**Abstract**

Alldredge, Mathew W. Avian point count surveys: Estimating components of the detection process. (Under the direction of Kenneth H. Pollock and Theodore R. Simons)

Point count surveys of birds are commonly used to provide indices of abundance or, in some cases, estimates of true abundance. The most common use of point counts is to provide an index of population abundance or relative abundance. To make spatial or temporal comparisons valid using this type of count requires the very restrictive assumption of equal detection probability for the comparisons being made.

We developed a multiple-independent observer approach to estimating abundance for point count surveys as a modification of the primary-secondary observer approach. This approach uses standard capture-recapture models, including models of inherent individual heterogeneity in detection probabilities and models using individual covariates to account for observable heterogeneity in detection probabilities. Two-observer models provided negatively biased estimates because they do not account for individual heterogeneity in detection probabilities. Models accounting for individual heterogeneity are always selected as the most parsimonious models for this data type.

We also developed a time of detection approach for estimating avian abundance when birds are detected aurally, which is a modification of the time of removal approach. This approach requires collecting detection histories of individual birds in consecutive time intervals and modeling the detection process using a capture-recapture framework. This approach incorporates both the probability a bird is available for detection and the probability of detection given availability. Analyses presented demonstrate the importance of models accounting for individual heterogeneity in detection probabilities.

We recommend time of detection point count surveys be designed with four or more equal intervals.

We also present a multiple species modeling strategy since many point count surveys collect data on multiple species and present the approach for distance sampling, multiple observer, and time of detection approaches. The purpose of using a multiple species modeling approach is to obtain more parsimonious models by exploiting similarities in the detection process among species. We present a method for defining species groups which leads to an *a priori* set of species groups and associated candidate models. Multiple species models worked well and in many cases gave more parsimonious models than a species specific modeling approach, especially for the multiple-observer and time of detection approaches. Parameter estimates for multiple species models are more precise than single species models. We recommend this approach for all situations where data on multiple species is collected.

Finally, we present a method for estimating the availability probability of birds during a point count based on singing rate or detailed singing time data. This approach requires data collected in conjunction with point count surveys that describe the singing rates or singing time distribution of the bird population of interest. The singing rate approach requires the assumption that an individual bird sings following a random process but rates may vary between birds. We modeled this using a finite-mixture Poisson model. The singing time approach is a nonparametric approach and does not require this restrictive assumption. Analysis of Ovenbird singing rate data demonstrates the importance of accounting for availability bias when estimating abundance, especially

as count lengths get short. We recommend this approach when "snapshot" type counts are necessary.

Analyses presented throughout this thesis demonstrate the importance of accurately modeling the detection process to estimate abundance. The importance of accounting for individual heterogeneity in detection probabilities was evident in every chapter. Using a point count method that accounts for individual heterogeneity is crucial to estimating abundance effectively and making valid spatial, temporal and species comparisons.

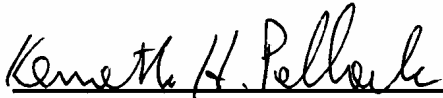# AVIAN POINT COUNT SURVEYS: ESTIMATING COMPONENTS OF THE DETECTION PROCESS
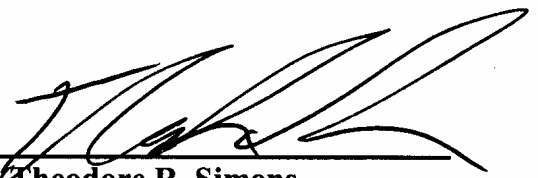
## by

## MATHEW W. ALLDREDGE

A dissertation submitted to the Graduate Faculty of
North Carolina State University
in partial fulfillment of the
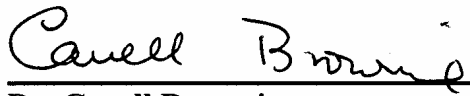requirements for the Degree of
Doctor of Philosophy

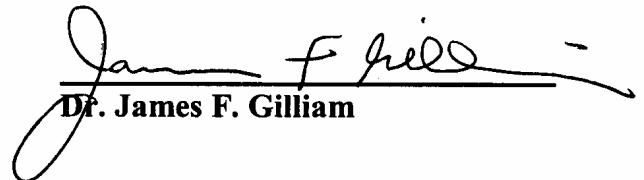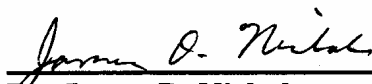Biomathematics/Zoology

Raleigh
2004

APPROVED BY:

_____
Dr. Kenneth H. Pollock
Chair of Advisory Committee

_____
Dr. Theodore R. Simons
Co-chair of Advisory Committee

_____
Dr. Cavell Brownie

_____
Dr. James F. Gilliam

_____
Dr. James D. Nichols

To my family:

I never would have been able to do this without their support and understanding.  Not only did my wife, Stacey, lovingly support me through this process, she also provided the strength, encouragement and motivation to keep me moving forward.  My son, Brennen, was an inspiration, always there to remind me of the things that are really important.  My unborn son (May 6, 2004), reminding me of the miracles and joys that life offers.  Special thanks to my mother who also provided support and encouragement.  Family is the one truly important thing in life.  I have been graciously blessed by my family, thank you.

iii

**Biography**

I grew up in the mountains of Colorado and Wyoming. My free time was always spent outside, skiing, hiking, fishing, hunting and otherwise goofing off in the mountains. I learned to enjoy wildlife at an early age and am always at my best when interacting with it and trying to learn and understand how all the pieces fit together. I also had the opportunity of growing up and interacting with several wildlife biologists, including my father, and gained many valuable experiences helping with their projects.

Seeing the stress and bureaucracy involved with wildlife research and management I thought it might be wise to follow a different career path and let my wildlife interests occupy my free time instead of my profession. I completed an undergraduate degree in mechanical engineering in 1994 and then quickly decided that office life was not for me. I started taking wildlife courses at Colorado State University and worked on a few fish projects. Then I took a job working on a grizzly bear study in Wyoming. While camped at the foot of the Tetons trapping and tracking bears I gained the experience of a lifetime which solidified my career path.

I attended the University of Idaho for my Master's of Science in Wildlife Resources. My master's project examined elk habitat and nutritional relationships on industrial timber lands. My advisor, Jim Peek, taught me more about wildlife and ecosystems than I could ever learn from books.

After completing my master's I took a job as a wildlife biologist and continued studying elk while my wife finished her degree. I really wanted to utilize my quantitative skills and learn more about the population sampling techniques I had started studying at CSU. Ken Pollock's reputation and interest in wildlife sampling led me to North

Carolina State University.  In 2002 I completed a master's in biomathematics.  Having an

interest in population dynamics and modeling as well, I decided to co-major in

Biomathematics and Zoology and threw in a Statistics minor for good measure.

**Acknowledgements**

  First I would like to thank my graduate committee for their support and advice they provided me.  Ken Pollock provided me many opportunities to work on a variety of projects during my four years at NCSU and helped me develop a broad background in sampling wildlife populations.  Ted Simons brought me on board the "Bird Radio" project and continuously advised me on the intricacies of point counts.  Jim Nichols not only was willing to fly down to NCSU for committee meetings but also helped me think through the sampling and modeling issues I proposed and dealt with many questions about model results (the ominous c-hat).  Always remember to "think hard" about something.  Cavell Brownie helped and encouraged me on the statistical development of my research and always had ideas on improving my techniques.  Jim Gilliam, another great thinker, always looked for applications to more general questions and challenged my model assumptions.  I owe the singing time model of Chapter 5 to this.  I would also like to thank Jaime Collazo for his assistance on this project.

  This project was funded by the U.S. Geological Survey and the National Park Service.

  I also appreciate all of the field work that went before me.  Without the accessibility of these data sets for example analyses of my methods my thesis would have been little more than statistical rambling.  I would thank each field technician by name but I know them only by the initials appearing on the data sheets.

  Finally, I would like to thank my family once again for all of their support and encouragement.

**Table of Contents**                                          **Page**

**CHAPTER 3**

**CHAPTER 4**

**CHAPTER 5**

**CHAPTER 6**

**List of Tables**

**Chapter 2:  ESTIMATING DETECTION PROBABILITIES FROM MULTIPLE OBSERVER POINT COUNTS**

## Chapter 3:  TIME OF DETECTION METHOD FOR ESTIMATING ABUNDANCE FROM POINT COUNT SURVEYS

**Chapter 4: MULTIPLE SPECIES ANALYSIS OF POINT COUNT DATA:
A MORE PARSIMONIOUS MODELING FRAMEWORK**

# Chapter 5: MODELING THE AVAILABILITY PROCESS
# FOR POINT COUNT SURVEYS USING AUXILIARY DATA

**List of Figures**

### Chapter 5: MODELING THE AVAILABILITY PROCESS FOR POINT COUNT SURVEYS USING AUXILIARY DATA

# Chapter 1

# Introduction

There are a wide variety of field and statistical techniques for assessing animal abundance, which include complete counts, partial counts, and capture methods (Seber 1982, Lancia et al. 1994, Williams et al. 2002). Rarely is it possible to conduct complete counts as only portions of the area of interest can actually be counted and generally not all animals in the sample areas will be observed. Such counts require that data are collected in a manner that allows for the estimation of the fraction of the population that is sampled. The actual sampling approach used is generally species and/or habitat specific and may depend on the specific research question (Seber 1982, Lancia et al. 1994).

The interest in estimating animal abundance is that it is commonly used as a measure of population health by ornithologists and other biologists (Lack 1954, 1966). Abundance estimates over successive years can provide information on population trends, which can be suggestive of population health (Ralph et al. 1995, Williams et al. 2002). Besides comparing abundance estimates between years it is also possible to compare between spatially distinct areas, which can provide information on habitat relationships or differences associated with management practices (Ralph et al. 1995). Comparisons that may be of interest are between unmanaged areas, such as National Parks, and actively managed areas, such as National Forests or state owned lands. These comparisons can be important tools in adaptive management (Walters and Hilborn 1978) and for understanding changes that occur in animal populations.

Although population estimates provide useful information about the state of animal populations and are the focus of this dissertation, it is worth noting that it is also necessary to obtain other demographic parameters to fully understand the dynamics of a

population. This involves a complete understanding of the losses and gains occurring in a population that are associated with birth, death and migration. Both abiotic and biotic factors can affect population process and these factors must also be examined to fully understand the dynamics of a population (Ricklefs and Miller 2000, Williams et al. 2002).

The lack of "good" quantitative measures of landbird (non-game bird species) abundance in the past comes from two sources; lack of interest, and difficulty in obtaining reasonable estimates (Nichols 1994). The lack of interest stems from the historical concern for game birds and waterfowl which have been actively managed for recreational use (Martin et al. 1979). Problems of obtaining reasonable estimates of landbird abundance have slowed the development of valid statistical techniques but with recent interest in landbird populations there has been a renewed interest in enhancing the available methods. Of the available methods point count surveys are the most widely used method for assessing abundance of landbirds (Ralph et al. 1995).

Recent interest in landbirds is due to concerns over possible declines of landbird populations (Robbins et al. 1986, Askins et al. 1990). These declines have been the motivation for programs such as Partners in Flight (Carter et al. 2000) that includes large scale monitoring programs. The original objective of Partners in Flight was inventory and monitoring of neotropical migrants, but this has been expanded to include other birds of concern.

The number of people participating in national monitoring programs is also evidence of the interest in bird populations. One such survey is the Breeding Bird Survey (BBS), which has been conducted since 1966 (Sauer et al. 1997, 2003). Currently the

BBS consists of about 3,700 active routes (nearly 2,900 surveyed annually) that are distributed across the continental U.S. and Canada.  Each route is 24.5 miles long and has 50 stops per route located at 0.5 mile intervals.  Surveys are conducted during the breeding season each year and are only done on days that satisfy a standardized protocol to try and ensure that detection probabilities are constant over time.  At each stop an observer counts the number of birds detected of all species that are either heard or seen during a three minute interval.

The BBS survey is representative of typical surveys for landbirds (Rosenstock et al. 2002).  A series of points are randomly placed over an area of interest.  Then, using a standardized approach, each point is surveyed for a set amount of time and all birds detected are recorded.  This can be done with either fixed radius plots or unlimited radius plots.  Such a count gives a measure of relative abundance of a population or provides an index to abundance but does not provide an estimate of true abundance.

There have been a number of studies that have demonstrated that both observer differences and environmental conditions affect the number of birds counted (Ralph and Scott 1981).  This led to the standardization of point counts as a means of providing comparable counts both temporally and spatially.  The general idea is that if counts are conducted by the same observer or by observers with similar ability and are always done on days with similar environmental conditions then the proportion of birds counted should be similar between counts.  Methods such as these have been highly criticized throughout the literature on abundance estimation because it is now widely believed that no amount of standardization can account for all of the variation associated with

detection of animals (Burnham 1981, Wilson and Bart 1985, Johnson 1995, Barker and

Sauer 1995, Nichols et al. 2000, Rosenstock et al. 2002, Thompson 2002).

The general model for the relationship between a count statistic ($C_i$) and the true

abundance ($N_i$) is given by (Lancia et al. 1994, Williams et al. 2002)

$$E\ (C_i) = N_i\, p_i \qquad\qquad (1)$$

where i denotes the location or time of the count and $p_i$ is the probability of detection.

The premise behind standardizing counts and obtaining an index to abundance is that the

detection probability is constant across space and time because of the standardization.

With this assumption comparisons of abundance across space and time are made using

the count as an index to abundance.

An alternative approach is to collect count data so that the associated detection

probability can be estimated (Nichols et al. 2000, Farnsworth et al. 2002, Rosenstock et

al. 2002, Thompson 2002).  With this additional information, direct estimation of

abundance is possible as

$$\hat{N}_i = \frac{C_i}{\hat{p}_i} \qquad\qquad (2)$$

Using this approach the resulting estimates of abundance can be used to draw inference

about population differences across space and/or time without the strict assumption of

constant probability of detection.

The detection process consists of two components; the probability that an individual is available for detection, and the probability that an individual is detected given that it is available. The $p_i$ given in equation 2 is the product of these two components (Marsh and Sinclair 1989) such that equation 2 becomes

$$\hat{N}_i = \frac{C_i}{\hat{p}_a \hat{p}_d} \qquad (3)$$

where $p_a$ is the probability that an individual is available for detection and $p_d$ is the probability that an individual is detected given that it is available. Marsh and Sinclair (1989) were concerned with aerial surveys of marine mammals where availability was associated with an individual's position in the water column and sea state. Most estimation methods of animal abundance assume that all animals are available and thus ignore this component. Examination of the original capture-recapture models (Otis et al. 1978, Seber 1982, Williams et al. 2002) shows that these models only estimate the capture probability of animals available for capture, although recent temporary emigration models (Kendall et al. 1997) do account for availability associated with spatial location of an individual. In some situations the availability process may be very important, including bird point counts which may require a bird producing a sound cue before it can be detected.

The general methods for estimating abundance from point count data are distance sampling, multiple observer approaches, the time of detection approach, double sampling and repeated count methods. Distance sampling models the decline in the probability of

detection associated with distance from the observer (Reynolds et al. 1980, Buckland et al. 1993). The multiple observer method uses capture-recapture models to estimate the detection probabilities of each observer (Nichols et al. 2000). The time of detection method also uses capture-recapture models to estimate detection probabilities associated with time intervals of a count (Farnsworth et al. 2002). Of these methods only the time of detection method estimates the product of availability and detection given availability while the other two approaches assume all animals are available. These methods will be reviewed in more detail as this dissertation develops. The double sampling approach (Bart and Earnst 2002) requires a "fast" method to obtain a count and then a "slow" method to resample a portion of the area initially sampled and is assumed to be a census. We do not believe that the double sampling method is appropriate in forested environments. Repeated count methods (Royle and Nichols 2003, Royle 2004) estimate both availability and detection given availability by repeated sampling of a set of points over time. Repeated count methods also include the probability that an individual is in the sample area as this may change between surveys.

The specific objectives of my dissertation are to examine the detection process associated with auditory detection of birds, present some alternative methods for estimating the detection process, and provide examples of these methods. In chapter 2, I will present the multiple independent-observer approach, the relevant models and provide an example analysis using this method. In chapter 3, I will present the time of detection model, of which Farnsworth et al.'s (2002) removal model is a special case, develop covariate models and heterogeneity models and present example analyses. In chapter 4, I will develop multiple species models that exploit similarities in the detection process

among similar species to model the detection probabilities which will give more parsimonious models with better precision. In chapter 5, I will examine the availability process more closely and present models that incorporate this directly from point count data and an approach that uses auxiliary information on singing frequency to model the availability process.

For the time of detection approach it is necessary to use program SURVIV (White 1983) to estimate model parameters when point count surveys are conducted with unequal time intervals. In chapters 3 and 4 we present example analyses from data collected with unequal time intervals. The SURVIV code used for these analyses is given in appendix 1. The single species code is easily modified by changing the values for new data. The multiple species code must be modified to fit the number of species in the analysis. We give the SURVIV code for a four species analysis and the relevant models.

To model the availability process using auxiliary data, it was also necessary to develop computer programs for the analysis. We do not give the code for the bootstrap since this is a standard analysis procedure. To analyze the singing time data we used MATLAB and have included the code in appendix 2. We have also written a Windows based computer program for collecting singing time data. This program can be used on Windows based PDA's. In appendix 3 we show the screen layout of the program and provide the source code.

Literature cited

Askins, R.A., J.F. Lynch, and R. Greenberg. 1990. Population declines of migratory birds of eastern North America. Current Ornithology 7:1057

Barker, R.J. and J.R. Sauer. 1995. Statistical aspects of point count sampling. Pages 125-130 in Monitoring Bird Populations by Point Counts (C.J. Ralph, J.R. Sauer, and S. Droege, Eds.). U.S. Department of Agriculture, Forest Service General Technical Report PSW-GTR-149.

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

Buckland, S.T., D.R. Anderson, K.P. Burnham, and J.L. Laake. 1993. Distance Sampling: Estimating Abundance of Biological Populations. Chapman and Hall, London.

Burnham, K.P. 1981. Summarizing remarks: Environmental influences. Pages 324-325 in Estimating numbers of terrestrial birds (C.J. Ralph and J.M. Scott, Eds.). Studies in Avian Biology No. 6.

Carter, M.F., W.C. Hunter, D.N. Pashley, and K.V. Rosenberg. 2000. Setting conservation priorities for landbirds in the United States: The Partners in Flight Approach. Auk 117:541-548.

Farnsworth, G.L., K.H. Pollock, J.D. Nichols, T.R. Simons, J.E. Hines, and J.R. Sauer. 2002. A removal model for estimating detection probabilities from point-count surveys. Auk 119:414-425.

Johnson, D.H. 1995. Point counts of birds: What are we estimating? Pages 117-123 in Monitoring Bird Populations by Point Counts (C.J. Ralph, J.R. Sauer, and S. Droege, Eds.). U.S. Department of Agriculture, Forest Service General Technical Report PSW-GTR-149.

Kendall, W.L., J.D. Nichols, and J.E. Hines. 1997. Estimating temporary emigration using capture-recapture data with Pollock's robust design. Ecology 78:563-578.

Lack, D. 1954. The Natural Regulation of Animal Numbers. Oxford University Press, London.

Lack, D. 1966. Population Studies of Birds. Clarendon Press, Oxford.

Lancia, R.A., J.D. Nichols, and K.H. Pollock. 1994. Estimating the number of animals in wildlife populations. Pages 215-253 in Research and management techniques for wildlife and habitats (T. Bookhout, Ed.). The Wildlife Society, Bethesda, Maryland.

Marsh, H. and D.F. Sinclair. 1989. Correcting for visibility bias in strip transect aerial surveys of aquatic fauna. Journal of Wildlife Management. 53:1017-1024.

Martin, F.W., R.S. Pospahala, and J.D. Nichols. 1979. Assessment and population management of North American migratory birds. Pages 187-239 in Environmental Biomonitoring, Assessment, Prediction, and Management—Certain Case Studies and Related Quantitative Issues. Statistical Ecology, Vol. S11 (J. Cairns, G.P. Patil, and W.E. Waters, eds.). International Coorperative Publication House, Fairland, MD.

Nichols, J.D. 1994. Capture-recapture methods for bird population studies. Proceedings of Italian Ornithological Congress 6:31-51.

Nichols, J.D., J.E. Hines, J.R. Sauer, F.W. Fallon, J.E. Fallon, and P.J. Heglund. 2000. A double-observer approach for estimating detection probability and abundance from point counts. Auk 117:393-408.

Otis, D.L., K.P. Burnham, G.C. White, and D.R. Anderson. 1978. Statistical inference from capture data on closed animal populations. Wildlife Monographs, No. 62.

Ralph, J.C., S. Droege, and J.R. Sauer. 1995. Managing and monitoring birds using point counts: Standards and applications. Pages 161-168 in Monitoring bird populations by point counts (J.C. Ralph, J.R. Sauer, and S. Droege (Eds.)). United States Forest Service General Technical Report PSW-GTR-149.

Ralph, J.C. and J.M. Scott. 1981. Eds. Estimating numbers of terrestrial birds. Studies in Avian Biology 6:630p.

Reynolds, R.T., J.M. Scott, and R.A. Nussbaum. 1980. A variable circular-plot method for estimating bird numbers. Condor 82:309-313.

Ricklefs, R.E. and G.L. Miller. 2000. Eds. Ecology 4th edition. W.H. Freeman and Company, New York.

Robins, C.S., D. Bystrak, and P.H. Geissler. 1986. The Breeding Bird Survey: Its first fifteen years, 1965-1979. United States Fish and Wildlife Service Resource Publication No. 157.

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

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

Royle, J.A. 2004. N-Mixture models for estimating population size from spatially replicated counts. Biometrics 60:108-115.

Sauer, J.R., J.E. Hines, G. Gough, I. Thomas, and B.G. Peterjohn. 1997. The North American Breeding Bird Survey results and analysis. Version 96.4. Patuxent Wildlife Research Center, Laurel, MD.

Sauer, J.R., J.E. Hines, and J. Fallon. 2003. The North American Breeding Bird Survey, Results and Analysis 1966-2002. Version 2003.1, USGS Patuxent Wildlife Research Center, Laurel, MD.

Seber, G.A.F. 1982. The Estimation of Animal Abundance and Related Parameters (2nd ed.). Edward Arnold, London.

Thompson, W.L. 2002. Towards reliable bird surveys: Accounting for individuals present but not detected. Auk. 119:18-25.
Walters, C.J. and R. Hilborn. 1978. Ecological optimization and adaptive management. Annual Review of Ecology and Systematics 9:157-188.

White, G.C. 1983. Numerical estimation of survival rates from band-recovery and biotelemetry data. Journal of Wildlife Management. 47:716-728.

Williams, B.K., J.D. Nichols, and M.J. Conroy. 2002. Analysis and Management of Animal Populations. Academic Press. San Diego, CA.

Wilson, D.M. and J. Bart. 1985. Reliability of singing bird surveys: Effects of song phenology during the breeding season. Condor 87:69-73.

Chapter 2

# ESTIMATING DETECTION PROBABILITIES FROM MULTIPLE OBSERVER POINT COUNTS

ESTIMATING DETECTION PROBABILITIES FROM MULTIPLE OBSERVER
POINT COUNTS

Mathew W. Alldredge, Biomathematics and Zoology, North Carolina State University,
Raleigh NC 27695.
Kenneth H. Pollock, Zoology, Biomathematics, and Statistics, North Carolina State
University, Raleigh, NC 27695.
Theodore R. Simons, USGS Cooperative Fish and Wildlife Research Unit,  Dept. of
Zoology, North Carolina State University, Raleigh, NC 27695.

Abstract.—Point counts are commonly used to obtain indices of bird population

abundance.  Recent methodological developments, including the dependent-observer

approach of Nichols et al. (2000) estimate detection probabilities which can reduce biases

associated with spatial and temporal variability in detection probability.  We present an

independent-observer point count approach, which is a generalization of the dependent-

observer approach.  The independent-observer approach is essentially a closed population

capture-recapture method.  Additional models can be parameterized using covariates,

such as detection distance, to account for heterogeneity associated with identified sources

of variation.  By comparing abundance estimates from two- and four-observer point

counts we demonstrate a negative bias in two-observer estimates. This negative bias,

caused by unobservable individual heterogeneity in detection probabilities, can be

accounted for when models with four independent observers are used.  In four out of five

data sets examined heterogeneity models were selected, producing abundance estimates

15% to 26% higher than models that did not account for heterogeneity.  The independent-

observer approach is more efficient (smaller variance) than the dependent-observer

approach because it uses the full detection history of an individual and not just first

detections. The method also allows the incorporation of detection distance estimates which account for situations where detection probabilities decline as a function of distance from the observer. Additionally, with four or more observers, the method accounts for individual heterogeneity in detection probabilities which reduces the bias of abundance estimates. Although independent observer methods are expensive and possibly impractical for large scale applications, we believe they can provide important insights into the sources and degree of perception bias (probability of detecting an individual given that it is available for detection) in avian point count estimates, and that they may be useful in a two stage sampling framework to calibrate single observer estimates.

## Introduction

Point counts are used extensively as indices of spatial and temporal differences in bird abundance, and to assess habitat relationships, responses to environmental change or management, and species diversity (Ralph et al. 1995a, Thompson 2002). They are used across a spectrum of scales from long term continental-scale surveys such as the Breeding Bird Survey (Robbins et al. 1986, Sauer et al. 1997, 2003) to short term site specific studies (Ralph et al. 1995a).

There are fundamentally two approaches to abundance estimation using count data. The first generates an abundance index using a standardized approach to control for known sources of bias (e.g. weather, observer skill, time of day and season) (Conroy 1996, Sauer et al. 1997, Williams et al. 2002). The second uses statistical methods that estimate how detection probabilities vary among observers and across space and time

(Nichols et al. 2000, Williams et al. 2002).   Estimated detection probabilities are used to adjust the raw counts to reduce the bias of abundance estimates.

In a review of 224 papers reporting sampling techniques used to draw inference about abundance, 95% relied on index counts (Rosenstock et al. 2002).  Comparisons of index counts across space or time require the strong assumption that the probability of detection is constant for all locations and/or times.  Assumptions of constant detection probability have long been questioned (Burnham 1981, Wilson and Bart 1985, Johnson 1995, Barker and Sauer 1995).  The weakness of these assumptions has motivated considerable research into statistical approaches that estimate detection probabilities directly for all study areas and time periods (Nichols et al. 2000, Bart and Earnst 2002, Farnsworth et al 2002, Rosenstock et al. 2002, Thompson 2002).  The general problem with index counts is that no amount of standardization can account for the unobservable or uncontrollable sources of variation that affect the raw count data (Burnham 1981, Johnson 1995).  Known sources that affect detection probabilities of birds are season (Ralph 1981, Skirvin 1981), time of day (Robbins 1981, Skirvin 1981), stage of nesting cycle (Wilson and Bart 1985), observer effects on singing frequency, habitat characteristics and local species densities (McShea and Rappole 1997), and differences among observers (Sauer et al. 1994).  Conceptually, if 2 count statistics differ across space or time it is not possible to distinguish if the difference is attributable to differences in detection probability (observer differences, effects of habitat structure or other factors affecting detection probability), or actual differences in abundance.  For a thorough comparison of these issues see Nichols et al. (2000) and Rosenstock et al. (2002).

The underlying model for estimating population size from count data is:

$$\hat{N}_i = \frac{C_i}{\hat{p}_i} \quad , \qquad\qquad (1)$$

where the $\hat{N}_i$ is the estimated abundance, $C_i$ is the count, $\hat{p}_i$ is the detection probability, and i denotes the time and location of the count (Lancia et al. 1994, Williams et al. 2002). The probability of detection has two components (Marsh and Sinclair 1989); the probability of being available for detection ($\hat{p}_a$) (i.e. if detections are auditory the probability that the bird sings during the count interval), and the probability of detecting ($\hat{p}_d$) a bird given that it is available. There are currently five methods for estimating detection probability for point count data. The methods employ; distance sampling, multiple observers, time of detection, double sampling, and repeated counts. The point transect distance or variable circular plot method (Reynolds et al. 1980, Buckland et al. 1993), and the dependent observer or primary-secondary observer approach (Nichols et al. 2000) only estimate the probability of detection given availability. The time of detection method (Farnsworth et al. 2002) estimates the product of availability and detection given availability but it cannot separate the two components. The double sampling approach requires a "rapid" sample and then a more intensive sub-sample of plots to correct for observability bias (Bart and Earnst 2002). The repeated counts method requires sampling the same plots over a period of time and estimates the product of the probability of being available and the probability of detection given that an individual is available but cannot separate these components (Royle and Nichols 2003).

The repeated count approaches also include the probability that an individual is in the sample area during the survey because individuals will move between successive surveys.

In this paper we focus on multiple observer methods of estimating detection probability from point counts. Nichols et al. (2000) suggested that a completely independent observer approach would provide more modeling flexibility and benefits over the dependent observer approach if independence between observers was possible. Our objectives are to; 1) present the two independent observer method and potential models for estimating detection probability, including the use of detection distance covariates, showing that the models are essentially closed capture-recapture models, including the use of distance covariates, 2) present a more general model using four independent observers showing that multiple observer models are essentially closed capture-recapture models that allow for individual heterogeneity, 3) compare the efficiency of the two independent observer approach to the primary-secondary observer approach of Nichols et al. (2002), 4) present a two independent observer example to demonstrate the procedure, 5) present a four independent observer example to model inherent heterogeneity in bird detection probabilities and demonstrate potential bias in two-observer estimates, and 6) simulate data under a heterogeneous model to illustrate the levels of heterogeneity typically present in data and the effect of heterogeneity on two-observer models and index counts.

**Methods**

Field methods.—The general sampling situation for the multiple observer methods is a point count survey where multiple points are surveyed from an area of interest. Point counts should be done using standardized guidelines that specify the time

of year and time of day to conduct counts, suitable weather conditions, duration of counts, spacing between points, etc. (Ralph et al. 1995b). This is a standard approach to point counts used to maximize detection probabilities and reduce extraneous variability among counts. For example the North American Breeding Bird Survey (BBS) requires a requisite level of observer expertise, uses the same routes and stops each year, specifies suitable weather conditions for counts, and uses a three minute count (Sauer et al. 1997). When areas of interest are large, stratification by similar habitat is necessary to account for differences in detection probabilities associated with habitat (Buckland et al. 1993, Ralph et al. 1995b, Nichols et al. 2000).

The dependent observer method of Nichols et al. (2000) uses two observers, one primary and one secondary, for each survey. The primary observer identifies all birds seen or heard and communicates this to the secondary observer. The secondary observer records birds identified by the primary observer and additional birds missed by the primary observer. The role of the primary and secondary observer must be switched during the survey, preferably so that one observer is primary for half the survey points. For each point, the data for each species are; the number detected by the primary observer, and the number missed by the primary but detected by the secondary observer.

The independent observer method uses essentially the same sampling design except that observers conduct each point count simultaneously but independently of the other observers. At the end of each point count observers combine their data and determine the detection history for each bird identified during the count. For a two-observer count the possible detection histories for each species are the number of birds

seen in common by both observers, the number of birds seen only by the first observer, and the number seen only by the second observer.

For both methods it is necessary to record the approximate direction and distance of all detections and to track the movement of birds during the count. Tracking movement avoids double counting of birds and minimizes matching errors with the independent observer method. Recording the location of each detection is necessary to match birds among observers using the independent observer method. Detection distance estimates are use to determine the effective area sampled during the survey which is necessary for making spatial or temporal comparisons (Ralph et al. 1995b). An alternative to estimating detection distance is the use of fixed radius plots, where only birds within a given radius are recorded.

Primary-secondary observer model.—The model proposed by Nichols et al. (2000) is a modification of the model used by Cook and Jacobson (1979) to correct for visibility bias in aerial surveys. The secondary observer only records detections not made by the primary observer. The additional requirement that observers switch primary and secondary roles creates two sets of data, which are equivalent to a generalization of a removal study (Zippin 1858, Seber 1982) with two groups. If we let $x_{ij}$ be the number of individuals counted by observer i (i = 1, 2) for points when observer j (j = 1, 2) is primary, then the probability that a bird in the sample area is detected by at least one-observer is,

$$\hat{p}_d = 1 - \frac{x_{12} x_{21}}{x_{22} x_{11}} . \qquad (2)$$

Note that this method, like all multiple observer methods, is estimating the probability of detection given availability of the animal. Abundance of the available portion of the population is then estimated using equation 1.

The assumptions for this method are:

1. The probability of detection by a particular observer for a given species is the same for all individuals of that species, regardless of whether the observer is primary or secondary.

2. The population within the effective search radius is closed during the count.

3. Birds are identified correctly and not double counted.

4. All detections made by the primary observer are independent of the secondary observer.

The first assumption will be violated if there is individual heterogeneity in detection probability because those missed by the primary observer would likely have a lower detection probability and also less likely to be detected by the secondary observer.

Two independent observer models.— If the same sampling technique is used but the observer's detections are independent, survey data are in the form of a Lincoln-Petersen closed-population capture-recapture model (Otis et al. 1978, Seber 1982).

$p_{d1}$ – probability of detection by the first observer

$p_{d2}$ – probability of detection by the second observer

$x_{11}$ – number detected by both observers

$x_{10}$ – number detected by the first observer only

$x_{01}$ – number detected by the second observer only

$n_1$ – total number detected by the first observer ($x_{10} + x_{11}$)

$n_2$ – total number detected by the second observer ($x_{01} + x_{11}$)

Using this notation the probability of detection by each observer is estimated by,

$$\hat{p}_{d1} = \frac{x_{11}}{n_2}$$

and (3)

$$\hat{p}_{d2} = \frac{x_{11}}{n_1} \quad ,$$

and the probability of detection by at least one-observer is,

$$\hat{p}_d = 1 - \left(1 - \hat{p}_{d1}\right)\left(1 - \hat{p}_{d2}\right) . \tag{4}$$

The estimate of the probability of detection by at least one-observer is then used with the observed count in equation 1 to estimate population size.

The assumptions for the independent observer models are:

1. Independence of observations among observers.

2. If a fixed radius plot is used, then counts within the fixed radius circle are accurate.

3. There are no matching errors among the observers so that assignments of detection histories $x_{11}$, $x_{10}$, and $x_{01}$ are accurate.

4. Detection probability for each species at all points is constant for each observer.

5. There is no undetected movement into or out of the fixed radius plot.

Relevant capture-recapture models for the two-observer case are model $M_o$ (equal detection probability between observers) and model $M_{obs}$ (unequal detection probability between observers; $M_t$ of Otis et al. 1978)  See Otis et al. (1978) and White et al. (1982)

for a description of these capture-recapture models.  Survey data can then be analyzed

using these models available in program CAPTURE (White et al. 1982) or program

MARK (White and Burnham 1999), which has the benefit of using information theoretic

model selection procedures.  Additional models can also be used that incorporate

individual bird covariates, such as radial detection distance from observers, can be

developed using a generalized Horvitz-Thompson estimator of population size (Huggins

1989, 1991, Alho 1990).  Using covariates in the models accounts for observable

heterogeneity in the detection probability of individual birds (Pollock 2002).

Modeling detection distance and other covariates requires conditioning the

probability of detecting a bird on availability (as before), and on the bird's detection

distance from the observer.  The probability of detection given availability for an

individual i, by observer j, can be represented as a function of detection distance as,

$$\log_e \left( p_{dji} \right) = \alpha_j + \beta_j \delta \left( r_i \right), \qquad (5)$$

where $\alpha_j$ is the intercept, $\beta_j$ is the slope, and $\delta(r_i)$ is a function of the detection distance

(such as detection distance squared).  The detection distance function allows for four

additional models:

Model $M_0^d$ :  equal intercept and slope terms between observers.

Model $M_0^{*d}$ :  equal intercept between observers but different slope.

Model $M_{obs}^d$ :  unequal intercept between observers but similar slope.

Model $M_{obs}^{*d}$ : unequal intercept and slope between observers.

When covariates are included in the model it is necessary to use the generalized

Horvitz-Thompson (Horvitz and Thompson 1952) estimator of population size (Huggins,

1989, 1991, Alho 1990) instead of equation 1:

$$\hat{N} = \sum_{i=1}^{n} \frac{1}{\hat{p}_i} \; ; \qquad\qquad (6)$$

where n is the number of birds detected and $p_i$ is the detection probability of an individual

bird. Program MARK provides the Horvitz-Thompson estimate of population size as a

'derived parameter,' when using the 'Huggins closed captures' data type.


Four or more independent observer models.—When four or more sampling

periods (in our case observers) are used in closed capture-recapture experiments, there

are conceptually eight models available for analysis (Otis et al. 1978, Pollock et al. 1990,

Williams et al. 2002). Only four of these models are reasonable models for independent

observer point count data:

$M_0$      Equal capture probability.

$M_{obs}$      Observer variation in capture probability.

$M_h$      Individual capture heterogeneity.

$M_{obs,h}$  Observer variation and individual capture heterogeneity.

Assumption four of the two independent-observer method is no longer necessary because

individual heterogeneity can be modeled with data from four or more observers. Using 3

observers may provide more precise estimates than two observers but it does not provide

the data necessary to fit the heterogeneity models.

Behavioral response models are probably not relevant to analysis of independent-observer point count data. A behavioral response is a response by an individual to capture (in our case, detection) that makes them either more or less likely to be captured after first capture. Because observations on point counts are done simultaneously and independently we assume that detections by one observer do not affect detections by the other observers. This reduces the number of capture-recapture models that are relevant for analysis of point count data to four, excluding covariate models.

The models available for analysis are based on assumptions about the sources of variability in the data. Model $M_0$ has the most restrictive assumptions by requiring that the probability of detection is the same for all individuals in the population and that there are no differences between observers in ability to detect individuals. Model $M_{obs}$ is less restrictive in that it allows for differences between observers but still requires that all individuals in the population have equal detection probabilities for a given observer.

Probably the most important models in this group are those that incorporate individual heterogeneity in capture probabilities. Individual heterogeneity indicates that each individual in the population has a unique capture (detection) probability. All of the other models identify the source of variation (temporal, behavioral) and model this process. Accounting for individual heterogeneity is important because ignoring it will cause a negative bias in population estimates caused by a positive bias in capture estimates (White et al. 1982, Johnson et al. 1986, Williams et al. 2002). Because it is not possible to account for unobservable or uncontrollable sources of variation that affect count data (Burnham 1981, Johnson 1995), models that incorporate these unobservable sources are potentially very useful.

Because it is possible to identify and model some sources of individual heterogeneity it is important to classify sources of heterogeneity as either observable or unobservable.  Observable heterogeneity includes differences due to factors like sex or age that can be accounted for in a statistical model by stratification (Johnson et al. 1986) or other factors which can be accounted for with covariates.  Detection distance is a covariate that could cause observable heterogeneity.  Incorporating observable individual heterogeneity into the independent observer models using covariates with four or more observers is identical to that described previously for the two-observer situation using the generalized Horvitz-Thompson estimator.

Unobservable heterogeneity has been attributed to age, social status, innate levels of activity, physical condition, and genetic variation (ie. covariates that are unknown for individual birds) (White et al. 1982, Johnson et al. 1986).  Heterogeneity in point counts results from individual differences in age, social status, and singing rates, or from site specific differences such as landscape structure, vegetative cover and background noise that affect the detection of auditory or visual cues.  Model $M_h$ represents heterogeneity in the detection probabilities of individual birds but no observer differences.  Model $M_{obs,h}$ represents both observer differences and heterogeneity in the detection probabilities of individual birds.

Three estimators are available to estimate abundance in the presence of individual heterogeneity; the Jackknife estimator for model $M_h$, Chao's estimator for model $M_h$ and model $M_{obs,h}$, and the finite mixture estimators for models $M_h$ and $M_{obs,h}$.  The Jackknife estimator is based on linear functions of the capture frequencies (Burnham and Overton 1978, 1979) and is not a likelihood based approach.  The Chao estimators are based on

sample coverage (Chao et al. 1992, Chao and lee 1992) and also are not likelihood based.

Both the Jackknife estimator and the Chao estimators can be run using program

CAPTURE but not in program MARK. Program CAPTURE uses a multivariate

discriminate function procedure to select the appropriate model for the data (Otis et al.

1978), but covariate models cannot be included in the suite of models. An alternative

likelihood based approach to the heterogeneity models are finite mixture models of

heterogeneity (Norris and Pollock 1996, Pledger 2000), which can be parameterized in

program MARK and are likelihood based approaches. By employing Akaike's

Information Criterion (AIC) model selection techniques (Burnham and Anderson 2002)

these approaches are applicable to an entire suite of models including the observable

heterogeneity and covariate models.

**Analyses**

Comparison of dependent and independent observer models.—Seber (1982)

compared the two-sample removal method with the Lincoln-Petersen method (two-

observer Model $M_t$) and demonstrated the greater efficiency (smaller variance) of the

Lincoln-Petersen method, especially for low capture probabilities. We performed a

similar comparison between the dependent and independent observer methods using

simulations over a range of detection probabilities (0.6, 0.7, 0.8, and 0.9) and true

population sizes of 20, 50, 100, and 200. For each detection probability and population

size, 1,000 data sets were generated and analyzed with both methods. The standard error

(SE) was then calculated from the 1,000 estimated population sizes and the ratio of the

SE of the dependent observer method to the independent observer method was

determined.  The same set of simulations was also run for the scenario when the assumption of independence was violated.  This was done by allowing 90% or 80% of the independent observer data to be independent but assuming for the remaining data  that if observer one detected a bird it was always detected by observer two.  This assumption violation caused a bias in the population estimate so we used mean-squared error (MSE) (MSE = Variance + Bias$^2$) instead of the standard error for comparison purposes.  Note that for the dependent observer method there was no bias in the simulated data and thus MSE is equivalent to the variance.  These simulations were used to determine the ratio of MSE$^{1/2}$ of the dependent observer method to the independent observer method as a measure of efficiency.


Field data.—Examples are provided for the independent-observer methods presented in this paper using data collected in Great Smoky Mountains National Park during June of 1999 (Simons unpublished data).  Counts were conducted at 70 points along low use hiking trails.  All observers had been conducting point counts on the study site for the previous six weeks during which their identification and distance estimation skills were periodically validated.   Before each count, observers estimated a 50-m radius circle by spotting landmarks using a laser range finder and began the count immediately thereafter.  Observers conducted variable circular plot 3 minute point counts (Reynolds et al. 1980) between dawn and 10:15 am and only in good weather (no rain or excessive wind) consistent with the recommendations for point count methodology detailed by Ralph et al. (1995b).  During each 3 minute count, observers recorded the number of breeding pairs of each bird species seen or heard.  Observers were separated and

instructed to not look at the other observers during the count.  At each point observers

recorded bird detections in all directions on an unlimited radius plot.  Points were spaced

a minimum of 250 m apart and the location and movement of all individual birds detected

were mapped in order to avoid double counting.  Following each count observers

compared their data sheets to determine the total number of birds detected and which

birds were seen in common.

We used the first two observers from the full four-observer data set to construct a

two-observer data set.  Two-observer analyses were also done for the other two observers

to confirm consistency but these will not be presented.  For illustrative purposes we will

present analyses for three species; Ovenbird (*Seiurus aurocpillus*), Tufted Titmouse

(*Parus bicolor*), and Scarlet Tanager (*Piranga olivacea*), and two species complexes

(Warblers and Vireos see appendix for list of species) using both two-observer and four-

observer methods.

Two independent observer examples.—The five two-observer data sets were

analyzed using program MARK (White and Burnham 1999) with the 'Huggins closed

captures' data type.  The *a priori* set of candidate models is:

1.  equal detectability between observers, Model $M_0$

2.   unequal detectability between observers, Model $M_{obs}$

3.  equal detectability between observers with distance function having the same

    slope for both observers, Model $M_0^d$

4.  equal detectability between observers at the point (equal intercept) with

    distance function having a different slope for each observer, Model $M_0^{*d}$

5.  unequal detectability between observers with distance function having the same slope for both observers, Model $M_{obs}^{d}$

6.  unequal intercept between observers with distance function having a different slope for each observer, Model $M_{obs}^{*d}$

The most parsimonious models were selected using second-order Aikaike's Information Criterion ($AIC_c$), an information-theoretic approach with an adjustment for small sample size (Burnham and Anderson 2002). Data were truncated following the recommendation of Buckland et al. (1993) by discarding 10% of the largest detection distances for each species.

Four independent observer examples.—Because this is a new approach and using four observers at a point is not typical, we present this separately from the two-observer example. The four-observer data set also allows for heterogeneity models of point count data, which have not previously been investigated. We start by examining models without heterogeneity and ones that model heterogeneity with detection distance covariates. Comparisons to the two-observer cases and examination of expected capture histories from this analysis provide added detail about the presence of heterogeneity.

Using the six models discussed previously for the two-observer examples, we estimate detection probabilities and population sizes using the four-observer data for the same three species and two species complexes. These models either do not allow for heterogeneity or model it only as a function of detection distance. Four-observer models were parameterized and run using program MARK with model selection based on $AIC_c$. For a given species we then compared the population estimates from two-observer and

four-observer data sets.  Differences were interpreted as reflecting heterogeneity in the data.

Examining the differences between the observed capture history and the predicted capture history from a model based on estimates of assumed homogeneous parameters is another method of detecting unobservable heterogeneity (Johnson et al. 1986).  For example, one could determine if the observed number of $X_{1100}$ records was similar to that predicted based on the selected model.  The model used to generate predicted capture histories was model $M_0$ for all data sets.  We determined the expected value of the count for the fifteen possible capture histories from all five data sets using the total number of observations and the estimated probability of detection.  We then compared the observed counts for the fifteen possible capture histories using a Chi-square goodness of fit test with fourteen degrees of freedom.  If differences between observed and predicted counts were evident we then looked at the standardized residuals to determine which capture histories were different.

Because heterogeneity was evident, we ran the Jackknife and Chao version of the capture-recapture heterogeneity model (Model $M_h$) using program CAPTURE (these estimators cannot be obtained in program MARK) and we ran the finite mixture heterogeneity models in program MARK.  In all cases the three methods gave similar estimates.  The estimates from the two-point mixture models are reported because model selection is based on AIC criteria from the entire suite of models including those with detection distance covariates.  For consistency we denote the model describing heterogeneity in detection probability as model $M_h$ and the model with observer differences and heterogeneity as model $M_{obs,h}$.

Heterogeneity simulations.—We hypothesized that heterogeneity might be caused by differences in calling/singing rates so we ran simulations to determine the effect of various levels of heterogeneity on the other available models. We predicted that birds with loud and frequent songs would have higher detection probabilities than birds with quiet or infrequent songs. To examine this, we simulated heterogeneous data with a three-point mixture model for populations of 100 and 200 birds. We used 3 levels of heterogeneity, low, moderate, and high, to give some understanding of the effect of heterogeneity on the other candidate models. All simulations had 20% of the population with probability of detection 1.0 and 60% of the population with detection probability 0.75. The remaining 20% of the population had detection probabilities of 0.5, 0.3, or 0.1 simulating low, moderate, or high heterogeneity, respectively. Using these parameterizations 1,000 four-observer data sets are generated and analyzed with model $M_0$ and model $M_h$. An additional 1,000 two-observer data sets are generated with this parameterization and analyzed with model $M_0$. A single observer count is also obtained for each simulation of the two-observer data set to represent an index count.

Additional simulations were run using a two-point mixture model for a population of 200 birds and parameterized so that the expected capture histories approximated the observer capture history for the "Warbler" group. The parameterizations that achieved this were 100 birds with detection probability of 0.9 and 100 birds with detection probability either 0.1 or 0.2. Again 1,000 simulations were run and analyzed for the four-observer, two-observer, and one-observer counts. This was done to examine the

performance of the non-heterogeneity models when heterogeneity is equivalent to that evident in actual field data.

## Results

Comparison of dependent and independent observer models.—Our simulations showed that, when model assumptions were met, the independent observer method was always more efficient than the dependent observer approach (Table 1). Detection probability could not be estimated with the dependent observer method when detection probabilities were low, and/or population sizes were small. Differences between the methods decreased with increasing detection probability because when detection probabilities are high almost all birds are detected by both methods. When the assumption of independence was violated the efficiency of a particular method was a function of the level of dependence, the detection probability, and the true population size.

Two independent observer examples.—The total number of birds detected for the two-observer data sets (after 10% truncation of observations with the largest detection distances) ranged from 31 (Titmouse) to 132 (Warbler) for the two-observer data sets. $AIC_c$ scores selected either model $M_0$ or model $M^d$ from all data sets (Table 2). When model $M_0$ was selected, model $M^d$ was always a reasonable alternative model based on differences in $AIC_c$ ($\Delta AIC_c < 2$). In contrast, when model $M^d$ was selected, model $M_0$ was not necessarily a reasonable alternative model (see for example the Ovenbird, Table 2). Models incorporating observer differences were never selected as the most

parsimonious model, although for some data sets these models may provide reasonable alternatives because $\Delta AIC_c$ is small.

Model $M_0$ detection probabilities for individual observers ranged from 0.77 (Tanager and Vireo) to 0.9 (Ovenbird) (Table 3). There were no differences in population estimates for a given data set between model $M_0$ and model $M^d$, but the standard errors were different.

Four independent observer examples.—Excluding the heterogeneity models, model $M_0^d$ was selected as the most parsimonious model of the four-observer data sets for all analyses except for the Tanager data set, for which model $M_0$ was selected. Model $M_0^d$ was a reasonable alternative to model $M_0$ for this case. All models not incorporating individual heterogeneity gave similar population estimates (Table 4). Species detection probabilities based on model $M_0$ (Table 5) were consistently lower for the four-observer analysis than for the two-observer analysis (Table 3). Estimated population size was > 10% higher for the four-observer data sets (Table 5) than for the corresponding two-observer data sets (Table 3). It is also worth noting that the raw count for the four-observer data (Table 5) was > the two-observer abundance estimates (Table 3), indicating that individual heterogeneity is present in the data for all species.

Observed capture histories were significantly different than expected under a null hypothesis of homogeneity in detection probabilities for all five data sets ($\chi^2 \geq 34.8$, $p \leq 0.005$, df=14). Examining the standardized residuals revealed that the number of observations was greater than expected when only one observer detected a bird or when

all four observers detected a bird, again, indicating that individual heterogeneity is present in the data.

Model $M_h$ (Jackknife estimator) was selected as the most parsimonious model for all five data sets using program CAPTURE. A two-point mixture model for model $M_h$ was selected as the most parsimonious model for all data sets except the Ovenbird data set using program MARK (Table 4).

The estimate for the finite mixture for all five species was close to 0.5, indicating that approximately half the population was in each detection group (Table 5). One group of the finite mixture was always highly detectable (probability of detection > 0.90), while the other group was generally much harder to detect (probability of detection $\leq$ 0.30) with the exception of the Ovenbird (probability of detection = 0.51). Population estimates (Table 5) from model $M_h$ (two-point mixture) ranged from 15% (Titmouse) to 26% (Warbler) higher than the estimates from the selected four-observer model without heterogeneity. The selected model for the Ovenbird using distance to model observable heterogeneity gave the same population estimate as the two-point mixture model, which was similar to estimates from models that did not account for heterogeneity.

The difference in model selection for the Ovenbird reflects the lack of covariate models in program CAPTURE. Model selection from CAPTURE ranked model $M_h$ first and model $M_0$ as the next best. If we ignore the distance covariate models we see that this ordering of models is the same using the AIC criterion in program MARK (Table 4). It is encouraging that the two different methods of model selection and the three types of heterogeneity estimators are in close agreement for these data.

With the exception of the Ovenbird, comparison of abundance estimates from the heterogeneity model to those obtained from models without heterogeneity demonstrates the large ($\geq 15\%$) negative bias in the estimates when heterogeneity is ignored. The negative bias associated with the abundance estimate for the Ovenbird was only 1%. It appears that the heterogeneity in Ovenbird detectability can be accounted for by using distance covariates. Heterogeneity in detection probabilities for Ovenbirds had minimal effects on abundance estimates, which is expected for this species because of its high singing rates and loud vocalization. The detectability of Ovenbirds based on model $M_0$ was 0.82, while detectabilities of all other birds based on this same model were $\leq 0.72$.

Heterogeneity simulations.—These simulations suggest that the estimation accuracy of the available models is highest for the four-observer model $M_h$, followed by the four-observer model $M_0$, the two-observer model $M_0$, and the one-observer count statistic (Table 6). True population size had no effect on the relative accuracy of various models. Both four-observer models were reasonably accurate (within 5% of the true value), except under the most extreme heterogeneity when both models substantially underestimated the true population. The two-observer model $M_0$ was reasonably accurate (within 4% of the true value) for the least heterogeneous data, but it did not perform well as heterogeneity became more extreme. The one-observer count statistic underestimated the true population by $\geq 25\%$ for all levels of heterogeneity.

Individual heterogeneity with a two-point mixture distribution with extremely different probabilities of detection between the two groups demonstrates greater differences in the accuracy of the potential models (Table 6). For these simulations only model $M_h$ performed well: all other estimators underestimated the true value by $\geq 20\%$.

The estimates given by each estimator are similar to the estimates given by the corresponding estimator for the Warbler data, which indicates how severe the heterogeneity in detection probabilities may be in the actual data.

**Discussion**

We have presented an independent observer point count method as an alternative to the primary-secondary observer approach of Nichols et al. (2000) and shown that the method provides reasonable estimates of bird abundance. This is not surprising given the similarities between the two methods and the additional information provided by additional observers. We have also demonstrated that when model assumptions are met, the method provides relatively more efficient estimates. The method also allows for greater model flexibility and complexity. Examples include the incorporation of individual covariates and models of individual heterogeneity when at least four observers are used. Theoretically, primary-secondary observer approaches could also account for individual heterogeneity. The method would require four observers with the added complexity that the third observer would record birds that the first two observers missed, and the fourth observer would record birds that the first three observers missed. While possible we feel that this approach would not be logistically feasible to accomplish in the field because the level of communication required between observers would become unmanageable.

Incorporating detection distance into the independent observer approach has the added benefit that a detection probability of one is not required at detection distance zero. This is a restrictive assumption with the distance sampling method (Buckland et al. 1993) and may not be reasonable for all surveys, especially when birds occur high in a forest

canopy. It is important to consider other characteristics that may affect the detection process, such as singing rate (Wilson and Bart 1985, McShea and Rappole 1997). McShea and Rappole (1997) found that singing rates were lower for birds closer to an observer. If singing rate does affect the probability of detection, then the distance sampling assumption of a monotonic decline in detection probability would also be violated.

Field application.—Nichols et al. (2000) thoroughly critiqued their primary-secondary observer approach and discussed potential problems with implementing their method in the field. Given the similarities between our methods we will attempt to recapitulate their main points.

The approach assumes that the population within the area being surveyed is closed to movement into or out of the survey area during the count, and that birds are not double counted. These issues are common to all point counts and they can be thought of as a function of count duration. Double counting arises when the same bird is observed at more than one location and is counted as two or more individuals. Reducing the count duration limits the probability of birds moving into, out of, or within the survey area during the count. Unfortunately reducing count duration reduces the total number of detections on the count. On our study sites, detections are generally auditory and thus, if birds have a low singing rate, short duration counts can significantly limit the proportion of birds that are available to count. Optimal count durations should be long enough to ensure that all birds are available for detection (sing at least once during the count) and short enough to minimize movement.

The independent observer approach requires that observations are independent among observers. The primary-secondary approach of Nichols et al. (2000) required only that the observations by the primary observer be independent of detections made by the secondary observer. Many of the issues involved with independence between observers are similar between the 2 methods. Independence may be violated if an observer obtains cues from other observers. This could occur if one-observer is writing down an observation (Nichols et al. 2000), estimating a distance, or moving in a manner that would draw the attention of other observers. Nichols et al. (2000) suggest that violation of the independence assumption is most likely when there are few birds at a point or when most observations are visual. In other words, it is harder to obtain cues from other observers when detections are auditory. Our data are based almost exclusively (>95%) on auditory detections.

Nichols et al. (2000) viewed differences between observers in their ability to detect birds at different distances as the most serious source of error in their method. We incorporate distance as covariates in the candidate models to account for this problem. Note that the primary-secondary method can also incorporate detection distance covariates, so this does not indicate that one method is better than the other. Because these covariate models, which use all birds detected, are available, we do not recommend using fixed radius plots. Modeling distance as a covariate using all detections provides more information, and presumably better estimates than ignoring detections outside a fixed radius. We do recommend 10% truncation of the farthest detections to remove outliers, as recommended by Buckland et al. (1993). We must emphasize though that

collecting distance data is essential to point count methods because even if it is not used to estimate detection probabilities, it does provide an estimate of the sample area.

A final source of error results from the process of matching observer's observations. After each point count the observers must agree on which birds were detected in common, and which birds were unique to an observer. This method requires observers to determine which birds were seen by each observer. If surveys are conducted during the breeding season when territories are fixed and birds are not in large flocks, then matching errors should be minimal. If many birds of one species are present at a point then it may be difficult to determine which birds are seen in common.

Both visual and auditory observations occur during point counts, and their detection probabilities are usually not equal, especially if covariates such as distance are used. Detection probabilities are also likely to vary among sex and age classes (Ralph et al. 1995b). For example, during the breeding season auditory detections of adult territorial males are likely to be much greater than auditory detections of females. Determining the best point count study design will require careful consideration the research objectives, and how species specific behavior and habitat characteristics might affect detection probabilities.

Our examples were derived from data collected during the height of the breeding season in Great Smokey Mountains National Park which is primarily composed of mature deciduous forest habitats. Observations were primarily auditory detections of singing males. Foraging females and non-territorial males were observed on rare occasions. In this context we recommend using only auditory detections of singing males for estimating breeding bird abundance. Including visual observations or auditory detections

of call notes that may include female and juvenile birds would overestimate breeding bird abundance.

Example analyses.—Model selection for the two-observer data sets suggests some important characteristics of these data. First, we notice that models that incorporated differences in detection probability between observers were not selected. This implies that detection probability based on auditory cues was similar between two highly trained observers. Models incorporating differences between observers were reasonable alternative models for several species. This suggests that these models should be included as candidates in most studies. These models may prove useful for species that are hard to detect or on surveys where observer's abilities vary more widely. Incorporation of detection distance as a covariate also provided important improvements to the model.

Model selection for the four-observer data sets generally selected the same models that were selected for the two-observer data sets, excluding the heterogeneity models. Again, estimates of detection probability and abundance were relatively precise but accuracy was doubtful given the differences between the two-observer and four-observer results. The two-observer abundance estimates were low and the detection probability estimates were high compared to the four-observer estimates, which was consistent with the expected biases associated with individual animal heterogeneity. Estimates based solely on a two-observer approach may appear reasonable (as in our examples) but caution is advised because no assessment of potential bias can be made with this method.

Heterogeneity models were generally selected as the "best" models. Potential causes of heterogeneity in detection probabilities for point count surveys are; violation of the independence assumption, combining types of detection (e.g. auditory and visual), estimating detection probability for a group of species, and individual variation in singing rates. Violation of the independence assumption caused by observer's cueing off of each other would affect the observed capture histories. This violation would result in more capture histories where the majority of observers detected an individual bird and fewer histories where only one-observer detected an individual bird than expected, under models of equal detection probability. This would result in a positive bias in detection probabilities and an underestimation of abundance. Comparisons of observed to expected capture histories, indicate that violations of the independence assumption are not the main factor associated with heterogeneity in these data.

Combining multiple types of detections is a potential source of individual heterogeneity that should also be considered carefully when studies are designed. For example, the detection probabilities of songs may be very different from the detection probabilities of calls or visual observations. Combined detection types may generate results similar to the simulations presented for mixture models where singing birds had high detection probabilities, calling birds had intermediate detection probabilities, and visual observations had low detection probabilities.

Variability in the singing rates of individual birds is another potential source of heterogeneity in detection probabilities (Wilson and Bart 1985, McShea and Rappole 1997). A bird that sings frequently is more likely to be detected than one that sings only once during a survey period. Differences in singing rates may be associated with

differences in pairing status and/or nest stage. For example, unpaired males may sing more frequently than paired males and paired males that have not nested are likely to sing at higher rates than a paired male that is incubating or caring for a brood (Wasserman 1977, Krebs et al. 1981, Lein 1981, Wilson and Bart 1985). This presents a situation where the use of finite mixture models is biologically reasonable because individual heterogeneity can be associated with biological groups based on an individual's reproductive stage.

So far we have described sources of heterogeneity and suggested possible ways to account for it, but there are other unaccountable sources of heterogeneity. These sources could arise from something as simple as whether the bird is facing toward or away from the observer, or more complex interactions between sound attenuation of the auditory cue and the surrounding landscape. Differences in song characteristics, vegetative cover/density, topographic features, and background noise all affect the intensity and attenuation of auditory cues (Richards 1981). Presence of conspecific territorial males, competitors, predators, mating status, and nesting status could all affect the singing behavior of individual birds. Stratification and careful survey design can provide partial control of some of these sources of heterogeneity, but they will always be present in survey data.

## Recommendations

We believe that the independent-observer approach should be strongly considered among the suite of statistical methods available for point-count studies. The benefits of this approach include; the number of candidate models available for analysis, the ease of

implementation with existing software, the ability to run both general and covariate models within program MARK, and the application of information theoretic approaches to model selection. In particular, the incorporation of detection distance, a relatively well studied methodology, as a covariate within the independent-observer framework may provide more precise abundance estimates. We are encouraged that our three different heterogeneity models gave similar results, suggesting that it may be useful to model unobservable individual heterogeneity. Because the finite-mixture models are likelihood based models and can be compared to other models using AIC model selection, we recommend their application when biologically meaningful phenomena that influence detection probability, such as changes in singing rates during the breeding season, are known.

We recommend conducting pilot studies involving multiple-observers collecting multiple types of data prior to any implementation of a large scale survey. Such a study will allow an assessment of the various sources of individual heterogeneity and selection of the most efficient methodology. Pilot study results can provide guidance for the allocation of single-observer distance sampling, double-observer sampling, and/or more than three-observer sampling based on factors determined to be affecting the detection process.

Finally, we feel that models combining the independent-observer and the time of detection approach (Farnsworth et al. 2002) would be a useful extension of this work. Combined models would allow separation of the probability of availability from the probability of detection given availability. For species with high singing rates, this may not provide much useful information because availability is often near one. For less

vocal species this additional component of the detection process may have significant effects on abundance estimates.

**Literature Cited**

Alho, J.M. 1990. Logistic regression in capture-recapture models. Biometrics 46:623-635.

Barker, R.J. and J.R. Sauer. 1995. Statistical aspects of point count sampling. Pages 125-130 in Monitoring Bird Populations by Point Counts (C.J. Ralph, J.R. Sauer, and S. Droege, Eds.). U.S. Department of Agriculture, Forest Service General Technical Report PSW-GTR-149.

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

Buckland, S.T., D.R. Anderson, K.P. Burnham, and J.L. Laake. 1993. Distance Sampling: Estimating Abundance of Biological Populations. Chapman and Hall, London.

Burnham, K.P. 1981. Summarizing remarks: Environmental influences. Studies in Avian Biology 6:324-325.

Burnham, K.P. and D.R. Anderson. 2002. Model selection and multimodel inference: A practical information-theoretic approach. Springer-Verlag, New York, NY.

Burnham. K.P. and W.S. Overton. 1978. Estimation of the size of a closed population when capture probabilities vary among animals. Biometrika 65:625-633.

Burnham. K.P. and W.S. Overton. 1979. Robust estimation of population size when capture probabilities vary among animals. Ecology 62:927-936.

Chao, A. and S.M. Lee. 1992. Estimating the number of classes via sample coverage. Journal of the American Statistical Association 87:210-217.

Chao, A., S.M. Lee, and S.L. Jeng. 1992. Estimation of population size for capture-recapture data when capture probabilities vary by time and individual animal. Biometrics 48:201-216.

Conroy, M.J. 1996. Abundance indices. Pages 179-192 in Measuring and monitoring biological diversity. Standard methods for mammals (D.E. Wilson, F.R. Cole, J.D. Nichols, R. Rudran and M. Foster, eds.). Smithsonian Institution Press, Washington, D.C.

Cook, R.D. and J.O. Jacobson. 1979. A design for estimating visibility bias in aerial surveys. Biometrics 35:735-742.

Farnsworth, G.L., K.H. Pollock, J.D. Nichols, T.R. Simons, J.E. Hines, and J.R. Sauer. 2002. A removal model for estimating detection probabilities from point-count surveys. Auk 119:414-425.

Horvitz, D.G. and D.J. Thompson. 1952. A generalization of sampling without replacement from a finite universe. Journal of the American Statistical Association 47:663-685.

Huggins, R.M. 1989. On the statistical analysis of capture experiments. Biometrika 76:133-140.

Huggins, R.M. 1991. Some practical aspects of a conditional likelihood approach to capture experiments. Biometrics 47:725-732.

Johnson, D.H., K.P. Burnham, and J.D. Nichols. 1986. The role of heterogeneity in animal population dynamics. Proceedings of the International Biometrics Conference, 13:1-15.

Johnson, D.H. 1995. Point counts of birds: What are we estimating? Pages 117-123 in Monitoring Bird Populations by Point Counts (C.J. Ralph, J.R. Sauer, and S. Droege, Eds.). U.S. Department of Agriculture, Forest Service General Technical Report PSW-GTR-149.

Krebs, J.R., M. Avery, and R.J. Cowie. 1981. Effect of removal of mate on the singing behavior of Great Tits. Animal Behavior 29:635-637.

Lancia, R.A., J.D. Nichols, and K.H. Pollock. 1994. Estimating the number of animals in wildlife populations. Pages 215-253 in Research and management techniques for wildlife and habitats (T. Bookhout, Ed.). The Wildlife Society, Bethesda, Maryland.

Lein, M.R. 1981. Display behavior of ovenbirds (*Seiurus aurocapillus*) II. Song variation and singing behavior. Wilson Bulletin 93:21-41.

Marsh, H. and D.F. Sinclair. 1989. Correcting for visibility bias in strip transect aerial surveys of aquatic fauna. Journal of Wildlife Management 53:1017-1024.

McShea, W.J. and J.H. Rappole. 1997. Variable song rates in three species of passerines and implications for estimating bird populations. Journal of Field Ornithology. 68:367-375.

Nichols, J.D., J.E. Hines, J.R. Sauer, F.W. Fallon, J.E. Fallon, and P.J. Heglund. 2000. A double-observer approach for estimating detection probability and abundance from point counts. Auk 117:393-408.

Norris, J.L., and K.H. Pollock. 1996. Nonparametric MLE under two closed capture-recapture models with heterogeneity. Biometrics 52:639-649.

Otis, D.L., K.P. Burnham, G.C. White, and D.R. Anderson. 1978. Statistical inference from capture data on closed animal populations. Wildlife Monographs, No. 62.

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

Pollock, K.H. 2002. The use of auxiliary variables in capture-recapture modeling: An overview. Journal of Applied Statistics 29:85-102.

Pollock, K.H., J.D. Nichols, C. Brownie, and J.E. Hines. 1990. Statistical inference for capture-recapture experiments. Wildlife Monograph, No. 107.

Ralph, J.C. 1981. An investigation of the effect of seasonal activity levels on avian censusing. Studies in Avian Biology 6:265-270.

Ralph, J.C., J.R. Sauer, S. Droege (Eds.). 1995a. Monitoring bird populations by point counts. United States Forest Service General Technical Report PSW-GTR-149.

Ralph, J.C., S. Droege, and J.R. Sauer. 1995b. Managing and monitoring birds using point counts: Standards and applications. Pages 161-168 in Monitoring bird populations by point counts (J.C. Ralph, J.R. Sauer, and S. Droege (Eds.)). United States Forest Service General Technical Report PSW-GTR-149.

Reynolds, R.T., J.M. Scott, and R.A. Nussbaum. 1980. A variable circular-plot method for estimating bird numbers. Condor 82:309-313.

Richards, D.G. 1981. Environmental acoustics and censuses of singing birds. Studies in Avian Biology 6:297-300.

Robbins, C.S. 1981. Effect of time of day on bird activity. Studies in Avian Biology 6:275-286.

Robbins, C.S., D. Bystrak, and P.H. Geissler. 1986. The Breeding Bird Sruvey: Its first fifteen years, 1965-1979. U.S. Fish and Wildlife Service Resource Pulicatin No. 157.

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

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

Sauer, J.R., J.E. Hines, G. Gough, I. Thomas, and B.G. Peterjohn. 1997. The North American Breeding Bird Survey results and analysis. Version 96.4. Patuxent Wildlife Research Center, Laurel, MD.

Sauer, J.R., J.E. Hines, and J. Fallon. 2003. The North American Breeding Bird Survey, Results and Analysis 1966-2002. Version 2003.1, USGS Patuxent Wildlife Research Center, Laurel, MD.

Sauer, J.R., B.G. Peterjohn, and W.A. Link. 1994. Observer differences in the North American Breeding Bird Survey. Auk 111:50-62.

Seber, G.A.F. 1982. The Estimation of Animal Abundance and Related Parameters (2nd ed.). Edward Arnold, London.

Skirvin, A.A. 1981. Effect of time of day and time of season on the number of observations and density estimates of breeding birds. Studies in Avian Biology 6:271-274.

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

Wasserman, F.E. 1977. Mate attraction function of song in the White-throated Sparrow. Condor 79:125-127.

White, G.C., D.R. Anderson, K.P. Burnham, and D.L. Otis. 1982. Capture-recapture removal methods for sampling closed populations. Los Alamos National Laboratory Publication LA-8787-NERP. Los Alamos, NM.

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

Williams, B.K., J.D. Nichols, and M.J. Conroy. 2002. Analysis and Management of Animal Populations. Academic Press. San Diego, CA.
Wilson, D.M., and J. Bart. 1985. Reliability of singing bird surveys: Effects of song phenology during the breeding season. Condor 87:69-73.

Wilson, D.M. and J. Bart. 1985. Reliability of singing bird surveys: Effects of song phenology during the breeding season. Condor 87:69-73.

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

Table 1:  Comparison of the dependent-observer approach to the independent-observer approach using simulations of 1,000 data sets for each population size, detectability and method.  For the unbiased scenario, the ratio of the SE of the dependent-observer method to the SE of the independent-observer method is presented.  The biased scenario represents 10% and 20% of the observations in the independent-observer data being dependent on the first observer and compares the ratio of the SE of the dependent-observer method to the $MSE^{1/2}$ for the independent observer method.  NE represents cases where the dependent observer estimate was frequently not estimable.

|  |  | Detection Probability | | | |
|  | True N | 0.6 | 0.7 | 0.8 | 0.9 |
| --- | --- | --- | --- | --- | --- |
|  | 20 | NE | NE | 3.21 | 1.43 |
|  | 50 | 4.89 | 2.51 | 1.37 | 1.26 |
| Unbiased | 100 | 1.92 | 1.56 | 1.21 | 1.10 |
|  | 200 | 1.63 | 1.42 | 1.28 | 1.14 |
|  | 20 | NE | NE | 2.38 | 0.74 |
| 10% | 50 | 3.58 | 1.77 | 0.90 | 0.68 |
| Bias | 100 | 1.43 | 1.03 | 0.80 | 0.63 |
|  | 200 | 1.10 | 0.87 | 0.69 | 0.54 |
|  | 20 | NE | NE | 1.93 | 0.63 |
| 20% | 50 | 2.98 | 1.34 | 0.71 | 0.51 |
| Bias | 100 | 1.02 | 0.75 | 0.54 | 0.43 |
|  | 200 | 0.76 | 0.56 | 0.45 | 0.36 |

Table 2: Model selection for the two-independent observer examples giving the $\Delta AIC_c$ values for all 6 candidate models. The smaller $\Delta AIC_c$ values indicate a more parsimonious model with 0 indicating the selected model. $AIC_c$ weights in parentheses.

| Date Set | Models | | | | | |
|---|---|---|---|---|---|---|
| | $M_0$ | $M_{obs}$ | $M_0^d$ | $M_0^{*d}$ | $M_{obs}^d$ | $M_{obs}^{*d}$ |
| Ovenbird | 5.93 (0.03) | 7.91 (0.01) | **0 (0.65)** | 3.26 (0.13) | 3.25 (0.13) | 5.35 (0.05) |
| Tanager | **0 (0.40)** | 2.03 (0.15) | 1.06 (0.24) | 2.76 (0.10) | 3.15 (0.08) | 4.80 (0.04) |
| Titmouse | **0 (0.33)** | 2.14 (0.11) | 0.10 (0.31) | 2.31 (0.10) | 2.31 (0.10) | 4.60 (0.03) |
| Vireo | **0 (0.33)** | 1.77 (0.14) | 0.63 (0.24) | 1.63 (0.15) | 2.42 (0.01) | 3.60 (0.05) |
| Warbler | **0 (0.31)** | 1.31 (0.16) | 0.53 (0.24) | 2.17 (0.11) | 1.86 (0.12) | 3.26 (0.06) |

Table 3: Abundance estimates (N) for the two-independent observer examples. Birds detected are the totals between the two observers. Model $M_0$ was selected as the most parsimonious for all data sets except the Ovenbird.

| Data set | Birds Detected | Model $M_0$ Detectability | $\hat{N}$ Model $M_0$ | $\hat{N}$ Model $M_0^d$ |
|---|---|---|---|---|
| Ovenbird | 72 | 0.90 (0.027) | 73 (0.94) | 73 (1.43) |
| Tanager | 45 | 0.77 (0.055) | 48 (2.09) | 48 (2.28) |
| Titmouse | 31 | 0.85 (0.052) | 32 (0.98) | 32 (1.22) |
| Vireo | 89 | 0.77 (0.039) | 94 (2.86) | 94 (3.03) |
| Warbler | 132 | 0.85 (0.026) | 135 (2.10) | 135 (2.19) |

Table 4: Model selection for the four-independent observer examples giving the $\Delta AIC_c$
values for all 8 candidate models. Model $M_{2h}$ and model $M_{obs,2h}$ are based on 2 point
mixture models of heterogeneity. The smaller $\Delta AIC_c$ values indicate a more
parsimonious model with 0 indicating the selected model. $AIC_c$ weights are in
parentheses.

| | Models | | | | | | | |
|---|---|---|---|---|---|---|---|---|
| Data Set | $M_0$ | $M_{obs}$ | $M_0^d$ | $M_0^{*d}$ | $M_{obs}^d$ | $M_{obs}^{*d}$ | $M_{2h}$ | $M_{obs,2h}$ |
| Ovenbird | 48.5 | 53.5 | **0.0** | 8.4 | 7.4 | 12.5 | 10.3 | 37.7 |
| | (0.0) | (0.0) | **(0.95)** | (0.01) | (0.02) | (0.0) | (0.01) | (0.0) |
| Tanager | 23.4 | 29.4 | 24.1 | 29.0 | 30.1 | 35.2 | **0.0** | 22.8 |
| | (0.0) | (0.0) | (0.0) | (0.0) | (0.0) | (0.0) | **(1.0)** | (0.0) |
| Titmouse | 28.9 | 34.1 | 26.0 | 30.3 | 31.2 | 35.8 | **0.0** | 6.2 |
| | (0.0) | (0.0) | (0.0) | (0.0) | (0.0) | (0.0) | **(0.96)** | (0.04) |
| Vireo | 71.2 | 77.2 | 60.2 | 65.6 | 66.3 | 71.7 | **0.0** | 36.4 |
| | (0.0) | (0.0) | (0.0) | (0.0) | (0.0) | (0.0) | **(1.0)** | (0.0) |
| Warbler | 99.0 | 102.4 | 83.9 | 88.5 | 87.3 | 92.1 | **0.0** | 4.2 |
| | (0.0) | (0.0) | (0.0) | (0.0) | (0.0) | (0.0) | **(0.89)** | (0.11) |

Table 5: Abundance estimates (N) for the four-independent observer examples. Birds detected are the totals among the four observers. Detection probabilities are given by p for model $M_0$ and $p^{group1}$ and $p^{group2}$ for model $M_{2h}$. The proportion of the population in group 1 is given by pr(group1). Standard errors are in parentheses.

| Data Set | Birds Detected | Model $M_0$ | | Model $M_{2h}$ | | | |
|---|---|---|---|---|---|---|---|
| | | p | N | pr(group1) | $p^{group1}$ | $p^{group2}$ | N |
| Ovenbird* | 81 | 0.82 | 81 | 0.41 | 0.51 | 1.0 | 82 |
| | | (0.022) | (0.305) | (0.059) | (0.060) | ($\approx 0.0$) | (0.889) |
| Tanager | 55 | 0.66 | 56 | 0.54 | 0.30 | 0.92 | 63 |
| | | (0.033) | (0.894) | (0.086) | (0.108) | (0.047) | (6.32) |
| Titmouse | 39 | 0.70 | 39 | 0.53 | 0.28 | 0.97 | 45 |
| | | (0.038) | (0.592) | (0.090) | (0.111) | (0.032) | (5.56) |
| Vireo | 111 | 0.67 | 112 | 0.55 | 0.25 | 0.93 | 134 |
| | | (0.023) | (1.243) | (0.055) | (0.067) | (0.026) | (12.00) |
| Warbler | 154 | 0.72 | 155 | 0.47 | 0.18 | 0.91 | 196 |
| | | (0.019) | (1.042) | (0.061) | (0.062) | (0.020) | (22.23) |

*Selected model was model $M_0^d$ and the estimated population size was 82 (SE = 0.889).

Table 6: Abundance estimates (averages over 1000 simulations) for four-observer and two-observer methods and a single observer count from simulated heterogeneous data from three- and two-point mixture distributions. For the three-point mixture distribution 20% of the population had detection probability 1, 60% had detection probability 0.75, and the remaining 20% had three different levels (low = 0.5, moderate = 0.3, and high = 0.1) of detection probabilities. For the two-point mixture distribution half the population had high detection probability (p = 0.9) and the other half low detection probability (p = 0.1 or p = 0.2), which gave capture histories similar to those observed in the "Warbler" data set.

| True Population | Level of Heterogeneity | 4 Observer Estimate | | 2 Observer Estimate | 1 Observer Count |
|---|---|---|---|---|---|
| | | $M_h$ | $M_0$ | $M_0$ | N |
| **3-point MIXTURE** | | | | | |
| 100 | Low | 100 (1.89) | 98 (1.26) | 96 (3.37) | 75 (4.04) |
| 100 | Moderate | 99 (3.52) | 95 (2.00) | 91 (3.49) | 71 (3.89) |
| 100 | High | 90 (3.99) | 87 (2.21) | 83 (3.11) | 67 (3.63) |
| 200 | Low | 200 (2.29) | 197 (1.59) | 192 (4.65) | 150 (5.61) |
| 200 | Moderate | 197 (4.51) | 190 (2.83) | 182 (5.16) | 142 (5.64) |
| 200 | High | 179 (5.20) | 174 (3.05) | 167 (4.43) | 134 (5.14) |
| **2-point MIXTURE** | | | | | |
| 200 | Highest (p = 0.1) | 188 (35.1) | 135 (4.82) | 122 (5.12) | 100 (4.32) |
| 200 | High (p = 0.2) | 205 (19.7) | 160 (5.20) | 142 (6.00) | 110 (4.74) |

Appendix 1:  Common and scientific names for the "Warbler" and "Vireo" species

groups used for analysis.

**Vireo Group**
   Red-eyed Vireo             *Vireo olivaceus*
   Solitary Vireo               *Vireo solitarius*

**Warbler Group**
   Black-throated Blue Warbler   *Dendroica caerulescens*
   Black-and-white Warbler   *Mniotilta varia*
   Black-throated Green Warbler   *Dendroica virens*
   Hooded Warbler   *Wilsonia citrina*
   Worm-eating Warbler   *Helmitheros vermivorus*
   Yellow-throated Warbler   *Dendroica dominica*
   Chestnut-sided Warbler   *Dendroica pensylvanica*
   Canada Warbler   *Wilsonia canadensis*

**Chapter 3**

**TIME OF DETECTION METHOD FOR ESTIMATING ABUNDANCE
FROM POINT COUNT SURVEYS**

# TIME OF DETECTION METHOD FOR ESTIMATING ABUNDANCE FROM POINT COUNT SURVEYS

Mathew W. Alldredge, Biomathematics and Zoology, North Carolina State University, Raleigh NC 27695.
Kenneth H. Pollock, Zoology, Biomathematics, and Statistics, North Carolina State University, Raleigh, NC 27695.
Theodore R. Simons, USGS Cooperative Fish and Wildlife Research Unit,  Dept. of Zoology, North Carolina State University, Raleigh, NC 27695.
Jaime Collazo, USGS Cooperative Fish and Wildlife Research Unit,  Dept. of Zoology, North Carolina State University, Raleigh, NC 27695.

Abstract.—Point count surveys are often used to collect data on the abundance and distribution of birds.  These data are generally used as an index or a measure of relative abundance.  Valid spatial or temporal comparisons of these measures require assuming a constant detection process.  Eliminating this restrictive assumption requires information about detection probabilities that allow for estimation of the detection process.  The time of removal approach for estimating detection probabilities is generalized to the time of detection approach, which uses the complete detection history of an individual and has fewer assumptions.  We apply this model to point count surveys where detections are aural and singing rates of birds are important components of the detection process.  Our model accounts for both availability bias and detection bias by modeling the combined probability that a bird sings during the count and that it is detected given that it sings.  The model requires dividing the count into several intervals and recording detections of individual birds in each interval.  We develop maximum likelihood estimators for three different forms of the detection process and provide a full suite of models based on capture-recapture models, including covariate models.  We present two examples of this method; one for four species of songbirds surveyed in Great Smokey Mountains National Park using three unequal time intervals, and one for the Pearly-eyed Thrasher

(*Margarops fuscatus*) surveyed in Puerto Rico using four equal time intervals.  Models incorporating individual heterogeneity were selected for all data sets using information-theoretic model selection techniques.  Detection probabilities varied among time intervals of the count, indicating a possible behavioral response to the observer.  We recommend applying this method to surveys with four or more equal intervals so that fewer assumptions are necessary and analyses can take full advantage of standard capture-recapture software.

## Introduction

Point count surveys are routinely used to determine animal abundance, particularly breeding birds (Thompson 2002), but statistically valid methods for collecting and analyzing such data are rarely used (Rosenstock et al. 2002).  Most studies do not correct for the detection process, relying instead on the raw count data as an index to abundance (Ralph et al. 1995, Nichols et al. 2000).  In general point count surveys involve counting all individuals seen or heard at a set of points under standardized conditions.  For example, the North American Breeding Bird Survey (BBS), standardizes the duration of counts, length of survey routes, distance between points, weather conditions, time of year, etc. (Robbins et al. 1986, Sauer et al. 1997), but it does not estimate detection probabilities directly.  Current point count methods that account for the detection process are; the distance sampling or variable circular plot method (Reynolds et al. 1980, Buckland et al. 1993), multiple-observer methods (Nichols et al. 2000, chapter 2), the time of detection method (Farnsworth et al. 2002), double sampling methods (Bart and Earnst 2002) and repeated counts (Royle and Nichols 2003).  All of

these methods are based on the general form for estimating abundance of species i as

(Lancia et al. 1994, Williams et al. 2002)

$$\hat{N}_i = \frac{C_i}{\hat{p}_i} \qquad\qquad (1)$$

where $C_i$ is the observed count and $\hat{p}_i$ is the estimated probability of detection for species

i.

The probability of detection $\hat{p}_i$ is actually the product of two probabilities: the

probability that an individual is available for detection $\hat{p}_{ai}$ and the probability that an

individual is detected given that it is available $\hat{p}_{di}$. Ignoring these probabilities gives rise

to availability bias and perception bias (Marsh and Sinclair 1989). When point count

surveys are conducted in dense habitat, where most detections are auditory, availability is

the probability that a bird occurring within the survey area sings during the count. Both

distance sampling and multiple-observer methods ignore the availability process,

assuming that all birds are available for detection during the count interval.

Farnsworth et al. (2002) presented a method based on the time of first detection

for estimating detection probabilities that estimated the total detection probability (ie. the

product of the availability probability and the probability of detection given availability).

This method is based on dividing a count into multiple intervals and recording the

interval in which an individual is first detected. In collecting data on the interval of first

detection data are in the form of a typical removal experiment (Seber 1982) where

animals are removed from the population upon capture, or in this case detection.

Removal experiments presume that the number of animals caught on subsequent

occasions will decline as a linear function of the population size, and they use the decline

to estimate initial population size. The removal method for point counts is similar in concept presuming that the number of new detections declines with time, and using the decline to estimate population size. The removal method is generalized to allow for individual heterogeneity of detection probabilities.

Model assumptions for this method are (Farnsworth et al. 2002):

1. There is no change in the population of birds within the detection radius during the point count (closed population).

2. There is no double-counting of individuals.

3. All members of one group are detected in the first interval (applies to the heterogeneity model for three time periods).

4. All members of the other group that have not yet been detected have a constant per minute probability of being detected.

5. If counts with limited-radius are used, observers accurately assign birds to within or beyond the radius used.

Assumption 2 requires observers to keep track of individual birds (usually by mapping) during the count to avoid double counting. This requires observers to decide whether a singing birds are new or whether they have been detected previously.

Capture-recapture methods (Otis et al. 1978, Williams et al. 2002) provide a flexible and efficient framework for analyzing these data. A capture-recapture modeling approach uses both first detections and all subsequent detections of an individual to estimate the probability of detection. Seber (1982) demonstrated that capture-recapture models are more efficient (smaller variance) than removal models. However, if

subsequent detections are not tracked accurately, errors in the capture history will bias the abundance estimates.

Our objective is to present a time of detection approach using a capture-recapture framework based on all detections of an individual as a more general alternative to the removal approach presented by Farnsworth et al. (2002). We first discuss field methods required to collect data suitable for this method. We then discuss the detection process and three ways of modeling it. A discussion of the candidate models and the general form of the time of detection model is then presented. Following this development of the general model, we present a finite-mixture model for individual heterogeneity and a covariate model for observable heterogeneity. Methods of analysis, including available software, are also discussed. We then illustrate our methods with an example for three unequal interval point counts and another for four equal interval point counts.

**Methods**

Field methods.—The sampling situation for the models we are proposing is the same as that described by Farnsworth et al. (2002). Multiple locations are selected for point counts within the area of interest, and at each point a count is conducted for a specified amount of time. Field conditions in which point counts are conducted should follow a standardized approach in order to maximize detection probabilities and minimize other sources of variability (Ralph et al. 1995, Sauer et al. 1997). Depending on the diversity of habitats and size of the area of interest it may also be necessary to stratify by habitat type (Nichols et al. 2002).

The method presented by Farnsworth et al. (2002) was developed to estimate detection probabilities for a single observer when a point count is divided into three or

more intervals of variable length. This was done to present a model that would work with existing types of point count data. A common approach, based on the recommendation of Ralph et al. (1995), records birds detected in the first 3 minutes of a count, the next 2 minutes and the final 5 minutes. The approach was recommended to permit comparisons among studies using single 3, 5, or 10 minute counting periods.

The full detection history data for a point count survey will consist of the number of birds observed with each of the $2^n-1$ possible detection histories. For example, if three time periods are used, then seven detection histories are possible ($x_w$: $x_{111}$, $x_{110}$, $x_{101}$, $x_{011}$, $x_{100}$, $x_{010}$, and $x_{001}$). Mapping detections using multicolored pens (Simons and Shriner 2000), where each color represents a time interval, is an effective method for tracking the time interval of detections.

Defining the sample area of a survey is necessary for making both temporal and spatial comparisons. One approach involves collecting count data from limited radius plots, so that the actual sample area is known (Nichols et al. 2000). An alternative approach is to use unlimited radius plots and estimate the detection distance for each individual observed during the count. This detection distance information can then be used to estimate the effective sample area, similar to the approach used in distance sampling (Buckland et al. 1993). Detection distance information can also be used as a covariate in models to account for differences in detection probabilities among individuals.

Detection process.—Estimating the detection process is key to obtaining reliable estimates of abundance (Nichols et al. 2000, Farnsworth et al. 2002, Thompson 2002). The detection process consists of both the probability that an individual is available for

detection and the probability that it is detected given that it is available for detection (Marsh and Sinclair 1989). Our method estimates the product of these two components. Most birds in forested environments are detected by their songs or calls. In this situation the availability process depends on the singing rates of individual birds and is thus a function of the duration of each interval used in the point count. We will present the availability process as the probability that a bird gives an auditory cue during the count, but the method could be applicable to other situations affecting availability.

One approach to model the detection process is by assuming that singing rates of a bird follow a Poisson process and that the probability of detecting all songs of a species is equal. Under these assumptions the instantaneous detection rate ($\varphi_i$) or the "Poisson detectability coefficient", is used to model the probability that a bird sings during a specified time period. The probability of detecting an individual in time interval i of length $t_i$ using an instantaneous rates formulation (Seber 1982, pg. 3 and 296) is then,

$$p_i = 1 - e^{-\varphi_i t_i} \qquad (2)$$

This formulation is consistent with that typically used for removal experiments where $\varphi_i$ corresponds to the "Poisson catchability coefficient" and $t_i$ corresponds to the effort on the $i^{th}$ occasion (Otis et al. 1978, Seber 1982 pg 296). Farnsworth et al. (2002) used an alternative formulation, the discrete rate formulation, which assumes a constant per minute detection rate ($\gamma_i$) so that the probability of detecting an individual in time interval i of length $t_i$ was given by

$$p_i = \gamma_i^{t_i} \qquad (3)$$

The interval length $t_i$ that occurs in both equations 2 and 3 has important consequences for the construction of candidate models and data analysis. Clearly if the

instantaneous detection rate or per minute detection rate differs between time intervals then the probability of detecting an individual during a given interval will be different for each interval. Note that there is insufficient information in the data to model this difference when using the removal approach. Equations 2 and 3 are particularly important when the instantaneous or discrete detection rates are constant for the duration of the count. If the time intervals are different then the probability of detecting an individual during an interval will be different for each interval even when the detection rates are constant. Using equations 2 or 3 can allow for unequal time intervals while requiring only a single parameter for constant detection rate models. In contrast, if the probability of detection is estimated for each interval directly, then a separate parameter is required for the detection probability of each interval. When point counts consist of equal intervals it is not necessary to use equations 2 or 3 because interval detection probabilities are lo longer a function of the interval length and interval detection probabilities are estimated directly.

Approach and candidate models.—A more general approach to the removal models presented by Farnsworth et al. (2002) uses a capture-recapture framework which tracks all time intervals in which individual birds are detected. The full time of detection approach provides a complete set of models, of which the removal model is a special case, for estimating detection probabilities. There are 8 general capture-recapture models (Otis et al. 1978, White et al. 1982, Pollock et al. 1990, Borchers et al. 2002), all of which may be applicable to the time of detection approach. Here we present each capture-recapture model and a description relevant to the time of detection method. We begin by assuming that all intervals are of equal length.

Model $M_0$ represents a constant detectability coefficient for all individuals across all time periods. This model is the simplest of all the models and requires estimation of only two parameters, the detection coefficient and N. In general this model may oversimplify the detection process because it assumes no individual heterogeneity and no differences between the first detection and subsequent detections. This model is generally not robust to heterogeneity or differences between first and subsequent detection probabilities. This model may be applicable when a species is very easily detected, for example, has very high rates of calling and very loud distinct calls. Such situations reduce the affects of individual heterogeneity in detection probabilities caused by variable singing rates because the probability of detecting any given call is high regardless of whether or not the bird was detected previously.

Model $M_t$ assumes equal detection coefficients for all individuals but different coefficients among time periods. For t periods the number of parameters in this model is t+1. The model does not account for individual heterogeneity in detection probabilities or differences between first and subsequent detections. For example, this model is applicable to situations where observer effects on singing rates (McShea and Rappole 1997) decline over time.

Model $M_b$ assumes equal probability of first detection for all individuals across all periods and a unique probability of subsequent detections that is equal for all individuals across all subsequent periods. This model requires the estimation of the detection coefficient for the first detection and one for all subsequent detections (plus one parameter for N), for a total of 3 parameters in the model. The model is applicable to situations in which the detection coefficient is constant over the duration of the count and

individual heterogeneity does not affect the detection process but where the requirement to track individuals detected during the count does affect the probability of subsequent detections. We assume if observers are properly tracking individuals that detection coefficients would be higher for subsequent detections than for first detections. However, if observers have a tendency to ignore individuals previously detected then the opposite would hold. It is important to note that under this model subsequent detections do not provide any information for estimating the detection coefficient of first detections or N. Model $M_b$ is equivalent to the removal method (Otis et al. 1978) because only first detections provide information about N. Model $M_b$ is equivalent to Farnsworth et al.'s (2002) model with no heterogeneity.

Model $M_{tb}$ assumes an equal detection coefficient of first detection for all individuals that differs between periods and a unique detection coefficient of subsequent detection that is equal for all individuals but differs between subsequent time periods. The full model contains more parameters than can actually be estimated so the simplifying assumption is made that there is only a single effect of first capture. This model estimates a detection coefficient of first capture for each time period and a single coefficient for the change in detection probability for subsequent detections. The probability of subsequent detections for a given time period is an additive effect to the probability of first detection for all time periods. Again, only first detections provide information about N. This is equivalent to the removal model, except that the removal model requires setting 2 of the parameters equal so that N is identifiable. The model is applicable in situations similar to those given for Model $M_b$ but where the observer also

has an effect on detections that diminishes over time. This model is over-parameterized and hard to fit to real data.

Model $M_h$ assumes a unique detection coefficient for each individual which remains constant across all time periods and is similar between first and all subsequent detections. The number of parameters involved in all of the heterogeneity models depends on how heterogeneity is modeled. This is discussed in the heterogeneity section of this paper. In general this model is applicable when the detectability coefficient is different among all members or groups of the population, but remains constant across time periods, and is not affected by first detection. For example, if some members of the population are unmated territorial males, some are mated territorial males, some are mated and incubating territorial males, and some territorial males are actually caring for a brood, we would expect singing rates and other behaviors to differ among groups. These differences would result in different detectability coefficients among groups. Because it is generally not possible to identify the status of individual birds differences in detection probabilities are evaluated using mixture models (Norris and Pollock 1996, Pledger 2000).

Model $M_{th}$ is similar to model $M_h$ in that there is a unique detection coefficient for each individual or group of individuals but this model also assumes that detection probabilities change between time periods. Applicable situations include those discussed for model $M_h$ but where observer effects on singing behavior diminish over time.

Model $M_{bh}$ assumes a unique detection coefficient of first capture for each individual that remains constant across time periods and a unique coefficient of subsequent detection that remains constant across time periods. Again, only first captures

provide information about N so this model is equivalent to the removal approach with heterogeneity. Farnsworth et al. (2002) present a 3 period removal model with a two point mixture model of heterogeneity. Because there were only 3 periods, a very strong assumption that all of the birds in one group are detected in the first interval is required. The full two point mixture models of Norris and Pollock (1996) and Pledger (2000) are available when four or more time periods are used. We believe that these models will be very useful because individual heterogeneity and situations where subsequent detections are affected by first detections are likely on avian point counts.

Model $M_{tbh}$ is the most general model. It assumes a unique probability of first capture for each individual that differs among time periods and a unique probability of subsequent detection for each individual that differs among time periods. Fitting this model requires several strong assumptions to reduce the number of parameters estimated. We did not attempt to fit this model to our data.

The individual heterogeneity models have been the slowest to develop because of their complexity, large number of parameters, and computing requirements. In fact only recently have likelihood based methods been developed using finite mixtures (Norris and Pollock 1996, Pledger 2000). The Jackknife (Burnham and Overton 1978, 1979) and Chao (Chao et al. 1992) heterogeneity model estimators are not likelihood based (Williams et al. 2002). The advantage to having maximum likelihood based estimators is that models can be evaluated with likelihood ratio tests for evaluating sources of variation and model selection techniques, such as Akaike's Information Criterion (AIC), can be used to select the most parsimonious model (Williams et al. 2002).

The Jackknife estimator for model $M_h$ is the original estimator for this model and probably is the most commonly used (Williams et al. 2002). This estimator is based on a generalized jackknife statistic using a linear function of capture frequencies for bias reduction (Burnham and Overton 1978, 1979). Several tests of this estimator using simulation have demonstrated that this estimator does work reasonably well under a variety of situations (Otis et al. 1978, Pollock et al. 1990, Williams et al. 2002). This estimator can be run using program CAPTURE (White et al. 1982).

Another approach commonly used for estimation under models $M_h$, $M_{th}$, and $M_{bh}$ is the sample coverage approach proposed by Chao et al. (1992), Chao and Lee (1992), and Lee and Chao (1994). This approach utilizes the canonical form given in equation 1 by dividing the observed count by an estimate of the sample coverage. The estimators for sample coverage are constructed using capture frequencies (Chao et al. 1992, Chao and Lee 1992, Lee and Chao 1994). The sample coverage approach works well in some situations but not in others (Williams et al. 2002). There are actually three estimators for sample coverage under model $M_{th}$ and Chao et al. (1992) provide guidelines on situations in which each is applicable. Simulation results for model $M_{bh}$ indicate that the jackknife estimator of Pollock and Otto performs better than the sample coverage estimator (Lee and Chao 1994). Both estimators $M_h$ and $M_{th}$ can be run using program CAPTURE.

The finite mixture heterogeneity estimators of Norris and Pollock (1996) and Pledger (2000) are maximum likelihood estimators, which has the advantage of providing a unified maximum likelihood approach of fitting and comparing the entire suite of models. The idea behind this approach is that the entire population of interest can be divided into a finite number of groups, each with unique probability of capture/detection.

The nonparametric approach of Norris and Pollock (1996) first conditions on N to determine the number of groups in the population and then N is varied to find the overall maximum likelihood. Pledger (2000) does not condition on N but instead simply specifies the number of groups at the outset and compares models with different numbers of specified groups. Both methods often lead to the same results (Pledger 2000). Simulation under the various heterogeneity models has demonstrated that the finite mixture approach provides competitive models to the other approaches (Norris and Pollock 1996, Pledger 2000, Williams et al. 2002). The finite mixture models can be fit using program MARK (White and Burnham 1999), which requires that the number of mixtures be specified.

It is worth noting that simulation studies have demonstrated that abundance estimates will depend on the method used and the number of groups used. Norris and Pollock (1996) provide estimates from their method along with the other commonly used approaches to modeling heterogeneity and clearly demonstrate that the resulting estimates of abundance can depend on the method used. Unfortunately there is no way of assessing which method is correct. We would suggest that there may be some biological basis for using finite mixtures for point count data since detections are likely to be a function of singing rates which are known to vary in relation to the breeding state (mated, incubating, brood rearing, etc.) of an individual bird. However, it is possible that this does not account for all the variation in the data and these assumed groups may not be sufficient to account for all the heterogeneity present in the data.

Other models relevant to estimating the detection process include those that incorporate auxiliary information, such as detection distance, which can be incorporated

as covariates. We develop these models in the "Covariate" section and demonstrate their use with an equal time interval example.

General form of model.—Models of equal time interval data can be fit using standard capture-recapture software (CAPTURE or MARK), but not for unequal time interval data. Because much existing point count data have been collected using unequal time intervals we present a general model and subsequent equations for both the instantaneous rates and discrete rates formulations. This model is also applicable to situations when equal time intervals are used, so we also present the equations for this case, which simply models the interval probabilities of first detections ($p_i$) and subsequent detections ($c_i$) directly.

Models without heterogeneity.—Model $M_{tb}$ is the most general model without heterogeneity. The remaining models without heterogeneity are just constrained forms of model $M_{tb}$. Point counts with i intervals are comprised of $2^i$ capture histories, one of which is not observable because it represents individuals that are never detected. Model $M_{tb}$ requires 2i parameters but the minimal sufficient statistic for the distribution has dimension 2i-1 (Otis et al. 1978, Borchers et al. 2002). Because of this the model is constrained by assuming that the change in detection probability after first capture is constant across all subsequent captures (Borchers et al. 2002). We will also make this assumption so that the probability of a subsequent detection ($c_i$) during period i is,

$$c_i = p_i + \delta \tag{4}$$

for equal intervals,

$$c_i = 1 - e^{-(\varphi_i + v)t_i} \tag{5}$$

for the instantaneous rates formulation, and

$$c_i = \left(\gamma_i + \eta\right)^{t_i} \tag{6}$$

for the discrete rates formulation, where $\delta$, $\nu$, and $\eta$ are the change in the detectability coefficients for all subsequent detections. Note that $c_i$ may change between periods even though $\delta$, $\nu$, or $\eta$ remain constant because either $p_i$, $\varphi_i$ or $\gamma_i$ and $t_i$ can vary among time periods.

We present the equations using the interval detection probabilities ($p_i$'s and $c_i$'s) for clarity in the following equations. These are replaced by one of the appropriate rates equations for unequal interval point counts. A count consisting of 4 intervals is comprised of 15 observable detection histories ($x_w$: $x_{1111}$, $x_{1110}$, ... , $x_{0001}$). The set of possible detection probabilities for the observable capture histories is denoted by $\pi_w$. The expected values of the counts for each detection history of the general model can be written as,

$$E(x_{1111}) = N p_1 c_2 c_3 c_4 = N \pi_{1111},$$

$$E(x_{1110}) = N p_1 c_2 c_3 (1 - c_4) = N \pi_{1110},$$

$$\vdots \tag{7}$$

$$E(x_{1010}) = N p_1 (1 - c_2) c_3 (1 - c_4) = N \pi_{1010},$$

$$\vdots$$

$$E(x_{0001}) = N(1 - p_1)(1 - p_2)(1 - p_3) p_4 = N \pi_{0001}$$

where the $p_i$'s and $c_i$'s are from either the instantaneous or discrete rates formulation.

From this it follows that the probability of detecting an individual at least once during the

entire count ($p_T$) is one minus the probability of never detecting an individual;

$$p_T = 1 - (1 - p_1)(1 - p_2)(1 - p_3)(1 - p_4) \tag{8}$$

The full multinomial likelihood can be written as;

$$L(N, p_i, c_i, | x_{w*}) = \frac{N!}{\prod_w x_w! \left(N - \sum_w \pi_w\right)!} \prod_w \pi_w^{x_w} \left(1 - \sum_w \pi_w\right)^{N - \sum_w x_w} \tag{9}$$

Where $x_w$ is the number of observations for each possible detection history, excluding the

history for the individuals never detected and $\pi_w$ (Equation 7) is the probability of

observing each capture history, excluding the probability of never seeing an individual.

Because N cannot be directly observed we condition on the total number of birds

counted ($x_T$) to make the problem more amenable to numerical methods. The likelihood

is decomposed into a marginal distribution $L_1$ and a conditional distribution $L_2$ so that

detection probabilities can be estimated from the conditional, and abundance can be

estimated from the marginal distributions. The relationship between these likelihoods is

$L = L_1 * L_2$. The marginal likelihood $L_1$ can be written as,

$$L_1(N \mid x_T) = \binom{N}{x_T} (\pi_.)^{x_T} (1 - \pi_.)^{N - x_T} \tag{10}$$

Where $\pi_.$ is the sum of the $\pi_w$. The likelihood $L_2$ is conditional on the observed capture

histories, given by,

$$L_2(p_i, c_i \mid x_w) = \binom{x_T}{x_{1111} \ldots x_{0001}} \prod_w \left(\frac{\pi_w}{\pi_.}\right)^{x_w} \tag{11}$$

From equation 11 the $p_i$'s and c's, $\varphi_i$'s and v's or $\gamma_i$'s and $\eta$'s (depending on which formulation is used) can be estimated by maximizing the likelihood for the observed data. From equation 8 the probability that an individual is detected at least once during the count ($\hat{p}_T$) can be calculated. For models with time variation $\hat{p}_T$ is given by;

$$\hat{p}_T = 1-(1-\hat{p}_1)(1-\hat{p}_2)(1-\hat{p}_3)(1-\hat{p}_4) \tag{12}$$

for equal time intervals, by

$$\hat{p}_T = 1-e^{-\hat{\varphi}_1 t_1}e^{-\hat{\varphi}_2 t_2}e^{-\hat{\varphi}_3 t_3}e^{-\hat{\varphi}_4 t_4} \tag{13}$$

for the instantaneous rates formulation or by

$$\hat{p}_T = 1-(1-\hat{\gamma}_1^{t_1})(1-\hat{\gamma}_2^{t_2})(1-\hat{\gamma}_3^{t_3})(1-\hat{\gamma}_4^{t_4}) \tag{14}$$

for the discrete rate formulation. The variance of $\hat{p}_T$ is obtained for each estimator by reparameterizing to estimate $\hat{p}_T$ directly. This can be done by solving one of the equations of $\hat{p}_T$ for the last interval parameter and using this in the likelihood. For example, if we solve equation 13 for $\hat{\varphi}_4$ we get

$$\hat{\varphi}_4 = \frac{1}{t_4}\ln\left(\frac{1-\hat{p}_t}{e^{\hat{\varphi}_1 t_1}e^{\hat{\varphi}_2 t_2}e^{\hat{\varphi}_3 t_3}}\right) \tag{15}$$

Replacing the final interval parameter with this equation will allow for direct estimation of the probability that an individual is detected at least once and the associated variance. An alternative approach is to use the delta method (Seber 1982) and obtain estimators for

the approximate variances. Abundance is estimated from the general form (equation 1) or the likelihood $L_1$ as,

$$\hat{N} = \frac{x_T}{\hat{p}_T} \tag{16}$$

The observed count ($x_T$) is one realization of a random variable and thus, has a variance associated with it. Assuming that the observed count ($x_T$) is from a binomial distribution an estimate of the variance of abundance is (Nichols et al. 2000, Williams et al. 2002);

$$\widehat{Var}(\hat{N}) = \frac{x_T^2 \widehat{Var}(\hat{p}_T)}{\hat{p}_T^4} + \frac{x_T(1-\hat{p}_T)}{\hat{p}_T^2} \tag{17}$$

With the framework of the model established, it is possible to determine the number of parameters for each potential model. Capture history probabilities or the expected values (equation 7) for four time periods indicate that there are potentially 8 parameters ($p_1$, $p_2$, $p_3$, $p_4$, $c_2$, $c_3$, $c_4$, and N) estimated for the general model. Recall that we must constrain the effect of subsequent capture to be equal across all time periods ($\delta_2 = \delta_3 = \delta_4$). The 4 time period model $M_{tb}$ thus, has 6 parameters. In general model $M_{tb}$ with t time periods has t+2 parameters to estimate. The remaining models are constrained from this form. Surveys with t intervals are shown for the instantaneous rates formulation.

Model $M_0$: Set all rate parameters equal ($\varphi_1 = \varphi_2 = \ldots = \varphi_t$, $v_2 = v_3 = \ldots = v_t = 0$, and N). Total parameters to estimate are 2.

Model $M_t$: Set rate parameters for subsequent detections equal to zero so that subsequent detections have the same probability as initial detections ($\varphi_1$, $\varphi_2$, …, $\varphi_t$, $v_2 = v_3 = \ldots = v_t = 0$ and N). Total parameters to estimate are t+1.

Model $M_b$: Set rate parameters for first detections equal and rate parameters for

subsequent detections equal ($\varphi_1=\varphi_2=\ldots=\varphi_t$, $v_2=v_3=\ldots=v_t$, and N). Total

parameters to estimate are 3.

Analysis of point count data with these models will depend on whether the data

are from equal interval or unequal interval point counts. If data are collected with

unequal interval point counts then one of the rate formulations is required. This is not

possible using one of the standard capture-recapture programs. Analyis of unequal time

interval point counts requires maximizing the likelihood for each candidate model using a

program such as SURVIV (White 1983) that allows the user to input the likelihood for

each model.

Analysis of point count data from equal interval point counts is much easier

because standard capture-recapture analysis programs, such as CAPTURE and MARK,

can be used. Program CAPTURE is an older program that will run the full suite of

models automatically but it does not use current methods of model selection. Program

MARK requires the user to parameterize all candidate models and uses Akaike's

information criterion (AIC), an information-theoretic approach, (Burnham and Anderson

2002) for model selection. Program MARK also accommodates covariate models.


Heterogeneity models.—Many factors create unobservable individual

heterogeneity in detection probabilities. These factors often relate to individual behavior

or where the individuals are located (Burnham 1981, Johnson et al. 1986). Factors

inherent to an individual that may cause individual heterogeneity include age, breeding

status (mated, unmated, incubating, etc.), and singing rate. Other causes of heterogeneity

are those associated with an individual's location because detection probabilities are being estimated from spatially distinct points. While stratification by habitat type can eliminate some differences among points (Buckland et al. 1993, Nichols et al. 2000) differences that cause heterogeneity exist in all data. These differences may include such factors as foliage density, local background noise, topographic features and the presence or absence of predators or competitors. Temporal differences in detection probabilities between points sampled early in the day compared to those sampled later in the day (Farnsworth et al. 2002) or between points done on consecutive days may also exist. Some of these factors can be minimized by standardizing the conditions under which point counts are conducted, but no amount of standardization will account for all of the heterogeneity present in the population (Burnham 1981).

One source of heterogeneity that may be especially important given that these models are based on singing rates, is individual variation in singing rate. Singing rates generally change during the breeding season as birds pair and begin incubation or caring for nestlings (Wasserman 1977, Lein 1981). Asynchronous breeding will cause heterogeneity in the data. If singing rates decline over the course of the breeding season, then sampling over this entire period would introduce heterogeneity into the data. Singing rates are also affected by habitat, local abundance, and the proximity of observers (McShea and Rappole 1997).

In general four or more time intervals are required to parameterize heterogeneity models, unless very strong assumptions are made, as in Farnsworth et al. (2002). We will develop the heterogeneity models based on four or more time intervals. The assumptions

required for the restricted heterogeneity models are given with the presentation of the three interval examples.

For point count surveys that use equal intervals it is possible to use several different heterogeneity models. For unequal time intervals we must use the rates equations and thus we will develop a finite mixture model of heterogeneity for this situation because the likelihood equation must be parameterized and maximized in a program such as SURVIV (White 1983). We use a finite mixture approach because it is a maximum likelihood approach and we believe the benefits of a maximum likelihood approach and information-theoretic model selection warrant the use of this approach. We present the two-point mixture model because as Pledger (2000) suggested, two-point mixtures often provide the most parsimonious models and estimators with good properties. This approach can be easily extended to more mixtures if appropriate. We present this model for four sampling intervals, because it is doubtful that data would be collected with more than four intervals.

Data from four interval point counts are summarized by the counts for the 15 observable detection histories $x_w$. Assuming that only two groups (two mixtures) comprise the population, the probabilities for each capture history are given by the sum of the product of the proportion of animals in each group times the group specific capture and recapture probabilities. For example, the expected value of the count for the 4 interval detection history for individuals detected in all intervals is

$$E(x_{1111}) = N(\lambda p_{11} c_{21} c_{31} c_{41} + (1-\lambda) p_{12} c_{22} c_{32} c_{42}) \qquad (18)$$

where $\lambda$ is the proportion of animals in heterogeneity group one and the proportion of animals in heterogeneity group two is $1-\lambda$ as they must occur in one of the groups and the

$p_{ij}$'s and $c_{ij}$'s are the probabilities of first detection and subsequent detection in the $i^{th}$ interval for individuals in the $j^{th}$ group. There are similar expressions for the other $x_w$. The conditional likelihood $L(p_{ij}, c_{ij}, \lambda \mid x_w)$ is similar to equation 9,

$$L_2(p_i, c_i \mid x_w) = \binom{x_T}{x_{1111}...x_{0001}} \prod_{\forall w} \left( \frac{\pi_{w1} + \pi_{w2}}{\pi_.} \right)^{x_w} \tag{19}$$

where $\pi_.$ is the probability of being detected at least once during the entire count, $\pi_{w1}$ is the probability of being in the first group and having capture history w, and $\pi_{w2}$ is the probability of being in the other group and having capture history w. Using this likelihood, the $p_{ij}$'s and c's, $\varphi_{ij}$'s and v's, or $\gamma_{ij}$'s and $\eta$'s (depending on the rate formulation used) can be estimated by maximizing the likelihood for the observed data.

The probability that an individual is detected at least once during the count ($\hat{p}_T$) is calculated using the estimated detection coefficients,

$$\hat{p}_T = 1 - \left[ \hat{\lambda}(1 - \hat{p}_{11})...(1 - \hat{p}_{t1}) + (1 - \hat{\lambda})(1 - \hat{p}_{12})...(1 - \hat{p}_{t2}) \right] \tag{20}$$

for the equal interval formulation. For one of the other rates formulations (equations 2 or 3) the $p_{ij}$ is replaced by the appropriate rate to estimate $\hat{p}_T$, as in equations 13 and 14. The variance for $\hat{p}_T$ can again be calculated by reparameterizing the model estimators to directly estimated $\hat{p}_T$. Equations 16 and 17 are then used to estimate abundance and the corresponding variance.

Covariates.—Previous examples describe cases in which sources of individual variation were not identifiable. There are situations where sources of variation are identifiable and can be quantified. Accounting for observable heterogeneity in capture

probabilities when individual covariates that explain the differences in capture probabilities are identified, is a special case of model $M_h$ (Huggins 1989, 1991, Alho 1990). Many covariates might explain differences in detection probabilities from point counts where detections are based on auditory cues. These include distance from the observer, singing rate, direction of singing relative to the observer, foliage density, pairing status, and so on. Unfortunately most of these are not quantifiable during a point count. For example, it is not possible to assess the pairing status of a bird during a point count and thus this source of heterogeneity cannot be accounted for with a covariate. Singing rate is another factor that could be used as a covariate, although this should be used as a categorical variable because not all songs are detected during a count.

Using detection distance as a covariate is very appealing because intuitively an observer's ability to detect a bird should decrease with increasing distance to the bird. Because distance sampling on point counts is frequently used to estimate abundance, using distance as a covariate seems even more reasonable. If observers can estimate the location of birds from their songs with reasonable accuracy then models incorporating a detection distance covariate will provide less biased estimates than those that do not. The ability of observers to accurately estimate the distance to singing birds is a key assumption and is one that is largely untested.

Covariates (including detection distance) are used to model detection probability by making detection probability a linear function of individual covariates. An intercept ($\beta_0$) and a slope ($\beta_1$) parameter are estimated to model the detection process. The original suite of models described can be viewed as intercept only models. The effect of the covariate term on the slope parameter is modeled as either a constant effect over time

periods or a variable effect over time periods. This adds two additional models to each of the original seven models described, or a total of 21 conceptual models.

Modeling detection distance as a covariate in model $M_0$ will demonstrate the approach. The original model $M_0$ assumes a constant detection probability over all periods and all detection distances. If we incorporate detection distance into the model with a constant slope across periods, then the model assumes a constant detection function over time periods. Alternatively we can model detection distance with a different slope for each interval, which assumes that the detection probability at the point (intercept) is constant across intervals but the effect of detection distance (slope) on the detection probability is not constant between intervals. This may occur when the observer's arrival at the point affects a bird's singing behavior, but the effect declines over time. The other models can be constructed similarly from the original suite of models by interpreting the original model parameters as reflecting the probability of detection at the point, and interpreting the covariate terms as reflecting how detection probabilities change with respect to covariates.

Covariate models of equal interval point counts can be parameterized in program MARK, using the "Huggins Closed Captures" data type. The ability of MARK to incorporate an entire set of candidate models into a single analysis using an information-theoretic approach to model selection is a compelling reason to collect data using equal time intervals.

Huggins (1989, 1991) and Alho (1990) used the linear-logistic function of individual covariates conditional on the total number of animals detected, to model detection probabilities as a function of the observed covariates. This model is used in

program MARK for modeling equal time interval cases. Under this model, the probability of detecting an individual, j, is given by,

$$\hat{p}_j = \frac{e^{\beta_0 + \beta_1 x_j}}{1 + e^{\beta_0 + \beta_1 x_j}} \tag{21}$$

where $\beta_0$ and $\beta_1$ are estimated parameters, and $x_j$ is the measured covariate value for the $j^{th}$ individual. We assume that movement during the point count is minimal so that $x_j$ is constant for each individual.

If we assume that the probability of detecting the $j^{th}$ individual is constant across time periods then the probability that this individual is detected at least once during a count with t intervals is

$$\hat{p}_j^* = 1 - (1 - \hat{p}_j)^t \tag{22}$$

It is also possible to estimate interval-specific detection probabilities for each individual by modifying equation 16 to have interval specific estimates of $\beta_0$ and $\beta_1$ and using the interval specific estimates to determine the probability of detecting each individual at least once during the count. The abundance estimator for a count with a total of M individuals detected is the Horvitz-Thompson estimator (Horvitz and Thompson 1952)

$$\hat{N} = \sum_{j=1}^{M} \frac{1}{\hat{p}_j^*} \tag{23}$$

Using time intervals of unequal length with covariate models becomes more difficult because existing capture-recapture software cannot be used. In this situation each individual in the sample has a unique likelihood that must be used to estimate the model parameters. This is done by modifying one of the rates formulations (equations 2

or 3) so that the detection probability is a function of the covariate. For example, modifying the instantaneous rates formulation gives the detection probability as,

$$p_{ij} = 1 - e^{-(\varphi_i + \beta\, x_j)t_i} \qquad (24)$$

such that the intercept (probability of detection at the point) is given by $\varphi_i$ and $\beta$ represents the change in the probability of detection with respect to detection distance. This form of the instantaneous rates equation is then used in the likelihood.

### Field Trials

We present examples from two different field studies, one from point counts using 3 unequal intervals, and the other from point counts using 4 equal time intervals. The first example is based on point count data collected in Great Smokey Mountains National Park from 1996 to 2000. We only used data from a single year (1998) and a single observer to avoid temporal and observer effects. Point counts were conducted along pre-established survey routes that followed trails in the park (Simons and Shriner 2000). All point counts were conducted in the first few hours of daylight on days meeting acceptable environmental conditions for point counts. All point counts were done during May and June of the breeding season. Point counts were divided into 3, 2, and 5 minute intervals and the complete detection history was recorded for each individual by using different colored pens for each time interval. Point counts were unlimited radius plots and detection distance was recorded for all detections.

Analysis is restricted to the four most frequently detected species to avoid the effects of sparse data on model selection. These species are the; Black-throated Green Warbler (*Dendroica virens*), Hooded Warbler (*Wilsonia citrine*), Ovenbird (*Seiurus*

*aurocapillus*) and Red-eyed Vireo (*Vireo olivaceus*), all of which have relatively loud calls and high singing rates. All models (except the covariate models) were used and analysis was done using program SURVIV. We truncated 10% of the data by omitting observations with the largest detection distances, as recommended by Buckland et al. (1993). Model selection was based on AIC (Burnham and Anderson 2002).

We modeled heterogeneity using constrained forms of the two-point mixture models. These models must be constrained so that all parameters are identifiable. We constrained the models by fixing the detection probabilities for one of the groups in the mixture. One model was to set all detection probabilities for one group equal to one, similar to Farnsworth et al. (2002). This parameterization assumes that all individuals in this group are detected in every interval. Alternatively, we set all interval detection probabilities equal to 0.9 for one of the groups in the mixture. This value of 0.9 was obtained from analysis of similar data from this same study area. A two-point mixture model of heterogeneity for four-independent observer point count data showed that for one of the mixtures the detection probability for Ovenbird was 1.0 and for Red-eyed Vireo was 0.89 (chapter 2). An analysis of singing rate data for the Ovenbird estimated the availability probability at 0.93 (chapter 5). By modeling one of the heterogeneity groups with a constant, model $M_{th}$ has a slightly different interpretation as only one of the groups will actually have time variation. The necessity of fixing parameters like this for three time interval point count data is the primary reason for using four or more intervals for point counts.

The other example is for the Pearly-eyed Thrasher based on point counts conducted in the karst belt of north-central Puerto Rico in 2003. Data were collected

during the breeding season (mid-February through May). Surveys were conducted using three teams of two experienced and trained observers each. Forty nine point count routes were used with one to 29 (average 9.7) point counts conducted on each route for a total of 477 points. Point count routes were located away from human habitation along "low use" trails. The first point was located 500 m from the start of the trail and subsequent points were located 200 m from the previous. Point locations were alternated from on the trail to 50 m off the trail (randomly left or right).

Point counts were conducted from 0400 hours to 0800 hours on days with suitable weather conditions. Each count was conducted for ten minutes. Counts were divided into four equal time intervals and the complete detection history was recorded for each individual by using different colored pens for each time interval. Point counts were fixed 100 m radius plots and detection distance within this plot was recorded for all detections. During the previous year distances at each plot were flagged at 10 m intervals.

The Pearly-eyed Thrasher data were analyzed using the full suite of models, including heterogeneity and detection distance covariate models. All analyses were done using program Mark. Data were not truncated since a fixed radius plot was used. Model selection was based on AIC (Burnham and Anderson 2002).

## Results

Three interval data set.—Heterogeneity models were the most parsimonious for the three unequal interval data sets from Great Smoky Mountains National Park (Table 1). The analysis clearly demonstrated the importance of heterogeneity models. The $\Delta AIC_c$ weights for all models without heterogeneity were always zero, indicating no evidence supporting these models. Model $M_{th(0.9)}$ was selected as the most parsimonious

for three of the data sets and Model $M_{h(0.9)}$ was selected for the other data set. Heterogeneity models with interval detection probabilities set to 0.9 for one heterogeneity group were selected for over models where a heterogeneity group had interval detection probabilities set to 1. This indicates that the assumption of complete detection is too restrictive for these data sets. All species except the Ovenbird showed evidence for time variation in instantaneous rates of detection, indicating that detection probabilities do not remain constant for the duration of counts. There was little support for the models incorporating both heterogeneity and behavior, indicating that detection probabilities for birds did not change after first detection.

Interval detection probabilities showed a consistent pattern; the shortest intervals had the smallest detection probabilities and the longest intervals had the highest detection probabilities (Table 2). This was not true for the Red-eyed Vireo which had a detection probability of 0.26 for the three minute interval and 0.29 for the two minute interval. This may indicate an observer effect that made this species less detectable during the first interval. Comparing the observed counts to the estimated abundance in the sample area showed differences of 5% (Ovenbird) to 23% (Hooded Warbler).

Four interval data set.—Sixteen of the 21 conceptual models for the four interval Pearly-eyed Thrasher data set gave reasonable results. The other 5 models gave results indicating one or more parameters were not identifiable. The most parsimonious model selected for this data set was model $M_{th}$ ($\Delta AIC_c$ weight = 0.70), with the remaining support for model $M_{th}$(detection distance consant slope) ($\Delta AIC_c$ weight = 0.30) (Table 3). For all models, the general form of the model and it's two covariate formulations were always ordered together in model selection. Heterogeneity models were always "better"

than models not accounting for heterogeneity. Of the models that did not account for unobservable heterogeneity, models incorporating detection distance always had lower $AIC_c$ values.

Under the selected model, 29% of the population had low detection probabilities while the remaining portion had high detection probabilities (Table 4). Low detection probabilities ranged from 0.09 to 0.56 and the high detection probabilities ranged from 0.70 to ≈1.00. Since the detection probability was essentially one for one of the heterogeneity groups during the second interval, abundance is equivalent to the count for within this heterogeneity group. A detection probability of one indicates that all individuals were seen. Note that the standard error is not estimable for parameter estimates near the boundary. The total number of observations for this data set was 520 and the estimated abundance for the sampled area was 547 (SE=8.6), 5% higher than the observed count.

## Discussion

The time of detection method of modeling point count data collected from consecutive time intervals is a less restrictive approach for estimating detection probabilities than the removal method. Both approaches are promising because they model the detection process and provide estimates of true abundance or density, and both permit spatial and temporal comparisons of data without the unrealistic assumptions necessary to compare index counts. These models can be applied to data sets with two or more time intervals of equal or unequal length.

Our approach is a more general approach than the removal method of Farnsworth et al. (2002). Recording the complete detection history of birds during a point count

provides a much larger suite of models that can incorporate more sources of variation, such as time variation, than the removal model. With the exception of the behavior models which are equivalent to the removal models (Seber 1982) this approach is more efficient (smaller variance) than the removal approach.

Differences between species in the estimated detection probabilities were expected and are probably due to differences in singing rates and sound intensity. The overall probability of detecting an individual at least once during a 10 minute count was 0.92 (SE = 0.006) for the Ovenbird, 0.79 (SE = 0.021) for the Black-throated Green Warbler, 0.71 (SE = 0.031) for the Red-eyed Vireo, and 0.65 (SE = 0.037) for the Hooded Warbler. Of these species Ovenbirds have the highest singing rates and thus, we would expect them to have the highest detection probability. Comparing estimates of detection probability to those given by Farnsworth et al. (2002) show very similar estimates for all four of these species.

The time of detection approach also reduces the number of assumptions required for the model. The removal method for three time intervals had five assumptions (Farnsworth et al. 2002), which could be relaxed to four if more than 3 intervals were used to collect data. The time of detection approach has only three assumptions for counts conducted with more than three time intervals, which are generally required for all analysis approaches to point count data. An additional assumption of both methods is that species are identified correctly.

Assumption 1: there is no change in the population within the detection radius during the point count (closed population). Violations of the closure assumptions are more likely to occur for longer duration point counts and for wide ranging species. This

method may not be applicable for wide ranging species, such as woodpeckers (Family *Picidae*) or crows (*Corvus spp.*) where movement during the count may be significant. Violations of the closure assumption are probably less of a problem for many small breeding songbirds, such as the ones illustrated here, as they have relatively small fixed territories during the breeding season (Farnsworth et al. 2002).

Short duration counts will reduce violations of the closure assumption. As the total length of a point count decreases so does the probability that birds will either move into or out of the sample area. This does not imply that all point counts should be arbitrarily short because a point count must be divided into four intervals to fit the full heterogeneity models. If the intervals are too short then interval detection probabilities will be small and the variance on abundance or density estimates will be large. Careful consideration must be given to the appropriate length of point counts for the species surveyed. This may imply different survey protocols for different species groups.

Assumption 2: there is no double-counting of individuals. Double-counting results in abundance or density estimates that are too large. Problems with double-counting are likely to increase as the length of the count increases because undetected movement of birds is more likely to occur. Our method requires observers to track individuals during the count which should minimize violations of this assumption. Proper training of observers to minimize double-counting is also required (Farnsworth et al. 2002).

Assumption 3: for unlimited radius plots, distance is measured accurately or for limited radius plots, individuals are accurately assigned to within or beyond the designated radius. We previously discussed the importance of knowing the sample area

so that comparisons can be made between studies. Observers can be trained to estimate detection distance or to assign birds to within a limited radius plot. The accuracy and precision of observer ability to estimate distance has not been rigorously assessed. We suspect that, even with training, observers tend to over-count individuals within fixed radius plots, and that the accuracy of observers in estimating detection distances of songs and calls may be poor.

The removal method also requires an assumption that the detection rate is constant for the duration of the count (Farnsworth et al. 2002). This assumption is not necessary with our method because there is sufficient information to model time variation from the full detection history. Evidence for a time effect was found in four out of the five data sets presented here. Therefore, it appears that the assumption of constant detection rates under the removal model may not be valid. McShea and Rappole (1997) found singing rates were affected by the presence of an observer. Movement of an individual during the count would also affect the detection process as distance from the observer may affect how likely an observer is to detect a call.

### Recommendations

We developed a general model to analyze point count data collected for two or more equal or unequal intervals. It can be used to analyze data in a variety of formats, including those recommended by Ralph et al. (1995). We recommend that future studies using this approach be designed with four or more equal intervals. The use of equal intervals simplifies analysis and permits the use of standard capture-recapture software. Use of four or more intervals allows for the application of full two-point mixture models.

Farnsworth et al. (2002) recommended combining the removal approach with the distance sampling approach to provide better estimates. We have done this by incorporating detection distance as a covariate in the models presented here. Further development is needed to incorporate alternative forms of the detection function. One obvious extension is to use distance squared, which is equivalent to the half-normal detection function applied in distance sampling. Given the number of factors affecting the aural detections of birds in point count surveys distance may not be a useful explanatory variable in all situations. For example, if singing rates of individual birds is a significant factor affecting the detection of birds then detection probabilities may have an initial increase with increasing detection distance before they decline at large distances. This is because birds near observers may sing less than birds farther from observers (McShea and Rappole 1997). Further investigation of the use of detection distance as an explanatory variable for estimating aural detection probabilities is required.

Point counts based on auditory detections rely on an observer's ability to accurately estimate the detection distance to singing birds, even on fixed radius plots. A rigorous assessment of the accuracy of detection distance estimation is needed before those methods become widely adapted. Such an evaluation should also identify the bias associated with fixed and unlimited radius plots.

For our examples we presented a single species modeling approach. Because point count data usually consists of multiple species we recommend using a multiple species modeling approach (chapter 4). If species with similar detection probabilities are modeled together then more parsimonious models can be used. This is especially useful for rare species or species that are hard to detect.

**Literature Cited**

Alho, J.M.  1990.  Logistic regression in capture-recapture models.  Biometrics 46:623-635.

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

Borchers, D.L., S.T. Buckland, and W. Zucchini.  2002.  Simple mark-recapture.  Pages 104-130 in Estimating animal abundance: Closed populations.  Springer, London.

Buckland, S.T., D.R. Anderson, K.P. Burnham, and J.L. Laake.  1993.  Distance Sampling: Estimating Abundance of Biological Populations.  Chapman and Hall, London.

Burnham, K.P.  1981.  Summarizing remarks: Environmental influences.  Pages 324-325 in Estimating numbers of terrestrial birds (C.J. Ralph and J.M. Scott, Eds.).  Studies in Avian Biology No. 6.

Burnham, K.P. and D.R. Anderson.  2002.  Model selection and multimodel inference: A practical information-theoretic approach.  Springer-Verlag, New York, NY.

Burnham. K.P. and W.S. Overton.  1978.  Estimation of the size of a closed population when capture probabilities vary among animals.  Biometrika. 65:625-633.

Burnham. K.P. and W.S. Overton.  1979.  Robust estimation of population size when capture probabilities vary among animals.  Ecology 62:927-936.

Chao, A. and S.M. Lee.  1992.  Estimating the number of classes via sample coverage.  Journal of the American Statistical Association 87:210-217.

Chao, A., S.M. Lee, and S.L. Jeng.  1992.  Estimation of population size for capture-recapture data when capture probabilities vary by time and individual animal.  Biometrics 48:201-216.

Farnsworth, G.L., K.H. Pollock, J.D. Nichols, T.R. Simons, J.E. Hines, and J.R. Sauer.  2002.  A removal model for estimating detection probabilities from point-count surveys.  Auk 119:414-425.

Horvitz, D.G. and D.J. Thompson.  1952.  A generalization of sampling without replacement from a finite universe.  Journal of the American Statistical Association. 47:663-685.

Huggins, R.M.  1989.  On the statistical analysis of capture experiments.  Biometrika 76:133-140.

Huggins, R.M. 1991. Some practical aspects of a conditional likelihood approach to capture experiments. Biometrics 47:725-732.

Johnson, D.H., K.P. Burnham, and J.D. Nichols. 1986. The role of heterogeneity in animal population dynamics. Proceedings of the International Biometrics Conference, 13:1-15.

Lancia, R.A., J.D. Nichols, and K.H. Pollock. 1994. Estimating the number of animals in wildlife populations. Pages 215-253 in Research and management techniques for wildlife and habitats (T. Bookhout, Ed.). The Wildlife Society, Bethesda, Maryland.

Lee, S.M. and A. Chao. 1994. Estimating population size via sample coverage for closed capture-recapture models. Biometrics 50:88-97.

Lein, M.R. 1981. Display behavior of ovenbirds (*Seiurus aurocapillus*) II. Song variation and singing behavior. Wilson Bulletin. 93:21-41.

Marsh, H. and D.F. Sinclair. 1989. Correcting for visibility bias in strip transect aerial surveys of aquatic fauna. Journal of Wildlife Management. 53:1017-1024.

McShea, W.J. and J.H. Rappole. 1997. Variable song rates in three species of passerines and implications for estimating bird populations. Journal of Field Ornithology. 68:367-375.

Nichols, J.D., J.E. Hines, J.R. Sauer, F.W. Fallon, J.E. Fallon, and P.J. Heglund. 2000. A double-observer approach for estimating detection probability and abundance from point counts. Auk 117:393-408.

Norris, J.L., and K.H. Pollock. 1996. Nonparametric MLE under two closed capture-recapture models with heterogeneity. Biometrics. 52:639-649.

Otis, D.L., K.P. Burnham, G.C. White, and D.R. Anderson. 1978. Statistical inference from capture data on closed animal populations. Wildlife Monographs, No. 62.

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

Pollock, K.H., J.D. Nichols, C. Brownie, and J.E. Hines. 1990. Statistical inference for capture-recapture experiments. Wildlife Monograph, No. 107.

Ralph, J.C., J.R. Sauer, S. Droege (Eds.). 1995. Monitoring bird populations by point counts. United States Forest Service General Technical Report PSW-GTR-149.

Reynolds, R.T., J.M. Scott, and R.A. Nussbaum. 1980. A variable circular-plot method for estimating bird numbers. Condor 82:309-313.

Robins, C.S., D. Bystrak, and P.H. Geissler. 1986. The Breeding Bird Survey: Its first fifteen years, 1965-1979. U.S. Fish and Wildlife Service Resource Publication No. 157.

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

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

Sauer, J.R., J.E. Hines, G. Gough, I. Thomas, and B.G. Peterjohn. 1997. The North American Breeding Bird Survey results and analysis. Version 96.4. Patuxent Wildlife Research Center, Laurel, MD.

Seber, G.A.F. 1982. The Estimation of Animal Abundance and Related Parameters (2$^{nd}$ ed.). Edward Arnold, London.

Simons, T.R. and S.A. Shriner. 2000. Ecology and conservation of neotropical migrants in the southern Appalachians. 1999 Annual report to the National Park Service.

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

Wasserman, F.E. 1977. Mate attraction function of song in the White-throated Sparrow. Condor. 79:125-127.

White, G.C., D.R. Anderson, K.P. Burnham, and D.L. Otis. 1982. Capture-recapture removal methods for sampling closed populations. Los Alamos National Laboratory Publication LA-8787-NERP. Los Alamos, NM.

White, G.C. 1983. Numerical estimation of survival rates from band-recovery and biotelemetry data. Journal of Wildlife Management. 47:716-728.

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

Williams, B.K., J.D. Nichols, and M.J. Conroy. 2002. Analysis and Management of Animal Populations. Academic Press. San Diego, CA.

Table 1: $\Delta AIC_c$ values for the 11 time of detection models fit to each data set. $\Delta AIC_c = 0.0$ for the most parsimonious model for each data set. $\Delta AIC_c$ weights (in parentheses) indicate the strength of the evidence for a given model compared to the other models (the larger number indicates more evidence for that model).

| Species | | | | | | Models | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | $M_0$ | $M_t$ | $M_b$ | $M_{tb}$* | $M_{h(.9)}$ | $M_{h(1)}$ | $M_{bh(.9)}$ | $M_{bh(1)}$ | $M_{th(.9)}$ | $M_{th(1)}$ |
| Black-throated Green Warbler | 45.7 | 42.9 | 47.7 | — | 6.8 | 4.3 | 8.8 | 6.3 | **0.0** | 4.1 |
| | (0.00) | (0.00) | (0.00) | | (0.03) | (0.09) | (0.01) | (0.03) | **(0.75)** | (0.10) |
| Hooded Warbler | 55.4 | 44.3 | 57.5 | — | 15.1 | 13.7 | 17.1 | 15.8 | **0.0** | 3.6 |
| | (0.00) | (0.00) | (0.00) | | (0.00) | (0.00) | (0.00) | (0.00) | **(0.86)** | (0.14) |
| Ovenbird | 56.6 | 40.7 | 58.6 | — | **0.0** | 12.2 | 2.0 | 14.3 | 15.2 | 14.9 |
| | (0.00) | (0.00) | (0.00) | | **(0.73)** | (0.00) | (0.27) | (0.00) | (0.00) | (0.00) |
| Red-eyed Vireo | 78.0 | 58.2 | 77.7 | — | 0.1 | 10.4 | 2.2 | 12.4 | **0.0** | 2.7 |
| | (0.00) | (0.00) | (0.00) | | (0.37) | (0.00) | (0.13) | (0.00) | **(0.39)** | (0.10) |

*Parameter estimates for model $M_{tb}$ were not reasonable. Standard errors were greater than 1.

Table 2. Parameter estimates from the selected model for the 3 interval point count data sets. $\lambda_1$ is the proportion of the population that is in group 1. Detection probability ($p_{ij}$) is the probability of detecting an individual from group j in interval i. The detection probabilities for group 2 ($p_{t2}$) were fixed. Standard errors in parentheses.

| Species | Observed | $\lambda_1$ | $p_{11}$ | $p_{21}$ | $p_{31}$ | $p_{t2}$ | N |
|---|---|---|---|---|---|---|---|
| Black-throated Green Warbler | 377 | 0.53 (0.038) | 0.40 (0.056) | 0.24 (0.056) | 0.54 (0.061) | 0.9 | 425 (15.1) |
| Hooded Warbler | 274 | 0.54 (0.050) | 0.33 (0.062) | 0.20 (0.056) | 0.34 (0.065) | 0.9 | 338 (17.4) |
| Ovenbird | 444 | 0.54 (0.054) | 0.53 (0.042) | 0.39 (0.054) | 0.71 (0.026) | 0.9 | 465 (13.1) |
| Red-eyed Vireo | 397 | 0.45 (0.044) | 0.26 (0.052) | 0.29 (0.062) | 0.44 (0.065) | 0.9 | 457 (15.9) |

Table 3: $\Delta AIC_c$ values for the four time interval Thrasher data set. A value of 0.0 indicates the most parsimonious model. $\Delta AIC_c$ weight is in parentheses where weight is nonzero. NE indicates models that were not included because of unreasonable parameter estimates.

| Model | No Distance | Constant Distance | Time Distance |
|-------|-------------|-------------------|---------------|
| $M_0$ | 170.2 | 165.2 | 164.2 |
| $M_t$ | 148.1 | 143.1 | 139.7 |
| $M_b$ | 102.3 | 99.5 | 97.4 |
| $M_{tb}$ | NE | NE | NE |
| $M_h$ | 72.2 | 70.3 | 53.4 |
| $M_{bh}$ | 23.3 | 13.7 | NE |
| $M_{th}$ | 0.0 (0.7) | 1.7 (0.3) | NE |

Table 4: Estimated detection probabilities $p_{ij}$, probability of group occurrence $\lambda_j$ and standard errors for interval i and group j of the four time interval Thrasher data set based on model $M_{th}$. Standard error for group two and time interval two is not estimable.

| Parameter | Group 1 | | Group 2 | |
|---|---|---|---|---|
| | Estimate $p_{ij}$ | SE($p_{ij}$) | Estimate $p_{ij}$ | SE($p_{ij}$) |
| $p_{1j}$ | 0.21 | 0.036 | 0.78 | 0.030 |
| $p_{2j}$ | 0.09 | 0.090 | $\approx 1.0$ | NE |
| $p_{3j}$ | 0.47 | 0.050 | 0.84 | 0.023 |
| $p_{4j}$ | 0.56 | 0.055 | 0.70 | 0.026 |
| $\lambda_j$ | 0.29 | 0.037 | 0.71 | 0.037 |

**Chapter 4**

**MULTIPLE SPECIES ANALYSIS OF POINT COUNT DATA: A MORE PARSIMONIOUS MODELING FRAMEWORK**

MULTIPLE SPECIES ANALYSIS OF POINT COUNT DATA: A MORE
PARSIMONIOUS MODELING FRAMEWORK

Mathew W. Alldredge, Biomathematics and Zoology, North Carolina State University,
Raleigh NC 27695.
Kenneth H. Pollock, Zoology, Biomathematics, and Statistics, North Carolina State
University, Raleigh, NC 27695.
Theodore R. Simons, USGS Cooperative Fish and Wildlife Research Unit, Dept. of
Zoology, North Carolina State University, Raleigh, NC 27695.

Abstract.—Although many population survey techniques provide information on multiple

species, these data are rarely analyzed with multiple species models. We develop a

general multiple species modeling strategy based on exploiting similarities in

capture/detection processes among species in order to more parsimoniously model

multiple species data. Data are derived from point count surveys of breeding birds using

distance sampling, time of detection, and multiple observer methods and the set of

candidate models for each method. A method of grouping species together based on

similarities in detection probabilities is discussed and a set of characteristics important in

developing these species groups for aural detection of songbirds is given. Examples are

presented for each analysis method using nineteen species and six species groups

surveyed in Great Smoky Mountains National Park. Species effect models were

generally the most parsimonious models for the distance sampling method but group

models without species effects were generally "better" for the other two methods.

Population estimates were more precise for group models than for single species models,

demonstrating the benefits of exploiting similarities among species when modeling

multiple species data. Partial species effect models and additive models were also useful

because they modeled similarities among species while allowing for species differences.

Although we present analyses for common species with relatively large observed counts, we believe that multiple species modeling will be particularly beneficial for modeling count data from rare species by "borrowing" information on the detection process from more common species. Partial species effect and additive models may be particularly useful in this aspect.

## Introduction

Most available methods that are available for sampling animal abundance are not species specific. Methods such as small mammal trapping (Webb 1965, Schwartz and Whitson 1986, Mengak and Guynn 1987), mist netting birds (Nur et al. 1999), electro fishing (Meador et al. 2003), and avian point counts (Ralph et al. 1997, Canterbury et al. 2000) are not species specific, and all provide information on multiple species. Nevertheless, most analyses of abundance data are done for individual species. Because the capture/detection process may be very similar among species, better estimate precision and more parsimonious models are possible through analyses that exploit species similarities.

The point count survey method is commonly used to estimate the relative abundance or density of bird populations (Ralph et al. 1995, Thompson 2002). Several recent papers have emphasized the necessity of understanding the detection process and the limitations of using counts as indexes of abundance (Nichols et al. 2000, Rosenstock et al. 2002, Farnsworth et al. 2002). The underlying estimator for estimating population size and density from point counts is (Nichols 1992, Lancia et al. 1994, Williams et al. 2002):

$$\hat{N} = \frac{C_i}{\hat{p}_i} \qquad\qquad (1)$$

and

$$\hat{D} = \frac{C_i}{a\hat{p}_i} \qquad\qquad (2)$$

where $\hat{N}$ and $\hat{D}$ are the population estimates, $C_i$ is the observed count for species i, $\hat{p}_i$ is the estimated probability of detection for species i and a is the total area sampled.

There are currently three methods available for estimating abundance from unrepeated point count surveys and modeling the detection process. The distance or variable circular plot method models the probability of detection as a function of distance from the observer, and typically provides an estimate of density (Ramsey and Scott 1979, Reynolds et al. 1980, Buckland et al. 1993). Multiple-observer methods estimate the probability of detection by each observer using a capture-recapture (independent observers) (Chapter 2) or removal (primary-secondary observers) framework (Nichols et al. 2000). Both methods assume all individuals in the sample area are available for detection. The third approach, the time of detection/depletion method, estimates the probability of detection over multiple time periods using a multiple detection capture-recapture or a first detection removal framework (Farnsworth et al. 2002, Chapter 3). This method estimates the combined probability that an individual is available for detection and that it is detected given that it is available.

The detection process modeled by all three methods is similar. It involves the bird making itself available for detection, and the ability of observers to detect available birds. In open habitats the detection of birds is primarily visual and thus availability

depends only on the bird being present and not hidden from view. In forested habitats detections are primarily auditory and availability depends on an individual bird singing or calling during the count (Farnsworth et al. 2002) as well as being present. In either case, the detection process is clearly based on an observer's ability to detect birds by sight or sound.

Because multiple species data are generally collected during point counts it is reasonable to assume that the actual detection process for a given observer is similar for similar bird species. Brightly colored birds that move regularly are more likely to be detected visually than more cryptically colored birds with secretive behaviors. Observers are more likely to detect birds with loud, frequent, high intensity songs then birds with occasional, quiet, low intensity songs.

The standard approach to analyzing point count data is to estimate the detection function for each species separately. This approach ignores any similarities among species (see for example Ralph et al. 1995). When similarities in the detection process do exist among species, then standard approaches will over-parameterize models of the detection process. Nichols et al. (2000) and Farnsworth et al. (2002) tested for species effects in their analyses, but they did not develop a multiple species modeling framework. Frameworks that exploit similarities among species in the detection process should produce more parsimonious models. These models should also provide more precise estimates for less common species that typically have small observed counts.

We present a multiple species modeling framework and explore its potential applications to three methods of analyzing point count data. We begin by discussing the importance of defining species groups and describe characteristics that can be used to

define these groups. We then briefly review model selection and the principle of parsimony, which provides the justification for using multiple species models. A general multiple species modeling strategy is then developed. A brief review of each point count estimation method is given along with a description of the multiple species candidate models and the number of parameters required for each. Examples are then presented for grouping species and analyzing grouped data using three current methods. We conclude with a discussion of the applicability and benefits of this approach to multiple species surveys.

## METHODS

Field methods.—We assume that standard point count survey techniques (Ralph et al. 1995) are used to sample multiple species populations. Specific point count protocols will depend on survey objectives, survey design and the methods of analysis. For details on field methods for distance sampling refer to Buckland et al. (1993), for multiple observer techniques see Nichols et al. (2000) and chapter 2, and for the time of detection approach see Farnsworth et al. (2002) and chapter 3.

Defining the area sampled is necessary for making spatial and temporal comparisons. This is generally accomplished by using fixed radius plots or estimating detection distances to all birds observed. Fixed radius plots may be difficult to apply in a multiple species framework because the appropriate plot radius or maximum detection distance will vary among species.

Defining species groups.—The objective of a multiple species modeling strategy is exploit similarities among species to produce more parsimonious models of the detection process. If only a few species are represented in the sample then analysis can

be done for the entire data set. When sampling methods provide information on a large number of species it will be beneficial to group species with similar characteristics together for analysis. For example, species with very loud vocalization detectable at large distances are likely to have a different detection process than species with quiet vocalizations.

We suggest constructing *a priori* groups of species with similar characteristics associated with the detection process. Using *a priori* groups will lead to biologically meaningful hypotheses that can be tested. Using *a posteriori* groups or "data dredging" for similar detection probabilities is not recommended because it does not necessarily produce biologically meaningful groups and models characteristics of a particular data set rather than the true underlying process. Species should be grouped together based on similarity in characteristics thought to be important to the detection process. These characteristics include sound intensity, singing rate, sound pitch, sound modulation, plumage color, and movement. Their relevance will vary among species and habitats and whether detections are primarily visual or auditory.

Detection distance, which is relatively easily measured may serve as a surrogate for some of the characteristics listed above. Using this approach species groups are constructed by combining species with similar maximum detection distance. In this case truncating data by 10% of the maximum detection distances is advisable to remove extreme observations (Buckland et al. 1993).

Sound intensity, or the energy content of a song (Gill 1995), is one of the most intuitive song characteristics influencing the detection process. Very loud songs are easily detected, whereas, quiet low intensity songs are less likely to be detected.

Measuring sound intensity during a point count would be difficult because the effect of distance would be confounded with any measurements taken.  Actual measurement of sound intensity should probably be done separate to the actual point count survey and should be done at some standardized distance.

Sound pitch, or frequency of a song, can affect the detection process in two ways.  First, an observer may be able to hear sounds in certain frequency ranges better than those in other ranges.  In this situation sounds of similar pitch would have similar detection probabilities for a given observer.  The other factor associated with sound pitch is how well a sound will travel in certain habitats.  Low-frequency sounds travel better or are less subject to attenuation in dense vegetation than are high-frequency sounds (Morton 1975, Wiley and Richards 1882).  Grouping species by similarity in sound pitch would create groups of species subject to similar sound attenuation.  Sound pitch may also be difficult to measure while conducting a point count survey.

Sound modulation, variation in either sound pitch or intensity (Gill 1995), may also affect the detection process.  The reasons sound modulation may affect the detection process are the same as those discussed for sound intensity and sound pitch.  Sound modulation should also be considered for its affect on a listener's ability to locate sound.  Songs that cover a broad frequency range are easier to locate than songs that cover a narrow frequency range (Wiley and Richards 1982, Gill 1995).  For example, alarm calls cover a narrow frequency range and are difficult to locate in comparison to songs that are used to attract mates.  This has important implications to point count surveys as birds must be located with some degree of accuracy for both fixed and unlimited radius point counts.

Singing rate is very important for modeling the detection process, especially when there are a large number of birds singing simultaneously (McShea and Rappole 1997). Because individual songs are detected with some probability, the more a bird sings the greater the probability that an observer will detect it. For example, if the probability of detecting any given song is 0.2 then the probability of detecting that bird at least once during a count is 0.2 if it only sings once, 0.67 if it sings 5 times and 0.89 if it sings 10 times. If the probability of detecting a single song is low then grouping species by similarity in singing rates may be useful because these groups will account for differences among species in the probability of detecting an individual at least once during the count. If the probability of detecting a single song is high, then this will not matter because the probability of detecting an individual at least once during a count approaches one, even with very low singing rates.

Plumage color and movement are important characteristics for point count surveys when detections are visual, but they may also be important when detections are primarily aural. Brightly colored birds, or birds that are actively moving during a count may attract an observer's attention to a specific area and increase the probability of detection by ear. More cryptically colored birds or sedentary birds would not provide this type of cue to an observer.

Using multiple characteristics to define species groups can be done by either averaging the values across groups for each species or by using a hierarchical approach. We prefer the hierarchical approach because it gives more weight to the characteristics thought more important. A hierarchical approach is to first group by the most important characteristic and then dividing groups further by using additional characteristics. These

hierarchical levels can then be used for between group comparisons as well by starting at the refined group levels for group comparisons.

Model selection.—Akaike's Information Criterion (AIC) is an information theoretic approach to selecting the most parsimonious models (Burnham and Anderson 2002). Models with a large number of parameters may have low bias but they have large estimation variance associated with every parameter. Models with fewer parameters have smaller variance associated with estimating parameters but have greater bias. Parsimonious models provide the "best" balance between estimation variance and bias.

The two components of AIC are deviance from the saturated model, which controls for bias, and the number of model parameters, which controls for estimation variance associated with the number of parameters in the model. Using single species models to analyze multiple species count data does not fully account for the number of parameters and may lead to over-parameterized models. For example, if a single species analysis is used for 5 species, and a two parameter model is selected as the most parsimonious model, then the total number of parameters estimated is ten. If the detection process is similar among all of these species then the most parsimonious multiple species model would not have species differences, it would only estimate two parameters, and estimates would be more precise than the single species approach.

Comparing multiple species models to more familiar models that account for animal level differences (e.g. comparing males versus females or juveniles versus adults) or observable heterogeneity will illustrate the procedure from a model selection framework. When capture-recapture experiments are used to estimate abundance of a single species, models that account for observable heterogeneity, also known as animal-

level stratification are generally used (Chapman and Junge 1956, Seber 1982, Lebreton et al. 1992, Borchers et al. 2002). For example, in some cases capture probabilities differ between males and females or between adults and juveniles. In this situation one would fit a model assuming males and females have the same capture probability and another model assuming differences in capture probability. The most parsimonious model would be selected based on AIC. Multiple species models are identical in concept to this capture-recapture modeling approach. One candidate model assumes that all species in a survey, or all species within a predefined group, have similar detection probabilities. The alternative model assumes species-level differences and accounts for them as observable differences in detection probabilities.

Multiple species modeling strategy.—The strategy we present is an *a priori* hierarchical approach of first examining the within group structure and then, if warranted, examining the between group structure. The first step in the modeling process is examining within group structure, testing the assumption that species within defined groups have similar detection probabilities. A set of candidate models for the within group structure of the detection process is developed for each point count methodology. If within group models suggest species differences then we do not proceed because the number of model parameters will be similar to those for individual species. If no differences in detection probability are found within a species group, we proceed to between group models.

Multiple observer and the time of detection methods can model the species effect as either an additive or an interaction effect. An interaction effect is equivalent to a species specific analysis where observer or time differences in the detection probabilities

vary among species. Using a linear models framework the interaction model has a unique slope and intercept for all species in the analysis. The additive species effect model assumes that there are differences between species but the differences in detection probabilities between observers or time intervals are consistent between species. Only a single parameter is necessary to model the difference between species using the additive model, which is really just an adjustment to the intercept for each species in the analysis but a common slope is used.

For consistency we will use similar notation to define the models for all of the methods. Effects seen in the original single species models, such as time (t), observer (obs), and difference in subsequent detection from first detection (b), will appear as subscripts. Superscripts will be used to denote models with interaction species effects ($^{*spp}$), additive species effects ($^{+spp}$) or without species affects ($^{0}$).

### Estimation methods and candidate models

Distance.—The distance sampling method (Ramsey and Scott 1979, Reynolds et al. 1980, Buckland et al. 1993) is probably the most commonly used statistical method for estimating abundance or density from point count data, primarily because the other two methods have only recently been developed. Distance sampling involves observers recording the radial distance to all species seen/heard. Distance sampling assumes that the probability of detection for a given species is a monotonically decreasing function of distance from the observer. If we let k be the number of points sampled with effective search radius w, then equation 2 can be rewritten as (Buckland et al. 1993),

$$\hat{D} = \frac{C_i}{k \pi w^2 \hat{p}_i} \qquad (3)$$

The probability of detection ($\hat{p}_i$) within a circle of radius w is,

$$\hat{p}_i = \frac{2}{w^2} \int_0^w rg(r)dr, \qquad (4)$$

where g(r) is the detection function, which gives the probability of detecting an individual given its radial distance (Buckland et al. 1993).

Distance methodology is based on the estimation of g(r), which determines the number of parameters in the model. Buckland et al. (1993) recommend a class of reliable models based on model robustness, shape criterion, and estimator efficiency. The recommended model is based on a key function and possibly a series expansion to adjust the key function, of the general form,

$$g(r) = key(r)[1 + series(r)] \qquad (5)$$

The recommended key functions for consideration are: the uniform, the half-normal, and the hazard-rate, which have 0, 1, and 2, parameters respectively. Each series adjustment term used in the model adds an additional parameter. Buckland et al. (1993) suggest that the key function is often adequate for properly truncated data. In other cases one or two adjustments may be necessary.

The assumptions for point count distance sampling are (Buckland et al. 1993):

1.  g(0) = 1, certain detection at the point.

2.  Objects are detected at their initial location.

3.  Distances are measured accurately or placed in the correct distance categories.

4.  Objects are detected independently.

An additional underlying assumption is that all individuals within the sampled area are available for detection visually or aurally or that inferences are only made about the available portion of the population (Marsh and Sinclair 1989, Farnsworth et al. 2002). This assumption is often overlooked but could be important for some species.

The distance sampling methodology has the most limited set of models for examining multiple species data, because the method only estimates a detection function relative to distance from the observer. Therefore, comparisons among species are constrained to comparisons of the functions used to estimate abundance.

Following the general approach discussed above, the two candidate models for the within group detection process are: 1) the same detection function for all species within group j (no species effect) and 2) different detection functions for all species within group j (species effect). If the "best" model for the groups indicates different detection functions for all species, then among group comparisons are not necessary. If detection functions are found to be the same for all species within a group, then among group comparisons are recommended.

The number of parameters used when examining each species separately is just the sum across all species of the number of parameters in each species specific detection function. With a multiple species approach, the number of parameters used is the same if the selected model represents different detection functions within groups. When a single detection function is selected for all species within a group, then the number of parameters is the sum of the number of parameters in each group-specific detection function. Parameters are further reduced if among group comparisons suggest that common detection functions are appropriate.

Time of detection.—The time of detection method was first proposed by Farnsworth et al. (2002) as a removal method in which the data consists of the times of first detection. A more general approach uses the full detection history of every individual, applying the full set of capture-recapture models (Otis et al. 1978) as a framework for analysis. Excluding the behavior models (Otis et al. 1978), the full time of detection approach is generally more efficient (smaller variance) than the constrained removal approach (Chapter 3 and see Seber 1982 pg.570). With this method, models determine the probability of detection between time intervals of a point count. Ralph et al. (1995) recommended using 3, 2, and 5 minute intervals so that future data are comparable to other studies using these time intervals. Unlike other methods of estimating detection probabilities from point count data, the time of detection method actually estimates the product of the probability that an individual is available for detection during a count, and the probability that an individual is actually detected given that it was available (Farnsworth et al. 2002, chapter 3). This can be seen by examining first detections in the second interval. These detections could either be of individuals that were not available for detection in the first interval but are in the second, or of individuals that were available in both the first and second intervals but were not detected in the first. Note that both movement and singing rates can affect the availability process. The time of detection approach assumes that movement is not a factor.

The assumptions for this approach are (Farnsworth et al. 2002, chapter 3):

1.      There is no change in the population of birds within the detection radius during the point count (closed population).

2.      There is no double-counting of individuals.

3.      Detection distances are measured accurately, or if limited radius counts

          are used, observers accurately assign birds to within or beyond the

          radius used.

There are 8 general capture-recapture models (Otis et al. 1978, White et al. 1982, Pollock et al. 1990), all of which may be applicable to the time of detection approach. The capture-recapture models and a description relevant to the time of detection method are (chapter 3):

Model $M_0$: Equal detection probability for all individuals among all time periods.

Model $M_t$: Equal detection probability for all individuals but different detection
        probabilities among time periods.

Model $M_b$: Equal probability of first detection for all individuals among all
        periods and a unique probability of subsequent detections that is equal for
        all individuals among all subsequent periods.

Model $M_{tb}$: Equal probability of first detection for all individuals but different
        among periods and a unique probability of subsequent detection that is
        equal for all individuals but different among subsequent time periods.

Model $M_h$: Unique probability of detection for each individual that remains
        constant among all time periods.

Model $M_{th}$: Unique probability of detection for each individual that differs
        among time periods.

Model $M_{bh}$: Unique probability of first capture for each individual that remains
        constant among time periods and a unique probability of subsequent
        detection that remains constant among time periods.

Model $M_{tbh}$:  Unique probability of first capture for each individual that differs

among time periods and a unique probability of subsequent detection that

differs among time periods.

Models $M_b$ and $M_{bh}$ are the removal models given by Farnsworth et al. (2002).

The full heterogeneity models (Burnham and Overton 1979, Chao et al. 1992,

Norris and Pollock 1996, Pledger 2000) are only applicable when four or more time

intervals are used.  With three time periods it is possible to fit constrained two-point

mixture models (Norris and Pollock 1996, Pledger 2000), by setting the detection

probability for one of the heterogeneity groups to a constant.  The assumption here is that

some birds in the point count are very easily detected while other birds are harder to

detect.  Farnsworth et al. (2002) made the assumption that the probability of detection for

one of the heterogeneity groups was one, and then estimated detection probabilities for

the other heterogeneity group.  An alternative is to use information from other studies,

such as a four interval study, to set a more realistic value for the high detection

probability (chapter 3).

The structure of the time intervals used in the survey will determine how the

detection process is modeled with this approach (Chapter 3).  If equal time intervals are

used, then the interval detection rate can be estimated directly using standard capture-

recapture software for analysis, such as programs CAPTURE (White et al. 1982) or

MARK (White and Burnham 1999).  When unequal time intervals are used then it is

necessary to model the detection rate as a function of time to allow for models with

constant detection rates.  The detection rate can be modeled as a discrete per minute rate

$(\gamma_i)$  (Farnsworth et al. 2002, chapter 3)

$$p_i = \gamma_i^{t_i} \qquad\qquad (6)$$

or as an instantaneous rate ($\varphi_i$) (chapter 3)

$$p_i = 1 - e^{-\varphi_i t_i} \qquad\qquad (7).$$

For these situations the model likelihoods are not in the form required by the standard capture-recapture programs. Therefore, it is necessary to parameterize the likelihood using either discrete or instantaneous rates formulation that is maximized using a program such as SURVIV (White 1983).

Models with no time effect assume a constant detection rate and interval probabilities varying as a function of the interval length. The examples provided for this section are from point counts with three unequal time intervals. We will also use the instantaneous rates formulation (see chapter 3 for more details).

The most general model is $M_{tbh}$, and we present the likelihood for this model using a two-point mixture for heterogeneity. All other models are just constrained forms of this general model. A point count with i intervals has $2^i$ capture histories, one of which is not observable because it represents those individuals never detected. For a four interval point count there are 15 observable detection histories (w) with observed counts ($x_w$). The expected value for the number of birds detected in every interval is given by:

$$E(x_{1111}) = N\left(\lambda\, p_{11} c_{21} c_{31} c_{41} + (1-\lambda)\, p_{12} c_{22} c_{32} c_{42}\right) \qquad\qquad (8)$$

where $\lambda$ is the proportion of animals in group one. The proportion of animals in group two is $1-\lambda$, and the $p_{ij}$'s and $c_{ij}$'s are the probabilities of first detection and subsequent detection in the $i^{th}$ interval for individuals in the $j^{th}$ group. The remaining expected values are written similarly by modifying equation 8 to account for intervals where individuals

are not detected. Using the rates formulation, the $p_{ij}$'s and $c_{ij}$'s are replaced by one of the rate equations (equations 6 or 7). The conditional likelihood $L(p_{ij}, c_{ij}, \lambda \mid x_w)$ for the two-point mixture model $M_{tbh}$ is given by (chapter 3):

$$L_2(p_i, c_i \mid x_w) = \binom{x_T}{x_{1111} \cdots x_{0001}} \prod_1^w \left( \frac{\pi_{w1} + \pi_{w2}}{\pi.} \right)^{x_w} \tag{9}$$

where $x_T$ is the total observed count, $\pi.$ is the probability of being detected at least once during the entire count, $\pi_{w1}$ is the probability of being in the first group and having capture history w, and $\pi_{w2}$ is the probability of being in the other group and having capture history w (chapter 3).

Maximizing the above likelihood for the observed data will give estimates of the the detection rates ($\varphi_{ij}$'s) and probability of being in one of the groups ($\lambda$), which can be used to estimate the probability that an individual is detected at least once during a four interval count as;

$$\hat{p}_T = 1 - \left[ \hat{\lambda} \left( 1 - e^{-\hat{\varphi}_{11} t_1} \cdots e^{-\hat{\varphi}_{41} t_4} \right) + \left( 1 - \hat{\lambda} \right) \left( 1 - e^{-\hat{\varphi}_{12} t_1} \cdots e^{-\hat{\varphi}_{42} t_4} \right) \right] \tag{10}$$

The variance of $\hat{p}_T$ can be obtained by reparameterizing the model to include $\hat{p}_T$ as a model parameter and thus, obtain an estimate of the variance for probability that an individual is detected at least once. Abundance can then be estimated using the total observed count ($x_T$) as;

$$\hat{N} = \frac{x_T}{\hat{p}_T} \tag{13}$$

The variance is

$$\hat{Var}(\hat{N}) = \frac{x_T^2 \hat{Var}(\hat{p}_T)}{\hat{p}_T^4} + \frac{x_T(1-\hat{p}_T)}{\hat{p}_T^2} \tag{14}$$

assuming that the observed count ($x_T$) is from a binomial distribution (Nichols et al.

2000, Williams et al. 2002). All other models are obtained by constraining this general

form.

The set of candidate models using the time of detection method for multiple

species analysis has 14 models, although we propose 6 additional models. The 14

candidate models are obtained by using the seven single species models (chapter 3) and

analyzing them with or without a species effect. Three of the additional models are

variations of time and species effect models. One approach is allowing time specific

detection probabilities to vary independently among species so that there is no

relationship between the detection probabilities of different species. A second approach

provides for species differences, but the variations in detection probabilities with time are

consistent across species. If time specific detection probabilities are known for one

species, then a single species adjustment is added to obtain the detection probabilities for

the other species. In other words the first approach assumes a time and species

interaction, and the second approach is an additive effect related to species differences.

The other models we propose are individual heterogeneity models, which allow

for differences in detection probabilities among individuals. Model $M_h^0$ assumes that all

detection probabilities and probabilities of being in the first heterogeneity group are the

same for all species. Models $M_h^{*spp}$ and $M_h^{+spp}$ assumes that detection probabilities and

probabilities of being in the first heterogeneity group are different among species. We

suggest that another reasonable model is $M_h^{part}$, which assumes that detection

probabilities are the same for all species but that the probability of being in the first group

is not the same for all species. Differences in singing rates related to breeding stage are

one source of individual heterogeneity. Model $M_h^{part}$ describes a process in which the

probability of detection is similar among species at various breeding stages but that the

proportion of the population of each species in a given breeding stage varies.

The number of parameters for each of these models is given in Table 1.

Comparing the number of parameters for models without species effects to those with

species effects demonstrates the reduced number of parameters that can be achieved by

using multiple species models.

Multiple observers.—There are two approaches to the multiple observer method;

the dependent observer approach (Nichols et al. 2001) and the independent observer

approach (Chapter 2). The two approaches are very similar except that the dependent

observer approach treats the data as a removal method and the independent observer

approach utilizes the full capture history. The removal method (Nichols et al. 2000) uses

2 observers, one primary and one secondary. The primary observer records all

individuals detected and communicates this to the secondary observer. The secondary

observer records individualsnot detected by the primary observer. The independent

observer approach (chapter 2) differs in that all observers conducting the count record all

detections independent of the other observers. For a 2 observer count there are 3

potential capture histories ($x_{11}$—seen by both, $x_{10}$ and $x_{01}$ seen by only one) instead of the

2 histories from the dependent observer approach.

The multiple observer method of estimating abundance from point count data is actually a direct application of capture-recapture models and can be analyzed with programs, such as CAPTURE (White et al. 1982) or MARK (White and Burnham 1999). Therefore, we will not present the likelihoods but simply describe the method and discuss potential models. For more details on capture-recapture models refer to Otis et al. (1978), Pollock et al. (1990), or Williams et al. (2002).

The assumptions for the multiple observer models are:

1. Counts within the fixed radius circle are measured accurately.

2. There are no matching errors between the observers so that the assignments to $x_{11}$, $x_{10}$, and $x_{01}$ are accurate.

3. Equal detection probability of each species at all points for each observer, except with heterogeneity models.

4. There is no undetected movement into or out of the fixed radius circle.

An additional assumption for the dependent observer approach is that the primary observer detects individuals independent of the secondary observer. The independent observer approach assumes that observers detect individuals independent of one another.

We will only use the independent observer approach because this is the more general method. The full suite of capture-recapture models presented in the previous section are also parameterized for these types of data. However, the models with behavioral effects are biologically unreasonable for this application (Chapter 2), because all observations are done simultaneously, and observers make detections independently. Therefore it is unreasonable to assume that detection by one observer affects detections by other observers.

For two observers the approach is in the form of a Lincoln-Petersen closed-population capture-recapture survey (chapter 2). If we let $x_{11}$ be the number of individuals detected by both observers, $n_1$ be the number detected by observer 1, and $n_2$ be the number detected by observer 2, we can write the probability of detection for each observer as;

$$\hat{p}_{d1} = \frac{x_{11}}{n_2}$$

and                                                                                            (15)

$$\hat{p}_{d2} = \frac{x_{11}}{n_1} \quad ,$$

The probability of detection by at least one-observer is,

$$\hat{p}_d = 1 - \left(1 - \hat{p}_{d1}\right)\left(1 - \hat{p}_{d2}\right) \cdot \tag{16}$$

Using the canonical form (equation 1) we can then estimate abundance. This is done using program MARK (White and Burnham 1999) with the Huggins Closed Captures data type, which estimates the probability of detection for each observer and abundance as a derived parameter. The number of parameters in a single species model for t independent observers is 1 for model $M_0$ and t for model $M_t$.

Individual heterogeneity in detection probabilities is presumably important in point count data regardless of the analysis method used (Farnsworth et al. 2002, Chapter 2 and 3). It is possible to account for observable heterogeneity for two or more observers using covariates, such as detection distance (Chapter 2). With four or more observers it is possible to account for unobservable heterogeneity using heterogeneity estimators such as the Jackknife (Burnham and Overton 1979), the Chao sample coverage estimator (Chao et al. 1992) or the finite mixture estimators of Norris and Pollock (1996) and

Pledger (2000). For illustrative purposes in this paper we will use a four-independent observer data set and two-point mixture estimators of individual heterogeneity in detection probabilities.

There are 12 candidate models for the multiple species independent observer approach, excluding any covariate models (Table 2). These models are obtained by adding a species effect to the four single species candidate models. Observer and species effects are modeled two different ways. The first approach, which is equivalent to the single species approach, assumes that observer effects differ across species and it models the probability of detection for each observer separately for each species (denoted with superscript $^{*spp}$ in the models). The second approach assumes that observer and species effects are additive. It assumes the probability of detection varies among observers but that the differences among observers is consistent across species (denoted with a superscript $^{+spp}$). This situation might arise when one observer is more skilled than the other and consistently detects a larger proportion of the birds at a point. Detection probabilities are estimated for each observer and a single parameter is used to adjust the detection probabilities for each subsequent species. We also consider a partial individual heterogeneity effect, similar to that used in the time of detection models. This partial heterogeneity effect allows detection probabilities to remain constant among species but the probability of being in the first heterogeneity group varies among species.

The number of parameters estimated for each model is provided in Table 2. Like the time of detection method, the number of parameters required for models without a species' effect is less than the number required for models with a species effect.

**Field Data**

To demonstrate the multiple species modeling approach we use point count data for songbirds collected in Great Smoky Mountains National Park during the breeding season. Survey points were located along low use hiking trails with a minimum 250 m separation between points. Surrounding vegetation was closed-canopy deciduous hardwood forest. Because of this detections were primarily aural (over 95%, Simons unpub. data). All point counts were conducted between dawn and 10:15 in the morning on days with good weather (no rain or excessive wind).

All observers were highly trained prior to conducting point counts. Training included identification skills, familiarity with birds occurring in the area, and distance estimation skills. Observers used laser range finders to delineated a 50 m radius circle prior to starting each point count and to verify distance estimates during the count.

The time of detection data were extracted from point count surveys conducted during May and June from 1996 to 1999. To avoid temporal and observer effects we only used data from 1998 collected by one of the most experienced observers. This restricted data set consisted of 323 survey points. We also omitted species with fewer than 50 observations from analysis. Observation periods were delineated by birds detected in the first three minutes, next two minutes, and final five minutes of a ten minute count. Multicolor pens were used to denote the time of initial detection and subsequent detections were recorded by underlining previous detections in the appropriate color. The example data set for the distance method was obtained by using only the observations recorded in the first 3 minutes of the time of detection data set.

A four-independent observer data set was collected in the same area during June of 1999. All observers were highly trained and had been conducting point counts for six weeks prior to this survey. Counts were conducted at 70 points and followed the same protocol for weather conditions and distance measurements. Observers conducted variable circular plot 3 minute point counts (Reynolds et al. 1980) and mapped the location of each bird detected at the point. Attempts were made to track each bird during the point count to avoid double counting and to simplify matching observations among observers. Following each point count observers combined their data to determine the total number of birds detected and identify birds seen in common.

Species groups.—Species groups were defined by obtaining rankings from 7 experts familiar with the species and study area used for our examples. Ranks were done on a scale of 1 to 5 with 1 indicating the lowest probability of detection or smallest value for a given characteristic and 5 indicating the highest probability of detection. These ranks were then averaged for the 7 experts that provided ranks. Maximum detection distance for each species was also used to define species groups.

We used the maximum detection distance, following 10% truncation of the largest distances, as the first criterion for defining species groups. Because many characteristics (sound intensity, pitch, modulation, etc.) can affect maximum detection distance, we defined three groups: 1) species with maximum detection distance $\leq 100$ m, 2) species with maximum detection distance $> 100$ m and $\leq 150$ m, and 3) species with maximum detection distance $> 150$ m.

Because maximum detection distance may contain information about sound intensity, sound pitch, and sound modulation, we chose to use singing rate to further

refine these groups. Birds with similar singing rates within a distance group were used as the final groups. Further refinement of groups was not necessary.

Distance analysis.—We used program DISTANCE (Thomas et al. 2002) to analyze the distance data. Separate analyses were done for each group to determine if a species effect was present within each group. Data format for DISTANCE was standard except that the species identification had to be entered as an observation level variable, which is the level at which detection distances are entered. This allowed for post stratification by species (Rosenstock et al. 2002).

Data were truncated at 10% of the maximum detection distances for all analyses, as recommended by Buckland et al. (1993). For each group an analysis was run for no species effects and for species effects. For each of these effects the following key functions and adjustments were run:

1. Half-normal key function – Cosine adjustment

2. Half-normal key function – Simple polynomial adjustment

3. Uniform key function – Cosine adjustment

4. Uniform key function – Simple polynomial adjustment

5. Hazard rate key function – Cosine adjustment

The appropriate key function and adjustment model are selected using $AIC_c$ (AIC corrected for small sample size) for both the species effect and no species effect models. $AIC_c$ was then used to choose between the species effect and no species effect models. Between group comparisons were not warranted. We use the effective detection radius (EDR) and the density estimate (D) to compare models with and without a species effect.

Time of detection analysis.—Time of detection data were analyzed with program SURVIV (White 1983) to estimate the detection parameters and two-point mixture heterogeneity parameters since the data are for three unequal time intervals and assumed the instantaneous rates formulation (equation 7) to model the interval detection probability. All 17 candidate models were initially run, but both of the models with time and behavior were omitted from the final analysis because all parameters were not identifiable. Selection of the most parsimonious model was done using $AIC_c$.

Population estimates and standard errors were derived from the estimated detection probabilities and heterogeneity parameters (equations 10-14). We report the probability of detecting an individual at least once during the count ($p_T$), the heterogeneity parameter ($\lambda$), and the population estimates for the selected model and the alternative species effect model.

Multiple observer analysis.—We used program MARK (White and Burnham 1999) with the "Huggins Closed Captures" and "Huggins Full Heterogeneity" data types to analyze the four-independent observer data. Two-point mixture models were used for the heterogeneity models because estimates from three-point mixture models were not reasonable. Model selection was based on $AIC_c$.

We report the heterogeneity parameter estimate, the observer specific detection probabilities, and the population estimate for the selected model. We also report the population estimate for "best" species effect model to demonstrate the improved precision when models with no species effect were selected.

**Results**

Species groups.—There were 19 species from the time of detection/distance data sets selected for analysis. Three groups were defined in the ≤ 100 m category, two groups in the 100 m to 150 m category, and one group in the over 150 m category, based on similarities in singing rates (Table 3). Group sizes ranged from two to four species.

Because fewer points were sampled using the independent observer method, only 8 species had sufficient sample size for analysis and species groups were modified. No analysis was done for group A, the Black-throated Blue Warbler and Indigo Bunting had to be omitted from groups B and C, respectively, and the Ovenbird, Scarlet Tanager, and Tufted Titmouse were the only species used from groups D, E, and F to form the combined group DEF.


Distance.—Species effects models were selected as the most parsimonious models based on comparison of $\Delta AIC_c$ values for five of six species groups (Table 4). These values reflect the difference between the $AIC_c$ value of a particular model and the model with the lowest $AIC_c$ value. The model with $\Delta AIC_c$ of zero was the most parsimonious model and competing models were those with $\Delta AIC_c <2$ (Burnham and Anderson 2002). When the species effect model was selected the no species effect model had $\Delta AIC_c >2$ indicating the species effect model was required to sufficiently explain the variation in the data. The single instance when the no species effect model was selected $\Delta AIC_c=1.76$ indicating that both the species effect and no species effect models had similar ability to explain the variation in the data.

All species in group A used a uniform key function, and three of the four had a simple polynomial adjustment term. All species in group B used the hazard rate key function, except the Dark-eyed Junco, which used a half-normal key function. In group C the hazard rate key function was used for the Black-throated Green Warbler and the uniform key function was used for the other two species. The uniform key function was used for all species in groups D and E, and the Hazard rate was used for all species in group F. In general the shape of the detection function was similar among all species in a group.

The effective detection radius did vary within groups except for group D which selected the no species effect model (Table 5). Differences in the effective detection radius were >10 for all other groups. These differences affected the shape of the detection function when data were pooled across species. For example, for group F the species specific detection function was a hazard rate with a cosine adjustment term for all species, but the detection function for the group was a half-normal with a cosine adjustment term. In contrast, a uniform detection function with a cosine adjustment term was selected for the species specific detection function and the group detection function for group D.

Group D demonstrates the benefits of the multiple species approach. The standard errors were generally smaller for the no species effect model for all groups but these estimates were biased for all groups except group D. Density estimates for group D were identical among species effect and no species effect models, but the standard errors were smaller for the no species effect model (Table 5). The increase in precision is especially obvious for the Veery, which had a smaller observed count.

Time of detection.—Individual heterogeneity models explained the data more parsimoniously than models that did not account for heterogeneity (Table 6). Individual heterogeneity models were selected for groups A and D with no time effect and no species effect on detection probabilities. Group A did not have a species effect on the heterogeneity parameter, but group D did. Selected models for all other groups were heterogeneity models with a time effect on the detection probabilities, groups B, C, E, and F. Species effects on detection probability and the heterogeneity parameter were also important in explaining the variation in the data for groups B and E. Species differences in detection probabilities and the heterogeneity parameter do not appear in the selected model but they may be important because the species effects model was very similar in its ability to explain variations in the data ($\Delta AIC_c$ weights 0.35 and 0.33, respectively). The selected model for group F had a species effect on the heterogeneity parameter but not on the detection probabilities. Models with both a time and behavior effect (change in detection probability following initial detection) never gave reasonable parameter estimates (standard errors >>1 for probabilities). Note that for some species groups, there are alternative competing models to explain the data based on similar $\Delta AIC_c$ values and similar $\Delta AIC_c$ weights. No models were analyzed to compare between groups because similar models were not selected for similar groups. For example, groups A, B, and C were all in a similar distance category but the models ranged from a heterogeneity model with no species or time effects to one that incorporated both of these effects.

The estimated detection probabilities ranged from 0.81 (group F) to 0.92 (group E) (Table 7). The estimated heterogeneity parameter ranged from 0.15 to 0.79, both of these estimates occurred in group F. Models that used common parameters among

species in a group showed increased precision for all parameter estimates. This is most clearly seen in the increased precision for the abundance estimates, especially for species with smaller observed counts.

Independent observer—Heterogeneity models explained the data in a more parsimonious manner for all three independent observer species groups than models not incorporating heterogeneity (Table 8). The heterogeneity model with no observer or species effects was the selected model for group B and was a reasonable alternative model for the other two groups ($\Delta AIC_c$ weights $\geq 0.22$). The heterogeneity model with no observer or species effect of the probabilities of detection and a species effect on the heterogeneity parameter was selected for group DEF and was also a reasonable alternative model for group C ($\Delta AIC_c$ weight $= 0.26$). An individual heterogeneity model with an observer effect on the probability of detection and no species effect was selected for group C. Selection for these simpler (fewer-parameter models) may have been partially because of the small observed counts in these data sets. Because the heterogeneity model with no observer or species effects was a reasonable model for both groups B and C, an additional model was run to determine if these groups could be combined. The group effect model was selected ($\Delta AIC_c = 13.8$) suggesting that group differences were important sources of variation in the data.

The heterogeneity parameter ranged from 0.34 to 0.58. The detection probability was generally $> 0.90$ for the high detection probability heterogeneity group for all species groups, with the exception of the detection probability of observer 3 in group C. The detection probabilities for the other heterogeneity group were $\leq 0.36$ for all species groups. Comparing the abundance estimates for the selected model to the abundance

estimates for the "best" species effect model shows very similar estimates but considerably smaller standard errors for the selected models with no or partial species effects.

## Discussion

Application of multiple species models to population surveys offers a promising approach to analyzing data when more than one species occurs in the sample. We have demonstrated that following a multiple species modeling approach will give more parsimonious models and better precision of estimates for species with similar detection processes. Evidence from the analyses presented here clearly indicates that in many cases single species analyses are over-parameterized.

The multiple species analysis procedure allows for direct comparison between models with and without species effects to determine if group based parameter estimates are warranted. Another approach to multiple species analyses simply assumes similarities among species and analyzes them as a group. This has been done for less common species when observed counts are not sufficient for a single species approach (Nichols et al. 2000). Testing this assumption in a multiple species analysis strengthens inferences about the population.

It is important to consider the effect of small counts on model selection. When observed counts are small, then estimation variance for model parameters will be larger and model selection will tend to select models without species effects. This results in greater bias but smaller estimation error. In other words, estimates may be precise but not accurate.

We feel that the additive species effect models are very important in situations where observed counts are small for some species. This is done by "sharing" information for estimation of parameters among species but still incorporates species effects. Consider a simple case for a common and a rare species. When a multiple species approach is used the precision of estimates will not be affected much for the common species, but they will be improved for the rare species. As an example, assume that the count was done using multiple observers and that the "true" underlying detection processes are different among species. An additive model that assumes a constant difference in the detection process among species accounts for this situation. Differences in detection probabilities among observers are estimated using the full data set (both species), which primarily is based on the species with a large count. In these cases only a single parameter is needed to estimate the detection probabilities for the rare species. The model "borrows" information from similar species to estimate the detection process and it gives more precise estimates than a single species approach. Similar data "sharing" was also demonstrated in models with a partial species effect where the detection probabilities were constant among species but the heterogeneity parameters were modeled with a species effect.

Individual heterogeneity models were always selected as the most parsimonious models to describe the data. There are many sources of variation that cause heterogeneity in animal surveys (Burnham 1981, Johnson et al. 1986). Temporal variation in singing rates (Wasserman 1977, Lein 1981) is a potential source of heterogeneity. Habitat, local abundance and proximity of observers has also been shown to affect singing rates of breeding songbirds (McShea and Rappole 1997). Further investigation is needed to

determine the effect of singing rate on detection probability and to identify and account for other sources of heterogeneity.

Multiple species models were not useful in conjunction with the distance sampling method. This is in part because these models have few parameters to begin with. The species effect models had at most 3 parameters, so the reduction in model parameters was not as great as in the other methods.

We feel that one of the problems with using the multiple species approach in conjunction with distance sampling is related to species differences in the detection radius. Pooling data for species with different effective detection radii changes the shape of the detection function and can make it harder to fit. In our analyses, the only model selected without a species effect was a group model where the effective detection radii where almost identical among species. If singing rates significantly affect detection rates then the assumption of declining detection probability with increasing detection radius may also be violated. For forest songbirds, including Ovenbirds and Wood Thrushes, singing rates have been shown to decrease with proximity of an observer (McShea and Rappole 1997). If singing rates are important to the detection process then the use of distance sampling may be problematic and multiple species models may not work if observer effects differ among species.

The multiple species approach was beneficial and provided more precise estimates of abundance for both the time of detection and multiple observer methods. These methods worked well because they were exploiting similarities among species in song structure and singing behavior and they were not as sensitive to detection distance. Individual heterogeneity models proved to be the most applicable to these data. If

singing rate is an important factor affecting aural detection of birds then mixture models of individual heterogeneity are particularly applicable because heterogeneity groups will reflect the proportions of the population in various breeding stages. Several studies have documented changes in singing rates relative to the breeding stage of an individual (Wasserman 1977, Lein 1981). Partial heterogeneity models may also be useful because they model similarities among species detection probabilities but they do not restrict the proportion of the population in each heterogeneity group.

We used both quantitative and qualitative information to define species groups. The use of maximum detection distance worked well to categorize species into broad groups but it did not allow complete classification of species groups. Singing rate information created reasonable species groups within the distance categories. Using assumed ranks is not the most desirable situation. People's perception of species detectability can be drastically different. For example, Farnsworth et al. (2002) suggested that the Acadian Flycatcher was one of the more detectable species in their surveys in Great Smoky Mountains National Park and attributed this to high singing rates. Our average ranking from seven experts suggests that the Acadian Flycatcher actually has one of the lower singing rates of the birds used in our analysis. A more direct measure of characteristics used in defining species groups would be beneficial to future studies, but would require extra field effort to collect the required data.

## Recommendations

We believe that a multiple species modeling approach should be strongly considered in the analysis of any data collected on multiple species. Because two of the

methods presented here are based on the closed population capture-recapture models, the benefits to analysis of similar data types and candidate models is clear. Development of candidate models and gains in precision need to be investigated for other data types, such as Cormack-Jolly-Seber (Seber 1982, Williams et al. 2002) and tag return models(Brownie et al. 1985, Williams et al. 2002).

Although the additive models were not selected in the analyses presented here there were some cases where they were reasonable alternative models. The partial heterogeneity models did demonstrate the benefit of "sharing" information to obtain more precise parameter estimates. Application of these types of models for multiple species analysis should be investigated further as they do not require the entire detection process to be similar among species. The use of covariates, such as detection distance, as additive effects in multiple species models should also be investigated.

# Literature Cited

Borchers, D.L., S.T. Buckland, and W. Zucchini. 2002. Simple mark-recapture. Pages 104-130 in Estimating animal abundance: Closed populations. Springer, London.

Brownie, C., D.R. Anderson, K.P. Burnham, and D.R. Robson. 1985. Statistical inference from band recovery data—A handbook, 2nd Ed. U.S. Fish and Wildlife Service Resource Puplication. 156.

Buckland, S.T., D.R. Anderson, K.P. Burnham, and J.L. Laake. 1993. Distance Sampling: Estimating Abundance of Biological Populations. Chapman and Hall, London.

Burnham, K.P. 1981. Summarizing remarks: Environmental influences. Pages 324-325 in Estimating numbers of terrestrial birds (C.J. Ralph and J.M. Scott, Eds.). Studies in Avian Biology No. 6.

Burnham, K.P. and D.R. Anderson. 2002. Model selection and multimodel inference: A practical information-theoretic approach. Springer-Verlag, New York, NY.

Burnham. K.P. and W.S. Overton. 1979. Robust estimation of population size when capture probabilities vary among animals. Ecology 62:927-936.

Canterbury, G.E., T.E. Martin, D.R. Petit, L.J. Petit, and D.F. Bradford. 2000. Bird communities and habitat as ecological indicators of forest condition in regional monitoring. Conservation Biology 14:544-558.

Chao, A., S.M. Lee, and S.L. Jeng. 1992. Estimation of population size for capture-recapture data when capture probabilities vary by time and individual animal. Biometrics 48:201-216.

Chapman, D.G., and C.O. Junge. 1956. The estimation of the size of a stratified animal population. Annals of mathematical statistics. 27:375-389.

Farnsworth, G.L., K.H. Pollock, J.D. Nichols, T.R. Simons, J.E. Hines, and J.R. Sauer. 2002. A removal model for estimating detection probabilities from point-count surveys. Auk 119:414-425.

Gill, F.B. 1995. Vocal communication. Pages 233-262 in Ornithology (2nd Ed.). W.H. Freeman and Company, New York.

Johnson, D.H., K.P. Burnham, and J.D. Nichols. 1986. The role of heterogeneity in animal population dynamics. Proceedings of the International Biometrics Conference, 13:1-15.

Lancia, R.A., J.D. Nichols, and K.H. Pollock. 1994. Estimating the number of animals in wildlife populations. Pages 215-253 in Research and management techniques for wildlife and habitats (T. Bookhout, Ed.). The Wildlife Society, Bethesda, Maryland.

Lebreton, J.D., K.P. Burnham, J. Clobert, and D.R. Anderson. 1992. Modeling survival and testing biological hypotheses using marked animals: A unified approach with case studies. Ecological Monographs. 62:67-118.

Lein, M.R. 1981. Display behavior of ovenbirds (*Seiurus aurocapillus*) II. Song variation and singing behavior. Wilson Bulletin. 93:21-41.

Marsh, H. and D.F. Sinclair. 1989. Correcting for visibility bias in strip transect aerial surveys of aquatic fauna. Journal of Wildlife Management. 53:1017-1024.

McShea, W.J. and J.H. Rappole. 1997. Variable song rates in three species of passerines and implications for estimating bird populations. Journal of Field Ornithology. 68:367-375.

Meador, M.R., J.P. McIntyre, and K.H. Pollock. 2003. Assessing the efficacy of single-pass backpack electrofishing to characterize fish community structure. Transactions of the American Fisheries Society 132:39-46.

Mengak, M.T. and D.C. Guynn, Jr. 1987. Pitfalls and snap traps for sampling small mammals and herpetofauna. American Midland Naturalist. 118:284-288.

Morton, E.S. 1975. Ecological sources of selection on avian sounds. American Naturalist. 109:17-34.

Nichols, J.D. 1992. Capture-recapture models: Using marked animals to study population dynamics. BioScience. 42:94-102.

Nichols, J.D., J.E. Hines, J.R. Sauer, F.W. Fallon, J.E. Fallon, and P.J. Heglund. 2000. A double-observer approach for estimating detection probability and abundance from point counts. Auk 117:393-408.

Norris, J.L., and K.H. Pollock. 1996. Nonparametric MLE under two closed capture-recapture models with heterogeneity. Biometrics. 52:639-649.

Nur, N., S.L. Jones, and G.R. Geupel. 1999. A statistical guide to data analysis of avian monitoring programs. U.S. Department of the Interior, Fish and Wildlife Service, BTP-R6001-1999, Washington, D.C.

Otis, D.L., K.P. Burnham, G.C. White, and D.R. Anderson. 1978. Statistical inference from capture data on closed animal populations. Wildlife Monographs, No. 62.

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

Pollock, K.H., J.D. Nichols, C. Brownie, and J.E. Hines. 1990. Statistical inference for capture-recapture experiments. Wildlife Monograph, No. 107.

Ralph, J.C., J.R. Sauer, S. Droege (Eds.). 1995. Monitoring bird populations by point counts. United States Forest Service General Technical Report PSW-GTR-149.

Ramsey, F.L., and J.M. Scott. 1979. Estimating population densities from variable circular plot surveys. Pages 155-181 in Sampling biological populations. Statistical ecology series, vol. 5 (R.M. Cormack, G.P. Patil, and D.S. Robson, Eds.). International Cooperative Publishing House, Fairland, Maryland.

Reynolds, R.T., J.M. Scott, and R.A. Nussbaum. 1980. A variable circular-plot method for estimating bird numbers. Condor 82:309-313.

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

Schwartz, O.A. and P.D. Whitson. 1986. A 12-year study of vegetation and mammal succession on a reconstructed tallgrass prairie in Iowa. American Midland Naturalist 117:240-249.

Seber, G.A.F. 1982. The Estimation of Animal Abundance and Related Parameters (2nd ed.). Edward Arnold, London.

Thomas, L., J.L. Laake, S. Strindberg, F.F.C. Marques, S.T. Buckland, D.L. Borchers, D.R. Anderson, K.P. Burnham, S.L. Hedley, and J.H. Pollard. 2002. Distance 4.0. Release 2. Research Unit for Wildlife Population Assessment, University of St. Andrews, UK. http://www.ruwpa.st-and.ac.uk/distance/

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

Wasserman, F.E. 1977. Mate attraction function of song in the White-throated Sparrow. Condor. 79:125-127.

Webb, W.L. 1965. Small mammal populations on islands. Ecology 46:479-488.

White, G.C., D.R. Anderson, K.P. Burnham, and D.L. Otis. 1982. Capture-recapture removal methods for sampling closed populations. Los Alamos National Laboratory Publication LA-8787-NERP. Los Alamos, NM.

White, G.C. 1983. Numerical estimation of survival rates from band-recovery and biotelemetry data. Journal of Wildlife Management. 47:716-728.

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

Wiley, R.H., and D.G. Richards.  1982.  Adaptations for acoustic communication in birds: Sound transmission and signal detection.  Pages 132-181 in Acoustic communication in birds (Vol. 1): Production, perception and design features of sounds (D.E. Kroodsma, E.H. Miller, and H. Ouellet Eds.).  Academic Press, New York.

Williams, B.K., J.D. Nichols, and M.J. Conroy.  2002.  Analysis and Management of Animal Populations. Academic Press.  San Diego, CA.

Table 1: Number of parameters in candidate models for the time of detection method with t time periods and s different species. Behavior models assume a single behavioral response and heterogeneity models are based on a 2 point mixture. Models with a superscript $^{+spp}$ indicate an additive effect between observers and species and models with a superscript $^{*spp}$ indicate an interaction effect between observers and species. Models with superscript $^{part}$ indicate those that have similar detection probabilities between species but different probabilities of being in the first group. An additional parameter is required for each model to estimate abundance.

| No Species Effect | | Species Effect | |
|---|---|---|---|
| Model | # Parameters | Model | # Parameters |
| $M_0^0$ | 1 | $M_0^{spp}$ | s |
| $M_t^0$ | t | $M_t^{+spp}$ | t+s-1 |
| | | $M_t^{*spp}$ | ts |
| $M_b^0$ | 2 | $M_b^{spp}$ | 2s |
| $M_{tb}^0$ | t+1 | $M_{tb}^{+spp}$ | t+s |
| | | $M_{tb}^{*spp}$ | s(t+1) |
| $M_h^0$ | 3 | $M_h^{part}$ | s+1 |
| | | $M_h^{spp}$ | 3s |
| $M_{bh}^0$ | 5 | $M_{bh}^{part}$ | 2s+1 |
| | | $M_{bh}^{spp}$ | 5s |
| | | $M_{th}^{part}$ | ts+1 |
| $M_{th}^0$ | 2t+1 | $M_{th}^{+spp}$ | 2t+2s-1 |
| | | $M_{th}^{*spp}$ | s(2t+1) |

Table 2: Number of parameters in candidate model for the multiple observer method using t observers, s species and heterogeneity models based on a 2 point mixture model. Models with superscript $^{part}$ indicate those that have similar detection probabilities between species but different probabilities of being in the first group. Models with a superscript $^{+spp}$ indicate an additive effect between observers and species and models with a superscript $^{*spp}$ indicate an interaction effect between observers and species. An additional parameter is required for each model to estimate abundance.

| No Species Effect | | Species Effect | |
|---|---|---|---|
| Model | # Parameters | Model | # Parameters |
| $M_0^0$ | 1 | $M_0^{spp}$ | $s$ |
| $M_{obs}^0$ | $t$ | $M_{obs}^{+spp}$ | $t+s-1$ |
| | | $M_{obs}^{*spp}$ | $ts$ |
| $M_h^0$ | 3 | $M_h^{part}$ | $s+2$ |
| | | $M_h^{spp}$ | $3s$ |
| | | $M_{obs,h}^{part}$ | $s+t$ |
| $M_{obs,h}^0$ | $2t+1$ | $M_{obs,h}^{+spp}$ | $3s+2t-2$ |
| | | $M_{obs,h}^{*spp}$ | $s(t+1)$ |

Table 3: Species groups for example analyses, grouped first into three maximum detection distance categories ($\leq$ 100 m, > 100 m and $\leq$ 150 m, or > 150 m) and then grouped by similarity in singing rates. Maximum detection distance is from the actual data and is truncated by 10% of the largest detection distances. All other categories are averages from rankings on a scale of 1 to 5 from seven experienced birders familiar with the study area. Higher ranks correspond to assumed higher values for each category.

| Group | Species | Maximum Distance | Singing Rate | Sound Intensity | Sound Pitch | Sound Modulation | Plumage Color | Movement |
|---|---|---|---|---|---|---|---|---|
| A | Acadian Flycatcher | 75 | 2.6 | 3.3 | 3.4 | 2.7 | 1.1 | 1.9 |
| | Black-and-white Warbler | 80 | 2.6 | 2.4 | 3.7 | 2.9 | 3.4 | 3.7 |
| | Golden-crowned Kinglet | 75 | 2.6 | 1.1 | 3.9 | 2.7 | 1.7 | 4.4 |
| | Worm-eating Warbler | 80 | 2.0 | 2.9 | 2.4 | 1.4 | 1.3 | 2.7 |
| B | Black-throated Blue Warbler | 80 | 3.6 | 3.1 | 2.9 | 3.1 | 2.9 | 3.3 |
| | Dark-eyed Junco | 100 | 3.7 | 3.1 | 3.0 | 1.7 | 2.3 | 2.9 |
| | Hooded Warbler | 100 | 3.4 | 4.0 | 3.0 | 4.0 | 3.9 | 3.0 |
| | Solitary Vireo | 85 | 3.3 | 3.4 | 3.3 | 3.4 | 2.0 | 2.6 |
| C | Black-throated Green Warbler | 100 | 4.0 | 3.4 | 3.3 | 3.9 | 3.1 | 3.0 |
| | Indigo Bunting | 70 | 4.0 | 3.9 | 3.0 | 3.4 | 3.7 | 2.7 |
| | Red-eyed Vireo | 100 | 4.7 | 3.6 | 3.0 | 2.9 | 1.4 | 2.4 |
| D | Scarlet Tanager | 130 | 3.1 | 3.9 | 2.4 | 2.7 | 5.0 | 2.6 |
| | Veery | 150 | 2.9 | 3.0 | 2.6 | 3.6 | 1.1 | 1.6 |
| E | Ovenbird | 110 | 4.1 | 5.0 | 2.4 | 2.3 | 1.4 | 2.9 |
| | Rufous-sided Towhee | 115 | 3.6 | 4.0 | 3.0 | 3.4 | 2.4 | 2.6 |
| F | Red-breasted Nuthatch | 200 | 2.0 | 2.7 | 3.0 | 1.4 | 2.6 | 4.1 |
| | Tufted Titmouse | 160 | 3.9 | 3.9 | 3.3 | 2.1 | 2.1 | 3.7 |
| | Winter Wren | 200 | 3.9 | 4.0 | 3.9 | 3.9 | 1.3 | 2.1 |
| | Wood Thrush | 200 | 2.9 | 4.3 | 3.0 | 4.3 | 1.9 | 2.1 |

Table 4: delta AICc for distance models using first 3 minute interval of time of detection data.

| Species Group | No Species Effect | Species Effect |
|---|---|---|
| A | 6.01 | 0.00 |
| B | 10.40 | 0.00 |
| C | 6.21 | 0.00 |
| D | 0.00 | 1.76 |
| E | 9.12 | 0.00 |
| F | 5.56 | 0.00 |

Table 5: Distance analysis for first 3 minutes of 10 minute point count. Results given for model with species effect and for model with no species effect. Observed count is after 10% truncation of largest observed detection distances, EDR is the effective detection radius and density is individuals per hectare. Standard errors are in parentheses.

| Group | Species | Obs. Count | Species Effect | | No Species Effect | |
|---|---|---|---|---|---|---|
| | | | EDR | Density | EDR | Density |
| A | Acadian Flycatcher | 58 | 53.9 (6.86) | 0.20 (0.041) | 48.2 (1.22) | 0.25 (0.022) |
| | Black-and-white Warbler | 90 | 43.5 (1.63) | 0.47 (0.072) | | 0.38 (0.035) |
| | Golden-crowned Kinglet | 64 | 47.1 (2.34) | 0.29 (0.061) | | 0.27 (0.025) |
| | Worm-eating Warbler | 56 | 46.9 (3.14) | 0.25 (0.056) | | 0.24 (0.021) |
| B | Black-throated Blue Warbler | 145 | 59.9 (2.16) | 0.40 (0.060) | 59.5 (6.57) | 0.40 (0.094) |
| | Dark-eyed Junco | 109 | 49.2 (2.75) | 0.44 (0.080) | | 0.30 (0.070) |
| | Hooded Warbler | 192 | 64.5 (4.41) | 0.46 (0.078) | | 0.53 (0.124) |
| | Solitary Vireo | 98 | 52.8 (4.88) | 0.35 (0.079) | | 0.27 (0.063) |
| C | Black-throated Green Warbler | 273 | 77.6 (2.98) | 0.45 (0.049) | 77.2 (1.96) | 0.45 (0.036) |
| | Indigo Bunting | 40 | 54.5 (2.03) | 0.13 (0.027) | | 0.07 (0.005) |
| | Red-eyed Vireo | 270 | 78.8 (6.11) | 0.43 (0.075) | | 0.45 (0.036) |
| D | Scarlet Tanager | 114 | 70.6 (1.85) | 0.23 (0.029) | 70.1 (1.51) | 0.23 (0.025) |
| | Veery | 39 | 69.0 (2.54) | 0.08 (0.018) | | 0.08 (0.009) |
| E | Ovenbird | 328 | 72.3 (1.69) | 0.62 (0.061) | 70.6 (1.40) | 0.65 (0.057) |
| | Rufous-sided Towhee | 54 | 54.5 (1.75) | 0.18 (0.037) | | 0.12 (0.009) |
| F | Red-breasted Nuthatch | 30 | 104.9 (15.86) | 0.03 (0.011) | 83.7 (6.22) | 0.04 (0.007) |
| | Tufted Titmouse | 79 | 100.7 (9.11) | 0.08 (0.018) | | 0.11 (0.018) |
| | Winter Wren | 80 | 105.7 (7.58) | 0.07 (0.014) | | 0.11 (0.019) |
| | Wood Thrush | 135 | 73.6 (10.84) | 0.25 (0.077) | | 0.19 (0.031) |

Table 6: $\Delta AIC_c$ for time of detection multiple species models for unlimited radius plots with 10% truncation of largest detection distances. Smaller values of $\Delta AIC_c$ indicate more parsimonious models. $\Delta AIC_c$ weights in parentheses. Larger weights indicate more support for a given model. Models with weights $\geq 0.20$ are in bold for each species indicating competing models. Models were omitted (NA) when parameter estimates were not realistic.

| | Groups | | | | | |
| Model | A | B | C | D | E | F |
|---|---|---|---|---|---|---|
| $M_0^0$ | 55.5 (0.00) | 145 (0.00) | 124 (0.00) | 41.1 (0.00) | 72.1 (0.00) | 180 (0.00) |
| $M_0^{spp}$ | 58.3 (0.00) | 123 (0.00) | 127 (0.00) | 36.9 (0.00) | 73.4 (0.00) | 84.0 (0.00) |
| $M_t^0$ | 43.6 (0.00) | 131 (0.00) | 106 (0.00) | 36.3 (0.00) | 59.8 (0.00) | 143 (0.00) |
| $M_t^{+spp}$ | 36.2 (0.00) | 79.3 (0.00) | 48.4 (0.00) | 36.4 (0.00) | 52.3 (0.00) | 44.4 (0.00) |
| $M_t^{*spp}$ | 51.4 (0.00) | 108 (0.00) | 107 (0.00) | 34.4 (0.00) | 61.2 (0.00) | 48.5 (0.00) |
| $M_b^0$ | 55.9 (0.00) | 147 (0.00) | 123 (0.00) | 43.1 (0.00) | 74.1 (0.00) | 182 (0.00) |
| $M_b^{spp}$ | 62.6 (0.00) | 115 (0.00) | 125 (0.00) | 40.7 (0.00) | 75.6 (0.00) | 92.2 (0.00) |
| $M_{tb}^0$ | NA | NA | NA | NA | NA | NA |
| $M_{tb}^{+spp}$ | NA | NA | NA | NA | NA | NA |
| $M_{tb}^{*spp}$ | NA | NA | NA | NA | NA | NA |
| $M_h^0$ | **0.0 (0.51)** | 20.3 (0.00) | 3.6 (0.04) | 3.2 (0.08) | 5.8 (0.03) | 75.2 (0.00) |
| $M_h^{part}$ | 4.5 (0.05) | 4.3 (0.09) | 5.0 (0.02) | **0.0 (0.40)** | 6.6 (0.02) | 5.9 (0.05) |
| $M_h^{spp}$ | 9.6 (0.00) | 8.4 (0.01) | 7.2 (0.01) | 1.9 (0.16) | 7.7 (0.01) | 8.9 (0.01) |
| $M_{bh}^0$ | 2.0 (0.18) | 22.3 (0.00) | 5.6 (0.02) | 5.2 (0.03) | 7.9 (0.01) | 77.3 (0.00) |
| $M_{bh}^{part}$ | 6.6 (0.02) | 6.3 (0.03) | 7.0 (0.01) | 2.1 (0.14) | 8.6 (0.01) | 8.0 (0.02) |
| $M_{bh}^{spp}$ | 18.1 (0.00) | 15.7 (0.00) | 11.1 (0.00) | 6.1 (0.02) | 11.8 (0.00) | 17.3 (0.00) |
| $M_{th}^0$ | **1.7 (0.21)** | 21.0 (0.00) | **0.0 (0.27)** | 15.1 (0.00) | 22.3 (0.00) | 69.4 (0.00) |
| $M_{th}^{part}$ | 6.4 (0.02) | 5.3 (0.06) | 1.3 (0.14) | 2.0 (0.15) | **0.8 (0.36)** | **0.0 (0.92)** |
| $M_{th}^{+spp}$ | 23.7 (0.00) | 19.2 (0.00) | **0.23 (0.24)** | 16.8 (0.00) | 12.9 (0.00) | 10.8 (0.00) |
| $M_{th}^{*spp}$ | 17.3 (0.00) | **0.0 (0.80)** | **0.14 (0.25)** | 6.5 (0.02) | **0.0 (0.55)** | 10.5 (0.00) |

Table 7: Parameter estimates from the time of detection method for each species. The Probability that an individual is detected at least once during the count $\hat{p}_T$ and the probability of being in the first heterogeneity group $\hat{\lambda}$ and the estimated abundance $\hat{N}$ are given for the selected model and for the selected model from a single species modeling approach. The instantaneous rates formulation was used to estimate detection probabilities. Standard errors are given in parentheses.

| Group | Species | Obs. Count | Selected Model | | | Alternative Single Species Model | | |
|---|---|---|---|---|---|---|---|---|
| | | | $\hat{p}_T$ | $\hat{\lambda}$ | $\hat{N}$ | $\hat{p}_T$ | $\hat{\lambda}$ | $\hat{N}$ |
| A | Acadian Flycatcher | 87 | 0.89 (0.049) | 0.58 (0.047) | 98 (6.4) | 0.91 (0.172) | 0.51 (0.115) | 95 (18.2) |
| | Black-and-white Warbler | 137 | 0.89 (0.049) | 0.58 (0.047) | 154 (9.5) | 0.86 (0.113) | 0.58 (0.078) | 160 (21.7) |
| | Golden-crowned Kinglet | 82 | 0.89 (0.049) | 0.58 (0.047) | 92 (6.1) | 0.92 (0.089) | 0.67 (0.155) | 89 (9.0) |
| | Worm-eating Warbler | 96 | 0.89 (0.049) | 0.58 (0.047) | 108 (6.9) | 0.92 (0.249) | 0.54 (0.098) | 104 (28.5) |
| B | Black-throated Blue Warbler | 197 | 0.90 (0.094) | 0.48 (0.081) | 220 (23.6) | | | |
| | Dark-eyed Junco | 189 | 0.83 (0.092) | 0.72 (0.058) | 227 (26.0) | | Same model | |
| | Hooded Warbler | 274 | 0.84 (0.062) | 0.65 (0.057) | 326 (25.4) | | | |
| | Solitary Vireo | 148 | 0.91 (0.440) | 0.42 (0.060) | 163 (79.1) | | | |

| | Species | N | | | | | | |
|---|---|---|---|---|---|---|---|---|
| C | Black-throated Green Warbler | 377 | 0.89 (0.053) | 0.53 (0.035) | 424 (26.2) | 0.9 (0.066) | 0.55 (0.057) | 419 (31.5) |
| | Indigo Bunting | 64 | 0.89 (0.053) | 0.53 (0.035) | 72 (5.2) | 0.93 (0.505) | 0.55 (0.092) | 69 (37.6) |
| | Red-eyed Vireo | 397 | 0.89 (0.053) | 0.53 (0.035) | 446 (27.5) | 0.87 (0.086) | 0.51 (0.048) | 454 (45.6) |
| D | Scarlet Tanager | 161 | 0.85 (0.067) | 0.52 (0.062) | 189 (16.0) | 0.86 (0.085) | 0.53 (0.068) | 186 (19.0) |
| | Veery | 67 | 0.85 (0.067) | 0.75 (0.088) | 79 (7.2) | 0.82 (0.111) | 0.74 (0.093) | 82 (11.9) |
| E | Ovenbird | 444 | 0.92 (0.053) | 0.51 (0.062) | 483 (28.6) | | | |
| | Rufous-sided Towhee | 79 | 0.89 (0.277) | 0.50 (0.078) | 89 (28.1) | Same model | | |
| F | Red-breasted Nuthatch | 54 | 0.81 (0.062) | 0.79 (0.087) | 67 (6.4) | 0.79 (0.139) | 0.80 (0.099) | 68 (12.7) |
| | Tufted Titmouse | 104 | 0.81 (0.062) | 0.78 (0.068) | 128 (11.2) | 0.77 (0.093) | 0.78 (0.068) | 135 (17.4) |
| | Winter Wren | 106 | 0.81 (0.062) | 0.50 (0.069) | 131 (11.4) | 0.87 (0.102) | 0.56 (0.107) | 121 (14.8) |
| | Wood Thrush | 153 | 0.81 (0.062) | 0.15 (0.047) | 188 (15.9) | 0.85 (0.331) | 0.13 (0.047) | 180 (70.4) |

Table 8: $\Delta AIC_c$ for the four independent observer multiple species models for unlimited radius plots with 10% truncation of largest detection distances. Smaller values of $\Delta AIC_c$ indicate more parsimonious models. $\Delta AIC_c$ weights in parentheses. Larger weights indicate more support for a given model. Models with weights $\geq 0.20$ are in bold for each species indicating competing models. The number of observations for each species was small for this data set so the groups have been modified for analysis.

| Models | Groups | | |
| --- | --- | --- | --- |
| | B | C | DEF |
| $M_0^0$ | 56.33 (0.00) | 103.52 (0.00) | 114.68 (0.00) |
| $M_0^{spp}$ | 55.94 (0.00) | 103.22 (0.00) | 107.66 (0.00) |
| $M_{obs}^0$ | 61.04 (0.00) | 105.35 (0.00) | 116.26 (0.00) |
| $M_{obs}^{+spp}$ | 60.69 (0.00) | 105.07 (0.00) | 109.22 (0.00) |
| $M_{obs}^{*spp}$ | 70.55 (0.00) | 109.37 (0.00) | 111.91 (0.00) |
| $M_h^0$ | **0.00 (0.72)** | **0.76 (0.22)** | **1.18 (0.26)** |
| $M_h^{part}$ | **2.10 (0.25)** | **0.40 (0.26)** | **0.00 (0.47)** |
| $M_h^{spp}$ | 7.73 (0.02) | 4.47 (0.03) | 4.61 (0.05) |
| $M_{obs,h}^0$ | 9.20 (0.01) | **0.00 (0.32)** | 5.91 (0.02) |
| $M_{obs,h}^{part}$ | 11.34 (0.00) | 1.30 (0.17) | 4.29 (0.06) |
| $M_{obs,h}^{+spp}$ | 17.42 (0.00) | * | 6.39 (0.02) |
| $M_{obs,h}^{*spp}$ | 25.54 (0.00) | * | 2.70 (0.12) |

*Models did not give realistic estimates. Standard errors for detection probabilities were much larger than one.

Table 9: Independent Observer results for group B without Black-throated Blue Warbler, group C without Indigo Bunting and a combined group of one species from groups D, E, and F. The probability of being in the low or high detectability groups is given by $\hat{\pi}$ and the probability of detection by one of the 4 observers is given by $\hat{p}_1, \hat{p}_2, \hat{p}_3$, and $\hat{p}_4$ These are reported based on the selected model. The abundance estimate $\hat{N}$ is given for the selected model and for the selected model from a single species analysis.

| Group | Species | Obs. Count | Group Probability and Observer Detection Probabilities | | | | | $\hat{N}_{selected}$ | $\hat{N}_{species}$ |
| | | | $\hat{\pi}$ | $\hat{p}_1$ | $\hat{p}_2$ | $\hat{p}_3$ | $\hat{p}_4$ | | |
|---|---|---|---|---|---|---|---|---|---|
| B | Dark-eyed Junco | 36 | | | | | | 41 (2.9) | 44 (5.7) |
| | Hooded Warbler | 38 | 0.34 (0.065) | | 0.24 (0.083) 0.90 (0.027) | | | 43 (3.0) | 40 (2.0) |
| | Solitary Vireo | 51 | | | | | | 57 (3.7) | 59 (5.9) |
| C | Black-throated Green Warbler | 47 | 0.40 (0.057) | 0.20 (0.091) 0.99 (0.016) | 0.13 (0.070) 0.95 (0.030) | 0.19 (0.072) 0.85 (0.054) | 0.11 (0.051) 0.91 (0.050) | 59 (5.4) | 60 (7.2) |
| | Red-eyed Vireo | 72 | | | | | | 90 (7.6) | 88 (7.9) |
| DEF | Ovenbird | 90 | 0.36 (0.066) | | | | | 96 (3.3) | 96 (3.7) |
| | Scarlet Tanager | 61 | 0.58 (0.094) | | 0.36 (0.062) 0.96 (0.021) | | | 68 (3.6) | 70 (5.6) |
| | Tufted Titmouse | 44 | 0.44 (0.097) | | | | | 48 (2.4) | 48 (3.3) |

**Chapter 5**

**MODELING THE AVAILABILITY PROCESS
FOR POINT COUNT SURVEYS USING AUXILIARY DATA**

# MODELING THE AVAILABILITY PROCESS FOR POINT COUNT SURVEYS USING AUXILIARY DATA

Mathew W. Alldredge, Biomathematics and Zoology, North Carolina State University, Raleigh NC 27695.

Kenneth H. Pollock, Zoology, Biomathematics, and Statistics, North Carolina State University, Raleigh, NC 27695.

Theodore R. Simons, USGS Cooperative Fish and Wildlife Research Unit, Dept. of Zoology, North Carolina State University, Raleigh, NC 27695.

Abstract.—Point count surveys based on aural detections of birds are commonly used to estimate abundance of bird populations, to make inferences about population health, and make spatial and temporal comparisons of populations. In order to make valid comparisons it is necessary to estimate true abundance from count data by correcting for the detection process. The detection process for point count surveys based on aurally detected birds consists of availability (the probability that an individual sings during the count) and perception (the probability an individual is detected given that it sings). Only the time of detection method models both of these components from a single point count survey but this method requires long survey periods which can be biased by movement of birds. Short duration "snapshot" type methods (e.g. variable circular plot or distance sampling and multiple observer counts) are not as subject to bias associated with birds moving during the count but these "snapshot" methods do not account for availability. Because availability bias increases as count length decreases it is important to include this component when modeling the detection process. We propose two methods of modeling the availability process; one that assumes an individual bird sings randomly following a Poisson process and another nonparametric approach that makes no

assumptions about the distribution of singing times for an individual bird. Data collection for both approaches may be "size" biased or over-represent birds with high singing rates, which must be corrected for when estimating the availability probability. For the first approach we develop likelihood based models to estimate the availability probability for homogeneous singing rates and finite-mixture Poisson models to account for individual heterogeneity in singing rates. The nonparametric approach to estimating the availability probability is based on re-sampling singing time data over the time interval used for a point count survey and determining whether an individual sings during the interval, an approach similar to doing a bootstrap. This approach does require that the true distribution of singing rates can be obtained. We applied these approaches to singing rate data collected for the Ovenbird (*Seiurus aurocpillus*) in Great Smoky Mountains National Park, and to simulated data that were comparable to field data. The homogeneous Poisson model did not fit the data well but a two-point Poisson mixture model did. The estimated availability probability for a three minute point count from the two-point Poisson mixture model was 0.87 (SE = 0.025), indicating that 13% of the population did not sing during the count and could not be detected. Analysis of the simulated singing time data indicates that if individuals sing following a random process, results are comparable to the singing rate approach, but if individuals sing in bouts, the availability probability will decrease or a smaller portion of the population actually sings during a point count survey. The benefit of these methods is that they can be used to correct for availability bias in conjunction with any point count method, including those that use very short survey intervals. This approach does require a significant amount of

field effort and may be most applicable to surveys intended to track spatial or temporal changes in the abundance of a few target species.

## Introduction

The importance of accurately modeling the detection process when estimating animal abundance has long been recognized (Burnham 1981, Barker and Sauer 1995, Johnson 1995, Williams et al. 2002). In spite of this, point count surveys for birds are continually used as indices or measures of relative bird abundance rather than measures of true abundance (Rosenstock et al. 2002). Recently methods for estimating the detection process for point count surveys have been emphasized, and new methods have been developed (Nichols et al. 2000, Bart and Earnst 2002, Farnsworth et al. 2002, Thompson 2002, Rosenstock et al. 2002, Royle and Nichols 2003).

When the detection process is not measured, abundance estimates are subject to both availability and perception biases (Marsh and Sinclair 1989). Given these two components of the detection process the conceptual estimator of abundance is

$$\hat{N} = \frac{C}{\hat{p}_d \hat{p}_a} \tag{1}$$

where C is the observed count, $\hat{p}_d$ is the probability of detection given availability and $\hat{p}_a$ is the availability probability. These parameters can be estimated jointly, as in some current point count methods, or separately. Estimating these parameters separately gives more flexibility in how the data are collected and the methods used for analysis.

Perception is the probability that an individual is detected given that it is available for detection. Both the distance sampling or variable circular plot method (Reynolds et al. 1980, Buckland et al. 1993) and the multiple observer methods (Nichols et al. 2000,

chapter 2) model the probability of detection given availability but do not account for the probability that an individual is available. These are "snapshot" type counts done over short time periods to avoid confounding effects associated with movement of birds during the count. Unless auxiliary information is collected to determine availability or it is reasonable to assume the probability of being available is one, these methods only estimate the available portion of the population,.

Availability refers to the portion of the population that has a nonzero probability of detection. This is similar to the problem of temporary emigration in capture-recapture studies where some portion of the population is in the sample area and can be observed (available) and others are not. Estimation of temporary emigration uses the robust capture-recapture design (Kendall et al. 1997) and was utilized for salamander surveys where individuals are unavailable for capture when below the surface (Bailey et al. 2004a).

Non-availability of birds in avian point counts is caused by birds not singing during a point count or birds moving into or out of the sample area during the count. The probability that a bird is available for detection during point counts based on auditory detections is the probability that an individual sings at least once during the count. Standard point count methods try to account for this by conducting counts early in the morning when singing rates are highest (Ralph et al. 1995). Some counts, such as the BBS (Sauer et al. 1997), also specify surveying early in the breeding season when singing rates are assumed highest. It is doubtful, however, that such standardization validates the assumption that all birds in the population sampled are available during a count. This is especially true for species with low singing rates.

The availability process is especially important in forested environments or dense vegetation (Wilson and Bart 1985). When visual identification is limited, it is common to detect most individuals aurally from calls and songs. Aural detections were made for over 95% of birds detected in point count surveys of breeding birds in Great Smoky Mountains National Park (Simons unpub. data). Therefore, the availability process in these situations is directly related to the singing rates of the individuals encountered. The availability process is of greatest importance for species with low singing rates relative to the duration of the point count.

Current point count methods that incorporate both the availability process and the detection process given availability are the time of detection method (Farnsworth et al. 2002, chapter 3) and repeated count methods (Bart and Earnst 2002, Royle and Nichols 2003, Royle 2004). All of these approaches estimate the joint probability of being available and detection given availability. Availability using the repeated counts method is related to both the probability of singing during a count and movement, because the population within the sample area can change between successive counts.

The time of detection method divides a count into consecutive time intervals and records all of the intervals in which an individual bird is detected. By constructing this detection history, it is possible to use the entire suite of capture-recapture models (Otis et al. 1978, Williams et al. 2002) to analyze these data (Chapter 3). A conceptual example of this method for an individual that is missed in one interval but then detected in the next interval will illustrate how availability is accounted for by the time of detection approach. Two possibilities exist for why an individual is not detected in the first interval; one is that the bird sang but the observer did not detect it, and the other is that the bird never

sang so the observer could not detect it. This method estimates the total detection process because both possibilities of not detecting a bird are represented.

The total survey length of each point count must be long (e.g. 10 minutes) to get reasonable parameter estimates when using the time of detection method and therefore, this method is subject to biases associated with movement. If movement occurs during the sample period, then the "true" population of birds being sampled or the actual sample area cannot be determined. Avoiding problems with movement requires the use of shorter counts, which may suggest the use of distance or multiple observer methods for estimating the probability of detection. Unfortunately no short count duration or "snapshot" type count approach accounts for availability in the detection process.

The objective of this paper is to propose new methods of estimating the probability that a bird sings during a point count. These methods are based on singing data, collected separately from the point count that are then applied to any point count method to correct for availability bias. One approach is to collect singing rate data and model the availability process with a Poisson model which assumes a bird sings at random times. The other approach is to record singing time data and estimate the availability probability by re-sampling the data, which does not require any assumptions about the distribution of songs beyond the data being representative of the population. All models are based on estimating the probability that a randomly selected individual from the population does not sing during a point count interval, the complement of which estimates the availability probability.

We will first discuss the field methods to collect data for either the singing rate approach or the singing time approach. A homogeneous Poisson singing rate model is

then presented as a simple case for modeling the availability process. Finite-mixture Poisson models are then presented as more biologically reasonable models with differences in singing rates associated with factors such as breeding phenology of an individual. In this section we also discuss "size" bias (the tendency of over representing birds with high singing rates in the sample) that may occur in the data and correct for this in these models. The final model is a nonparametric approach that uses the actual singing times of an individual to estimate the availability probability and variances by re-sampling the singing time data. We present an example of the singing rate models using data collected on the Ovenbird and select between models using information theoretic model selection. Examples for the singing time analysis method are given using simulated data that are comparable to the Ovenbird data set. Two data sets are simulated; one that has random singing times and one that assumes birds sing in bouts of five songs each. We conclude with a discussion of this approach and compare this method with the time of detection method of accounting for availability bias.

## Methods

Field methods.—We assume that data collected to model the availability process for point counts are collected at the same times and locations as the point count data. This is important to ensure that the availability process being modeled from auxiliary singing data are representative of the population of birds sampled with point counts. These data could be collected at each point following the conclusion of a point count or during travel to subsequent point counts. If point counts are stratified by habitat type then singing data must be collected and analyzed for these same strata. Singing rates or

times should be recorded over long time periods to accurately represent the singing

distribution for an individual bird (e.g. 10 to 20 minutes).

Ideally birds are sampled by visually locating birds in the absence of auditory

cues (possibly wearing headphones while locating birds).  This will avoid issues of "size"

bias in the sample or over-representing birds with high singing rates.  Recording singing

rates or singing times can begin immediately after a bird is identified visually.

In some situations and for some species it is only feasible to locate birds by

sound.  In these cases it is important to ignore the song or singing bout that was used to

locate the bird.  If this is not done then the singing rate of each individual is also biased

because the probability of a bird never singing during an observation period is zero.

Waiting a set period of time (e.g. one minute) after locating a bird is probably the best

approach to sampling when birds are identified by song.

Collecting data on singing rates only requires recording the number of songs or

calls given by an individual bird during the observation interval.  Some effort is

necessary to ensure that a bird is present during the observation interval and has not

flown off.  Singing time data require recording the exact start and stop times of each song

during an observation period.  This can be done with either a stop watch or by using a

personal digital assistant (PDA) (we have developed a program to do this).

Availability assuming homogeneous singing rates.—Although we doubt that

average singing rates across the population of interest are homogeneous or identical

among individuals, this provides a reasonable starting place.  The assumption for this

model is that all birds in the population of interest have the same singing rate parameter

($\lambda$) so that on average every bird calls the same number of times during a time interval.  If

we let the random variable X represent the number of times an individual sings during a time interval (t) and assume that singing rates are from a Poisson process then X is distributed Poisson($\lambda$t). From a sample of n individuals we would have $x_1, x_2, \ldots, x_n$ realizations of the random variable X.

The likelihood for the set of n observations $x_i$ (i=1,…,n) under the model assuming homogeneous singing rates from a Poisson process is:

$$L(\lambda|x_i,t)=\prod_{i=1}^{n}(\lambda t)^{x_i}\exp(-\lambda t) \qquad\qquad (2)$$

Maximizing this likelihood or minimizing the negative natural log of the likelihood will give the estimated Poisson parameter $\hat{\lambda}$ given the data. This is equivalent to the moment estimator for $\hat{\lambda}$ or the mean of the observed singing rates. All models of singing rate data are likelihood based and therefore, can be compared using information-theoretic approaches, such as Akaike's Information Criterion (Burnham and Anderson 2002). The AIC value for this homogeneous one parameter model is:

$$AIC=-2\ln(L(\lambda|x,t))+2 \qquad\qquad (3)$$

Modeling the singing rate parameter as a function of time ($\lambda$t) is important because it allows us to estimate the probability of availability for point counts of different length. The probability of being available for detection ($\hat{p}_a$) is the probability that an individual sings at least once during the count, $p(X \geq 0)$. The probability that an individual sings at least once during a count of duration t* is estimated as the complement of the probability that a bird never sings during a count such that:

$$\hat{p}_a=1-p(X=0)=1-\exp(-\hat{\lambda}t^*) \qquad\qquad (4)$$

Availability assuming heterogeneous singing rates.—When the singing rate

parameter is not homogeneous across the entire population then singing rate data are

biased towards those individuals with higher singing rates. This is because individuals

with higher singing rates are more likely to be detected than individuals with lower

singing rates. This is known as a "size" biased sample (Patil and Ord 1976, Patil and Rao

1977, 1978) and results in a weighted distribution. This results in a positive bias or

overestimation of the availability probability.

When individual heterogeneity exists then the actual data being collected are for

the "size" biased distribution $f^w(x)$ and not the true distribution $f(x)$. The relationship

between these two distributions is (Patil and Rao 1977, 1978):

$$f^w(x|\theta) = \frac{w(x)f(x|\theta)}{w} \tag{5}$$

where $w(x)$ is the weighting function and $w$ is a constant given by:

$$w = \int_0^\infty w(x)f(x|\theta)dx \tag{6}$$

for X continuous or

$$w = \sum_{\forall x} w(x)p(x|\theta) \tag{7}$$

for X discrete. The constant $w$ scales the cumulative distribution for the "size" biased

data $F^w(x)$ to sum to one.

The weighting function $w(x)$ is any function of the random variable that gives the

same chance of including any observation produced by the original distribution $f(x)$. It is

reasonable to assume that including a bird in a sample is a function of singing rate. We

will let the weighting function $w(x)=x$, which gives a linear relationship such that the

probability of including a bird with a singing rate twice that of another is twice that of the

bird with the lower singing rate. Patil and Rao (1977) found that this weighting function was the most commonly used relationship, including wildlife surveys where group size created a "size" bias.

Observations based on auditory cues often results in "size" biased data. Therefore, we will present individual heterogeneity models of singing rate that incorporate the correction for size bias. When data are collected so the data are not size biased (e.g. visually locating birds or using a marked population of birds) or the true singing rate distribution is represented by the data, then letting $w(x)=1$ for all singing rates will give the appropriate model.

If a finite set of biological factors determine singing rates, then finite mixture models (Lindsay 1983, 1986, Norris and Pollock 1996, Pledger 2000) of individual heterogeneity are reasonable. For example, if singing rates are determined by pairing status and nesting stage (Wilson and Bart 1985), then a finite set of factors affecting singing rate is identifiable. Essentially this involves dividing the population into groups based on an individual's current breeding status and claiming that all individuals within a group have on average the same singing rates. If a set of k groups are identified then this would require estimating singing rate parameters $\lambda_1, \lambda_2, \ldots, \lambda_k$, plus the proportion of the total population in each group $\delta_1, \delta_2, \ldots, \delta_{k-1}, 1 - \sum_1^{k-1} \delta_j$. Using equation 5 and the assumption that individual singing rates follow a Poisson process, the likelihood for a set of n observations $x_i$ (i=1,…,n) under the finite-mixture model of singing rates assuming a k-point mixture of Poisson processes is:

$$L(\lambda_1,...,\lambda_k,\delta_1,...,\delta_{k-1}|x,t)= \tag{8}$$

$$\prod_{i=1}^{n}\frac{1}{wx_i!}\left(\sum_{j=1}^{k-1}\delta_j x_i(\lambda_j t)^{x_i}\exp(-\lambda_j t)+\left(1-\sum_{j=1}^{k-1}\delta_j\right)x_i(\lambda_k t)^{x_i}\exp(-\lambda_k t)\right)$$

Because the sum of the proportions across all groups must be one, the final term in this probability represents the final group and it is equal to one minus the sum of the proportions of the population that are in all the other groups.

Maximizing the likelihood or minimizing the negative natural log of the likelihood of the observed singing rates gives the maximum likelihood estimates (MLE's) for $\lambda_1$, $\lambda_2$, …, $\lambda_k$ and the proportion of the population in each group $\delta_1$, $\delta_2$, …, $\delta_{k-1}$. The probability of being available for detection during a point count of length t* is given by

$$\hat{p}_a=1-p(X=x)=\left(1-\sum_{j=1}^{k-1}\delta_j\exp(-\lambda_j t^*)-\left(1-\sum_{j=1}^{k-1}\delta_j\right)\exp(-\lambda_k t)\right) \tag{9}$$

The AIC value is similar to the homogeneous AIC with a larger "penalty" for the number of parameters. For a k-point mixture model there are 2k-1 estimated parameters, which leads to the following equation for AIC:

$$AIC=-2ln(L(\lambda_1,...\lambda_k,\delta_1,...\delta_{k-1}|x,t))+4k-2 \tag{10}$$

Estimating variance and confidence intervals.—The estimated availability probability is a function of random variables ($\hat{\lambda}_j$'s and $\hat{\delta}_j$'s) that each have an associated estimation variance. Because of this an exact derivation of variance can be difficult to obtain (Williams et al. 2002), especially a general form for the heterogeneous singing rate models when the number of mixtures varies. One approach is to use the delta method (Seber 1982) to obtain the large sample approximation of the variance. We

used the bootstrap (Efron 1979, Efron and Gong 1983, Manly 1998) to obtain empirical estimates of variance and distribution free confidence intervals.

The bootstrap procedure treats the observed data of n observations as the population and generates k new samples by selecting n observations from the data with replacement. Model parameters are then estimated for each of the k samples. The mean and variance of each estimate are estimated by the mean and variance of the set of bootstrap estimates. Confidence intervals are constructed assuming that the estimates are distributed normally and using the mean and variance to compute symmetric normal theory confidence intervals. An alterative method uses the percentile approach which is distribution free (Manly 1998). This approach involves ordering the k estimates for each parameter and then selecting the endpoints of the confidence intervals as the values that contain the central $(1-\alpha)100\%$ of the estimates.

It is also possible to use the bootstrap samples to calculate the variance of the estimate of the availability probability, instead of deriving variance estimators based on the variances of the parameter estimates. The probability of being available during a count of length t is calculated (equations 4 or 9) for each of the k bootstrap samples using the parameter estimates from that sample. The variance of this estimate across all bootstrap samples is then used as the estimate of the variance for $\hat{p}_a$. Confidence intervals are then constructed as discussed above.

Availability using singing times.—This nonparametric approach to estimating the availability probability for detection during a point count uses the exact singing times of individual birds and is not restricted by the assumption that the singing rate of an individual bird is based on a Poisson process. This approach is much more robust to

differences among individual birds and it is applicable in situations where birds sing in

singing bouts separated by periods of silence. This singing pattern clearly violates the

Poisson assumption. In many ways this approach is similar to the bootstrap just

discussed. Use of this approach requires that singing times are resampled from the

observed data based on the true distribution of singing rates.

As previously discussed, over-representing birds with high singing rates will lead

to a "size" biased sample. When sampling periods are sufficiently long so that all birds

in the sample sing at least once this bias can be corrected without assuming any

distribution of singing rates for the population. The data are corrected for "size" bias

based on the observed singing rates for each individual. Correcting the biased data is

done by rearranging equation 5 and using the probability mass function for the corrected

distribution of X, $p(X=x)$ and the biased distribution of X, $p^w(X=x)$ as:

$$p(X = x) = \frac{wp^w(X = x)}{w(x)} \qquad (11)$$

To solve for the weight, w, we use the fact that the sum of $p(X=x)$ over all observable

singing rates, r, must equal one, such that:

$$w = \frac{1}{\displaystyle\sum_{\forall r} \frac{p^w(X = r)}{w(r)}} \qquad (12)$$

If not all of the birds in the sample actually sing during the sampling period then

the above formulas to correct for "size" bias do not work. One option is to sample birds

from the population differently so that an unbiased sample is obtained, possibly using

marked birds. If this is not possible then it will probably be necessary to assume a

distributional form for the singing rates of the population (e.g. a Poisson distribution) or use the singing rate method.

We assume that a sample of singing times from the population is representative and treat it as the "population" of possible singing behaviors. In other words the sample of singing times is assumed to represent what is observed during a point count, including birds that do not sing during a specified interval, those that sing frequently, and those that sing sporadically or in bouts. A single bird in the sample can provide much information about what may occur in a point count. For example, if a bird is observed for 20 minutes then examining any smaller portion of this may be a period when a bird does not sing, sings constantly, or sings sporadically. This assumption implies that the singing rates involved with this sample represent the true distribution of singing rates.

The number of observations to include in each sample replicate must be equal to the population size estimate obtained from the point count survey. Sampling the data set an infinite number of times will give a good estimate of the average availability probability for the population but does not provide any information about the variance of the estimate in relation to correcting for availability bias in a point count survey. Because we are assuming that singing time data are representative of the population, then the number of observations to include for each sample should correspond to the abundance of birds estimated in the point count survey area. This estimated abundance in the survey area corresponds to the portion of the population that is actually detected. In other words each replicate sample is a potential realization of the availability probability for a randomly selected sample of size N.

To estimate the abundance in the sample area it is necessary to correct the observed count for both availability and detection bias. Correction for detection bias is based on the type of point count survey. To correct for availability bias the singing time data is sampled sufficiently to obtain the average availability probability using the algorithm described below. The variance of the availability probability is then estimated by generating samples based on the estimated abundance.

Having estimated the abundance in the survey area it is now possible to estimate the availability probability and its variance. Again, note that using the estimated abundance does not affect the average availability probability but it does affect the variance of this estimate. The following algorithm is used to generate the sample and estimate the availability probability of each replicate (m) based on a point count survey length of t* minutes from singing time data collected for t minutes on each individual.

1. Generate each observation, $y_i$ (i = 1,…,N), in the sample.

    a. Randomly select a singing rate (x) for observation i based on the probability mass function for $p(X=x)$.

    b. Randomly select an individual bird, i, (with replacement) with the singing rate x.

    c. Randomly select a starting time, $t_s$ from 0 to t.

    d. Search the singing time data for bird i from $t_s$ to $t_s+t*$ and record if any part of a song is within this interval. This is done by using starting and ending times for each song.

e. If the specified bird i sings during the interval then record a one, otherwise record a zero. Let $y_i$ be the event that bird i sings, such that if the bird sings $y_i = 1$, otherwise $y_i = 0$.

2. Repeat steps a through e for each observation in the sample of size N.

3. Estimate the availability probability for the current sample (replicate j) $\hat{p}_{aj}$ as,

$$\hat{p}_{aj} = \frac{\sum_{1}^{N} y_i}{N} \qquad (13)$$

4. Repeat steps 1 through 3 to obtain m replicate estimates of $\hat{p}_{aj}$.

The average availability probability is the sum of the availability probability estimates $\hat{p}_{aj}$ divided by the number of replicates m, and the variance of the sample measurements $\hat{p}_{aj}$, is the sum of the square difference between the estimates and their mean, divided by m-1. Confidence intervals are generated by assuming that $\hat{p}_{aj}$ is normally distributed or by using the percentile method, described for the bootstrap.

**Analysis of Field Data**

To demonstrate the use of auxiliary information to estimate the availability probability we use singing rate data collected from 269 male Ovenbirds in Great Smoky Mountains National Park during May of 2000 and 2001 (A. Podolsky unpub. data). Birds included in the sample were located from auditory cues and the number of songs sung by each bird were counted over a five minute interval. Observations were made in the first few hours of the morning during the peak of daily singing activity.

These data were used to estimate the availability probability using the homogeneous singing rate model and two- and three-point Poisson mixture models of heterogeneous singing rates. All estimates are maximum likelihood estimates (MLE's) and model comparison is based on AIC (Burnham and Anderson 2002). For each model 1,000 bootstraps were run to obtain a variance estimate and percentile confidence intervals for all parameters and estimates of the availability probability. The availability probability was estimated for one, two, and three minute point counts to demonstrate the effects of the time interval used for point counts on availability.

We simulated comparable data to the Ovenbird data to demonstrate the singing time method. This was possible because all birds in our sample sang at least once during the sample period. Simulations assumed that songs were five seconds in length. Two different data sets were simulated; one simulation assumed individuals sang at random times and the other assumed individuals sang in singing bouts. The data set simulated with random singing times meets the assumptions of the singing rate models, and should have similar availability probabilities. The other data set was simulated assuming that birds sing in bouts of five songs per bout. For birds with singing rates of five or less then that individual would have a single singing bout corresponding to the specified singing rate. Birds with singing rates greater than five would have multiple singing bouts but the total number of songs would correspond to the specified singing rate. Songs within a singing bout were simulated with no intersong interval. Multiple singing bouts were simulated with random starting times for non-overlapping singing bouts. We expect lower availability probability when birds sing in bouts because there are larger time periods when birds do not sing.

**Results**

Singing rates for the Ovenbird data ranged from one to 23 songs during a five minute interval. Plotting the frequency or probability of observing each singing rate shows a bimodal distribution of the data (Figure 1). The average singing rate was 8.56 songs per five minute interval or 1.71 songs per minute.

Singing rate models.—Both the homogeneous Poisson and the two-point Poisson mixture models gave maximum likelihood estimates (MLE's) for the model parameters. The three-point Poisson mixture model was problematic due to local minima, which made finding the MLE's for the model parameters difficult. Bootstrap estimates for the three-point Poisson model suffered similar convergence problems giving unrealistic parameter estimates. Because of these unrealistic parameter estimates we will not present the results of this model. The two-point mixture Poisson model was the most parsimonious model for these data based on AIC values (Table 1). The difference in AIC for the homogeneous Poisson model was large ($\Delta$AIC = 182.7), indicating it is not a reasonable alternative model for these data.

The homogeneous Poisson model did not fit the data well (Pearson $\chi^2$ = 719, df =22, P < 0.0001), but the two-point mixture Poisson model fit reasonably well (Pearson $\chi^2$ = 29.6, df =20, P = 0.08)(Figure 1). The estimate of lambda for the homogeneous Poisson model was 1.71, which corresponds to the average per minute singing rate. The parameter estimates for the two-point Poisson model indicate about 66% of the birds were in a low singing rate group and the remainder sung at almost four times this rate (Table 1). Parameter estimates for both the homogeneous and two-point Poisson mixture

models were reasonably precise based on the bootstrap variance estimates (Table 1). The

bootstrap distribution of $\hat{\lambda}$ for the homogeneous Poisson model was approximately

normal (Figure 2) as were the distributions of $\hat{\lambda}_1, \hat{\lambda}_2$ and $\hat{\delta}$ for the two-point Poisson

model (Figure 3). The 95% confidence intervals using the percentile approach were

nearly symmetric for parameter estimates from both the homogeneous model and the

two-point Poisson mixture model.

Estimates of the availability probability were 0.82 (SE = 0.009) for a one minute

count, 0.97 (SE = 0.003) for a two minute count, and 0.99 (SE = 0.001) for a three minute

count based on the homogeneous Poisson model. Availability probability estimates for

the two-point Poisson mixture model were 0.57 (SE = 0.029) for a one minute count, 0.77

(SE = 0.030) for a two minute count, and 0.87 (SE = 0.025) for a three minute count

(Table 1).

Singing time model.—As expected the estimates of the availability probability

using the singing time approach were higher for the random singing data set than for the

singing bout data set (Table 2). All estimates were reasonably precise using a resample

size of 269 individuals. The difference between the availability probabilities for one and

two minute point counts was about 0.1 but it was only 0.06 for the three minute count.

This occurs as the length of the point count interval approaches the length of the singing

time interval.

Estimates of the availability probability for the random singing time data set

(Table 2) were similar to the estimates obtained from the two-point Poisson mixture

model (Table 1). This is expected because the Poisson model assumes random singing

times. This was just one realization of a random process, so we would not expect exact correspondence between the two models.

The distribution of availability probability estimates is slightly skewed to the right for both the random singing time data and the singing bout data (Figures 4 and 5). This can also be seen in the confidence intervals that are slightly larger to the left of the mean for most estimates (Table 2). We recommend the percentile confidence intervals in this case. The overall shapes of the distributions are similar between the random singing time analyses and the singing bout analyses.

## Discussion

We have presented a method for accounting for availability bias in point count surveys where individual birds are detected by song or call. This method allows for correction of availability bias using auxiliary data, so it is applicable to any point count method, even those that only correct for perception bias. We have provided two methods of analysis, both of which incorporate individual heterogeneity in singing behaviors. One method relies on the assumption that an individual bird's singing times are randomly distributed following a Poisson process. Individual heterogeneity in singing rates is incorporated into this model by using a mixture of Poisson distributions. The other approach is a nonparametric approach that does not require any assumptions about the structure of an individual bird's singing patterns. The approach is based on re-sampling observed singing time data in relation to the time interval and abundance estimate from a point count survey. It incorporates individual heterogeneity in singing behavior through the re-sampling process.

The field methods for these approaches are relatively straightforward but may lead to a "size" biased sample. We have presented methods to account for this bias or tendency to over-represent individuals with high singing rates in the sample. Correcting for "size" bias presents additional problems for the singing time approach if individuals in the sample do not sing at least once during the observation period. If this occurs it may be necessary to assume a distribution for singing rates or use the singing rate method. Alternatively unbiased data could be collected by using birds identified by some other means. If birds are found by visual detections or by randomly sampling a marked population, then "size" bias is not a problem.

One concern with these approaches is censoring data when birds move undetected during the sample period. If this occurs no songs will be recorded for the end of the observation period. Unfortunately, in this situation it is not know if the bird is there and not singing or has left the area. It is necessary to monitor birds visually or at close distances to avoid these problems. Actually following birds as they move during the sample period may be necessary for some species. If territory size is small for the species of interest and movement distances are correspondingly small then this will be realistic. However, if movement of birds is excessive and birds can make undetected movements out of hearing range during the sample period, these methods will not work. In other words, a key assumption is that the amount of time that a bird is monitored is known without error. Radio-telemetry would be useful in testing this assumption.

For this paper the probability of being available for detection is related to whether an individual actually sings during a point count survey. Point count data are primarily collected for singing males. In fact, it may be advisable to ignore other types of

detections because they will have different detection probabilities.  If females or non-breeding males also sing regularly and are likely to be included in the point count survey then auxiliary singing data must also include these birds.

Singing rate data must be collected at times comparable to the point count surveys.  Factors such as the time of day (Robbins 1981a, Skirvin 1981), weather conditions (Robbins 1981b), season (Best 1981, Skirvin 1981), breeding stage (Best 1981, Wilson and Bart 1985), local abundance of conspecifics (McShea and Rappole 1997) and local habitat characteristics (McShea and Rappole 1997) have been documented to affect singing rates of birds.  Because singing rates are so dependent on these factors it is critical that singing rate data be collected across the same area as the point count surveys, on the same days and over the same time period of the day.  If this is not done then the estimated probability of singing during a point count will be biased.  For example, collecting singing data in late morning following the conclusion of point counts would lead to a negative bias in the availability probability if singing rates decline later in the morning.  This would then have a positive bias on the abundance estimate.  The best approach is to collect singing information using an additional observer, following the conclusion of each point count or during travel between points to ensure that the singing information is representative of the availability process involved with the detection process in the point count.

Evidence that singing rates change during the breeding season highlights the importance of correcting point count surveys for availability bias.  Wilson and Bart (1985) report that the probability that a House Wren (*Trogolodytes aedon*) sings at least

once during a three-minute period was 0.7 before mating, 0.5-0.6 from mating to completion of egg-laying, 0.7 during incubation, and 0.5 or less following incubation.

Observation periods should be representative of the singing behavior of the individual sampled. These observation periods should be longer than the length of the point count survey that is being conducted, possibly more than twice as long. If birds sing frequently during the entire period in which point counts are conducted and do not have long intersong intervals then the duration of the sample can be shorter. As the length of intersong intervals increases or becomes more variable then the duration of the sample will also have to increase to accurately represent the singing characteristics of the individual.

Collecting these data will require a large amount of time and may lead to attempts to collect data on multiple birds at the same time. This should not be done for conspecifics because they may counter sing and thus, data would not be independent. This could be done for different species at the same time if movement of birds is minimal. To collect this information may require an observer to follow a bird and this may limit collection on multiple birds at one time.

The individual bird must be considered as well as the entire population of birds to model the availability process using singing rate data. The individual must be considered because observations during a point count are made at the individual level, which is clear based on defining $p_a$ as the probability that an individual will sing during a count. We make the assumption that the singing rate of an individual (i) follows a Poisson process, with singing rate parameter $\lambda_i$. A Poisson process is a memoryless process (Tuckwell 1995), which means that the probability that a bird sings at some future time does not

depend on the last time that it sang. In other words this assumes singing times are completely random events. The average time between calls is given by the rate parameter $\lambda_i$, which is the expected value of the Poisson process.

Because it is not possible to know the singing rate parameters of the individuals detected during a point count, the probability of being available must be based on the average availability of the population. We assume birds occurring in the sample area of a point count survey represent a random sample from the population. In order to determine the average probability of availability it is necessary to make assumptions about the distribution of the singing rate parameter across the population. The simplest assumption is that all individuals in the population have the same singing rate parameter (i.e. the population is homogeneous with regard to singing rate). This assumption could be true during certain periods of the breeding season (e.g. if all individuals have the same pairing status or are in the same reproductive stage) but in general it is too restrictive for typical applications.

It is more likely that there is individual heterogeneity in singing rates among individuals of a population. Pairing status and breeding stage are major components that effect singing behavior of males during the breeding season (Best 1981, Skirvin 1981, Wilson and Bart 1985). Singing rates of unpaired territorial males are often higher because these individuals are trying to attract mates (Best 1981). Singing rates generally decline during the breeding season as males that are incubating or rearing broods sing less (Wilson and Bart 1985).

If singing rates are determined primarily by pairing status and breeding stage then it is reasonable to model individual heterogeneity with a finite mixture model (Lindsay

1983, 1986), similar to the finite mixture models of heterogeneity for capture-recapture data (Norris and Pollock 1996, Pledger 2000). If heterogeneity in singing rates is associated with innate differences among birds, then biological interpretations of the mixture groups is not possible.

The probability that an individual is available for detection during a count and the probability of detection given availability are considered a nuisance parameter, because these parameters only provide information used to estimate population abundance. Because these probabilities are not of use to estimate abundance it is not necessary to estimate the availability process separately if the time of detection method works reasonably well. Unfortunately a second component of the availability process is movement of individuals during the count. The time of detection method is biased if movement that affects detectability occurs during the count.

Three types of movement can occur during a point count: movement in relation to the observer at the point, small scale movements within the sample area, and larger scale movements into and out of the sample area. Movement in relation to the observer is a problem inherent with point counts and it can bias the count (Conant et al. 1981, Buckland et al. 1993). If movement removes birds from the sample area, then observed counts will always be too small. Alternatively, some species of birds may be attracted to observers, in which case the counts will be too large. Minor responses by birds to an observer will have less effect on the count but would bias distance estimates if this is being used in the model of detection. One approach to account for observer related movements is to wait at the point for a specified amount of time before beginning the count, but the effectiveness of this has never been tested.

Small scale random movements within the sample area are of less concern unless they have a significant effect on the detection process. This type of movement may affect detection distance estimates, but random movement should average out across the sample of observed birds. Small scale movements may also result in double counting. This occurs when a bird is counted, then moves and is recounted as a new individual. The result is a positive bias in the count.

Large scale movements in and out of the count are also important. These movements create a positive bias in the count because not all birds are in the sample area at the same time. This makes inference about the population difficult because the sample area or proportion of the population sampled cannot be determined (Scott and Ramsey 1981). These movements also affect the availability process when the time of detection method is used (Chapter 3). Now a zero in the detection history could also represent an individual that was not present in the sample area during that interval. The probability of availability due to singing is confounded by movement. The best approach for avoiding the confounding effects of movement is to shorten the count duration (Scott and Ramsey 1981, Buckland et al. 1993).

Shortening the duration of a point count minimizes movement problems (Scott and Ramsey 1981), including those related to double counting, but increases the importance of the availability component of the detection process. As the duration of the count gets shorter the probability that an individual actually sings during the count also gets smaller. The time of detection method does not work well with very short duration counts. "Snapshot" type counts (e.g. distance sampling and multiple observer methods)

minimize problems associated with movement. Availability probabilities on snapshot counts must be estimated using auxiliary information.

The singing time model requires more detailed data and analysis is more complex because it requires writing a computer program to resample the data. The overall field effort between the two methods is the same. We have developed a computer program for collecting this data which allows the user to input relevant details about the observation period (date, observer, weather, etc.), the species, the clock time of the observation, and the start and stop times of each song. Records of start and stop times are collected by clicking on start and stop buttons. We have designed this so that it can be loaded on a PC based pda for use in the field, and output is an excel file. The benefit of this is that accuracy is maintained because the observer does not have to try and write down times from a stop watch and data entry is not required following data collection.

The example presented was based on simulated singing time data and the true singing rate distribution was known. In practice it may be necessary to correct singing time data for "size" bias in singing rates, as seen for the Poisson models. Correcting data for "size" bias is straightforward if all birds in the sample sing at least once. If not all birds sing during the sampling period then correcting for "size" bias may require assuming a distribution for singing rates.

The importance of correcting for availability bias is often overlooked in abundance surveys or inference is only made about the portion of the population that is available. Analysis of the Ovenbird data set demonstrates how important availability bias can be in estimating animal abundance. The Ovenbird has a very loud song and a very high singing rate so high detection probabilities are expected. High singing rates imply

high availability probabilities for the Ovenbird. Probabilities (0.87 from the two-point Poisson mixture model) were high compared to species with lower singing rates. Based on the estimate from the two-point Poisson mixture model for a three minute count 13% of the population was missed on field surveys because they were not available for detection. This estimate is probably low because Ovenbirds sing in bouts. An estimate of 20% (from the 3 minute singing time analysis) may be more realistic.

Correction of availability bias is as important as correction for perception bias or the probability that an individual is detected given that it is available for detection. Estimated probability of detection given availability of 0.80 for the Ovenbird was determined from analysis of a four-independent observer data set using a two-point heterogeneity model (Chapter 2). The magnitudes of both types of bias are comparable and thus, estimates of animal abundance should account for both.

Such corrections for availability have been done for aerial surveys of marine mammals (Marsh and Sinclair 1989), capture-recapture surveys of salamanders (Bailey et al. 2004a, 2004b) and point count surveys of birds (Farnsworth et al. 2002, chapter 3). In these surveys the availability process was related to individuals not being visible to observers because of dive depths, not being susceptible to capture because individuals remained underground to avoid desiccation, and individuals could not be detected because they did not sing during the count. The importance of estimating the availability process is seen in the salamander surveys where only 13% of the population was available for capture and in surveys of dugongs where 30 to 100% of the population was available for detection depending on water conditions and the position of an individual in the water column.

We recognize that accounting for availability bias in point count surveys using auxiliary information on singing behavior of birds significantly increases the effort required to obtain valid abundance estimates. Implementation of this technique should be assessed with respect to the amount of availability bias thought to be present in the data and the mobility of the species of interest. It may be a more efficient use of survey effort to use the time of detection method when species are sedentary. Recall that this requires a longer count period for the point count. It is better to use a very short point count for more mobile species and spend the extra time collecting data to account for availability bias. It is difficult to account for availability bias for broad scale monitoring studies given the spatial range of the surveys and the multitude of species being monitored. In these cases it may be necessary to select a few species of interest to examine the availability process.

Data on the singing behavior of birds may be useful for understanding other ecological questions as well. Spatial comparisons in changes in singing rates across the breeding season may be indicative of the quality of breeding habitats. If better habitats are occupied earlier in the year then changes in singing rates relative to breeding stage could be used to assess differences in breeding habitats. Yearly differences in singing rates may also indicate changes in populations or habitats. McShea and Rappole (1997) found that singing rates varied within a species based on habitat characteristics and local abundance of conspecifics.

### Recommendations

The biological interpretation of mixture models in this application makes using these models desirable. Modeling differences in singing rates that may be associated

with breeding phenology lends support to these models and also has ecological application for temporal and spatial comparison of singing rate differences. The limitation of this approach is assuming that singing patterns of an individual are random events. This method could be modified to model singing bouts as random events and the number of songs for each bout as a random variable. Further investigation of the use of mixture models and an evaluation of singing rate distributions is warranted.

The problems with the three-point Poisson mixture model were discouraging but expected because of the number of parameters in the model. Problems with convergence of Poisson mixture models have been noted before and the use of truncated Poissons has been recommended (Lindsay 1986). Further investigation of Poisson mixture models with more than two mixtures is needed.

Requiring that data are either unbiased or that all birds in the sample sing is a serious limitation to the singing time approach. Ideally the sample period would be sufficient so that all birds sing at least once during the period. This was the case for the Ovenbird data we presented but may not be realistic for all species. Further investigation of the treatment of birds that do not sing during a sample period is necessary. This includes an assessment of the weighting function used to correct for "size" biased data. Comparing data based on auditory detections of birds, to data collected by other means (e.g. visual detection of birds or sampling a marked population) is necessary to determine the validity of the weighting function used here, and if it is reasonable to assume a distributional form for singing rates.

**Literature cited**

Bailey, L.L., T.R. Simons and K.H. Pollock. 2004a. Estimating detection probability parameters for plethodon salamanders using the robust capture-recapture design. Journal of Wildlife Management 68:1-13.

Bailey, L.L., T.R. Simons and K.H. Pollock. 2004b. Spatial and temporal variation in detection probability of plethodon salamanders using the robust capture-recapture design. Journal of Wildlife Management 68:14-24.

Barker, R.J. and J.R. Sauer. 1995. Statistical aspects of point count sampling. Pages 125-130 in Monitoring Bird Populations by Point Counts (C.J. Ralph, J.R. Sauer, and S. Droege, Eds.). U.S. Department of Agriculture, Forest Service General Technical Report PSW-GTR-149.

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

Best, L.B. 1981. Seasonal changes in detection of individual bird species. Studies in Avain Biology 6:252-261.

Buckland, S.T., D.R. Anderson, K.P. Burnham, and J.L. Laake. 1993. Distance Sampling: Estimating Abundance of Biological Populations. Chapman and Hall, London.

Burnham, K.P. 1981. Summarizing remarks: Environmental influences. Pages 324-325 in Estimating numbers of terrestrial birds (C.J. Ralph and J.M. Scott, Eds.). Studies in Avian Biology No. 6.

Burnham, K.P. and D.R. Anderson. 2002. Model selection and multimodel inference: A practical information-theoretic approach (2nd ed). Springer-Verlag, New York, NY.

Conant, S., M.S. Collins, and C.J. Ralph. 1981. Effects of observers using different methods upon the total population estimates of two resident island birds. Studies in Avian Biology 6:377-381.

Efron, B. 1979. Bootstrap methods: Another look at the jackknife. The Annals of Statistics 7:1-26.

Efron, B. and G. Gong. 1983. A leisurely look at the bootstrap, the jackknife, and cross-validation. The American Statistician 37:36-48.

Farnsworth, G.L., K.H. Pollock, J.D. Nichols, T.R. Simons, J.E. Hines, and J.R. Sauer. 2002. A removal model for estimating detection probabilities from point-count surveys. Auk 119:414-425.

Johnson, D.H. 1995. Point counts of birds: What are we estimating? Pages 117-123 in Monitoring Bird Populations by Point Counts (C.J. Ralph, J.R. Sauer, and S. Droege, Eds.). U.S. Department of Agriculture, Forest Service General Technical Report PSW-GTR-149.

Kendall, W.L., J.D. Nichols, and J.E. Hines. 1997. Estimating temporary emigration using capture-recapture data with Pollock's robust design. Ecology 78:563-578.

Lindsay, B.G. 1983. The geometry of mixture likelihoods: A general theory. The Annals of Statistics 11:86-94.

Lindsay, B.G. 1986. Exponential family mixture models (with least-squares estimators). The Annals of Statistics 14:124-137.

Manly, B.F.J. 1998. Randomization, bootstrap and Monte Carlo methods in biology. Chapman and Hall, New York.

Marsh, H. and D.F. Sinclair. 1989. Correcting for visibility bias in strip transect aerial surveys of aquatic fauna. Journal of Wildlife Management. 53:1017-1024.

McShea, W.J. and J.H. Rappole. 1997. Variable song rates in three species of passerines and implications for estimating bird populations. Journal of Field Ornithology. 68:367-375.

Nichols, J.D., J.E. Hines, J.R. Sauer, F.W. Fallon, J.E. Fallon, and P.J. Heglund. 2000. A double-observer approach for estimating detection probability and abundance from point counts. Auk 117:393-408.

Norris, J.L., and K.H. Pollock. 1996. Nonparametric MLE under two closed capture-recapture models with heterogeneity. Biometrics. 52:639-649.

Otis, D.L., K.P. Burnham, G.C. White, and D.R. Anderson. 1978. Statistical inference from capture data on closed animal populations. Wildlife Monographs, No. 62.

Patil, G.P. and J.K. Ord. 1976. On size-biased sampling and related form-invariant weighted distributions. Sankhya: The Indian Journal of Statistics 38:48-61.

Patil, G.P. and C.R. Rao. 1977. The weighted distributions: A survey of their applications. Pages 383-405 in Applications of Statistics (P.R. Krishnaiah ed.). North Holland Publishing Company.

Patil, G.P. and C.R. Rao. 1978. Weighted distributions and size-biased sampling with applications to wildlife populations and human families. Biometrics 34:179-189.

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

Ralph, J.C., S. Droege, and J.R. Sauer. 1995b. Managing and monitoring birds using point counts: Standards and applications. Pages 161-168 in Monitoring bird populations by point counts (J.C. Ralph, J.R. Sauer, and S. Droege (Eds.)). United States Forest Service General Technical Report PSW-GTR-149.

Reynolds, R.T., J.M. Scott, and R.A. Nussbaum. 1980. A variable circular-plot method for estimating bird numbers. Condor 82:309-313.

Robbins, C.S. 1981a. Effect of time of day on bird activity. Studies in Avian Biology 6:275-286.

Robbins, C.S. 1981b. Bird activity levels related to weather. Studies in Avain Biology 6:301-310.

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

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

Royle, J.A. 2004. N-Mixture models for estimating population size from spatially replicated counts. Biometrics 60:108-115.

Sauer, J.R., J.E. Hines, G. Gough, I. Thomas, and B.G. Peterjohn. 1997. The North American Breeding Bird Survey results and analysis. Version 96.4. Patuxent Wildlife Research Center, Laurel, MD.

Scott, J.M. and F.L. Ramsey. 1981. Length of count period as a possible source of bias in estimating bird densities. Studies in Avian Biology 6:409-413.

Seber, G.A.F. 1982. The Estimation of Animal Abundance and Related Parameters (2nd ed.). Edward Arnold, London.

Skirvin, A.A. 1981. Effect of time of day and time of season on the number of observations and density estimates of breeding birds. Studies in Avian Biology 6:271-274.

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

Tuckwell, H.C. Elementary applications of probability theory (2nd ed.) Chapman and Hall, London.

Williams, B.K., J.D. Nichols, and M.J. Conroy. 2002. Analysis and Management of Animal Populations. Academic Press. San Diego, CA.

Wilson, D.M. and J. Bart. 1985. Reliability of singing bird surveys: Effects of song phenology during the breeding season. Condor 87:69-73.

Table 1: Parameter estimates for the homogeneous Poisson model and the two-point Poisson mixture models. Standard errors and percentile 95% confidence intervals obtained with 1,000 bootstrap samples. $\Delta$AIC value of zero indicates the selected model. Availability probability estimates $\hat{p}_a(1)$, $\hat{p}_a(2)$, and $\hat{p}_a(3)$ are for 1, 2, and 3 minute point count surveys, respectively.

| Model | $\Delta$AIC | $\hat{\lambda}_1$ | $\hat{\lambda}_2$ | $\hat{\delta}$ | $\hat{p}_a(1)$ | $\hat{p}_a(2)$ | $\hat{p}_a(3)$ |
|---|---|---|---|---|---|---|---|
| Homogeneous | 182.7 | 1.71(0.052) | | | 0.82(0.009) | 0.97(0.003) | 0.99(0.001) |
| 95% CI | | 1.60-1.81 | | | 0.80-0.84 | 0.96-0.97 | 0.99-1.00 |
| 2-Point | 0.0 | 0.54(0.061) | 2.02(0.069) | 0.66(0.036) | 0.57(0.029) | 0.77(0.030) | 0.87(0.025) |
| 95% CI | | (0.42-0.66) | (1.88-2.16) | (0.59-0.73) | (0.51-0.63) | (0.70-0.82) | (0.81-0.91) |

Table 2: Availability probability estimates for one, two, and three minute point count surveys using two simulated data sets for a five minute observation period. One data set uses completely random singing times and the other assumes birds sing in bouts of five songs. For one iteration a sample of 100 birds is drawn with replacement and 1,000 iterations are done for each analysis. For each data set analyses are done for one, two, and three minute point counts giving the availability probabilities $\hat{p}_a(1)$, $\hat{p}_a(2)$ and $\hat{p}_a(3)$, respectively. Percentile 95% confidence intervals are reported and standard errors are in parentheses.

| Data Set | $\hat{p}_a(1)$ | $\hat{p}_a(2)$ | $\hat{p}_a(3)$ |
|---|---|---|---|
| Random Singing | 0.48 (0.049) | 0.71 (0.046) | 0.86 (0.036) |
| 95% CI | (0.38-0.57) | (0.62-0.79) | (0.79-0.92) |
| Singing Bouts | 0.37 (0.051) | 0.61 (0.051) | 0.80 (0.040) |
| 95% CI | (0.27-0.47) | (0.51-0.70) | (0.72-0.88) |

**List of Figures**

Figure 1: Homogeneous Poisson model and two-point Poisson mixture model fit to Ovenbird singing rate data. Poisson mixture model is corrected for "size" bias that occurs in this data, which is not a factor under the assumptions of the homogeneous Poisson model.

Figure 2: Distribution of $\hat{\lambda}$ from 1,000 bootstrap estimates for the homogeneous Poisson model fit to the Ovenbird data set. Points within the percentiled 95% confidence intervals are in black.

Figure 3: Distribution of $\hat{\lambda}_1$, $\hat{\lambda}_2$ and $\hat{\delta}$ from 1,000 bootstrap estimates for the two-point Poisson mixture model fit to the Ovenbird data set. Points within the percentiled 95% confidence intervals are in black.

Figure 4: Distribution of the availability probability from simulated data with random singing times using the singing time analysis approach for one, two and three minute point counts. Data was simulated to be comparable to the Ovenbird data set.

Figure 5: Distribution of the availability probability from simulated data assuming birds sing in bouts of five songs. Analysis was based on the singing time approach for one, two, and three minute point counts. Data was simulated to be comparable to the Ovenbird data set.
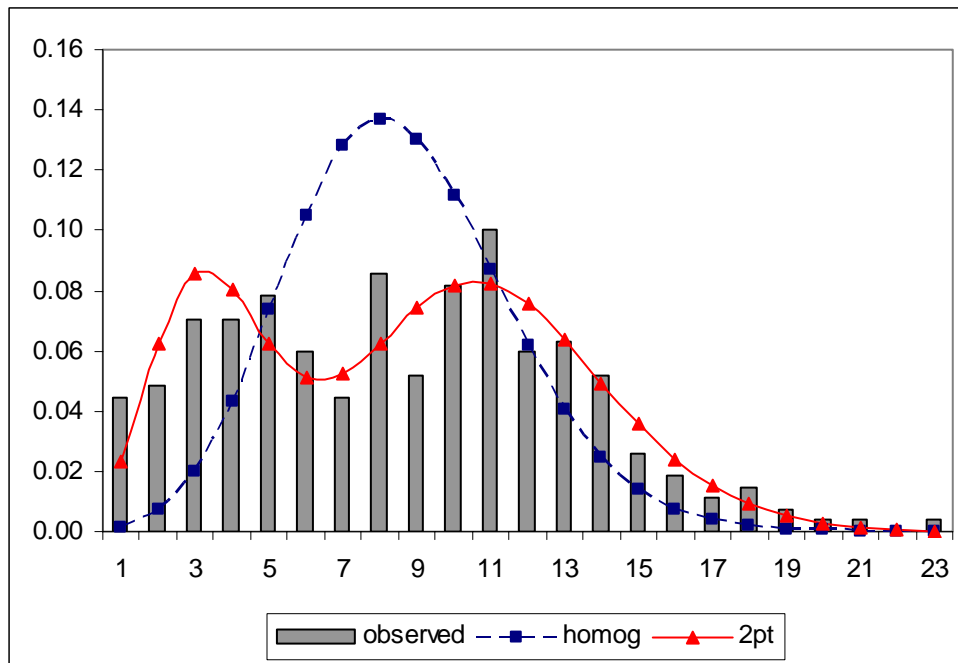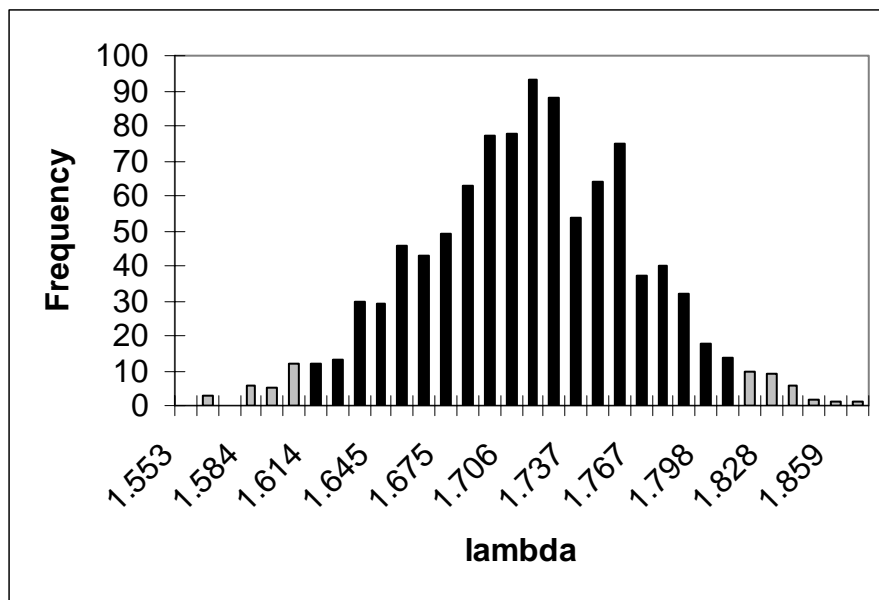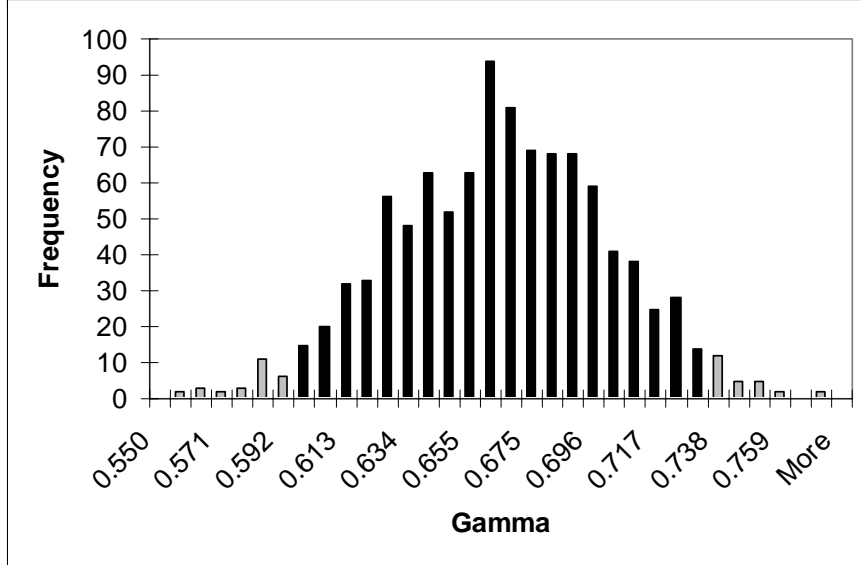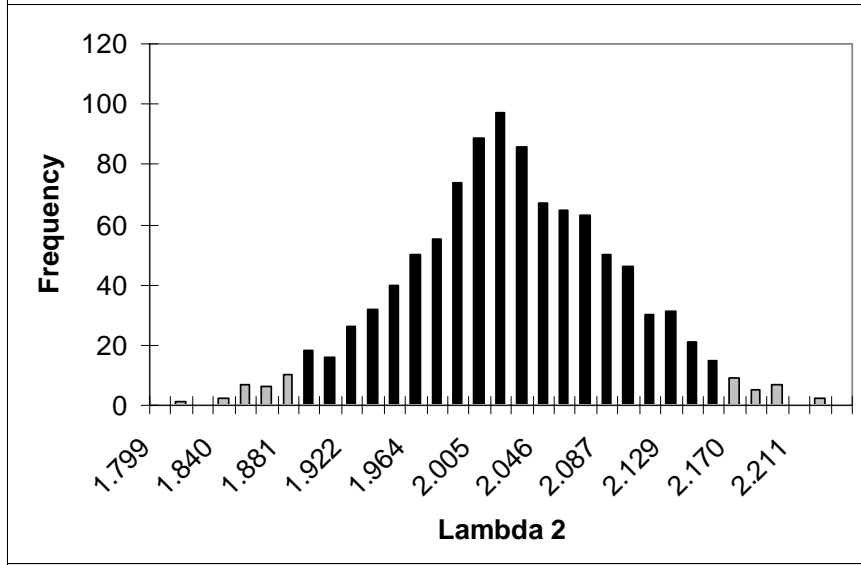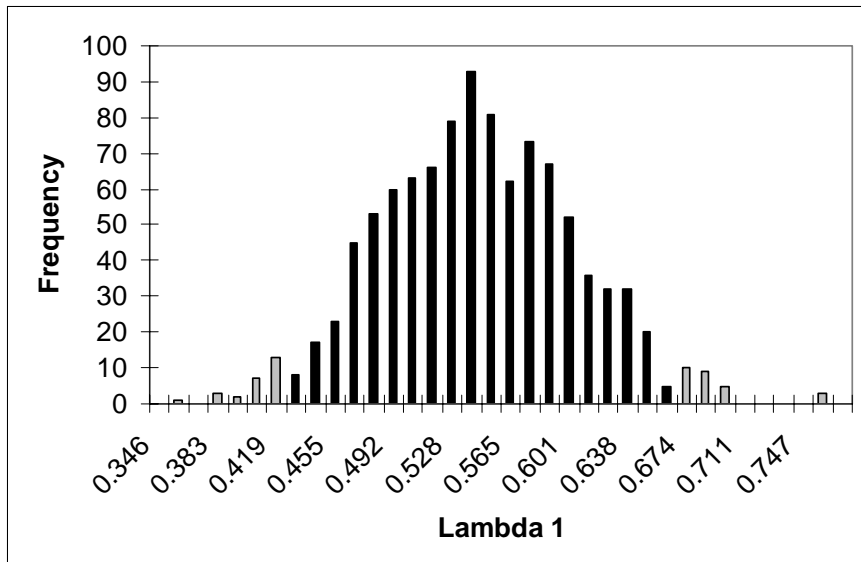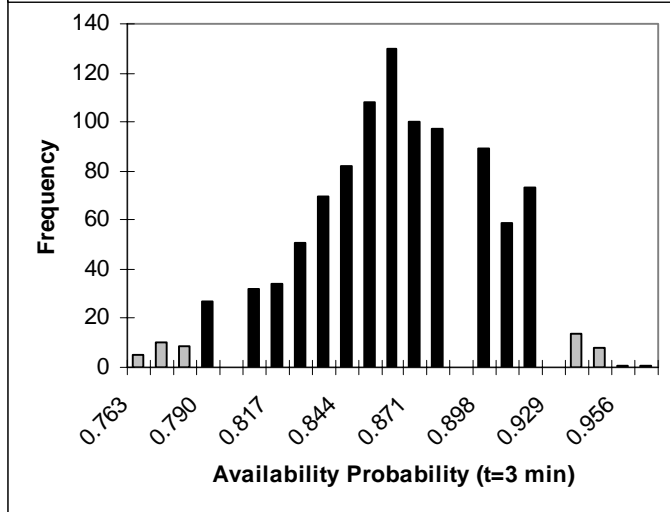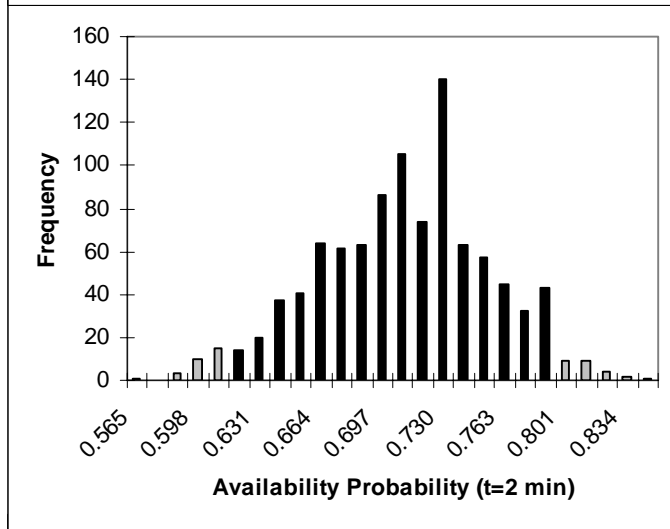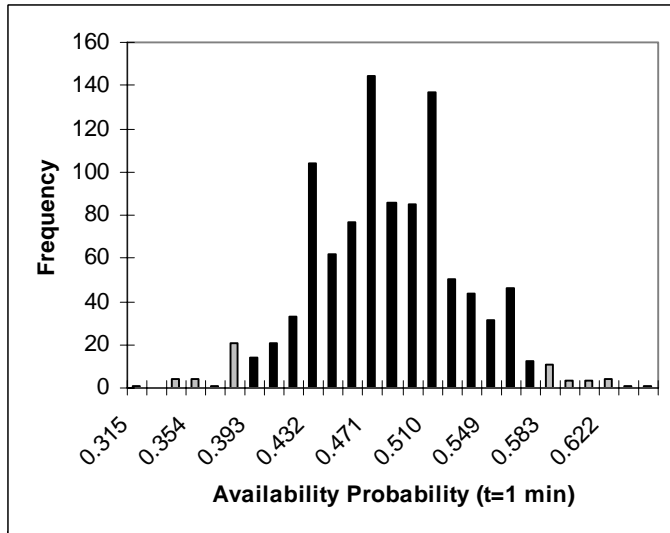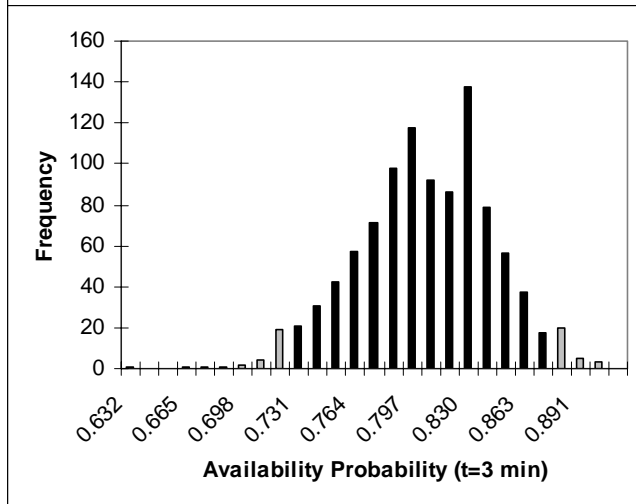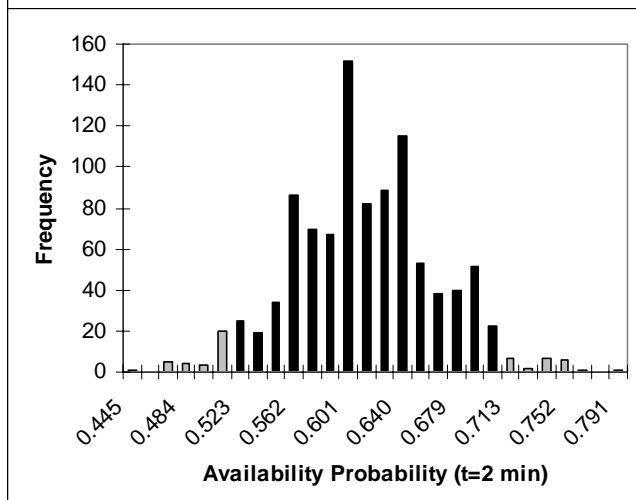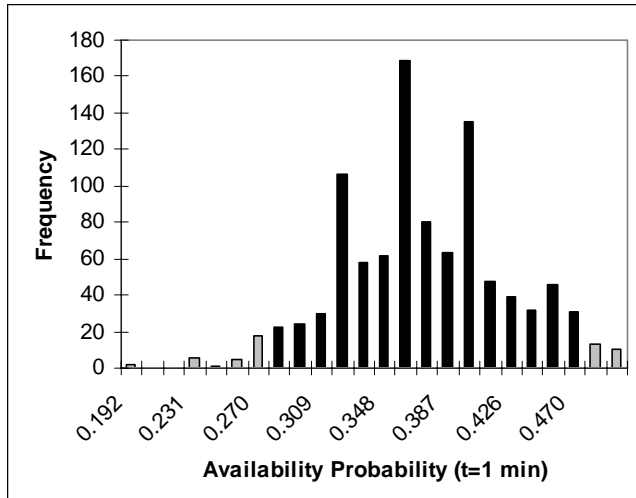
**Chapter 6**

**EXECUTIVE SUMMARY**

**Introduction**

Population abundance is one of the most common parameters used to assess the current state or health of a wildlife population. Abundance is used to make spatial and temporal comparisons of populations, to assess population trajectories and differences between populations. This information is used to infer habitat relationships and judge the effectiveness of management actions on wildlife populations of interest.

The benefit of using population abundance as a measure of population status and for spatial and temporal comparisons is that this single variable contains information about several population processes that contribute to population level changes. A series of abundance estimates provide an estimate of population change which combines information on birth, death, immigration and emigration. Population change also allows for spatial comparison between areas of interest to provide information about habitat quality or differences in management practices. For example, comparing abundance estimates or rates of population change between managed (e.g. National forests) and unmanaged areas (National Parks) provide a measure of management effectiveness.

Point count surveys are commonly used to provide indices of wildlife population abundance or, in some cases, estimates of true abundance. The ease of application and the limited effort required for these surveys compared to other methods of estimating abundance makes them very practical to use in a wide variety of field situations. Point count surveys are also used to collect information on multiple species which saves time and effort for monitoring multiple populations of interest. This type of survey is used to estimate local abundance of species of interest and for large scale monitoring of bird populations in general.

The most common use of point counts is to provide an index of population abundance or relative abundance. To make spatial or temporal comparisons valid using this type of count requires the very restrictive assumption of equal detection probability for the comparisons being made. There are several statistical approaches available to estimate the detection process from point count data and provide estimates of true abundance. Each of these statistical approaches has a set of assumptions that must be met and require additional data to be collected and/or extra field effort. The characteristics of a particular field situation will dictate which point count method is applicable based on the validity of particular assumptions.

In this thesis we have reviewed several of the available approaches to estimating true abundance from point count data and provided more general modeling frameworks for the multiple observer and time of detection approaches. We have discussed the components of the detection process and factors that will affect it. We have presented a multiple species modeling framework based on modeling similarities between species in the detection process. We have also provided a method for estimating availability bias in point counts using auxiliary data from song counts and distributions. We will review the objectives of each chapter and the findings for each. We will then provide general conclusions and implications about point count surveys.

Chapter 2: ESTIMATING DETECTION PROBABILITIES FROM MULTIPLE OBSERVER POINT COUNTS

**Objectives**

1. Present and illustrate the two independent observer method and potential models for estimating detection probability, including the use of detection distance

covariates, showing that the models are essentially closed capture-recapture models.

2. Present and illustrate a more general model using four independent observers showing that multiple observer models are essentially closed capture-recapture models that allow for individual heterogeneity.

3. Compare the efficiency of the two independent observer approach to the primary-secondary observer approach of Nichols et al. (2002).

4. Simulate data under a heterogeneous model to illustrate the levels of heterogeneity typically present in data and the effect of heterogeneity on two-observer models and index counts.

**Implications and Findings**

1. The independent observer approach gives more efficient (smaller variance) estimates than the primary-secondary observer approach.

2. Two-independent observer models appear to work well in practice and give reasonably precise estimates.

3. Four-independent observer model estimates indicate a negative bias in the two-independent observer model estimates.

4. Individual heterogeneity in detection probabilities is important, can be extreme and leads to negative bias in abundance estimates when models fail to account for it.

Chapter 3: TIME OF DETECTION METHOD FOR ESTIMATING ABUNDANCE FROM POINT COUNT SURVEYS

**Objectives**

1. Present a time of detection approach using a capture-recapture framework based on all detections of an individual as a more general alternative to the removal approach presented by Farnsworth et al. (2002).

2. Discuss field methods required to collect data suitable for this method.

3. Present a finite-mixture model for individual unobservable heterogeneity and a covariate model for observable heterogeneity.

4. Illustrate our methods with an example for three-unequal interval point counts and another for a four-equal interval point count survey.

**Implications and Findings**

1. The time of removal method (Farnsworth et al. 2002) is a special case of the time of detection method.

2. Using the full detection history in a capture-recapture framework is more efficient (smaller variance) than the time of removal approach using only first detections.

3. Time effects can be an important source of variation in point count surveys, which cannot be modeled with the removal method.

4. Individual heterogeneity is important and leads to negative bias in abundance estimates when models fail to account for it.

5. Modeling the detection process is important to estimate abundance but the contribution of the availability component and detection given availability to the overall probability of detection is unknown.

Chapter 4: MULTIPLE SPECIES ANALYSIS OF POINT COUNT DATA: A MORE PARSIMONIOUS MODELING FRAMEWORK

**Objectives**

1. Present a multiple species modeling framework to achieve more parsimonious models and explore its potential applications to three methods (distance sampling, time of detection and multiple observer) of analyzing point count data.

2. Discuss the importance of defining species groups and describe characteristics that can be used to define these groups.

3. Provide examples of grouping species and an analysis for each point count method using these species groups.

**Implications and Findings**

1. Defining species groups based on similarities in the detection process is critical to multiple species modeling because this leads to a biologically reasonable set of *a priori* candidate models.

2. Using categorical variables or subjective rankings to define species groups is possible but may lead to poorly defined species groups as not all observers agree on the rankings for certain species.

3. Multiple species modeling with the distance sampling approach did not provide better estimates than a single species approach in my examples. Further investigation of distance models of multiple species data is necessary.

4. Multiple species models for the multiple-observer and time of detection methods did provide better estimates (more precise) than a single species approach.

5. Accounting for individual heterogeneity in detection probabilities is important in multiple species models. Failure to account for this leads to serious negative bias in abundance estimates.

6. This approach can be extended to modeling rare or uncommon species that typically cannot be modeled using a single species approach because of insufficient data. Additive species effect models will be particular useful in this respect.

## Chapter 5: MODELING THE AVAILABILITY PROCESS FOR POINT COUNT SURVEYS USING AUXILIARY DATA

**Objectives**

1. Investigate models for estimating the probability of a bird being available for detection in relation to singing rates or times using data collected separately from point count surveys. These estimates can be applied to any point count method to adjust for the proportion of birds that are unavailable during the point count survey.

2. Present a homogeneous Poisson singing rate model as a simple case for modeling the availability process.

3. Present finite-mixture Poisson models as biologically reasonable models for the availability process associated with differences in singing rates relative to breeding phenology of an individual.

4. Discuss "size" bias that may occur in the data and how to correct for this.

5. Present a nonparametric approach that uses the actual singing times of an individual, corrects for "size" bias in the data and then estimates the availability probability and variances by re-sampling the singing time data.

6. Present an example of the singing rate models using data collect on the Ovenbird and present an example of the singing time approach using simulated data.

## Implications and Findings

1. Homogeneous Poisson and finite Poisson mixture models can be fit to singing rate data using a maximum likelihood approach.

2. Obtaining maximum likelihood estimates for the three-point Poisson mixture model was problematic for our data.

3. The two-point Poisson mixture model fits the data reasonably well and gives reasonable estimates of availability probabilities. Changes in singing rate associated with breeding/nesting stage provide biological justification for the use of mixture models.

4. Individual heterogeneity in singing rates is important to model when estimating the probability a bird sings during a point count. Failure to account for individual heterogeneity in singing rates will lead to a positive bias in availability probability estimates or a negative bias in abundance estimates.

5. The nonparametric approach using singing time data also seems to work well based on simulation. The singing time approach is robust to nonrandom singing behavior and will generally be more applicable to estimating the availability process.

6. This approach works with all point count survey methods which allows for "snapshot" type approaches when movement is thought to be a problem in point counts.

## General Conclusions

Two of the chapters in this thesis provided generalizations of existing point count models (Chapter 2 and 3). These generalizations give a more complete set of candidate models that can account for potential sources of variation more completely. In general the models presented in these two chapters are more efficient (smaller variance) than the existing models.

Species differences in the detection process do exist, as seen in the examples presented throughout this thesis, and failure to account for this will give invalid conclusions about species differences. For example, species with low counts may actually be more abundant than species with high counts but are just less detectable. Exploiting species similarities in the detection process is also important and can lead to more parsimonious models when done properly.

In every chapter of this thesis we found that individual heterogeneity in detection probability or singing rates was a necessary component in the models to accurately represent the detection and/or availability process. Given this we feel that it is extremely important to use point count methods that will be able to account for individual heterogeneity in detection probability in order to obtain realistic estimates of abundance. Failure to account for individual heterogeneity will lead to negatively biased abundance estimates and will give misleading spatial, temporal and species comparisons.

It is possible that individual heterogeneity in detection probabilities is related to differences in breeding stage among individuals. Singing rates or the number of times a bird sings during a point count will influence the probability that an individual is detected. There is sufficient evidence to support changes in singing rate of an individual with respect to changes in breeding stage (Wilson and Bart 1985, Best 1981). The effect of singing rate on the detection process is not well understood and the relationship of the proportions of the population at various breeding stages to the finite-mixture models of heterogeneity is speculative. Further investigation of these relationships is necessary.

**Appendix 2**


**SURVIV Code:**

**Time of Detection Analysis**

**SURVIV code for single species time of detection analysis of three-unequal interval point count survey data.**

```
proc title TIME OF DETECTION, POINT COUNT--3 PERIODS;
proc model npar=7;
/* SPECIES 1, BTNW */
cohort=444;
/* COHORT = # OF SPECIES 1 SEEN DURING COUNT
        ENTER THE NUMBER OF BIRDS THAT HAD EACH
        HISTORY, THE ORDER IS X111, X100, X010, X001
      X110, X101, AND X011*/
191: (1-s(6))*s(7)*s(7)*s(7)+s(6)*(((1-exp(-3.*s(1)))*
                (1-exp(-2.*(s(2)+s(4))))*
                (1-exp(-5.*(s(3)+s(5))))))/
                (1-((1-s(6))*(1-s(7))*(1-s(7))*(1-s(7))+s(6)*
                ((exp(-3.*s(1)))*(exp(-2.*s(2)))*
                (exp(-5.*s(3))))));
34: (1-s(6))*s(7)*(1-s(7))*(1-s(7))+s(6)*((1-exp(-3.*s(1)))*
                (exp(-2.*(s(2)+s(4))))*
                (exp(-5.*(s(3)+s(5)))))/
                (1-((1-s(6))*(1-s(7))*(1-s(7))*(1-s(7))+s(6)*
                ((exp(-3.*s(1)))*(exp(-2.*s(2)))*
                (exp(-5.*s(3))))));
26: (1-s(6))*(1-s(7))*s(7)*(1-s(7))+s(6)*(((exp(-3.*s(1)))*
                (1-exp(-2.*s(2)))*
                (exp(-5.*(s(2)+s(5))))))/
                (1-((1-s(6))*(1-s(7))*(1-s(7))*(1-s(7))+s(6)*
                ((exp(-3.*s(1)))*(exp(-2.*s(2)))*
                (exp(-5.*s(3))))));
46: (1-s(6))*(1-s(7))*(1-s(7))*s(7)+s(6)*(((exp(-3.*s(1)))*
                (exp(-2.*s(2)))*
                (1-exp(-5.*s(3)))))/
                (1-((1-s(6))*(1-s(7))*(1-s(7))*(1-s(7))+s(6)*
                ((exp(-3.*s(1)))*(exp(-2.*s(2)))*
                (exp(-5.*s(3))))));
44: (1-s(6))*s(7)*s(7)*(1-s(7))+s(6)*(((1-exp(-3.*s(1)))*
                (1-exp(-2.*(s(2)+s(4))))*
                (exp(-5.*(s(3)+s(5))))))/
                (1-((1-s(6))*(1-s(7))*(1-s(7))*(1-s(7))+s(6)*
                ((exp(-3.*s(1)))*(exp(-2.*s(2)))*
                (exp(-5.*s(3))))));
66: (1-s(6))*s(7)*(1-s(7))*s(7)+s(6)*(((1-exp(-3.*s(1)))*
                (exp(-2.*(s(2)+s(4))))*
                (1-exp(-5.*(s(3)+s(5)))))/
                (1-((1-s(6))*(1-s(7))*(1-s(7))*(1-s(7))+s(6)*
                ((exp(-3.*s(1)))*(exp(-2.*s(2)))*
```

```
                     (exp(-5.*s(3))))));
37: (1-s(6))*s(7)*s(7)*(1-s(7))+s(6)*(((exp(-3.*s(1)))*
                     (1-exp(-2.*s(2)))*
                     (1-exp(-5.*(s(3)+s(5))))))/
                     (1-((1-s(6))*(1-s(7))*(1-s(7))*(1-s(7))+s(6)*
                     ((exp(-3.*s(1)))*(exp(-2.*s(2)))*
                     (exp(-5.*s(3))))));
Labels /* TO HELP IDENTIFY PARAMETERS */;
s(1)=p11 (p1 species 1);
s(2)=p21 (p2 species 1);
s(3)=p31 (p3 species 1);
s(4)=c21 (c2 species 1);
s(5)=c31 (c3 species 1);
proc estimate name=b;
initial;
constraints; s(1)=s(2); s(1)=s(3); s(4)=s(5);
s(6)=1.; s(7)=0.;
proc estimate name=t;
initial;
constraints; s(4)=0.; s(5)=0.;
s(6)=1.; s(7)=0.;
proc estimate name=dot;
initial;
constraints; s(1)=s(2); s(1)=s(3); s(4)=0.; s(5)=0.;
s(6)=1.; s(7)=0.;
proc estimate name=h1;
initial;
constraints; s(1)=s(2); s(1)=s(3); s(4)=0.; s(5)=0.;
S(7)=1.;
proc estimate name=h9;
initial;
constraints; s(1)=s(2); s(1)=s(3); s(4)=0.; s(5)=0.;
S(7)=0.9;
proc estimate name=bh1;
initial;
constraints; s(1)=s(2); s(1)=s(3); s(4)=s(5);
s(7)=1.;
proc estimate name=bh9;
initial;
constraints; s(1)=s(2); s(1)=s(3); s(4)=s(5);
s(7)=0.9;
proc test;
```

**SURVIV code for four-species time of detection analysis of three-unequal interval point count survey data.**

```
proc title TIME OF DETECTION, POINT COUNT--3 PERIODS;
proc model npar=28;
/* SPECIES 1, RN */
cohort=54;
/* COHORT = # OF SPECIES 1 SEEN DURING COUNT
        ENTER THE NUMBER OF BIRDS THAT HAD EACH
        HISTORY, THE ORDER IS X111, X100, X010, X001
      X110, X101, AND X011*/
13: (1-s(6))*s(7)*s(7)*s(7)+s(6)*(((1-exp(-3.*s(1)))*
              (1-exp(-2.*(s(2)+s(4))))*
              (1-exp(-5.*(s(3)+s(5))))))/
              (1-((1-s(6))*(1-s(7))*(1-s(7))*(1-s(7))+s(6)*
              ((exp(-3.*s(1)))*(exp(-2.*s(2)))*
              (exp(-5.*s(3)))))));
8: (1-s(6))*s(7)*(1-s(7))*(1-s(7))+s(6)*((1-exp(-3.*s(1)))*
              (exp(-2.*(s(2)+s(4))))*
              (exp(-5.*(s(3)+s(5)))))/
              (1-((1-s(6))*(1-s(7))*(1-s(7))*(1-s(7))+s(6)*
              ((exp(-3.*s(1)))*(exp(-2.*s(2)))*
              (exp(-5.*s(3)))))));
5: (1-s(6))*(1-s(7))*s(7)*(1-s(7))+s(6)*(((exp(-3.*s(1)))*
              (1-exp(-2.*s(2)))*
              (exp(-5.*(s(2)+s(5))))))/
              (1-((1-s(6))*(1-s(7))*(1-s(7))*(1-s(7))+s(6)*
              ((exp(-3.*s(1)))*(exp(-2.*s(2)))*
              (exp(-5.*s(3)))))));
14: (1-s(6))*(1-s(7))*(1-s(7))*s(7)+s(6)*(((exp(-3.*s(1)))*
              (exp(-2.*s(2)))*
              (1-exp(-5.*s(3)))))/
              (1-((1-s(6))*(1-s(7))*(1-s(7))*(1-s(7))+s(6)*
              ((exp(-3.*s(1)))*(exp(-2.*s(2)))*
              (exp(-5.*s(3)))))));
6: (1-s(6))*s(7)*s(7)*(1-s(7))+s(6)*(((1-exp(-3.*s(1)))*
              (1-exp(-2.*(s(2)+s(4))))*
              (exp(-5.*(s(3)+s(5))))))/
              (1-((1-s(6))*(1-s(7))*(1-s(7))*(1-s(7))+s(6)*
              ((exp(-3.*s(1)))*(exp(-2.*s(2)))*
              (exp(-5.*s(3)))))));
3: (1-s(6))*s(7)*(1-s(7))*s(7)+s(6)*(((1-exp(-3.*s(1)))*
              (exp(-2.*(s(2)+s(4))))*
              (1-exp(-5.*(s(3)+s(5)))))/
              (1-((1-s(6))*(1-s(7))*(1-s(7))*(1-s(7))+s(6)*
              ((exp(-3.*s(1)))*(exp(-2.*s(2)))*
```

```
                 (exp(-5.*s(3))))));
5: (1-s(6))*s(7)*s(7)*(1-s(7))+s(6)*(((exp(-3.*s(1)))*
                 (1-exp(-2.*s(2)))*
                 (1-exp(-5.*(s(3)+s(5))))))/
                 (1-((1-s(6))*(1-s(7))*(1-s(7))*(1-s(7))+s(6)*
                 ((exp(-3.*s(1)))*(exp(-2.*s(2)))*
                 (exp(-5.*s(3))))));
/* SPECIES 2, TT */
cohort=104;
/* COHORT = # OF SPECIES 2 SEEN DURING COUNT
       ENTER THE NUMBER OF BIRDS THAT HAD EACH
       HISTORY, THE ORDER IS X111, X100, X010, X001
     X110, X101, AND X011*/
28: (1-s(13))*s(14)*s(14)*s(14)+s(13)*(((1-exp(-3.*s(8)))*
                 (1-exp(-2.*(s(9)+s(11))))*
                 (1-exp(-5.*(s(10)+s(12))))))/
                 (1-((1-s(13))*(1-s(14))*(1-s(14))*(1-s(14))+s(13)*
                 ((exp(-3.*s(8)))*(exp(-2.*s(9)))*
                 (exp(-5.*s(10))))));
24: (1-s(13))*s(14)*(1-s(14))*(1-s(14))+s(13)*((1-exp(-3.*s(8)))*
                 (exp(-2.*(s(9)+s(11))))*
                 (exp(-5.*(s(10)+s(12)))))/
                 (1-((1-s(13))*(1-s(14))*(1-s(14))*(1-s(14))+s(13)*
                 ((exp(-3.*s(8)))*(exp(-2.*s(9)))*
                 (exp(-5.*s(10))))));
8: (1-s(13))*(1-s(14))*s(14)*(1-s(14))+s(13)*(((exp(-3.*s(8)))*
                 (1-exp(-2.*s(9)))*
                 (exp(-5.*(s(9)+s(12))))))/
                 (1-((1-s(13))*(1-s(14))*(1-s(14))*(1-s(14))+s(13)*
                 ((exp(-3.*s(8)))*(exp(-2.*s(9)))*
                 (exp(-5.*s(10))))));
22: (1-s(13))*(1-s(14))*(1-s(14))*s(14)+s(13)*(((exp(-3.*s(8)))*
                 (exp(-2.*s(9)))*
                 (1-exp(-5.*s(10)))))/
                 (1-((1-s(13))*(1-s(14))*(1-s(14))*(1-s(14))+s(13)*
                 ((exp(-3.*s(8)))*(exp(-2.*s(9)))*
                 (exp(-5.*s(10))))));
10: (1-s(13))*s(14)*s(14)*(1-s(14))+s(13)*(((1-exp(-3.*s(8)))*
                 (1-exp(-2.*(s(9)+s(11))))*
                 (exp(-5.*(s(10)+s(12))))))/
                 (1-((1-s(13))*(1-s(14))*(1-s(14))*(1-s(14))+s(13)*
                 ((exp(-3.*s(8)))*(exp(-2.*s(9)))*
                 (exp(-5.*s(10))))));
9: (1-s(13))*s(14)*(1-s(14))*s(14)+s(13)*(((1-exp(-3.*s(8)))*
                 (exp(-2.*(s(9)+s(11))))*
                 (1-exp(-5.*(s(10)+s(12))))))/
```

```
                   (1-((1-s(13))*(1-s(14))*(1-s(14))*(1-s(14))+s(13)*
                   ((exp(-3.*s(8)))*(exp(-2.*s(9)))*
                   (exp(-5.*s(10))))));
3: (1-s(13))*s(14)*s(14)*(1-s(14))+s(13)*(((exp(-3.*s(8)))*
                   (1-exp(-2.*s(9)))*
                   (1-exp(-5.*(s(10)+s(12)))))/
                   (1-((1-s(13))*(1-s(14))*(1-s(14))*(1-s(14))+s(13)*
                   ((exp(-3.*s(8)))*(exp(-2.*s(9)))*
                   (exp(-5.*s(10))))));
/* SPECIES 3, WO */
cohort=106;
/* COHORT = # OF SPECIES 3 SEEN DURING COUNT
        ENTER THE NUMBER OF BIRDS THAT HAD EACH
        HISTORY, THE ORDER IS X111, X100, X010, X001
     X110, X101, AND X011*/
43: (1-s(20))*s(21)*s(21)*s(21)+s(20)*(((1-exp(-3.*s(15)))*
                   (1-exp(-2.*(s(16)+s(18))))*
                   (1-exp(-5.*(s(17)+s(19)))))/
                   (1-((1-s(20))*(1-s(21))*(1-s(21))*(1-s(21))+s(20)*
                   ((exp(-3.*s(15)))*(exp(-2.*s(16)))*
                   (exp(-5.*s(17))))));
14: (1-s(20))*s(21)*(1-s(21))*(1-s(21))+s(20)*((1-exp(-3.*s(15)))*
                   (exp(-2.*(s(16)+s(18))))*
                   (exp(-5.*(s(17)+s(19)))))/
                   (1-((1-s(20))*(1-s(21))*(1-s(21))*(1-s(21))+s(20)*
                   ((exp(-3.*s(15)))*(exp(-2.*s(16)))*
                   (exp(-5.*s(17))))));
0: (1-s(20))*(1-s(21))*s(21)*(1-s(21))+s(20)*(((exp(-3.*s(15)))*
                   (1-exp(-2.*s(16)))*
                   (exp(-5.*(s(16)+s(19)))))/
                   (1-((1-s(20))*(1-s(21))*(1-s(21))*(1-s(21))+s(20)*
                   ((exp(-3.*s(15)))*(exp(-2.*s(16)))*
                   (exp(-5.*s(17))))));
19: (1-s(20))*(1-s(21))*(1-s(21))*s(21)+s(20)*(((exp(-3.*s(15)))*
                   (exp(-2.*s(16)))*
                   (1-exp(-5.*s(17)))))/
                   (1-((1-s(20))*(1-s(21))*(1-s(21))*(1-s(21))+s(20)*
                   ((exp(-3.*s(15)))*(exp(-2.*s(16)))*
                   (exp(-5.*s(17))))));
15: (1-s(20))*s(21)*s(21)*(1-s(21))+s(20)*(((1-exp(-3.*s(15)))*
                   (1-exp(-2.*(s(16)+s(18))))*
                   (exp(-5.*(s(17)+s(19)))))/
                   (1-((1-s(20))*(1-s(21))*(1-s(21))*(1-s(21))+s(20)*
                   ((exp(-3.*s(15)))*(exp(-2.*s(16)))*
                   (exp(-5.*s(17))))));
8: (1-s(20))*s(21)*(1-s(21))*s(21)+s(20)*(((1-exp(-3.*s(15)))*
```

```
                          (exp(-2.*(s(16)+s(18))))*
                          (1-exp(-5.*(s(17)+s(19))))))/
                          (1-((1-s(20))*(1-s(21))*(1-s(21))*(1-s(21))+s(20)*
                          ((exp(-3.*s(15)))*(exp(-2.*s(16)))*
                          (exp(-5.*s(17)))))));
7: (1-s(20))*s(21)*s(21)*(1-s(21))+s(20)*(((exp(-3.*s(15)))*
                          (1-exp(-2.*s(16)))*
                          (1-exp(-5.*(s(17)+s(19))))))/
                          (1-((1-s(20))*(1-s(21))*(1-s(21))*(1-s(21))+s(20)*
                          ((exp(-3.*s(15)))*(exp(-2.*s(16)))*
                          (exp(-5.*s(17)))))));
/* SPECIES 4, WW */
cohort=153;
/* COHORT = # OF SPECIES 4 SEEN DURING COUNT
        ENTER THE NUMBER OF BIRDS THAT HAD EACH
        HISTORY, THE ORDER IS X111, X100, X010, X001
    X110, X101, AND X011*/
103: (1-s(27))*s(28)*s(28)*s(28)+s(27)*(((1-exp(-3.*s(22)))*
                          (1-exp(-2.*(s(23)+s(25))))*
                          (1-exp(-5.*(s(24)+s(26))))))/
                          (1-((1-s(27))*(1-s(28))*(1-s(28))*(1-s(28))+s(27)*
                          ((exp(-3.*s(22)))*(exp(-2.*s(23)))*
                          (exp(-5.*s(24)))))));
8: (1-s(27))*s(28)*(1-s(28))*(1-s(28))+s(27)*((1-exp(-3.*s(22)))*
                          (exp(-2.*(s(23)+s(25))))*
                          (exp(-5.*(s(24)+s(26)))))/
                          (1-((1-s(27))*(1-s(28))*(1-s(28))*(1-s(28))+s(27)*
                          ((exp(-3.*s(22)))*(exp(-2.*s(23)))*
                          (exp(-5.*s(24)))))));
1: (1-s(27))*(1-s(28))*s(28)*(1-s(28))+s(27)*(((exp(-3.*s(22)))*
                          (1-exp(-2.*s(23)))*
                          (exp(-5.*(s(23)+s(26))))))/
                          (1-((1-s(27))*(1-s(28))*(1-s(28))*(1-s(28))+s(27)*
                          ((exp(-3.*s(22)))*(exp(-2.*s(23)))*
                          (exp(-5.*s(24)))))));
11: (1-s(27))*(1-s(28))*(1-s(28))*s(28)+s(27)*(((exp(-3.*s(22)))*
                          (exp(-2.*s(23)))*
                          (1-exp(-5.*s(24)))))/
                          (1-((1-s(27))*(1-s(28))*(1-s(28))*(1-s(28))+s(27)*
                          ((exp(-3.*s(22)))*(exp(-2.*s(23)))*
                          (exp(-5.*s(24)))))));
11: (1-s(27))*s(28)*s(28)*(1-s(28))+s(27)*(((1-exp(-3.*s(22)))*
                          (1-exp(-2.*(s(23)+s(25))))*
                          (exp(-5.*(s(24)+s(26))))))/
                          (1-((1-s(27))*(1-s(28))*(1-s(28))*(1-s(28))+s(27)*
                          ((exp(-3.*s(22)))*(exp(-2.*s(23)))*
```

```
                    (exp(-5.*s(24))))));
9: (1-s(27))*s(28)*(1-s(28))*s(28)+s(27)*(((1-exp(-3.*s(22))))*
                    (exp(-2.*(s(23)+s(25)))))*
                    (1-exp(-5.*(s(24)+s(26)))))))/
                    (1-((1-s(27))*(1-s(28))*(1-s(28))*(1-s(28))+s(27)*
                    ((exp(-3.*s(22))))*(exp(-2.*s(23)))*
                    (exp(-5.*s(24))))));
10: (1-s(27))*s(28)*s(28)*(1-s(28))+s(27)*(((exp(-3.*s(22))))*
                    (1-exp(-2.*s(23)))*
                    (1-exp(-5.*(s(24)+s(26)))))))/
                    (1-((1-s(27))*(1-s(28))*(1-s(28))*(1-s(28))+s(27)*
                    ((exp(-3.*s(22))))*(exp(-2.*s(23)))*
                    (exp(-5.*s(24))))));
Labels /* TO HELP IDENTIFY PARAMETERS */;
s(1)=p11 (p1 species 1);
s(2)=p21 (p2 species 1);
s(3)=p31 (p3 species 1);
s(4)=c21 (c2 species 1);
s(5)=c31 (c3 species 1);
s(6)=pi1 (pi species 1);
s(7)=gc1 (p grp spp 1);
s(8)=p12 (p1 species 2);
s(9)=p22 (p2 species 2);
s(10)=p32 (p3 species 2);
s(11)=c22 (c2 species 2);
s(12)=c32 (c3 species 2);
s(13)=pi2 (pi species 2);
s(14)=gc2 (p grp spp 2);
s(15)=p13 (p1 species 3);
s(16)=p23 (p2 species 3);
s(17)=p33 (p3 species 3);
s(18)=c23 (c2 species 3);
s(19)=c33 (c3 species 3);
s(20)=pi3 (pi species 3);
s(21)=gc3 (p grp spp 3);
s(22)=p14 (p1 species 4);
s(23)=p24 (p2 species 4);
s(24)=p34 (p3 species 4);
s(25)=c24 (c2 species 4);
s(26)=c34 (c3 species 4);
s(27)=pi4 (pi species 4);
s(28)=gc4 (p grp spp 4);
proc estimate name=M();
initial;
constraints; s(1)=s(2); s(1)=s(3);
s(1)=s(8); s(1)=s(9); s(1)=s(10);
```

```
s(1)=s(15); s(1)=s(16); s(1)=s(17);
s(1)=s(22); s(1)=s(23); s(1)=s(24);
s(4)=0.; s(5)=0.;
s(11)=0.; s(12)=0.;
s(18)=0.; s(19)=0.;
s(25)=0.; s(26)=0.;
s(6)=1.; s(13)=1.; s(20)=1.; s(27)=1.;
s(7)=0.; s(14)=0.; s(21)=0.; s(28)=0.;
proc estimate name=M()grp;
initial;
constraints; s(1)=s(2); s(1)=s(3);
s(8)=s(9); s(8)=s(10);
s(15)=s(16); s(15)=s(17);
s(22)=s(23); s(22)=s(24);
s(4)=0.; s(5)=0.;
s(11)=0.; s(12)=0.;
s(18)=0.; s(19)=0.;
s(25)=0.; s(26)=0.;
s(6)=1.; s(13)=1.; s(20)=1.; s(27)=1.;
s(7)=0.; s(14)=0.; s(21)=0.; s(28)=0.;
proc estimate name=M(t);
initial;
constraints;
s(1)=s(8); s(2)=s(9); s(3)=s(10);
s(1)=s(15); s(2)=s(16); s(3)=s(17);
s(1)=s(22); s(2)=s(23); s(3)=s(24);
s(4)=0.; s(5)=0.;
s(11)=0.; s(12)=0.;
s(18)=0.; s(19)=0.;
s(25)=0.; s(26)=0.;
s(6)=1.; s(13)=1.; s(20)=1.; s(27)=1.;
s(7)=0.; s(14)=0.; s(21)=0.; s(28)=0.;
proc estimate name=M(t)grp;
initial;
constraints;
s(4)=0.; s(5)=0.;
s(11)=0.; s(12)=0.;
s(18)=0.; s(19)=0.;
s(25)=0.; s(26)=0.;
s(6)=1.; s(13)=1.; s(20)=1.; s(27)=1.;
s(7)=0.; s(14)=0.; s(21)=0.; s(28)=0.;
proc estimate name=M(b);
initial;
constraints; s(1)=s(2); s(1)=s(3);
s(1)=s(8); s(1)=s(9); s(1)=s(10);
s(1)=s(15); s(1)=s(16); s(1)=s(17);
```

```
s(1)=s(22); s(1)=s(23); s(1)=s(24);
s(4)=s(5);
s(4)=s(11); S(4)=s(12);
s(4)=s(18); s(4)=s(19);
s(4)=s(25); s(4)=s(26);
s(6)=1.; s(13)=1.; s(20)=1.; s(27)=1.;
s(7)=0.; s(14)=0.; s(21)=0.; s(28)=0.;
proc estimate name=M(b)grp;
initial;
constraints; s(1)=s(2); s(1)=s(3);
s(8)=s(9); s(8)=s(10);
s(15)=s(16); s(15)=s(17);
s(22)=s(23); s(22)=s(24);
s(4)=s(5);
S(11)=s(12);
s(18)=s(19);
s(25)=s(26);
s(6)=1.; s(13)=1.; s(20)=1.; s(27)=1.;
s(7)=0.; s(14)=0.; s(21)=0.; s(28)=0.;
proc estimate name=M(h);
initial;
constraints; s(1)=s(2); s(1)=s(3);
s(1)=s(8); s(1)=s(9); s(1)=s(10);
s(1)=s(15); s(1)=s(16); s(1)=s(17);
s(1)=s(22); s(1)=s(23); s(1)=s(24);
s(4)=0.; s(5)=0.;
s(11)=0.; s(12)=0.;
s(18)=0.; s(19)=0.;
s(25)=0.; s(26)=0.;
s(7)=0.9; s(14)=0.9; s(21)=0.9; s(28)=0.9;
proc estimate name=M(h)grp;
initial;
constraints; s(1)=s(2); s(1)=s(3);
s(8)=s(9); s(8)=s(10);
s(15)=s(16); s(15)=s(17);
s(22)=s(23); s(22)=s(24);
s(4)=0.; s(5)=0.;
s(11)=0.; s(12)=0.;
s(18)=0.; s(19)=0.;
s(25)=0.; s(26)=0.;
s(7)=0.9; s(14)=0.9; s(21)=0.9; s(28)=0.9;
proc estimate name=M(bh);
initial;
constraints; s(1)=s(2); s(1)=s(3);
s(1)=s(8); s(1)=s(9); s(1)=s(10);
s(1)=s(15); s(1)=s(16); s(1)=s(17);
```

```
s(1)=s(22); s(1)=s(23); s(1)=s(24);
s(4)=s(5);
s(4)=s(11); S(4)=s(12);
s(4)=s(18); s(4)=s(19);
s(4)=s(25); s(4)=s(26);
s(7)=0.9; s(14)=0.9; s(21)=0.9; s(28)=0.9;
proc estimate name=M(bh)grp;
initial;
constraints; s(1)=s(2); s(1)=s(3);
s(8)=s(9); s(8)=s(10);
s(15)=s(16); s(15)=s(17);
s(22)=s(23); s(22)=s(24);
s(4)=s(5);
S(11)=s(12);
s(18)=s(19);
s(25)=s(26);
s(7)=0.9; s(14)=0.9; s(21)=0.9; s(28)=0.9;
proc estimate name=M(th);
initial;
constraints;
s(1)=s(8); s(2)=s(9); s(3)=s(10);
s(1)=s(15); s(2)=s(16); s(3)=s(17);
s(1)=s(22); s(2)=s(23); s(3)=s(24);
s(4)=0.; s(5)=0.;
s(11)=0.; s(12)=0.;
s(18)=0.; s(19)=0.;
s(25)=0.; s(26)=0.;
s(7)=0.9; s(14)=0.9; s(21)=0.9; s(28)=0.9;
proc estimate name=M(th)grp;
initial;
constraints;
s(4)=0.; s(5)=0.;
s(11)=0.; s(12)=0.;
s(18)=0.; s(19)=0.;
s(25)=0.; s(26)=0.;
s(7)=0.9; s(14)=0.9; s(21)=0.9; s(28)=0.9;
proc estimate name=M(h*);
initial;
constraints; s(1)=s(2); s(1)=s(3);
s(1)=s(8); s(1)=s(9); s(1)=s(10);
s(1)=s(15); s(1)=s(16); s(1)=s(17);
s(1)=s(22); s(1)=s(23); s(1)=s(24);
s(4)=0.; s(5)=0.;
s(11)=0.; s(12)=0.;
s(18)=0.; s(19)=0.;
s(25)=0.; s(26)=0.;
```

```
s(6)=s(13); s(6)=s(20); s(6)=s(27);
s(7)=0.9; s(14)=0.9; s(21)=0.9; s(28)=0.9;
proc estimate name=M(bh*);
initial;
constraints; s(1)=s(2); s(1)=s(3);
s(1)=s(8); s(1)=s(9); s(1)=s(10);
s(1)=s(15); s(1)=s(16); s(1)=s(17);
s(1)=s(22); s(1)=s(23); s(1)=s(24);
s(4)=s(5);
s(4)=s(11); S(4)=s(12);
s(4)=s(18); s(4)=s(19);
s(4)=s(25); s(4)=s(26);
s(6)=s(13); s(6)=s(20); s(6)=s(27);
s(7)=0.9; s(14)=0.9; s(21)=0.9; s(28)=0.9;
proc estimate name=M(th*);
initial;
constraints;
s(1)=s(8); s(2)=s(9); s(3)=s(10);
s(1)=s(15); s(2)=s(16); s(3)=s(17);
s(1)=s(22); s(2)=s(23); s(3)=s(24);
s(4)=0.; s(5)=0.;
s(11)=0.; s(12)=0.;
s(18)=0.; s(19)=0.;
s(25)=0.; s(26)=0.;
s(6)=s(13); s(6)=s(20); s(6)=s(27);
s(7)=0.9; s(14)=0.9; s(21)=0.9; s(28)=0.9;
proc test;
```

# Appendix 3


# Singing Time Programs

**MATLAB code for singing time analysis.**

```
cdf = [0.229381331, 0.353629552, 0.474691922, 0.565488698, 0.645772164,
0.696745794, 0.729514555, ...
0.784470499, 0.814205116, 0.85625836, 0.903177269, 0.928664083,
0.953660767, 0.972775878, 0.981696263, ...
0.987669735, 0.99104299, 0.995290792, 0.997302909, 0.998258665,
0.999168908, 0.999168908,1];  % cdf is : 1*23

sample = 3*60; % sampling interval
interval = 5*60; % total duration for experiment data

rate_freq = zeros(23,2); rate_freq(1:23,1) = (1:1:23)';  % data entry numbers for
each rate from 1 to 23

rate_freq(1:23,2) = [12, 13, 19, 19, 21, 16, 12, 23, 14, 22, 27, 16, 17, 14, 7, 5, 3, 4,
2, 1, 1, 0, 1]'; % need to change for new data

rate_details =
[1,293,298,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0
,0,0,0,0,0,0;
1,288,293,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,
0,0,0,0,0,0;
1,0,5,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,0,
0,0,0;  … Continue data matrix];

rand('state',sum(100*clock));
loop1 = 1000; % outer loop
loop2 = 100; % inner loop
for i = 1:1:loop1,
   counts = 0; % total number of the overlapping counts
   for j = 1:1:loop2,
    seed = rand(1);  % determine direction
    if seed <=0.5,
     direction = 1;
    else
     direction = -1;
    end

    ratef = 0;
    while (ratef ==0),
    seed = rand(1);  % determine rate
    for k=1:1:22,    % pick a rate
      if seed <=cdf(1),
        rate = 1;
        break;
```

```
    elseif (seed > cdf(k) & seed <= cdf(k+1))
       rate = k+1;
       break;
    end
  end
  % pick a record, need the information of how many records for each rate
  ratef = rate_freq(rate,2);
  end

  if rate > 1,
    rate_end = sum(rate_freq(1:rate-1,2));
  else
    rate_end = 0;
  end
  rate_start = rate_end + 1; % locate rate_details
  new_rate_end = rate_start + ratef - 1;  % now we need to pick a record from
rate_details(rate_start:new_rate_end,:)
  dh = 1.0/ratef;
  new_cdf = 0:dh:1;
  seed = rand(1);
  for k=1:1:ratef,
    if seed <=new_cdf(1) | ratef ==1,
       record = 1;
       break;
    elseif (seed > new_cdf(k) & seed <= new_cdf(k+1))
       record = k+1;
       break;
    end
  end
  % now the record is the number inside the subset of the choosen rate
  index1 = rate_start + record - 1;
  if direction ==1,
     start1 = (interval-sample)*rand(1);  % generate start point from [0,interval-
sample]
     end1 = start1 + sample;
     m = 2;
     while rate_details(index1,m) ~=0 & m <=11
        if rate_details(index1,m)>=start1 & rate_details(index1,m)<=end1,
           counts = counts + 1;
           break
        end
        m = m + 1;
     end
  else
     start1 = sample + (interval-sample)*rand(1);  % generate start point from
[sample,interval]
```

```
        end1 = start1 - sample;
        m = 2;
        while rate_details(index1,m) ~=0 & m <=11
            if rate_details(index1,m)>=end1 & rate_details(index1,m)<=start1,
                counts = counts + 1;
                break
            end
            m = m + 1;
        end
    end
  end
  est_p(i) = counts/loop2;
end
out=est_p';
save sing_rand3.out out -ascii -double
mean(est_p)
var(est_p)
```

**Singing Time Program for a Pocket PC.**