For example, several participants noted that some movement of observers within a plot for point counts, rather than standing at a single point, would likely increase detection probabilities of birds in a sample unit. However, it was suggested that another consequence of such observer movement would likely be a decrease in ability to keep track of individual birds, with subsequent increases in the problems of double counting and phantom birds. Observer movement might induce unwanted bird movement as well. For multiple-observer approaches, a suggestion was made to place observers at different edges of a sample plot as a means of reducing the correlation in detection probabilities of individual birds by the different observers, thus decreasing the problems associated with individual heterogeneity. Once again, though, this approach could lead to an increase in matching and other errors associated with localizing individual birds and determining which birds were detected by one or multiple observers.

Recommended durations of point counts have long been discussed and debated. The explicit recognition of a probability that a bird vocalizes and becomes available for detection in auditory surveys (p_a) leads to the natural suggestion to increase point count duration in order to increase p_a. However, increases in survey duration increase such problems as bird movement into and out of the study plot and difficulties in tracking individual birds over time. These problems can be reduced to some degree by employing a "snapshot" approach (Buckland et al. 2006).

A relatively new approach to collection of field data was that proposed by M.G. Efford (with D.K. Dawson) of using acoustic recording devices (e.g., Hobson et al. 2002) to collect bird vocalization data for subsequent use with estimation approaches. Software would then be developed to help extract the relevant information (species and individual identification) about bird vocalizations. An array of such recording devices could be used with models of the type developed for other passive detectors (e.g., Efford 2004; Efford et al. 2005; Borchers and Efford in press) to directly estimate bird density.

During discussion, Thomas speculated that it may be possible to develop a highly portable acoustic array that could be used to provide real-time localizations of received signals (such as bird vocalizations). Such a system could then be used as a field aid, to provide better estimation of distance in aural surveys than is possible using human ears. More reliable estimation of location than is possible for humans may be achieved if more than two sensors are used, or if the sensors are located farther apart than human ears.

This proposed work by Efford and Dawson on passive arrays of recording devices was among the top proposals for future work recommended by workshop participants. This work will include the recording devices themselves, computer software for resulting sound data, and estimation approaches for computing density estimates from these data. Additional work on combining data and methods was another recommendation of virtually all participants. Research on singing rates and patterns was identified as a priority research problem. For example, current time of detection approaches typically assume a random process in which all birds in the survey area exhibit some probability of vocalizing during each time interval. However, it is likely that this probability is better viewed as resulting from a Markov process in which vocalization during one interval is partially dependent on whether the bird vocalized during the previous interval(s). The likely influence of vocalizations of other birds on singing rates and patterns was also noted, potentially leading to very complicated models of the stochastic process of bird vocalization. Additional research on models of individual bird misidentification (Section 5.3) was also judged to be an important direction of future work. There is no question that double counting and phantom individuals occur and can influence all of the recommended estimation approaches, and formally accounting for the additional uncertainty associated with these errors will represent an important advance.

Finally, there were some very pragmatic recommendations for future work. The most obvious was to field test the various proposed methods using such testing platforms as the bird radio system. The objective of these tests would simply be to determine which estimation approaches “work best” under various conditions and for various objectives. Such field testing could also be used to establish an approximate relationship between estimator performance and number of species considered in the sampling. Knowledge of such a relationship could then lead to the development of rules of thumb about how many focal species can be surveyed with the expectation of drawing reasonable inferences. Recommendations also included work on survey design and allocation of effort based on survey objectives. Practical considerations about effort devoted to visiting many sample units versus estimating components of detection probability for a unit are important and may lead to recommendations about the use of double sampling (sensu Pollock et al. 2002).

Substantial research has been conducted on point counts during the last 5 years. The workshop and this paper represent an attempt to summarize key results of that work and to point to key areas of future research. We offer the opinion that the various recommendations for future research should not be used as reasons for postponing consideration of changes to existing point count programs. Although future work will certainly lead to more refined recommendations, we believe that many existing programs that use point counts can be improved based on the current state of knowledge.

Acknowledgments We are extremely grateful to the workshop participants for attending, self funded, and for sharing so freely their research findings and ideas. We are also grateful for their comments and suggestions on earlier drafts of this manuscript, which have led to a much improved paper. Participants were: Matthew W. Alldredge, David L. Borchers, Stephen T. Buckland, Paul B. Conn, Duane R. Diefenbach, Paul F. Doherty, Murray G. Efford, Douglas H. Johnson, Marc Kéry, William A. Link, Tiago A. Marques, James D. Nichols, Krishna Pacifici, Kenneth H. Pollock, J. Andrew Royle, Michael Runge, Theodore R. Simons, Len Thomas and Ray Webster. We thank Kenneth P. Burnham, Rachel M. Fewster, Marshall Howe, and Bruce Peterjohn for reviews and constructive comments.

References

Alldredge MW, Pollock KH, Simons TR, Shriner SA (2007b) Multiple species analysis of point count data: a more parsimonious modeling framework. J. Appl. Ecol. 44:281–290.

Alldredge MW, Simons TR, Pollock KH (2007c) Factors affecting aural detections of songbirds. Ecol. Appl. 17:948-955.

Alldredge MW, Pollock KH, Simons TR (2006) Estimating detection probabilities from multiple-observer point counts. Auk 123:1172–1182.

Alldredge MW, Pollock KH, Simons TR, Collazo JA, Shriner SA (2007a) Time of detection method for estimating abundance from point count surveys. Auk 124:653-664.

Alpizar-Jara R, Pollock KH (1996) A combination line transect and capture–recapture sampling model for multiple observers in aerial surveys. Environ. Ecol. Stat. 3:311–327.

Alpizar-Jara R, Pollock KH (1999) Combining line transect and capture–recapture for mark-resighting studies. Pages 99–114 *in* Garner GW, Amstrup SC, Laake JL, Manly BFJ, McDonald LL, Robertson DG (eds.) Marine mammal survey and assessment methods. Balkema, Rotterdam.

Anderson DR (2001) The need to get the basics right in wildlife field studies. Wildl. Soc. Bull. 29:1294–1297.

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

Besbeas P, Freeman SN, Morgan BJT (2005) The potential for integrated population modeling. Aust. N.Z. J. Stat. 47:35–48.

Blondel J, Ferry C, Frochot B (1970) La methode des indices ponctuels d'abondance (IPA) ou des releves d'avifaune par "stations d'ecoute." Alauda 41:63–84.

Borchers DL (1999) Composite mark-recapture line transect surveys. Pages115–126 *in* Garner GW, Amstrup SC, Laake JL, Manly BFJ, McDonald LL, Robertson DG (eds.) Marine mammal survey and assessment methods. Balkema, Rotterdam.

Borchers DL, Buckland ST, Goedhart PW, Clarke ED, Hedley SL (1998a) Horvitz–Thompson etsimators for double-platform line transect surveys. Biometrics 54:1221–1237.

Borchers DL, Buckland ST, Zucchini W (2002) Estimating animal abundance. Springer, New York.

Borchers DL, Efford MG. Spatially explicit maximum likelihood methods for capture-recapture studies. Biometrics (in press).

Borchers DL, Laake JL, Southwell C, Paxton CGM (2006) Accommodating unmodeled heterogeneity in double-observer distance sampling surveys. Biometrics 62:372–378.

Borchers DL, Marques TA, Gunnlaugsson T, Víkingsson GA. Distance sampling with measurement errors. (in prep.).

Borchers DL, Zucchini W, Fewster RM (1998b) Mark-recapture models for line transect surveys. Biometrics 54:1207–1220.

Bravington MV, Hedley SL, Wood SN. A general approach for modelling clustered line transect data using gamma random fields. (in prep.)

Buckland ST (2006) Point transect surveys for songbirds: robust methodologies. Auk 123: 345–357.

Buckland ST, Anderson DR, Burnham KP, Laake JL (1993) Distance sampling: estimating abundance of biological populations. Chapman and Hall, London, UK.

Buckland ST, Anderson DR, Burnham KP, Laake JL, Borchers DL, Thomas L (2001) Introduction to distance sampling. Oxford University Press, Oxford, UK.

Buckland ST, Borchers DL, Johnston A, Henrys PA, Marques TA (2007a) Line transect methods for plant surveys. Biometrics 63:989–998.

Buckland ST, Newman KB, Fernández C, Thomas L, Harwood J (2007b) Embedding population dynamics models in inference. Stat. Sci. 22:44–58.

Buckland ST, Summers RW, Borchers DL, Thomas L (2006) Point transect sampling with traps or lures. J. Appl. Ecol. 43:377–384.

Burnham KP, Anderson DR, Laake JL (1980) Estimation of density from line transect sampling of biological populations. Wildl. Monogr. 72:1–202.

Burnham KP, Anderson DR, Laake JL (1981) Line transect estimation of bird population density using a Fourier series. Stud. Avian Biol. 6:466–482.

Burnham KP, Buckland ST, Laake JL, Borchers DL, Marques TA, Bishop JRB, Thomas L (2004) Further topics in distance sampling. Pages 307–392 *in* Buckland ST, Anderson DR, Burnham KP, Laake JL, Borchers DL, Thomas L, eds. Advanced distance sampling. Oxford University Press, Oxford, UK.

Carothers AD (1973) The effects of unequal catchability on Jolly-Seber estimates. Biometrics 29:79–100.

Carothers AD (1979) Quantifying unequal catchability and its effect on survival estimates in an actual population. J. Anim. Ecol. 48:863–869.

Carroll RJ, Lombard F (1985) A note on *N* estimators for the binomial distribution. J. Amer. Stat. Assoc.80:423–426.

Cochran WG (1977) Sampling techniques. Third ed. Wiley, New York, USA.

Conn PB, Bailey LL, Sauer JR (2004) Indexes as surrogates to abundance for low-abundance species. Pages 59–74 *in* Thompson WL (ed.) Sampling rare or elusive species. Island Press, Washington, DC, USA.

Cook RD, Jacobsen JO (1979) A design for estimating visibility bias in aerial surveys. Biometrics 35:735–742.

Diefenbach DR, Marshall MR, Mattice JA, Brauning DW (2007) Incorporating availability for detection in estimates of bird abundance. Auk 124:96-106.

Dorazio RM, Royle JA (2005) Estimating the size and composition of biological communities by modeling the occurrence of species. J. Amer. Stat. Assoc. 100:389–398.

Efford MG (2004) Density estimation in live-trapping studies. Oikos 106:598–610.

Efford MG, Dawson DK. The effect of distant - related heterogeneity on population size estimates from point counts (in prep.)

Efford MG, Warburton B, Coleman MC, Barker RJ (2005) A field test of two methods for density estimation. Wildl. Soc. Bull. 33:731–738.

Ellingson AR, Lukacs PM (2003) Improving methods for regional landbird monitoring: a reply to Hutto and Young. Wildl. Soc. Bull. 31:896–902.

Emlen JT (1971) Population densities of birds derived from transect counts. Auk 88:323–342.

Farnsworth GL, Nichols JD, Sauer JR, Fancy SG, Pollock KH, Shriner SA, Simons TR (2005) Statistical approaches to the analysis of point count data: a little extra information can go a long way. Pages 736–743 *in* Ralph CJ, Rich TD (eds.) Bird Conservation Implementation and Integration in the Americas: Proceedings of the 3rd International Partners in Flight Conference. Volume 2. Gen. Tech. Rep. PSW-GTR-191. Pacific Southwest Research Station, Forest Service, U.S. Dept. Agriculture: Albany, CA.

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

Gilbert RO (1973) Approximations of the bias in the Jolly-Seber capture–recapture model. Biometrics 29:501–526.

Hedley SL (2000) Modelling heterogeneity in cetacean surveys. Ph.D. thesis, Univ. St. Andrews, St. Andrews, Scotland, UK.

Hedley SL, Buckland ST (2004) Spatial models for line transect sampling. J. Agric. Biol. Environ. Stat. 9:181–199.

Hedley SL, Buckland ST, Borchers DL (2004) Spatial distance sampling models. Pages 48–70 *in* Buckland ST, Anderson DR, Burnham KP, Laake JL, Borchers DL, Thomas L (eds.) Advanced distance sampling. Oxford University Press, Oxford, UK.

Hiby L, Lovell P (1998) Using aircraft in tandem formation to estimate abundance of harbour porpoise. Biometrics 54:1280–1289.

Hobson KA, Rempel RS, Greenwood H, Turnbull B, Van Wilgenburg SL (2002) Acoustic surveys of birds using electronic recordings: new potential from an omnidirectional microphone system. Wildl. Soc. Bull. 30:709–720.

Hutto RL, Young JS (2002) Regional landbird monitoring: perspectives from the northern Rocky Mountains. Wildl. Soc. Bull. 30:738–750.

Hutto RL, Young JS (2003) On the design of monitoring programs and the use of population indices: a reply to Ellingson and Lukacs. Wildl. Soc. Bull. 31:903–910.

Jarvinen O, Vaisanen RA (1975) Estimating relative densities of breeding birds by the line transect method. Oikos 26:316–322.

Jolly GM, Dickson JM (1983) The problem of unequal catchability in mark-recapture estimation of small mammal populations. Can. J. Zool. 61:922–927.

Kendall WL, Nichols JD, Hines JE (1997) Estimating temporary emigration and breeding proportions using capture–recapture data with Pollock's robust design. Ecology 78:563–578.

Kery M, Royle JA, Schmid H (2005) Modeling avian abundance from replicated counts using binomial mixture models. Ecol. Appl. 1450–1461.

Kissling ML, Garton EO (2006) Estimating detection probability and density from point-count surveys: a combination of distance and double-observer sampling. Auk 123: 735–752.

Laake JL, Borchers DL (2004) Methods for incomplete detection at distance zero. Pages 108–189 *in* Buckland ST, Anderson DR, Burnham KP, Laake JL, Borchers DL, Thomas L (eds.) Advanced distance sampling. Oxford University Press, Oxford, UK.

Lancia RA, Kendall WL, Pollock KH, Nichols JD (2005) Estimating the number of animals in wildlife populations. Pages 106–153 *in* Braun CE ed. Research and management techniques for wildlife and habitats. The Wildlife Society, Bethesda, MD, USA.

Lancia RA, Nichols JD, Pollock KH (1994) Estimating the number of animals in wildlife populations. Pages 215–253 *in* Bookhout T (ed.) Research and management techniques for wildlife and habitats. The Wildlife Society, Bethesda, MD, USA.

Link WA (2003) Nonidentifiability of population size from capture–recapture data with heterogeneous detection probabilities. Biometrics 59:1123–1130.

Link WA, Sauer JR (1997) Estimation of population trajectories from count data. Biometrics 53:499–497.

Link WA, Sauer JR (1998) Estimating relative abundance from count data.Austrian J. Stat. 27: 83–97.

Link WA, Sauer JR (2002) A hierarchical analysis of population change with application to Cerulean Warblers. Ecology 83:2832–2840.

Link WA, Sauer JR (2007) Seasonal components of avian population change: joint analysis of two large-scale monitoring programs. Ecology 88:49–55.

Mackenzie DI, Nichols JD, Lachman GB, Droege S, Royle JA, Langtimm CA (2002) Estimating site occupancy when detection probabilities are less than one. Ecology 83: 2248–2255.

Mackenzie DI, Nichols JD, Royle JA, Pollock KH, Bailey LA, Hines JE (2006) Occupancy modeling and estimation. Academic Press, San Diego, CA, USA.

MacKenzie DI, Royle JA (2005) Designing occupancy studies: general advice and allocating survey effort. J. Appl. Ecol. 42:1105–1114.

Manly, B.F.J., McDonald LL, Garner GW (1996) Maximum likelihood estimation for the double-count method with independent observers. J. Agric. Biol. Environ. Stat. 1: 170–189.

Marques TA (2004) Predicting and correcting bias caused by measurement error in line transect sampling using multiplicative error models. Biometrics 60:757–763.

Marques TA (2008) Incorporating measurement error density gradients in distance sampling surveys. Ph.D. thesis, Univ. St. Andrews, St. Andrews, Scotland, UK.

Marques TA, Thomas L, Fancy SG, Buckland ST (2007) Improving estimates of bird density using multiple covariate distance sampling. Auk (124:1229-1245).

Marsh H, Sinclair DF (1989) Correcting for visibility bias in strip transect aerial surveys of aquatic fauna. J. Wildl. Manage. 53:1017–1024.

Newman KB, Buckland ST, Lindley ST, Thomas L, Fernández C (2006) Hidden process models for animal population dynamics. Ecol. Appl. 16:74–86.

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

Nichols JD, Pollock KH (1983) Estimation methodology in contemporary small mammal capture–recapture studies. J. Mammal. 64:253–260.

Norris III JL, Pollock KH (1996) Nonparametric MLE under two closed capture–recapture models with heterogeneity. Biometrics 52:639–649.

Okamura H, Minamikawa S, Kitakado T (2006) Effect of surfacing patterns on abundance estimates of long-diving animals. Fish. Sci. 72:631–638.

Otis DL, Burnham KP, White GC, Anderson DR (1978) Statistical inference from capture data on closed animal populations. Wildl. Monogr. 62:1–35.

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

Pollock KH (1982) A capture–recapture design robust to unequal probability of capture. J. Wildl. Manage. 46:757–760.

Pollock KH, Nichols JD, Simons TR, Farnsworth GR, Bailey LL, Sauer JR (2002) Large scale wildlife monitoring studies: statistical methods for design and analysis. Environmetrics 13: 1–15.

Ralph CJ, Sauer JR, Droege S (eds.) (1995) Monitoring bird populations by point counts. Gen. Tech., Rep. PSW-GTR-149. U.S. Forest Service Pacific Southwest Research Station, Albany, CA, USA.

Ralph CJ, Scott JM (eds.) (1981) Estimating numbers of terrestrial birds. Stud. Avian Biol. No. 6:1–630.

Ramsey FL, Scott JM (1979) Estimating population densities from variable circular plot surveys. Pages 155–181 in Cormack RM, Patil GP, Robson DS (eds.) Sampling biological populations. Statistical Ecology Series, Vol. 5, International Cooperative Publication House, Fairland, MD, USA.

Rosenstock SS, Anderson DR, Giesen KM, Leukering T, Carter MF (2002) Landbird counting techniques: current practices and an alternative. Auk 119:46–53.

Royle JA (2004) N-mixture models for estimating population size from spatially replicated counts. Biometrics 60:108–115.

Royle JA, Nichols JD (2003) Estimating abundance from repeated presence absence data or point counts. Ecology 84:777–790.

Royle JA, Nichols JD, Kery M (2005) Modeling occurrence and abundance of species when detection is imperfect. Oikos 110:353–359.

Sauer JR, Link WA (2002) Hierarchical modeling of population stability and species group attributes from survey data. Ecology 86:1743–1751.

Seber GAF (1982) The estimation of animal abundance and related parameters. Second ed. MacMillian Publication Co., Inc., New York, USA.

Simons TR, Alldredge MW, Pollock KH, Wettroth JM (2007) Experimental analysis of the auditory detection process on avian point counts. Auk 124:986-999.

Skalski JR, Robson DS (1992) Techniques for wildlife investigations. Academic Press, San Diego, USA.

Smith GW (1995) A critical review of the aerial and ground surveys of breeding waterfowl in North America. U.S. Dept. Interior, Biological Science Report 5. Washington, DC, USA.

Spiegelhalter DJ, Thomas A, Best NG, Gilks WR (1995) BUGS: Bayesian inference using Gibbs sampling. Version 0.50. MRC Biostatistics Unit, Cambridge, UK.

Strindberg S, Buckland ST, Thomas L (2004) Design of distance sampling surveys and geographic information systems. Pages 190–228 *in* Buckland ST, Anderson DR, Burnham KP, Laake JL, Borchers DL, Thomas L (eds.) Advanced distance sampling. Oxford University Press, Oxford, UK.

Thomas L, Borchers DL, Buckland ST, Hammond IE. Estimation of population size from distance sampling data with heterogenous detection probabilities. (in prep.)

Thomas L, Buckland ST, Newman KB, Harwood J (2005) A unified framework for modelling wildlife population dynamics. Aust. N. Z. J. Stat. 47:19–34.

Thomas L, Burnham KP, Buckland ST (2004) Temporal inferences from distance sampling surveys. Pages 71–107 *in* Buckland ST, Anderson DR, Burnham KP, Laake JL, Borchers DL, Thomas L (eds.) Advanced distance sampling. Oxford University Press, Oxford, UK.

Thompson SK (2002a) Sampling. Wiley, New York, USA.

Thompson WL (2002b) Towards reliable bird surveys: accounting for individuals present but not detected. Auk 119:18–25.

Williams BK, Nichols JD, Conroy MJ (2002) Analysis and management of animal populations. Academic Press, San Diego, USA.
Wood SN, Bravington MV, Hedley SL. Soap film smoothing. J. Royal Stat. Soc. Ser. B. (in press).
Yoccoz NG, Nichols JD, Boulinier T (2001) Monitoring of biological diversity in space and time. Trends Ecol. Evol. 16:446–453.