CONSERVATION AND MANAGEMENT OF RARE SPECIES: THE DEVELOPMENT OF ADAPTIVE MODELS TO REDUCE UNCERTAINTY INFLUENCING DECISION MAKING

by

JAMIAN KRISHNA PACIFICI

(Under the Direction of Michael J. Conroy and Robert J. Cooper)

ABSTRACT

Conservation and management of rare species is one of the most challenging tasks confronting natural resource managers. Species are classified as rare for several reasons: (1) very few individuals are known to exist, (2) the species is widely distributed resulting in low densities, (3) the species has a clumped distribution and/or (4) the species has very low detection rates (elusive behavior, difficult to catch/observe). They are often most negatively affected by environmental perturbation (more specifically human alterations) making conservation and management extremely challenging. The Ivory-billed Woodpecker (*Campephilus principalis*), if extant (Fitzpatrick et al. 2005; Hill et al. 2006; Jackson 2006), may be the most rare and elusive bird species in the United States and thus presents a great challenge for designing efficient and effective surveys. In this dissertation I present results from a large-scale effort to estimate occupancy rates for the Ivory-billed Woodpecker. In addition I used this case study to highlight several important problems and shortfalls common to many studies involving rare species. These shortfalls motivated the development of several new approaches that provide advances in rare species modeling. First, I developed a framework for allocating effort that provides a probabilistic approach to sampling, allowing for improved accuracy in estimating occupancy probability. This approach was found to have a much lower predictive error rate compared to traditional approaches such as single-season occupancy estimation especially when there was a large amount of spatial heterogeneity in habitat and detection probability was low. Second, I developed a hierarchical model that integrates adaptive cluster sampling and occupancy estimation, which allowed for additional effort to be placed at adjacent sites after a known detection. I found this model to outperform traditional occupancy modeling and provide excellent coverage under a variety of conditions. Future improvements in conservation and management of rare species will be accomplished through a variety of techniques and approaches. Ultimately, I believe the most operative approach will be the integration of unique and innovative methods of data collection coupled with models that identify and subsequently estimate the most important vital rates responsible for driving population dynamics.

INDEX WORDS: Adaptive sampling, Bayesian hierarchical modeling, Design-based estimation, Detection probability, Ivory-billed woodpecker, Model-based estimation, Occupancy estimation, Rare species

CONSERVATION AND MANAGEMENT OF RARE SPECIES: THE DEVELOPMENT OF ADAPTIVE MODELS TO REDUCE UNCERTAINTY INFLUENCING DECISION MAKING

by

JAMIAN KRISHNA PACIFICI

B.S., North Carolina State University, 2003

M.S., North Carolina State University, 2007

A Dissertation Submitted to the Graduate Faculty of The University of Georgia in Partial Fulfillment of the Requirements for the Degree

DOCTOR OF PHILOSOPHY

ATHENS, GEORGIA

2011

CONSERVATION AND MANAGEMENT OF RARE SPECIES: THE DEVELOPMENT OF ADAPTIVE MODELS TO REDUCE UNCERTAINTY INFLUENCING DECISION MAKING

by

JAMIAN KRISHNA PACIFICI

Major Professors: Michael J. Conroy
Robert J. Cooper

Committee: John Drake
Nicole Lazar
James Peterson

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
August 2011

DEDICATION

I dedicate this dissertation to my son, Samson.

ACKNOWLEDGEMENTS

I would like to acknowledge my advisors Mike Conroy, and Bob Cooper for all of their support and guidance throughout this process. I would especially like to acknowledge them for all of the different opportunities they have provided. I believe through these diverse opportunities I have truly grown as a scientist. I would also like to acknowledge my committee members, John Drake, Nicole Lazar, and Jim Peterson for all their support and patience. I believe they provided me with a diverse set of skills which continually nurtured my development. I owe a special thanks to John Drake for his financial support during the last year of my dissertation. I also owe a special thanks to Bob Dorazio for his guidance and the time he spent to further develop my quantitative skills while hosting me in Florida. I would like to thank the UGA Research Computing Center and especially Shan-Ho Tsai for the help with implementing thousands of simulations.

I would like to thank all of my fellow graduate students for their support. I have thoroughly enjoyed my time at UGA and will remember all of the laughs we have had. I would like to give a special recognition to my wife Lara, and for everything she does to make our lives special.

TABLE OF CONTENTS

Page

ACKNOWLEDGEMENTS v

LIST OF FIGURES ix

LIST OF TABLES xvi

CHAPTER

1 INTRODUCTION AND LITERATURE REVIEW 1

Chapter Description 8

Literature Cited 10

2 INSIGHTS FROM A LARGE SCALE IVORY-BILLED WOODPECKER (CAMPEPHILUS PRINCIPALIS) SEARCH EFFORT WITH APPLICATIONS TO RARE WIDE-RANDING AVIAN SPECIES 14

Methods 18

Results 23

Discussion 27

Literature Cited 32

3 A TWO-PHASE SAMPLING DESIGN FOR INCREASING DETECTIONS OF RARE SPECIES IN OCCUPANCY SURVEYS 54

Methods...58

Results...65

Discussion 67

Literature Cited 71

4 OCCUPANCY ESTIMATION WITHIN AN ADAPTIVE-SAMPLING DESIGN: EVALUATION OF A BAYESIAN HIERARCHICAL MODEL FOR RARE OR ELUSIVE SPECIES 96

Methods...102

Results...112

Discussion 114

Literature Cited 120

5 CONCLUSION...141

Literature Cited 147

APPENDICES

A R CODE FOR ANALYSES AND SIMULATIONS..148

B ADDITIONAL RESULTS FROM CHAPTER 3 190

C ADDITIONAL RESULTS FROM CHAPTER 4..267

LIST OF FIGURES

Page

Figure 2.1: River basins in the United States within the former range (cross-hatched) of the Ivory-billed woodpecker (*Campephilus principalis*)..38

Figure 2.2: Power of detection for the survey method given that the patch is occupied under three different levels of detection probability and over a range of repeat visits to the patch (*K*) and with 95% confidence limits obtained from the delta method 39

Figure 2.3: Probability of nonoccurrence conditional on no detections at a patch for two levels of occupancy ψ=0.75 (top) and ψ=0.3 (bottom) over a range of number of visits to a patch (*K*) 40

Figure 2.4: Distribution of maximum likelihood estimates obtained from evaluating estimator bias, variance, and mean squared error for a range of true values for occupancy and detection using 10000 simulations...42

Figure 3.1: Habitat covariates (row 1) and associated occupancy data (rows 2 and 3) for three different types of simulated environments..79

Figure 3.2: Plots of test statistic and associated quartiles (25% and 75%) measuring lack of fit (lower is better) for two-phase adaptive approach (circles) and traditional single-season occupancy approach (squares) when n=100, J=3, and for three different levels of detection, p=0.25, 0.5, and 0.75...81

Figure 3.3: Plots of test statistic and associated quartiles (25% and 75%) measuring lack of fit (lower is better) for two-phase adaptive approach (circles) and traditional single-season occupancy approach (squares) when n=200, J=3, and for three different levels of detection, p=0.25, 0.5, and 0.75 83

Figure 3.4: Plots of test statistic and associated quartiles (25% and 75%) measuring lack of fit (lower is better) for two-phase adaptive approach (circles) and traditional single-season occupancy approach (squares) when there is no simulated relationship between habitat and occupancy for n=100, J=3, and for three different levels of detection, p=0.25, 0.5, and 0.75 85

Figure 3.5: Distribution from 1000 simulations of the estimates of *Ntot* for the two-phase adaptive approach (empty dotted triangles) and the traditional single-season occupancy approach (gray filled rectangles) in Habitat 1 (extreme spatial correlation, see text for more details) with the number of repeat visits, J, equal to 3 and detection probability, p, equal to 0.25 87

Figure 4.1: Habitat covariates (row 1) and associated occupancy data (row 2) for three different types of simulated environments 130

Figure 4.2: Plots of the relative root mean squared error (RMSE) comparing three different models: ACSOCC (Adaptive-cluster sampling occupancy, solid lines and squares), SSOCC (Single-season occupancy, dashed lines with circles), and ACS (Adaptive-cluster sampling, dotted lines with triangles) 132

Figure 4.3: Plots of relative bias (RBIAS) comparing three different models: ACSOCC (Adaptive-cluster sampling occupancy, solid lines), SSOCC (Single-season occupancy, dashed lines), and ACS (Adaptive cluster sampling, dotted lines)................................134

Figure 4.4: Total cost for SSOCC and ACSOCC models under a range of scenarios 136

Figure B.1: Plots of test statistic and associated quartiles (25% and 75%) measuring lack of fit (lower is better) for two-phase adaptive approach (circles) and traditional single-season occupancy approach (squares) when n=100, J=5, and for three different levels of detection, p=0.25, 0.5, and 0.75..192

Figure B.2: Plots of test statistic and associated quartiles (25% and 75%) measuring lack of fit (lower is better) for two-phase adaptive approach (circles) and traditional single-season occupancy approach (squares) when n=200, J=5, and for three different levels of detection, p=0.25, 0.5, and 0.75..194

Figure B.3: Plots of test statistic and associated quartiles (25% and 75%) measuring lack of fit (lower is better) for two-phase adaptive approach (circles) and traditional single-season occupancy approach (squares) when there is no simulated relationship between habitat and occupancy for n=100, J=5, and for three different levels of detection, p=0.25, 0.5, and 0.75..196

Figure B.4: Plots of test statistic and associated quartiles (25% and 75%) measuring lack of fit (lower is better) for two-phase adaptive approach (circles) and traditional single-season occupancy approach (squares) when there is no simulated relationship between habitat and occupancy for n=200, J=3, and for three different levels of detection, p=0.25, 0.5, and 0.75..198

Figure B.5: Plots of test statistic and associated quartiles (25% and 75%) measuring lack of fit (lower is better) for two-phase adaptive approach (circles) and traditional single-season occupancy approach (squares) when there is no simulated relationship between habitat and occupancy for n=200, J=5, and for three different levels of detection, p=0.25, 0.5, and 0.75 200

Figure B.6: Distribution from 1000 simulations of the estimates of Ntot for the two-phase adaptive approach (empty dotted rectangles) and the traditional single-season occupancy approach (gray filled rectangles) in Habitat 1 (Extreme spatial correlation, see text for more details) with the number of repeat visits, J, equal to 3 and detection probability, p, equal to 0.75 202

Figure B.7: Distribution from 1000 simulations of the estimates of N^{tot} for the two-phase adaptive approach (empty dotted rectangles) and the traditional single-season occupancy approach (gray filled rectangles) in Habitat 2 (Moderate spatial correlation, see text for more details) with the number of repeat visits, J, equal to 3 and detection probability, p, equal to 0.25 204

Figure B.8: Distribution from 1000 simulations of the estimates of N^{tot} for the two-phase adaptive approach (empty dotted rectangles) and the traditional single-season occupancy approach (gray filled rectangles) in Habitat 2 (Moderate spatial correlation, see text for more details) with the number of repeat visits, J, equal to 3 and detection probability, p, equal to 0.75 206

Figure B.9: Distribution from 1000 simulations of the estimates of N^{tot} for the two-phase adaptive approach (empty dotted rectangles) and the traditional single-season occupancy approach (gray filled rectangles) in Habitat 3 (No spatial correlation, see text for more details) with the number of repeat visits, *J*, equal to 3 and detection probability, *p*, equal to 0.25 208

Figure B.10: Distribution from 1000 simulations of the estimates of N^{tot} for the two-phase adaptive approach (empty dotted rectangles) and the traditional single-season occupancy approach (gray filled rectangles) in Habitat 3 (No spatial correlation, see text for more details) with the number of repeat visits, *J*, equal to 3 and detection probability, *p*, equal to 0.75 210

Figure B.11: Distribution from 1000 simulations of the bias in estimates of N^{tot} for the two-phase adaptive approach (empty dotted rectangles) and the traditional single-season occupancy approach (gray filled rectangles) in Habitat 1 (Extreme spatial correlation, see text for more details) with the number of repeat visits, *J*, equal to 3 and detection probability, *p*, equal to 0.25....................212

Figure B.12: Distribution from 1000 simulations of the bias in estimates of N^{tot} for the two-phase adaptive approach (empty dotted rectangles) and the traditional single-season occupancy approach (gray filled rectangles) in Habitat 1 (Extreme spatial correlation, see text for more details) with the number of repeat visits, *J*, equal to 3 and detection probability, *p*, equal to 0.75....................214

Figure B.13: Distribution from 1000 simulations of the bias in estimates of N^{tot} for the two-phase adaptive approach (empty dotted rectangles) and the traditional single-season occupancy approach (gray filled rectangles) in Habitat 2 (Moderate spatial correlation, see text for more details) with the number of repeat visits, *J*, equal to 3 and detection probability, *p*, equal to 0.25 216

Figure B.14: Distribution from 1000 simulations of the bias in estimates of N^{tot} for the two-phase adaptive approach (empty dotted rectangles) and the traditional single-season occupancy approach (gray filled rectangles) in Habitat 2 (Moderate spatial correlation, see text for more details) with the number of repeat visits, *J*, equal to 3 and detection probability, *p*, equal to 0.75 218

Figure B.15: Distribution from 1000 simulations of the bias in estimates of N^{tot} for the two-phase adaptive approach (empty dotted rectangles) and the traditional single-season occupancy approach (gray filled rectangles) in Habitat 3 (No spatial correlation, see text for more details) with the number of repeat visits, *J*, equal to 3 and detection probability, *p*, equal to 0.25 220

Figure B.16: Distribution from 1000 simulations of the bias in estimates of N^{tot} for the two-phase adaptive approach (empty dotted rectangles) and the traditional single-season occupancy approach (gray filled rectangles) in Habitat 3 (No spatial correlation, see text for more details) with the number of repeat visits, *J*, equal to 3 and detection probability, *p*, equal to 0.75 222

Figure C.1: Plots of the relative root mean squared error (RMSE) comparing three different models: ACSOCC (Adaptive-cluster sampling occupancy, solid lines and squares), SSOCC (Single-season occupancy, dashed lines with circles), and ACS (Adaptive-cluster sampling, dotted lines with triangles) 269

Figure C.2: Plots of the relative bias (RBIAS) comparing three different models: ACSOCC (Adaptive-cluster sampling occupancy, solid lines), SSOCC (Single-season occupancy, dashed lines), and ACS (Adaptive-cluster sampling, dotted lines) 271

LIST OF TABLES

Page

Table 2.1: Information about the survey design and detection data collected from 2007-2008 for the Ivory-billed woodpecker (*Campephilus principalis*) across its former range in the United States 43

Table 2.2: The deviance (-2*Loglikelihood), number of parameters, Akaike's Information Criteria adjusted for small sample size (AIC_c), the difference between the model and the top model (Delta AIC_c), model weights (*w*), and likelihood of candidate detection (p) and occupancy (*psi*) models for the Ivory-billed woodpecker (*Campephilus principalis*) in two states (SC and FL)..44

Table 2.3: Model-averaged real parameter estimates at the mean value of the patch-level covariate and model-averaged beta coefficients obtained from the candidate set of 16 models (>10% of the weight of the top model) 46

Table 2.4: Total number of occupied patches out of 595 total patches (N^{occ}), number of repeat visits (K), occupancy probability conditional on no detections at a patch ($\psi condl$), and occupancy rate or proportion of patches occupied out of 595 (ψ^{fs}) and their estimated variances following Royle and Dorazio (2008)...47

Table 2.5: Results from conducting 10,000 simulations to evaluate bias, variance, and mean squared error loss (MSE) for both occupancy and detection under a range of scenarios

characterized by the number of visits to a patch (*K*), the number of patches, true occupancy probability (0.3, 0.75, 0.9; lower confidence bound, mean, and upper confidence bound of maximum likelihood estimate), and a fixed detection probability (0.046; mean of maximum likelihood estimate) 49

Table 2.6: The optimal design given the true occupancy probability (0.3, 0.75, 0.9; lower confidence bound, mean, and upper confidence bound of maximum likelihood estimate), and a fixed detection probability (0.046; mean of maximum likelihood estimate) 51

Table 3.1: Average value from 1000 simulations of test statistic measuring lack of fit, $\boldsymbol{d}$, test statistic measuring goodness of fit, ***GOF***, estimate of *Ntot*, bias associated in estimating *Ntot*, and mean squared error, ***MSE***, associated with estimating *Ntot* 88

Table 3.2: Average value from 1000 simulations of test statistic measuring lack of fit, $\boldsymbol{d}$, test statistic measuring goodness of fit, ***GOF***, estimate of *Ntot*, bias associated in estimating *Ntot*, and mean squared error, ***MSE***, associated with estimating *Ntot* 90

Table 3.3: Average value from 1000 simulations for estimates of coefficients in the logit-linear model for occupancy probability, $\widehat{\boldsymbol{\beta}_0}$ and $\widehat{\boldsymbol{\beta}_1}$, along with average bias and mean squared error (**MSE**) 92

Table 3.4: Average value from 1000 simulations for estimates of coefficients in the logit-linear model for occupancy probability, $\widehat{\boldsymbol{\beta}_0}$ and $\widehat{\boldsymbol{\beta}_1}$, along with average bias and mean squared error (**MSE**) 94

Table 4.1: Simulation results from 500 synthetic datasets with difference design criteria comparing three different models: ACSOCC (Adaptive-cluster sampling occupancy), SSOCC (Single-season occupancy), and ACS (Adaptive-cluster sampling) 137

Table B.1: Average value from 1000 simulations of test statistic measuring lack of fit, ***d***, test statistic measuring goodness of fit, ***GOF***, estimate of N^{tot}, bias associated in estimating N^{tot}, and mean squared error, ***MSE***, associated with estimating N^{tot} 223

Table B.2: Average value from 1000 simulations of test statistic measuring lack of fit, ***d***, test statistic measuring goodness of fit, ***GOF***, estimate of N^{tot}, bias associated in estimating N^{tot}, and mean squared error, ***MSE***, associated with estimating N^{tot} 234

Table B.3: Average value from 1000 simulations of test statistic measuring lack of fit, ***d***, test statistic measuring goodness of fit, ***GOF***, estimate of N^{tot}, bias associated in estimating N^{tot}, and mean squared error, ***MSE***, associated with estimating N^{tot} 245

Table B.4: Average value from 1000 simulations of test statistic measuring lack of fit, ***d***, test statistic measuring goodness of fit, ***GOF***, estimate of N^{tot}, bias associated in estimating N^{tot}, and mean squared error, ***MSE***, associated with estimating N^{tot} 256

Table C.1: Simulation results from 500 synthetic datasets with different design criteria for the ACSOCC model (Adaptive-cluster sampling occupancy) 272

Table C.2: Simulation results from 500 synthetic datasets with different design criteria for the SSOCC model (Single-season occupancy)..287

Table C.3: Simulation results from 500 synthetic datasets with different design criteria for the ACS model (Adaptive-cluster sampling using the modified Horvitz-Thompson estimator) 302

Table C.4: Simulation results from 500 synthetic datasets with different design criteria for the ACS model (Adaptive-cluster sampling using the modified Hansen-Hurwitz estimator) 317

CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

Conservation and management of rare species is an important component to maintaining ecosystem health and community dynamics. Species are classified as rare for several reasons: (1) very few individuals known to exist, (2) widely distributed resulting in low densities, and/or (3) very low detection rates (elusive behavior, difficult to catch/observe). They are often most negatively affected by environmental perturbation (more specifically human alterations) making conservation and management extremely challenging. Rare species are often simultaneously the species for which strong inference about state variables and vital rates are most needed and the species for which such information is most difficult to obtain (MacKenzie et al. 2005). Difficulties arise when designing surveys for rare species because obtaining adequate samples of information to be used in analysis can be demanding. Often a very large proportion of zeros or non-detections exist in the dataset creating difficulty not only in parameter estimation (few data, convergence issues), but interpreting the non-detection as a true absence. Unfortunately there is often a lack of thought devoted to fundamental questions associated with sound sampling programs (Yoccoz et al. 2001), which adds to the difficulty in conserving and managing rare or elusive species.

Decisions about the management and conservation of rare or threatened species are made in the face of considerable uncertainty. This uncertainty arises from a lack of knowledge about the populations themselves and the dynamic processes driving those populations. Often basic state variables of interest such as population size and occupancy are not known, which are critical to help managers and conservationists make appropriate decisions. This uncertainty is exacerbated by the poor understanding of the influence of other environmental variables on population dynamics. Ultimately, decisions will need to be made regardless of the quantity or quality of the information available.

Formal decision making frameworks such as Structured Decision Making (SDM; Lindley 1985; Clemen 1996) and Adaptive Resource Management (ARM; Williams et al. 2002; Moore and Conroy 2006) provide a foundation to evaluate and optimize the decision process in an integrated framework. Elements of the decision process include stating objectives, specifying decision alternatives, recognizing consequences or outcomes, identifying models that describe how we think the decision will influence outcomes, and a monitoring program to follow the system's evolution and response to management. Each decision opportunity relies on explicitly stating and differentiating these elements of the process.

The use of a formal decision making process allows for the reduction of uncertainty involved in making that decision. In reality there are multiple types of uncertainty influencing each decision. Environmental stochasticity involves the uncertainty related to environmental factors beyond the control of the decision maker leading to stochastic or non-deterministic outcomes. For example, a drought year could potentially have a severe impact on an expected outcome. Partial controllability is the uncertainty associated with the realization of a decision. For example, 100 ha of forest are proposed to be burned, but instead 150 ha are actually burned. Because we rarely if ever observe the true state of the system and instead rely on a sample of the population, statistical uncertainty corrupts our ability to effectively determine current conditions and evaluate the results of our conservation actions. Statistical uncertainty manifests itself in estimates of the parameters or variables of interest and can lead to bias, imprecision or both. Finally, structural uncertainty presents itself in the underlying assumptions about how a system will respond to our decision. Structural uncertainty is akin to having very little information and insight into how a population is influenced by its environment, common practice for rare or threatened species. Therefore the response of the population to changes in the environment is poorly understood.

The goal of decision making is ultimately to make the best or optimal decision in light of multiple sources of uncertainty. Therefore one nested goal is to potentially reduce the uncertainty associated with each decision. One approach then is to gain as much information about the system and species so that structural and statistical uncertainties are minimized while accounting for environmental stochasticity. This can be accomplished by using appropriate methods of survey design, estimation and modeling to reduce statistical uncertainty and to understand the basic relationships that drive population dynamics to reduce structural uncertainty. This is often the most critical step to effective decision making and therefore is the main focus of this dissertation.

Rarely are ecologists or wildlife biologists able to collect a complete enumeration of animals present (i.e. census) over some specified area. Instead a sample, collected over space or time, must be used that relates information from the collected sample to the larger population of interest. Although this is at the foundation of statistics (statistical inference), the relationship between the evidence in the sample of animals and the connection to the larger population of animals is often times very difficult to infer. This is due in part to two major sources of variation, 1) imperfect detection of animals and, 2) spatial coverage of the sample. Parameters of interest such as density, survival rate, reproduction rate, or colonization/extinction rates are therefore very difficult to estimate because we rarely observe all of the animals in the population and we rarely collect data over the entire spatial extent of that population.

The problem of sampling becomes even more difficult when discussing species that are described as rare or elusive. These animals are usually of great concern to managers and conservationists because there is little information about population sizes and how they respond to environmental perturbation (including human induced changes) and ultimately are at higher risk (Thompson 2004). Sampling is difficult because these animals exhibit specific characteristics that reduce the ability to collect a sufficient amount of information to estimate parameters of interest. These characteristics include being hard to capture or observe, occurring in low numbers, being patchily distributed, and having low detection rates (see Chapter 2 *in* Thompson 2004 for lengthy discussion). Even species that have a large total population size may occur sparsely over a very large area making sampling very difficult. It is therefore necessary to devote effort to developing specific modeling and estimation techniques that can apply to these unique and difficult circumstances to reduce uncertainty.

Current approaches to the estimation of finite population parameters can broadly be classified into two main categories: design-based approach and model-based approach. A third distinction can be made which separates Bayesian inference from a likelihood model-based approach. Other authors have somewhat similar distinctions, for instance Rubin (1976) roughly outlines three approaches to statistical inference: 1) sampling-distribution inference (similar to design-based), 2) direct-likelihood inference (comparison of ratios of the likelihood function for the various values of the parameters of interest), and 3) Bayesian inference (inference based solely from posterior distributions corresponding to specified prior distributions). Thompson and

Seber (1996) make three distinctions as well: 1) design-based, 2) frequentist model-based, and 3) likelihood model-based (includes Bayesian inference).

In the design-based approach probability only enters through the use of design-induced probabilities to select one sample over another. Nothing is assumed about the underlying population and inference is only based on hypothetical repetition of selecting sample units. The frequentist model-based approach suggests that the values of the variable of interest from the population are viewed as a realization of a set of random variables. A "superpopulation model" (stochastic model) is assumed describing the distribution of possible realizations of the population values. Inference procedures are based on having good properties over a hypothetically repeated realization of population values. In the likelihood model-based approach inference is based on the likelihood functions of the unknowns given the sample data. The Bayesian approach extends this by allowing for an assignment of subjective prior distributions on the population or its parameters, thus inference is based on the posterior distribution which is a combination of the priors and likelihood function. I felt these categories encompass and elucidate the main differences among the current approaches, although they are neither mutually exclusive nor independent. I will therefore mainly make the distinction between design-based and model-based inference at this stage.

Current methods that account for variation in detection such as capture-recapture (Seber 1982; Williams et al. 2002) and occupancy estimation (MacKenzie et al. 2006) are two examples of model-based approaches that have been used extensively to estimate parameters of interest for animal populations. Design-based approaches have also seen widespread use especially in relation to rare species (e.g. Smith et al. 2003; Thompson 2004), but often do not allow for incorporation of imperfect detection. Bayesian approaches are relatively newer in their application, but have already made an impact in the wildlife literature (Link et al. 2002) and methods involving Bayesian inference that account for imperfect detection continue to evolve (e.g. see Royle and Dorazio 2008).

Difficulty in estimating parameters of interest for rare or threatened species can be due to a lack of data and very low detection rates. Methods such as Bayesian hierarchical modeling (e.g. Royle and Kery 2007), which permit the leveraging of information at larger scales to better estimate smaller scale parameters, can be of great use when there is very little data at scales of interest. Other techniques such as the use of spatial modeling to accommodate spatial dependencies in the data can be used to more accurately predict quantities of interest (e.g. Webster et al. 2008). Techniques from other fields including the use of small area estimation methods (Rao 2003) from survey sampling, which were developed to estimate parameters of interest when very little or no data were collected from the area of interest, seem a natural fit to the problem of estimation for rare or threatened species. Small area methods improve upon direct design-based estimators by including suitable linking models with other sources of data, for example collected covariates or in our case data from similar species.

An important avenue of research brought forth by small area estimation is the investigation of combining design-based and model-based (both likelihood and Bayesian) inference. Often specific sampling designs (e.g. adaptive cluster sampling, stratified sampling, disproportionate sampling, systematic sampling, sequential sampling) have been favored for rare or elusive species (see Part Two *of* Thompson 2004), but rarely are they able to account for imperfect detection without the use of independent estimates of detectability (see Chapter 9 *of* Thompson and Seber 1996). The use of these designs is obvious because of the geographically clustered nature of many rare populations (Christman 2000), but a second use of these specific designs is to accommodate observer behavior. Many rare or threatened species are so rare or endangered that any information about them is extremely important (e.g. Ivory-billed Woodpecker). Therefore the ability to mimic observer behavior by putting more effort in areas where individuals have been detected is a potentially important component of an effective sampling design. Thus, integrating the sampling design and model-based inference which allows for the incorporation of imperfect detection and spatial dependency is potentially optimal.

In Bayesian inference the incorporation of data collection or design within the modeling framework is favored by some (e.g. see Chapter 7 *in* Gelman et al. 2004) alluding to the complete definition of observed data including how the observed values arose. Acknowledging that the complete definition of observed data does indeed include information on how the observed values arose has direct bearing on inference. Therefore it is necessary to incorporate all of the information regarding the data (observed values and data collection process) in the probability model used for analysis. This general view of the problem allows for the separation of “observed data” and “missing data” (together make “complete data”) in which inference is conditional on observed data and also on the pattern of observed and missing data. Missing data can include unintentional missing data due to unfortunate circumstance and intentional missing data such as data from units not sampled in a survey. Data collected for rare species usually consists of a large proportion of “missing” data thus this mode of inference seems appropriate when attempting to marriage design-based and model-based inference while accounting for imperfect detection.

Chapter Description

Chapter 2 –

The Ivory-billed Woodpecker (IBWO), if extant (Fitzpatrick et al. 2005; Hill et al. 2006; Jackson 2006), may be the most rare and elusive bird species in the United States and thus presents a great challenge for designing efficient and effective surveys. The species once existed at low densities in the southeastern U.S. from Florida to Texas and as far north as Illinois and Indiana and is thought to have used extensive forested areas with very large trees and many dead trees (Jackson 2002). In 1938, Tanner (1942) took the last universally accepted photograph of this species in the U.S.; however, intriguing sightings continued throughout the 20th century (Jackson 2002; Fitzpatrick et al. 2005; Hill et al. 2006). Recent evidence that the Ivory-billed Woodpecker persists in both Arkansas (Fitzpatrick et al. 2005) and Florida (Hill et al. 2006) has reinvigorated the hope that this species can be saved from extinction. The putative rediscovery of the Ivory-billed Woodpecker in the Cache-lower White River Basins initiated a new search effort. The primary objective of the search has been to find the bird and document its existence, mostly searching only those locations that were believed to be optimal based mostly on the limited data provided by Tanner (1942).

The objective of Chapter 2 is to present a study design for the new search effort for IBWO that permits occupancy and analyze data collected under this design. In addition I focus on evaluating the evidence obtained from the survey in relation to the effort expended. I discuss complications that arose during the study and are representative of many large-scale surveys for rare avian species. These complications provide suggestions to improve future studies where many logistical constraints exist and provide motivation for Chapters 3 and 4.

Chapter 3 –

One of the main complications that arose from the IBWO survey was the inefficient use of resources specifically related to allocating effort for occupancy-based surveys. Therefore in Chapter 3 I develop a framework for efficiently allocating effort that profits from placing more effort in areas with higher probability of occupancy. This allows for areas with high predicted probability of occupancy to have a higher inclusion probability for sampling thus more effort can be concentrated in these areas. I compare this approach with the traditional simple-random sample associated with occupancy estimation and use simulations to evaluate the predictive performance of each approach.

Chapter 4 –

Again motivated by the findings in Chapter 2, I develop a model that allows for the augmentation of traditional occupancy estimation to allow for additional effort to be allocated to adjacent sites once a detection has occurred. I create a statistically rigorous approach that integrates adaptive cluster sampling and occupancy estimation and permits inference under maximum likelihood or Bayesian flavors. I focus on the development of a Bayesian hierarchical model and assess model fit using a Bayesian p-value, evaluate the frequentist properties of the model under a range of design scenarios and compare the performance of the new model to traditional occupancy estimation and adaptive cluster sampling.

Chapter 5 –

I provide a synthesis and conclusion of the previous chapters. In addition I highlight major contributions of each chapter and discuss future research needs.

Literature Cited

Christman, M. C. 2000. A review of quadrat-based sampling of rare, geographically clustered populations. Journal of Agricultural, Biological, and Environmental Statistics 5: 168-201.

Clemen, R. T. 1996. Making Hard Decisions. Duxbury. Belmont, CA.

Fitzpatrick, J. W., M. Lammertink, M. D. Luneau, T. W. Gallagher, B. R. Harrison, G. M. Sparling, K. V. Rosenberg, R. W. Rohrbaugh, E. C. H. Swarthout, P. H. Wrege, S. B. Swarthout, M. S. Dantzker, R. A. Charif, T. R. Barksdale, J. V. Remsen, S. D. Simon, and D. Zollner. 2005. Ivory-billed woodpecker (Campephilus principalis) persists in continental North America. Science 308:1460-1462.

Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin. 2004. Bayesian Data Analysis. Chapman and Hall. Boca Raton, FL.

Hill, G. E., D. J. Mennill, B. W. Rolek, T. L. Hicks, and K. A. Swiston. 2006. Evidence Suggesting that Ivory-billed Woodpeckers (Campephilus principalis) Exist in Florida. Avian Conservation and Ecology - Écologie et conservation des oiseaux 1:URL: http://www.ace-eco.org/vol1/iss3/art2/.

Jackson, J. A. 2002. Ivory-billed Woodpecker: Campephilus Principalis. Birds of North America, Inc.

Jackson, J. A. 2006. The public perception of science and reported confirmation of the Ivory-billed Woodpecker in Arkansas. Auk 123:1185-1189.

Lindley, D. V. 1985. Making Decisions. Wiley. London, England.

Link, W. A., E. Cam, J. D. Nichols, and E. G. Cooch. 2002. Of bugs and birds: Markov chain Monte carlo for hierarchical modeling in wildlife research. Journal of Wildlife Management 66: 277-291.

Little, R. A. 2006. Calibrated Bayes: A Bayes/Frequentist roadmap. The American Statistician 60:213-223.

MacKenzie, D.I., J.D. Nichols, N.S. Sutton, K. Kawanishi, and L.L. Bailey. 2005. Improving inferences in population studies of rare species that are detected imperfectly. Ecology 86: 1101-1113.

MacKenzie, D. I., J. D. Nichols, J. A. Royle, K. H. Pollock, L. L. Bailey, and J. E. Hines. 2006. Occupancy Estimation and Modeling. Elsevier-Academic. San Diego, CA.

Moore, C. T., and M. J. Conroy. 2006. Optimal regeneration planning for old-growth forest: addressing scientific uncertainty in endangered species recovery through adaptive management. Forest Science 52: 155-172.

Rao, J. N. K. 2003. Small Area Estimation. Wiley. Hoboken, New Jersey.

Royle, J. A. and M. Kery. 2007. A Bayesian state-space formulation of dynamic occupancy models. Ecology 88: 1813-1823.

Royle, J. A., and R. M. Dorazio. 2008. Hierarchical Modeling and Inference in Ecology. Elsevier-Academic. San Diego, CA.

Rubin, D. B. 1976. Inference and missing data. Biometrika 63: 581-592.

Rubin, D. B. 1984. Bayesianly justifiable and relevant frequency calculations for the applied statistician. The Annals of Statistics 12:1151-1172.

Seber, G. A. F. The Estimation of Animal Abundance. The Blackburn Press. Caldwell, New Jersey.

Smith, D. R., R. F. Villella, and D. P. Lemarie. 2003. Application of adaptive cluster sampling to low-density populations of freshwater mussels. Environmental and Ecological Statistics 10: 7-15.

Tanner, J. T. 1942. The Ivory-Billed Woodpecker. Dover Publications, Inc., Mineola.

Thompson, S. K., and G. A. F. Seber. 1996. Adaptive Sampling. Wiley. New York, NY.

Thompson, W. L. ed. 2004. Sampling Rare or Elusive Species. Island Press. Washington, D.C.

Webster, R. A., and K. H. Pollock. 2008. Bayesian spatial modeling of data from unit-count surveys of fish in streams. Transactions of the American Fisheries Society 137: 438-453.

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

Yoccoz, N. G., J. D. Nichols, and T. Boulinier. 2001. Monitoring of biological diversity in space and time. Trends in Ecology and Evolution 16:446-453.

CHAPTER 2

INSIGHTS FROM A LARGE SCALE IVORY-BILLED WOODPECKER (CAMPEPHILUS PRINCIPALIS) SEARCH EFFORT WITH APPLICATIONS TO RARE WIDE-RANGING AVIAN SPECIES[1]

---

[1]Pacifici, K., M. J. Conroy, R. J. Cooper, J. T. Peterson, and R. S. Mordecai. To be submitted to *Avian Conservation and Ecology*.

Abstract

Rare or elusive species present many conservation challenges, particularly when species occur over a very large range as in many rare avian species. Recently, a large scale survey was implemented to obtain information about the Ivory-billed Woodpecker (*Campephilus principalis*), previously thought to be extinct. The objectives of the survey were to document the species' existence and describe habitat relationships and overall distribution. Of 595 patches included in the historical range of the bird, 180 patches were surveyed, resulting in 28 total putative detections of evidence of Ivory-billed Woodpecker presence with 27 occurring within two river basins: Congaree Swamp, South Carolina and Lower Choctawhatchee, Florida. Detection probabilities were estimated at 0.045 ($\pm$ 0.012) for South Carolina and 0.047 ($\pm$ 0.016) for Florida. Occupancy probabilities were estimated at 0.73 ($\pm$ 0.199) for South Carolina and 0.77 ($\pm$ 0.249) for Florida. We found no relationship between occupancy or detection with two habitat variables, number of snags (i.e. dead trees) and number of big trees (> 24'' dbh). Model-averaged estimates for these covariates had very large standard errors (CVs > 100%) suggesting no effect on detection or occupancy. The power to detect occurrence given our estimates of detection ranged from 0.55 to 1 for a survey with 40 visits per patch. Regardless of whether or not the Ivory-billed Woodpecker actually exists, we believe that our implementation of a single-season occupancy model assisted by the use of historical information provides a practical and scientifically rigorous framework for monitoring many wide-ranging avian species.

Conservation and management of rare species is one of the most challenging tasks confronting natural resource managers. Collecting data is difficult because these animals have specific characteristics that reduce the ability to accumulate a sufficient amount of information to estimate parameters of interest. Characteristics include elusiveness, low densities, and patchy distribution, often resulting in low detection probability (see Chapter 2 in Thompson 2004 for lengthy discussion). Even species that have a large total population size may occur sparsely over a very large area making sampling very difficult. Often this is the case for avian species that can occur over a wide range creating problems with sampling a very large area (Pollock et al. 2002). Unfortunately, rare species are often simultaneously the species for which accurate estimates of population status and vital rates are most needed and the species for which such information is most difficult to obtain (MacKenzie et al. 2005).

The Ivory-billed Woodpecker (IBWO; *Campephilus principalis*), if extant (Fitzpatrick et al. 2005; Hill et al. 2006; Jackson 2006), may be the rarest and most elusive bird species in the United States. Consequently, designing efficient and effective surveys for this species presents a great challenge. The species once existed at low densities in the southeastern U.S. from Florida to Texas and as far north as Illinois and Indiana and is thought to have used extensive forested areas with very large trees and many dead trees (Jackson 2002). Recent evidence that the IBWO persists in both Arkansas (Fitzpatrick et al. 2005) and Florida (Hill et al. 2006) has reinvigorated the hope that this species can be saved from extinction and resulted in new search efforts beginning in 2006 (Fitzpatrick et al. 2005). The primary objective of these searches was to find the bird and document its existence. Locations that were believed to contain optimal habitat based on the limited data provided by Tanner (1942) were included in the search effort. Meanwhile, searches were initiated in locations within the former range where unsubstantiated sightings were reported in recent decades, again focusing on places believed to have the best chance of being occupied. Some evidence of IBWO presence and other promising data were collected in these places, but there also was evidence of presence in areas that were not consistent with prior expectations of habitat affinities (i.e., smaller tracts, few large trees and snags). The inability to obtain a photograph or other definitive evidence led to a new search effort that was based on probabilistic survey sampling methodology (Thompson 1992).

Our goals were to describe the design and application of the new search effort for IBWO and analyze data collected with the design. Our specific objectives of analysis were to: estimate occupancy probability (probability of site being occupied) and the occupancy rate ($psi^{(fs)}$; proportion of area occupied), while accounting for imperfect detection; assess relationships between occupancy and habitat characteristics; and assess design characteristics (bias, variance, mean squared error) and tradeoffs of the current design to suggest improvements for future studies involving rare wide-ranging avian species.

Recently much debate has arisen about the validity of the "observations" and we recognize the controversial nature of these data and whether or not they actually constitute IBWO observations (Elphick et al. 2010, Fitzpatrick et al. 2006, Jackson 2006, Jackson 2006b, Jones et al. 2007, McKelvey et al. 2008, Roberts et al. 2009,). To avoid using cumbersome terminology such as "putative observation," and to avoid confusion with standard wildlife survey terminology such as observation and detection, we use these latter terms with the understanding that they do not constitute confirmed detections, but were instead sounds or other evidence that may or may not have actually been IBWO.

Methods

Study Design

In 2006, a two-stage (following Thompson 1992) occupancy-based survey design occurring at two spatial scales, a primary level and a secondary level, was created for the IBWO as follows. Consider the population of spatial units as composed of *N* primary units from which we take a random sample of *n*. At the second stage, we take an initial simple random sample of m_i units without replacement from the primary unit *i* for $i = 1 \ldots n$ associated with the *jth* secondary unit of the *ith* primary unit is a variable of interest y_{ij}. In this case, y_{ij} represents detection of an IBWO from a presence-absence survey with *k* visits. The *N* primary units are individual river basins within the former range of the IBWO (Figure 2.1). Many of those were eliminated from further consideration due to their (believed) complete lack of suitability. River basins with consistent sightings and/or sound recordings (i.e., high quality evidence) were always selected to survey. Those were the Cache/lower White in Arkansas, the Choctawhatchee in Florida, and the Congaree/Wateree in South Carolina. Other river basins were also selected non-randomly based on recent reported sightings. Remaining basins in the sampling frame were randomly selected with weights based on the subjective probability of IBWO occurring in the area. These selection weights resulted in basins with an assumed high occupancy probability being frequently selected and those with assumed low occupancy probability being rarely selected.

The *M* secondary units were defined as approximately 2-km^2 patches of land within the selected river basins. The 2-km^2 size was chosen because it is currently in use as part of the Lower Mississippi Valley Joint Venture habitat survey (LMVJV refers to these patches as stands, which are subunits of management compartments on public land in the survey). These patches were generally squares or a similar shape on a grid, but were occasionally modified to follow existing features of the landscape, such as water features or management compartments. Patches that were inaccessible due to logistics or landowner permission were omitted from the sampling frame. Patches were randomly selected with weights based on the perceived (by the investigator) probability of IBWO use.

Habitat surveys

We followed the Lower Mississippi Valley Joint Venture habitat measurement protocol (http://www.lmvjv.org/IBWO_habitat_inventory_&_assessment.htm), which included measurements on 4 transects of 5 plots each, or n=20 plots per patch of density of large (>24” dbh) trees, density of snags (dead trees), and the diameter class of dominant tree species within a 16-m radius. Habitat surveys were only done once for each patch unless in the view of the survey team the patch had undergone significant change since the last survey.

Data and Modeling

Data were collected over seventeen months (1/2007 – 5/2008) and consisted of detection data suitable for occupancy analysis (MacKenzie et al. 2006). Following advice from MacKenzie et al. (2005) we treated the seventeen months as a “single-season”, which allowed us to borrow information temporally across the entire study to estimate parameters. All detections were auditory and consisted of two commonly described sounds produced by IBWO: a kent call and a double knock (Jackson 2002). In 1935, Arthur Allen made the only known recording of a kent call (Allen and Kellogg 1937), a call often described as sounding like a toy trumpet or clarinet. There is no known recording of the IBWO double knock; however, historic descriptions of “double resounding whacks” produced by IBWO (Allen and Kellogg 1937) agree well with double knocks produced by other woodpeckers in the genus *Campephilus* (Ron Rohrbaugh, Jr. *unpubl. data*). Although there is some dispute about the validity of the “double-knocks” as being unique to IBWO (Jones et al. 2007), we again treated these data as valid “detections” for the analysis.

All detections occurred within three river basins (Figure 2.1 and Table 2.1; Congaree Swamp in SC, Lower Choctawhatchee in FL, and Lower Trinity in TX), but the third river basin (TX) only contained a single detection. Instead of modeling occupancy and detection across all 595 patches, we modeled occupancy and detection conditional on the river basins containing information. This results in biased high estimates of detection and occupancy; we acknowledge this bias and discuss the implications further in the discussion. Furthermore, because only a single detection occurred in the Lower Trinity river basin in Texas, we excluded it from the analysis and focused on the two river basins in South Carolina and Florida.

We used a single-season occupancy model in Program MARK (White and Burnham 1999) to fit several models to examine heterogeneity in detection and occupancy along with covariate relationships. We treated SC and FL as separate groups because of the slightly different habitat characteristics in each river basin, but we examined the utility of borrowing information across groups in an attempt to estimate parameters more precisely (MacKenzie et al. 2005). Because of the limited amount of information contained in the data (27 total detections over 17 month span, see Table 2.1) we were only able to fit a simple set of models that included variation in detection and occupancy as a function of number of snags, number of big trees and both in combination in an additive logit-linear model for each of the groups and for both groups together (resulted in 25 total models). We used Akaike's Information Criterion adjusted for small sample size (AIC_c) to conduct model selection (Burnham and Anderson 2002). We then produced model-averaged estimates of detection and occupancy to be used for inference (Burnham and Anderson 2002). Unfortunately, we were not able to conduct a Goodness-of-Fit Test (MacKenzie and Bailey 2004) due to a large number of missing values created by variability in the number of visits to particular patches (Table 2.1). MacKenzie and Bailey (2004) suggest treating patches (referred to as sites) with different combinations of missing observations as separate cohorts, but this would require >15 cohorts (>15 different numbers of visits to patches) which would result in < 2 detections per cohort in many cases.

We calculated the power of the survey design (p*), $\Pr(y>0|z=1)$ (Royle and Dorazio 2008; probability of at least one detection given that the patch is occupied, where y is the detection and z is the "latent" or unobserved occupancy state) and estimated the variance of this quantity using the delta method (Seber 1982). We also estimated ψ_{condl} (MacKenzie et al. 2006), or the probability of occupancy conditional on no detections having occurred at that patch, and estimated the variance of ψ_{condl} using the bootstrap method (Efron and Tibshirani 1993) with 1,000,000 iterations. We predicted ψ_{condl} and related quantities for other values of occupancy (e.g. lower confidence bound of maximum likelihood estimate) and calculated the variance of these quantities by back calculating a variance assuming the estimate has the same coefficient of variation as the original estimate of occupancy. We estimated ψ^{fs} (finite-sample occupancy rate, or proportion of area occupied), and N_{occ} , the number of occupied patches in the total study area, along with a variance of these quantities using the approach described by Royle and Dorazio (2008) for likelihood-based inference assuming constant homogeneous detection (also see

Dorazio and Royle 2005 for a more complicated model which resorts to empirical Bayes estimators). Alternatively one could estimate this quantity in a more intuitive manner by resorting to Bayesian inference (see Royle and Kery 2007 for an example).

We evaluated estimator performance under a range of scenarios concentrated around our observed results. We estimated bias, variance and Mean Squared Error (MSE) for the probability of detection and probability of occupancy when true occupancy covered the estimated range (lower confidence bound to upper confidence bound), detection probability was fixed at its maximum likelihood estimate and the number of patches and visits varied (patches: 50, 200, and 600, visits: 4, 10, 20, and 40). We simulated detection histories under the true values and used 10,000 simulations to estimate the properties of interest (bias, variance, MSE). Following MacKenzie and Royle (2005b), Bailey et al. (2007), and Guillera-Arroita et al. (2010) we chose to use simulations instead of expected values because of the low sample size that we observed. We present results including the percentage of boundary estimates ($p = 1$) obtained under each scenario. We used Program R (2009) to conduct the simulations, but acknowledge that Program GENPRES (Bailey et al. 2007) could also be used. We also attempted to find an optimal design given our estimates of occupancy and detection and under a range of patch and visit combinations assuming homogeneous constant detection probability using the software package SODA (Guillera-Arroita et al. 2010). We were able to investigate three criteria for optimization. One was finding the design that minimized Mean Squared Error for the probability of occupancy estimator. Two other approaches incorporated the variance of the probability of detection into the design criteria, which is useful when the probability of detection is a parameter of interest, which is often the case for rare species (e.g., interest lies not only in occupancy, but in amount of effort needed to detect species). The first of these approaches, A-optimality, minimizes the trace of the variance-covariance matrix (i.e., the sum of the variances of the parameters) and gives equal weight to the two variances (probability of occupancy and probability of detection) rather than minimizing the variance of each of the parameters separately. The second approach, D-optimality, minimizes the determinant of the variance-covariance matrix. For a more detailed description of these approaches in the context of occupancy estimation see Guillera-Arroita et al. (2010).

Results

180 total patches were surveyed out of a possible 595 patches in the historic range of the IBWO that included 13 states (Figure 2.1 and Table 2.1). There were 28 total detections from 1/2007 – 5/2008. These detections were all classified as auditory and consisted of 17 “double-knocks”, 7 “kent” calls, and 4 auditory detections with no description. Twenty-seven of the detections occurred in two river basins, Congaree Swamp, SC and Lower Choctawhatchee, FL (Figure 2.1 and Table 2.1). There were 32 total patches in the Congaree Swamp river basin in SC and 18 total patches in the Lower Choctawhatchee river basin in FL and these patches were visited anywhere between 3-40 times with one patch visited 59 times (Table 2.1). Mean number of snags on all sampled patches was 0.9922 ($\pm$ 0.335) with a range of 0.14 to 2.37. Mean number of big trees (>24’’ dbh) was 1.75 ($\pm$ 2.13) with a range of 0 to 6.03.

A total of 25 models were run of which three models (psi(group)p(snags), psi(group)p(trees), psi(group)p(snags+trees)) were excluded because of inestimable parameters (standard errors estimated as 0) (Table 2.2). We found very little support in the data for covariates influencing detection or occupancy (Table 2.2). The top model was the constant (no variation) model for detection and occupancy with an AIC_c weight of 0.2284 and was $\approx$2.5 times more likely than the second best model (psi(.)p(snags)). Out of the 22 models fit to the data, 16 had weights within 10% of the top model suggesting the confidence set of models includes almost all of the *a priori* models. It was clear the models with more complexity (i.e., parameters; covariates and group effects) did little to reduce the deviance (Table 2.2).

Model-averaged parameter coefficients from the candidate set of models (excludes models with weights below 10% of top model) had very large estimated standard errors with the majority of the confidence intervals ranging over the entire parameter space (Table 2.3). This result suggests there was no evidence of a discernible group effect on occupancy or detection either (difference between SC and FL) (Table 2.3). Model-averaged estimates (at the mean patch level covariate value) of detection were slightly higher in FL as was occupancy (Table 2.3). The naïve estimate of occupancy (number of sites with detection/total number of sites sampled) for the two river basins with detections was 0.28, which was substantially lower than the estimated occupancy probabilities. Interestingly, detection probabilities were estimated precisely and the confidence intervals were small compared with the estimates of occupancy probability (Table 2.3).

Figure 2.2 shows the power of the survey method given three different values of detection (lower confidence bound, mean and upper confidence bound of maximum likelihood estimate) over a range of visits. Here, power can either be interpreted as power of rejecting the null hypothesis (non-occurrence) given that the alternative is true (occurrence), or as the probability of at least one detection given the patch is occupied. Depending on what value of detection is used the power of the survey for 40 visits ranged from 0.5 to 0.99, suggesting that at sites with 40 visits and no detections there is strong evidence the site is unoccupied. This statement is strengthened for the few sites that were visited >40 times as power tends to one. These estimates are conditional on the two river basins with detections (Figure 2.1 and Table 2.1). Applying these estimates to the other patches with no detections suggests that there is sufficient power to conclude that these patches were unoccupied as many of the patches were visited frequently (> 40 visits); thus the power to detect the species given occurrence is near one.

The estimated probability of nonoccurrence given no detections at a patch (1- ψ_{condl}) varied with the number of visits to the patch (K) at both the lower bound and mean of the maximum likelihood estimate of occupancy (0.3 and 0.75 respectively) and, assuming fixed detection probability of 0.0461, the mean maximum likelihood estimate (Figure 2.3). We estimate that it takes approximately 70 visits before probability of a "true absence" was 0.9 or higher although the confidence bound for this estimate was very large and includes 0.9 in as little as 20 visits. Using the lower bound of occupancy, the estimate of a "true absence" was 0.9 or higher in as few as 30 visits and the confidence bound included 0.9 with less than 10 visits. If we take the lower bound of the occupancy estimate (ψ=0.3) and assume it was a proxy for the probability of occupancy for the patches outside of the two river basins used to fit the models we conclude that there was strong evidence that those patches were indeed unoccupied. Using maximum likelihood estimates, we estimate that between 141 – 445 patches out of 595 that should be occupied depending on which occupancy probability and number of visits to a patch is used (Table 2.4).

Table 2.5 displays the bias, variance, and MSE for a range of designs given a fixed detection probability (0.0461, mean of maximum likelihood estimate) and three different estimates of the probability of occupancy (lower confidence bound, mean, and upper confidence bound of maximum likelihood estimate). There appears to be little bias in estimating detection probability regardless of the occupancy parameter although in many scenarios estimates tend to the boundary. MSE for detection probability is relatively low regardless of the true parameter values and ranges from 0 to 0.0076 (ψ=0.3, patches=50, visits=4). Bias in estimating occupancy, on the other hand, can be substantial. The current design (patches = 50, visits = 4 to 40) exhibits a considerable amount of bias (minimum = 0.0019 and maximum = 0.5508) with bias decreasing as the number of visits to a patch increases and decreasing as the true occupancy value increases. MSE in estimating occupancy follows a similar pattern as it decreases as the number of visits to a patch increases and it decreases as the true occupancy value increases. There is a greater reduction of MSE by increasing the number of visits as opposed to increasing the number of patches regardless of the true occupancy value.

Figure 2.4 displays the distribution of maximum likelihood estimates for a select group of combinations of true occupancy, number of patches, and number of visits. It is clear with a low number of visits to a patch that there is substantial bias in estimating both occupancy and detection although there appears to be more difficulty estimating occupancy. Sampling more sites while holding the number of visits fixed results in substantial improvement in reducing bias while estimating both detection and occupancy. This improvement is less pronounced when increasing the number of visits and holding the number of patches fixed.

The optimal design according to the D-opt criterion for the current study ($\hat{\psi}$ =0.75, 200 total patches) was to sample only 4 patches with 50 visits. The other two criteria found no optimal design for the current study (Table 2.6). Note that it did not search all possible combinations of criteria, but instead started with the design suggested as being optimal by asymptotic evaluations and then iterates through different combinations. In many cases there was no optimal design found when looking at all of the possible combinations. Often, the optimal design consisted of a large number of visits or a very large number of patches. The D-optimal criteria appeared to always find an optimal design while the other two criteria appeared to come to the same conclusion.

Discussion

The putative rediscovery of the IBWO and the surrounding doubt associated with the discovery has highlighted the need for rigor and standards in science and conservation (Jackson 2006, Jones et al. 2007, McKelvey et al. 2008). Although much of the debate centers on the validity of the observations, an important issue is raised regarding the methodological approaches used to validate (or invalidate) the existence and distribution of critically rare species (Elphick 2008, Elphick et al. 2010, Roberts et al. 2009, Scott et al. 2008). Often, data on critically rare species are collected in a haphazard manner limiting the ability to make credible inference about the species of interest. Presumably, this is primarily due to the difficulty in designing surveys for rare species as many common standard methods are usually not suitable (Thompson 2004). We believe that our implementation of a single-season occupancy model assisted by the use of historical information provides a practical and scientifically rigorous framework for monitoring wide-ranging avian species. We acknowledge that in our particular case our estimates of occupancy and detection were biased high because we needed to exclude the sampled patches within the former range of the IBWO in which IBWO were not detected to fit our models. Nonetheless, we believe the overall framework we presented should be useful for making valid inference about a species of interest.

We were unable to find evidence of a relationship among several habitat covariates and IBWO apparent presence and distribution although several relationships were thought to exist (Tanner 1942). This result highlights a critical misconception about sampling extremely rare species. Often sampling only occurs in areas where the species is "thought" to be or based on a preconceived habitat relationship, but by using a statistically valid sampling design other areas can be sampled, permitting inference to a larger area. In addition, the information gained at other putatively "unsuitable" or "non-ideal" locations can potentially provide confirmation and support of the pre-existing relationships thought to exist or provide evidence of new relationships. This is not a minor point for species where there is very little biological information available. Relying solely on a small number of observations and subsequently restricting search effort to those particular areas can severely limit the understanding of the resource requirements and affect the effectiveness of management.

Our occupancy estimates were high (>0.75), but were estimated with a considerable amount of uncertainty (large standard error) and often ranged over the entire parameter space (0 to 1). Because of this poor precision it makes it difficult to make confident statements about occupancy rates. This uncertainty is elevated when we acknowledge the bias in our estimates. As our simulation results showed, bias could exceed 0.5 in our current study. Unfortunately, there was no feasible optimal design for our situation either. Observers would not be interested in sampling only 4 patches 50 times unless there was substantial evidence at those patches and only at those patches. Although this design would be optimal in one sense of the word it would not permit making inference to a larger area possible as no new locations could be searched. Another source of bias that was introduced into our estimates was a result of relying on data only from two river basins with detections instead of using all of the patches that were sampled.

Given our calculated values of the proportion of area occupied based on our estimates of occupancy and detection, we assume that there should have been many more detections across the former range of the IBWO. There were no detections in the other 130 patches, suggesting that our estimates of detection and occupancy were likely biased by non-representative sampling, and may call into question the validity of these localized auditory detections. On the other hand it also suggests that the IBWO is absent from the majority of its former range while some corroboration of evidence exists in Florida (Hill et al. 2006).

The use of not only the probability of occupancy (ψ), but the proportion of area occupied (PAO or ψ^{fs}) illustrates an important, yet subtle point for conservation. For many rare species, both of the quantities (ψ and ψ^{fs}) are essential and knowing only one may not be sufficient. Consider a situation in which a rare species is only known to exist in a very small geographic region. Simply estimating occupancy probability may provide a theoretical measure of the probability of sites occupied based on a random sample of an infinitely large population of sites, but if there are no other known study areas then its use is limited. Instead, a manager would be more interested in knowing how many sites are actually occupied in the study area. Royle and Kery (2007) point out that although the expected values of the two quantities may be equivalent, the levels of uncertainty are typically different and in their example ψ^{fs} resulted in more precise estimates. This was also exemplified in our own results and stems from the information in the sample data that contains sites where the state is actually known, thus reducing the uncertainty in the estimate of ψ^{fs}. By simply estimating and using the probability of occupancy (infinite sample) the variation in the estimate would be much larger than the finite sample estimate and would not use all of the information in the data. This reduction in uncertainty is critical for managing and conserving rare species in small or restricted geographic ranges. Admittedly, the estimation of such quantities can be challenging and preferably a Bayesian hierarchical modeling approach allows for a natural and intuitive option to estimation (see MacKenzie et al. 2006 and Royle and Dorazio 2008 for examples).

Although we provided a statistically valid and rigorous approach to estimate parameters of interest for a wide-ranging rare avian species, there was one critical deficiency in our design that should be addressed in future work. In the IBWO survey searchers were highly motivated (some might say obsessed) to find a bird. When a putative detection (i.e., sound) occurred in a particular patch, it was extremely difficult to get searchers to survey anywhere else, which is human nature. But the simple random sampling design required them to sample in another randomly selected location. Not only does this create antipathy towards the use of a formal design it can also result in missing important information and data collection opportunities for species that are clustered, or in this case for an individual bird that might use a cluster of patches in its home range. We found that observers were constantly trying to find ways to bypass the use of a simple random sample to reallocate sampling effort in locations they thought were more suitable because of a known detection or strong prior belief. This problem can be generalized to many situations involving rare species when information is limited and highly valued and can be exemplified by the behavior of birders when a rare species has been sighted. A potential solution would be to implement an adaptive design which permits augmentation of the design to accommodate clustered individuals while still allowing for statistically valid inference. One approach to augment the sample design would be to use design-based estimators based on adaptive cluster sampling (Thompson and Seber 1996) that have been used quite frequently for rare species (Thompson 2004). In such a scenario any detection could act as a “trigger” to permit augmentation of the sampling design such that adjacent patches are sampled. This approach would allow an increase in effort around areas where detections are observed thus allowing the potential for more evidence to be collected. Unfortunately, the only way to account for imperfect detection with this method would be to have independent estimates from another survey (Thompson and Seber 1996) which may not be feasible. Alternatively one could develop a hierarchical model that incorporated the adaptive cluster sampling design into the modeling effort (see Rapley and Welsh 2008 for a Bayesian treatment of this approach). Ideally a design that uses repeat visits to the patches could be implemented thus alleviating the need for independent estimates of detection. This would permit inference of occupancy and detection while accommodating the adaptive augmentation of the sampling design.

Our results come from a large multi-state and multi-agency effort to collect information on a rare species. In the future, these sorts of cooperative approaches will likely play a major role in the conservation of rare avian species because of the large areas that birds traverse and the difficulty in collecting information at one particular site. Not only does our design provide a framework for rigorous inference at a large scale, it displays the benefits of working across geographic boundaries when a common goal is to be reached. Yoccoz et al. (2001) and Pollock et al. (2002) identified three major points of interest for any long-term and/or large scale monitoring program: (1) Why? (2) What? and (3) How? These points must be addressed and communicated effectively among all cooperators to ensure the success and credibility of large scale long-term monitoring programs for rare avian species. We see the development and application of new statistically rigorous survey methods and models along with the cooperation of partners across large landscapes as providing the critical steps in conservation and management of rare species.

Literature Cited

Allen, A. A., and P. P. Kellogg. 1937. Recent observations on the ivory-billed woodpecker. Auk 54:164-184.

Bailey L, Hines J, Nichols J, MacKenzie D. 2007. Sampling design trade-offs in occupancy studies with imperfect detection: examples and software. Ecological Applications 17: 281 – 290.

Dorazio, R, Royle, J. 2005. Estimating size and composition of biological communities by modeling the occurrence of species. Journal of the American Statistical Association 100: 389-398.

Efron, B, Tibshirani R. 1993. An introduction to the boostrap. Chapman and Hall, Boca Raton, Florida.

Elphick, C .2008. How you count counts: the importance of methods research in applied ecology. Journal of Applied Ecology 45:1313-1320.

Elphick, C, Roberts, D, Reed, M. 2010. Estimated dates of recent extinctions in North America and Hawaiian birds. Biological Conservation 143:617-624.

Fitzpatrick J et al. 2005. Ivory-billed woodpecker (*Campephilus principalis*) persists in continental North America. Science 308:1460-1462.

Fitzpatrick J, Lammertink, M, Luneau Jr., M, Gallagher, T, Rosenberg, K. 2006. Response to Comment on "Ivory-billed Woodpecker (*Campephilus principalis*) persists in continental North America. Science 311:1555b

Guillera-Arroita, G, Ridout, M, Morgan B. 2010. Design of occupancy studies with imperfect detection. Methods in Ecology and Evolution 1: 131-139.

Hill G, Mennill D, Rolek B, Hicks T, Swiston K. 2006. Evidence Suggesting that Ivory-billed Woodpeckers (Campephilus principalis) Exist in Florida. Avian Conservation Ecology 1:URL: http://www.ace-eco.org/vol1/iss3/art2/.

Jackson J. 2002. Ivory-billed Woodpecker: Campephilus Principalis. Birds of North America, Inc.

Jackson J. 2006. The public perception of science and reported confirmation of the Ivory-billed Woodpecker in Arkansas. Auk 123:1185-1189.

Jackson J. 2006b. Ivory-billed Woodpecker (*Campephilus principalis*): hope, and the interfaces of science, conservation, and politics. Auk 123:1-15.

Jones, C, Troy, J, Pomara, L. 2007. Similarities between *Campephilus* woodpecker double raps and mechanical sounds produced by duck flocks. Wilson J. Ornithology 119:259-262.

MacKenzie D, Nichols J, Lachman G, Droege S, Royle J, Langtimm C. 2002. Estimating site occupancy rates when detection probabilities are less than one. Ecology 83: 2248 – 2255.

MacKenzie, D, Bailey, L. 2004. Assessing the fit of site-occupancy models. Journal of Agricultural, Biological, and Environmental Statistics 9:300-318.

MacKenzie D, Nichols J, Sutton N, Kawanishi K, and Bailey L. 2005. Improving inferences in population studies of rare species that are detected imperfectly. Ecology 86: 1101 – 1113.

MacKenzie D, Nichols J, Royle J, Pollock K, Bailey L, Hines J. 2006. Occupancy estimation and modeling. Elsevier, London.

MacKenzie D, Royle J. 2005b. Designing occupancy studies: general advice and allocating survey effort. Journal of Applied Ecology 42: 1105 – 1114.

McKelvey, K, Aubry, K, Schwartz M. 2008. Using anecdotal occurrence data for rare or elusive species: the illusion of reality and a call for evidentiary standards. Bioscience 58:549-555.

O'Connell, A, Talancy, N, Bailey, L, Sauer, J, Cook, R, Gilbert, A. 2006. Estimating site Occupancy and detection probability parameters for meso- and large mammals in a Coastal ecosystem. Journal of Wildlife Management 70:1625-1633.

Pollock, K, Nichols, J, Simons, T, Farnsworth, G, Bailey, L, Sauer, J. 2002. Large scale Wildlife monitoring studies: statistical methods for design and analysis. Environmetrics 13:105-119.

R Development Core Team. 2009. . R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. ISBN 3-900051-07-0, URL http://www.R-project.org.

Rapley, V, Welsh, A. 2008. Model-based inferences from adaptive cluster sampling. Bayesian. Analysis 3:717-736.

Roberts, D, Elphick, C, Reed, M. 2009. Identifying anomalous reports of putatively extinct species and why it matters. Conservation Biology 24: 189-196.

Royle J, Dorazio R. 2008. Hierarchical modeling and inference in ecology. Academic Press, London.

Royle J, Kery M. 2007. A Bayesian state-space formulation of dynamic occupancy models. Ecology 88: 1813 – 1823.

Scott, J, Ramsey, F, Lammertink, M, Rosenberg, K, Rohrbaugh, R, Wiens, J, Reed, J. 2008. When is an "extinct" species really extinct? Gauging the search efforts for Hawaiian forest birds and the Ivory-billed Woodpecker. Avian Conservation Ecology 3 http://www.ace-eco.org/vol3/iss2/art3/

Seber, G. 1982. The estimation of animal abundance and related parameters. The Blackburn Press, New Jersey.

Tanner, J. 1942. The Ivory-billed Woodpecker. Research Report No. 1, National Audubon Society, New York.

Thompson, S. 1992. Sampling. Wiley Interscience, New York.

Thompson, S, Seber G. 1996. Adaptive sampling. Wiley Interscience, New York.

Thompson ,W. 2004. Sampling rare or elusive species. Island Press, Washington

White G, Burnham K. 1999. Program MARK: Survival estimation from populations of marked animals. Bird Study 46 Supplement 120 – 138.

Williams K, Nichols J, Conroy M. 2002. Analysis and management of animal populations. Academic Press, San Diego.

Yoccoz, N, Nichols J, Boulinier T. 2001. Monitoring of biological diversity in space and time. Trends in Ecology and Evolution 16: 446 – 453.

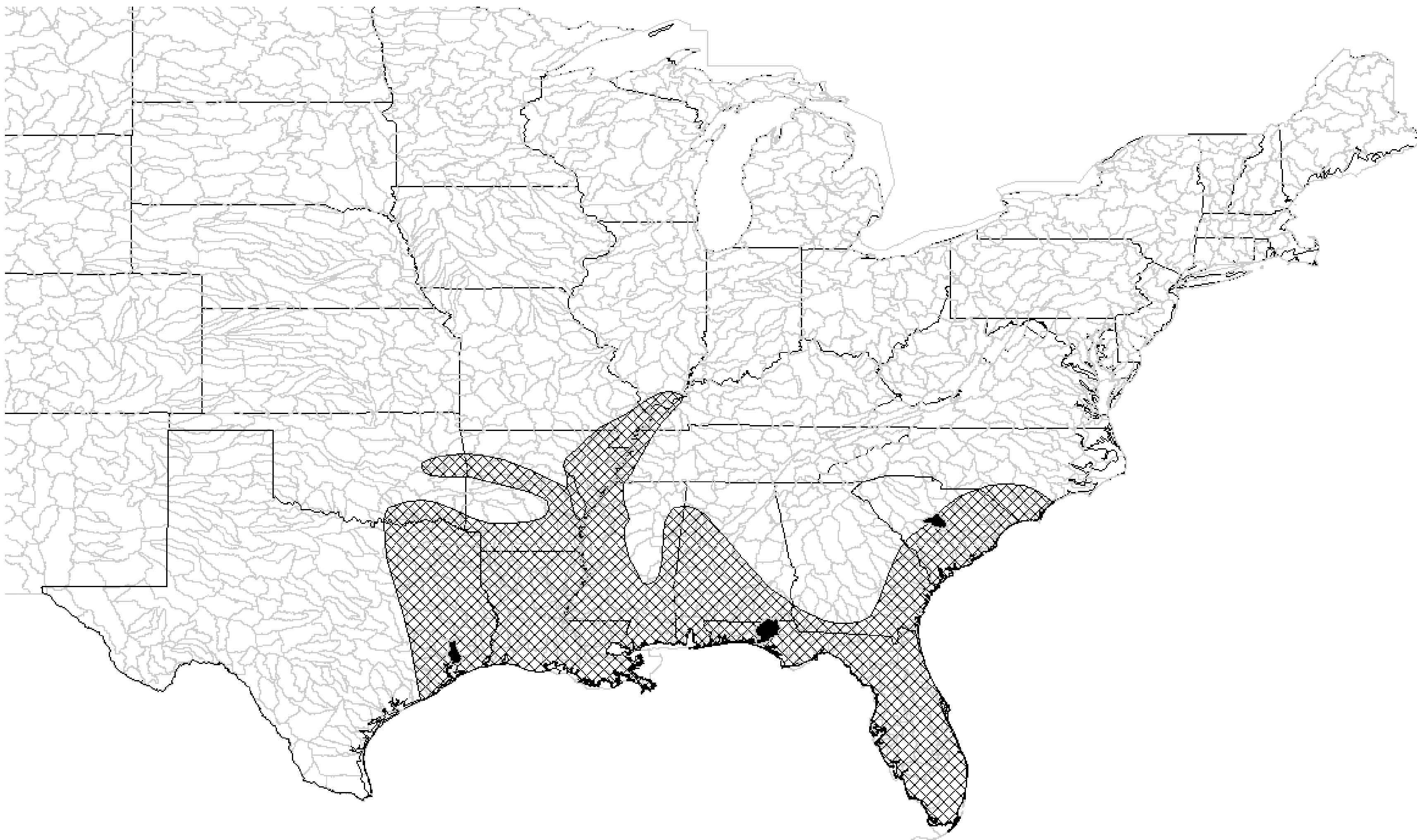


Figure 2.1. River basins in the United States within the former range (cross-hatched) of the Ivory-billed Woodpecker (*Campephilus principalis*). Three darkened areas represent basins that contained at least a single auditory detection observed during the survey in 2007-2008.

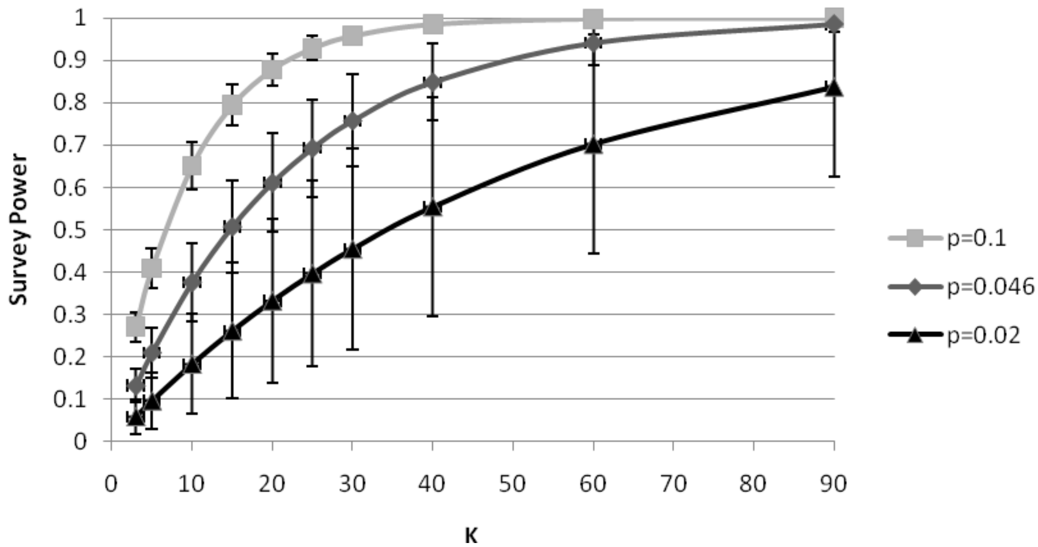


Figure 2.2. Power of detection for the survey method given that the patch is occupied under three different levels of detection probability and over a range of repeat visits to the patch (K) and with 95% confidence limits obtained from the delta method. This can also be interpreted as the probability of at least one detectoin conditional on the patch being occupied.

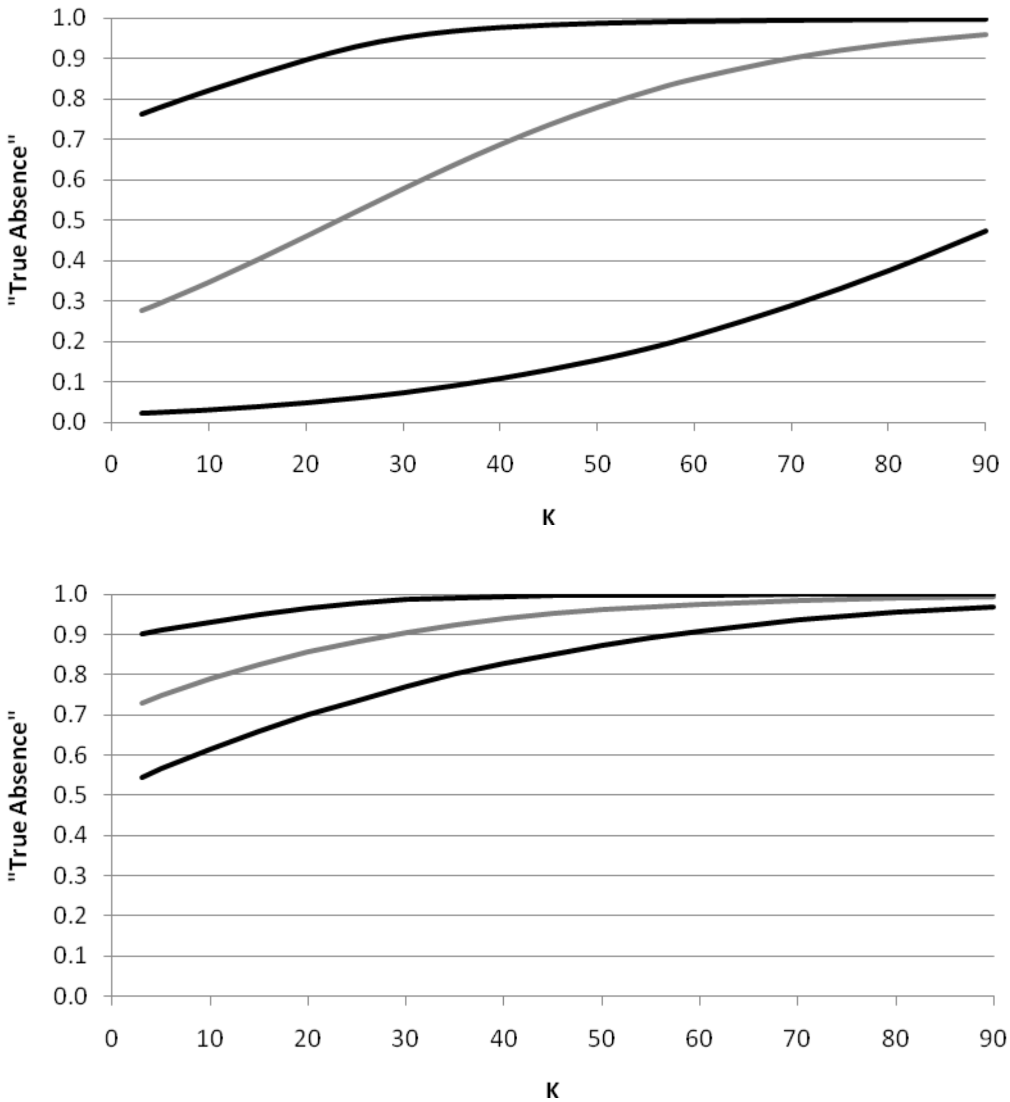


Figure 2.3. Probability of nonoccurrence conditional on no detections at a patch for two levels of occupancy, ψ=0.75 (top) and ψ=0.3 (bottom) over a range of number of visits to a patch (*K*). The dark lines represent the 95% confidence bounds obtained from bootstrapping with 1,000,000 iterations and the two levels of occupancy correspond to the mean and lower confidence bound of the maximum likelihood estimate from the data.

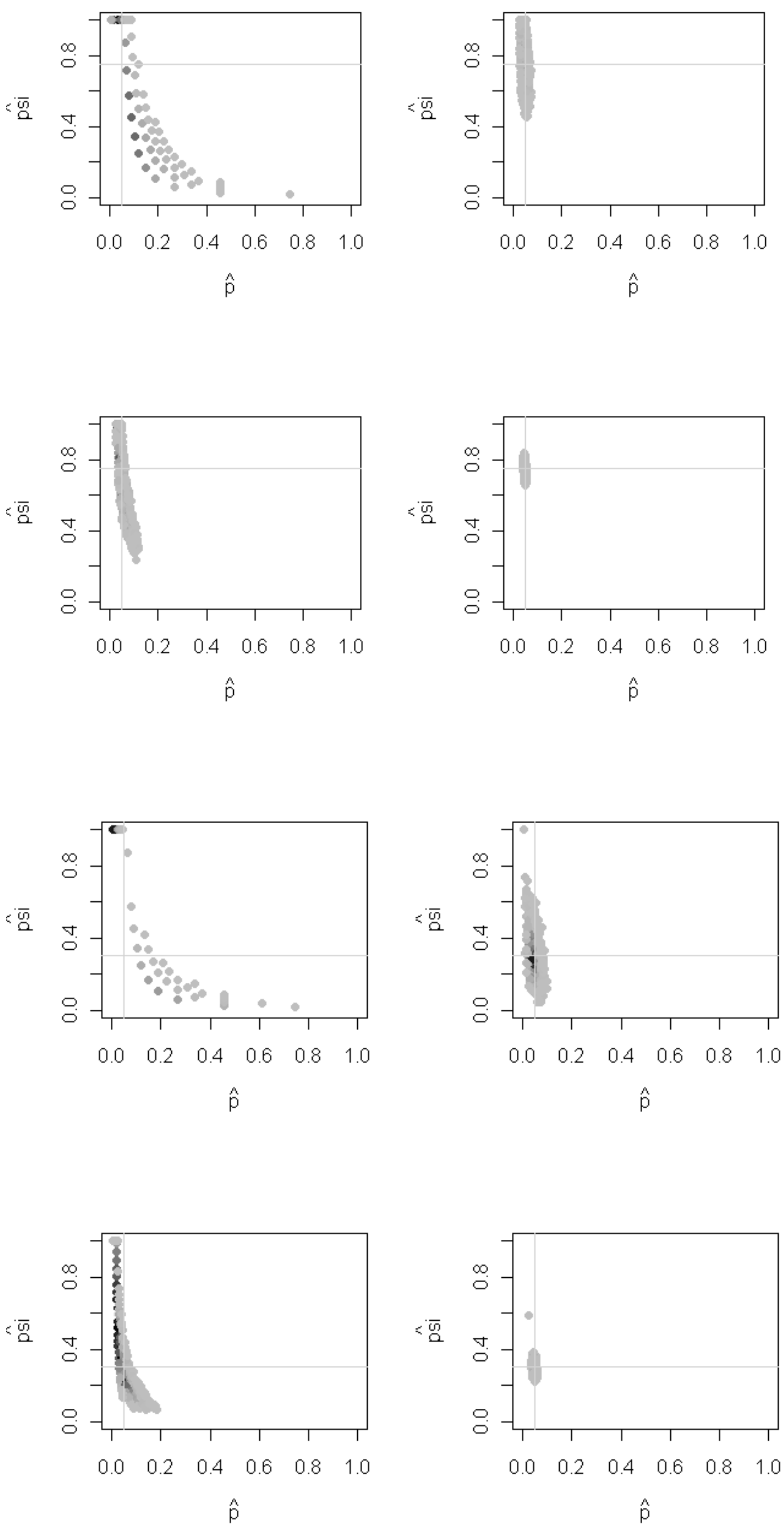

$\hat{psi}$
$\hat{p}$
0.0 0.2 0.4 0.6 0.8 1.0
0.0 0.4 0.8

Figure 2.4. Distribution of maximum likelihood estimates obtained from evaluating estimator bias, variance and Mean Squared Error for a range of true values for occupancy and detection using 10,000 simulations. Values of occupancy and detection were obtained from maximum likelihood estimates of the data and the range of conditions (# of patches and visits) were from the range of conditions used in the field survey. Light gray cross hairs represent the true simulated value of occupancy ($\hat{\psi}$) and detection ($\hat{p}$). From left to right and then down: ψ=0.75, p=0.046, patches = 50, visits = 4; ψ=0.75, p=0.046, patches = 50, visits = 40; ψ=0.75, p=0.0461, patches = 600, visits = 4; ψ=0.75, p=0.046, patches = 600, visits = 40; ψ=0.3, p=0.046, patches = 50, visits = 4; ψ=0.3, p=0.046, patches = 50, visits = 40; ψ=0.3, p=0.046, patches = 600, visits = 4; and ψ=0.3, p=0.046, patches = 600, visits = 40.

Table 2.1. Information about the survey design and detection data collected from 2007-2008 for the Ivory-billed Woodpecker (*Campephilus principalis*) across its former range in the United States. Patches were approximately 2 km^2. *One patch was visited 59 times in the Congaree River Basin in South Carolina.

State	# River Basins	Range of # of Patches Within Each River Basin	Number of Patches in a River Basin With a Detection	Number of Visits to the Patches	# of Auditory Detections
Alabama	39	1-81			0
Arkansas	49	1-37			0
Florida	54	1-81	4 (Lower Choctawhatchee)	3-14	5
Georgia	26	1-23			0
Kentucky	9	1			0
Louisiana	56	1			0
Missouri	10	1-23			0
Mississippi	53	1-37			0
North Carolina	10	1			0
Oklahoma	17	1			0
South Carolina	22	1-44	10 (Congaree)	4-40*	22
Tennessee	11	1			0
Texas	48	1-36	1 (Lower Trinity)		1

Table 2.2. The deviance (-2*log likelihood), number of parameters, Akaike's Information Criteria adjusted for small sample size (AIC_c), the difference between the model and the top model (ΔAIC_c), model weights (w), and likelihood of candidate detection (p) and occupancy (psi) models for the Ivory-billed Woodpecker (*Campephilus principalis*) in two states (SC and FL). The dotted line represents the cutoff for the candidate set of models (>10% of the top model's weight). The group effect corresponds to differences between the 2 states.

Model[1]	Deviance	Num. Par	AICc	ΔAICc	*w*	Likelihood
{psi(.)p(.)}	229.95	2	234.20	0.00	0.228	1.000
{psi(.)p(snags)}	229.50	3	236.02	1.82	0.092	0.403
{psi(group)p(.)}	229.53	3	236.06	1.85	0.090	0.396
{psi(.)p(big_trees)}	229.73	3	236.25	2.05	0.082	0.359
{psi(.)p(group)}	229.82	3	236.34	2.14	0.078	0.344
{psi(snags)p(.)}	229.90	3	236.42	2.21	0.075	0.331
{psi(big_trees)p(.)}	229.94	3	236.46	2.26	0.074	0.323
{psi(.)p(snags+big_trees)}	229.42	4	238.31	4.11	0.029	0.128
{psi(snags)p(snags)}	229.50	4	238.39	4.18	0.028	0.124
{psi(big_trees)p(snags)}	229.50	4	238.39	4.18	0.028	0.123
{psi(group)p(group)}	229.50	4	238.39	4.19	0.028	0.123
{psi(snags)p(group)}	229.66	4	238.55	4.34	0.026	0.114
{psi(snags)p(big_trees)}	229.70	4	238.59	4.38	0.025	0.112
{psi(big_trees)p(big_trees)}	229.73	4	238.62	4.41	0.025	0.110
{psi(big_trees)p(group)}	229.80	4	238.69	4.48	0.024	0.106

{psi(snags+big_trees)p(.)}	229.89	4	238.78	4.58	0.023	0.101
{psi(snags)p(snags+big_trees)}	229.42	5	240.79	6.58	0.008	0.037
{psi(big_trees)p(snags+big_trees)}	229.42	5	240.79	6.58	0.008	0.037
{psi(snags+big_trees)p(snags)}	229.50	5	240.86	6.66	0.008	0.036
{psi(snags+big_trees)p(group)}	229.65	5	241.01	6.81	0.008	0.033
{psi(snags+big_trees)p(big_trees)}	229.68	5	241.04	6.84	0.007	0.033
{psi(snags+big_trees) p(snags+big_trees)}	229.42	6	243.37	9.17	0.002	0.010

[1](.) indicates that the parameter was modeled as a constant.

Table 2.3. Model averaged real parameter estimates at the mean value of the patch-level covariate and model-averaged Beta coefficients obtained from the candidate set of 16 models (>10% of the weight of top model). Naïve estimate of occupancy corresponds to the number of known occupied patches (> 1 detection) divided by the number of patches sampled (50 total from SC and FL).

Real Parameter Estimates	**Model averaged Estimate**	**Unconditional SE**
Naïve	*0.28*	
p SC	0.045	0.012
p FL	0.047	0.016
psi SC	0.733	0.199
psi FL	0.768	0.249
Beta parameters		
p Intercept	-2.3500	0.9844
p Group effect (Florida)	-0.2951	0.7054
p Snags	0.0454	0.3411
p Big Trees	0.0105	0.1859
psi Intercept	-4.3519	2.6873
psi Group effect (Florida)	0.8138	1.2508
psi Snags	0.0435	0.5061
psi Big Trees	0.0034	0.2947

Table 2.4. Total number of occupied patches out of 595 total patches (N_{occ}), number of repeat visits (K), occupancy probability conditional on no detections at a patch (ψ_{condl}), and occupancy rate or proportion of patches occupied out of 595 (ψ^{fs}) and their estimated variances following Royle and Dorazio (2008). The two levels of occupancy (ψ) represent the values used for the calculation of these quantities and correspond to the mean and lower confidence bound of the maximum likelihood estimates.

$\psi=0.75$	N_{occ}	$Var(N_{occ})$	K	ψ_{condl}	$Var(\psi_{condl})$	ψ^{fs}	$Var(\psi^{fs})$
	445	1115.22	3	0.72	0.0397	0.75	0.00315
	442	1171.46	5	0.70	0.0418	0.74	0.00331
	434	1299.38	10	0.65	0.0464	0.73	0.00367
	425	1421.99	15	0.60	0.0508	0.71	0.00402
	415	1522.79	20	0.54	0.0544	0.70	0.00430
	405	1587.90	25	0.48	0.0568	0.68	0.00449
	396	1608.15	30	0.42	0.0576	0.67	0.00454
	386	1585.67	35	0.36	0.0568	0.65	0.00448
	378	1523.29	40	0.31	0.0547	0.63	0.00430
	370	1439.82	45	0.26	0.0517	0.62	0.00407
	363	1338.49	50	0.22	0.0481	0.61	0.00378
	356	1225.76	55	0.18	0.0441	0.60	0.00346
	351	1110.46	60	0.15	0.0400	0.59	0.00314
	343	884.15	70	0.10	0.0319	0.58	0.00250
	337	685.56	80	0.06	0.0248	0.57	0.00194
	333	511.90	90	0.04	0.0186	0.56	0.00145

ψ=0.3	184	259.34	3	0.27	0.0083	0.31	0.00073
	181	242.32	5	0.25	0.0078	0.30	0.00068
	174	204.56	10	0.21	0.0065	0.29	0.00058
	168	173.85	15	0.17	0.0055	0.28	0.00049
	163	147.42	20	0.14	0.0047	0.27	0.00042
	159	123.32	25	0.12	0.0039	0.27	0.00035
	155	100.36	30	0.09	0.0032	0.26	0.00028
	152	79.61	35	0.08	0.0025	0.26	0.00022
	150	62.04	40	0.06	0.0019	0.25	0.00018
	148	47.20	45	0.05	0.0015	0.25	0.00013
	146	35.69	50	0.04	0.0011	0.25	0.00010
	145	26.57	55	0.03	0.0008	0.24	0.00008
	144	19.61	60	0.02	0.0006	0.24	0.00006
	142	10.49	70	0.02	0.0003	0.24	0.00003
	141	5.50	80	0.01	0.0001	0.24	0.00002
	141	2.86	90	0.01	0.0001	0.24	0.00001

Table 2.5. Results from conducting 10,000 simulations to evaluate bias, variance and mean squared error (MSE) for both occupancy and detection under a range of scenarios characterized by the number of visits to a patch (*K*), the number of patches, true occupancy probability (0.3, 0.75, 0.9; lower confidence bound, mean and upper confidence bound of maximum likelihood estimate), and a fixed detection probability (0.046; mean of maximum likelihood estimate). P empty corresponds to the percentage of simulated detection histories that were empty while p bound represents the percentage of estimates of *p* that were at the boundary (*p*=1).

ψ	K	Patches	p Bias	p Var	p MSE	p Empty	p Bound	ψ Bias	ψ Var	ψ MSE
0.9	4	50	0.0187	0.002	0.0023	0	0.629	-0.0912	0.0781	0.0864
0.9	4	200	0.0087	0.0004	0.0005	0	0.473	-0.0716	0.0446	0.0497
0.9	4	600	0.0044	0.0002	0.0002	0	0.417	-0.0424	0.0257	0.0275
0.9	10	50	0.005	0.0003	0.0003	0	0.455	-0.0449	0.0323	0.0343
0.9	10	200	0.002	0.0001	0.0001	0	0.343	-0.0185	0.0153	0.0157
0.9	10	600	0.0006	0	0	0	0.226	-0.0026	0.0077	0.0077
0.9	20	50	0.0012	0.0001	0.0001	0	0.351	-0.0114	0.0137	0.0139
0.9	20	200	0.0004	0	0	0	0.156	-0.0004	0.0056	0.0056
0.9	20	600	0	0	0	0	0.038	0.002	0.0025	0.0025
0.9	40	50	0.0001	0	0	0	0.129	0.0019	0.0054	0.0054
0.9	40	200	-0.0001	0	0	0	0.08	0.002	0.0017	0.0017
0.9	40	600	0	0	0	0	0	0.0011	0.0006	0.0006
0.75	4	50	0.0153	0.0028	0.0031	0.001	0.661	0.0506	0.0923	0.0948
0.75	4	200	0.0068	0.0006	0.0007	0	0.405	0.0118	0.063	0.0631

0.75	4	600	0.0023	0.0002	0.0002	0	0.293	0.0209	0.0397	0.0401
0.75	10	50	0.0029	0.0004	0.0004	0	0.349	0.026	0.0487	0.0494
0.75	10	200	0.0006	0.0001	0.0001	0	0.151	0.0213	0.0242	0.0247
0.75	10	600	0	0	0	0	0.032	0.013	0.0107	0.0109
0.75	20	50	-0.0001	0.0001	0.0001	0	0.155	0.0257	0.0233	0.0239
0.75	20	200	-0.0001	0	0	0	0.009	0.0096	0.0074	0.0075
0.75	20	600	0	0	0	0	0	0.0028	0.0025	0.0025
0.75	40	50	-0.0002	0	0	0	0.008	0.0085	0.0083	0.0084
0.75	40	200	-0.0001	0	0	0	0	0.002	0.0019	0.0019
0.75	40	600	0	0	0	0	0	0.001	0.0006	0.0006
0.3	4	50	0.0019	0.0076	0.0076	0.071	0.767	0.5508	0.1075	0.4109
0.3	4	200	0.0034	0.0022	0.0022	0	0.489	0.3052	0.1584	0.2515
0.3	4	600	0.0005	0.0008	0.0008	0	0.126	0.1318	0.0782	0.0956
0.3	10	50	0.0003	0.0014	0.0014	0.001	0.332	0.2388	0.1279	0.185
0.3	10	200	-0.0008	0.0004	0.0004	0	0.03	0.0678	0.0353	0.0399
0.3	10	600	-0.0002	0.0001	0.0001	0	0	0.0173	0.0067	0.007
0.3	20	50	-0.0008	0.0004	0.0004	0	0.04	0.0671	0.0379	0.0424
0.3	20	200	-0.0004	0.0001	0.0001	0	0	0.012	0.0044	0.0046
0.3	20	600	-0.0001	0	0	0	0	0.0033	0.0012	0.0012
0.3	40	50	-0.0003	0.0001	0.0001	0	0	0.0098	0.0068	0.0069
0.3	40	200	-0.0002	0	0	0	0	0.0022	0.0015	0.0015
0.3	40	600	0	0	0	0	0	0.0009	0.0005	0.0005

Table 2.6. The optimal design given the true occupancy probability (0.3, 0.75, 0.9; lower confidence bound, mean and upper confidence bound of maximum likelihood estimate), and a fixed detection probability (0.046; mean of maximum likelihood estimate). Total patches corresponds to the number of visits times the number of patches and can represent total effort. *K* represents the number of visits and s represents the number of sites (patches) that need to be visited for the optimal design. "None" refers to no optimal design being found given the constraints. Optimal is defined as minimizing MSE of occupancy (ψ MSE), A-opt minimizes the trace of the variance-covariance matrix (between occupancy and detection) and D-opt minimizes the determinant of the variance-covariance matrix. The search process for an optimal design does not include searching all possible combinations of criteria, but instead starts with the design suggested as being optimal by asymptotic evaluations and then iterates through different combinations from there.

ψ	Total Patches	ψ MSE	A-opt	D-opt
0.9	200	none	none	k=49,s=4
0.9	800	none	none	k=66,s=12
0.9	2400	k=70,s=34	k=70,s=34	k=84,s=28
0.9	500	none	none	k=55,s=9
0.9	2000	k=74,s=27	k=62,s=32	k=80,s=25
0.9	6000	k=50,s=120	k=50,s=120	k=50,s=120
0.9	1000	none	none	k=71,s=14
0.9	4000	k=78,s=51	k=50,s=80	k=50,s=80
0.9	12000	k=50,s=240	k=50,s=240	k=50,s=240

0.9	2000	k=62,s=32	k=66,s=30	k=62,s=32
0.9	8000	k=50,s=160	k=50,s=160	k=50,s=160
0.9	24000	k=50,s=480	k=50,s=480	k=50,s=480
0.75	200	none	none	k=50,s=4
0.75	800	none	none	k=61,s=13
0.75	2400	none	none	k=75,s=32
0.75	500	none	none	k=50,s=10
0.75	2000	none	none	k=50,s=40
0.75	6000	k=49,s=122	k=49,s=122	k=50,s=120
0.75	1000	none	none	k=55,s=18
0.75	4000	k=49,s=81	k=49,s=81	k=50,s=80
0.75	12000	k=49,s=244	k=49,s=244	k=50,s=240
0.75	2000	none	none	k=51,s=39
0.75	8000	k=49,s=163	k=49,s=163	k=50,s=160
0.75	24000	k=49,s=489	k=49,s=489	k=50,s=480
0.3	200	none	none	k=49,s=4
0.3	800	none	none	k=61,s=13
0.3	2400	k=39,s=61	k=38,s=63	k=39,s=61
0.3	500	none	none	k=50,s=10
0.3	2000	none	none	k=42,s=47
0.3	6000	k=32,s=187	k=32,s=187	k=41,s=146
0.3	1000	none	none	k=50,s=20
0.3	4000	k=38,s=105	k=40,s=100	k=48,s=83

0.3	12000	k=32,s=375	k=32,s=375	k=41,s=292
0.3	2000	none	none	k=50,s=40
0.3	8000	k=32,s=250	k=32,s=250	k=41,s=195
0.3	24000	k=32,s=750	k=32,s=750	k=41,s=585

CHAPTER 3

A TWO-PHASE SAMPLING DESIGN FOR INCREASING DETECTIONS OF RARE SPECIES IN OCCUPANCY SURVEYS[1]

[1]Pacifici, K., R. M. Dorazio, and M. J. Conroy. To be submitted to *Methods in Ecology and Evolution.*

Abstract

Occupancy estimation is a valuable tool to examine the effects of habitat characteristics, evaluate the outcome of large scale perturbations, determine the influence of habitat landscape variation on species richness, explore processes driving population dynamics, and understand community dynamics and interactions. Recently, there has been a focus on developing sampling protocols using occupancy estimation. Here, we develop a two-phase sampling approach for rare species. We evaluate this new approach compared to the traditional single-season occupancy approach under a range of conditions exhibiting variation in population structure, detection probability, and spatial relatedness. We develop an intuitive measure of predictive error and use simulations to evaluate the two approaches. We found that our new approach outperformed the traditional approach under almost every scenario. We believe that our new approach will be valuable for managers and conservationists interested in rare species and will provide a more efficient allocation of effort.

Occupancy modeling/estimation (MacKenzie et al. 2006) has become a valuable tool to examine effects of localized habitat characteristics (Dorazio et al. 2006), evaluate the outcome of large scale perturbations (Saracco et al. 2011), determine the influence of habitat and landscape variation on species richness (Kery and Royle 2009), explore processes driving population dynamics (MacKenzie et al. 2003), and understand community dynamics and interactions (Dorazio and Royle 2005; Zipkin et al. 2009) all while accounting for imperfect detection (MacKenzie et al. 2002). The recent increase of occupancy studies relying on binary presence-absence data is due in part to the ease at which this type of data can be collected in the field while still allowing for explicit discrimination of biological hypotheses and full investigation of process dynamics (MacKenzie et al. 2006). Although the theory supporting occupancy estimation is sound, often the greatest resistance to the use of occupancy modeling is the complications arising when designing and implementing such approaches. The amount and allocation of effort needed to survey many sites over a large scale and the necessity of spatial and/or temporal replication (required to estimate detection probability) can sometimes be challenging (MacKenzie and Royle 2005; Bailey et al. 2007). These challenges become more significant and often magnified when dealing with rare species (MacKenzie et al. 2005).

Rare species are usually of great concern to managers and conservationists because there is little information about relevant demographic state variables and how they respond to environmental perturbation (including human induced changes), and they are ultimately at higher risk (Thompson 2004). In general sampling is difficult because these animals are often hard to observe or capture, occur in low numbers and are patchily distributed (Thompson 2004). Difficulties arise when designing surveys for rare species because obtaining adequate information to be used in analysis can be demanding. Often a very large proportion of zeros or non-detections exist in the dataset creating problems not only with parameter estimation (few data, convergence issues) and the interpretation of non-detections as true absences (Martin et al. 2005), but result in wasted field effort and squandered resources.

Recently there has been an increase in studies focusing on the implementation and logistical constraints imposed when collecting occupancy-type data, with much of this effort concentrated on designing efficient surveys. MacKenzie and Royle (2005) explored the optimal number of sites to survey and repeat visits to each site along with investigating different methods of sampling (i.e., double sampling, removal sampling or a standard design) and their relative efficiency, Guillera-Arroita et al. (2010) incorporated the precision of the estimated detection parameters into the optimality of the design and Field et al. (2005) evaluated the statistical power for alternative survey designs to detect a decline in occupancy over a 3-yr. period when false-negative errors (imperfect detection) were present.

One area of research that has received little attention is the efficient allocation and distribution of effort for occupancy surveys. Often managers and conservationists want to maximize the amount of information gained with each survey especially when working with rare or endangered species where severe logistical and financial constraints exist (Wilson et al. 2006). Frequently, rare species are found in spatially correlated patches (Prendergast et al. 1993) and thus a more efficient use of resources would be to allocate additional effort to these areas in an attempt to increase the amount of information collected. An adaptive design that permits reallocation of effort towards areas of higher occupancy would provide the following benefits, 1) identify hotspots or clusters of individuals, 2) result in more information collected via more positive detections, and 3) be a more economical use of resources.

Here we provide a two-phase framework for adaptive reallocation of effort in occupancy studies. In the first phase a simple random sample is taken and occupancy probability is estimated and predicted at non-sampled locations as a function of habitat covariates thought to influence occupancy. The predicted occupancy probabilities at non-sampled locations then form the inclusion probabilities for the second phase of sampling. Thus in the second phase additional locations are sampled relative to their predicted occupancy probability. Locations with a high predicted probability of occupancy have a higher probability of inclusion such that once a "good quality" habitat is found (i.e., sampled) predicted probability of occurrence will be higher and effort will be more concentrated in these habitats. We compare this adaptive strategy for effort allocation to traditional simple-random-sampling used in single-season occupancy estimation. We develop an intuitive measure of error and use simulations to assess the overall error rate and predictive ability of the two methods. We show that our new adaptive approach is highly robust outperforming the traditional approach in almost every scenario. Even in the worst cases our approach results in similar error rates to those that are found when using the traditional simple-random-sampling single-season occupancy approach.

Methods

We describe a two-phase sampling protocol that permits informed allocation of effort. Our framework for adaptation is predicated on two important premises: 1) that variation in occurrence is associated with spatially varying habitat information such that interest lies in examining the relationship between a covariate or suite of covariates and their influence on occupancy, and 2) the availability of fully observable covariate information throughout the study area obtained through possibly remotely-sensed data (e.g., GIS). In addition we assume that although there is interest in the relationship between specific habitat covariates and occurrence, this relationship is not completely understood or quantified before the study. For example there are often a suite of covariates believed to influence occurrence and possibly explain spatial heterogeneity, but the specific habitat covariates as well as the direction and magnitude of influence is not necessarily known.

Two-phase sampling approach

The first phase of our approach is the traditional single-season occupancy model as described by MacKenzie et al. (2006) in which a simple random sample of size n_1 is taken from N sites. On each of the $i = 1 \ldots n_1$ sites J Bernoulli detection samples are taken resulting in $y_i = 0, 1, 2, \ldots, J_i$ detections per site. A simple model for estimating occupancy is thus:

$$y_i | z_i, p \sim Binomial(J, z_i p)$$
$$z_i | \psi \sim Bernoulli(\psi)$$

Where z_i is the underlying latent occupancy state at site i, $i=1,\ldots n_1$, p is the probability of detection which we can assume to be constant across sites and visits, J is the number of repeat visits to an individual site and ψ is the probability of occupancy. Variation in occupancy probability can be modeled as a function of site-specific covariates thus inducing heterogeneity across the different locations in the study area. To do so we can use a logit-linear model with site-specific habitat covariates x_i, influencing ψ_i:

$$logit(\psi_i) = \beta_0 + \beta_1 * x_i \tag{3.1}$$

Here we will focus on only one covariate influencing occupancy, but our approach can be expanded to accommodate multiple covariates such that:

$$logit(\psi_i) = \beta_0 + \beta_1 * x_{1i} + \beta_2 * x_{2i} + \cdots + \beta_k * x_{ki} \tag{3.2}$$

Under conditional independence among sites the marginal or integrated likelihood for the observations (integrating out the latent process variable z_i) is

$$\Pr(y_1, y_2, \dots, y_{n_1}) = \prod_{i=1}^{n_1} \Pr(y_i | \psi_i, p)$$

$$= \prod_{i:y_i>0} \binom{J}{y_i} \psi_i p^{y_i} (1-p)^{J-y_i} \prod_{i:y_i=0} \psi_i (1-p)^J + (1-\psi_i) \qquad (3.3)$$

This is often referred to as a zero-inflated binomial (ZIB) distribution because it is a mixture of a binomial distribution and a point-mass at zero. Estimation is done through maximizing the above likelihood providing parameter estimates for p and the β_ks which establishes a quantifiable relationship between occurrence, ψ_i and site-specific covariates, x_i.

The estimated relationship between occurrence and site-specific covariates is used to inform the next phase of sampling. We first predict occupancy probability at the remaining non-sampled locations relying on the assumption that covariates are fully observable at all locations within the study area. Define m as the remaining non-sampled locations such that $m = N - n_1$ and $\widetilde{\psi_J}$ as the predicted probability of occurrence at site j, $j = 1 \dots m$ with observable covariate x_j such that:

$$\widetilde{\psi_J} = \frac{\exp(\widehat{\beta_0} + \widehat{\beta_1} * x_j)}{1 + \exp(\widehat{\beta_0} + \widehat{\beta_1} * x_j)}. \qquad (3.4)$$

We next calculate the probability of inclusion for each site j as:

$$\pi_j = \frac{\widetilde{\psi_J}}{\sum_{j=1}^{m} \widetilde{\psi_J}} \qquad (3.5)$$

The second phase of sampling consists of selecting a sample size n_2 from the remaining m sites where each site j has the associated weight π_j. Estimation of the associated parameters p and the β_ks is then done by maximizing the likelihood using the augmented data ($n_1 + n_2$).

Proportion of Area Occupied or Finite-Sample Estimation

We are particularly interested in providing support and guidance for studies involving rare or endangered species where interest lies in making sound inference about a local population (finite-sample). Therefore we focus on the setting where our geographic extent consists of a finite number of sites and we want to predict the proportion of total sites that are occupied (proportion of area occupied, PAO) or the finite-sample occupancy rate as opposed to the occurrence probability associated with a theoretically infinite population from which a selection of sites has been sampled (Royle and Dorazio 2008). This is a trivial calculation when using Bayesian estimation, where we can sum up the latent occupancy states, z_i directly:

$\psi^{fs} = \frac{1}{n}\sum_{i=1}^{n} z_i$, where “fs” is for finite sample. However it is not so trivial when using maximum likelihood estimation. One approach suggested by Royle and Dorazio (2008) is to expand the ZIB likelihood to include a new parameter for the total number of occupied sites, N^{tot} and estimate it directly. This results in a joint likelihood based on the trinomial distribution for the three parameters ψ, p, and N^{tot}:

$$L(\psi, p, N^{tot} | \boldsymbol{y}, N, n) = \left[\frac{N^{tot}!}{(N^{tot}-n)!} p^{\sum_{i=1}^{n} y_i} (1-p)^{J*N-\sum_{i=1}^{n} y_i}\right] \text{x}$$

$$\left[\frac{N!}{N^{tot}!(N-N^{tot})!} \psi^{N^{tot}} (1-\psi)^{N-N^{tot}}\right].$$

The joint likelihood can be maximized to obtain estimates of the three parameters ψ, p, and N^{tot}.

We chose to use a slightly different approach, wherein we work directly with the latent states z_i, and the probability of occupancy at each site. It is therefore necessary to differentiate among three different types of sites: 1) those that were sampled and a detection occurred, 2) those that were sampled and no detections occurred, and 3) those that were not sampled. For the first type of site we know that the underlying latent occupancy state $z = 1$ because a detection occurred. The second type of site could result from either the site being unoccupied or the site being occupied but no detections occurred. We can show this probabilistically using Bayes rule:

$$\Pr(z_i | y_i = 0) = \frac{\psi(1-p)^J}{\psi(1-p)^J + (1-\psi)}, \tag{3.6}$$

where $\psi = \frac{\exp(\widehat{\beta_0} + \widehat{\beta_1} * x_i)}{1 + \exp(\widehat{\beta_0} + \widehat{\beta_1} * x_i)}$ for the i^{th} site. The third type of site is simply a prediction to a non-sampled location j, given a new covariate value and estimates of the βs so that the probability of occupancy $\widetilde{\psi_J}$, is (Eq. 4):

$$\widetilde{\psi_J} = \frac{\exp(\widehat{\beta_0} + \widehat{\beta_1} * x_j)}{1 + \exp(\widehat{\beta_0} + \widehat{\beta_1} * x_j)}. \tag{3.4}$$

We can calculate an estimate of N^{tot} by summing up all of the different probabilities at each of the three different types of sites:

$$N^{tot} = \sum_N \{\Pr(z_i = 1 | y_i > 1, x_i) + \Pr(z_i | y_i = 0, x_i) + \Pr(z_j | x_j)\}. \tag{3.7}$$

Note that this approach is similar to the Bayesian approach because we are summing up the latent states, but by working directly with the probabilities we avoid a second layer of stochastic variation because we do not calculate the z_is directly (1s or 0s from a Bernoulli trial).

Predictive ability and error rate

We develop a measure of predictive ability to compare our new two-phase approach with the traditional single-season approach. We define the following test statistic, *d*, to quantity the lack of fit. Let N = the total number of sites in the sampling frame, x_i be the observable covariate for site *i* and let z_i be the true latent state of the i^{th} site while $\widetilde{z_i}$ is the estimated/predicted latent occupancy state (from above). Note that $\Pr(z_i = 1|\psi) = \psi$ and $\Pr(z_i = 0|\psi) = 1 - \psi$. Next define the following:

$$t_{11} = \sum_{i=1}^{N} \Pr(\widetilde{z_i} = 1 \mid x_i, y_i) * I(z_i = 1) \quad (3.8)$$

$$t_{10} = \sum_{i=1}^{N} \Pr(\widetilde{z_i} = 1 \mid x_i, y_i) * I(z_i = 0) \quad (3.9)$$

$$t_{01} = \sum_{i=1}^{N} \Pr(\widetilde{z_i} = 0 \mid x_i, y_i) * I(z_i = 1) \quad (3.10)$$

$$t_{00} = \sum_{i=1}^{N} \Pr(\widetilde{z_i} = 0 \mid x_i, y_i) * I(z_i = 0). \quad (3.11)$$

Note that the *I*() is the indicator argument taking the value of 1 if the argument is true and 0 otherwise. Because we are interested in the lack of fit we can define the following test statistic:

$$d = t_{10} + t_{01} \quad (3.12)$$

such that

$$d = \sum_{i=1}^{N} \Pr(\widetilde{z_i} = 1 \mid x_i, y_i) * I(z_i = 0) + \sum_{i=1}^{N} \Pr(\widetilde{z_i} = 0 \mid x_i, y_i) * I(z_i = 1)$$

$$= \sum_{i=1}^{N} \Pr(\widetilde{z_i} = 1 \mid x_i, y_i) * I(z_i = 0) + \sum_{i=1}^{N} [1 - \Pr(\widetilde{z_i} = 1 \mid x_i, y_i)] * I(z_i = 1). \quad (3.13)$$

Again we need to differentiate among the three different types of sites, 1) sites where a detection occurred:

$$\Pr(\widetilde{z_i} = 1 \mid x_i, y_i > 0) = 1, \quad (3.14)$$

2) sites that were sampled, but no detection occurred:

$$\Pr(\tilde{z}_i = 1 \mid x_i, y_i = 0) = \frac{\hat{\psi}(1-\hat{p})^J}{\hat{\psi}(1-\hat{p})^J + 1 - \hat{\psi}}, \tag{3.15}$$

and, 3) sites that were not sampled:

$$\Pr(\tilde{z}_i = 1 \mid x_i) = \hat{\psi}. \tag{3.16}$$

Simulations

We used simulations to compare our two-phase approach with the traditional single-season occupancy approach under a range of known patterns in occupancy and design criteria. We were interested in exploring the usefulness of our design when patterns in occupancy are spatially correlated through the use of habitat covariates and overall occupancy rate is relatively low. We therefore created three different habitat types with varying degrees of spatial correlation on a 20 x 20 grid (Figure 3.1). The first consisted of three blocks of habitat: low, medium, and high quality with a small amount of random noise added (standard normal deviate) to each block. We considered habitat 1 as an example of extreme spatial correlation. The second type of habitat was generated using a Matérn cluster process (Matérn 1986; Møller and Waagepetersen 2003; Baddeley and Turner 2005). The Matérn cluster process is a doubly-stochastic or two-stage model for point generation and consists of first defining a parent Poisson process with some mean intensity κ. Next, within a radius r of each parent process a second Poisson process with mean μ is generated. This creates a clustering of points around each of the parents with a fine scale of control to manipulate the amount and degree of clustering. We used this to generate spatially correlated habitat by defining $\kappa = 0.09$, $r = 2.3$, and $\mu = 22$. We considered habitat 2 as an example of moderate spatial correlation. The third habitat type had no spatial correlation and was generated as completely random. The true occupancy rates for each habitat were calculated by specifying a logit-linear model with the simulated habitat covariates (Eq. 1):

$$logit(\psi_i) = \beta_0 + \beta_1 * x_i + \varepsilon,$$

where x_i is the habitat covariate at each site, $\varepsilon \sim N(0,1)$ and we fixed $\beta_0 = -2$ and $\beta_1 = 2$. A second set of habitats were created that were identical to the first three habitats except we set $\beta_1 = 0$ to mimic the situation when the proposed covariate has no effect on occupancy (Figure 3.1).

To explore the new approach under different design criteria we varied the overall detection probability p (0.25, 0.5, or 0.75), the number of repeat visits to a site J (3, 5), and the overall sample size. We fixed the sample size n for the traditional single-season occupancy model (100 or 200) and then varied the proportion allocated to either the first or second phase of sampling for the two-phase approach (25%, 50%, or 75% to the first phase). We generated 1000 synthetic data sets for each of the combinations of parameters and habitat structures. We explored the overall error rate as calculated by the test statistic d along with bias and mean-squared error for the parameters p, β_0, β_1, and N^{tot} (the total number of occupied sites out of the 400 sites). All simulations and analysis were done using the software package R v. 2.12 (R Development Core Team 2010) and all code is provided in Appendix A.

Results

We present selected results and the remainder of the results can be found in Appendix B. The true occupancy rate for the three different habitats ranged from 0.33 to 0.39. The overall error rate regardless of sampling approach decreased as sample size increased and as detection probability increased. There was little effect of increasing the number of repeat visits (J) on the overall error rate. Allocating 25% to the first phase of sampling for the two-phase approach resulted in the lowest overall error rate and the difference was the largest between the two approaches in Habitat 1 (Figures 3.2 and 3.3, Appendix B). In an environment with extreme spatial correlation (Habitat 1) the gain in predictive ability for the two-phase approach was most noticeable when detection probability was low (0.25) and 25% of the sample size was allocated to the first phase of sampling. In an environment with moderate spatial correlation (Habitat 2) there was little improvement for the two-phase approach compared to the traditional single-season occupancy approach regardless of the sample size and detection probability (Figures 3.2 and 3.3, Appendix B). In the random habitat with no simulated spatial correlation (Habitat 3) the two-phase approach outperformed the traditional single-season occupancy approach under most circumstances (Figures 3.2 and 3.3, Appendix B). The greatest difference between the two approaches in Habitat 3 was when the overall sample size was 200 and detection probability was high (0.75). When there was no simulated effect of habitat on occupancy, the two approaches had very similar overall error rates which were much higher relative to the scenarios with a simulated effect of habitat on occupancy (Figure 3.4, Appendix B).

The overall distributions of estimates for the total number of occupied sites, N^{tot}, were very similar for the two approaches (Figure 3.5, Appendix B). The two-phase approach had a tighter distribution around the true value for lower sample size (n=100), but this effect was generally non-existent when the sample size increased (Figure 3.5, Appendix B). This effect was most noticeable when 50% of the sample size was allocated to each phase of sampling although this was highly variable and depended on the habitat (Appendix B).

The overall bias and mean-squared error (MSE) tended to decrease for both approaches as sample size, number of repeat visits, and detection probability all increased (Tables 3.1-3.4, Appendix B). When estimating N^{tot} , Bias and MSE were lowest when using the two-phase approach with either 25% or 50% allocated to the first phase of sampling regardless of the sample size, habitat, or detection probability. Estimation in habitat 1 resulted in the lowest bias and MSE for both approaches. Bias in estimating N^{tot} was generally positive for the traditional single-season occupancy approach and generally negative for the two-phase approach. Overall estimation of detection probability, p, was very accurate regardless of the approach or habitat (Appendix B). Both approaches did a better job of estimating β_1 than β_0 from the logit-linear model describing occupancy (Tables 3.3 and 3.4), but there was little difference between the two approaches.

Discussion

Our goal is to provide a logical and coherent approach to adaptively allocating resources to improve the predictive ability of occupancy modeling and estimation. We found that our two-phase sampling approach has the potential to reduce the overall error rate by a large amount under certain scenarios. When spatial correlation was extreme or when there was no spatial correlation, we found our approach to be an improvement over the traditional simple-random-sampling involved with the single-season occupancy approach although we found no benefit of the two-phase approach when there was only moderate spatial correlation. This suggests that the benefit is highly dependent upon the structure of the habitat and the amount and type of spatial correlation. Fortunately, our new approach did not result in consistently higher error rates under any circumstance. This provides evidence that the approach is fairly robust to a broad range of conditions and design factors and merits use under a wide variety of settings. Under the worst case scenarios the two-phase approach resulted in similar error rates as the traditional single-season occupancy model with simple-random-sampling.

Although we explored the merit of our approach under a range of conditions and design criteria we kept total occupancy rate relatively low. As the occupancy rate increases we posit that the overall benefit of this approach may decrease under certain conditions. If the overall occupancy rate is high, we believe a simple-random-sample will do a sufficient job of estimating occupancy and reallocating effort will likely not improve the overall error rate. We therefore see our method as being most applicable for studies of rare species that tend to cluster or aggregate across the landscape making sampling difficult (Thompson 2004).

Our approach is dependent upon the availability of covariate information across the entire landscape. Although this initially may seem restrictive we believe that such information is readily available for a wide range of locations. Remotely-sensed data is becoming increasingly more available and several recent studies show the diversity of questions that involve remotely-sensed data. These include the assessment of gene flow and genetic differentiation (Weigel et al. 2003; Alberto et al. 2010), determining the influence of stream characteristics on introgression in trout (Bennett et al. 2010), evaluating the efficacy of marine protected areas (Friedlander et al. 2007), predicting species distributions (Raxworthy et al. 2003, Spens et al. 2007), and mapping species' habitat (Rotenberry et al. 2006). This suggests that there is available geographic information across a diverse set of landscapes and therefore we do not see the requirement of available covariate information throughout the study area as restricting the utility of our approach.

Our simulations suggest that approximately 25% of the total sample size should be allocated to the first phase of sampling. This design resulted in the greatest improvement over the traditional approach under a wide range of conditions. Not only is this potentially useful for reallocating effort, it is also easy to implement because many studies begin with a pilot study in which a small number of sites are surveyed. The use of a pilot study can not only provide information about required sample sizes to meet a predetermined level of precision, but by using our approach it can provide a more efficient use of resources. The reallocation of effort to areas where more individuals exist can increase the information content of the sample thus improving the overall knowledge about the system of interest. This can often lead to a reduction of uncertainty which can inform the potential conservation actions that must be initiated with limited resources (McDonald-Madden et al. 2008).

There has been an increase in the awareness and utility of incorporating decision making into conservation and natural resource management (Possingham 1997; Conroy et al. 2002; Williams et al. 2002; Dorazio and Johnson 2003; Nichols and Williams 2006; Martin et al. 2009; McDonald-Madden 2010; Conroy et al. 2011) although much of the literature has focused on implementing insulated decisions that involve manipulating or eliciting a response from the species of interest. We believe our framework further advances the marriage of monitoring, decision making and conservation/management by integrating scientific objectives with the decisions necessary to allocate resources. Although decisions that elicit a direct response of the system are the most rewarding there are also a set of important decisions regarding study design and allocation of resources. Other authors have found that study design can have a significant effect on parameter estimation and inference (Bailey et al. 2007; MacKenzie and Royle 2005; Guillera-Arroita et al. 2010) thus influencing the decisions to be made. Therefore decisions about study design should be included in objective-driven science or management creating a more robust set of available decision alternatives to be implemented.

Our approach provides a framework that facilitates explicitly defining and stating *a priori* hypotheses about habitat-occupancy relationships before the study starts. It forces biologists/ecologists and natural resource managers to clearly define such relationships which are often based on prior information regarding the distribution and spatial variation of individual species. This in turn provides a mechanism for listing competing models which describe variation in occupancy and allows for the exploration of these relationships during the first phase of our approach. The second phase provides guidance on effectively concentrating resources to the information richest locations thus forming a focused monitoring effort for conservation which can be more effective (Nichols and Williams 2006). In addition the two phase approach permits the integration of *a priori* information into the study in a natural and coherent manner.

We believe our two-phase approach provides a structured and coherent framework allowing for a more effective use of resources for a wide variety of occupancy based studies. Our approach places a premium on clearly defining objectives and biological hypotheses before a study begins thus echoing others who have stressed the importance of objective-driven science and management (Yoccoz et al. 2001). In addition our approach forces biologists/ecologists and natural resource managers to acknowledge the fundamental constraints limiting available resources by potentially incorporating design criteria into decision making. Not only does this bring attention to the usefulness of conservation planning, in addition, the adaptive phase of reallocation can potentially concentrate resources to obtain the highest conservation value per unit of effort.

Literature Cited

Alberto, F., P. T. Raimondi, D. C. Reed, N. C. Coelho, R. Leblois, A. Whitmer, and E. Serrao. 2010. Habitat continuity and geographic distance predict population genetic differentiation in giant kelp. Ecology 91: 49-56.

Baddeley, A. and R. Turner. 2005. Spatstat: an R package for analyzing spatial point patterns. Journal of Statistical Software 12: 1-42.

Bailey L, Hines J, Nichols J, MacKenzie D. 2007. Sampling design trade-offs in occupancy studies with imperfect detection: examples and software. Ecological Applications 17: 281 – 290.

Bennett, S. N., J. R. Olson, J. L. Kershner, and P. Corbett. 2010. Propagule pressure and stream characteristics influence introgression: cutthroat and rainbow trout in British Columbia. Ecological Applications 20: 263-277.

Conroy, M. J., M. W. Miller, and J. E. Hines. 2002. Identification and synthetic modeling of factors affecting American black duck populations. Wildlife Monographs 150.

Conroy, M. J., M. C. Runge, J. D. Nichols, K. W. Stodola, and R. J. Cooper. 2011. Conservation in the face of climate change: the roles of alternative models, monitoring, and adaptation in confronting and reducing uncertainty. Biological Conservation 144: 1204-1213.

Dorazio, R. M., and F. A. Johnson. 2003. Bayesian inference and decision theory- a framework for decision making in natural resource management. Ecological Applications 13: 556-563.

Dorazio, R. M., and J. A. Royle. 2005. Estimating size and composition of biological communities by modeling the occurrence of species. Journal of the American Statistical Association 100: 389-398.

Dorazio, R. M., J. A. Royle, B. Soderstrom, and A. Glimskar. 2006. Estimating species richness and accumulation by modeling species occurrence and detectability. Ecology 87: 842-854.

Field, S. A., A. J. Tyre, and H. P. Possingham. 2005. Optimizing allocation of monitoring effort under economic and observational constraints. Journal of Wildlife Management 69: 473-482.

Friedlander, A. M., E. K. Brown, and M. E. Monaco. 2007. Coupling ecology and GIS to evaluate efficacy of marine protected areas in Hawaii. Ecological Applications 17: 715-730.

Guillera-Arroita, G, Ridout, M, Morgan B. 2010. Design of occupancy studies with imperfect detection. Methods in Ecology and Evolution 1: 131-139.

Kery, M., and J. A. Royle. 2009. Inference about species richness and community structure using species-specific occupancy models in the National Swiss Breeding Bird Survey MHB. In "Modeling Demographic Processes in Marked Populations. Eds. D. L. Thompson. E. G. Cooch, and M. J. Conroy. Pgs. 639-656. Springer, New York, New York.

MacKenzie, D. I., J. D. Nichols, G. B. Lachman, S. Droege, J. A. Royle, and C. A. Langtimm. 2002. Estimating site occupancy rates when detection probabilities are less than one. Ecology 83: 2248-2255.

MacKenzie, D. I., J. D. Nichols, J. E. Hines, M. G. Knutson, and A. B. Franklin. 2003. Estimating site occupancy, colonization, and local extinction when a species is detected imperfectly. Ecology 84: 2200-2207.

MacKenzie, D. I., J. D. Nichols, N. Sutton, K. Kawanishi, and L. L. Bailey. 2005. Improving inference in population studies of rare species that are detected imperfectly. Ecology 86: 1101-1113.

MacKenzie, D. I., J. D. Nichols, J. A. Royle, K. H. Pollock, L. L. Bailey, and J. E. Hines. 2006. Occupancy estimation and modeling. Academic Press, New York, New York.

MacKenzie, D. I., and J. A. Royle. 2005. Designing occupancy studies: general advice and allocating survey effort. Journal of Applied Ecology 42:1105-1114.

Martin, J., M. C. Runge, J. D. Nichols, B. C. Lubow, and W. L. Kendall. 2009. Structured decision making as a conceptual framework to identify thresholds for conservation and management. Ecological Applications 19:1079-1090.

Martin, T. G., B. A. Wintle, J. R. Rhodes, P. M. Kuhnert, S. A. Field, S. J. Low-Choy, A. J. Tyre, and H. P. Possingham. 2005. Zero tolerance ecology: improving ecological inference by modeling the source of zero observations. Ecology Letters 8: 1235-1246.

Matérn B. 1986. Spatial Variation. Lecture Notes in Statistics, Vol. 36. Springer-Verlag. New York, New York.

McDonald-Madden, E., W. J. M. Probert, C. E. Hauser, M. C. Runge, H. P. Possingham, M. E. Jones, J. L. Moore, T. M. Rout, P. A. Vesk, and B. A. Wintle. 2010. Active adaptive conservation of threatened species in the face of uncertainty. Ecological Applications 20: 1476-1489.

McDonald-Madden, E., P. W. J. Baxter, and H. P. Possingham. 2008. Making robust decisions for conservation with restricted money and knowledge. Journal of Applied Ecology 45: 1630-1638.

Møller J. and R. P. Waagepetersen. 2003. Statistical Inference and Simulation for Spatial Point Processes. Chapman Hall/CRC, Boca Raton, Florida.

Nichols, J. D., and B. K. Williams. 2006. Monitoring for conservation. Trends in Ecology and Evolution 21:668-673.

Possingham, H. P. 1997. State-dependent decision analysis for conservation biology. In "The Ecological Basis of Conservation: Heterogeneity, Ecosystems, and Biodiversity". Eds. S. T. A. Pickett, R. S. Ostfeld, M. Shachak and G. E. Likens. Pp. 298-304. Chapman and Hall, New York, New York.

Prendergast, J. R., R. M. Quinn, J. H. Lawton, B. C. Eversham, and D. W. Gibbons. 1993. Rare species, the coincidence of diversity hotspots and conservation strategies. Nature 365: 335-337.

R Development Core Team. 2010. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org.

Raxworthy, C. J., E. Martinez-Meyer, N. Horning, R. A. Nussbaum, G. E. Schneider, M. A. Ortega-Huerta, and A. T. Peterson. 2003. Predicting distributions of known and unknown reptile species in Madagascar. Nature 426: 837-841.

Rotenberry, J. T., K. L. Preston, and S. T. Knick. 2006. GIS-based niche modeling for mapping species' habitat. Ecology 87: 1458-1464.

Royle J. A. and R. M. Dorazio. 2008. Hierarchical modeling and inference in ecology. Academic Press, London.

Saracco, J. F., R. B. Siegel, and R. L. Wilkerson. 2011. Occupancy modeling of Black-backed Woodpeckers on burned Sierra Nevada forests. Ecosphere 2: 1-17.

Spens, J., G. Englund, and H. Lundqvist. 2007. Network connectivity and dispersal barriers: using geographical information system (GIS) tools to predict landscape scale distribution of a key predator (*Esox Lucius*) among lakes. Journal of Applied Ecology 44: 1127-1137.

Thompson, W. L, editor. 2004. Sampling Rare or Elusive Species. Island Press, Washington, D. C.

Weigel, D. E., J. T. Peterson, and P. Spruell. 2003. Introgressive hybridization between native cutthroat trout and introduced rainbow trout. Ecological Applications 13: 38-50.

Williams, B. K., J. D. Nichols, and M. J. Conroy. 2002. Analysis and management of animal populations. Academic Press, San Diego.

Wilson, K. A., M. F. McBride, M. Bode, and H. P. Possingham. 2006. Prioritizing global conservation efforts. Nature 440: 337-340.

Yoccoz, N. G., J. D. Nichols, and T. Boulinier. 2001. Monitoring of biological diversity in space and time. Trends in Ecology and Evolution 16: 446-453.

Zipkin, E. F., A. DeWan, and J. A. Royle. 2009. Impacts of forest fragmentation on species richness: a hierarchical approach to community modeling. Journal of Applied Ecology 46: 815-822.

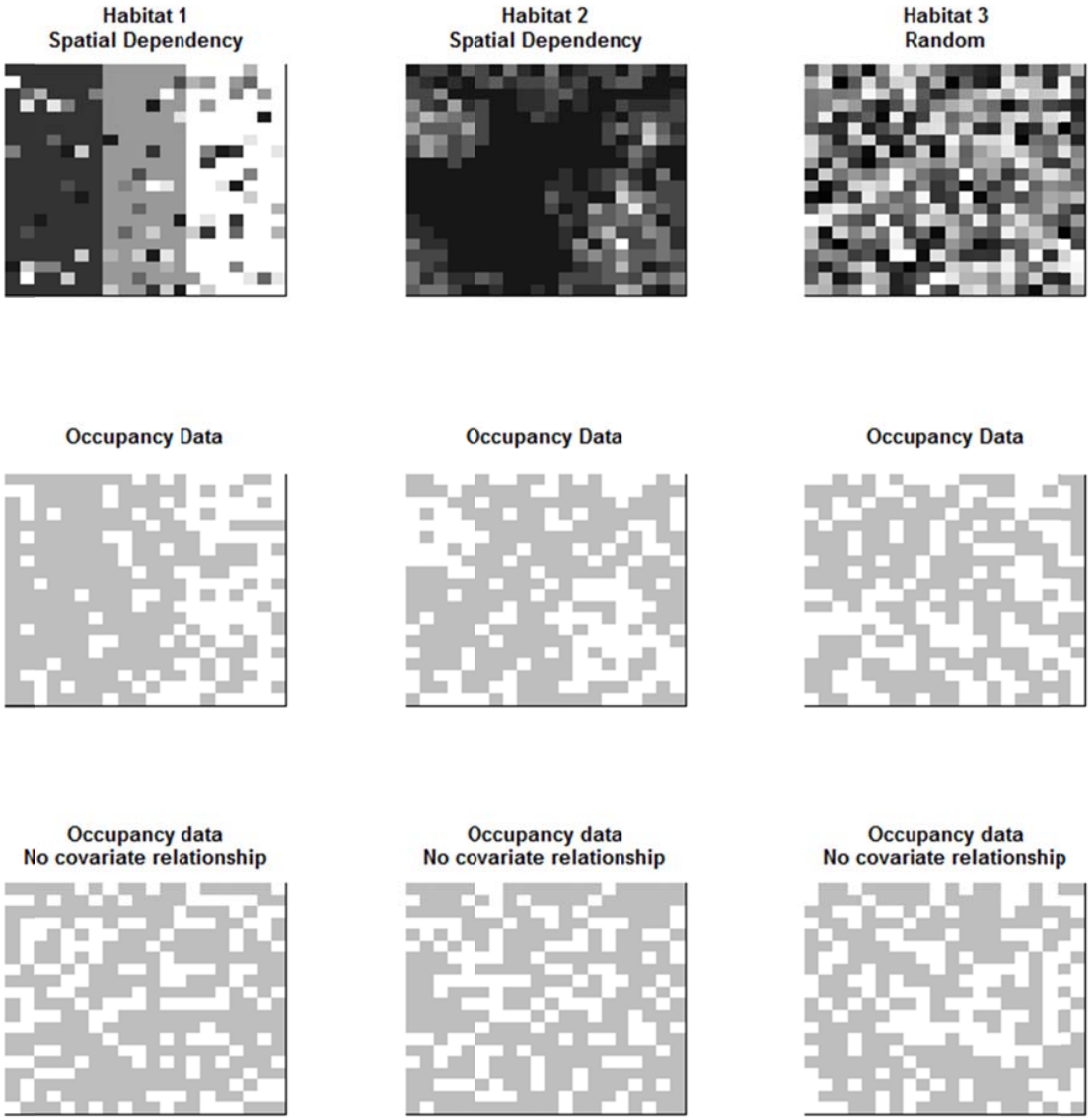

Habitat 1
Spatial Dependency
Habitat 2
Spatial Dependency
Habitat 3
Random
Occupancy Data
Occupancy Data
Occupancy Data
Occupancy data
No covariate relationship
Occupancy data
No covariate relationship
Occupancy data
No covariate relationship

Figure 3.1. Habitat covariates (row 1) and associated occupancy data (rows 2 and 3) for three different types of simulated environments. The first habitat represents extreme spatial correlation, the second habitat represents moderate spatial correlation and the third habitat was generated randomly and contains no spatial correlation. The true occupancy rates for each habitat (rows 2 and 3) were calculated by specifying a logit-linear model with the simulated habitat covariates with the slope in the equation equal to 2 (row 2) or 0 (row 3; mimics no habitat effect on occupancy). Lighter colors represent higher quality habitat for row 1 and occupied for rows 2 and 3.

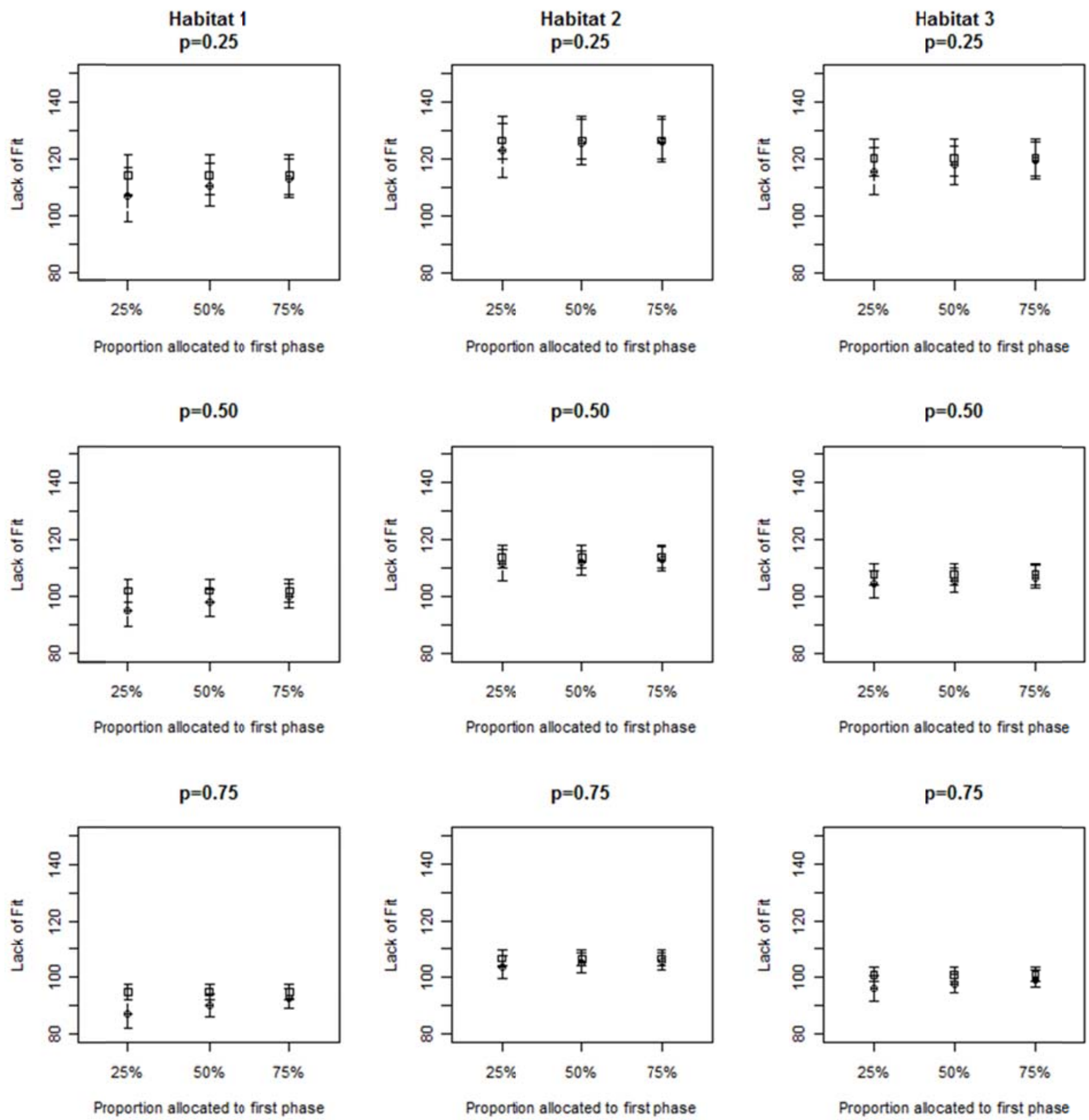

Habitat 1
p=0.25
Habitat 2
p=0.25
Habitat 3
p=0.25
p=0.50
p=0.50
p=0.50
p=0.75
p=0.75
p=0.75
Lack of Fit
80
100
120
140
25%
50%
75%
Proportion allocated to first phase

Figure 3.2. Plots of test statistic and associated quartiles (25% and 75%) measuring lack of fit (lower is better) for two-phase adaptive approach (circles) and traditional single-season occupancy approach (squares) when n=100, J=3, and for three different levels of detection, p=0.25, 0.5, 0.75. X-axis represents proportion of sample allocated to phase one for the two-phase adaptive approach. Habitat 1 is a simulated habitat with extreme spatial correlation while habitat 2 has moderate spatial correlation and habitat 3 is randomly generated and contains no spatial correlation.

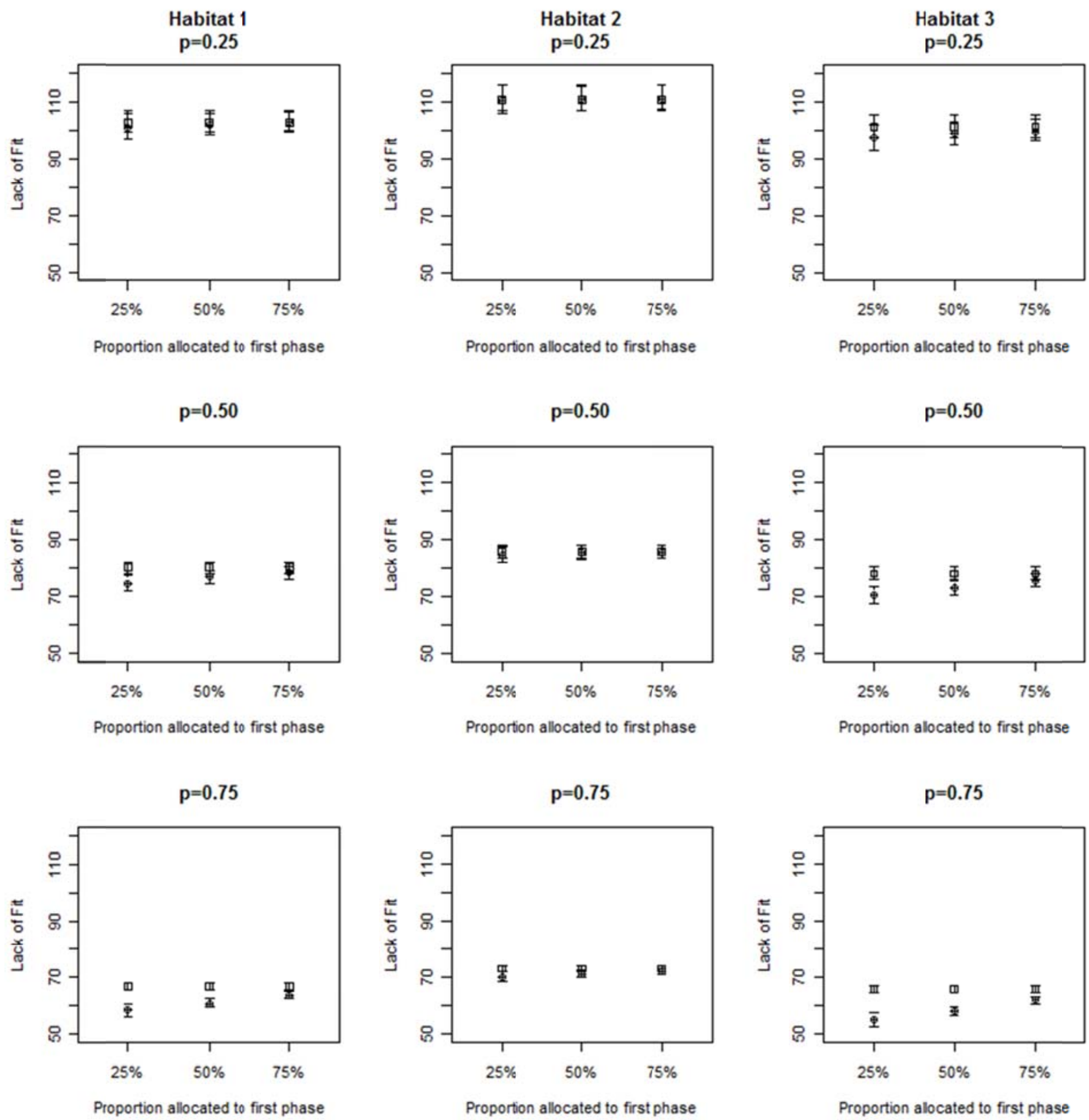
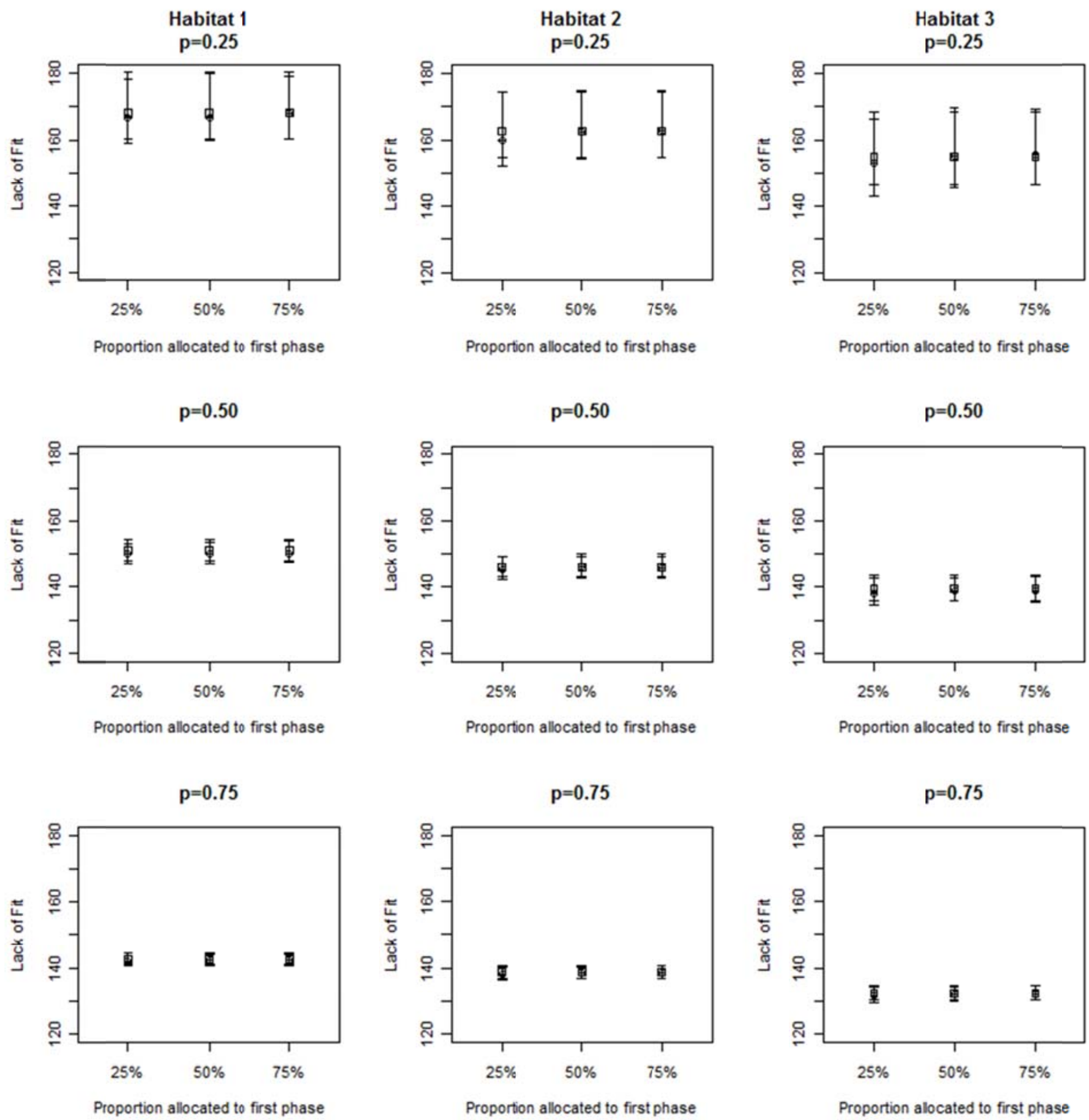

Habitat 1
p=0.25
Habitat 2
p=0.25
Habitat 3
p=0.25
p=0.50
p=0.50
p=0.50
p=0.75
p=0.75
p=0.75
Lack of Fit
50
70
90
110
25%
50%
75%
Proportion allocated to first phase

Figure 3.3. Plots of test statistic and associated quartiles (25% and 75%) measuring lack of fit (lower is better) for two-phase adaptive approach (circles) and traditional single-season occupancy approach (squares) when n=200, J=3, and for three different levels of detection, p=0.25, 0.5, 0.75. X-axis represents proportion of sample allocated to phase one for the two-phase adaptive approach. Habitat 1 is a simulated habitat with extreme spatial correlation while habitat 2 has moderate spatial correlation and habitat 3 is randomly generated and contains no spatial correlation.



Habitat 1
p=0.25
Habitat 2
p=0.25
Habitat 3
p=0.25
p=0.50
p=0.50
p=0.50
p=0.75
p=0.75
p=0.75
Lack of Fit
120
140
160
180
25%
50%
75%
Proportion allocated to first phase

Figure 3.4. Plots of test statistic and associated quartiles (25% and 75%) measuring lack of fit (lower is better) for two-phase adaptive approach (circles) and traditional single-season occupancy approach (squares) when there is no simulated relationship between habitat and occupancy for n=100, J=3, and for three different levels of detection, p=0.25, 0.5, 0.75. X-axis represents proportion of sample allocated to phase one for the two-phase adaptive approach. Habitat 1 is a simulated habitat with extreme spatial correlation while habitat 2 has moderate spatial correlation and habitat 3 is randomly generated and contains no spatial correlation.

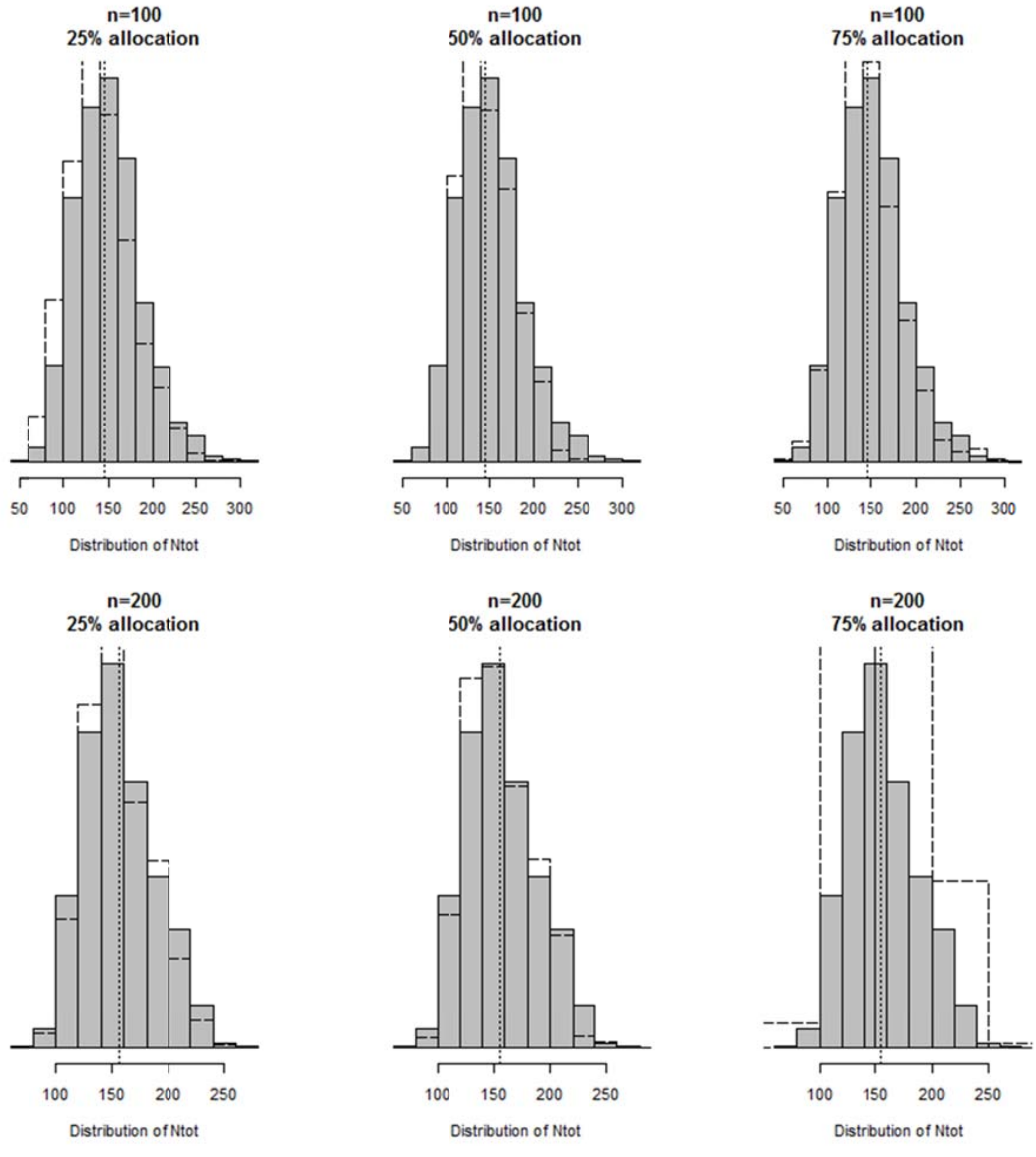

n=100
25% allocation
n=100
50% allocation
n=100
75% allocation
50 100 150 200 250 300
Distribution of Ntot
n=200
25% allocation
n=200
50% allocation
n=200
75% allocation
100 150 200 250
Distribution of Ntot

Figure 3.5. Distribution from 1000 simulations of the estimates of N^{tot} for the two-phase adaptive approach (empty dotted rectangles) and the traditional single-season occupancy approach (gray filled rectangles) in Habitat 1 (extreme spatial correlation, see text for more details) with the number of repeat visits, J, equal to 3 and detection probability, p, equal to 0.25. Percent allocation refers to the percentage of the total sample size, n, allocated to the first phase of sampling for the two-phase adaptive approach. The dotted line represents the true value of N^{tot}.

Table 3.1. Average value from 1000 simulations of test statistic measuring lack of fit, ***d***, test statistic measuring goodness of fit, ***GOF***, estimate of N^{tot}, $\widehat{N^{tot}}$, bias associated in estimating N^{tot}, and mean-squared-error, ***MSE***, associated with estimating N^{tot}. N^{tot} represents the true value of the total number of sites occupied out of 400. n_1 and n_2 represent the sample size at each phase of sampling. n_2 = 0 represents the traditional single-season occupancy approach. ***Habitat*** represents three different habitats with varying degrees of spatial correlation (1-extreme spatial correlation, 2-moderate spatial correlation, 3-no spatial correlation).

n_1	n_2	p	*Habitat*	d	*GOF*	$\widehat{N^{tot}}$	*Bias* $\widehat{N^{tot}}$	*MSE* $\widehat{N^{tot}}$	N^{tot}
100	0	0.25	1	114.755	285.245	150.480	5.340	1407.308	145.140
100	0	0.5	1	102.159	297.841	147.745	2.605	320.458	145.140
100	0	0.75	1	95.062	304.938	147.913	2.773	218.215	145.140
25	75	0.25	1	107.554	292.446	140.491	-4.649	1314.674	145.140
25	75	0.5	1	95.307	304.693	142.073	-3.066	409.661	145.140
25	75	0.75	1	87.063	312.937	141.484	-3.655	305.384	145.140
50	50	0.25	1	111.088	288.912	144.188	-0.952	1035.041	145.140
50	50	0.5	1	98.113	301.887	146.220	1.080	310.262	145.140
50	50	0.75	1	90.353	309.647	146.240	1.100	201.638	145.140
75	25	0.25	1	113.701	286.299	145.722	0.582	1206.157	145.140
75	25	0.5	1	100.375	299.625	147.991	2.852	295.876	145.140
75	25	0.75	1	92.861	307.139	147.135	1.995	201.181	145.140
100	0	0.25	2	129.117	270.883	149.873	8.873	2032.738	141.000
100	0	0.5	2	114.089	285.911	142.655	1.655	366.643	141.000

100	0	0.75	2	106.980	293.020	141.005	0.005	214.669	141.000
25	75	0.25	2	124.113	275.887	139.564	-1.436	1541.001	141.000
25	75	0.5	2	111.311	288.689	139.684	-1.316	418.092	141.000
25	75	0.75	2	103.762	296.238	139.074	-1.926	272.805	141.000
50	50	0.25	2	127.072	272.928	146.771	5.771	1773.656	141.000
50	50	0.5	2	112.168	287.832	141.025	0.025	331.710	141.000
50	50	0.75	2	105.305	294.695	140.119	-0.881	227.736	141.000
75	25	0.25	2	127.516	272.484	146.480	5.480	1584.779	141.000
75	25	0.5	2	113.423	286.577	141.053	0.053	313.394	141.000
75	25	0.75	2	105.973	294.027	140.368	-0.632	218.706	141.000
100	0	0.25	3	121.151	278.849	157.188	12.976	2044.173	144.211
100	0	0.5	3	108.290	291.710	151.054	6.843	462.077	144.211
100	0	0.75	3	101.308	298.692	149.554	5.342	244.966	144.211
25	75	0.25	3	116.458	283.542	152.960	8.749	1607.953	144.211
25	75	0.5	3	104.595	295.405	148.170	3.959	354.152	144.211
25	75	0.75	3	96.273	303.727	147.198	2.987	236.908	144.211
50	50	0.25	3	118.334	281.666	155.361	11.150	1593.905	144.211
50	50	0.5	3	106.114	293.886	149.212	5.000	369.732	144.211
50	50	0.75	3	98.096	301.904	147.857	3.646	212.632	144.211
75	25	0.25	3	120.279	279.721	158.513	14.302	1877.900	144.211
75	25	0.5	3	107.152	292.848	149.964	5.753	351.897	144.211
75	25	0.75	3	99.690	300.310	149.013	4.802	225.244	144.211

Table 3.2. Average value from 1000 simulations of test statistic measuring lack of fit, ***d***, test statistic measuring goodness of fit, ***GOF***, estimate of N^{tot}, $\widehat{N^{tot}}$, bias associated in estimating N^{tot}, and mean-squared-error, ***MSE***, associated with estimating N^{tot}. N^{tot} represents the true value of the total number of sites occupied out of 400. n_1 and n_2 represent the sample size at each phase of sampling. n_2 = 0 represents the traditional single-season occupancy approach. ***Habitat*** represents three different habitats with varying degrees of spatial correlation (1-extreme spatial correlation, 2-moderate spatial correlation, 3-no spatial correlation).

n_1	n_2	p	*Habitat*	d	*GOF*	$\widehat{N^{tot}}$	*Bias* $\widehat{N^{tot}}$	*MSE* $\widehat{N^{tot}}$	N^{tot}
200	0	0.25	1	103.380	296.620	156.644	0.977	998.222	155.667
200	0	0.5	1	80.143	319.857	148.410	-7.257	197.677	155.667
200	0	0.75	1	66.856	333.144	147.750	-7.917	134.989	155.667
50	150	0.25	1	101.870	298.130	155.348	-0.319	859.942	155.667
50	150	0.5	1	74.834	325.166	148.241	-7.426	185.475	155.667
50	150	0.75	1	58.440	341.560	148.144	-7.523	126.438	155.667
100	100	0.25	1	102.566	297.434	155.041	-0.626	908.924	155.667
100	100	0.5	1	76.823	323.177	147.958	-7.709	179.998	155.667
100	100	0.75	1	61.024	338.976	147.128	-8.539	138.919	155.667
150	50	0.25	1	103.795	296.205	156.860	1.194	1108.037	155.667
150	50	0.5	1	78.430	321.570	148.429	-7.238	186.988	155.667
150	50	0.75	1	63.884	336.116	147.245	-8.421	138.432	155.667
200	0	0.25	2	112.331	287.669	148.367	5.367	912.685	143.000
200	0	0.5	2	86.082	313.918	143.285	0.285	157.894	143.000

200	0	0.75	2	73.242	326.758	143.263	0.263	78.417	143.000
50	150	0.25	2	111.983	288.017	146.835	3.835	1022.134	143.000
50	150	0.5	2	84.731	315.269	143.027	0.027	126.740	143.000
50	150	0.75	2	70.290	329.710	143.038	0.038	80.748	143.000
100	100	0.25	2	112.390	287.610	146.105	3.105	876.839	143.000
100	100	0.5	2	85.803	314.197	143.414	0.414	138.902	143.000
100	100	0.75	2	71.421	328.579	142.920	-0.080	80.428	143.000
150	50	0.25	2	112.167	287.833	145.877	2.877	756.609	143.000
150	50	0.5	2	85.981	314.019	142.878	-0.122	141.002	143.000
150	50	0.75	2	72.336	327.664	143.227	0.227	78.675	143.000
200	0	0.25	3	102.148	297.852	141.142	-9.541	949.983	150.682
200	0	0.5	3	78.291	321.709	134.026	-16.656	412.159	150.682
200	0	0.75	3	66.029	333.971	133.620	-17.063	361.748	150.682
50	150	0.25	3	98.022	301.978	135.821	-14.862	831.152	150.682
50	150	0.5	3	70.565	329.435	133.313	-17.370	408.793	150.682
50	150	0.75	3	54.980	345.020	132.493	-18.189	389.696	150.682
100	100	0.25	3	99.413	300.587	136.957	-13.726	776.537	150.682
100	100	0.5	3	73.198	326.802	133.335	-17.347	417.079	150.682
100	100	0.75	3	58.192	341.808	132.519	-18.163	390.058	150.682
150	50	0.25	3	100.779	299.221	139.387	-11.296	909.738	150.682
150	50	0.5	3	75.740	324.260	133.412	-17.270	411.439	150.682
150	50	0.75	3	62.053	337.947	132.916	-17.767	381.537	150.682

Table 3.3. Average value from 1000 simulations for estimates of coefficients in the logit-linear model for occupancy probability, $\widehat{\boldsymbol{\beta_0}}$ and $\widehat{\boldsymbol{\beta_1}}$, along with average bias and mean-squared-error (***MSE***). $\boldsymbol{n_1}$ and $\boldsymbol{n_2}$ represent the sample size at each phase of sampling. $\boldsymbol{n_2}$ = 0 represents the traditional single-season occupancy approach. ***Habitat*** represents three different habitats with varying degrees of spatial correlation (1-extreme spatial correlation, 2-moderate spatial correlation, 3-no spatial correlation).

n_1	n_2	p	*Habitat*	$\widehat{\beta_0}$	*Bias* $\widehat{\beta_0}$	*MSE* $\widehat{\beta_0}$	$\widehat{\beta_1}$	*Bias* $\widehat{\beta_1}$	*MSE* $\widehat{\beta_1}$
100	0	0.25	1	-0.734	1.266	3.034	2.016	0.016	6.456
100	0	0.5	1	-0.835	1.165	1.461	1.527	-0.473	0.370
100	0	0.75	1	-0.843	1.157	1.421	1.553	-0.447	0.302
25	75	0.25	1	-3.148	-1.148	41.485	4.090	2.090	42.623
25	75	0.5	1	-1.482	0.518	8.417	2.131	0.131	5.417
25	75	0.75	1	-1.496	0.504	9.969	2.113	0.113	6.292
50	50	0.25	1	-1.327	0.673	8.549	2.365	0.365	11.155
50	50	0.5	1	-0.953	1.047	1.431	1.698	-0.302	0.405
50	50	0.75	1	-0.933	1.067	1.287	1.668	-0.332	0.256
75	25	0.25	1	-0.909	1.091	4.774	2.061	0.061	7.362
75	25	0.5	1	-0.859	1.141	1.418	1.604	-0.396	0.645
75	25	0.75	1	-0.881	1.119	1.355	1.598	-0.402	0.271
100	0	0.25	2	-0.397	1.603	5.441	1.767	-0.233	6.242
100	0	0.5	2	-0.701	1.299	1.764	1.247	-0.753	0.690
100	0	0.75	2	-0.728	1.272	1.665	1.245	-0.755	0.648

25	75	0.25	2	-1.164	0.836	6.062	2.516	0.516	15.251
25	75	0.5	2	-0.790	1.210	1.613	1.361	-0.639	0.648
25	75	0.75	2	-0.778	1.222	1.576	1.288	-0.712	0.598
50	50	0.25	2	-0.514	1.486	5.930	2.142	0.142	10.294
50	50	0.5	2	-0.747	1.253	1.659	1.306	-0.694	0.623
50	50	0.75	2	-0.748	1.252	1.624	1.247	-0.753	0.644
75	25	0.25	2	-0.572	1.428	3.413	1.706	-0.294	4.805
75	25	0.5	2	-0.734	1.266	1.681	1.248	-0.752	0.704
75	25	0.75	2	-0.742	1.258	1.635	1.255	-0.745	0.628
100	0	0.25	3	0.046	2.046	32.855	3.769	1.769	133.278
100	0	0.5	3	-0.766	1.234	1.632	1.522	-0.478	0.406
100	0	0.75	3	-0.784	1.216	1.541	1.480	-0.520	0.356
25	75	0.25	3	-2.232	-0.232	65.776	6.486	4.486	301.096
25	75	0.5	3	-0.884	1.116	2.067	1.615	-0.385	1.111
25	75	0.75	3	-0.852	1.148	1.432	1.524	-0.476	0.340
50	50	0.25	3	-0.743	1.257	34.414	5.166	3.166	233.646
50	50	0.5	3	-0.819	1.181	1.526	1.543	-0.457	0.454
50	50	0.75	3	-0.822	1.178	1.468	1.488	-0.512	0.349
75	25	0.25	3	-0.338	1.662	19.262	3.444	1.444	93.541
75	25	0.5	3	-0.788	1.212	1.564	1.524	-0.476	0.434
75	25	0.75	3	-0.800	1.200	1.511	1.487	-0.513	0.362

Table 3.4. Average value from 1000 simulations for estimates of coefficients in the logit-linear model for occupancy probability, $\widehat{\boldsymbol{\beta_0}}$ and $\widehat{\boldsymbol{\beta_1}}$, along with average bias and mean-squared-error (***MSE***). $\boldsymbol{n_1}$ and $\boldsymbol{n_2}$ represent the sample size at each phase of sampling. $\boldsymbol{n_2} = 0$ represents the traditional single-season occupancy approach. ***Habitat*** represents three different habitats with varying degrees of spatial correlation (1-extreme spatial correlation, 2-moderate spatial correlation, 3-no spatial correlation).

n_1	n_2	*p*	***Habitat***	$\widehat{\beta_0}$	***Bias*** $\widehat{\beta_0}$	***MSE*** $\widehat{\beta_0}$	$\widehat{\beta_1}$	***Bias*** $\widehat{\beta_1}$	***MSE*** $\widehat{\beta_1}$
200	0	0.25	1	-0.630	1.370	2.305	1.870	-0.130	3.279
200	0	0.5	1	-0.779	1.221	1.524	1.467	-0.533	0.327
200	0	0.75	1	-0.782	1.218	1.502	1.434	-0.566	0.342
50	150	0.25	1	-0.702	1.298	1.914	2.387	0.387	12.811
50	150	0.5	1	-0.770	1.230	1.551	1.411	-0.589	0.390
50	150	0.75	1	-0.765	1.235	1.551	1.383	-0.617	0.400
100	100	0.25	1	-0.667	1.333	2.470	2.207	0.207	8.918
100	100	0.5	1	-0.777	1.223	1.529	1.422	-0.578	0.373
100	100	0.75	1	-0.786	1.214	1.497	1.404	-0.596	0.376
150	50	0.25	1	-0.636	1.364	2.313	2.063	0.063	6.953
150	50	0.5	1	-0.773	1.227	1.541	1.442	-0.558	0.351
150	50	0.75	1	-0.790	1.210	1.486	1.427	-0.573	0.350
200	0	0.25	2	-0.529	1.471	3.041	1.346	-0.654	0.802
200	0	0.5	2	-0.647	1.353	1.863	1.215	-0.785	0.666
200	0	0.75	2	-0.649	1.351	1.840	1.217	-0.783	0.637

50	150	0.25	2	-0.494	1.506	5.085	1.396	-0.604	1.399
50	150	0.5	2	-0.655	1.345	1.835	1.220	-0.780	0.653
50	150	0.75	2	-0.655	1.345	1.826	1.204	-0.796	0.653
100	100	0.25	2	-0.475	1.525	7.555	1.356	-0.644	2.242
100	100	0.5	2	-0.647	1.353	1.858	1.207	-0.793	0.672
100	100	0.75	2	-0.656	1.344	1.824	1.204	-0.796	0.652
150	50	0.25	2	-0.585	1.415	2.293	1.333	-0.667	0.718
150	50	0.5	2	-0.655	1.345	1.838	1.221	-0.779	0.654
150	50	0.75	2	-0.651	1.349	1.837	1.216	-0.784	0.635
200	0	0.25	3	-0.773	1.227	8.451	2.046	0.046	50.721
200	0	0.5	3	-1.007	0.993	1.019	1.463	-0.537	0.335
200	0	0.75	3	-1.011	0.989	1.000	1.455	-0.545	0.325
50	150	0.25	3	-1.055	0.945	1.151	1.710	-0.290	2.095
50	150	0.5	3	-1.026	0.974	0.995	1.462	-0.538	0.341
50	150	0.75	3	-1.026	0.974	0.979	1.435	-0.565	0.348
100	100	0.25	3	-0.907	1.093	8.973	1.834	-0.166	26.199
100	100	0.5	3	-1.027	0.973	0.990	1.471	-0.529	0.329
100	100	0.75	3	-1.030	0.970	0.969	1.450	-0.550	0.330
150	50	0.25	3	-0.919	1.081	1.836	1.866	-0.134	11.036
150	50	0.5	3	-1.024	0.976	0.989	1.474	-0.526	0.326
150	50	0.75	3	-1.019	0.981	0.987	1.442	-0.558	0.335

CHAPTER 4

OCCUPANCY ESTIMATION WITHIN AN ADAPTIVE SAMPLING DESIGN: EVALUATION OF A BAYESIAN HIERARCHICAL MODEL FOR RARE OR ELUSIVE SPECIES[1]

[1]Pacifici, K., R. M. Dorazio, and M. J. Conroy. To be submitted to *Ecological Applications*.

Abstract

Monitoring state variables of interest such as animal abundance, density or site-occupancy rate is a critical component to many large scale and long-term conservation efforts. Often this information is very difficult to collect especially when working with rare or elusive species. These species exhibit specific characteristics such as low detection rates, patchily distributed, spatially aggregated or clumped distributions, creating unique challenges for these species. Two recent methodological developments have been introduced to circumvent many of these problems: 1) adaptive cluster sampling, and 2) occupancy estimation. Here, we attempt to combine those two approaches leveraging the advantages of each approach to better estimate site-occupancy rate. We develop a Bayesian hierarchical model that integrates traditional adaptive cluster sampling with occupancy estimation allowing for the augmentation of adjacent sites during sample, but still accommodating imperfect detection. We use simulations to evaluate our new model and compare it to traditional occupancy estimation and traditional adaptive cluster sampling under a range of known conditions in population spatial structure, detection probability, and design criteria. We use a simulated environment that mimics a range of spatially correlated habitat characteristics. We found that our new approach outperformed traditional occupancy estimation when detection rates were extremely low and the population was highly spatially clustered. Our new approach provided robust coverage under all conditions making it useful for a variety of species and situations with highly aggregated species. We believe it has potential for improving our ability to make accurate inference for rare species.

Monitoring state variables of interest such as animal abundance, density or site-occupancy rate, is a critical component to many large scale and long-term conservation efforts. Unfortunately, this information can be very challenging to obtain due to limited financial resources and logistical constraints imposed by complex environmental and geographic conditions (Possingham et al. 2001). These problems become amplified when the species of interest does not occupy a large proportion of area, is patchily distributed, occurs in very low numbers, or is very difficult to observe or capture (Thompson 2004; MacKenzie et al. 2005). These specific characteristics, possessed by many rare or elusive species, create unique challenges not only for designing effective surveys, but for conducing analyses as well (Thompson 2004; MacKenzie et al. 2005; Cunningham and Lindenmayer 2005).

Recently there has been an increase in developing specific approaches aimed at providing robust parameter estimates for rare or elusive species. A somewhat arbitrary, but noticeable distinction in many of the current approaches is at what stage of the study the advancement in methodology is concentrated. For example, several recent approaches aim to borrow or leverage information across different species (Alldredge et al. 2007; MacKenzie et al. 2005), multiple spatial scales (Dixon et al. 2005; Nichols et al. 2008), or community characteristics (Zipkin et al. 2009) which can be done at the analysis stage of the study. This is in contrast to other approaches that tailor data collection to profit from a specific behavior of the species of interest. For instance the use of different sampling designs such as stratification (Thompson 2002; Edwards et al. 2005), sequential sampling (Thompson 2002; Thompson 2004), multi-phase sampling (Thompson 2002; Thompson 2004), or adaptive sampling (Thompson 1990; Thompson 2004) can potentially increase the information content in a particular sample as well as provide more efficient estimation by accommodating the non-uniform spatial structure of the species.

Although a focus on either approach (focus on analysis and modeling vs. focus on design and data collection) is equally valuable it is important to consider both in the context of a scientific study (Yoccoz et al. 2001).

The distinction between the modeling stage and the design stage can further be articulated as it bears resemblance to a once-standing division in the field of statistics: model-based inference versus design-based inference (Smith 1994; Kish 1995; Little 2004). Fortunately this division is all but gone as many current approaches benefit from an incorporation of both modes of inference to obtain estimates (e.g., model-assisted survey sampling Särndal et al. 1992, small area estimation Rao 2003, Bayesian modeling Ch. 7 *in* Gelman et al. 2004). Therefore it is worth providing sufficient detail to highlight the advantages and disadvantages of each with the aim to combine these two modes of inference to improve estimation in ecological studies of rare species.

Following Dorazio (1998) and Thompson (1992) model-based inference suggests that the values of the variable of interest from the population are viewed as a realization of a set of random variables. A "superpopulation model" (stochastic model) is assumed describing the distribution of possible realizations of the observed population values. Once the data are collected and thought to be representative of the population, the probabilities of sample selection become unnecessary in matters of inference. Inference is based on the likelihood functions of the unknowns given the sample data and follow the likelihood principle (Berger and Wolpert 1984), requiring all conclusions about the population to be based solely on the observed data in the sample. Advantages of model-based inference include the ability to make use of auxiliary information, evaluation of competing hypotheses about the theoretical relationship among variables, efficient use of sample data, and dealing explicitly with multiple sources of error.

There are a variety of well-developed model-based approaches for parameter estimation including mark-recapture (Williams et al. 2002) and site-occupancy modeling (MacKenzie et al. 2006), both of which incorporate the ability to account for survey or detection bias explicitly. Several disadvantages include the overreliance on often tenuous assumptions (e.g., identically and independently distributed random variables), and the possibility of model misspecification (Williams et al. 2002).

In the design-based approach, probability only enters the estimation process through the use of design-induced probabilities to select one sample over another. Nothing is assumed about the underlying population and inference is only based on hypothetical repetition of selecting sample units. The observable characteristics of a population are regarded as fixed constants, and the idea is to choose a sampling design that will improve the precision of a parameter estimate if anticipated differences in the population are actually realized in the sample. Estimators are closely linked to the sampling design and are usually unbiased regardless of the nature of the population. No assumptions about the data collected are needed to guarantee the unbiasedness of the estimators. In addition the use of design-based inference alleviates to some degree the potential disastrous effects of important but unknown auxiliary variables. Design-based approaches have seen much use in ecology because there is a potential gain in estimator efficiency and performance when used with geographically clustered or rare populations (Brown 2003; Christman 2000; Smith et al. 2003). A potential second advantage of certain design-based estimators is to accommodate observer behavior (Pacifici Chapter 2). Many rare or threatened species are so rare or endangered that any information about them is extremely important (e.g. Ivory-billed Woodpecker, *Campephilus principalis*). Therefore, the ability to mimic observer behavior by putting more effort in areas where individuals have been detected is a potentially important component of an effective sampling design for rare species. Unfortunately, the use of design-based approaches requires an independent estimate of detection probability to account for survey or detection bias (Thompson and Seber 1996; Smith et al. 2010).

Given that both design-based and model-based approaches have unique advantages and disadvantages, an approach that combined properties of both techniques should ideally take advantage of features of both approaches. We acknowledge that others have explored a combined approach (Edwards et al. 2004) with Hines et al. (2010) developing an occupancy-based model that accounted for the induced spatial dependency among adjacent transects that were sampled in a linear fashion. Our motivation originated from a suggestion by Pacifici et al. (Chapter 2) which posited the potential advantages of augmenting the sample design used in occupancy estimation to allow for the inclusion of adjacent sites into the sample and therefore allocating effort to areas surrounding an observed detection. They describe a scenario that would have required the combination of occupancy estimation (MacKenzie et al. 2006) and adaptive cluster sampling (Thompson 1990) and point out the potential advantages of such an approach. MacKenzie and Royle (2005) also suggested the possibility of selecting sites by adaptive sampling leading to reliable inference about occupancy probability. Rapley and Welsh (2008) developed a model-based approach to adaptive cluster sampling that would accommodate the benefits of both approaches, but did not incorporate estimates of detection probability. We therefore develop a statistically rigorous approach that integrates adaptive cluster sampling and occupancy estimation thus allowing for the estimation of detectability into the model. Although a frequentist or Bayesian approach is possible we agree with Little (2004) and Gelman et al. (2004) that Bayesian hierarchical modeling provides a natural, flexible way of incorporating the data collection process into our model structure and thus is an appropriate avenue to develop a model that integrates both design and model-based components. Therefore our objectives were to:

1) develop a Bayesian hierarchical model that integrates occupancy estimation within an adaptive cluster sampling framework while accounting for imperfect detection,

2) assess model diagnostics and model fit using a Bayesian p-value,

3) evaluate the frequentist properties of the model under a range of design scenarios and,

4) compare the performance of the developed model with the traditional single-season occupancy model and the traditional adaptive cluster sampling approach.

Methods

Sampling Overview

We provide a general overview of the sampling framework which combines adaptive cluster sampling with occupancy estimation. We envision N specific sites where an initial simple random sample of size n is taken without replacement from the N sites. On each of the $i = 1 \ldots n$ sites J Bernoulli detection samples are taken resulting in $y_i = 0, 1, 2, \ldots, J_i$ detections per site. Here we assume that J is identical at all of the n sites, but a balanced design is not essential. The J samples can be a result of J independent visits to a particular site or J independent observers surveying a single site. Detections can be visual, by detection of sign (visual or aural), physical captures or any other approach as long as there is no discrepancy about the positive determination of a detection.

Next we define a condition C such that if a particular site satisfies the condition, $y_i > C$, the sites within the neighborhood of y_i are added to the sample. Because we are working with occupancy data we define C =0 although this condition can be generalized to have a higher threshold depending on the characteristics of the species and the objectives of the study (see Discussion). The condition C can be satisfied by the joint observations of the J Bernoulli detection samples such that any one of the J samples can trigger adaptation. Note that the neighborhoods can be defined to have a variety of shapes and do not have to be contiguous, but the neighborhood relationship must be symmetric in that if unit *j'* is in the neighborhood of unit *i'* then unit *i'* must be in the neighborhood of unit *j'*. After the condition has been satisfied, at each of the neighboring sites J Bernoulli detection samples are taken and if the condition C is satisfied at these additional sites then their neighborhoods are added as well. This process is continued until a cluster of units is obtained that contains a boundary of sites that do not satisfy the condition C. The data is of the form $y_{ij} = 0,1,2,\ldots J_{ij}$, i=1…n, j=1…k_i where k_i is the size of the cluster which includes all of the adapted sites associated with site i. If a site from the initial simple random sample (size n) did not meet the criteria C then that site is part of a cluster of size 1 (k = 1).

As with traditional occupancy estimation (MacKenzie et al. 2002) we assume that the duration of sampling is sufficiently short enough that each site's occupancy status remains fixed during the time required to complete the survey. We also require that each site belongs to one and only one cluster. We believe this is not a strict requirement as most sampling protocols are developed to occur in a sequential fashion such that it is possible to identify sites that have already been sampled (and thus are already a part of a cluster) and will not be sampled again. In the event that a site could potentially belong to two different clusters we suggest that it be assigned to whichever cluster was sampled first.

Statistical Model

To develop the statistical model it is necessary to reiterate that we assume each site belongs to one and only one cluster and therefore once a site has been sampled (completion of J visits) it cannot be sampled again even if it is adjacent to a different site with $y_{ij}>0$. Therefore we can assume that the clusters are conditionally independent and once the data has been collected we only need to model the spatial process underlying the clustering.

We use a state-space approach in which we express the model by its two component processes: a submodel for the unobserved or partially observed state process (z_{ij}); i=1,2, …, n, j=1,2, …, k_i, and a submodel for the observations conditional on the unobserved state process ($y_{ij}|z_{ij}$).

State model:

$$z_{ij}|\psi_{ij} \sim Bernoulli(\psi_{ij}) \text{ for } i=1,2,\ldots,n, j=1,2,\ldots,k_i, \quad (4.1)$$

$$\text{logit}(\psi_{ij}) = b_{i0} + \beta_1 * x_{ij}, \quad (4.2)$$

$$b_{i0}|\beta_0, \sigma^2 \sim Normal(\beta_0, \sigma^2), \quad (4.3)$$

where z_{ij} is the latent occupancy state at site ij taking the value 1 if the site is occupied and 0 if the site is unoccupied, x_{ij} is an observed site-specific covariate thought to influence occupancy probability, and k_i is the size of the i^{th} cluster. Additional site-specific covariates could be specified to influence the probability of occupancy, but will not be considered further.

Observation model:

$$y_{ij}|J, p, \psi_{ij} \sim Binomial(J, pz_{ij}) \text{ for } i=1,2,\ldots,n, j=1,2,\ldots,k_i, \quad (4.4)$$

where J is the number of replicate observations at each site and p is the probability of detection which is assumed to be constant here, and $\text{logit}(\psi_{ij}) = b_{i0} + \beta_1 * x_{ij}$ as above. Thus, if a site is occupied then the data are Binomial with probability p and J trials and if the site is unoccupied then the data are Binomial with $\Pr(y_{ij}=1) = 0$. Additional variation in the probability of detection could be modeled through the use of replicate-level covariates by substituting logistic regression formulations for p (MacKenzie et al. 2002).

We can express the likelihood as follows:

$$[\boldsymbol{y_{ij}}|p, \beta_0\beta_1, \sigma^2] = \prod_{i=1}^{n}[b_{i0}|\beta_0, \sigma^2]\left\{\prod_{j=1}^{k_i}[y_{ij}|J, p, b_{i0}, \beta_1]\right\}, \tag{4.5}$$

noting that the joint probability of the counts $\boldsymbol{y_i} = (y_{i1}, \dots, y_{ik_i})$ detected within the i^{th} adaptive cluster is

$$[\boldsymbol{y_i}|p, \beta_0\beta_1, \sigma^2] = \int_{-\infty}^{\infty}[b_{i0}|\beta_0, \sigma^2]\prod_{j=1}^{k_i}[y_{ij}|J, p, b_{i0}, \beta_1]\,db_{i0}. \tag{4.6}$$

It is possible to use maximum likelihood to obtain estimates as the integration could be approximated with an adaptive form of Gauss-Hermite quadrature (Pinheiro and Bates 1995; Dorazio and Royle 2005) or with stochastic methods such as Monte Carlo integration (Press et al. 2007) although such an approach can be computationally intensive to implement (we found that in our application Monte Carlo integration required an approximately equal amount of computational time as the suggested Bayesian approach).

Bayesian Analysis

We chose to conduct estimation and inference in a Bayesian framework using conventional methods of Markov chain Monte Carlo (MCMC). The model proposed is a relatively simple random-effects model and can be implemented in the freely available software package OpenBUGS v. 3.2.1 (Lunn et al. 2009). We chose noninformative priors for p (Uniform(0,1)) and σ (Uniform(0,20) following Gelman 2006). For the coefficients (β_0, β_1) we chose to use priors that followed a t-distribution with specified scale, location, and degrees of freedom (location = 0, scale = 1.56, d.f. = 7.76) to ensure they would be approximately uniform on the logit scale. We ran the model using the package R2OpenBUGS in program R v. 2.12 (R Development Core Team 2010) and evaluated convergence by examining trace plots, autocorrelation, and R-hat values (comparison of within-chain and between-chain variances) for each parameter estimate (Gelman et al. 2004).

Simulations

We used simulations to evaluate our new adaptive cluster sampling-occupancy model (ACSOCC) and to compare our new model to traditional adaptive cluster sampling (ACS) and single-season occupancy estimation (SSOCC) under a range of known patterns in occupancy and design criteria. We were interested in exploring the usefulness of our model when patterns in occupancy are spatially correlated through the use of habitat covariates and overall occupancy rate is relatively low. We therefore created three different habitat types with varying degrees of spatial correlation on a 20 x 20 grid (Figure 4.1). The first consisted of three blocks of habitat: low, medium, and high quality with a small amount of random noise added (standard normal deviate) to each block. We considered habitat 1 as an example of extreme spatial correlation. The second type of habitat was generated using a Matérn cluster process (Matérn 1986; Møller and Waagepetersen 2003; Baddeley and Turner 2005). The Matérn cluster process is a doubly-stochastic or two-stage model for point generation and consists of first defining a parent Poisson process with some mean intensity κ. Next, within a radius r of each parent process a second

Poisson process with mean μ is generated. This creates a clustering of points around each of the parents with a fine scale of control to manipulate the amount and degree of clustering. We used this to generate spatially correlated habitat by defining $\kappa = 0.03$, $r = 2.5$, and $\mu = 19$. We considered habitat 2 as an example of moderate spatial correlation. The third habitat type had no spatial correlation and was generated as completely random. The true occupancy rates for each habitat were calculated by specifying a logit-linear model with the simulated habitat covariates:

$$logit(\psi_i) = \beta_0 + \beta_1 * x_i + \varepsilon, \tag{4.7}$$

where x_i is the habitat covariate at each site, $\varepsilon \sim N(0,1)$ and we fixed $\beta_0 = -2$ and $\beta_1 = 2$.

Model Evaluation

We were able to assess the fit of our model using posterior predictive checks by simulating replicated data under the fitted model from the posterior predictive distribution (Gelman et al. 2004). This allowed an assessment and comparison of the “observed data” (data simulated from spatially correlated habitat) to the replicated data (data simulated from the fitted model); ideally there would be very little discrepancy between the two. This approach allows for a check of possible model misfit and any systematic differences between observed data and replicated data suggest a failing in the model structure.

We specifically used an approach suggested by Gelman et al. (1996) referred to as a Bayesian p-value. We defined a discrepancy measure or test quantity $D = \sum_{i,j}(y_{ij} - E[y_{ij}|p, \beta_0, \beta_1, \sigma^2])^2$ for the observations y_{ij} and their expected values under the model. This discrepancy statistic is computed at each iteration of the MCMC algorithm. A reference distribution is computed from the replicated data by simulating data sets from the posterior predictive distribution and a discrepancy measure, D^{sim}, is calculated for the replicated data. The Bayesian p-value is defined as the probability: $\Pr(D > D^{sim})$. Extreme values (e.g., less than 0.05 or greater than 0.95) indicate that the model is inadequate.

Model Comparisons

We compared our model to the single-season occupancy model (SSOCC; MacKenzie et al. 2006), and the traditional Adaptive Cluster Sampling design-based estimators that assume perfect detection (ACS; Thompson and Seber 1996). We calculated the total sample size for the ACSOCC model (includes both primary and secondary sites) and used this as the sample size for the SSOCC model to account for the discrepancy in sample sizes. In addition we compared the ACSOCC model to traditional adaptive cluster sampling (ACS) except that we purposely fixed p (detection probability) to be < 1 under all scenarios to examine the influence of imperfect detection on the ACS design-based estimators. We chose to use the modified Horvitz-Thompson estimator for ACS as several others have acknowledged the improved performance when compared to the modified Hansen-Hurwitz estimator (Thompson and Seber 1996; Christman 2000; Salehi 2003). The modified Horvitz-Thompson estimator for the population mean is:

$$\hat{\mu} = \frac{1}{N}\sum_{k=1}^{\kappa}\frac{y_k^*}{\alpha_k}, \tag{4.8}$$

where y_k^* is the sum of the y-values for the k^{th} network, κ is the number of distinct networks in the sample, N is the total number of sites in the study area (equal to 400 here), and α_k is defined for x_k units in the k^{th} network as: $\alpha_k = 1 - \left[\binom{N-x_k}{n}\Big/\binom{N}{n}\right]$, where n is the sample size for the initial simple random sample. Here a network is defined as a cluster with the edge units removed. Note that for the ACSOCC model we require working with the entire cluster (edge units included) while traditional ACS utilizes the networks for estimation not the entire cluster. Following Thompson and Seber (1996) an estimator for the variance is:

$$\widehat{var}(\hat{\mu}) = \frac{1}{N^2}\left[\sum_{j=1}^{\kappa}\sum_{k=1}^{\kappa}\frac{y_j^* y_k^*}{\alpha_{jk}}\left(\frac{\alpha_{jk}}{\alpha_j \alpha_k} - 1\right)\right], \tag{4.9}$$

where $\alpha_{jk} = 1 - \left[\binom{N-x_j}{n} + \binom{N-x_k}{n} - \binom{N-x_j-x_k}{n}\right]/\binom{N}{n}$. We are particularly interested in comparing our new model (ACSOCC) to the current approaches (SSOCC, ACS) in the context of studies focused on rare species where interest lies in making sound inference about a local population (finite-sample). Therefore we focus on the setting where our geographic extent consists of a finite number of sites and we want to predict the proportion of total sites that are occupied (proportion of area occupied, PAO) or the finite-sample occupancy rate (ψ^{fs}) as opposed to the occurrence probability associated with a theoretically infinite population from which a selection of sites has been sampled (Royle and Dorazio 2008). For the ACSOCC model we can calculate this by simply summing up the latent occupancy states at each site directly (Royle and Kery 2007; Royle and Dorazio 2008). This problem is slightly more complicated for traditional maximum likelihood estimation with the SSOCC model and we refer readers to Royle and Dorazio (2008) for more details. We chose to follow an approach outlined by Pacifici et al. (Chapter 3) to estimate the total number of occupied sites and finite-sample occupancy rate for SSOCC. For both the SSOCC model and the ACSOCC model we assumed that covariate information was completely observable at all sites within the study. This was only done for convenience and is not necessary for implementation of either model. For the ACS design-based estimator we calculated the total number of occupied sites in the population as $\hat{\tau} = N\hat{\mu}$ with $\widehat{var}(\hat{\tau}) = N^2\widehat{var}(\hat{\mu})$ and then used this quantity to calculate the finite-sample occupancy rate.

We explored frequentist properties of all three models (ACSOCC, SSOCC, and ACS) under a range of different design criteria. We varied the overall detection probability p (0.25, 0.5, 0.75), the number of repeat visits to a site J (3, 5), the initial sample size n (20, 50, 100, 150), and the spatial structure of the habitat (1, 2, 3 from above). We calculated 500 synthetic datasets for each combination of design criteria (72 total) and report relative bias, relative mean squared error, and interval coverage for the parameters. We evaluated relative bias (RBIAS) and relative root mean-squared error (RMSE) as

$$\text{RBIAS} = \frac{\frac{1}{l}\sum_{i}^{m}(\hat{\theta}_i - \theta_i)}{\bar{\theta}}$$

and

$$\text{RMSE} = \frac{\sqrt{\frac{1}{l}\sum_{i}^{m}(\hat{\theta}_i - \theta_i)^2}}{\bar{\theta}}$$

where θ_i is the value of the parameter of interest at the i^{th} simulation trial and $\hat{\theta}_i$ is the mean for that parameter. For the ACSOCC model we used the posterior mean from the MCMC samples for the specified parameter. Each synthetic dataset (out of 500) consisted of running 2 MCMC chains each of length 40,000 with a 10,000 burn-in period and thinned by 10 for the ACSOCC model. Coverage for the ACSOCC model consisted of computing the proportion of 95% Bayesian Credible Intervals in 500 simulation trials that included the true parameter in the interval.

Cost Analysis

To examine the potential benefit in terms of total cost for the ACSOCC model compared to the SSOCC model we developed a simple cost analysis. We specified two cost functions, one for SSOCC and one for ACSOCC that included start-up costs, travel costs, and sampling costs. We constrained the total number of samples to be equal for SSOCC and ACSOCC assuming that information collected was a linear function of number of sites visited. This analysis permitted us to investigate the total cost of each design under a range of scenarios. The cost function for SSOCC is:

$$\text{Total cost}_{\text{SSOCC}} = c_0 + c_1 * \bar{x}_1 * n + c_2 * n * J$$

where c_0 = initial cost or startup cost (20; same for SSOCC and ACSOCC), c_1= cost of moving to a new site in study area (10,20,30,or 40), c_2 = cost of collecting sample at each site (2,4,6, or 8), n = number of sample sites (25,30,40, or 50), J = number of repeat visits to each site (5, 10), and $\bar{x}_1$= average # of sites moved between locations (equal to 5.3). This was calculated by simulating data on 10x10 grid and calculating Euclidean distance between sites, where Euclidean distance is

$$d = \sqrt{(x_1 - x_2)^2 + (y_1 - y_2)^2} \text{ for the pair } (x_1, y_1), (x_2, y_2)$$

The cost function for adaptive sampling is:

$$\text{Total cost}_{\text{Adaptive}} = c_0 + c_1 * \bar{x}_1 * n_1 + c_2 * n_1 * J + c_1 * \bar{x}_2 * n_2 + c_2 * n_2 * J$$

where the first part of the equation is identical to SSOCC except only n_1 sites are sampled ($n = n_1 + n_2$), $\bar{x}_2$ = average # of sites moved between locations in adaptive sampling (always equal to 1 because adaptive sites are adjacent), and n_2 = number of sites sampled in adaptive phase.

All code for the data generation, simulations, and analysis are presented in Appendix A while additional results are presented in Appendix C.

Results

The overall true occupancy rate for the simulated habitats ranged from 0.32 – 0.43. The Bayesian p-value averaged over the 500 synthetic datasets ranged from 0.49 – 0.65 for all of the different scenarios suggesting that there was little discrepancy between the replicated data from the fitted model and the observed data according to our measure of discrepancy.

We computed RMSE, RBIAS, and coverage for all parameters, but present here only the parameter of interest ψ^{fs} (finite-sample occupancy rate). Overall both occupancy models (ACSOCC and SSOCC) were unbiased in estimating detection probability when p was high (0.5 or 0.75). There was a small amount of negative bias for both models when p=0.25 and the initial sample size, n, was small (50 or less), but the bias decreased as n increased (Appendix C).

RMSE for ψ^{fs} was generally lower for the ACSOCC model compared to the SSOCC model when initial sample size, n, was low and detection probability, p, was low, but this difference became negligible as p increased and/or n increased (Figures 4.2). This pattern of RMSE occurred across all three types of habitats and levels of spatial correlation. In general there was very little effect of habitat on the RMSE for any of the three approaches. There was little effect on RMSE from increasing the number of repeat visits especially when detection probability was high or the initial sample size was high (Appendix C). RMSE was largest for

ACS in all cases as much of the error was driven by the large amounts of bias induced by imperfect detection (Figures 4.3).

RBIAS for ψ^{fs} was smaller for the SSOCC model compared to the ACSOCC model under almost all design criteria, but the difference between the two models was very small when the probability of detection was 0.5 or 0.75. There was a small amount of positive bias for both the ACSOCC and SSOCC models when estimating ψ^{fs} and detection probability was low, but the bias became negative when detection probability was larger than 0.25 (Figures 4.3). The ACS approach always exhibited substantial negative bias. Both the ACSOCC and SSOCC models exhibited a noticeable drop in RBIAS when increasing the initial sample size in Habitat 1 with low detection probability (Figure 4.3). This pattern is not present in Habitat 2 or in Habitat 3 and is not as pronounced when detection probability is increased or the number of visits is increased. Otherwise there is very little variation in RBIAS for any of the three approaches when moving among the three habitat types.

Coverage for the ACSOCC model was almost always near the nominal 95% level under all of the design criteria even when n and p were low (Table 4.1). There was little observed variation in coverage even as patterns in detection probability, habitat, or sample size fluctuated. Coverage for the SSOCC model was much lower than the ACSOCC model and only under a few occasions did it meet or exceed the nominal level (Table 4.1). We expect coverage for the ACSOCC model to be slightly higher than the SSOCC model because the Bayesian approach can accommodate the propagation of uncertainty as opposed to the SSOCC approach. Interestingly, the variation in estimates of occupancy rate over the 500 simulations was always smaller for the ACSOCC model compared to the SSOCC model (Table 4.1). The SSOCC model exhibited large amounts of variation in the observed estimates especially when detection probability was low and the sample size was low. This suggests that although the SSOCC shows relatively little RBIAS and on average is unbiased the variation is much larger among the estimates compared to the ACSOCC model. The ACSOCC model exhibits smaller variation among its estimates of mean occupancy rate even if there is slightly more bias overall (Table 4.1). Coverage for the ACS approach was very poor and never met or exceeded the nominal level (Table 4.1).

Under all scenarios we found ACSOCC to be more cost efficient than SSOCC. We found the average cost of SSOCC across all possible scenarios to be 6665.46 $\pm$ 3036.35 while the average cost of ACSOCC was 5199.55 $\pm$ 2205.58. The difference between SSOCC and ACSOCC was always positive (average difference = 1465.91 $\pm$ 1220.09, minimum = 43, maximum = 5160) suggesting that ACSOCC was always more economical when compared to SSOCC. Figure 4.4 shows a plot of the total cost for SSOCC and ACSOCC under a range of conditions highlighting the pattern in cost differences as a function of sample sizes. The smallest difference between SSOCC and ACSOCC occurs when total sample size = 25 and in ACSOCC, the initial sample (n_1) had a size of 24 and n_2 = 1 (excluding the obvious case when n_1 = 40 and the cost is identical). The largest difference occurred when n_1 = 20 and n_2 = 30.

Discussion

The conservation and management of rare species is one of the most daunting challenges natural resource managers and ecologists face. It is important that methods are developed that permit accurate estimation and inference for these unique scenarios. We have developed an approach that augments the traditional single-season occupancy design to leverage information from adjacent sites when a known detection has occurred. We have thus provided a relatively simple model that integrates adaptive-cluster sampling into an occupancy estimation framework. Our simulations show a stark improvement in interval coverage for the ACSOCC model compared to SSOCC and traditional ACS most notably when detection probability is low and there is extreme spatial correlation in occupancy.

Several other authors have explored the use of different survey designs for occupancy-based studies. MacKenzie and Royle (2005) explored two common sampling designs and their influence on estimator performance. Double-sampling, where repeat surveys are conducted at a subset of sites only, was found to have little advantage over the traditional approach while removal sampling, where surveying of a site stops once the species is detected or J surveys have been conducted, was found to be more efficient in terms of obtaining a smaller standard error for estimating occupancy. MacKenzie and Royle (2005) went on further to say that this gain in efficiency for removal sampling was only realized when a greater maximum number of visits to a particular site is conducted. This suggests that the use of these designs is not always warranted except under specific circumstances. We found similar results as our model, and thus the use of adaptive-cluster sampling, showed very little improvement in RMSE except when detection probability was extremely low and the initial sample size was low. In addition our model was biased high when detection probability was low, as was SSOCC but to a lesser degree.

Our results suggest that our model may only be useful under certain conditions that relate to specific characteristics of the population. This is not surprising as the benefits of traditional adaptive sampling are only realized for very specific circumstances as well. Several authors have shown that the gain in efficiency for adaptive-cluster sampling depends on many factors including the condition to adapt, the number of sites, and the aggregation and distribution of the population (Smith et al. 1995; Thompson and Seber 1996). Smith et al. (1995) also has shown that for adaptive-cluster sampling to be more efficient than simple-random sampling the final sample size should not be much larger than the initial sample size. In addition Thompson (1990) has shown that the within-network variance should be a high proportion of the total variance in the population. For binary data (only 1's and 0's) the within network variance is 0, since every network in the population consists of either a single unit (0 detections) or a group of one or more units with one detection. It is worth noting, however, that Thompson and Seber (1996) identified a threshold for the initial sample size for which the modified Horvitz-Thompson estimator was more efficient than simple-random sampling for binary data ($n > 50$).

Our simulations also support many of the previous findings from the adaptive sampling literature. For example, in our simulations the final sample size was much greater than the initial sample size and this difference was inflated when detection probability was high. In some cases the final sample size was greater than seven times the initial sample size. This has several implications. First it may seem daunting for the field biologist who initially plans to sample 20 sites and ends up sampling over 150, which could be a logistical nightmare. This is a common problem for adaptive-cluster sampling and has led other authors to develop approaches that provide specific stopping rules or other ways to define a fixed sample size (Christman and Lan 1998; Christman and Lan 2001; Rocco 2003). These variations of adaptive sampling may be useful to consider in such cases. Second, an argument could be made that differentiates between the 150 sites sampled in ACSOCC with the 150 sites sampled in SSOCC because sampling adjacent sites in ACSOCC can be more economical and logistically more feasibly than complete simple-random sampling. This was supported by our simple cost analysis and has been suggested by other authors (Thompson and Seber 1996). Regardless, the drastic difference between initial and final sample size could explain the lack of an advantage for ACSOCC over SSOCC under certain conditions.

Although Thompson and Seber (1996) found little evidence of improved adaptive sampling estimator performance compared to simple-random sampling for binary data this should not impede the use of such a design with occupancy estimation. The findings of Thompson and Seber (1996) suggest that there is no gain in precision of the ACS estimator, but one benefit is the increase in the number of sites sampled thus increasing the likelihood of sampling more individuals. This gain in the likelihood of observing a particular species has been noted by several other authors as well (Salehi and Brown 2010; Thompson 2004) and can play a critical role in some study objectives when finding a species is equally, if not, more important than estimation (e.g., Pacifici et al. Chapter 2). In addition we found such a stark improvement in interval coverage with minimal variation in model performance over the 500 trials for each simulation it suggests there is a clear advantage in performance compared to SSOCC.

Similar to traditional adaptive cluster sampling we envision the logistics of conducting adaptation to be complex and require excellent communication among all parties involved. Of course the exact procedure to allocate effort will ultimately depend on the size of each site, the definition and arrangement of the neighborhood, and the available resources. We therefore envision several different approaches to conducting searches on adjacent sites where we define the sites in the initial simple random sample as primary and adjacent sites that are sampled as a part of adaptation as secondary. We foresee adaptation occurring in one of two possible ways. The adaptation occurs by using the same field biologist(s) that sampled the initial site to sample adjacent secondary sites within the neighborhood of the primary site once the condition has been satisfied. An alternative approach is that adaptation is accommodated by allocating additional effort (other field biologists) to the secondary sites once the condition has been satisfied. Much of this will depend on the order to which sampling will occur. For instance, it would be possible in some scenarios to complete sampling on all of the primary sites before surveys begin on the secondary sites. Alternatively it may be possible to immediately conduct secondary surveys once the condition has been met on a primary site. Again this will depend on the particular study and the available resources, but these issues should be clearly identified and resolved beforehand.

Although we have focused solely on occupancy-type data we believe that our model can be easily extended. Occupancy estimation has seen many variations as needed to accommodate different objectives and constraints for ecological studies. We believe that many of these same approaches could be easily integrated into our model. For example, the use of auxiliary information collected at each site (e.g., counts of individuals Royle and Nichols 2003; Royle 2004) could easily be integrated into our model by focusing specifically on abundance instead of occupancy. This would require a different state model in which abundance was directly modeled instead of occupancy or a model that explicitly relied on the occupancy-abundance relationship (Royle and Nichols 2003; Conroy et al. 2008). There already exists a wide variety of occupancy-based modeling flavors focused on modeling spatial variation in abundance that could be suggested (Dorazio et al. 2005; Royle et al. 2007; Post van der Burg et al. 2011; Webster et al. 2008). As MacKenzie and Royle (2005) found, a removal-based approach may be useful to reduce the logistical effort required to conduct repeat visits while still obtaining reasonable estimates of occupancy. For the ACSOCC model this may be even more advantageous because sampling adjacent sites and conducting repeat visits can be logistically taxing. Thus a removal-type design could still provide the benefits of augmenting the design, but would reduce the overall effort.

We envision other areas of expansion that should be investigated as well. As noted in the adaptive sampling literature, changing the definition of the condition to adapt (trigger) can provide valuable changes in estimation and inference (see overview by Turk and Borkowski 2005). We imagine many possible definitions for the trigger in occupancy-based studies which would ultimately depend on the overall objectives of the study. For example, we can conceive of a situation where the use of auxiliary information (e.g., counts of individuals) could be used as the trigger. A second suggestion is to use a combination of species or an index of multiple species (i.e., measure of diversity) as the trigger for adaptation especially if interest is in community composition or species richness. A separate area of expansion involves the exploration of various neighborhood structures. Christman (1996) found physically contiguous neighborhoods to be most efficient for classical adaptive sampling and this may be relevant for ACSOCC as well.

Although the incorporation of model-based and design-based approaches is not new we believe our approach is unique and potentially useful for a variety of studies interested in patchily distributed, clustered or rare species exhibiting spatial variation. This model builds on both the strength of occupancy modeling and adaptive sampling and performs at least as well, and often better than occupancy modeling alone. In addition it benefits from incorporating observer behavior by allowing for extra effort to be included in areas with known detections while permitting statistically rigorous estimates of occupancy and detection probability. We see the continuation of research focusing on integrating sample design and data collection into the modeling framework as a much needed and critical component to rare species conservation and management. Approaches that allow for the flexibility of combining designs and modeling can provide a critical and informative step in conserving and managing rare species.

Literature Cited

Alldredge, M. W., K. H. Pollock, T. R. Simons, and S. A. Shriner. 2007. Multiple-species analysis of point count data: a more parsimonious modeling framework. Journal of Applied Ecology 44: 281-290.

Baddeley, A. and R. Turner. 2005. Spatstat: an R package for analyzing spatial point patterns. Journal of Statistical Software 12: 1-42.

Berger, J. O., and R. L. Wolpert. 1984. The Likelihood Principle. Institute of Mathematical Statistics, Hayward, California.

Brown, J. A. 2003. Designing an efficient adaptive cluster sample. Environmental and Ecological Statistics 10: 95-105.

Christman, M. C. 1996. Comparison of efficiency of adaptive sampling in some spatially clustered populations. In ASA Proceedings of the Section on Statistics and the Environment, pp. 122-126.

Christman, M. C. 2000. A review of quadrat-based sampling for rare, geographically clustered populations. Journal of Agricultural, Biological, and Environmental Statistics 5: 168-201.

Christman, M. C., and F. Lan. 1998. Sequential adaptive sampling designs to estimate abundance in rare populations. In ASA Proceedings of the Section on Statistics and the Environment, pp. 87-96.

Christman, M. C., and F. Lan. 2001. Inverse adaptive cluster sampling. Biometrics 57: 1096-1105.

Conroy, M. J., J. P. Runge, R. J. Barker, M. R. Schofield, and C. J. Fonnesbeck. 2008. Efficient estimation of abundance for patchily distributed populations via two-phase, adaptive sampling. Ecology 89: 3362-3370.

Cunningham, R. B., and D. B. Lindenmayer. 2005. Modeling count data of rare species: some statistical issues. Ecology 86: 1135-1142.

Dixon, P. M., A. M. Ellison, and N. J. Gotelli. 2005. Improving the precision of estimates of the frequency of rare events. Ecology 86: 1114-1123.

Dorazio, R. M. 1998. Design-based and model-based inference in surveys of freshwater mollusks. Journal of the American Benthological Society 18: 118-131.

Dorazio, R. M., and J. A. Royle. 2005. Estimating size and composition of biological communities by modeling the occurrence of species. Journal of the American Statistical Association 100: 389-398.

Dorazio, R. M., H. L. Jelks, and F. Jordan. 2005. Improving removal-based estimates of abundance by sampling a population of spatially distinct subpopulations. Biometrics 61: 1093-1101.

Edwards Jr., T. C., D. R. Cutler, L. Geiser, J. Alegria, and D. McKenzie. 2004. Assessing rarity of species with low detectability: lichens in Pacific northwest forests. Ecological Applications 14: 414-424.

Edwards Jr., T. C., D. R. Cutler, N. E. Zimmerman, L. Geiser, and J. Alegria. 2005. Model-based stratifications for enhancing the detection of rare ecological events. Ecology 86: 1081-1090.

Gelman, A., X. L. Meng, and H. Stern. 1996. Posterior predictive assessment of model fitness via realized discrepancies. Statistica Sinica 6: 733-807.

Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin. 2004. Bayesian Data Analysis. Chapman and Hall, Boca Raton, Florida.

Gelman, A. 2006. Prior distributions for variance parameters in hierarchical models. Bayesian Analysis 1: 515-533.

Hines, J. E., J. D. Nichols, J. A. Royle, D. I. MacKenzie, A. M. Gopalaswamy, N. Samba Kumar, and K. U. Karanth. 2010. Tigers on trails: occupancy modeling for cluster s sampling. Ecological Applications 20: 1456-1466.

Kish, L. 1995. The hundred years' wars of survey sampling. Statistics in Transition 2: 813-830.

Little, R. J. 2004. To model or not to model? Competing modes of inference for finite population sampling. Journal of the American Statistical Association 99: 546-556.

Lunn, D., D. Spiegelhalter, A. Thomas, and N. Best. 2009. The BUGS project: evolution, critique, and future directions (with discussion). Statistics in Medicine 28: 3049- 3082.

MacKenzie, D. I., J. D. Nichols, G. B. Lachman, S. Droege, J. A. Royle, and C. A. Langtimm. 2002. Estimating site occupancy rates when detection probabilities are less than one. Ecology 83: 2248-2255.

MacKenzie, D. I., and J. A. Royle. 2005. Designing occupancy studies: general advice and allocating survey effort. Journal of Applied Ecology 42:1105-1114.

MacKenzie, D. I., J. D. Nichols, N. Sutton, K. Kawanishi, and L. L. Bailey. 2005. Improving inference in population studies of rare species that are detected imperfectly. Ecology 86: 1101-1113.

MacKenzie, D. I., J. D. Nichols, J. A. Royle, K. H. Pollock, L. L. Bailey, and J. E. Hines. 2006. Occupancy estimation and modeling. Academic Press, New York, New York.

Matérn B. 1986. Spatial Variation. Lecture Notes in Statistics, Vol. 36. Springer-Verlag. New York, New York.

Møller J. and R. P. Waagepetersen. 2003. Statistical Inference and Simulation for Spatial Point Processes. Chapman Hall/CRC, Boca Raton, Florida.

Nichols, J. D., L. L. Bailey, A. F. O'Connell Jr., N. W. Talancy, E. H. Campbell Grant, A. T. Gilbert, E. M. Annand, T. P. Husband, and J. E. Hines. 2008. Multi-scale occupancy estimation and modeling using multiple detection methods. Journal of Applied Ecology 45: 1321-1329.

Pacifici, K., M. J. Conroy, R. J. Cooper, J. T. Peterson, and R. S. Mordecai. Insights from a Large scale ivory-billed woodpecker (Campephilus principalis) search effort with applications to rare wide-ranging avian species. Ph.D. Dissertation, Chpater 2.

Pacifici, K., R. M. Dorazio, and M. J. Conroy. Efficient adaptation: a framework for allocating resources in occupancy studies for rare species. Ph.D. Dissertation, Chapter 3.

Pinheiro, J. C., and D. M. Bates. 1995. Approximations to the log-likelihood function in the nonlinear mixed-effects model. Journal of Computational and Graphical Statistics 4: 12-35.

Possingham, H. P., S. J. Andelman, B. R. Noon, S. Trombulak, and H. R. Pulliam. 2001. Making smart conservation decisions *in* Conservation Biology: Research Priorities for the next decade. Eds. M. E. Soule and G. H. Orians, pp. 225-244. Island Press, Washington, D.C..

Post van der Burg, M., B. Bly, T. VerCauteren, and A. J. Tyre. 2011. Making better sense of monitoring data from low density species using a spatially explicit modeling approach. Journal of Applied Ecology 48: 47-55.

Press, W. H., S A. Teukolsky, W. T. Vetterling, and B. P. Flannery. 2007. Numerical Recipes 3rd edition: the art of scientific computing. Cambridge University Press, New York, New York.

R Development Core Team. 2010. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org.

Rao, J. N. K. 2003. Small Area Estimation. Wiley-Interscience, New York, New York.

Rapley, V. E., and A. H. Welsh. 2008. Model-based inferences from adaptive cluster sampling. Bayesian Analysis 3: 717-736.

Rocco, E. 2003. Constrained inverse adaptive cluster sampling. Journal of Official Statistics 19: 45-57.

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

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., and M. Kery. 2007. A Bayesian state-space formulation of dynamic occupancy models. Ecology 88: 1813-1823.

Royle, J. A., M. Kery, R. Gautier, and H. Schmid. 2007. Hierarchical spatial models of abundance and occurrence from imperfect survey data. Ecological Monographs 77: 465-481.

Royle J. A., and R. M. Dorazio. 2008. Hierarchical modeling and inference in ecology. Academic Press, London.

Salehi, M. M. 2003. Comparison between Hansen-Hurwitz and Horvitz-Thompson estimators for adaptive cluster sampling. Environmental and Ecological Statistics 10: 115-127.

Salehi, M., and J. A. Brown. 2010. Complete allocation sampling: an efficient and easily implemented adaptive sampling design. Population Ecology 52: 451-456.

Särndal, C. E., B. Swensson, J. Wretman. 2003. Model Assisted Survey Sampling. Springer-Verlag, New York, New York.

Smith, T. M. F. 1994. Sample Surveys 1975-1990: An age of reconciliation? International Statistical Review 62: 5-34.

Smith, D. R., M. J. Conroy, and D. H. Brakhage. 1995. Efficiency of adaptive cluster sampling for estimating density of wintering waterfowl. Biometrics 51: 777-788.

Smith, D. R., R. F. Villella, and D. P. Lemarie. 2003. Application of adaptive cluster sampling to low-density populations of freshwater mussels. Environmental and Ecological Statistics 10: 7-15.

Smith, D. R., B. R. Gray, T. J. Newton, and D. Nichols. 2010. Effect of imperfect detectability on adaptive and conventional sampling: simulated sampling of freshwater mussels in the upper Mississippi River. Environmental Monitoring and Assessment 170: 499-507.

Thompson, S. K. 1990. Adaptive cluster sampling. Journal of the American Statistical Association 85: 1050-1059.

Thompson, S. K. 2002. Sampling. Wiley-Interscience, New York, New York.

Thompson, S. K., and G. A. F. Seber. 1996. Adaptive Sampling. Wiley-Interscience, New York, New York.

Thompson, W. L, editor. 2004. Sampling Rare or Elusive Species. Island Press, Washington, D. C.

Turk, P., and J. J. Borkowski. 2005. A review of adaptive cluster sampling: 1990-2003. Environmental and Ecological Statistics 12: 55-94.

Williams B. K., J. D. Nichols, and M. J. Conroy. 2002. Analysis and management of animal populations. Academic Press, San Diego.

Webster, R. A., K. H. Pollock, T. R. Simons. 2008. Bayesian spatial modeling of data from avian point count surveys. Journal of Agricultural, Biological, and Environmental Statistics 13: 121-139.

Yoccoz, N. G., J. D. Nichols, and T. Boulinier. 2001. Monitoring of biological diversity in space and time. Trends in Ecology and Evolution 16: 446-453.

Zipkin, E. F., A. DeWan, and J. A. Royle. 2009. Impacts of forest fragmentation on species richness: a hierarchical approach to community modeling. Journal of Applied Ecology 46: 815-822.

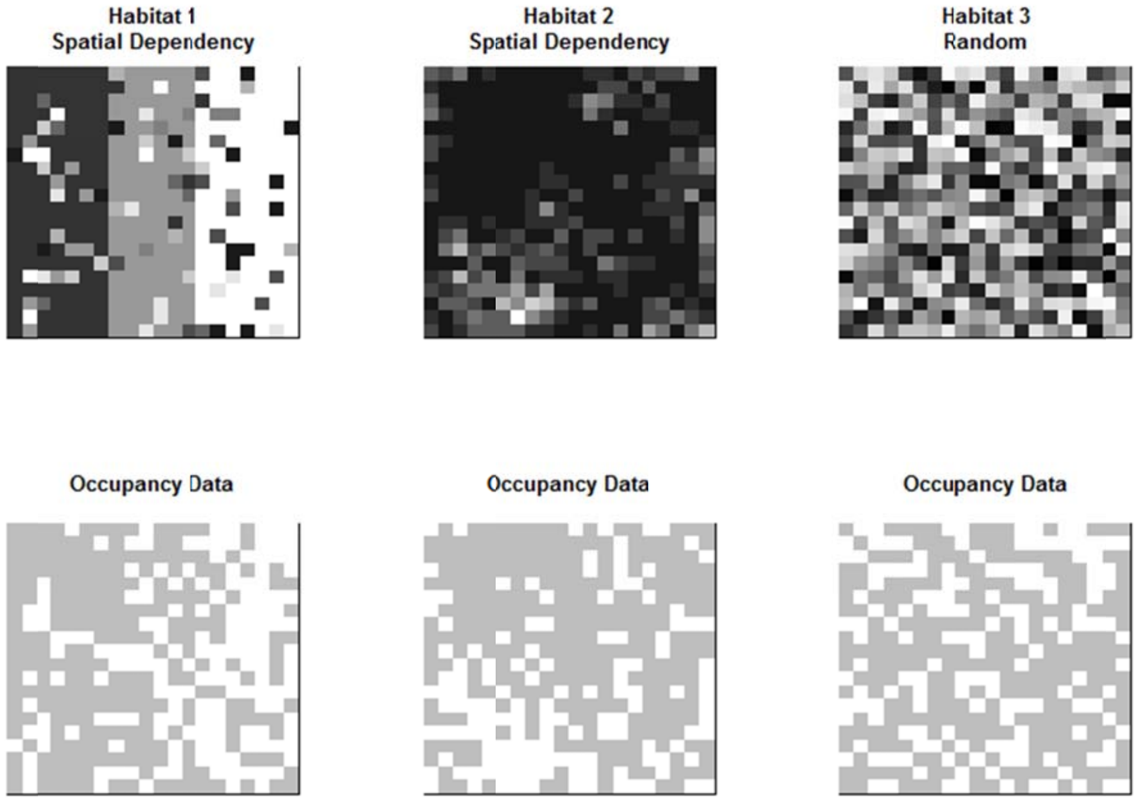

Habitat 1
Spatial Dependency
Habitat 2
Spatial Dependency
Habitat 3
Random
Occupancy Data
Occupancy Data
Occupancy Data

Figure 4.1. Habitat covariates (row 1) and associated occupancy data (row 2) for three different types of simulated environments. The first habitat represents extreme spatial correlation, the second habitat represents moderate spatial correlation and the third habitat was generated randomly and contains no spatial correlation. The true occupancy rates for each habitat (rows 2 and 3) were calculated by specifying a logit-linear model with the simulated habitat covariates (see text for more details). Lighter colors represent higher quality habitat for row 1 and occupied for row 2.

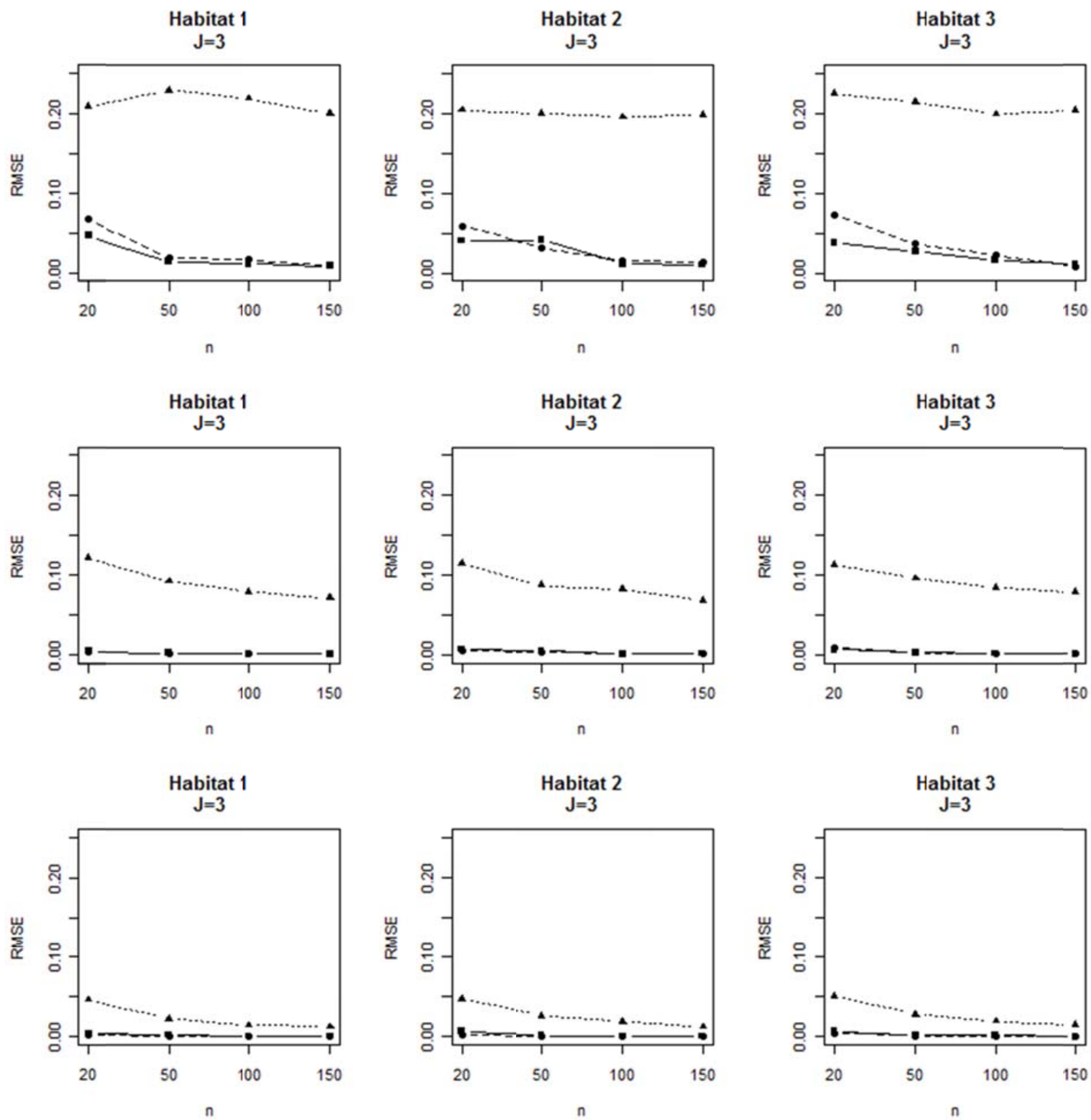

Habitat 1
J=3
Habitat 2
J=3
Habitat 3
J=3
RMSE
0.00
0.10
0.20
20
50
100
150
n

Figure 4.2. Plots of the relative root mean-squared error (RMSE) comparing three different models: ACSOCC (Adaptive-cluster sampling occupancy, solid lines and squares), SSOCC (Single-season occupancy, dashed lines with circles), and ACS (Adaptive-cluster sampling, dotted lines with triangles). The columns differentiate among the three different habitats (1,2, and 3) with varying levels of spatial correlation (high, med, none, respectively). The detection probability is fixed at $p = 0.25$ for the first row, $p = 0.5$ for the second row, and $p = 0.75$ for the third row. All cases have three repeat visits to a site (J). The x-axis represents an increase in the initial sample size, n (20, 50, 100, 150), for each of the three different models.

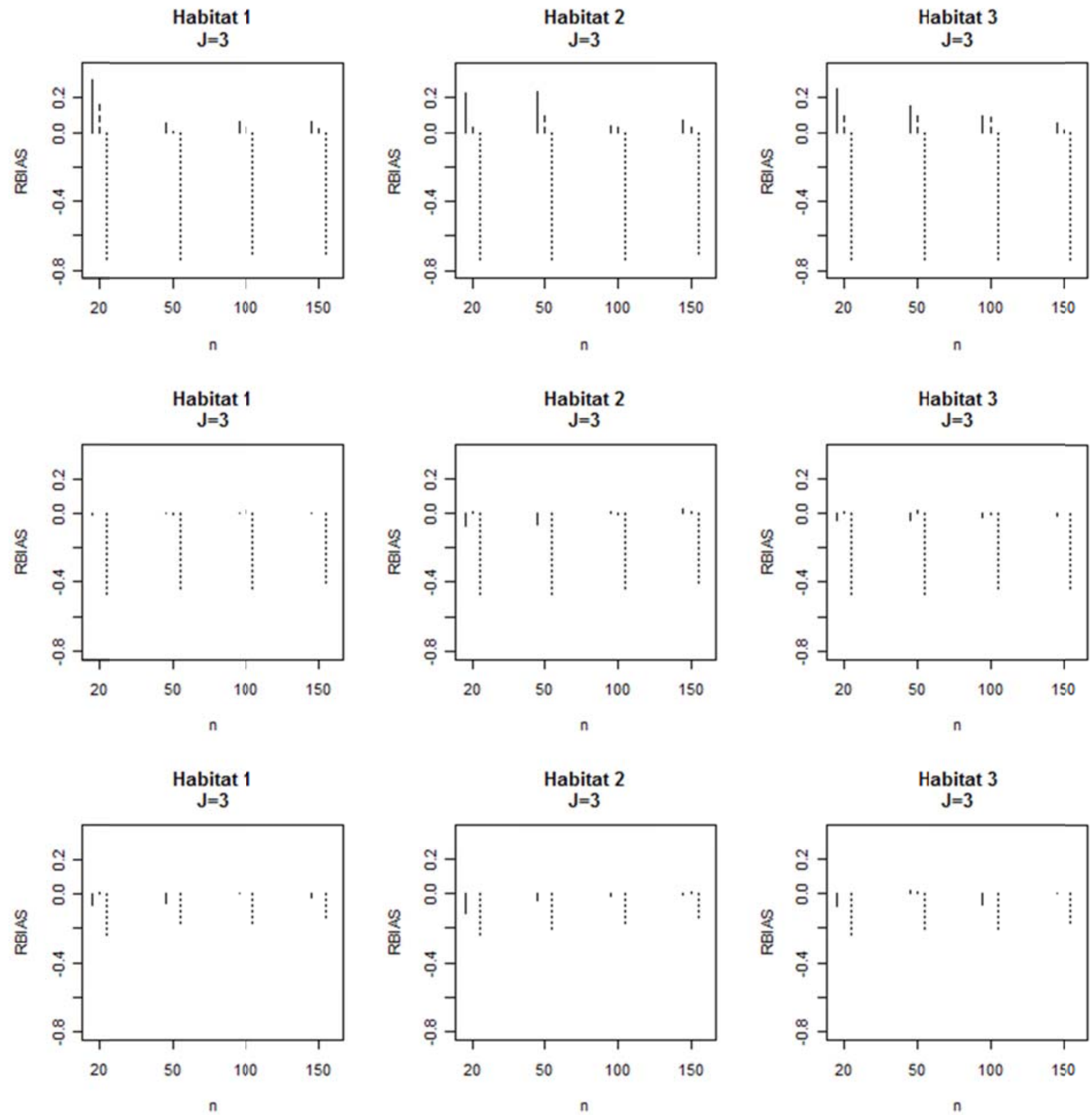

Habitat 1
J=3
Habitat 2
J=3
Habitat 3
J=3
RBIAS
0.2
0.0
-0.4
-0.8
20
50
100
150
n
Habitat 1
J=3
Habitat 2
J=3
Habitat 3
J=3
RBIAS
0.2
0.0
-0.4
-0.8
20
50
100
150
n
Habitat 1
J=3
Habitat 2
J=3
Habitat 3
J=3
RBIAS
0.2
0.0
-0.4
-0.8
20
50
100
150
n

Figure 4.3. Plots of the relative bias (RBIAS) comparing three different models: ACSOCC (Adaptive-cluster sampling occupancy, solid lines), SSOCC (Single-season occupancy, dashed lines), and ACS (Adaptive-cluster sampling, dotted lines). The columns differentiate among the three different habitats (1,2, and 3) with varying levels of spatial correlation (high, med, none, respectively). The detection probability is fixed at $p = 0.25$ for the first row, $p = 0.5$ for the second row, and $p = 0.75$ for the third row. All cases have three repeat visits to a site (J). The x-axis represents an increase in the initial sample size, n (20, 50, 100, 150), for each of the three different models.

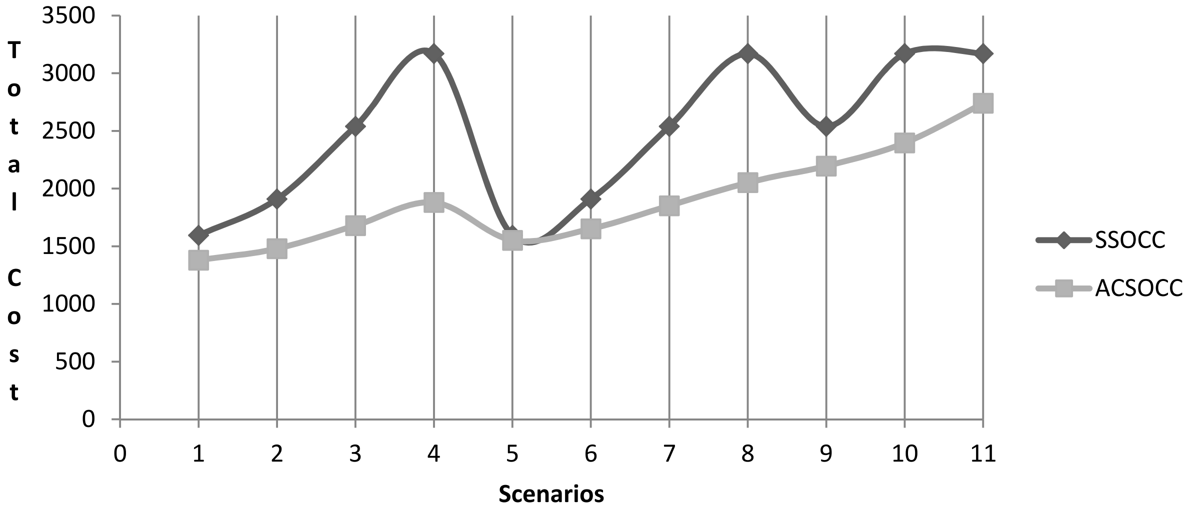

3500
3000
2500
2000
1500
1000
500
0
T o t a l C o s t
0
1
2
3
4
5
6
7
8
9
10
11
Scenarios
SSOCC
ACSOCC

Figure 4.4. Total cost for SSOCC and ACSOCC models under a range of scenarios. Scenarios represent plausible values included in the cost function (see method section for more details). Scenario 1: n_1 = 20, n_2 =5; Scenario 2: n_1 = 20, n_2 =10; Scenario 3: n_1 = 20, n_2 =20; Scenario 4: n_1 = 20, n_2 =30; Scenario 5: n_1 = 24, n_2 =1; Scenario 6: n_1 = 24, n_2 =6; Scenario 7: n_1 = 24, n_2 =16; Scenario 8: n_1 = 24, n_2 =26; Scenario 9: n_1 = 32, n_2 =8; Scenario 10: n_1 = 32, n_2 =18; Scenario 11: n_1 = 40, n_2 =10. In all cases $n = n_1 + n_2$ (SSOCC sample size equals total from ACSOCC).

Table 4.1. Simulation results from 500 synthetic datasets with different design criteria comparing three different models: ACSOCC (Adaptive-cluster sampling occupancy), SSOCC (Single-season occupancy), ACS (Adaptive-cluster sampling). Habitat refers to the amount of generated spatial correlation where Habitat 1 has the most spatial correlation and Habitat 3 has no spatial correlation (see text for more details). Parameters are *p*-detection probability, *n*- initial sample size for adaptive-sampling models (ACSOCC, ACS), and *J*-number of repeat visits is fixed at 3. *Sample Size* is the average realized number of sites sampled for the adaptive-sampling models and the total sample size for the SSOCC model. *True* ψ is the actual finite-sample occupancy rate while $\widehat{\psi^{fs}}$ is the estimated occupancy-rate for each model averaged over the 500 synthetic datasets. *Range* represents the minimum and maximum estimated occupancy rate out of the 500 synthetic datasets for each set of design criteria. *Coverage* represents the percent of confidence intervals that contained the true occupancy rate out of the 500 synthetic data sets. For the ACSOCC model the 95% Bayesian Credible Interval was used to calculate coverage.

	n	Sample Size	True ψ	ACSOCC $\widehat{\psi}^{fs}$	ACSOCC Range		ACSOCC Coverage	SSOCC $\widehat{\psi}^{fs}$	SSOCC Range		SSOCC Coverage	ACS $\widehat{\psi}^{fs}$	ACS Range		ACS Coverage
Habitat 1	20	58	0.36	0.46	0.21	0.63	0.93	0.42	0.09	0.99	0.16	0.09	0.00	0.51	0.09
p=0.25	50	123	0.41	0.43	0.18	0.63	0.99	0.41	0.20	0.78	0.34	0.10	0.00	0.25	0.00
	100	200	0.41	0.43	0.26	0.63	0.97	0.42	0.22	0.73	0.38	0.11	0.03	0.20	0.00
	150	244	0.38	0.40	0.26	0.52	0.97	0.38	0.23	0.59	0.41	0.10	0.04	0.18	0.00
p=0.5	20	151	0.43	0.42	0.27	0.52	0.99	0.43	0.29	0.56	0.65	0.22	0.00	0.52	0.36
	50	207	0.40	0.40	0.26	0.48	0.96	0.40	0.31	0.48	0.73	0.21	0.07	0.44	0.09
	100	255	0.38	0.37	0.30	0.43	0.98	0.38	0.29	0.47	0.69	0.21	0.10	0.34	0.01
	150	291	0.38	0.38	0.32	0.43	0.93	0.38	0.31	0.45	0.77	0.22	0.13	0.32	0.00
p=0.75	20	151	0.37	0.35	0.16	0.40	0.98	0.37	0.29	0.51	0.75	0.28	0.00	0.60	0.83
	50	222	0.43	0.40	0.31	0.45	0.90	0.42	0.36	0.48	0.83	0.35	0.20	0.51	0.67
	100	288	0.39	0.39	0.33	0.42	0.97	0.39	0.34	0.44	0.88	0.33	0.21	0.43	0.48
	150	309	0.39	0.38	0.34	0.40	0.91	0.39	0.34	0.42	0.89	0.32	0.24	0.39	0.26
Habitat 2	20	51	0.35	0.43	0.13	0.63	0.96	0.37	0.06	1.00	0.25	0.09	0.00	0.31	0.09
p=0.25	50	113	0.35	0.44	0.14	0.64	0.89	0.39	0.15	0.73	0.29	0.09	0.00	0.21	0.00
	100	181	0.35	0.37	0.19	0.56	0.96	0.37	0.19	1.00	0.38	0.09	0.02	0.22	0.00

	150	248	0.37	0.40	0.24	0.55	0.94	0.39	0.25	1.00	0.41	0.10	0.03	0.17	0.00
p=0.5	20	104	0.39	0.36	0.18	0.51	0.97	0.40	0.20	0.57	0.59	0.20	0.00	0.52	0.33
	50	147	0.34	0.31	0.19	0.40	0.94	0.34	0.23	0.53	0.67	0.17	0.06	0.35	0.15
	100	262	0.38	0.39	0.30	0.44	0.95	0.38	0.31	0.44	0.81	0.21	0.09	0.35	0.01
	150	294	0.35	0.36	0.27	0.43	0.92	0.35	0.29	0.42	0.78	0.20	0.11	0.29	0.00
p=0.75	20	148	0.37	0.33	0.13	0.40	0.93	0.37	0.28	0.52	0.77	0.28	0.05	0.65	0.80
	50	208	0.36	0.35	0.27	0.39	0.96	0.36	0.29	0.44	0.84	0.28	0.11	0.45	0.68
	100	297	0.41	0.41	0.36	0.44	0.96	0.41	0.36	0.46	0.91	0.34	0.23	0.44	0.42
	150	303	0.35	0.35	0.31	0.37	0.98	0.35	0.32	0.40	0.89	0.29	0.18	0.37	0.30
Habitat 3	20	51	0.38	0.47	0.22	0.63	0.95	0.42	0.11	0.88	0.14	0.09	0.00	0.31	0.10
p=0.25	50	114	0.39	0.45	0.16	0.65	0.95	0.43	0.20	0.83	0.29	0.10	0.00	0.29	0.00
	100	184	0.36	0.40	0.21	0.57	0.93	0.39	0.22	0.77	0.32	0.10	0.02	0.20	0.00
	150	243	0.37	0.39	0.22	0.57	0.95	0.38	0.24	0.62	0.43	0.10	0.03	0.18	0.00
p=0.5	20	90	0.40	0.38	0.20	0.52	0.99	0.40	0.23	0.67	0.51	0.20	0.00	0.51	0.35
	50	190	0.39	0.37	0.25	0.49	0.95	0.40	0.30	0.50	0.70	0.20	0.06	0.38	0.10
	100	251	0.38	0.37	0.30	0.43	0.96	0.38	0.31	0.46	0.80	0.21	0.08	0.32	0.00

	150	289	0.36	0.35	0.29	0.42	0.94	0.36	0.30	0.41	0.81	0.19	0.10	0.28	0.00
p=0.75	20	96	0.37	0.34	0.15	0.42	0.98	0.36	0.25	0.50	0.63	0.28	0.00	0.68	0.79
	50	212	0.39	0.40	0.29	0.46	0.96	0.39	0.33	0.46	0.83	0.31	0.13	0.49	0.65
	100	272	0.38	0.35	0.30	0.39	0.75	0.38	0.34	0.42	0.91	0.30	0.19	0.41	0.46
	150	333	0.41	0.41	0.38	0.44	0.98	0.42	0.38	0.45	0.93	0.34	0.26	0.45	0.22

CHAPTER 5

CONCLUSION

Natural resource managers are faced with the difficult task of conserving and managing rare or elusive species. It is critical that substantial effort is put forth to understand what proximate and ultimate factors are influencing and driving their population dynamics. Regardless of the amount and quality of information available to managers, decisions will be made that can have a substantial impact on rare species. It is therefore paramount that methods and approaches are developed to reduce the uncertainty associated with such decisions. In this dissertation I have focused on developing methods and models that improve our ability to conserve and manage rare species. I have used a case study on Ivory-billed Woodpeckers to provide insight into common problems and shortfalls when working with extremely rare species. The case study has provided motivation to create new methods that provide potential solutions for other studies working with rare species and are robust to a wide variety of scenarios and circumstances.

In chapter 2 I presented a large scale occupancy survey for an extremely rare species and analyzed the associated data collected under this survey. I found that with extreme effort in certain locations, the power of the given survey to detect an individual is very high (>0.95). I estimated that it takes approximately 70 visits to a particular site before the probability of a "true absence" reaches a probability of 0.9 given the estimated detection probability (MLE 0.046). In addition using the maximum likelihood estimates of occupancy from the two river basins, I estimated that between 141-445 sites out of the 595 sites should be occupied depending on the number of visits to a site. All of these results suggest that there is strong evidence against the presence of the IBWO in the two river basins in Florida and South Carolina. Additionally we found that the distribution of effort for the survey created substantial bias in estimating occupancy although there was little bias associated with estimating detection probability. I also found that increasing the number of visits to a site reduced the MSE at a faster rate than increasing the number of sites visited when occupancy and detection are extremely low. MacKenzie and Royle (2005) found it is more beneficial to increase the number of sites visited when detection probability is low, but did not explore cases when occupancy and detection are both extremely low as in our case. Given the current survey design we could find no optimal design that would permit accurate estimates of occupancy and detection.

Chapter 2 highlighted the importance of *a priori* thought devoted to the allocation and distribution of effort. The implementation of the occupancy survey was not done in a standard manner resulting in large heterogeneity in the number of visits to a particular site. This was mainly the cause of observer behavior and observer beliefs about suitable and unsuitable habitat. Two main problems were that observers wanted to put substantially more effort in areas where they observed or thought they observed an individual and observers only wanted to search areas they thought contained suitable habitat as opposed to adhering to a probabilistic sampling framework. These two problems motivated the development of the following two chapters. The heterogeneity in effort and confusion among observers about where to place effort was a serious problem that can potentially affect many other studies and therefore I developed an approach to efficiently allocate effort for occupancy surveys.

In Chapter 3 I developed a framework for allocating effort that provides a probabilistic approach to sampling, allowing for improved accuracy in estimating occupancy probability. The statistical sampling literature already contains many methods for selecting sites in a heterogeneous manner (i.e., stratified sampling, sequential sampling, multi-phase sampling), but these methods do not explicitly allow for the estimation of important quantities in a model-based framework and thus allowing for imperfect detection. The approach I have developed allows for heterogeneity in inclusion probability in addition to reducing the error rate associated with estimating occupancy probability. I found the greatest reduction in predictive error for the new approach when there was a large amount of spatial heterogeneity in habitat and detection probability was low. The optimal approach for the new method was to allocate a relatively small (25%) proportion of sites to the first phase of sampling. Surprisingly, I found improved performance for the new approach even when the habitat had no spatial dependency and was completely random.

Chapter 3 highlighted the need to consider survey design criteria as a part of the decision process. Often there are many options regarding the distribution of effort: how many sites to visit, how many visits to a particular site, which sites to sample first, how to select sites, and these should be utilized as decision alternatives in a decision framework. Using different forms of sample designs and effectively allocating resources can potentially result in the quickest reduction of uncertainty.

In Chapter 4 I addressed the other major issue raised by Chapter 2, namely that observers want to sample more intensively in areas with known detections. As suggested in Chapter 2 I created a model that integrated adaptive-cluster sampling and occupancy estimation, which allowed for additional effort to be placed at adjacent sites after a known detection. I found this model to outperform traditional occupancy modeling under certain conditions. This result was very similar to the advantages obtained when using traditional adaptive cluster sampling compared to simple-random sampling which is that the characteristics of the population often dictate what the appropriate model or estimator is to use.

The use of a hierarchical model allowed for an explicit recognition and differentiation of factors influencing the detection process and the state process. I believe this way of viewing the problem (hierarchical model) provides a generic and robust framework for many additional scenarios and can be applied to a large and diverse set of problems in ecology. I therefore see several important extensions to the new model that would not require restructuring the framework. Two important areas of improvement are the expansion to address multi-species studies. Specifically, the focus on community dynamics and species richness is an important aspect of ecology and thus the model should be augmented to accommodate this. I briefly addressed this in Chapter 4, but believe that this would be a worthy future research need.

The second major area of development that I believe is important to address in the near future, is allowing for the model to accommodate dynamics. I see this development as a critical next step, much like the progression from the original occupancy model (MacKenzie et al. 2002), to a model that incorporated dynamical parameters (MacKenzie et al. 2003). This would be a difficult endeavor and require much thought because of several important limitations. In the original occupancy model parameters were used that relate the state of a site at one time period to the state of a site at the following time period thus allowing for colonization or extinction of a specific site. This is more difficult with adaptive-cluster sampling because not all sites will be sampled uniformly in each time period. For instance there will be sites that were not sampled at time t, but are sampled at time $t+1$ because they were adjacent to an occupied site. Therefore an additional layer of uncertainty exists because new sites are added to the list of sites without any information about their previous state. Although this is akin to problems in capture-recapture (Williams et al. 2002) and could be worked out probabilistically using conditioning, it may be easier to work with the entire cluster as we have in our model. By treating the cluster as the individual unit for which dynamics are influencing, several things become apparent. First, estimation may be more tractable because the size of the unit will be approximately stable (under some conditions) and therefore modification to our existing model could accommodate the advent of dynamical parameters. Second, we can focus on the cluster itself as the unit of interest for which changes in occupancy can be attributed to directly and subsequently changes in the overall population. This would allow for focus to be placed on the individual clusters as themselves smaller populations and therefore there could potentially be interactions among the clusters in a similar fashion as metapopulation dynamics (Hanski 1998) thus allowing for an additional scale of resolution.

Future improvements in conservation and management of rare species will be accomplished through a variety of techniques and approaches. I believe the acknowledgement and ultimately incorporation of spatial dependence and autocorrelation into modeling efforts is critical to our understanding of spatial heterogeneity in the distribution and abundance of rare species. Models should account for this form of dependence whether it is directly by modeling the covariance structure or through the use of more complex survey designs. In addition the use of innovative techniques such as global positioning systems (GPS), non-invasive genetic marking, passive integrated transponder (PIT) tags, camera traps, and photographic identification lend themselves to new approaches of collecting and analyzing data. Ultimately, the most operative approach will be the integration of unique and innovative methods of data collection coupled with models that identify and subsequently estimate the most important vital rates responsible for driving population dynamics.

Literature Cited

Hanski, I. 1998. Metapopulation dynamics. Nature 396: 41-49.

MacKenzie, D. I., J. D. Nichols, G. B. Lachman, S. Droege, J. A. Royle, and C. A. Langtimm. 2002. Estimating site occupancy rates when detection probabilities are less than one. Ecology 83: 2248-2255.

MacKenzie, D. I., J. D. Nichols, J. E. Hines, M. G. Knutson, and A. B. Franklin. 2003. Estimating site occupancy, colonization, and local extinction when a species is detected imperfectly. Ecology 84: 2200-2207.

MacKenzie, D. I., and J. A. Royle. 2005. Designing occupancy studies: general advice and allocating survey effort. Journal of Applied Ecology 42:1105-1114.

Williams B. K., J. D. Nichols, and M. J. Conroy. 2002. Analysis and management of animal populations. Academic Press, San Diego.

APPENDIX A

Please find select R and OpenBUGS code used in the design and analysis (simulations) for Chapters 3 and 4. The first section labeled Chapter 3 contains code specific to chapter 3. The second section labeled Chapter 4 contains code specific to chapter 4.

Chapter 3

```
#Approach #1

#Simple random sample, traditional single-season occupancy analysis with fixed n

#source("functions.r")

#source("data_generation.r")

approach1<-function(n,p,k,d,occ.data,habitat){

n=n

k=k

p.detect=p

d=d

flat=flatten(occ.data,d)

#take SRS of size n
```

```
primary.samples=sample(1:(d*d),n,replace=F)

#get data in encounter history format

observed.data.ssocc=matrix(NA,nrow=length(primary.samples),ncol=k)

for(i in 1:length(primary.samples)){

observed.data.ssocc[i,]=sample.ssocc.history(flat[primary.samples[i],2],flat[primary.samples[i],3
],occ.data,k=k,p=p.detect)

  }

#data vector y

y=observed.data.ssocc

M=length(primary.samples)

#get covariates at sampled sites

sample.cov=matrix(NA,ncol=1,nrow=length(primary.samples))

for(i in 1:length(primary.samples)){

  sample.cov[i]=habitat[flat[primary.samples[i],2],flat[primary.samples[i],3]]

  }

#likelihood function including covariates influencing psi, p still constant

lik.zib.cov<-function(parms,vars){

#calculates likelihood for zero-inflated binomial
```

```
#with constant detection p and occupancy prob psi with one covariate

 tmp<-c(0,0,0)

 names(tmp)<-c("pconstant","psiconstant","psicov")

 tmp[vars]<-parms

 ones<-rep(1,M)

 pmat<-invlgt(tmp[1])

 psi<-invlgt(tmp[2]*ones+tmp[3]*sample.cov)

 loglik<-rep(NA,M)

 for(i in 1:n){

  yvec<-y[i,]

  nd<-sum(yvec)

  pvec<-pmat

  cp<-(pvec^yvec)*((1-pvec)^(1-yvec))

  loglik[i]<-log(prod(cp)*psi[i] + ifelse(nd==0,1,0)*(1-psi[i]))

 }

 sum(-1*loglik)

}
```

```
#calculate maximum likelihood estimates

x=optim(c(0.5,0.5,0.5),lik.zib.cov,method="BFGS",hessian=T)

#get parameter estimates for psi with covariate value = 0 and p

ssocc.estimates=invlgt(x$par)

#ssocc.psi=ssocc.estimates[2]

ssocc.p=ssocc.estimates[1]

#ssocc.psi.se=sqrt(1/x$hessian)[2,2]

#ssocc.p.se=sqrt(1/x$hessian)[1,1]

#predicted occupancy at all sites given observed covariate and estimated coefficients

predicted.sites=matrix(NA,nrow=d,ncol=d)

for(xxx in 1:length(primary.samples)){

 if(sum(y[xxx,1:k]) > 0){

   predicted.sites[flat[primary.samples[xxx],2],flat[primary.samples[xxx],3]]=1}

 else if(sum(y[xxx,1:k]) == 0){

tmp1=invlgt(x$par[2]+x$par[3]*habitat[flat[primary.samples[xxx],2],flat[primary.samples[xxx],
3]])

   tmp2=(1-invlgt(x$par[1]))^k
```

```
predicted.sites[flat[primary.samples[xxx],2],flat[primary.samples[xxx],3]]=(tmp1*tmp2)/((tmp1
*tmp2)+ (1-tmp1))
    }
  }
for(i in 1:d){
  for(j in 1:d){
    if(is.na(predicted.sites[i,j])==T){
      predicted.sites[i,j]=invlgt(x$par[2]+x$par[3]*habitat[i,j])
    }
  }
}
predicted.sites.vector=as.vector(predicted.sites)
#calculate variance (z_i | psi_hat)
predicted.sites.var=matrix(NA,nrow=d,ncol=d)
for(i in 1:d){
  for(j in 1:d){
    predicted.sites.var[i,j]=predicted.sites[i,j]*(1-predicted.sites[i,j])
  }
```

```
}

predicted.sites.var.vector=as.vector(predicted.sites.var)

ntot.hat=sum(predicted.sites.vector)

#          CALCULATE SCALAR COMPARISON STATISTIC T                    #

#.xx is true/model so .00 indicates truth = 0 and model =0

t.stat.11=c()

t.stat.01=c()

t.stat.10=c()

t.stat.00=c()

occ.data.vector=as.vector(occ.data)

temp=cbind(occ.data.vector,predicted.sites.vector,1-predicted.sites.vector)

t.stat.11=sum(temp[temp[,1]==1,2])

t.stat.01=sum(temp[temp[,1]==0,2])

t.stat.10=sum(temp[temp[,1]==1,3])

t.stat.00=sum(temp[temp[,1]==0,3])

t.stat=c(t.stat.11,t.stat.01,t.stat.10,t.stat.00)

return(list(t.stat,predicted.sites.vector,predicted.sites.var.vector,x,ntot.hat))
```

```
}
#Approach #2
#Simple random sample, predict occupancy, sample in proportion to high probability
# of occupancy; adapt by habitat
#source("functions.r")
#source("data_generation.r")
approach2<-function(n1,n2,p,k,d,occ.data,habitat){
n1=n1
n2=n2
p.detect=p
d=d
flat=flatten(occ.data,d)
#take SRS of size n
primary.samples.n1=sample(1:(d*d),n1,replace=F)
#get data in encounter history format
observed.data.ssocc.n1=matrix(NA,nrow=length(primary.samples.n1),ncol=k)
for(i in 1:length(primary.samples.n1)){
```

```
observed.data.ssocc.n1[i,]=sample.ssocc.history(flat[primary.samples.n1[i],2],flat[primary.sampl
es.n1[i],3],occ.data,k=k,p=p.detect)
  }
#data vector y
y.n1=observed.data.ssocc.n1
M.n1=length(primary.samples.n1)
#get covariates at sampled sites
sample.cov.n1=matrix(NA,ncol=1,nrow=length(primary.samples.n1))
for(i in 1:length(primary.samples.n1)){
  sample.cov.n1[i]=habitat[flat[primary.samples.n1[i],2],flat[primary.samples.n1[i],3]]
  }
#likelihood function including covariates influencing psi, p still constant
lik.zib.cov.n1<-function(parms,vars){
#calculates likelihood for zero-inflated binomial
#with constant detection p and occupancy prob psi with one covariate
  tmp<-c(0,0,0)
  names(tmp)<-c("pconstant","psiconstant","psicov")
```

```
  tmp[vars]<-parms
  ones.n1<-rep(1,M.n1)
  pmat.n1<-invlgt(tmp[1])
  psi.n1<-invlgt(tmp[2]*ones.n1+tmp[3]*sample.cov.n1)
  loglik.n1<-rep(NA,M.n1)
  for(i in 1:n1){
    yvec.n1<-y.n1[i,]
    nd.n1<-sum(yvec.n1)
    pvec.n1<-pmat.n1
    cp.n1<-(pvec.n1^yvec.n1)*((1-pvec.n1)^(1-yvec.n1))
    loglik.n1[i]<-log(prod(cp.n1)*psi.n1[i] + ifelse(nd.n1==0,1,0)*(1-psi.n1[i]))
  }
  sum(-1*loglik.n1)
}
#calculate maximum likelihood estimates
x.n1=optim(c(0.5,0.5,0.5),lik.zib.cov.n1,method="BFGS",hessian=T)
#get parameter estimates for psi with covariate value = 0 and p
```

```
ssocc.estimates.n1=invlgt(x.n1$par)

#ssocc.psi.n1=ssocc.estimates.n1[2]

ssocc.p.n1=ssocc.estimates.n1[1]

#ssocc.psi.se.n1=sqrt(1/x.n1$hessian)[2,2]

#ssocc.p.se.n1=sqrt(1/x.n1$hessian)[1,1]

#predicted occupancy at all sites given observed covariate and estimated coefficients

predicted.sites.n1=matrix(NA,nrow=d,ncol=d)

for(xxx in 1:length(primary.samples.n1)){

  if(sum(y.n1[xxx,1:k]) > 0){

    predicted.sites.n1[flat[primary.samples.n1[xxx],2],flat[primary.samples.n1[xxx],3]]=1}

  else if(sum(y.n1[xxx,1:3]) == 0){

tmp1=invlgt(x.n1$par[2]+x.n1$par[3]*habitat[flat[primary.samples.n1[xxx],2],flat[primary.sam
ples.n1[xxx],3]])

    tmp2=(1-invlgt(x.n1$par[1]))^k

predicted.sites.n1[flat[primary.samples.n1[xxx],2],flat[primary.samples.n1[xxx],3]]=(tmp1*tmp
2)/((tmp1*tmp2)+ (1-tmp1))

    }

  }
```

```
for(i in 1:d){
  for(j in 1:d){
    if(is.na(predicted.sites.n1[i,j])==T){
      predicted.sites.n1[i,j]=invlgt(x.n1$par[2]+x.n1$par[3]*habitat[i,j])
    }
  }
}
predicted.sites.vector.n1=as.vector(predicted.sites.n1)
#calculate variance (z_i | psi_hat)
predicted.sites.var.n1=matrix(NA,nrow=d,ncol=d)
for(i in 1:d){
  for(j in 1:d){
    predicted.sites.var.n1[i,j]=predicted.sites.n1[i,j]*(1-predicted.sites.n1[i,j])
  }
}
predicted.sites.var.vector.n1=as.vector(predicted.sites.var.n1)
#weights for sampling at second stage
```

```
wts.n1=c()

temporary=predicted.sites.vector.n1

temporary[primary.samples.n1]<-0

#normalizes the weights without sites included in n1

for(i in 1:length(temporary)){

  wts.n1[i]=temporary[i]/sum(temporary)

}

temp=1:(d*d)

n2.index=sample(temp,size=n2,replace=F,prob=wts.n1)

#data in encounter history format

observed.data.n2=matrix(NA,nrow=length(n2.index),ncol=k)

for (i in 1:length(n2.index)){

observed.data.n2[i,]=sample.ssocc.history(flat[n2.index[i],2],flat[n2.index[i],3],occ.data,k=k,p=p
.detect)

}

#augmented data y2

y.n2=c()

y.n2=rbind(y.n1,observed.data.n2)
```

```
augment.sample=c()

augment.sample=append(augment.sample,primary.samples.n1)

augment.sample=append(augment.sample,n2.index)

M.n2=dim(y.n2)[1]

#get covariates at sampled sites

sample.cov.n2=matrix(NA,ncol=1,nrow=length(n2.index))

for(i in 1:length(n2.index)){

  sample.cov.n2[i]=habitat[flat[n2.index[i],2],flat[n2.index[i],3]]

}

#get total (n1+n2) covariates at sampled sites

sample.cov.tot=append(sample.cov.n1,sample.cov.n2)

lik.zib.cov.n2<-function(parms,vars){

#calculates likelihood for zero-inflated binomial

#with constant detection p and occupancy prob psi with one covariate

  tmp<-c(0,0,0)

  names(tmp)<-c("pconstant","psiconstant","psicov")

  tmp[vars]<-parms
```

```
  ones<-rep(1,M.n2)
  pmat.n2<-invlgt(tmp[1])
  psi.n2<-invlgt(tmp[2]*ones+tmp[3]*sample.cov.tot)
  loglik.n2<-rep(NA,M.n2)
  for(i in 1:M.n2){
    yvec.n2<-y.n2[i,]
    nd.n2<-sum(yvec.n2)
    pvec.n2<-pmat.n2
    cp.n2<-(pvec.n2^yvec.n2)*((1-pvec.n2)^(1-yvec.n2))
    loglik.n2[i]<-log(prod(cp.n2)*psi.n2[i] + ifelse(nd.n2==0,1,0)*(1-psi.n2[i]))
  }
  sum(-1*loglik.n2)
}
#calculate mle's for betas
x.n2=optim(c(0.5,0.5,0.5),lik.zib.cov.n2,method="BFGS",hessian=T)
#get parameter estimates for psi with covariate value = 0 and p
ssocc.estimates.n2=invlgt(x.n2$par)
```

```
#ssocc.psi.n2=ssocc.estimates.n2[2]

ssocc.p.n2=ssocc.estimates.n2[1]

#ssocc.psi.se.n2=sqrt(1/x.n2$hessian)[2,2]

#ssocc.p.se.n2=sqrt(1/x.n2$hessian)[1,1]

#predicted occupancy at all sites given observed covariate and estimated coefficients

predicted.sites.n2=matrix(NA,nrow=d,ncol=d)

for(xxx in 1:length(augment.sample)){

  if(sum(y.n2[xxx,1:k]) > 0){

    predicted.sites.n2[flat[augment.sample[xxx],2],flat[augment.sample[xxx],3]]=1}

  else if(sum(y.n2[xxx,1:k]) == 0){

tmp1=invlgt(x.n2$par[2]+x.n2$par[3]*habitat[flat[augment.sample[xxx],2],flat[augment.sample

                        [xxx],3]])

    tmp2=(1-invlgt(x.n2$par[1]))^k

predicted.sites.n2[flat[augment.sample[xxx],2],flat[augment.sample[xxx],3]]=(tmp1*tmp2)/((tm

                        p1*tmp2)+ (1-tmp1))

    }

  }
```

```
for(i in 1:d){
  for(j in 1:d){
    if(is.na(predicted.sites.n2[i,j])==T){
      predicted.sites.n2[i,j]=invlgt(x.n2$par[2]+x.n2$par[3]*habitat[i,j])
    }
  }
}
predicted.sites.vector.n2=as.vector(predicted.sites.n2)
#calculate variance (z_i | psi_hat)
predicted.sites.var.n2=matrix(NA,nrow=d,ncol=d)
for(i in 1:d){
  for(j in 1:d){
    predicted.sites.var.n2[i,j]=predicted.sites.n2[i,j]*(1-predicted.sites.n2[i,j])
  }
}
predicted.sites.var.vector.n2=as.vector(predicted.sites.var.n2)
ntot.hat=sum(predicted.sites.vector.n2)
```

```
#              CALCULATE SCALAR COMPARISON STATISTIC T                     #

#.xx is true/model so .00 indicates truth = 0 and model =0

t.stat.2.11=c()

t.stat.2.01=c()

t.stat.2.10=c()

t.stat.2.00=c()

occ.data.vector=as.vector(occ.data)

temp=cbind(occ.data.vector,predicted.sites.vector.n2,1-predicted.sites.vector.n2)

t.stat.2.11=sum(temp[temp[,1]==1,2])

t.stat.2.01=sum(temp[temp[,1]==0,2])

t.stat.2.10=sum(temp[temp[,1]==1,3])

t.stat.2.00=sum(temp[temp[,1]==0,3])

t.stat.2=c(t.stat.2.11,t.stat.2.01,t.stat.2.10,t.stat.2.00)

return(list(t.stat.2,predicted.sites.vector.n2,predicted.sites.var.vector.n2,x.n2,ntot.hat))

}
```

Chapter 4

OpenBUGS code:

```
model {

for (i in 1:n.data){

        for(j in 1:k.vector[i]){

                z[i,j]~dbern(psi[i,j])

                logit(psi[i,j])<-b0[i] + beta1*x.data[i,j]

                mu[i,j]<-z[i,j]*p.det

                y.data[i,j]~dbin(mu[i,j],k)

                }

        b0[i]~dnorm(beta0,tau)

        }

for(i in 1:n.data){

        for(j in 1:k.vector[i]){

                y.rep[i,j]~dbin(mu[i,j],k)                #posterior predictive dist

                exp.y[i,j]<-mu[i,j]*k                     #E(y_ij|parms)

                #var.y[i,j]<-(mu[i,j]*(1-mu[i,j]))/k      #var(y_ij|pars)

                tmp.data[i,j]<-(y.data[i,j]-exp.y[i,j])   #squared error loss

                tmp2.data[i,j]<-pow(tmp.data[i,j],2)      #data
```

```
                #tmp.var[i,j]<-pow(var.y[i,j],-1)

                loss.data[i,j]<-tmp2.data[i,j]

                tmp.rep[i,j]<-(y.rep[i,j]-exp.y[i,j])     #replicated data

                tmp2.rep[i,j]<-pow(tmp.rep[i,j],2)

                loss.rep[i,j]<-tmp2.rep[i,j]

                abstmp.data[i,j]<-abs(y.data[i,j]-exp.y[i,j])   #absolute error loss

                abstmp.rep[i,j]<-abs(y.rep[i,j]-exp.y[i,j])      #replicated data

                }

        }

#get sum of test quantity (squared error loss or absolute error loss)

for(i in 1:n.data){

        data.tmp[i]<-sum(loss.data[i,1:k.vector[i]])

        rep.tmp[i]<-sum(loss.rep[i,1:k.vector[i]])

        absdata.tmp[i]<-sum(abstmp.data[i,1:k.vector[i]])

        absrep.tmp[i]<-sum(abstmp.rep[i,1:k.vector[i]])

        }

test.data<-sum(data.tmp[1:n.data])
```

```
test.rep<-sum(rep.tmp[1:n.data])

abs.data<-sum(absdata.tmp[1:n.data])

abs.rep<-sum(absrep.tmp[1:n.data])

test<-step(test.data-test.rep)

test3<-step(abs.data-abs.rep)

#used to get latent states

for(i in 1:n.data){

        zsum[i]<-sum(z[i,1:k.vector[i]])

        }

#sum of latent states for all observed/sampled sites

ztot.obs<-sum(zsum[1:n.data])

#prediction to new locations that were not sampled

for(xx in 1:n.predict){

        b0.new[xx]~dnorm(beta0,tau)

        logit(psi.new[xx])<-b0.new[xx]+beta1*cov.predict[xx]

        z.new[xx]~dbern(psi.new[xx])

        }
```

```
#sum of latent states for unobserved/unsampled locations

ztot.new<-sum(z.new[1:n.predict])

#latent states for all sites observed + unobserved

ztot<- ztot.obs + ztot.new

#overall occupancy prob

psi.tot<- ztot/400

#priors and other definitions

tau<-1/(sigma*sigma)

p.det~dunif(0,1)

t.nu <- 7.763179      # Uniform prior on logit scale

t.sigma <- 1.566267   # Uniform prior on logit scale

tmp<-pow(t.sigma,-2)

beta0 ~ dt(0,tmp,t.nu)

beta1~dt(0,tmp,t.nu)

sigma~dunif(0,20)

}
```

R code:

```
#adaptive-sampling occupancy approach

#source("functions.r")

#source("data_generation.r")

ACSocc<-function(n,p.detect,k,d,occ.data,habitat,n.iter,n.burn,n.thin,n.chain){

##############          ACS-occupancy model          ###########

n=n

k=k

p=p.detect

d=d

n.iter=n.iter

n.burn=n.burn

n.thin=n.thin

n.chain=n.chain

habitat=habitat

occ.data=occ.data

ntot=sum(occ.data)

psitot=ntot/(d*d)
```

```
#take SRS of size n

primary.samples=sample(1:(d*d),n,replace=F)

#flatten out matrix to keep indices stored

flat=flatten(occ.data,d)

#flatten out matrix of covariates

flat.cov=flatten(habitat,d)

#maximum number of adapted sites

max.adapt.sites=200

#create matrices to store data

primary.data=matrix(NA,nrow=n,ncol=max.adapt.sites)

#matrix to store covariate data

covariate.data=matrix(NA,nrow=n,ncol=max.adapt.sites)

#matrix to store locations "1t1" is row 1, col 1

loc.data=matrix(NA,nrow=n,ncol=max.adapt.sites)

#create empty vector to keep track of sites that have been sampled

storage=c()

#loop through primary sites, check if it has already been sampled, get adjacent
```

```
#sites and proceed with sampling those sites.  Condition for adaptation is single detection
for(xx in 1:n){
  #create empy queue to hold list of sites for each primary site
  queue=c()
  #adding location of primary site to list
  primary.paste=paste(flat[primary.samples[xx],2],"t",flat[primary.samples[xx],3],sep="")
  #making sure primary site has not already been sampled
  if(primary.paste %in% storage==F){
  storage=append(storage,primary.paste)
  #getting data (both detections and covariates) at primary site location
  primary.data[xx,1]=rbinom(1,k,flat[primary.samples[xx],1]*p)
  covariate.data[xx,1]=flat.cov[primary.samples[xx],1]
  loc.data[xx,1]=paste(flat[primary.samples[xx],2],"t",flat[primary.samples[xx],3],sep="")
  #if get a detection do adaptive sampling to rooks neighbors
  if(primary.data[xx,1]>0){
    temp=adjacency.rooks(flat[primary.samples,2][xx],flat[primary.samples,3][xx],occ.data)
    queue=rbind(queue,temp)
```

```
temp1=queue[,1:2]

unique.queue=unique(temp1)

unique.queue.paste=paste(unique.queue[,1],"t",unique.queue[,2],sep="")

queue.1=unique.queue[unique.queue.paste %in% storage==F,]

#manually iterating through the process of checking neighbors, getting detections...

second=adapt2(queue.1,occ.data,storage,k,p,flat.cov)

third=adapt2(second$queue.2,occ.data,second$storage,k,p,flat.cov)

fourth=adapt2(third$queue.2,occ.data,third$storage,k,p,flat.cov)

fifth=adapt2(fourth$queue.2,occ.data,fourth$storage,k,p,flat.cov)

sixth=adapt2(fifth$queue.2,occ.data,fifth$storage,k,p,flat.cov)

seventh=adapt2(sixth$queue.2,occ.data,sixth$storage,k,p,flat.cov)

eighth=adapt2(seventh$queue.2,occ.data,seventh$storage,k,p,flat.cov)

ninth=adapt2(eighth$queue.2,occ.data,eighth$storage,k,p,flat.cov)

tenth=adapt2(ninth$queue.2,occ.data,ninth$storage,k,p,flat.cov)

eleven=adapt2(tenth$queue.2,occ.data,tenth$storage,k,p,flat.cov)

twelve=adapt2(eleven$queue.2,occ.data,eleven$storage,k,p,flat.cov)

thirteen=adapt2(twelve$queue.2,occ.data,twelve$storage,k,p,flat.cov)
```

```
fourteen=adapt2(thirteen$queue.2,occ.data,thirteen$storage,k,p,flat.cov)

fifteen=adapt2(fourteen$queue.2,occ.data,fourteen$storage,k,p,flat.cov)

sixteen=adapt2(fifteen$queue.2,occ.data,fifteen$storage,k,p,flat.cov)

seventeen=adapt2(sixteen$queue.2,occ.data,sixteen$storage,k,p,flat.cov)

eighteen=adapt2(seventeen$queue.2,occ.data,seventeen$storage,k,p,flat.cov)

nineteen=adapt2(eighteen$queue.2,occ.data,eighteen$storage,k,p,flat.cov)

twenty=adapt2(nineteen$queue.2,occ.data,nineteen$storage,k,p,flat.cov)

twenty1=adapt2(twenty$queue.2,occ.data,twenty$storage,k,p,flat.cov)

twenty2=adapt2(twenty1$queue.2,occ.data,twenty1$storage,k,p,flat.cov)

twenty3=adapt2(twenty2$queue.2,occ.data,twenty2$storage,k,p,flat.cov)

twenty4=adapt2(twenty3$queue.2,occ.data,twenty3$storage,k,p,flat.cov)

twenty5=adapt2(twenty4$queue.2,occ.data,twenty4$storage,k,p,flat.cov)

twenty6=adapt2(twenty5$queue.2,occ.data,twenty5$storage,k,p,flat.cov)

twenty7=adapt2(twenty6$queue.2,occ.data,twenty6$storage,k,p,flat.cov)

twenty8=adapt2(twenty7$queue.2,occ.data,twenty7$storage,k,p,flat.cov)

twenty9=adapt2(twenty8$queue.2,occ.data,twenty8$storage,k,p,flat.cov)

thirty=adapt2(twenty9$queue.2,occ.data,twenty9$storage,k,p,flat.cov)
```

```
thirty1=adapt2(thirty$queue.2,occ.data,thirty$storage,k,p,flat.cov)

thirty2=adapt2(thirty1$queue.2,occ.data,thirty1$storage,k,p,flat.cov)

thirty3=adapt2(thirty2$queue.2,occ.data,thirty2$storage,k,p,flat.cov)

thirty4=adapt2(thirty3$queue.2,occ.data,thirty3$storage,k,p,flat.cov)

thirty5=adapt2(thirty4$queue.2,occ.data,thirty4$storage,k,p,flat.cov)

thirty6=adapt2(thirty5$queue.2,occ.data,thirty5$storage,k,p,flat.cov)

thirty7=adapt2(thirty6$queue.2,occ.data,thirty6$storage,k,p,flat.cov)

thirty8=adapt2(thirty7$queue.2,occ.data,thirty7$storage,k,p,flat.cov)

thirty9=adapt2(thirty8$queue.2,occ.data,thirty8$storage,k,p,flat.cov)

forty=adapt2(thirty9$queue.2,occ.data,thirty9$storage,k,p,flat.cov)

forty1=adapt2(forty$queue.2,occ.data,forty$storage,k,p,flat.cov)

#adding everything to storage and getting unique sites that need to be sampled

storage=append(storage,forty1$storage)

storage=append(storage,forty1$queue.2)

storage=unique(storage)

#getting all of the data for the adaptively added sites associated with primary site

y=rbind(second$secondary.data,third$secondary.data,fourth$secondary.data,
```

```
  fifth$secondary.data,sixth$secondary.data,seventh$secondary.data,eighth$secondary.data,
    ninth$secondary.data,tenth$secondary.data,eleven$secondary.data,twelve$secondary.data,
 thirteen$secondary.data,fourteen$secondary.data,fifteen$secondary.data,sixteen$secondary.data,
seventeen$secondary.data,eighteen$secondary.data,nineteen$secondary.data,twenty$secondary.d
ata,
twenty1$secondary.data,twenty2$secondary.data,twenty3$secondary.data,twenty4$secondary.da
ta,
twenty5$secondary.data,twenty6$secondary.data,twenty7$secondary.data,twenty8$secondary.da
ta,
twenty9$secondary.data,thirty$secondary.data,thirty1$secondary.data,thirty2$secondary.data,
thirty3$secondary.data,thirty4$secondary.data,thirty5$secondary.data,thirty6$secondary.data,
thirty7$secondary.data,thirty8$secondary.data,thirty9$secondary.data,forty$secondary.data,
          forty1$secondary.data)
  #getting locations for all of the adaptively added sites associated with primary site
  locs=rbind(second$location,third$location,fourth$location,
  fifth$location,sixth$location,seventh$location,eighth$location,
    ninth$location,tenth$location,eleven$location,twelve$location,
      thirteen$location,fourteen$location,fifteen$location,sixteen$location,
```

```
seventeen$location,eighteen$location,nineteen$location,twenty$location,
twenty1$location,twenty2$location,twenty3$location,twenty4$location,
twenty5$location,twenty6$location,twenty7$location,twenty8$location,
twenty9$location,thirty$location,thirty1$location,thirty2$location,
thirty3$location,thirty4$location,thirty5$location,thirty6$location,
thirty7$location,thirty8$location,thirty9$location,forty$location,
forty1$location)
#getting data and putting it in the associated row for the primary site
if(is.null(y)==F){
primary.data[xx,2:(dim(y)[1]+1)]=y[,1]
covariate.data[xx,2:(dim(y)[1]+1)]=y[,2]
loc.data[xx,2:(dim(locs)[1]+1)]=locs[,1] }
}
}
}
#get rid of NAs in primary.data[,1] that were a result of the primary sample already
#sampled during one of the other clusters.  Need to augment the size of primary data and n
```

```
y.data=primary.data[!(is.na(primary.data[,1])),]

location.data=loc.data[!(is.na(loc.data[,1])),]

x.data=covariate.data[!(is.na(covariate.data[,1])),]

#k.vector gives list of size of each cluster

k.vector=c()

for(i in 1:dim(y.data)[1]){

  temp=0

  for(j in 1:dim(y.data)[2]){

    if(is.na(y.data[i,j])==F){

      temp=temp+1 }

    }

  k.vector=append(k.vector,temp)

}

#get number of primary locations

n.data=dim(y.data)[1]

#get true sample size (this includes all of the adaptively added sites)

sample.size=sum(k.vector)
```

```
#create matrix that identifies which sites have been sampled
zeros=matrix(0,ncol=ncol(occ.data),nrow=nrow(occ.data))
for(i in 1:dim(location.data)[1]){
  for(j in 1:dim(location.data)[2]){
    if(is.na(location.data[i,j])==F){
      temp=strsplit(location.data[i,j],"t")
      temp1=as.numeric(temp[[1]][1])
      temp2=as.numeric(temp[[1]][2])
      zeros[temp1,temp2]=1
    }
  }
}
#need vector of sites that have not been sampled
zeros.flat=flatten(zeros,d)
n.predict=(d*d)-sample.size
#need vector of covariate values for those sites that have not been sampled
cov.predict=c()
```

```
cov.predict.locs=c()

for(i in 1:(d*d)){

  if(zeros.flat[i,1]==0){

    cov.predict=append(cov.predict,flat.cov[i,1])

    cov.predict.locs=append(cov.predict.locs,paste(zeros.flat[i,2],"t",zeros.flat[i,3],sep=""))

  }

  else i=i+1

}

#Bayes inference

library(R2OpenBUGS)

data=list("n.data","y.data","k.vector","k","x.data","n.predict","cov.predict")

parameters<-c("p.det","ztot","psi.tot","test","test3")

par.inits<-function(){

  list(p.det=runif(1),beta1=rnorm(1),beta0=rnorm(1),sigma=runif(1),b0=rnorm(n.data))}

adaptocc.bayes<-bugs(data,inits=par.inits,parameters,"ACSocc.txt",
n.thin=n.thin,n.chains=n.chain,
n.burnin=n.burn,n.iter=n.iter,debug=F,codaPkg=F,DIC=T,OpenBUGS.pgm="/usr/local/OpenBU
GS-3.1.2/bin/OpenBUGS")
```

```
attach.bugs(adaptocc.bayes)

ztot.median=median(ztot)

ztot.mean=mean(ztot)

psi.tot.median=median(psi.tot)

psi.tot.mean=mean(psi.tot)

test.mean=mean(test)

test3.mean=mean(test3)

p.median=median(p.det)

p.mean=mean(p.det)

quant.ztot=quantile(ztot,probs=c(0.025,0.975))

low.ztot=quant.ztot[1]

high.ztot=quant.ztot[2]

coverage=0

if(ntot<quant.ztot[2] & ntot>quant.ztot[1]){

  coverage=1}

bias.p=mean(p.det)-p

mse.p=(mean(p.det)-p)^2
```

```
bias.psi=mean(psi.tot)- psitot

mse.psi=(mean(psi.tot) - psitot)^2

bias.z=mean(ztot)-ntot

mse.z=(mean(ztot)-ntot)^2

detach.bugs()

return(list(sample.size=sample.size,ztot.median=ztot.median,ztot.mean=ztot.mean,

low.ztot=low.ztot,high.ztot=high.ztot,psi.tot.median=psi.tot.median,

psi.tot.mean=psi.tot.mean,test.mean=test.mean,test3.mean=test3.mean,p.median=p.median,

p.mean=p.mean,coverage=coverage,bias.p=bias.p,mse.p=mse.p,bias.psi=bias.psi,

mse.psi=mse.psi,bias.z=bias.z,mse.z=mse.z,true.z=ntot,true.psi=psitot))

}

#traditional adaptive cluster sampling Thompson 1990

##      HORVITZ-THOMPSON ESTIMATOR

#get unique network.data (each network only listed once)

network.data.unique=unique(network.data)

#calculate estimate of mean and total

big.N=(d*d)
```

```
alpha.k=c()
y.k=c()
for(i in 1:dim(network.data.unique)[1]){
  alpha.k[i]=1-((choose(big.N-network.data.unique[i,2],n))/(choose(big.N,n)))
  y.k[i]=network.data.unique[i,3]
}
mu.hat.ht=(1/big.N)*sum(y.k/alpha.k)
ntot.hat.ht=sum(y.k/alpha.k)
#variance estimation
ht.denom=choose(big.N,n)
vhat.ht=0
alphakh=0
tnum=0
for(j in 1:dim(network.data.unique)[1]){
  for(k in 1:dim(network.data.unique)[1]){
    if(j==k){
      alphakh=alpha.k[j]
```

```
    }
    else {
      alphakh=1-((choose((big.N-network.data.unique[j,2]),n)+choose((big.N-
network.data.unique[k,2]),n)-
        choose((big.N-network.data.unique[j,2]-network.data.unique[k,2]),n))/ht.denom)
    }
    tnum=network.data.unique[j,3]*network.data.unique[k,3]
    tnum=tnum*(alphakh-alpha.k[j]*alpha.k[k])/(alphakh*alpha.k[j]*alpha.k[k])
    vhat.ht=vhat.ht+tnum
  }
}
vhat.ht=vhat.ht/(big.N^2)
var.mu.hat.ht=vhat.ht
var.ntot.hat.ht=(big.N^2)*var.mu.hat.ht
se.ntot.hat.ht=sqrt(var.ntot.hat.ht)
#check to see if variance is negative
var.ht.0=0
if(var.mu.hat.ht<0){
```

```
  var.ht.0=1}

#95% Asymptotic confidence interval

ci.95.ht=1.96*se.ntot.hat.ht

lower.bound.ht=ntot.hat.ht - ci.95.ht

upper.bound.ht=ntot.hat.ht + ci.95.ht

####    HANSEN-HURWITZ ESTIMATOR FOR MEAN, TOTAL AND VARIANCE

wi=network.data[,3]/network.data[,2]

mu.hat.hh=mean(wi)

ntot.hat.hh=mu.hat.hh*big.N

piece.1.hh=((big.N-n)/(big.N*n*(n-1)))

temp=c()

for(i in 1:n){

  temp[i]=(wi[i]-mu.hat.hh)^2

}

var.mu.hat.hh=piece.1.hh*sum(temp)

var.ntot.hat.hh=var.mu.hat.hh*big.N^2

se.ntot.hat.hh=sqrt(var.ntot.hat.hh)
```

```
#check to see if variance is negative
var.hh.0=0
if(var.mu.hat.hh<0){
  var.hh.0=1}
ci.95.hh=1.96*se.ntot.hat.hh
lower.bound.hh=ntot.hat.hh-ci.95.hh
upper.bound.hh=ntot.hat.hh+ci.95.hh
#create matrix that identifies which sites have been sampled
zeros=matrix(0,ncol=ncol(occ.data),nrow=nrow(occ.data))
for(i in 1:dim(loc.data)[1]){
  for(j in 1:dim(loc.data)[2]){
    if(is.na(loc.data[i,j])==F){
      temp=strsplit(loc.data[i,j],"t")
      temp1=as.numeric(temp[[1]][1])
      temp2=as.numeric(temp[[1]][2])
      zeros[temp1,temp2]=1
    }
```

```
  }
}
#need vector of sites that have not been sampled
zeros.flat=flatten(zeros,d)
#get true sample size (this includes all of the adaptively added sites)
sample.size=sum(zeros)
coverage.ht=0
if(ntot<upper.bound.ht & ntot > lower.bound.ht){
  coverage.ht=1}
coverage.hh=0
if(ntot<upper.bound.hh & ntot > lower.bound.hh){
  coverage.hh=1}
psi.hat.ht=ntot.hat.ht/(d*d)
psi.hat.hh=ntot.hat.hh/(d*d)
bias.z.ht=ntot-ntot.hat.ht
bias.z.hh=ntot-ntot.hat.hh
mse.z.ht=(ntot-ntot.hat.ht)^2
```

```
mse.z.hh=(ntot-ntot.hat.hh)^2

bias.psi.ht=psi.hat.ht - psitot

mse.psi.ht=(psi.hat.ht - psitot)^2

bias.psi.hh=psi.hat.hh - psitot

mse.psi.hh=(psi.hat.hh - psitot)^2

return(list(sample.size=sample.size,ntot.hat.ht=ntot.hat.ht,var.ntot.hat.ht=var.ntot.hat.ht,

ntot.hat.hh=ntot.hat.hh,var.ntot.hat.hh=var.ntot.hat.hh,bias.z.ht=bias.z.ht,

bias.z.hh=bias.z.hh,mse.z.ht=mse.z.ht,mse.z.hh=mse.z.hh,bias.psi.ht=bias.psi.ht,

mse.psi.ht=mse.psi.ht,bias.psi.hh=bias.psi.hh,mse.psi.hh=mse.psi.hh,

coverage.ht=coverage.ht,coverage.hh=coverage.hh,var.ht.0=var.ht.0,

var.hh.0=var.hh.0,true.z=ntot))

}

#Monte Carlo Integration of ACSOCC likelihood

M=1000

lik.zib.cov.rand.effects<-function(parms){

#calculates likelihood for zero-inflated binomial with random effects

#with constant detection p and occupancy prob psi with one covariate and random effects
```

```
pconstant=invlgt(parms[1])

beta0=parms[2]

sigma=exp(parms[3])

beta1=parms[4]

negl1=0 #counter for likelihood

for(i in 1:n){

 temp2=matrix(nrow=M,ncol=k[i])

 b0=rnorm(M,beta0,sigma)

 for(j in 1:k[i]){

  xval=x[i,j]

  psi=invlgt(b0)

  temp=psi*dbinom(y[i,j],J,pconstant)+ifelse(y[i,j]==0,1,0)*(1-psi)

  temp2[,j]=temp

  }

 temp3=apply(temp2,1,prod)

 integral1=mean(temp3)

 negl1=negl1-log(integral1)
```

```
   }
return(negl1)
} #close function for lik.zib.cov.rand.effects
xx=optim(c(1,0,0),lik.zib.cov.rand.effects,hessian=F)
yy=optim(c(xx$par[1],xx$par[2],xx$par[3]),lik.zib.cov.rand.effects,method="BFGS",hessian=F)
```

APPENDIX B

Please find additional figures and tables of results from the simulations conducted in Chapter 3.

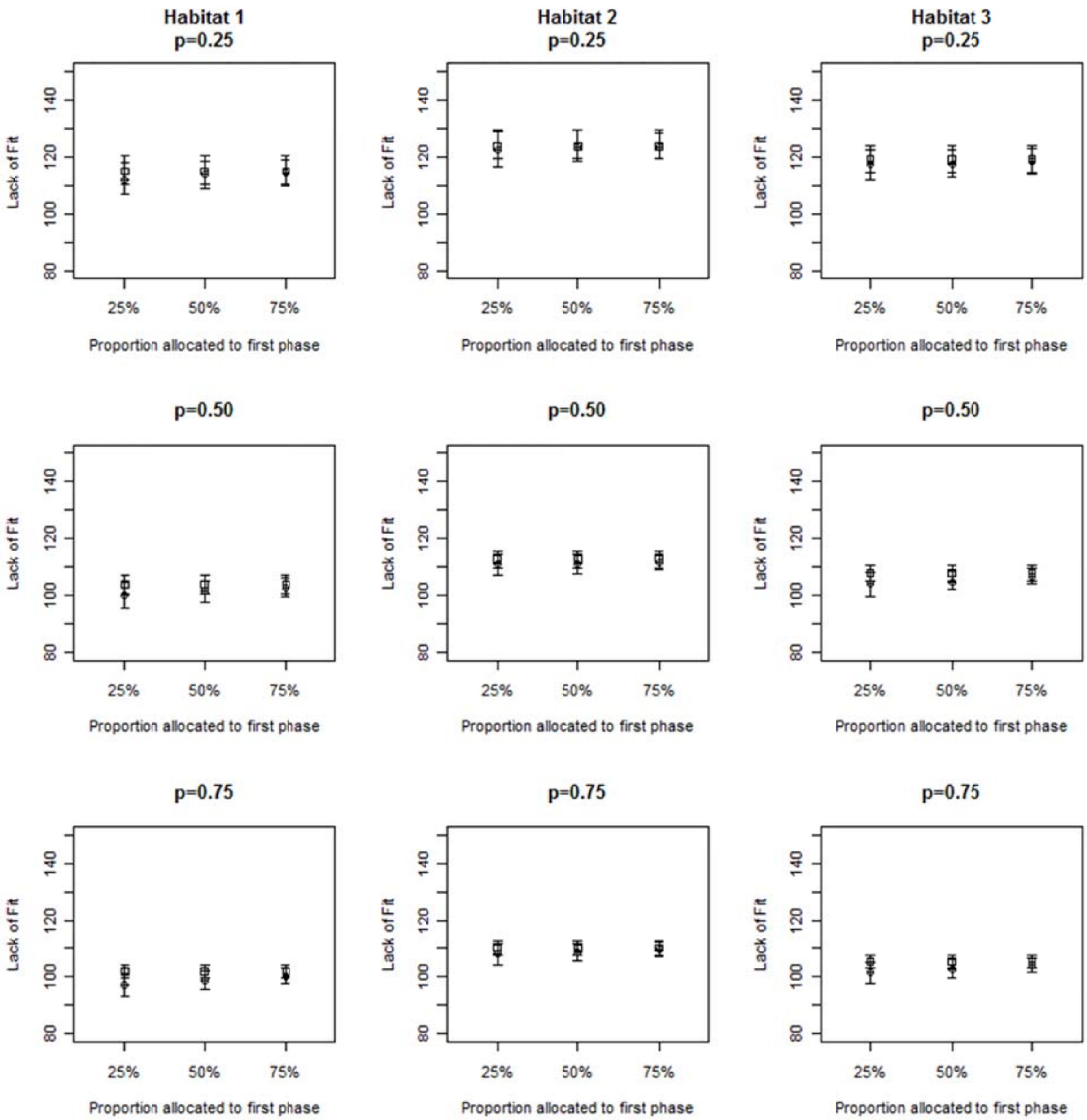

Habitat 1
p=0.25
Habitat 2
p=0.25
Habitat 3
p=0.25
p=0.50
p=0.50
p=0.50
p=0.75
p=0.75
p=0.75
Lack of Fit
80
100
120
140
25%
50%
75%
Proportion allocated to first phase

Figure B.1. Plots of test statistic and associated quartiles (25% and 75%) measuring lack of fit (lower is better) for two-phase adaptive approach (circles) and traditional single-season occupancy approach (squares) when n=100, J=5, and for three different levels of detection, p=0.25, 0.5, 0.75. X-axis represents proportion of sample allocated to phase one for the two-phase adaptive approach. Habitat 1 is a simulated habitat with extreme spatial correlation while habitat 2 has moderate spatial correlation and habitat 3 is randomly generated and contains no spatial correlation.

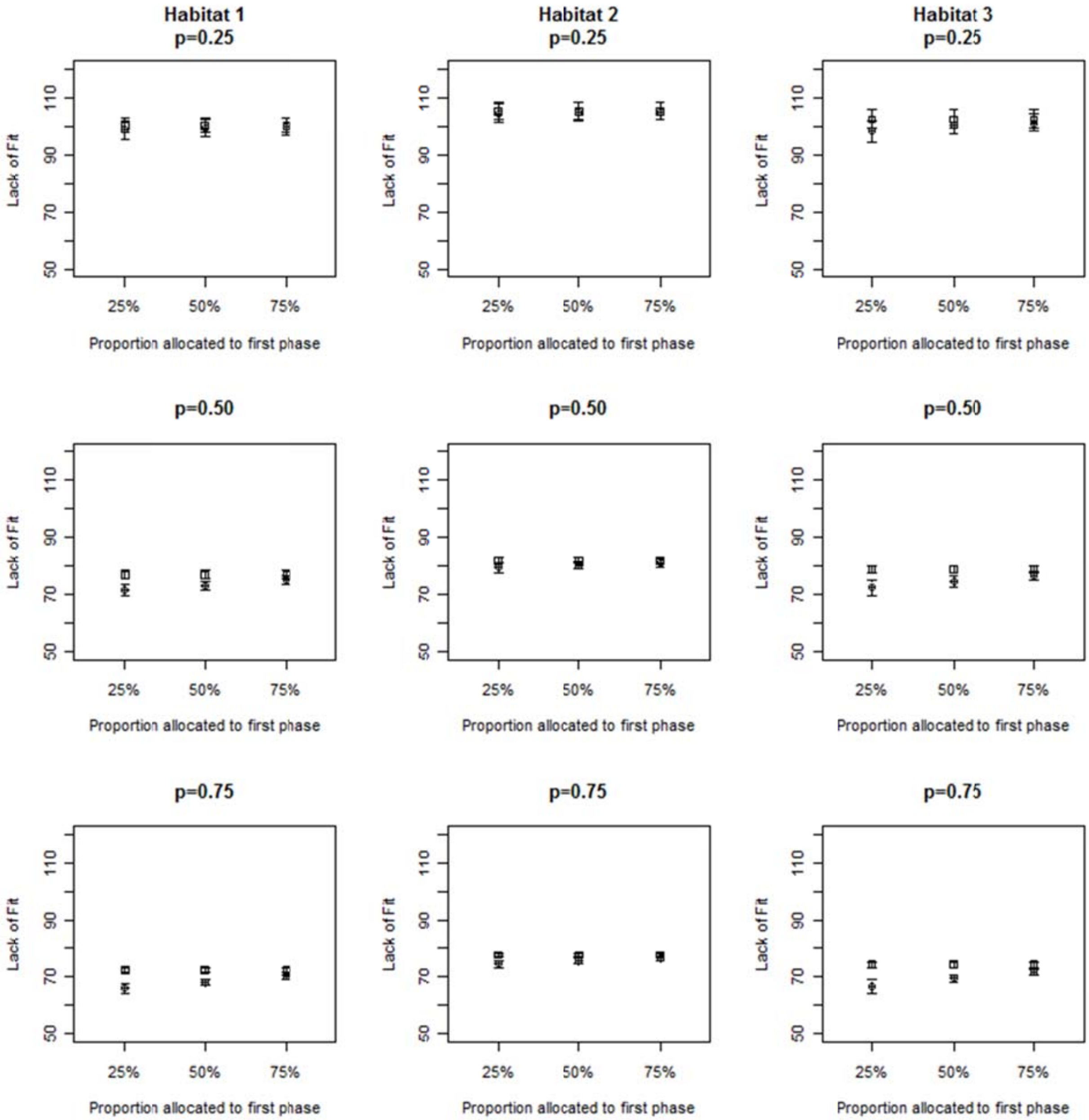
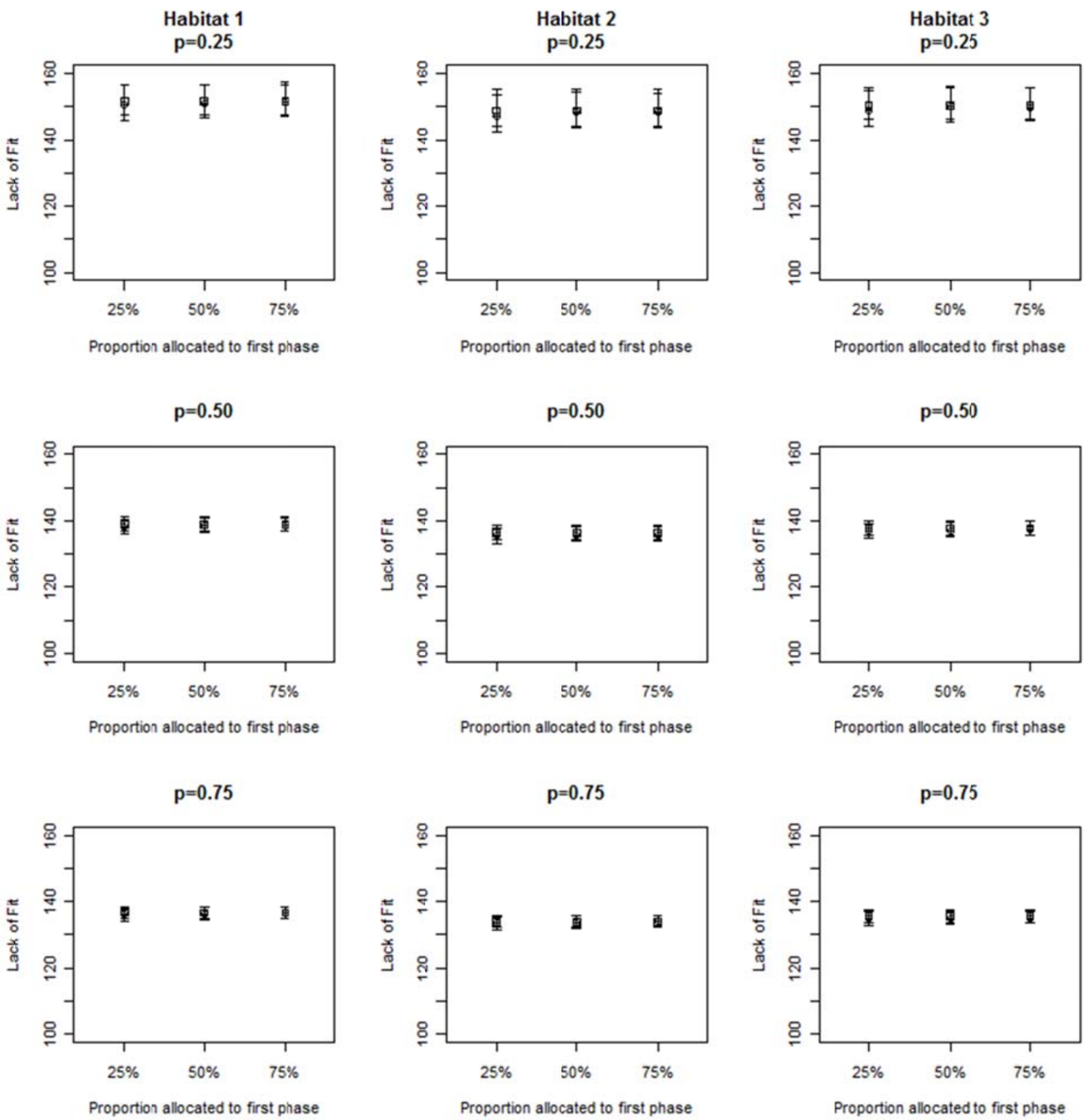

Habitat 1
p=0.25
Habitat 2
p=0.25
Habitat 3
p=0.25
p=0.50
p=0.75
Lack of Fit
Proportion allocated to first phase
25%
50%
75%
50
70
90
110

Figure B.2. Plots of test statistic and associated quartiles (25% and 75%) measuring lack of fit (lower is better) for two-phase adaptive approach (circles) and traditional single-season occupancy approach (squares) when n=200, J=5, and for three different levels of detection, p=0.25, 0.5, 0.75. X-axis represents proportion of sample allocated to phase one for the two-phase adaptive approach. Habitat 1 is a simulated habitat with extreme spatial correlation while habitat 2 has moderate spatial correlation and habitat 3 is randomly generated and contains no spatial correlation.



Habitat 1
p=0.25
Habitat 2
p=0.25
Habitat 3
p=0.25
p=0.50
p=0.50
p=0.50
p=0.75
p=0.75
p=0.75
Lack of Fit
100
120
140
160
25%
50%
75%
Proportion allocated to first phase

Figure B.3. Plots of test statistic and associated quartiles (25% and 75%) measuring lack of fit (lower is better) for two-phase adaptive approach (circles) and traditional single-season occupancy approach (squares) when there is no simulated relationship between habitat and occupancy for n=100, J=5, and for three different levels of detection, p=0.25, 0.5, 0.75. X-axis represents proportion of sample allocated to phase one for the two-phase adaptive approach. Habitat 1 is a simulated habitat with extreme spatial correlation while habitat 2 has moderate spatial correlation and habitat 3 is randomly generated and contains no spatial correlation.

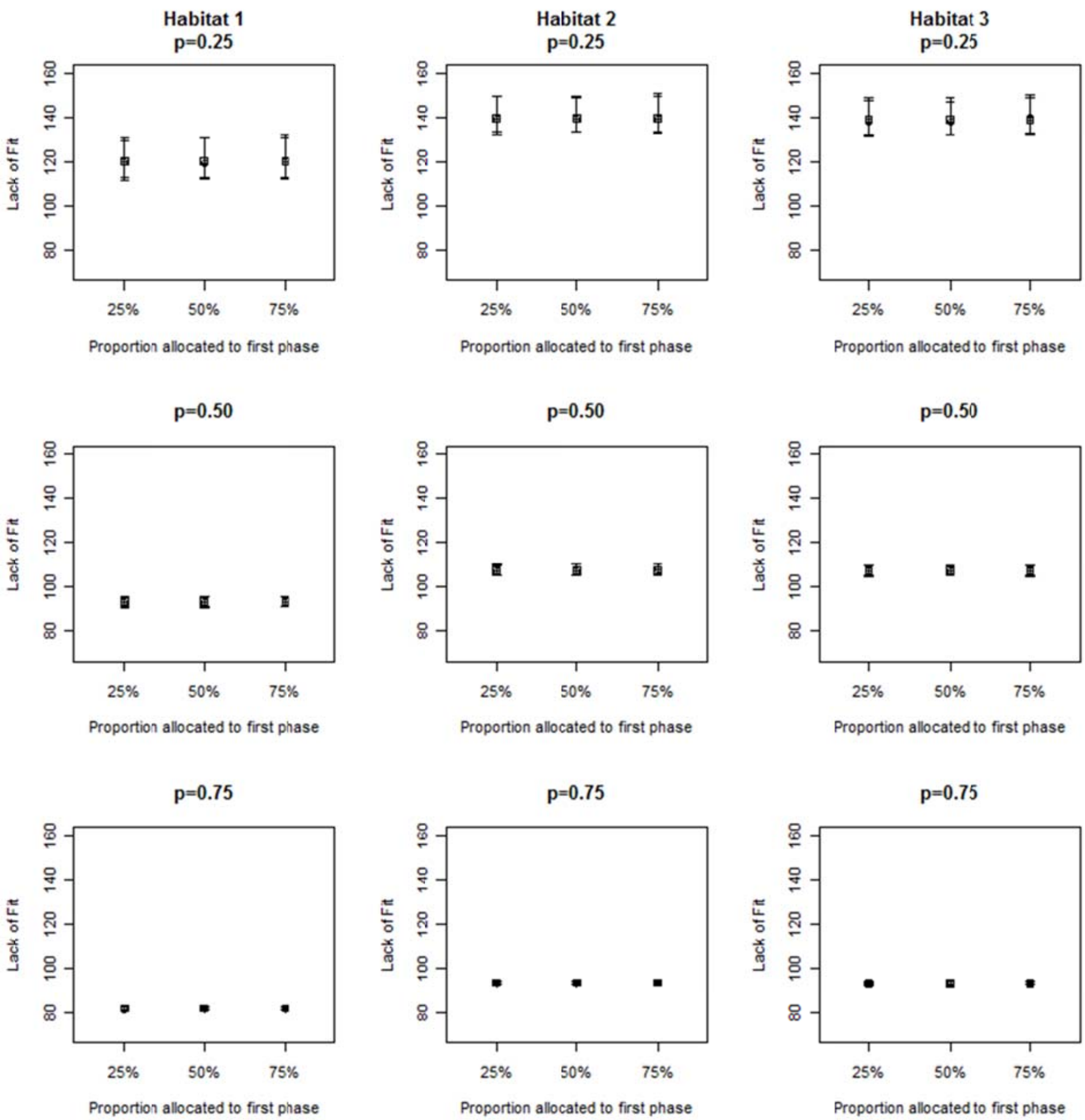

Habitat 1
Habitat 2
Habitat 3
p=0.25
p=0.50
p=0.75
Lack of Fit
80
100
120
140
160
25%
50%
75%
Proportion allocated to first phase

Figure B.4. Plots of test statistic and associated quartiles (25% and 75%) measuring lack of fit (lower is better) for two-phase adaptive approach (circles) and traditional single-season occupancy approach (squares) when there is no simulated relationship between habitat and occupancy for n=200, J=3, and for three different levels of detection, p=0.25, 0.5, 0.75. X-axis represents proportion of sample allocated to phase one for the two-phase adaptive approach. Habitat 1 is a simulated habitat with extreme spatial correlation while habitat 2 has moderate spatial correlation and habitat 3 is randomly generated and contains no spatial correlation.

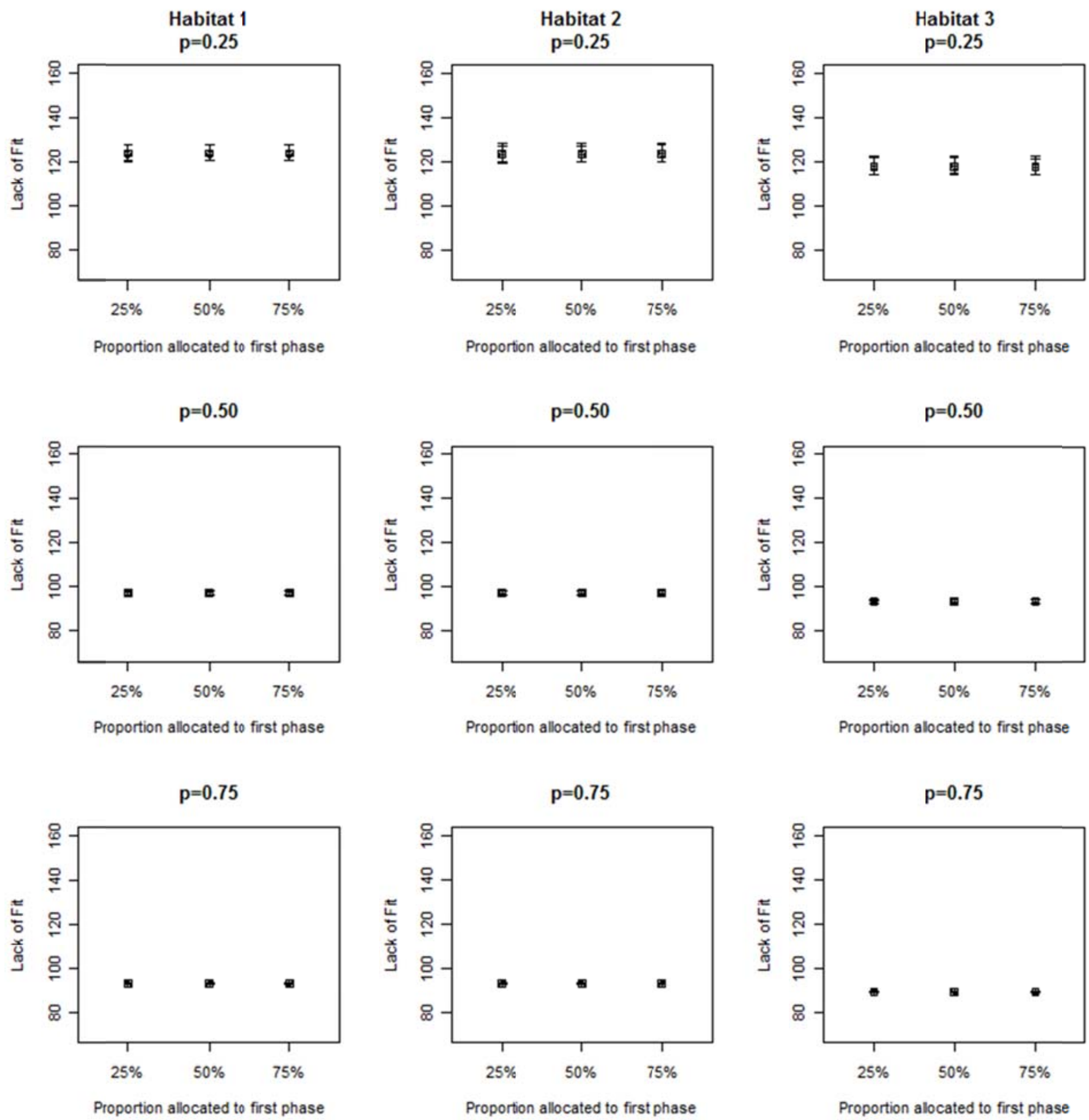

Habitat 1
p=0.25
Habitat 2
p=0.25
Habitat 3
p=0.25
p=0.50
p=0.50
p=0.50
p=0.75
p=0.75
p=0.75
Lack of Fit
Proportion allocated to first phase
25%
50%
75%
80
100
120
140
160

Figure B.5. Plots of test statistic and associated quartiles (25% and 75%) measuring lack of fit (lower is better) for two-phase adaptive approach (circles) and traditional single-season occupancy approach (squares) when there is no simulated relationship between habitat and occupancy for n=200, J=5, and for three different levels of detection, p=0.25, 0.5, 0.75. X-axis represents proportion of sample allocated to phase one for the two-phase adaptive approach. Habitat 1 is a simulated habitat with extreme spatial correlation while habitat 2 has moderate spatial correlation and habitat 3 is randomly generated and contains no spatial correlation.

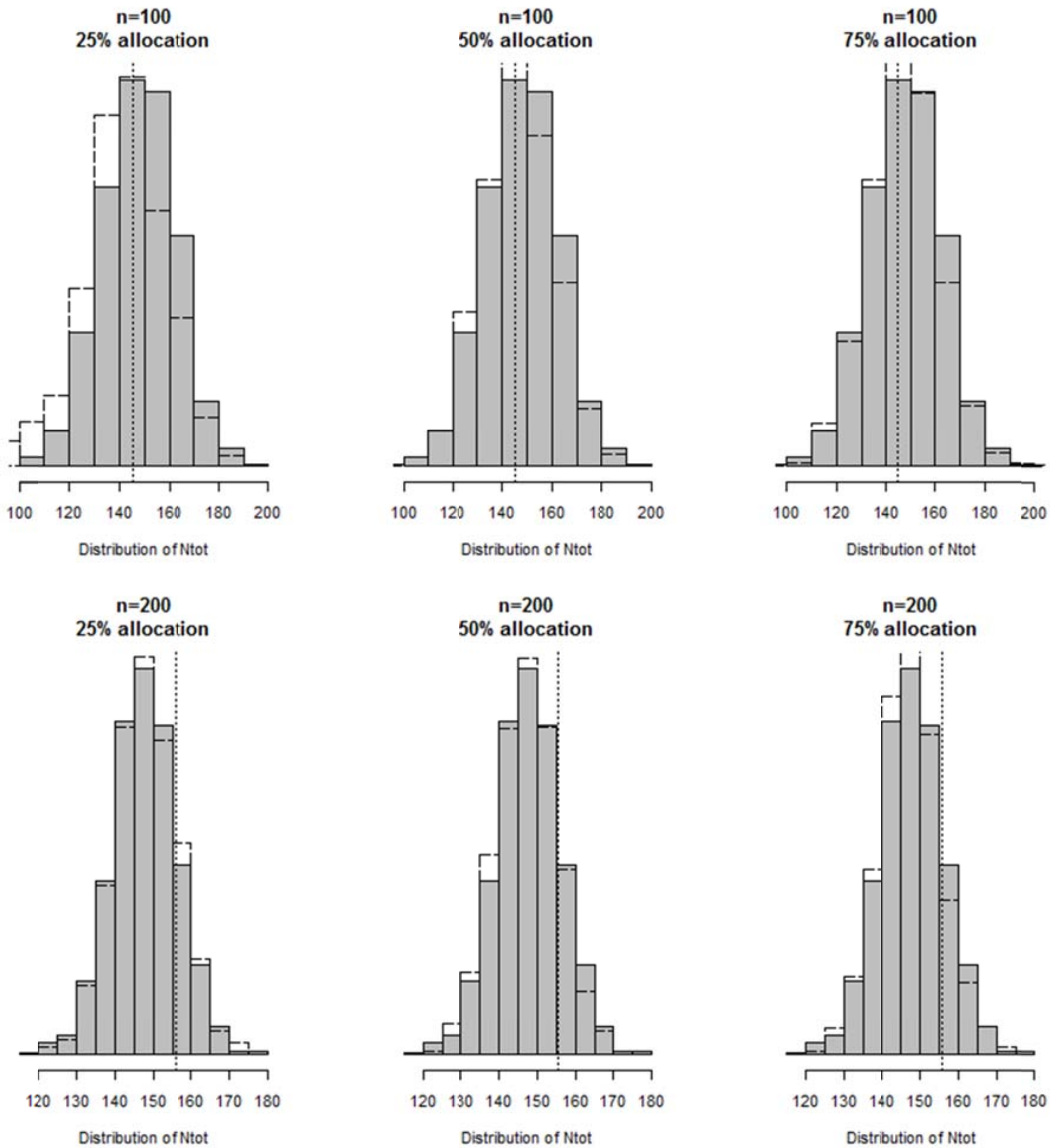

n=100
25% allocation
n=100
50% allocation
n=100
75% allocation
100 120 140 160 180 200
Distribution of Ntot
n=200
25% allocation
n=200
50% allocation
n=200
75% allocation
120 130 140 150 160 170 180
Distribution of Ntot

Figure B.6. Distribution from 1000 simulations of the estimates of N^{tot} for the two-phase adaptive approach (empty dotted rectangles) and the traditional single-season occupancy approach (gray filled rectangles) in Habitat 1 (extreme spatial correlation, see text for more details) with the number of repeat visits, J, equal to 3 and detection probability, p, equal to 0.75. Percent allocation refers to the percentage of the total sample size, n, allocated to the first phase of sampling for the two-phase adaptive approach. The dotted line represents the true value of N^{tot}.

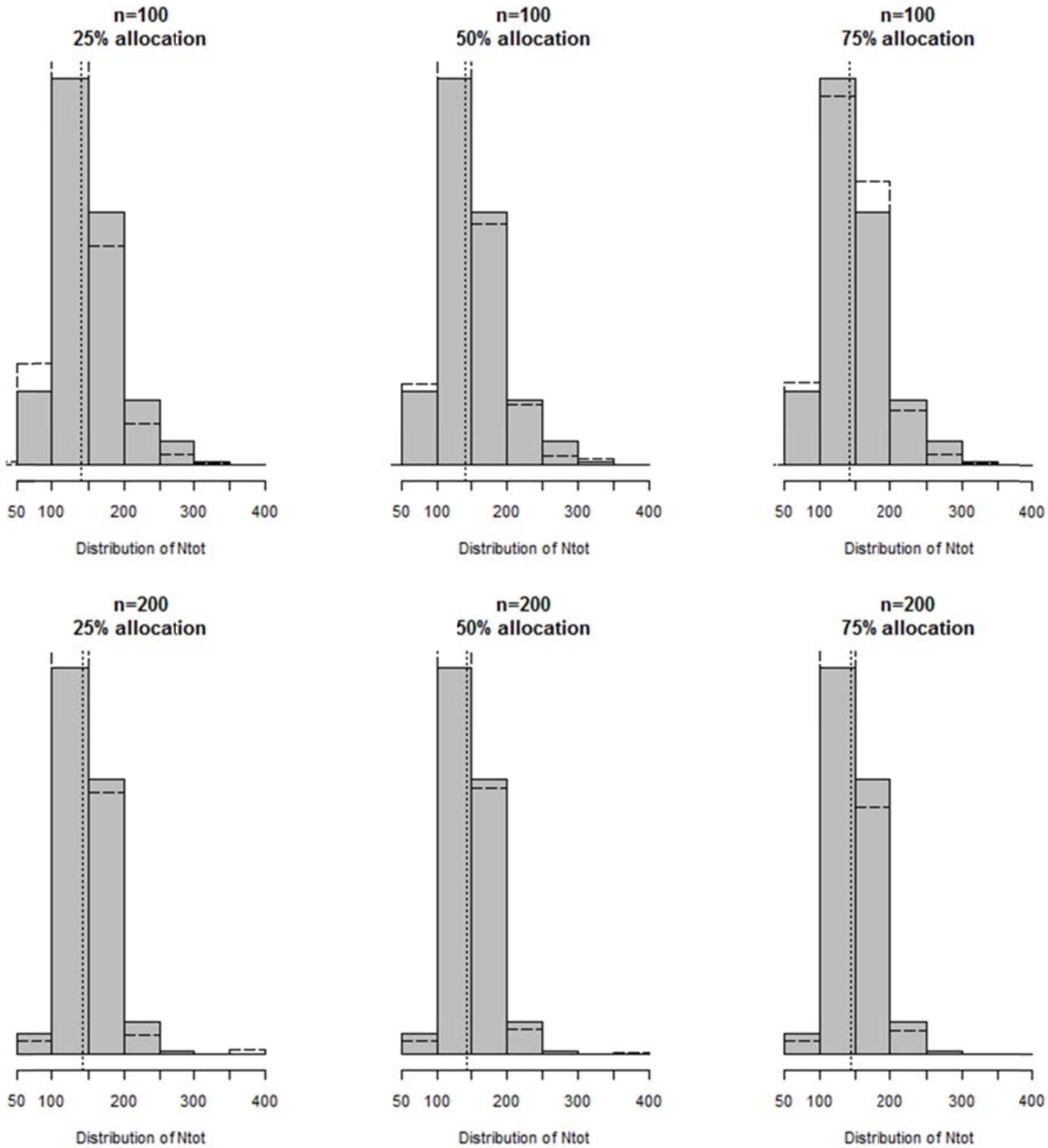

n=100
25% allocation
50 100 200 300 400
Distribution of Ntot
n=100
50% allocation
50 100 200 300 400
Distribution of Ntot
n=100
75% allocation
50 100 200 300 400
Distribution of Ntot
n=200
25% allocation
50 100 200 300 400
Distribution of Ntot
n=200
50% allocation
50 100 200 300 400
Distribution of Ntot
n=200
75% allocation
50 100 200 300 400
Distribution of Ntot

Figure B.7. Distribution from 1000 simulations of the estimates of N^{tot} for the two-phase adaptive approach (empty dotted rectangles) and the traditional single-season occupancy approach (gray filled rectangles) in Habitat 2 (moderate spatial correlation, see text for more details) with the number of repeat visits, J, equal to 3 and detection probability, p, equal to 0.25. Percent allocation refers to the percentage of the total sample size, n, allocated to the first phase of sampling for the two-phase adaptive approach. The dotted line represents the true value of N^{tot}.

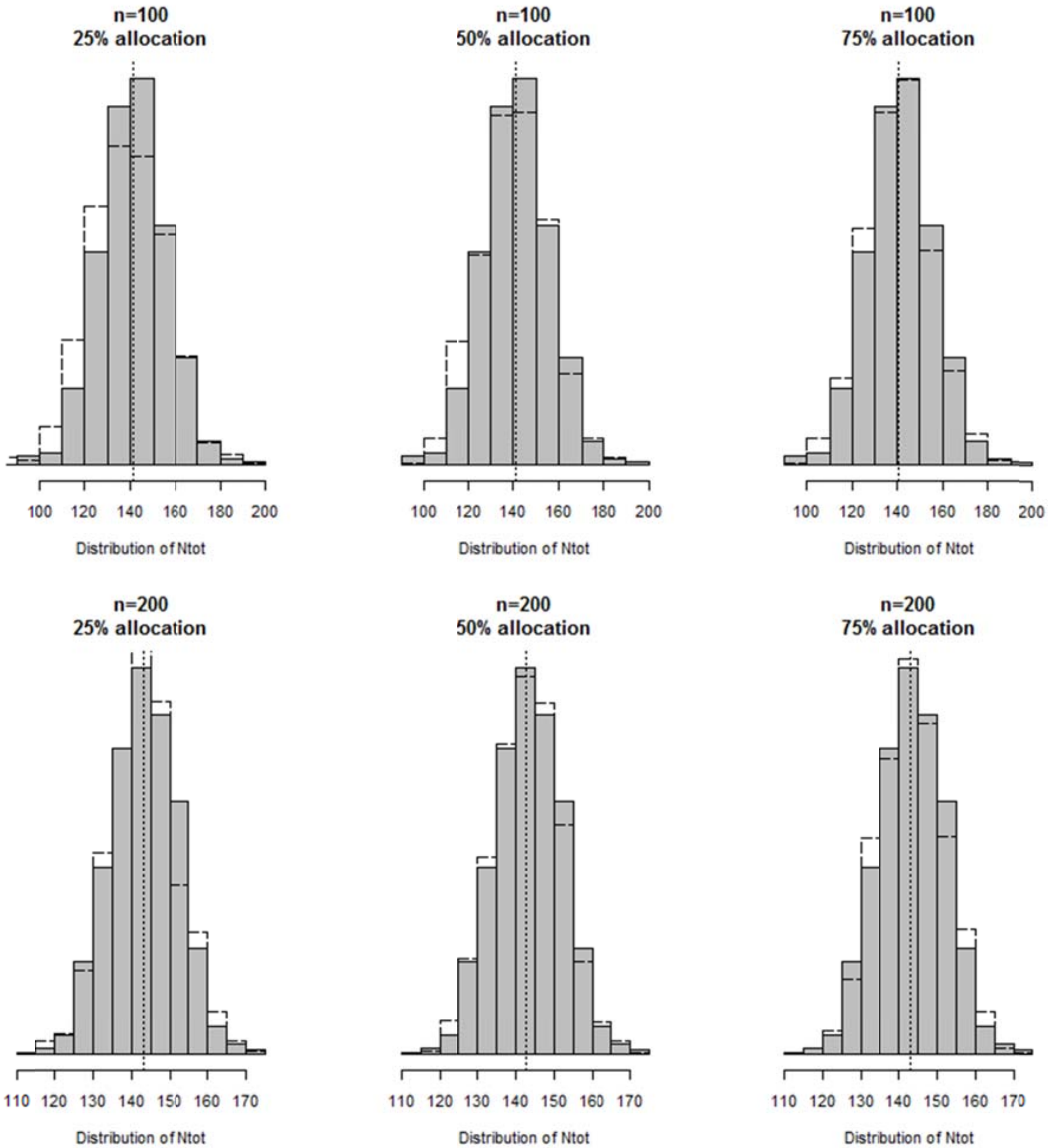
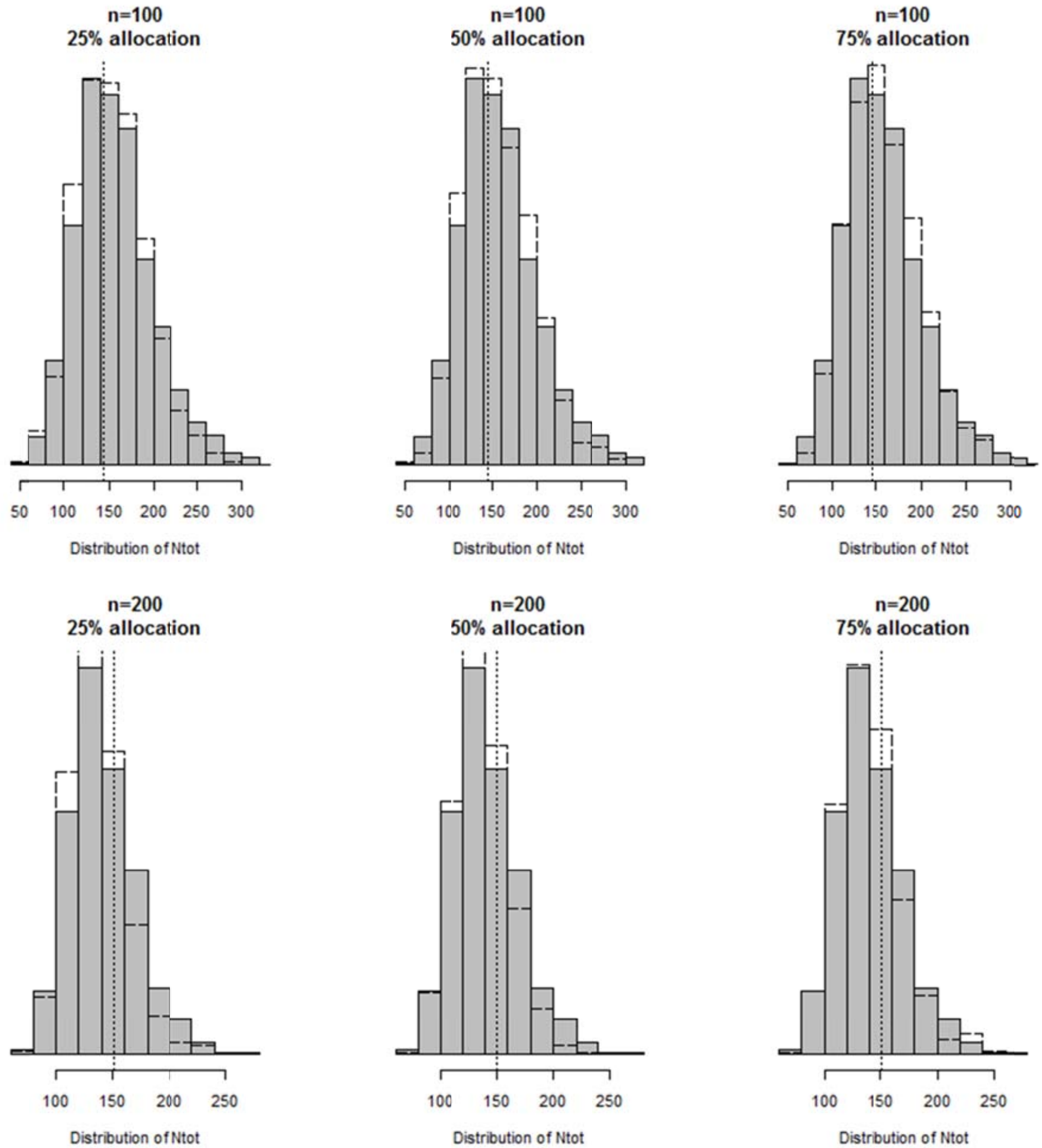

n=100
25% allocation
100 120 140 160 180 200
Distribution of Ntot
n=100
50% allocation
100 120 140 160 180 200
Distribution of Ntot
n=100
75% allocation
100 120 140 160 180 200
Distribution of Ntot
n=200
25% allocation
110 120 130 140 150 160 170
Distribution of Ntot
n=200
50% allocation
110 120 130 140 150 160 170
Distribution of Ntot
n=200
75% allocation
110 120 130 140 150 160 170
Distribution of Ntot

Figure B.8. Distribution from 1000 simulations of the estimates of N^{tot} for the two-phase adaptive approach (empty dotted rectangles) and the traditional single-season occupancy approach (gray filled rectangles) in Habitat 2 (moderate spatial correlation, see text for more details) with the number of repeat visits, J, equal to 3 and detection probability, p, equal to 0.75. Percent allocation refers to the percentage of the total sample size, n, allocated to the first phase of sampling for the two-phase adaptive approach. The dotted line represents the true value of N^{tot}.



n=100
25% allocation
n=100
50% allocation
n=100
75% allocation
50 100 150 200 250 300
Distribution of Ntot
n=200
25% allocation
n=200
50% allocation
n=200
75% allocation
100 150 200 250
Distribution of Ntot

Figure B.9. Distribution from 1000 simulations of the estimates of N^{tot} for the two-phase adaptive approach (empty dotted rectangles) and the traditional single-season occupancy approach (gray filled rectangles) in Habitat 3 (no spatial correlation, see text for more details) with the number of repeat visits, J, equal to 3 and detection probability, p, equal to 0.25. Percent allocation refers to the percentage of the total sample size, n, allocated to the first phase of sampling for the two-phase adaptive approach. The dotted line represents the true value of N^{tot}.

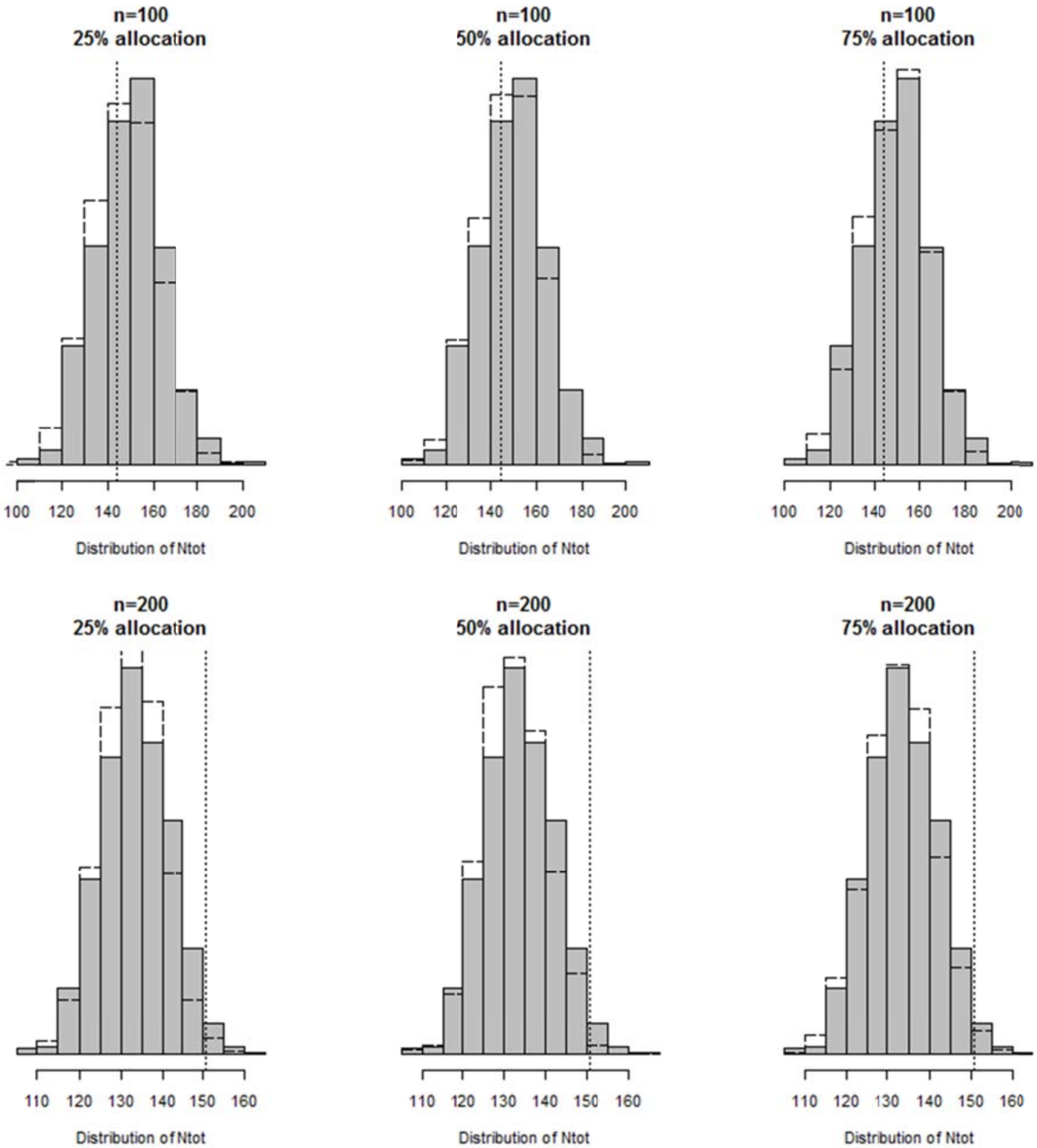

n=100
25% allocation
n=100
50% allocation
n=100
75% allocation
100 120 140 160 180 200
Distribution of Ntot
n=200
25% allocation
n=200
50% allocation
n=200
75% allocation
110 120 130 140 150 160
Distribution of Ntot

Figure B.10. Distribution from 1000 simulations of the estimates of N^{tot} for the two-phase adaptive approach (empty dotted rectangles) and the traditional single-season occupancy approach (gray filled rectangles) in Habitat 3 (no spatial correlation, see text for more details) with the number of repeat visits, J, equal to 3 and detection probability, p, equal to 0.75. Percent allocation refers to the percentage of the total sample size, n, allocated to the first phase of sampling for the two-phase adaptive approach. The dotted line represents the true value of N^{tot}.

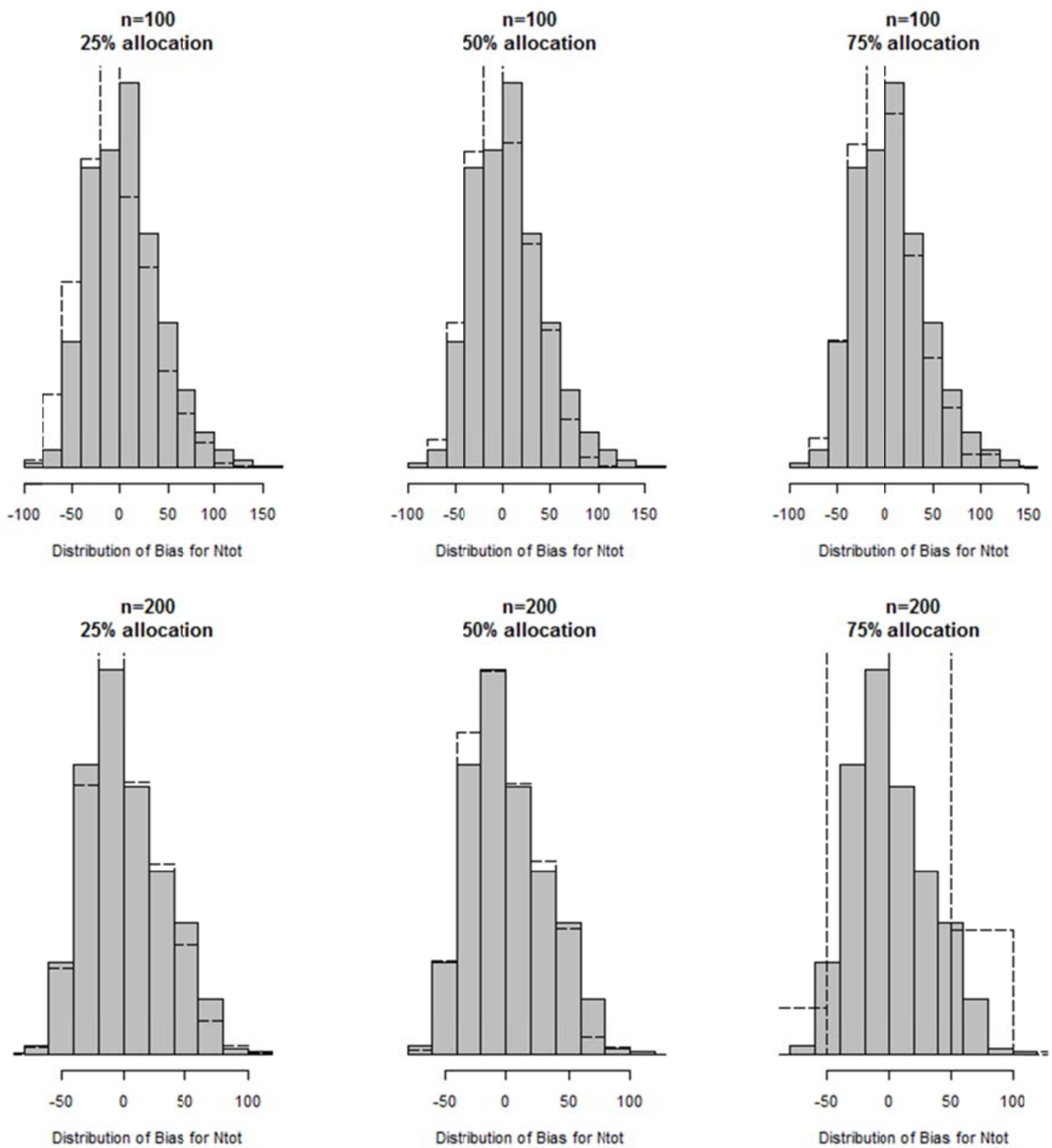

n=100
25% allocation
n=100
50% allocation
n=100
75% allocation
-100 -50 0 50 100 150
Distribution of Bias for Ntot
n=200
25% allocation
n=200
50% allocation
n=200
75% allocation
-50 0 50 100
Distribution of Bias for Ntot

Figure B.11. Distribution from 1000 simulations of the bias in estimates of N^{tot} for the two-phase adaptive approach (empty dotted rectangles) and the traditional single-season occupancy approach (gray filled rectangles) in Habitat 1 (extreme spatial correlation, see text for more details) with the number of repeat visits, J, equal to 3 and detection probability, p, equal to 0.25. Percent allocation refers to the percentage of the total sample size, n, allocated to the first phase of sampling for the two-phase adaptive approach.

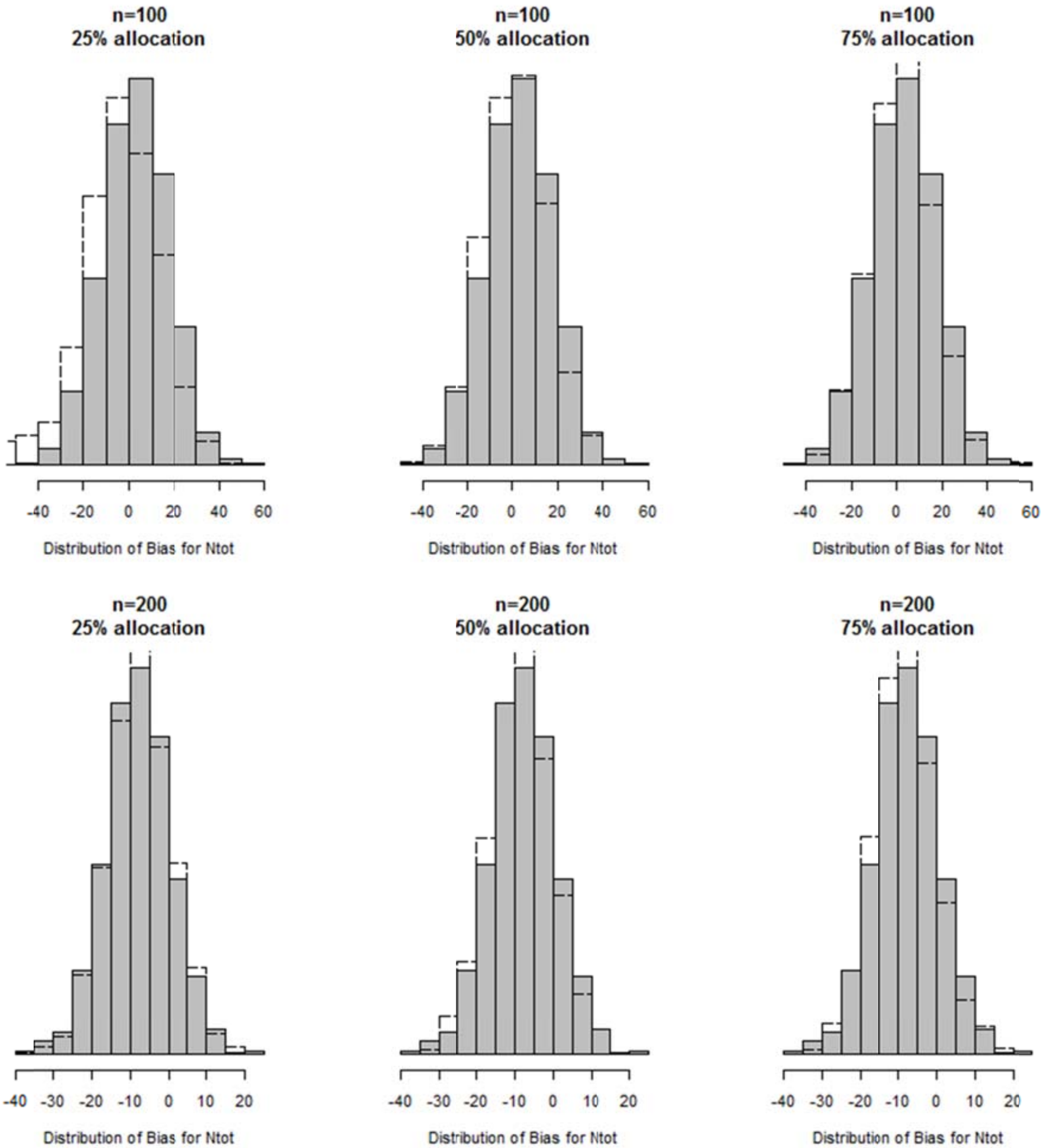
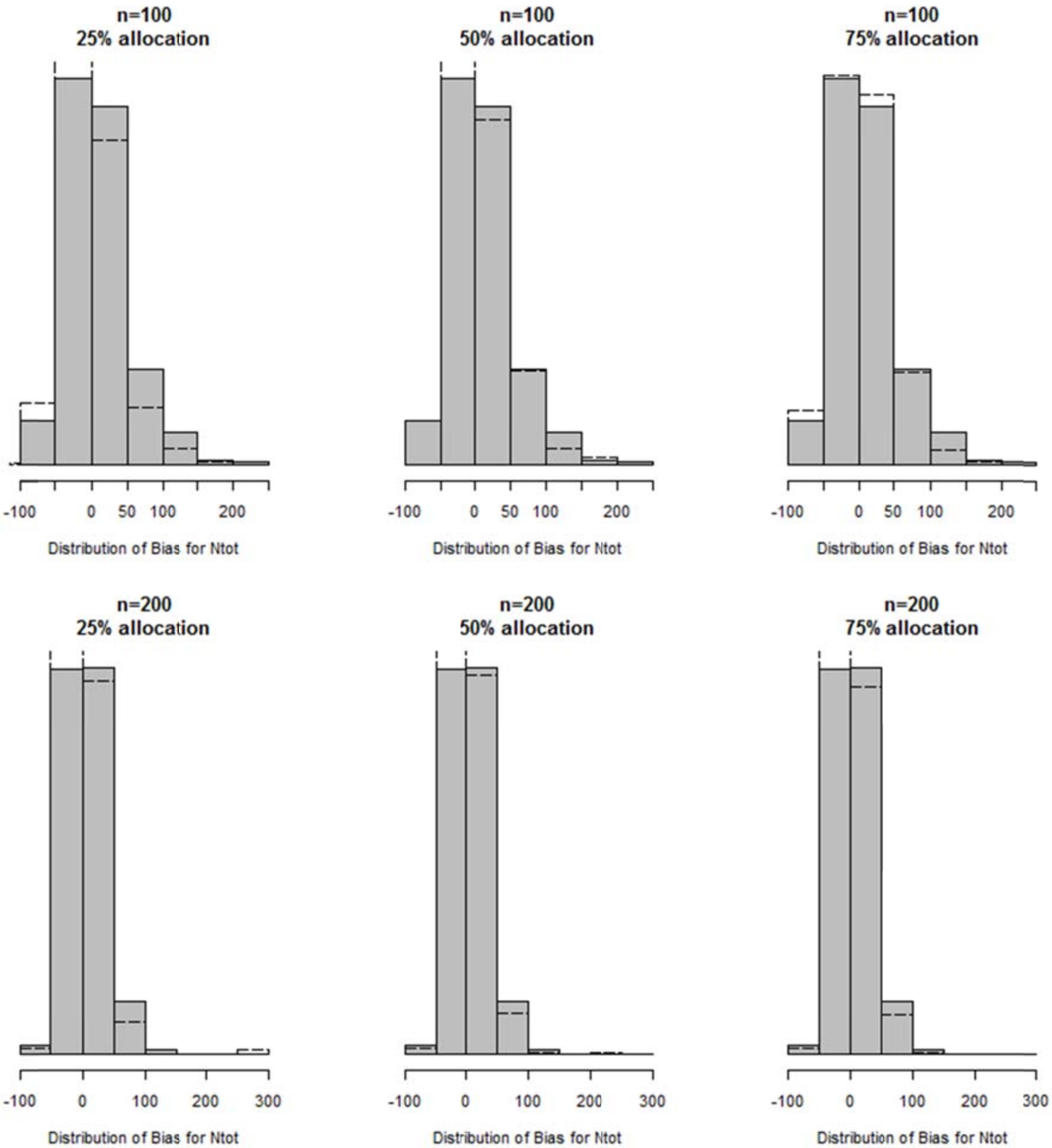
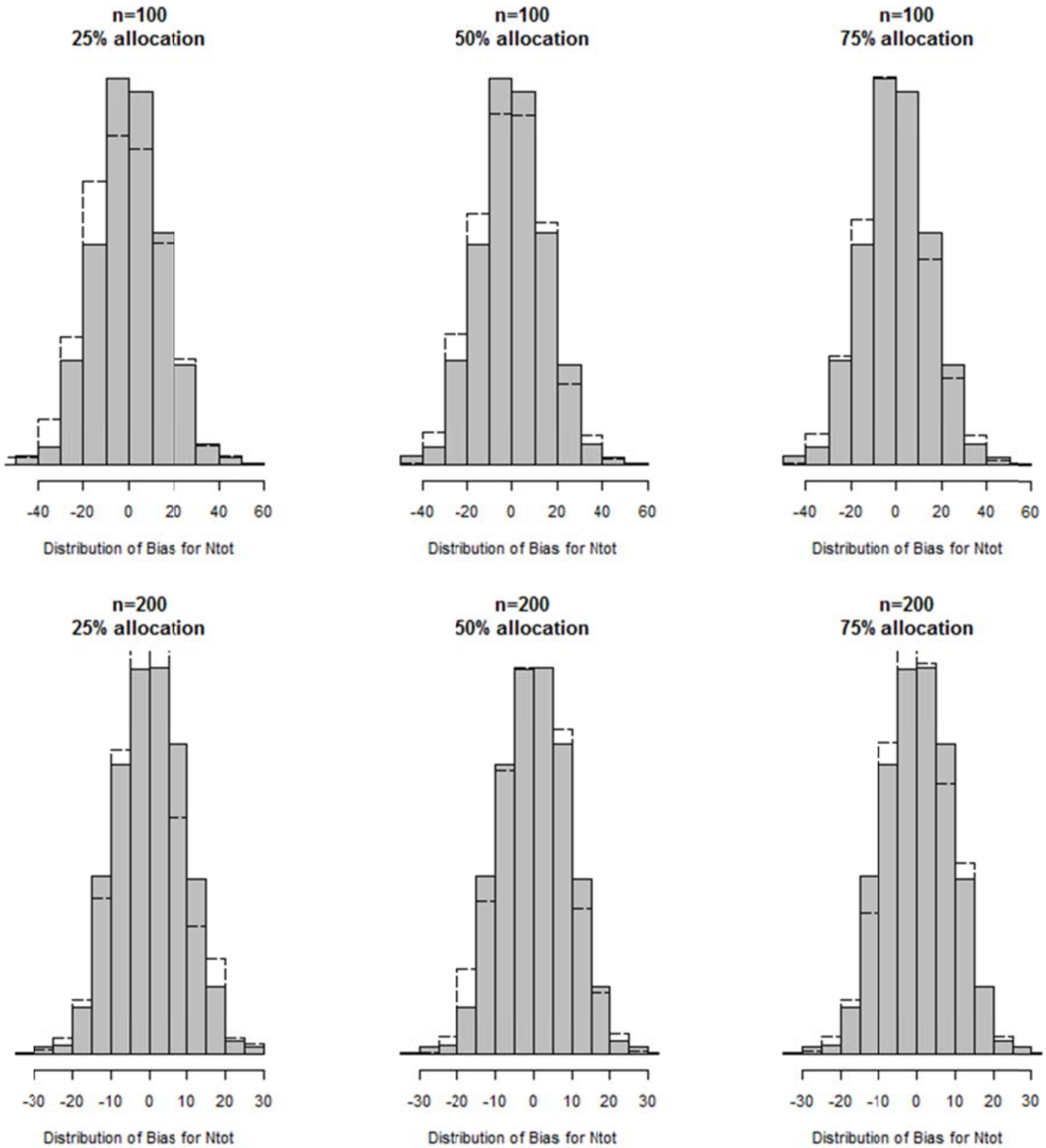

n=100
25% allocation
n=100
50% allocation
n=100
75% allocation
-40 -20 0 20 40 60
Distribution of Bias for Ntot
n=200
25% allocation
n=200
50% allocation
n=200
75% allocation
-40 -30 -20 -10 0 10 20
Distribution of Bias for Ntot

Figure B.12. Distribution from 1000 simulations of the bias in estimates of N^{tot} for the two-phase adaptive approach (empty dotted rectangles) and the traditional single-season occupancy approach (gray filled rectangles) in Habitat 1 (extreme spatial correlation, see text for more details) with the number of repeat visits, J, equal to 3 and detection probability, p, equal to 0.75. Percent allocation refers to the percentage of the total sample size, n, allocated to the first phase of sampling for the two-phase adaptive approach.



n=100
25% allocation
-100
0
50
100
200
Distribution of Bias for Ntot
n=100
50% allocation
-100
0
50
100
200
Distribution of Bias for Ntot
n=100
75% allocation
-100
0
50
100
200
Distribution of Bias for Ntot
n=200
25% allocation
-100
0
100
200
300
Distribution of Bias for Ntot
n=200
50% allocation
-100
0
100
200
300
Distribution of Bias for Ntot
n=200
75% allocation
-100
0
100
200
300
Distribution of Bias for Ntot

Figure B.13. Distribution from 1000 simulations of the bias in estimates of N^{tot} for the two-phase adaptive approach (empty dotted rectangles) and the traditional single-season occupancy approach (gray filled rectangles) in Habitat 2 (moderate spatial correlation, see text for more details) with the number of repeat visits, J, equal to 3 and detection probability, p, equal to 0.25. Percent allocation refers to the percentage of the total sample size, n, allocated to the first phase of sampling for the two-phase adaptive approach.



n=100
25% allocation
-40 -20 0 20 40 60
Distribution of Bias for Ntot
n=100
50% allocation
-40 -20 0 20 40 60
Distribution of Bias for Ntot
n=100
75% allocation
-40 -20 0 20 40 60
Distribution of Bias for Ntot
n=200
25% allocation
-30 -20 -10 0 10 20 30
Distribution of Bias for Ntot
n=200
50% allocation
-30 -20 -10 0 10 20 30
Distribution of Bias for Ntot
n=200
75% allocation
-30 -20 -10 0 10 20 30
Distribution of Bias for Ntot

Figure B.14. Distribution from 1000 simulations of the bias in estimates of N^{tot} for the two-phase adaptive approach (empty dotted rectangles) and the traditional single-season occupancy approach (gray filled rectangles) in Habitat 2 (moderate spatial correlation, see text for more details) with the number of repeat visits, J, equal to 3 and detection probability, p, equal to 0.75. Percent allocation refers to the percentage of the total sample size, n, allocated to the first phase of sampling for the two-phase adaptive approach.

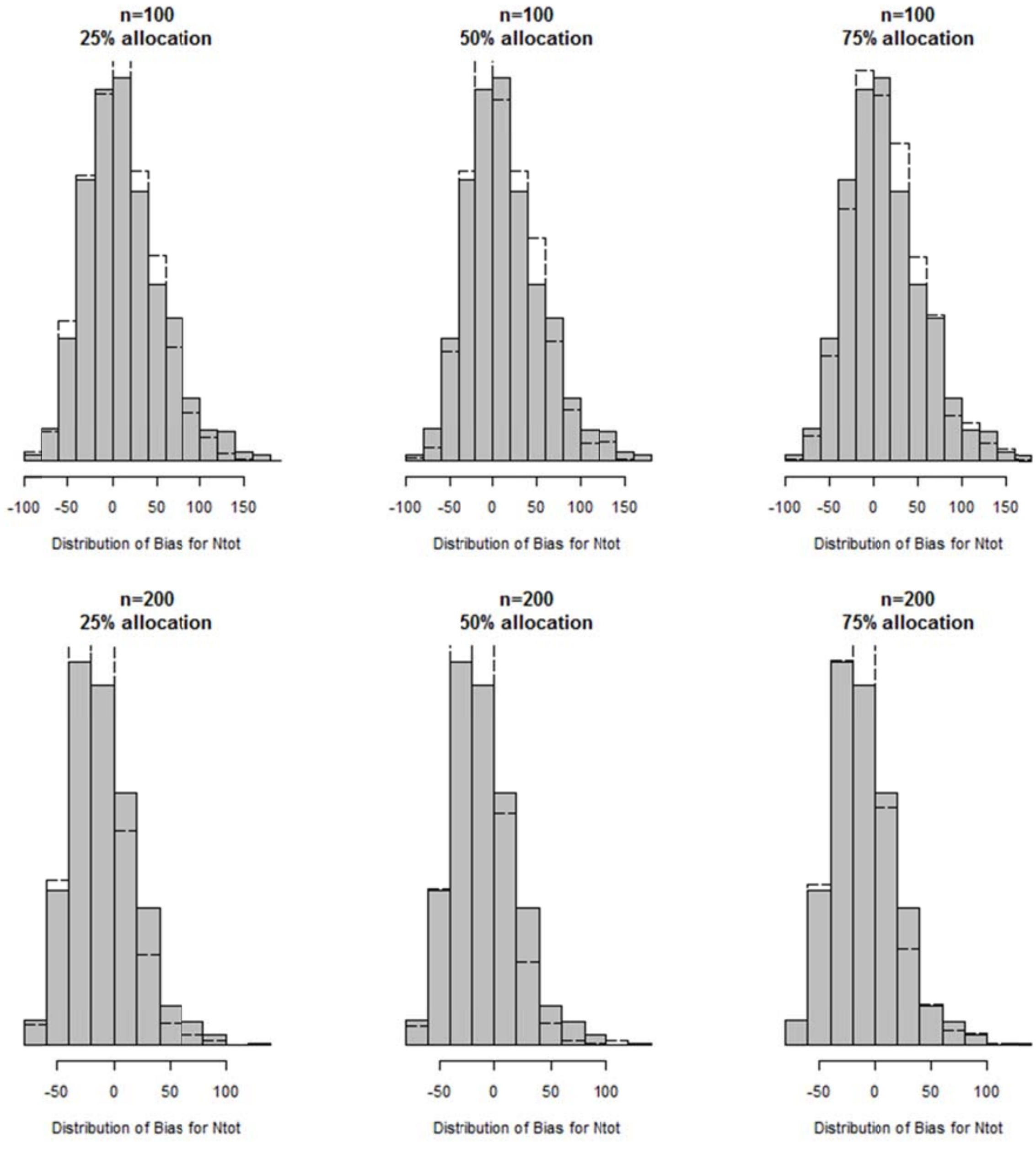

n=100
25% allocation
n=100
50% allocation
n=100
75% allocation
-100 -50 0 50 100 150
Distribution of Bias for Ntot
n=200
25% allocation
n=200
50% allocation
n=200
75% allocation
-50 0 50 100
Distribution of Bias for Ntot

Figure B.15. Distribution from 1000 simulations of the bias in estimates of N^{tot} for the two-phase adaptive approach (empty dotted rectangles) and the traditional single-season occupancy approach (gray filled rectangles) in Habitat 3 (no spatial correlation, see text for more details) with the number of repeat visits, J, equal to 3 and detection probability, p, equal to 0.25. Percent allocation refers to the percentage of the total sample size, n, allocated to the first phase of sampling for the two-phase adaptive approach.

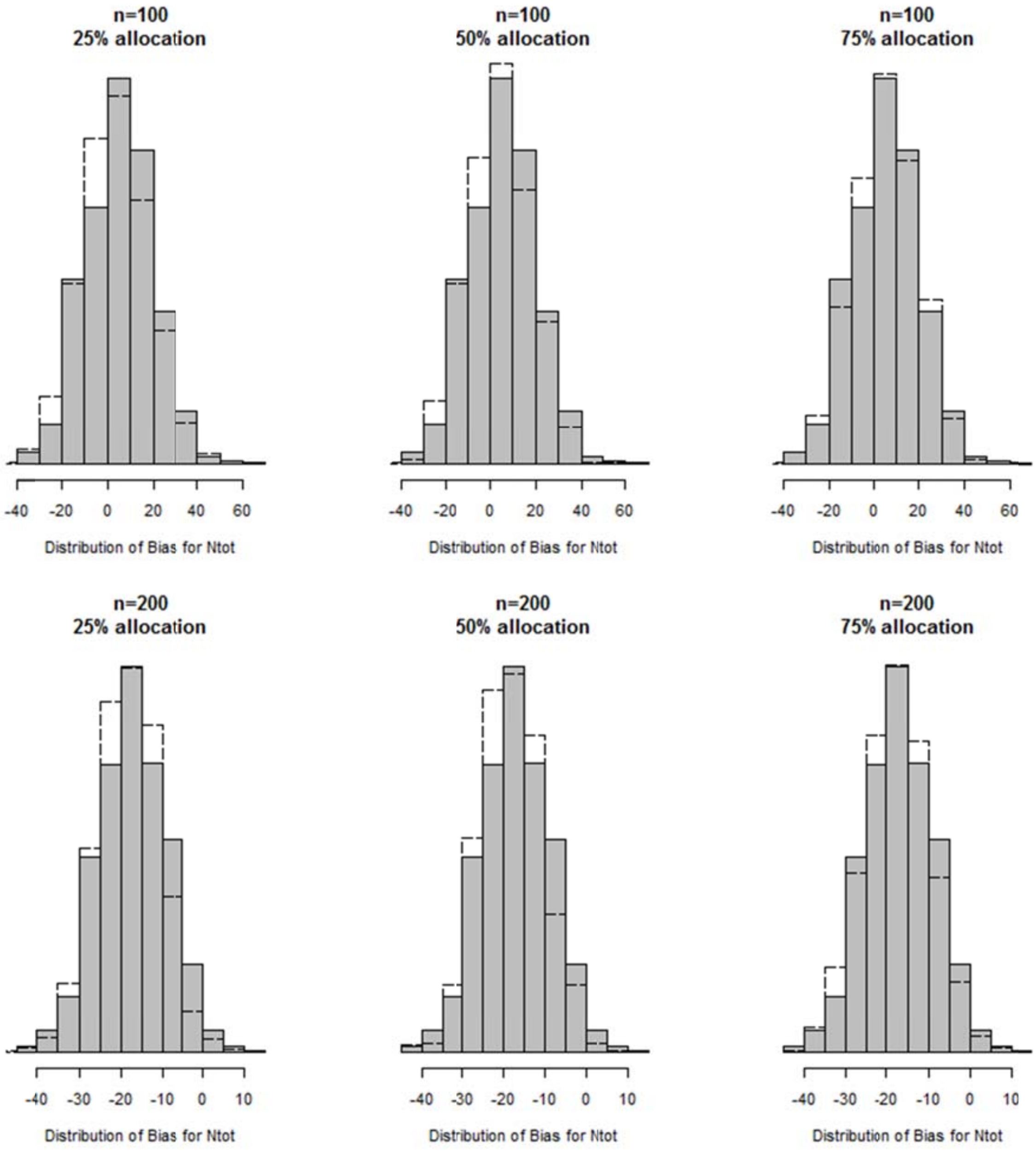

n=100
25% allocation
-40 -20 0 20 40 60
Distribution of Bias for Ntot
n=100
50% allocation
-40 -20 0 20 40 60
Distribution of Bias for Ntot
n=100
75% allocation
-40 -20 0 20 40 60
Distribution of Bias for Ntot
n=200
25% allocation
-40 -30 -20 -10 0 10
Distribution of Bias for Ntot
n=200
50% allocation
-40 -30 -20 -10 0 10
Distribution of Bias for Ntot
n=200
75% allocation
-40 -30 -20 -10 0 10
Distribution of Bias for Ntot

Figure B.16. Distribution from 1000 simulations of the bias in estimates of N^{tot} for the two-phase adaptive approach (empty dotted rectangles) and the traditional single-season occupancy approach (gray filled rectangles) in Habitat 3 (no spatial correlation, see text for more details) with the number of repeat visits, J, equal to 3 and detection probability, p, equal to 0.75. Percent allocation refers to the percentage of the total sample size, n, allocated to the first phase of sampling for the two-phase adaptive approach.

Table B.1. Average value from 1000 simulations of estimates of test statistic measuring lack of fit, $\boldsymbol{d}$, test statistic measuring goodness of fit, $\boldsymbol{GOF}$, estimate of N^{tot}, $\widehat{\boldsymbol{N}^{tot}}$, bias associated in estimating N^{tot}, and mean-squared-error, ***MSE***, associated with estimating N^{tot}. Estimates of coefficients in the logit-linear model for occupancy probability, $\widehat{\boldsymbol{\beta_0}}$ and $\widehat{\boldsymbol{\beta_1}}$, along with average bias and mean-squared-error (***MSE***). $\boldsymbol{N^{tot}}$ represents the true value of the total number of sites occupied out of 400. $\boldsymbol{n_1}$ and $\boldsymbol{n_2}$ represent the sample size at each phase of sampling. $\boldsymbol{n_2} = 0$ represents the traditional single-season occupancy approach. ***Habitat*** represents three different habitats with varying degrees of spatial correlation (1-extreme spatial correlation, 2-moderate spatial correlation, 3-no spatial correlation).

N	n_1	n_2	*p*	*J*	*habitat*	*approach*	*d*	*GOF*	N^{tot}*-hat*	*Bias* N^{tot}	*MSE* N^{tot}	N^{tot}	ψ
100	50	50	0.25	3	1	1	114.755	285.245	150.480	5.340	1407.308	145.140	0.370
100	50	50	0.5	3	1	1	102.159	297.841	147.745	2.605	320.458	145.140	0.370
100	50	50	0.75	3	1	1	95.062	304.938	147.913	2.773	218.215	145.140	0.370
100	50	50	0.25	3	1	2	111.088	288.912	144.188	-0.952	1035.041	145.140	0.370
100	50	50	0.5	3	1	2	98.113	301.887	146.220	1.080	310.262	145.140	0.370
100	50	50	0.75	3	1	2	90.353	309.647	146.240	1.100	201.638	145.140	0.370
100	25	75	0.25	3	1	2	107.554	292.446	140.491	-4.649	1314.674	145.140	0.370
100	25	75	0.5	3	1	2	95.307	304.693	142.073	-3.066	409.661	145.140	0.370
100	25	75	0.75	3	1	2	87.063	312.937	141.484	-3.655	305.384	145.140	0.370
100	75	25	0.25	3	1	2	113.701	286.299	145.722	0.582	1206.157	145.140	0.370
100	75	25	0.5	3	1	2	100.375	299.625	147.991	2.852	295.876	145.140	0.370
100	75	25	0.75	3	1	2	92.861	307.139	147.135	1.995	201.181	145.140	0.370
100	50	50	0.25	5	1	1	115.546	284.454	158.422	2.702	657.654	155.720	0.387
100	50	50	0.5	5	1	1	103.911	296.089	155.394	-0.325	224.154	155.720	0.387
100	50	50	0.75	5	1	1	101.807	298.193	155.928	0.208	209.604	155.720	0.387

100	50	50	0.25	5	1	2	113.960	286.040	156.317	0.598	567.511	155.720	0.387
100	50	50	0.5	5	1	2	101.544	298.456	154.722	-0.998	216.550	155.720	0.387
100	50	50	0.75	5	1	2	98.945	301.055	153.762	-1.958	228.373	155.720	0.387
100	25	75	0.25	5	1	2	112.653	287.347	156.331	0.611	675.991	155.720	0.387
100	25	75	0.5	5	1	2	100.075	299.925	152.981	-2.739	275.358	155.720	0.387
100	25	75	0.75	5	1	2	97.223	302.777	152.927	-2.793	266.763	155.720	0.387
100	75	25	0.25	5	1	2	114.782	285.218	157.002	1.282	625.160	155.720	0.387
100	75	25	0.5	5	1	2	102.768	297.232	155.650	-0.070	219.510	155.720	0.387
100	75	25	0.75	5	1	2	100.593	299.407	155.238	-0.482	226.265	155.720	0.387
200	100	100	0.25	3	1	1	103.380	296.620	156.644	0.977	998.222	155.667	0.368
200	100	100	0.5	3	1	1	80.143	319.857	148.410	-7.257	197.677	155.667	0.368
200	100	100	0.75	3	1	1	66.856	333.144	147.750	-7.917	134.989	155.667	0.368
200	100	100	0.25	3	1	2	102.566	297.434	155.041	-0.626	908.924	155.667	0.368
200	100	100	0.5	3	1	2	76.823	323.177	147.958	-7.709	179.998	155.667	0.368
200	100	100	0.75	3	1	2	61.024	338.976	147.128	-8.539	138.919	155.667	0.368
200	50	150	0.25	3	1	2	101.870	298.130	155.348	-0.319	859.942	155.667	0.368

200	50	150	0.5	3	1	2	74.834	325.166	148.241	-7.426	185.475	155.667	0.368
200	50	150	0.75	3	1	2	58.440	341.560	148.144	-7.523	126.438	155.667	0.368
200	150	50	0.25	3	1	2	103.795	296.205	156.860	1.194	1108.037	155.667	0.368
200	150	50	0.5	3	1	2	78.430	321.570	148.429	-7.238	186.988	155.667	0.368
200	150	50	0.75	3	1	2	63.884	336.116	147.245	-8.421	138.432	155.667	0.368
200	100	100	0.25	5	1	1	100.665	299.335	159.250	2.831	319.297	156.419	0.387
200	100	100	0.5	5	1	1	77.152	322.848	155.533	-0.886	86.793	156.419	0.387
200	100	100	0.75	5	1	1	72.579	327.421	155.151	-1.268	70.004	156.419	0.387
200	100	100	0.25	5	1	2	99.721	300.279	157.264	0.845	272.283	156.419	0.387
200	100	100	0.5	5	1	2	73.278	326.722	154.590	-1.830	84.281	156.419	0.387
200	100	100	0.75	5	1	2	68.036	331.964	155.275	-1.144	69.946	156.419	0.387
200	50	150	0.25	5	1	2	98.889	301.111	156.764	0.345	262.498	156.419	0.387
200	50	150	0.5	5	1	2	71.484	328.516	155.553	-0.866	82.876	156.419	0.387
200	50	150	0.75	5	1	2	66.143	333.857	155.174	-1.245	79.626	156.419	0.387
200	150	50	0.25	5	1	2	100.107	299.893	157.104	0.685	280.170	156.419	0.387
200	150	50	0.5	5	1	2	75.162	324.838	155.145	-1.274	81.886	156.419	0.387

200	150	50	0.75	5	1	2	70.183	329.817	155.012	-1.407	78.988	156.419	0.387
100	50	50	0.25	3	2	1	129.117	270.883	149.873	8.873	2032.738	141.000	0.353
100	50	50	0.5	3	2	1	114.089	285.911	142.655	1.655	366.643	141.000	0.353
100	50	50	0.75	3	2	1	106.980	293.020	141.005	0.005	214.669	141.000	0.353
100	50	50	0.25	3	2	2	127.072	272.928	146.771	5.771	1773.656	141.000	0.353
100	50	50	0.5	3	2	2	112.168	287.832	141.025	0.025	331.710	141.000	0.353
100	50	50	0.75	3	2	2	105.305	294.695	140.119	-0.881	227.736	141.000	0.353
100	25	75	0.25	3	2	2	124.113	275.887	139.564	-1.436	1541.001	141.000	0.353
100	25	75	0.5	3	2	2	111.311	288.689	139.684	-1.316	418.092	141.000	0.353
100	25	75	0.75	3	2	2	103.762	296.238	139.074	-1.926	272.805	141.000	0.353
100	75	25	0.25	3	2	2	127.516	272.484	146.480	5.480	1584.779	141.000	0.353
100	75	25	0.5	3	2	2	113.423	286.577	141.053	0.053	313.394	141.000	0.353
100	75	25	0.75	3	2	2	105.973	294.027	140.368	-0.632	218.706	141.000	0.353
100	50	50	0.25	5	2	1	124.837	275.163	148.568	2.568	621.482	146.000	0.365
100	50	50	0.5	5	2	1	112.790	287.210	145.993	-0.007	258.456	146.000	0.365
100	50	50	0.75	5	2	1	110.361	289.639	145.383	-0.617	228.508	146.000	0.365

100	50	50	0.25	5	2	2	124.206	275.794	145.364	-0.636	582.411	146.000	0.365
100	50	50	0.5	5	2	2	111.324	288.676	144.487	-1.513	257.627	146.000	0.365
100	50	50	0.75	5	2	2	108.798	291.202	144.440	-1.560	245.665	146.000	0.365
100	25	75	0.25	5	2	2	122.967	277.033	142.326	-3.674	693.525	146.000	0.365
100	25	75	0.5	5	2	2	110.830	289.170	143.053	-2.947	278.444	146.000	0.365
100	25	75	0.75	5	2	2	107.788	292.212	142.888	-3.112	286.378	146.000	0.365
100	75	25	0.25	5	2	2	124.425	275.575	146.203	0.203	524.876	146.000	0.365
100	75	25	0.5	5	2	2	111.947	288.053	145.023	-0.977	235.125	146.000	0.365
100	75	25	0.75	5	2	2	109.754	290.246	144.949	-1.051	219.971	146.000	0.365
200	100	100	0.25	3	2	1	112.331	287.669	148.367	5.367	912.685	143.000	0.358
200	100	100	0.5	3	2	1	86.082	313.918	143.285	0.285	157.894	143.000	0.358
200	100	100	0.75	3	2	1	73.242	326.758	143.263	0.263	78.417	143.000	0.358
200	100	100	0.25	3	2	2	112.390	287.610	146.105	3.105	876.839	143.000	0.358
200	100	100	0.5	3	2	2	85.803	314.197	143.414	0.414	138.902	143.000	0.358
200	100	100	0.75	3	2	2	71.421	328.579	142.920	-0.080	80.428	143.000	0.358
200	50	150	0.25	3	2	2	111.983	288.017	146.835	3.835	1022.134	143.000	0.358

200	50	150	0.5	3	2	2	84.731	315.269	143.027	0.027	126.740	143.000	0.358
200	50	150	0.75	3	2	2	70.290	329.710	143.038	0.038	80.748	143.000	0.358
200	150	50	0.25	3	2	2	112.167	287.833	145.877	2.877	756.609	143.000	0.358
200	150	50	0.5	3	2	2	85.981	314.019	142.878	-0.122	141.002	143.000	0.358
200	150	50	0.75	3	2	2	72.336	327.664	143.227	0.227	78.675	143.000	0.358
200	100	100	0.25	5	2	1	106.290	293.710	148.074	3.074	602.967	145.000	0.363
200	100	100	0.5	5	2	1	82.106	317.894	145.193	0.193	91.525	145.000	0.363
200	100	100	0.75	5	2	1	77.775	322.225	144.683	-0.317	75.237	145.000	0.363
200	100	100	0.25	5	2	2	105.423	294.577	146.595	1.595	250.471	145.000	0.363
200	100	100	0.5	5	2	2	80.361	319.639	144.457	-0.543	88.251	145.000	0.363
200	100	100	0.75	5	2	2	75.619	324.381	144.468	-0.532	82.780	145.000	0.363
200	50	150	0.25	5	2	2	104.725	295.275	145.969	0.969	252.891	145.000	0.363
200	50	150	0.5	5	2	2	79.389	320.611	144.689	-0.311	83.968	145.000	0.363
200	50	150	0.75	5	2	2	74.514	325.486	144.640	-0.360	80.562	145.000	0.363
200	150	50	0.25	5	2	2	105.434	294.566	146.172	1.172	256.158	145.000	0.363
200	150	50	0.5	5	2	2	81.180	318.820	144.781	-0.219	86.370	145.000	0.363

200	150	50	0.75	5	2	2	76.602	323.398	144.338	-0.662	79.442	145.000	0.363
100	50	50	0.25	3	3	1	121.151	278.849	157.188	12.976	2044.173	144.211	0.373
100	50	50	0.5	3	3	1	108.290	291.710	151.054	6.843	462.077	144.211	0.373
100	50	50	0.75	3	3	1	101.308	298.692	149.554	5.342	244.966	144.211	0.373
100	50	50	0.25	3	3	2	118.334	281.666	155.361	11.150	1593.905	144.211	0.373
100	50	50	0.5	3	3	2	106.114	293.886	149.212	5.000	369.732	144.211	0.373
100	50	50	0.75	3	3	2	98.096	301.904	147.857	3.646	212.632	144.211	0.373
100	25	75	0.25	3	3	2	116.458	283.542	152.960	8.749	1607.953	144.211	0.373
100	25	75	0.5	3	3	2	104.595	295.405	148.170	3.959	354.152	144.211	0.373
100	25	75	0.75	3	3	2	96.273	303.727	147.198	2.987	236.908	144.211	0.373
100	75	25	0.25	3	3	2	120.279	279.721	158.513	14.302	1877.900	144.211	0.373
100	75	25	0.5	3	3	2	107.152	292.848	149.964	5.753	351.897	144.211	0.373
100	75	25	0.75	3	3	2	99.690	300.310	149.013	4.802	225.244	144.211	0.373
100	50	50	0.25	5	3	1	119.784	280.216	154.118	3.878	862.466	150.240	0.373
100	50	50	0.5	5	3	1	107.921	292.079	150.660	0.420	246.166	150.240	0.373
100	50	50	0.75	5	3	1	105.580	294.420	149.062	-1.179	214.183	150.240	0.373

100	50	50	0.25	5	3	2	118.080	281.920	152.491	2.251	675.478	150.240	0.373
100	50	50	0.5	5	3	2	105.301	294.699	148.045	-2.195	230.096	150.240	0.373
100	50	50	0.75	5	3	2	103.023	296.977	149.226	-1.014	219.928	150.240	0.373
100	25	75	0.25	5	3	2	117.387	282.613	149.216	-1.024	671.097	150.240	0.373
100	25	75	0.5	5	3	2	104.048	295.952	147.589	-2.651	271.295	150.240	0.373
100	25	75	0.75	5	3	2	101.547	298.453	148.134	-2.106	252.970	150.240	0.373
100	75	25	0.25	5	3	2	118.863	281.137	152.443	2.203	677.885	150.240	0.373
100	75	25	0.5	5	3	2	106.866	293.134	149.128	-1.112	250.355	150.240	0.373
100	75	25	0.75	5	3	2	104.431	295.569	149.471	-0.769	220.591	150.240	0.373
200	100	100	0.25	3	3	1	102.148	297.852	141.142	-9.541	949.983	150.682	0.332
200	100	100	0.5	3	3	1	78.291	321.709	134.026	-16.656	412.159	150.682	0.332
200	100	100	0.75	3	3	1	66.029	333.971	133.620	-17.063	361.748	150.682	0.332
200	100	100	0.25	3	3	2	99.413	300.587	136.957	-13.726	776.537	150.682	0.332
200	100	100	0.5	3	3	2	73.198	326.802	133.335	-17.347	417.079	150.682	0.332
200	100	100	0.75	3	3	2	58.192	341.808	132.519	-18.163	390.058	150.682	0.332
200	50	150	0.25	3	3	2	98.022	301.978	135.821	-14.862	831.152	150.682	0.332

200	50	150	0.5	3	3	2	70.565	329.435	133.313	-17.370	408.793	150.682	0.332
200	50	150	0.75	3	3	2	54.980	345.020	132.493	-18.189	389.696	150.682	0.332
200	150	50	0.25	3	3	2	100.779	299.221	139.387	-11.296	909.738	150.682	0.332
200	150	50	0.5	3	3	2	75.740	324.260	133.412	-17.270	411.439	150.682	0.332
200	150	50	0.75	3	3	2	62.053	337.947	132.916	-17.767	381.537	150.682	0.332
200	100	100	0.25	5	3	1	102.861	297.139	156.057	0.129	325.133	155.928	0.387
200	100	100	0.5	5	3	1	78.748	321.252	154.831	-1.097	98.137	155.928	0.387
200	100	100	0.75	5	3	1	74.316	325.684	154.737	-1.191	81.937	155.928	0.387
200	100	100	0.25	5	3	2	100.570	299.430	154.829	-1.099	235.785	155.928	0.387
200	100	100	0.5	5	3	2	74.623	325.377	154.511	-1.417	90.037	155.928	0.387
200	100	100	0.75	5	3	2	69.431	330.569	154.500	-1.428	69.763	155.928	0.387
200	50	150	0.25	5	3	2	98.516	301.484	153.602	-2.326	221.242	155.928	0.387
200	50	150	0.5	5	3	2	72.212	327.788	153.411	-2.517	100.899	155.928	0.387
200	50	150	0.75	5	3	2	66.663	333.337	153.259	-2.669	84.461	155.928	0.387
200	150	50	0.25	5	3	2	101.446	298.554	155.216	-0.712	239.744	155.928	0.387
200	150	50	0.5	5	3	2	76.568	323.432	155.175	-0.753	84.006	155.928	0.387

200	150	50	0.75	5	3	2	71.770	328.230	154.869	-1.059	72.971	155.928	0.387

Table B.2. Average value from 1000 simulations of estimates of test statistic measuring lack of fit, $\boldsymbol{d}$, test statistic measuring goodness of fit, $\boldsymbol{GOF}$, estimate of N^{tot}, $\widehat{\boldsymbol{N^{tot}}}$, bias associated in estimating N^{tot}, and mean-squared-error, ***MSE***, associated with estimating N^{tot}. Estimates of coefficients in the logit-linear model for occupancy probability, $\widehat{\boldsymbol{\beta_0}}$ and $\widehat{\boldsymbol{\beta_1}}$, along with average bias and mean-squared-error (***MSE***). $\boldsymbol{N^{tot}}$ represents the true value of the total number of sites occupied out of 400. $\boldsymbol{n_1}$ and $\boldsymbol{n_2}$ represent the sample size at each phase of sampling. $\boldsymbol{n_2} = 0$ represents the traditional single-season occupancy approach. ***Habitat*** represents three different habitats with varying degrees of spatial correlation (1-extreme spatial correlation, 2-moderate spatial correlation, 3-no spatial correlation).

N	n_1	n_2	*p*	*J*	*habitat*	*approach*	*p-hat*	*Bias p*	*MSE p*	β_0	*Bias* β_0	*MSE* β_0	β_1	*Bias* β_1	*MSE* β_1
100	50	50	0.25	3	1	1	0.261	0.011	0.005	-0.734	1.266	3.034	2.016	0.016	6.456
100	50	50	0.5	3	1	1	0.508	0.008	0.004	-0.835	1.165	1.461	1.527	-0.473	0.370
100	50	50	0.75	3	1	1	0.751	0.001	0.002	-0.843	1.157	1.421	1.553	-0.447	0.302
100	50	50	0.25	3	1	2	0.261	0.011	0.004	-1.327	0.673	8.549	2.365	0.365	11.155
100	50	50	0.5	3	1	2	0.502	0.002	0.003	-0.953	1.047	1.431	1.698	-0.302	0.405
100	50	50	0.75	3	1	2	0.750	0.000	0.002	-0.933	1.067	1.287	1.668	-0.332	0.256
100	25	75	0.25	3	1	2	0.254	0.004	0.004	-3.148	-1.148	41.485	4.090	2.090	42.623
100	25	75	0.5	3	1	2	0.503	0.003	0.003	-1.482	0.518	8.417	2.131	0.131	5.417
100	25	75	0.75	3	1	2	0.752	0.002	0.001	-1.496	0.504	9.969	2.113	0.113	6.292
100	75	25	0.25	3	1	2	0.263	0.013	0.005	-0.909	1.091	4.774	2.061	0.061	7.362
100	75	25	0.5	3	1	2	0.502	0.002	0.003	-0.859	1.141	1.418	1.604	-0.396	0.645
100	75	25	0.75	3	1	2	0.749	-0.001	0.002	-0.881	1.119	1.355	1.598	-0.402	0.271
100	50	50	0.25	5	1	1	0.248	-0.002	0.002	-0.608	1.392	2.103	1.542	-0.458	1.604
100	50	50	0.5	5	1	1	0.501	0.001	0.001	-0.659	1.341	1.858	1.381	-0.619	0.460
100	50	50	0.75	5	1	1	0.749	-0.001	0.001	-0.646	1.354	1.888	1.366	-0.634	0.467

100	50	50	0.25	5	1	2	0.249	-0.001	0.002	-0.689	1.311	2.018	1.639	-0.361	2.277
100	50	50	0.5	5	1	2	0.500	0.000	0.001	-0.682	1.318	1.812	1.395	-0.605	0.454
100	50	50	0.75	5	1	2	0.750	0.000	0.001	-0.695	1.305	1.782	1.384	-0.616	0.457
100	25	75	0.25	5	1	2	0.245	-0.005	0.002	-0.894	1.106	4.120	1.971	-0.029	6.895
100	25	75	0.5	5	1	2	0.501	0.001	0.001	-0.744	1.256	1.890	1.431	-0.569	0.554
100	25	75	0.75	5	1	2	0.749	-0.001	0.001	-0.824	1.176	3.352	1.491	-0.509	1.492
100	75	25	0.25	5	1	2	0.248	-0.002	0.002	-0.645	1.355	1.996	1.581	-0.419	1.812
100	75	25	0.5	5	1	2	0.500	0.000	0.001	-0.661	1.339	1.857	1.394	-0.606	0.448
100	75	25	0.75	5	1	2	0.750	0.000	0.001	-0.661	1.339	1.858	1.367	-0.633	0.467
200	100	100	0.25	3	1	1	0.245	-0.005	0.003	-0.630	1.370	2.305	1.870	-0.130	3.279
200	100	100	0.5	3	1	1	0.498	-0.002	0.002	-0.779	1.221	1.524	1.467	-0.533	0.327
200	100	100	0.75	3	1	1	0.749	-0.001	0.001	-0.782	1.218	1.502	1.434	-0.566	0.342
200	100	100	0.25	3	1	2	0.243	-0.007	0.002	-0.667	1.333	2.470	2.207	0.207	8.918
200	100	100	0.5	3	1	2	0.498	-0.002	0.002	-0.777	1.223	1.529	1.422	-0.578	0.373
200	100	100	0.75	3	1	2	0.749	-0.001	0.001	-0.786	1.214	1.497	1.404	-0.596	0.376
200	50	150	0.25	3	1	2	0.241	-0.009	0.002	-0.702	1.298	1.914	2.387	0.387	12.811

200	50	150	0.5	3	1	2	0.499	-0.001	0.002	-0.770	1.230	1.551	1.411	-0.589	0.390
200	50	150	0.75	3	1	2	0.749	-0.001	0.001	-0.765	1.235	1.551	1.383	-0.617	0.400
200	150	50	0.25	3	1	2	0.241	-0.009	0.003	-0.636	1.364	2.313	2.063	0.063	6.953
200	150	50	0.5	3	1	2	0.498	-0.002	0.002	-0.773	1.227	1.541	1.442	-0.558	0.351
200	150	50	0.75	3	1	2	0.748	-0.002	0.001	-0.790	1.210	1.486	1.427	-0.573	0.350
200	100	100	0.25	5	1	1	0.247	-0.003	0.001	-0.558	1.442	2.139	1.249	-0.751	0.620
200	100	100	0.5	5	1	1	0.499	-0.001	0.001	-0.604	1.396	1.967	1.194	-0.806	0.669
200	100	100	0.75	5	1	1	0.750	0.000	0.000	-0.608	1.392	1.951	1.190	-0.810	0.674
200	100	100	0.25	5	1	2	0.247	-0.003	0.001	-0.585	1.415	2.060	1.236	-0.764	0.989
200	100	100	0.5	5	1	2	0.500	0.000	0.001	-0.615	1.385	1.937	1.176	-0.824	0.696
200	100	100	0.75	5	1	2	0.750	0.000	0.000	-0.603	1.397	1.968	1.164	-0.836	0.716
200	50	150	0.25	5	1	2	0.248	-0.002	0.001	-0.595	1.405	2.029	1.325	-0.675	2.645
200	50	150	0.5	5	1	2	0.500	0.000	0.001	-0.600	1.400	1.982	1.160	-0.840	0.726
200	50	150	0.75	5	1	2	0.750	0.000	0.000	-0.605	1.395	1.968	1.158	-0.842	0.726
200	150	50	0.25	5	1	2	0.249	-0.001	0.001	-0.586	1.414	2.054	1.234	-0.766	0.656
200	150	50	0.5	5	1	2	0.501	0.001	0.001	-0.608	1.392	1.957	1.181	-0.819	0.690

200	150	50	0.75	5	1	2	0.750	0.000	0.000	-0.608	1.392	1.957	1.172	-0.828	0.703
100	50	50	0.25	3	2	1	0.251	0.001	0.005	-0.397	1.603	5.441	1.767	-0.233	6.242
100	50	50	0.5	3	2	1	0.501	0.001	0.004	-0.701	1.299	1.764	1.247	-0.753	0.690
100	50	50	0.75	3	2	1	0.748	-0.002	0.002	-0.728	1.272	1.665	1.245	-0.755	0.648
100	50	50	0.25	3	2	2	0.247	-0.003	0.005	-0.514	1.486	5.930	2.142	0.142	10.294
100	50	50	0.5	3	2	2	0.499	-0.001	0.003	-0.747	1.253	1.659	1.306	-0.694	0.623
100	50	50	0.75	3	2	2	0.752	0.002	0.002	-0.748	1.252	1.624	1.247	-0.753	0.644
100	25	75	0.25	3	2	2	0.247	-0.003	0.004	-1.164	0.836	6.062	2.516	0.516	15.251
100	25	75	0.5	3	2	2	0.491	-0.009	0.003	-0.790	1.210	1.613	1.361	-0.639	0.648
100	25	75	0.75	3	2	2	0.751	0.001	0.001	-0.778	1.222	1.576	1.288	-0.712	0.598
100	75	25	0.25	3	2	2	0.251	0.001	0.005	-0.572	1.428	3.413	1.706	-0.294	4.805
100	75	25	0.5	3	2	2	0.498	-0.002	0.004	-0.734	1.266	1.681	1.248	-0.752	0.704
100	75	25	0.75	3	2	2	0.750	0.000	0.002	-0.742	1.258	1.635	1.255	-0.745	0.628
100	50	50	0.25	5	2	1	0.251	0.001	0.002	-0.562	1.438	2.241	1.226	-0.774	0.997
100	50	50	0.5	5	2	1	0.500	0.000	0.002	-0.608	1.392	1.987	1.180	-0.820	0.751
100	50	50	0.75	5	2	1	0.752	0.002	0.001	-0.617	1.383	1.959	1.198	-0.802	0.715

100	50	50	0.25	5	2	2	0.250	0.000	0.002	-0.625	1.375	2.019	1.270	-0.730	1.170
100	50	50	0.5	5	2	2	0.498	-0.002	0.001	-0.637	1.363	1.913	1.239	-0.761	0.665
100	50	50	0.75	5	2	2	0.749	-0.001	0.001	-0.639	1.361	1.907	1.247	-0.753	0.640
100	25	75	0.25	5	2	2	0.250	0.000	0.002	-0.732	1.268	2.221	1.313	-0.687	1.477
100	25	75	0.5	5	2	2	0.503	0.003	0.001	-0.661	1.339	1.858	1.224	-0.776	0.701
100	25	75	0.75	5	2	2	0.751	0.001	0.001	-0.667	1.333	1.846	1.261	-0.739	0.626
100	75	25	0.25	5	2	2	0.252	0.002	0.002	-0.603	1.397	2.062	1.220	-0.780	0.818
100	75	25	0.5	5	2	2	0.501	0.001	0.002	-0.626	1.374	1.939	1.218	-0.782	0.691
100	75	25	0.75	5	2	2	0.751	0.001	0.001	-0.625	1.375	1.936	1.205	-0.795	0.696
200	100	100	0.25	3	2	1	0.249	-0.001	0.003	-0.529	1.471	3.041	1.346	-0.654	0.802
200	100	100	0.5	3	2	1	0.501	0.001	0.002	-0.647	1.353	1.863	1.215	-0.785	0.666
200	100	100	0.75	3	2	1	0.750	0.000	0.001	-0.649	1.351	1.840	1.217	-0.783	0.637
200	100	100	0.25	3	2	2	0.249	-0.001	0.002	-0.475	1.525	7.555	1.356	-0.644	2.242
200	100	100	0.5	3	2	2	0.499	-0.001	0.002	-0.647	1.353	1.858	1.207	-0.793	0.672
200	100	100	0.75	3	2	2	0.751	0.001	0.001	-0.656	1.344	1.824	1.204	-0.796	0.652
200	50	150	0.25	3	2	2	0.247	-0.003	0.002	-0.494	1.506	5.085	1.396	-0.604	1.399

200	50	150	0.5	3	2	2	0.500	0.000	0.002	-0.655	1.345	1.835	1.220	-0.780	0.653
200	50	150	0.75	3	2	2	0.749	-0.001	0.001	-0.655	1.345	1.826	1.204	-0.796	0.653
200	150	50	0.25	3	2	2	0.250	0.000	0.002	-0.585	1.415	2.293	1.333	-0.667	0.718
200	150	50	0.5	3	2	2	0.499	-0.001	0.002	-0.655	1.345	1.838	1.221	-0.779	0.654
200	150	50	0.75	3	2	2	0.750	0.000	0.001	-0.651	1.349	1.837	1.216	-0.784	0.635
200	100	100	0.25	5	2	1	0.249	-0.001	0.001	-0.474	1.526	6.258	0.965	-1.035	1.323
200	100	100	0.5	5	2	1	0.500	0.000	0.001	-0.629	1.371	1.896	0.958	-1.042	1.105
200	100	100	0.75	5	2	1	0.750	0.000	0.001	-0.635	1.365	1.876	0.957	-1.043	1.105
200	100	100	0.25	5	2	2	0.248	-0.002	0.001	-0.614	1.386	1.965	0.979	-1.021	1.114
200	100	100	0.5	5	2	2	0.500	0.000	0.001	-0.640	1.360	1.865	0.961	-1.039	1.097
200	100	100	0.75	5	2	2	0.748	-0.002	0.000	-0.640	1.360	1.865	0.958	-1.042	1.099
200	50	150	0.25	5	2	2	0.250	0.000	0.001	-0.622	1.378	1.942	0.979	-1.021	1.103
200	50	150	0.5	5	2	2	0.500	0.000	0.001	-0.637	1.363	1.874	0.953	-1.047	1.113
200	50	150	0.75	5	2	2	0.748	-0.002	0.001	-0.638	1.362	1.871	0.956	-1.044	1.104
200	150	50	0.25	5	2	2	0.249	-0.001	0.001	-0.618	1.382	1.952	0.981	-1.019	1.092
200	150	50	0.5	5	2	2	0.500	0.000	0.001	-0.634	1.366	1.880	0.952	-1.048	1.117

200	150	50	0.75	5	2	2	0.751	0.001	0.000	-0.641	1.359	1.860	0.965	-1.035	1.088
100	50	50	0.25	3	3	1	0.252	0.002	0.005	0.046	2.046	32.855	3.769	1.769	133.278
100	50	50	0.5	3	3	1	0.498	-0.002	0.004	-0.766	1.234	1.632	1.522	-0.478	0.406
100	50	50	0.75	3	3	1	0.751	0.001	0.002	-0.784	1.216	1.541	1.480	-0.520	0.356
100	50	50	0.25	3	3	2	0.246	-0.004	0.004	-0.743	1.257	34.414	5.166	3.166	233.646
100	50	50	0.5	3	3	2	0.495	-0.005	0.003	-0.819	1.181	1.526	1.543	-0.457	0.454
100	50	50	0.75	3	3	2	0.752	0.002	0.002	-0.822	1.178	1.468	1.488	-0.512	0.349
100	25	75	0.25	3	3	2	0.241	-0.009	0.004	-2.232	-0.232	65.776	6.486	4.486	301.096
100	25	75	0.5	3	3	2	0.495	-0.005	0.003	-0.884	1.116	2.067	1.615	-0.385	1.111
100	25	75	0.75	3	3	2	0.746	-0.004	0.002	-0.852	1.148	1.432	1.524	-0.476	0.340
100	75	25	0.25	3	3	2	0.249	-0.001	0.005	-0.338	1.662	19.262	3.444	1.444	93.541
100	75	25	0.5	3	3	2	0.498	-0.002	0.004	-0.788	1.212	1.564	1.524	-0.476	0.434
100	75	25	0.75	3	3	2	0.748	-0.002	0.002	-0.800	1.200	1.511	1.487	-0.513	0.362
100	50	50	0.25	5	3	1	0.247	-0.003	0.002	-0.589	1.411	2.953	1.684	-0.316	15.912
100	50	50	0.5	5	3	1	0.496	-0.004	0.002	-0.709	1.291	1.724	1.325	-0.675	0.544
100	50	50	0.75	5	3	1	0.752	0.002	0.001	-0.731	1.269	1.664	1.315	-0.685	0.539

100	50	50	0.25	5	3	2	0.247	-0.003	0.002	-0.653	1.347	3.263	1.791	-0.209	17.416
100	50	50	0.5	5	3	2	0.501	0.001	0.001	-0.754	1.246	1.625	1.315	-0.685	0.551
100	50	50	0.75	5	3	2	0.749	-0.001	0.001	-0.729	1.271	1.679	1.294	-0.706	0.563
100	25	75	0.25	5	3	2	0.246	-0.004	0.002	-0.802	1.198	3.610	1.763	-0.237	18.641
100	25	75	0.5	5	3	2	0.500	0.000	0.001	-0.771	1.229	1.610	1.320	-0.680	0.557
100	25	75	0.75	5	3	2	0.750	0.000	0.001	-0.754	1.246	1.643	1.299	-0.701	0.570
100	75	25	0.25	5	3	2	0.248	-0.002	0.002	-0.665	1.335	5.667	1.908	-0.092	60.656
100	75	25	0.5	5	3	2	0.500	0.000	0.002	-0.730	1.270	1.681	1.301	-0.699	0.566
100	75	25	0.75	5	3	2	0.750	0.000	0.001	-0.723	1.277	1.690	1.298	-0.702	0.559
200	100	100	0.25	3	3	1	0.246	-0.004	0.003	-0.773	1.227	8.451	2.046	0.046	50.721
200	100	100	0.5	3	3	1	0.499	-0.001	0.002	-1.007	0.993	1.019	1.463	-0.537	0.335
200	100	100	0.75	3	3	1	0.749	-0.001	0.001	-1.011	0.989	1.000	1.455	-0.545	0.325
200	100	100	0.25	3	3	2	0.247	-0.003	0.002	-0.907	1.093	8.973	1.834	-0.166	26.199
200	100	100	0.5	3	3	2	0.495	-0.005	0.002	-1.027	0.973	0.990	1.471	-0.529	0.329
200	100	100	0.75	3	3	2	0.748	-0.002	0.001	-1.030	0.970	0.969	1.450	-0.550	0.330
200	50	150	0.25	3	3	2	0.247	-0.003	0.002	-1.055	0.945	1.151	1.710	-0.290	2.095

200	50	150	0.5	3	3	2	0.499	-0.001	0.002	-1.026	0.974	0.995	1.462	-0.538	0.341
200	50	150	0.75	3	3	2	0.749	-0.001	0.001	-1.026	0.974	0.979	1.435	-0.565	0.348
200	150	50	0.25	3	3	2	0.247	-0.003	0.003	-0.919	1.081	1.836	1.866	-0.134	11.036
200	150	50	0.5	3	3	2	0.498	-0.002	0.002	-1.024	0.976	0.989	1.474	-0.526	0.326
200	150	50	0.75	3	3	2	0.750	0.000	0.001	-1.019	0.981	0.987	1.442	-0.558	0.335
200	100	100	0.25	5	3	1	0.251	0.001	0.001	-0.579	1.421	2.517	1.147	-0.853	0.864
200	100	100	0.5	5	3	1	0.501	0.001	0.001	-0.614	1.386	1.940	1.163	-0.837	0.726
200	100	100	0.75	5	3	1	0.751	0.001	0.000	-0.614	1.386	1.938	1.157	-0.843	0.732
200	100	100	0.25	5	3	2	0.249	-0.001	0.001	-0.640	1.360	1.904	1.242	-0.758	0.667
200	100	100	0.5	5	3	2	0.500	0.000	0.001	-0.635	1.365	1.889	1.211	-0.789	0.656
200	100	100	0.75	5	3	2	0.750	0.000	0.000	-0.632	1.368	1.892	1.203	-0.797	0.662
200	50	150	0.25	5	3	2	0.249	-0.001	0.001	-0.688	1.312	1.786	1.332	-0.668	0.592
200	50	150	0.5	5	3	2	0.499	-0.001	0.001	-0.665	1.335	1.815	1.255	-0.745	0.598
200	50	150	0.75	5	3	2	0.750	0.000	0.000	-0.662	1.338	1.817	1.242	-0.758	0.609
200	150	50	0.25	5	3	2	0.251	0.001	0.001	-0.622	1.378	1.945	1.210	-0.790	0.703
200	150	50	0.5	5	3	2	0.500	0.000	0.001	-0.617	1.383	1.933	1.185	-0.815	0.692

200	150	50	0.75	5	3	2	0.749	-0.001	0.000	-0.621	1.379	1.921	1.184	-0.816	0.691

Table B.3. Average value from 1000 simulations of estimates of test statistic measuring lack of fit, $\boldsymbol{d}$, test statistic measuring goodness of fit, $\boldsymbol{GOF}$, estimate of N^{tot}, $\widehat{\boldsymbol{N^{tot}}}$, bias associated in estimating N^{tot}, and mean-squared-error, $\boldsymbol{MSE}$, associated with estimating N^{tot}. Estimates of coefficients in the logit-linear model for occupancy probability, $\widehat{\boldsymbol{\beta_0}}$ and $\widehat{\boldsymbol{\beta_1}}$, along with average bias and mean-squared-error (***MSE***). $\boldsymbol{N^{tot}}$ represents the true value of the total number of sites occupied out of 400. $\boldsymbol{n_1}$ and $\boldsymbol{n_2}$ represent the sample size at each phase of sampling. $\boldsymbol{n_2} = 0$ represents the traditional single-season occupancy approach. ***Habitat*** represents three different habitats with varying degrees of spatial correlation (1-extreme spatial correlation, 2-moderate spatial correlation, 3-no spatial correlation) and no covariate relationship (see text for more details).

N	n_1	n_2	*p*	*J*	*habitat*	*approach*	*d*	*GOF*	N^{tot}*-hat*	*Bias* N^{tot}	*MSE* N^{tot}	N^{tot}	ψ
100	50	50	0.25	3	1.0	1	173.647	226.353	171.496	20.496	4100.754	151.000	0.377
100	50	50	0.5	3	1.0	1	151.635	248.365	153.869	2.869	511.722	151.000	0.377
100	50	50	0.75	3	1.0	1	142.912	257.088	151.263	0.263	297.576	151.000	0.377
100	50	50	0.25	3	1.0	2	173.258	226.742	169.485	18.485	4320.183	151.000	0.377
100	50	50	0.5	3	1.0	2	150.988	249.012	152.340	1.340	520.729	151.000	0.377
100	50	50	0.75	3	1.0	2	142.689	257.311	150.147	-0.853	309.433	151.000	0.377
100	25	75	0.25	3	1.0	2	171.093	228.907	160.777	9.777	3741.166	151.000	0.377
100	25	75	0.5	3	1.0	2	150.700	249.300	150.719	-0.281	521.817	151.000	0.377
100	25	75	0.75	3	1.0	2	142.419	257.581	148.686	-2.314	346.604	151.000	0.377
100	75	25	0.25	3	1.0	2	173.452	226.548	170.121	19.121	4477.118	151.000	0.377
100	75	25	0.5	3	1.0	2	151.167	248.833	152.930	1.930	482.336	151.000	0.377
100	75	25	0.75	3	1.0	2	142.660	257.340	149.940	-1.060	287.618	151.000	0.377
100	50	50	0.25	5	1.0	1	152.907	247.093	144.541	11.541	871.365	132.999	0.353
100	50	50	0.5	5	1.0	1	139.197	260.803	141.140	8.141	378.050	132.999	0.353
100	50	50	0.75	5	1.0	1	136.893	263.107	140.349	7.349	332.978	132.999	0.353

100	50	50	0.25	5	1.0	2	152.801	247.199	142.608	9.609	905.640	132.999	0.353
100	50	50	0.5	5	1.0	2	138.860	261.140	140.624	7.624	369.467	132.999	0.353
100	50	50	0.75	5	1.0	2	136.585	263.415	140.072	7.073	348.762	132.999	0.353
100	25	75	0.25	5	1.0	2	151.458	248.542	138.349	5.349	920.239	132.999	0.353
100	25	75	0.5	5	1.0	2	138.334	261.666	138.442	5.442	385.228	132.999	0.353
100	25	75	0.75	5	1.0	2	136.095	263.905	138.568	5.569	331.587	132.999	0.353
100	75	25	0.25	5	1.0	2	153.245	246.755	144.303	11.303	949.438	132.999	0.353
100	75	25	0.5	5	1.0	2	138.931	261.069	140.250	7.250	345.215	132.999	0.353
100	75	25	0.75	5	1.0	2	136.824	263.176	141.084	8.085	328.818	132.999	0.353
200	100	100	0.25	3	1.0	1	126.371	273.629	120.353	-12.612	2161.577	132.965	0.275
200	100	100	0.5	3	1.0	1	93.799	306.201	111.330	-21.635	605.068	132.965	0.275
200	100	100	0.75	3	1.0	1	81.899	318.101	110.258	-22.707	593.260	132.965	0.275
200	100	100	0.25	3	1.0	2	126.821	273.179	120.851	-12.115	2269.514	132.965	0.275
200	100	100	0.5	3	1.0	2	93.598	306.402	110.050	-22.916	671.588	132.965	0.275
200	100	100	0.75	3	1.0	2	81.652	318.348	109.756	-23.209	623.566	132.965	0.275
200	50	150	0.25	3	1.0	2	126.773	273.227	119.918	-13.048	2756.791	132.965	0.275

200	50	150	0.5	3	1.0	2	93.388	306.612	109.641	-23.324	690.231	132.965	0.275
200	50	150	0.75	3	1.0	2	81.482	318.518	108.958	-24.007	664.249	132.965	0.275
200	150	50	0.25	3	1.0	2	128.047	271.953	123.206	-9.759	2518.567	132.965	0.275
200	150	50	0.5	3	1.0	2	93.743	306.257	111.255	-21.710	602.513	132.965	0.275
200	150	50	0.75	3	1.0	2	81.774	318.226	109.934	-23.032	618.619	132.965	0.275
200	100	100	0.25	5	1.0	1	127.496	272.504	154.200	10.517	1750.531	143.683	0.365
200	100	100	0.5	5	1.0	1	97.732	302.268	145.767	2.084	111.058	143.683	0.365
200	100	100	0.75	5	1.0	1	93.288	306.712	146.672	2.989	102.261	143.683	0.365
200	100	100	0.25	5	1.0	2	126.583	273.417	151.418	7.735	1369.161	143.683	0.365
200	100	100	0.5	5	1.0	2	97.617	302.383	145.654	1.971	108.440	143.683	0.365
200	100	100	0.75	5	1.0	2	93.206	306.794	145.304	1.621	101.471	143.683	0.365
200	50	150	0.25	5	1.0	2	126.845	273.155	151.821	8.139	1580.684	143.683	0.365
200	50	150	0.5	5	1.0	2	97.613	302.387	145.200	1.517	111.181	143.683	0.365
200	50	150	0.75	5	1.0	2	93.118	306.882	145.238	1.555	94.843	143.683	0.365
200	150	50	0.25	5	1.0	2	126.308	273.692	150.343	6.660	1298.626	143.683	0.365
200	150	50	0.5	5	1.0	2	97.656	302.344	145.596	1.913	113.588	143.683	0.365

200	150	50	0.75	5	1.0	2	93.216	306.784	146.171	2.489	99.247	143.683	0.365
100	50	50	0.25	3	2.0	1	168.306	231.694	158.241	16.241	3694.796	142.000	0.355
100	50	50	0.5	3	2.0	1	146.744	253.256	142.716	0.716	445.371	142.000	0.355
100	50	50	0.75	3	2.0	1	139.037	260.963	141.511	-0.489	282.792	142.000	0.355
100	50	50	0.25	3	2.0	2	169.045	230.955	159.441	17.441	4092.388	142.000	0.355
100	50	50	0.5	3	2.0	2	146.688	253.312	142.481	0.481	437.392	142.000	0.355
100	50	50	0.75	3	2.0	2	138.722	261.278	140.918	-1.082	278.161	142.000	0.355
100	25	75	0.25	3	2.0	2	165.787	234.213	149.616	7.616	3554.107	142.000	0.355
100	25	75	0.5	3	2.0	2	145.867	254.133	140.730	-1.270	506.723	142.000	0.355
100	25	75	0.75	3	2.0	2	138.273	261.727	139.499	-2.501	341.638	142.000	0.355
100	75	25	0.25	3	2.0	2	168.462	231.538	158.115	16.115	3748.338	142.000	0.355
100	75	25	0.5	3	2.0	2	146.867	253.133	142.660	0.660	488.935	142.000	0.355
100	75	25	0.75	3	2.0	2	138.946	261.054	141.669	-0.331	282.767	142.000	0.355
100	50	50	0.25	5	2.0	1	150.469	249.531	138.399	4.399	795.483	134.000	0.335
100	50	50	0.5	5	2.0	1	136.647	263.353	134.749	0.749	293.774	134.000	0.335
100	50	50	0.75	5	2.0	1	134.380	265.620	133.944	-0.056	267.502	134.000	0.335

100	50	50	0.25	5	2.0	2	150.217	249.783	137.508	3.508	739.186	134.000	0.335
100	50	50	0.5	5	2.0	2	136.203	263.797	133.170	-0.830	313.749	134.000	0.335
100	50	50	0.75	5	2.0	2	134.021	265.979	132.813	-1.187	289.727	134.000	0.335
100	25	75	0.25	5	2.0	2	148.801	251.199	132.808	-1.192	900.676	134.000	0.335
100	25	75	0.5	5	2.0	2	135.582	264.418	131.255	-2.745	359.613	134.000	0.335
100	25	75	0.75	5	2.0	2	133.684	266.316	131.908	-2.092	314.297	134.000	0.335
100	75	25	0.25	5	2.0	2	150.075	249.925	136.944	2.944	755.416	134.000	0.335
100	75	25	0.5	5	2.0	2	136.321	263.679	133.434	-0.566	294.786	134.000	0.335
100	75	25	0.75	5	2.0	2	134.283	265.717	133.964	-0.036	267.374	134.000	0.335
200	100	100	0.25	3	2.0	1	146.570	253.430	156.098	15.098	3328.519	141.000	0.353
200	100	100	0.5	3	2.0	1	108.265	291.735	141.711	0.711	167.412	141.000	0.353
200	100	100	0.75	3	2.0	1	93.559	306.441	140.483	-0.517	95.031	141.000	0.353
200	100	100	0.25	3	2.0	2	146.463	253.537	154.713	13.713	3210.288	141.000	0.353
200	100	100	0.5	3	2.0	2	108.095	291.905	140.772	-0.228	174.089	141.000	0.353
200	100	100	0.75	3	2.0	2	93.437	306.563	140.972	-0.028	95.136	141.000	0.353
200	50	150	0.25	3	2.0	2	146.986	253.014	154.637	13.637	3803.403	141.000	0.353

200	50	150	0.5	3	2.0	2	107.799	292.201	139.855	-1.145	167.802	141.000	0.353
200	50	150	0.75	3	2.0	2	93.311	306.689	139.623	-1.377	104.729	141.000	0.353
200	150	50	0.25	3	2.0	2	145.775	254.225	153.619	12.619	2767.346	141.000	0.353
200	150	50	0.5	3	2.0	2	108.174	291.826	142.202	1.202	161.768	141.000	0.353
200	150	50	0.75	3	2.0	2	93.479	306.521	140.650	-0.350	97.996	141.000	0.353
200	100	100	0.25	5	2.0	1	126.416	273.584	151.649	6.649	1226.644	145.000	0.363
200	100	100	0.5	5	2.0	1	97.424	302.576	145.332	0.332	97.852	145.000	0.363
200	100	100	0.75	5	2.0	1	93.033	306.967	145.318	0.318	97.997	145.000	0.363
200	100	100	0.25	5	2.0	2	125.740	274.260	150.280	5.280	1140.624	145.000	0.363
200	100	100	0.5	5	2.0	2	97.280	302.720	144.364	-0.636	107.606	145.000	0.363
200	100	100	0.75	5	2.0	2	92.993	307.007	144.474	-0.526	94.817	145.000	0.363
200	50	150	0.25	5	2.0	2	124.937	275.063	147.807	2.807	958.867	145.000	0.363
200	50	150	0.5	5	2.0	2	97.338	302.662	144.210	-0.790	110.013	145.000	0.363
200	50	150	0.75	5	2.0	2	92.931	307.069	144.220	-0.780	98.071	145.000	0.363
200	150	50	0.25	5	2.0	2	125.924	274.076	150.005	5.005	1023.003	145.000	0.363
200	150	50	0.5	5	2.0	2	97.329	302.671	143.972	-1.028	105.258	145.000	0.363

200	150	50	0.75	5	2.0	2	92.996	307.004	144.905	-0.095	96.247	145.000	0.363
100	50	50	0.25	3	3.0	1	161.176	238.824	144.092	11.778	3339.318	132.314	0.322
100	50	50	0.5	3	3.0	1	140.275	259.725	130.095	-2.219	410.806	132.314	0.322
100	50	50	0.75	3	3.0	1	132.798	267.202	128.696	-3.618	306.794	132.314	0.322
100	50	50	0.25	3	3.0	2	162.118	237.882	145.485	13.171	3859.154	132.314	0.322
100	50	50	0.5	3	3.0	2	139.678	260.322	129.568	-2.746	429.972	132.314	0.322
100	50	50	0.75	3	3.0	2	132.382	267.618	127.806	-4.508	318.139	132.314	0.322
100	25	75	0.25	3	3.0	2	158.419	241.581	135.490	3.176	3309.773	132.314	0.322
100	25	75	0.5	3	3.0	2	138.835	261.165	126.585	-5.729	585.933	132.314	0.322
100	25	75	0.75	3	3.0	2	131.860	268.140	125.980	-6.334	420.318	132.314	0.322
100	75	25	0.25	3	3.0	2	162.406	237.594	146.539	14.225	3677.485	132.314	0.322
100	75	25	0.5	3	3.0	2	139.842	260.158	129.264	-3.050	466.062	132.314	0.322
100	75	25	0.75	3	3.0	2	132.581	267.419	128.182	-4.132	287.977	132.314	0.322
100	50	50	0.25	5	3.0	1	151.679	248.321	142.720	-0.646	701.384	143.366	0.350
100	50	50	0.5	5	3.0	1	138.162	261.838	141.347	-2.019	278.382	143.366	0.350
100	50	50	0.75	5	3.0	1	135.807	264.193	140.170	-3.196	284.499	143.366	0.350

100	50	50	0.25	5	3.0	2	151.754	248.246	143.227	-0.139	784.634	143.366	0.350
100	50	50	0.5	5	3.0	2	137.459	262.541	139.133	-4.234	325.299	143.366	0.350
100	50	50	0.75	5	3.0	2	135.309	264.691	139.487	-3.879	284.605	143.366	0.350
100	25	75	0.25	5	3.0	2	150.350	249.650	138.791	-4.575	941.566	143.366	0.350
100	25	75	0.5	5	3.0	2	136.804	263.196	136.847	-12.793	500.437	143.366	0.350
100	25	75	0.75	5	3.0	2	134.713	265.287	137.634	-5.733	353.925	143.366	0.350
100	75	25	0.25	5	3.0	2	151.568	248.432	142.271	-1.095	726.533	143.366	0.350
100	75	25	0.5	5	3.0	2	137.807	262.193	139.887	-3.480	309.863	143.366	0.350
100	75	25	0.75	5	3.0	2	135.365	264.635	139.077	-4.289	291.694	143.366	0.350
200	100	100	0.25	3	3.0	1	145.426	254.574	152.337	18.672	3185.668	133.665	0.348
200	100	100	0.5	3	3.0	1	107.895	292.105	139.596	5.930	196.017	133.665	0.348
200	100	100	0.75	3	3.0	1	93.367	306.633	139.200	5.535	132.118	133.665	0.348
200	100	100	0.25	3	3.0	2	144.354	255.646	149.901	16.236	3041.162	133.665	0.348
200	100	100	0.5	3	3.0	2	107.753	292.247	139.037	5.371	200.129	133.665	0.348
200	100	100	0.75	3	3.0	2	93.184	306.816	138.349	4.683	120.024	133.665	0.348
200	50	150	0.25	3	3.0	2	145.292	254.708	151.048	17.383	3585.795	133.665	0.348

200	50	150	0.5	3	3.0	2	107.747	292.253	139.163	5.498	210.792	133.665	0.348
200	50	150	0.75	3	3.0	2	93.020	306.980	138.129	4.464	116.175	133.665	0.348
200	150	50	0.25	3	3.0	2	146.653	253.347	154.758	21.093	3620.127	133.665	0.348
200	150	50	0.5	3	3.0	2	107.681	292.319	139.452	5.787	202.371	133.665	0.348
200	150	50	0.75	3	3.0	2	93.237	306.763	139.127	5.462	119.198	133.665	0.348
200	100	100	0.25	5	3.0	1	123.301	276.699	144.131	6.891	2439.896	137.241	0.335
200	100	100	0.5	5	3.0	1	93.547	306.453	133.808	-3.433	115.269	137.241	0.335
200	100	100	0.75	5	3.0	1	89.502	310.498	133.711	-3.530	104.479	137.241	0.335
200	100	100	0.25	5	3.0	2	121.267	278.733	139.613	2.373	1451.576	137.241	0.335
200	100	100	0.5	5	3.0	2	93.381	306.619	133.185	-4.055	116.491	137.241	0.335
200	100	100	0.75	5	3.0	2	89.323	310.677	133.166	-4.075	114.971	137.241	0.335
200	50	150	0.25	5	3.0	2	121.596	278.404	139.991	2.750	1888.622	137.241	0.335
200	50	150	0.5	5	3.0	2	93.205	306.795	132.615	-4.626	127.266	137.241	0.335
200	50	150	0.75	5	3.0	2	89.133	310.867	132.768	-4.472	119.428	137.241	0.335
200	150	50	0.25	5	3.0	2	121.883	278.117	141.391	4.150	1952.309	137.241	0.335
200	150	50	0.5	5	3.0	2	93.451	306.549	133.959	-3.282	110.097	137.241	0.335

200	150	50	0.75	5	3.0	2	89.388	310.612	133.599	-3.642	109.709	137.241	0.335

Table B.4. Average value from 1000 simulations of estimates of test statistic measuring lack of fit, $\boldsymbol{d}$, test statistic measuring goodness of fit, $\boldsymbol{GOF}$, estimate of N^{tot}, $\widehat{\boldsymbol{N^{tot}}}$, bias associated in estimating N^{tot}, and mean-squared-error, ***MSE***, associated with estimating N^{tot}. Estimates of coefficients in the logit-linear model for occupancy probability, $\widehat{\boldsymbol{\beta_0}}$ and $\widehat{\boldsymbol{\beta_1}}$**,** along with average bias and mean-squared-error (***MSE***). $\boldsymbol{N^{tot}}$ represents the true value of the total number of sites occupied out of 400. $\boldsymbol{n_1}$ and $\boldsymbol{n_2}$ represent the sample size at each phase of sampling. $\boldsymbol{n_2} = 0$ represents the traditional single-season occupancy approach. ***Habitat*** represents three different habitats with varying degrees of spatial correlation (1-extreme spatial correlation, 2-moderate spatial correlation, 3-no spatial correlation) and no covariate relationship (see text for more details).

N	n_1	n_2	p	J	*habitat*	*approach*	*p-hat*	*Bias p*	*MSE p*	β_0	*Bias* β_0	*MSE* β_0	β_1	*Bias* β_1	*MSE* β_1
100	50	50	0.25	3	1.0	1	0.244	-0.006	0.006	0.127	2.127	12.497	0.231	0.231	5.091
100	50	50	0.5	3	1.0	1	0.492	-0.008	0.004	-0.482	1.518	2.363	0.057	0.057	0.051
100	50	50	0.75	3	1.0	1	0.748	-0.002	0.002	-0.506	1.494	2.266	0.066	0.066	0.038
100	50	50	0.25	3	1.0	2	0.243	-0.007	0.006	0.326	2.326	19.613	0.488	0.488	11.968
100	50	50	0.5	3	1.0	2	0.497	-0.003	0.004	-0.499	1.501	2.315	0.055	0.055	0.048
100	50	50	0.75	3	1.0	2	0.747	-0.003	0.002	-0.518	1.482	2.232	0.047	0.047	0.035
100	25	75	0.25	3	1.0	2	0.247	-0.003	0.006	-0.357	1.643	17.798	0.514	0.514	18.892
100	25	75	0.5	3	1.0	2	0.497	-0.003	0.004	-0.522	1.478	2.268	0.047	0.047	0.064
100	25	75	0.75	3	1.0	2	0.750	0.000	0.002	-0.537	1.463	2.190	0.056	0.056	0.045
100	75	25	0.25	3	1.0	2	0.244	-0.006	0.006	0.292	2.292	17.647	0.222	0.222	8.783
100	75	25	0.5	3	1.0	2	0.497	-0.003	0.004	-0.491	1.509	2.333	0.047	0.047	0.044
100	75	25	0.75	3	1.0	2	0.748	-0.002	0.002	-0.521	1.479	2.223	0.054	0.054	0.037
100	50	50	0.25	5	1.0	1	0.252	0.002	0.002	-0.591	1.409	2.100	-0.172	-0.172	0.105
100	50	50	0.5	5	1.0	1	0.501	0.001	0.002	-0.622	1.378	1.939	-0.157	-0.157	0.066
100	50	50	0.75	5	1.0	1	0.749	-0.001	0.001	-0.630	1.370	1.912	-0.164	-0.164	0.067

100	50	50	0.25	5	1.0	2	0.247	-0.003	0.002	-0.617	1.383	2.024	-0.164	-0.164	0.116
100	50	50	0.5	5	1.0	2	0.499	-0.001	0.002	-0.629	1.371	1.922	-0.161	-0.161	0.068
100	50	50	0.75	5	1.0	2	0.750	0.000	0.001	-0.635	1.365	1.903	-0.165	-0.165	0.069
100	25	75	0.25	5	1.0	2	0.250	0.000	0.002	-0.712	1.288	2.223	-0.177	-0.177	0.766
100	25	75	0.5	5	1.0	2	0.498	-0.002	0.002	-0.668	1.332	1.974	-0.157	-0.157	0.173
100	25	75	0.75	5	1.0	2	0.749	-0.001	0.001	-0.655	1.345	1.854	-0.173	-0.173	0.080
100	75	25	0.25	5	1.0	2	0.247	-0.003	0.002	-0.572	1.428	2.832	-0.142	-0.142	0.529
100	75	25	0.5	5	1.0	2	0.501	0.001	0.002	-0.632	1.368	1.910	-0.158	-0.158	0.066
100	75	25	0.75	5	1.0	2	0.750	0.000	0.001	-0.621	1.379	1.935	-0.160	-0.160	0.061
200	100	100	0.25	3	1.0	1	0.248	-0.002	0.005	-0.568	1.432	10.915	0.340	0.340	5.494
200	100	100	0.5	3	1.0	1	0.499	-0.001	0.003	-0.963	1.037	1.097	0.074	0.074	0.024
200	100	100	0.75	3	1.0	1	0.749	-0.001	0.001	-0.973	1.027	1.067	0.066	0.066	0.018
200	100	100	0.25	3	1.0	2	0.246	-0.004	0.005	-0.593	1.407	9.546	0.306	0.306	4.793
200	100	100	0.5	3	1.0	2	0.499	-0.001	0.003	-0.979	1.021	1.065	0.063	0.063	0.024
200	100	100	0.75	3	1.0	2	0.750	0.000	0.001	-0.980	1.020	1.055	0.065	0.065	0.019
200	50	150	0.25	3	1.0	2	0.245	-0.005	0.005	-0.605	1.395	8.854	0.331	0.331	5.145

200	50	150	0.5	3	1.0	2	0.498	-0.002	0.003	-0.986	1.014	1.053	0.072	0.072	0.029
200	50	150	0.75	3	1.0	2	0.750	0.000	0.001	-0.991	1.009	1.034	0.073	0.073	0.021
200	150	50	0.25	3	1.0	2	0.244	-0.006	0.005	-0.436	1.564	12.527	0.391	0.391	6.439
200	150	50	0.5	3	1.0	2	0.497	-0.003	0.003	-0.963	1.037	1.096	0.065	0.065	0.022
200	150	50	0.75	3	1.0	2	0.749	-0.001	0.001	-0.978	1.022	1.060	0.067	0.067	0.018
200	100	100	0.25	5	1.0	1	0.245	-0.005	0.002	0.040	2.040	18.584	0.424	0.424	6.960
200	100	100	0.5	5	1.0	1	0.499	-0.001	0.001	-0.560	1.440	2.086	0.047	0.047	0.015
200	100	100	0.75	5	1.0	1	0.751	0.001	0.001	-0.550	1.450	2.114	0.042	0.042	0.014
200	100	100	0.25	5	1.0	2	0.246	-0.004	0.001	-0.129	1.871	13.880	0.291	0.291	4.145
200	100	100	0.5	5	1.0	2	0.500	0.000	0.001	-0.561	1.439	2.082	0.051	0.051	0.016
200	100	100	0.75	5	1.0	2	0.751	0.001	0.001	-0.565	1.435	2.071	0.047	0.047	0.015
200	50	150	0.25	5	1.0	2	0.247	-0.003	0.002	-0.067	1.933	15.292	0.336	0.336	5.694
200	50	150	0.5	5	1.0	2	0.498	-0.002	0.001	-0.567	1.433	2.068	0.046	0.046	0.018
200	50	150	0.75	5	1.0	2	0.750	0.000	0.001	-0.566	1.434	2.068	0.047	0.047	0.017
200	150	50	0.25	5	1.0	2	0.247	-0.003	0.001	-0.148	1.852	14.572	0.283	0.283	4.132
200	150	50	0.5	5	1.0	2	0.502	0.002	0.001	-0.562	1.438	2.081	0.047	0.047	0.015

200	150	50	0.75	5	1.0	2	0.750	0.000	0.001	-0.555	1.445	2.098	0.048	0.048	0.014
100	50	50	0.25	3	2.0	1	0.248	-0.002	0.006	-0.135	1.865	10.247	-0.301	-0.301	4.193
100	50	50	0.5	3	2.0	1	0.497	-0.003	0.004	-0.605	1.395	2.001	-0.080	-0.080	0.055
100	50	50	0.75	3	2.0	1	0.748	-0.002	0.002	-0.614	1.386	1.956	-0.064	-0.064	0.043
100	50	50	0.25	3	2.0	2	0.244	-0.006	0.007	-0.051	1.949	11.511	-0.344	-0.344	7.507
100	50	50	0.5	3	2.0	2	0.495	-0.005	0.004	-0.613	1.387	1.982	-0.101	-0.101	0.081
100	50	50	0.75	3	2.0	2	0.748	-0.002	0.002	-0.625	1.375	1.928	-0.094	-0.094	0.065
100	25	75	0.25	3	2.0	2	0.246	-0.004	0.007	-0.647	1.353	10.893	-0.744	-0.744	12.221
100	25	75	0.5	3	2.0	2	0.497	-0.003	0.004	-0.686	1.314	2.390	-0.175	-0.175	0.698
100	25	75	0.75	3	2.0	2	0.748	-0.002	0.002	-0.650	1.350	1.889	-0.117	-0.117	0.101
100	75	25	0.25	3	2.0	2	0.245	-0.005	0.006	-0.246	1.754	8.303	-0.319	-0.319	3.771
100	75	25	0.5	3	2.0	2	0.494	-0.006	0.004	-0.608	1.392	1.999	-0.087	-0.087	0.064
100	75	25	0.75	3	2.0	2	0.748	-0.002	0.002	-0.615	1.385	1.954	-0.090	-0.090	0.058
100	50	50	0.25	5	2.0	1	0.248	-0.002	0.002	-0.660	1.340	1.911	-0.051	-0.051	0.101
100	50	50	0.5	5	2.0	1	0.500	0.000	0.002	-0.691	1.309	1.753	-0.034	-0.034	0.045
100	50	50	0.75	5	2.0	1	0.751	0.001	0.001	-0.699	1.301	1.728	-0.021	-0.021	0.042

100	50	50	0.25	5	2.0	2	0.246	-0.004	0.002	-0.673	1.327	1.904	-0.068	-0.068	0.316
100	50	50	0.5	5	2.0	2	0.497	-0.003	0.002	-0.711	1.289	1.705	-0.042	-0.042	0.049
100	50	50	0.75	5	2.0	2	0.747	-0.003	0.001	-0.714	1.286	1.693	-0.051	-0.051	0.049
100	25	75	0.25	5	2.0	2	0.249	-0.001	0.003	-0.801	1.199	2.269	-0.168	-0.168	1.068
100	25	75	0.5	5	2.0	2	0.500	0.000	0.002	-0.805	1.195	2.370	-0.150	-0.150	1.219
100	25	75	0.75	5	2.0	2	0.750	0.000	0.001	-0.745	1.255	1.877	-0.086	-0.086	0.403
100	75	25	0.25	5	2.0	2	0.249	-0.001	0.003	-0.678	1.322	1.854	-0.040	-0.040	0.086
100	75	25	0.5	5	2.0	2	0.500	0.000	0.002	-0.706	1.294	1.713	-0.042	-0.042	0.047
100	75	25	0.75	5	2.0	2	0.749	-0.001	0.001	-0.699	1.301	1.728	-0.034	-0.034	0.041
200	100	100	0.25	3	2.0	1	0.245	-0.005	0.004	-0.135	1.865	11.085	0.209	0.209	1.856
200	100	100	0.5	3	2.0	1	0.500	0.000	0.002	-0.607	1.393	1.960	0.137	0.137	0.035
200	100	100	0.75	3	2.0	1	0.751	0.001	0.001	-0.620	1.380	1.917	0.147	0.147	0.035
200	100	100	0.25	3	2.0	2	0.243	-0.007	0.004	-0.202	1.798	8.559	0.211	0.211	1.660
200	100	100	0.5	3	2.0	2	0.500	0.000	0.002	-0.619	1.381	1.929	0.137	0.137	0.041
200	100	100	0.75	3	2.0	2	0.747	-0.003	0.001	-0.615	1.385	1.931	0.145	0.145	0.035
200	50	150	0.25	3	2.0	2	0.243	-0.007	0.004	-0.027	1.973	15.574	0.059	0.059	2.954

200	50	150	0.5	3	2.0	2	0.501	0.001	0.002	-0.630	1.370	1.900	0.128	0.128	0.044
200	50	150	0.75	3	2.0	2	0.751	0.001	0.001	-0.630	1.370	1.890	0.133	0.133	0.036
200	150	50	0.25	3	2.0	2	0.244	-0.006	0.004	-0.275	1.725	6.583	0.232	0.232	1.187
200	150	50	0.5	3	2.0	2	0.499	-0.001	0.002	-0.602	1.398	1.974	0.138	0.138	0.037
200	150	50	0.75	3	2.0	2	0.750	0.000	0.001	-0.618	1.382	1.923	0.139	0.139	0.033
200	100	100	0.25	5	2.0	1	0.246	-0.004	0.002	-0.242	1.758	10.145	0.062	0.062	1.747
200	100	100	0.5	5	2.0	1	0.500	0.000	0.001	-0.564	1.436	2.073	0.013	0.013	0.012
200	100	100	0.75	5	2.0	1	0.750	0.000	0.001	-0.564	1.436	2.074	0.018	0.018	0.010
200	100	100	0.25	5	2.0	2	0.248	-0.002	0.002	-0.277	1.723	9.495	-0.008	-0.008	1.631
200	100	100	0.5	5	2.0	2	0.502	0.002	0.001	-0.575	1.425	2.043	0.014	0.014	0.013
200	100	100	0.75	5	2.0	2	0.750	0.000	0.000	-0.573	1.427	2.046	0.005	0.005	0.012
200	50	150	0.25	5	2.0	2	0.249	-0.001	0.001	-0.337	1.663	9.039	-0.012	-0.012	1.733
200	50	150	0.5	5	2.0	2	0.501	0.001	0.001	-0.577	1.423	2.039	0.005	0.005	0.013
200	50	150	0.75	5	2.0	2	0.749	-0.001	0.001	-0.576	1.424	2.039	0.013	0.013	0.011
200	150	50	0.25	5	2.0	2	0.246	-0.004	0.001	-0.307	1.693	8.203	0.092	0.092	1.817
200	150	50	0.5	5	2.0	2	0.502	0.002	0.001	-0.579	1.421	2.031	0.009	0.009	0.013

200	150	50	0.75	5	2.0	2	0.749	-0.001	0.001	-0.569	1.431	2.061	0.011	0.011	0.010
100	50	50	0.25	3	3.0	1	0.248	-0.002	0.007	0.005	2.005	37.701	0.133	0.133	22.940
100	50	50	0.5	3	3.0	1	0.495	-0.005	0.005	-0.750	1.250	1.619	0.017	0.017	0.060
100	50	50	0.75	3	3.0	1	0.747	-0.003	0.002	-0.760	1.240	1.577	0.023	0.023	0.042
100	50	50	0.25	3	3.0	2	0.243	-0.007	0.008	0.005	2.005	37.723	0.113	0.113	38.985
100	50	50	0.5	3	3.0	2	0.499	-0.001	0.004	-0.757	1.243	1.605	0.020	0.020	0.061
100	50	50	0.75	3	3.0	2	0.750	0.000	0.002	-0.772	1.228	1.550	0.013	0.013	0.046
100	25	75	0.25	3	3.0	2	0.245	-0.005	0.007	-0.618	1.382	19.789	-0.271	-0.271	43.231
100	25	75	0.5	3	3.0	2	0.500	0.000	0.005	-0.808	1.192	1.536	0.015	0.015	0.095
100	25	75	0.75	3	3.0	2	0.749	-0.001	0.002	-0.805	1.195	1.511	0.027	0.027	0.075
100	75	25	0.25	3	3.0	2	0.244	-0.006	0.007	-0.008	1.992	25.621	0.048	0.048	19.293
100	75	25	0.5	3	3.0	2	0.499	-0.001	0.005	-0.760	1.240	1.600	-0.002	-0.002	0.058
100	75	25	0.75	3	3.0	2	0.749	-0.001	0.002	-0.767	1.233	1.560	0.009	0.009	0.043
100	50	50	0.25	5	3.0	1	0.250	0.000	0.002	-0.617	1.383	2.002	-0.232	-0.232	0.117
100	50	50	0.5	5	3.0	1	0.498	-0.002	0.002	-0.621	1.379	1.935	-0.216	-0.216	0.084
100	50	50	0.75	5	3.0	1	0.750	0.000	0.001	-0.636	1.364	1.897	-0.225	-0.225	0.089

100	50	50	0.25	5	3.0	2	0.246	-0.004	0.002	-0.604	1.396	2.197	-0.339	-0.339	4.604
100	50	50	0.5	5	3.0	2	0.501	0.001	0.002	-0.651	1.349	1.861	-0.234	-0.234	0.105
100	50	50	0.75	5	3.0	2	0.748	-0.002	0.001	-0.645	1.355	1.871	-0.236	-0.236	0.099
100	25	75	0.25	5	3.0	2	0.249	-0.001	0.002	-0.688	1.312	1.886	-0.277	-0.277	0.219
100	25	75	0.5	5	3.0	2	0.498	-0.002	0.002	-0.683	1.317	1.790	-0.246	-2.244	5.100
100	25	75	0.75	5	3.0	2	0.747	-0.003	0.001	-0.670	1.330	1.814	-0.248	-0.248	0.110
100	75	25	0.25	5	3.0	2	0.248	-0.002	0.002	-0.625	1.375	1.987	-0.236	-0.236	0.131
100	75	25	0.5	5	3.0	2	0.499	-0.001	0.002	-0.641	1.359	1.887	-0.229	-0.229	0.096
100	75	25	0.75	5	3.0	2	0.750	0.000	0.001	-0.650	1.350	1.860	-0.246	-0.246	0.099
200	100	100	0.25	3	3.0	1	0.246	-0.004	0.004	0.547	2.547	86.923	0.492	0.492	36.553
200	100	100	0.5	3	3.0	1	0.498	-0.002	0.002	-0.630	1.370	1.898	0.066	0.066	0.022
200	100	100	0.75	3	3.0	1	0.748	-0.002	0.001	-0.633	1.367	1.882	0.068	0.068	0.019
200	100	100	0.25	3	3.0	2	0.249	-0.001	0.004	0.175	2.175	46.247	0.476	0.476	19.764
200	100	100	0.5	3	3.0	2	0.498	-0.002	0.002	-0.637	1.363	1.880	0.071	0.071	0.025
200	100	100	0.75	3	3.0	2	0.751	0.001	0.001	-0.642	1.358	1.856	0.072	0.072	0.020
200	50	150	0.25	3	3.0	2	0.246	-0.004	0.004	0.705	2.705	74.160	0.920	0.920	43.675

200	50	150	0.5	3	3.0	2	0.497	-0.003	0.002	-0.636	1.364	1.884	0.075	0.075	0.026
200	50	150	0.75	3	3.0	2	0.750	0.000	0.001	-0.645	1.355	1.849	0.067	0.067	0.019
200	150	50	0.25	3	3.0	2	0.242	-0.008	0.004	0.438	2.438	59.127	0.432	0.432	23.346
200	150	50	0.5	3	3.0	2	0.500	0.000	0.002	-0.632	1.368	1.894	0.062	0.062	0.022
200	150	50	0.75	3	3.0	2	0.748	-0.002	0.001	-0.633	1.367	1.880	0.061	0.061	0.017
200	100	100	0.25	5	3.0	1	0.245	-0.005	0.002	0.100	2.100	32.440	0.199	0.199	9.136
200	100	100	0.5	5	3.0	1	0.499	-0.001	0.001	-0.694	1.306	1.720	0.101	0.101	0.023
200	100	100	0.75	5	3.0	1	0.750	0.000	0.001	-0.695	1.305	1.716	0.099	0.099	0.023
200	100	100	0.25	5	3.0	2	0.245	-0.005	0.002	-0.175	1.825	28.496	0.255	0.255	8.923
200	100	100	0.5	5	3.0	2	0.500	0.000	0.001	-0.702	1.298	1.699	0.107	0.107	0.028
200	100	100	0.75	5	3.0	2	0.749	-0.001	0.001	-0.701	1.299	1.699	0.104	0.104	0.025
200	50	150	0.25	5	3.0	2	0.246	-0.004	0.002	-0.110	1.890	27.149	0.352	0.352	7.076
200	50	150	0.5	5	3.0	2	0.498	-0.002	0.001	-0.709	1.291	1.681	0.115	0.115	0.032
200	50	150	0.75	5	3.0	2	0.750	0.000	0.001	-0.707	1.293	1.685	0.115	0.115	0.031
200	150	50	0.25	5	3.0	2	0.247	-0.003	0.002	-0.038	1.962	41.816	0.120	0.120	12.500
200	150	50	0.5	5	3.0	2	0.500	0.000	0.001	-0.693	1.307	1.722	0.110	0.110	0.026

200	150	50	0.75	5	3.0	2	0.750	0.000	0.001	-0.696	1.304	1.712	0.107	0.107	0.025

APPENDIX C

Please find additional results from Chapter 4. All code is provided in Appendix A.

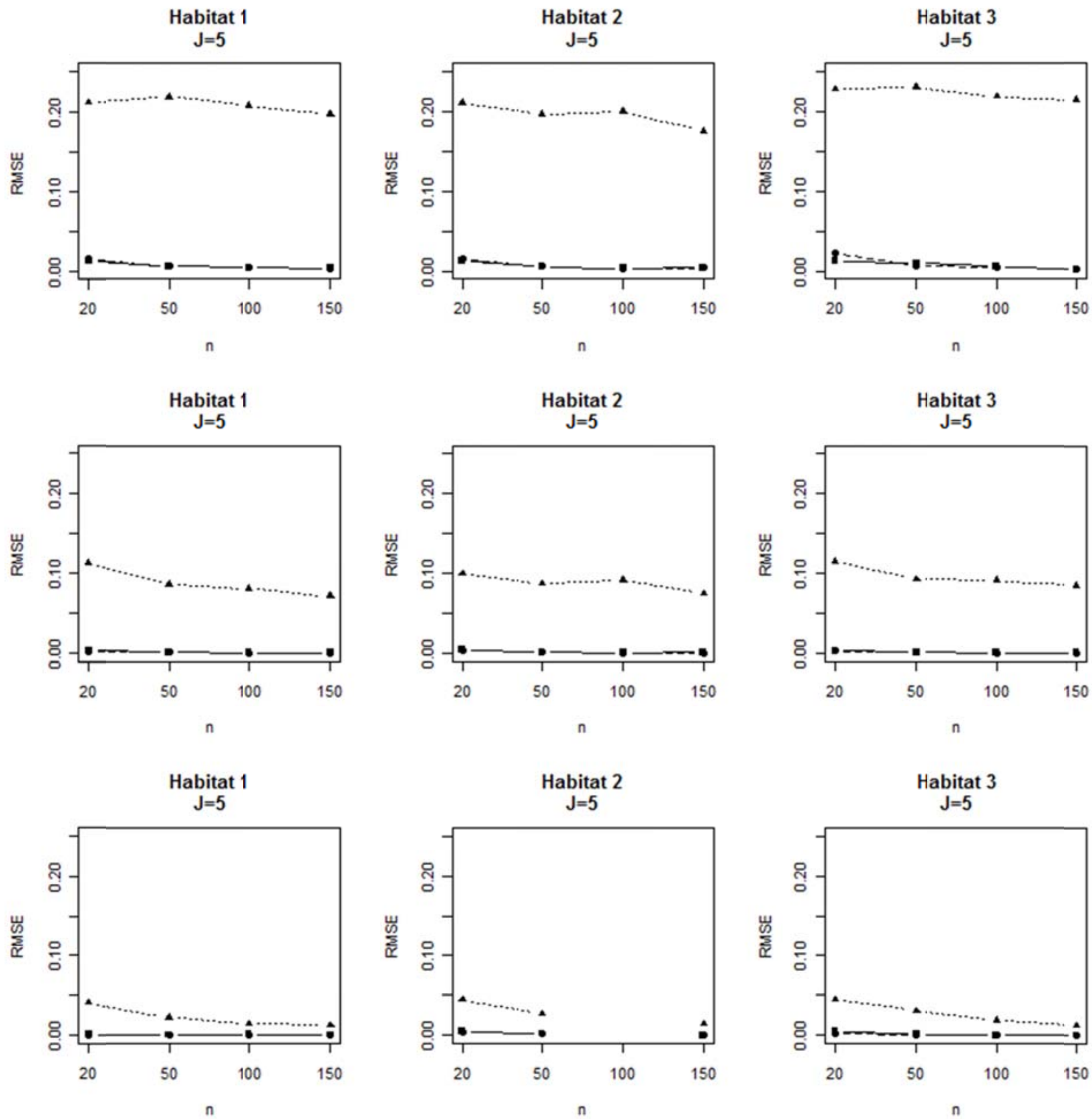
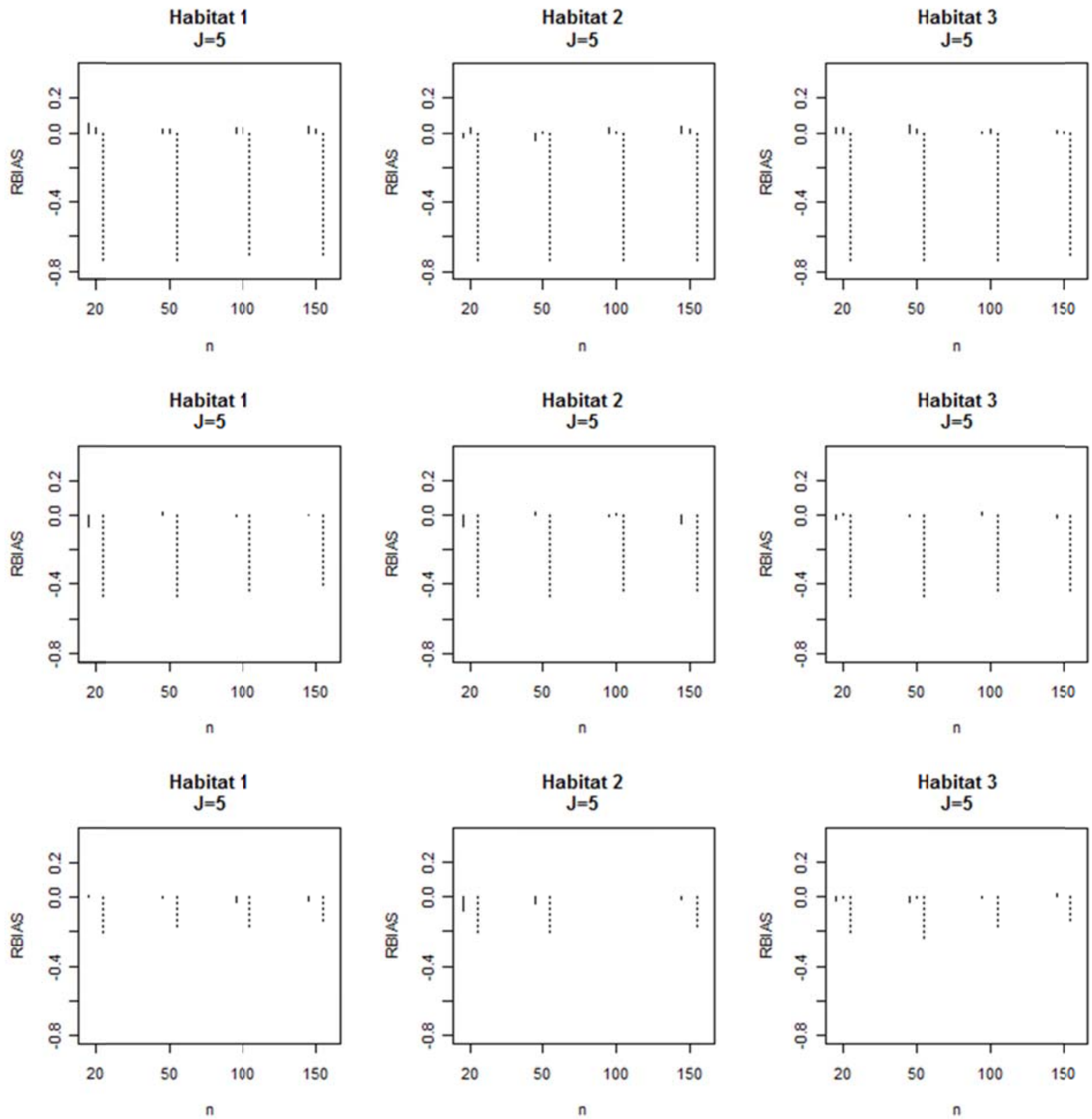

Habitat 1
J=5
Habitat 2
J=5
Habitat 3
J=5
Habitat 1
J=5
Habitat 2
J=5
Habitat 3
J=5
Habitat 1
J=5
Habitat 2
J=5
Habitat 3
J=5
RMSE
0.00
0.10
0.20
20
50
100
150
n

Figure C.1. Plots of the relative root mean-squared error (RMSE) comparing three different models: ACSOCC (Adaptive-cluster sampling occupancy, solid lines and squares), SSOCC (Single-season occupancy dashed lines with circles), and ACS (Adaptive-cluster sampling dotted lines with triangles). The columns differentiate among the three different habitats (1, 2, and 3) with varying levels of spatial correlation (high, med, none, respectively). The detection probability is fixed at $p = 0.25$ for the first row, $p = 0.5$ for the second row, and $p = 0.75$ for the third row. All cases have five repeat visits to a site (J). The x-axis represents an increase in the initial sample size, n (20, 50, 100, 150), for each of the three different models. Note that one simulation trial encountered fatal errors and is not represented ($n = 100$, $p = 0.75$, $J = 5$, Habitat 2).



Habitat 1
J=5
Habitat 2
J=5
Habitat 3
J=5
RBIAS
0.2
0.0
-0.4
-0.8
20
50
100
150
n
Habitat 1
J=5
Habitat 2
J=5
Habitat 3
J=5
RBIAS
0.2
0.0
-0.4
-0.8
20
50
100
150
n
Habitat 1
J=5
Habitat 2
J=5
Habitat 3
J=5
RBIAS
0.2
0.0
-0.4
-0.8
20
50
100
150
n

Figure C.2. Plots of the relative bias (RBIAS) comparing three different models: ACSOCC (Adaptive-cluster sampling occupancy, solid lines), SSOCC (Single-season occupancy dashed lines), and ACS (Adaptive-cluster sampling dotted lines). The columns differentiate among the three different habitats (1,2, and 3) with varying levels of spatial correlation (high, med, none, respectively). The detection probability is fixed at $p = 0.25$ for the first row, $p = 0.5$ for the second row, and $p = 0.75$ for the third row. All cases have five repeat visits to a site (J). The x-axis represents an increase in the initial sample size, n (20, 50, 100, 150), for each of the three different models. Note that one simulation trial encountered fatal errors and is not represented ($n = 100$, $p = 0.75$, $J = 5$, Habitat 2).

Table C.1. Simulation results from 500 synthetic datasets with different design criteria for the ACSOCC model (Adaptive-cluster sampling occupancy). Habitat refers to the amount of generated spatial correlation where Habitat 1 has the most spatial correlation and Habitat 3 has no spatial correlation (see text for more details). Parameters are *p*-detection probability, *n*- initial sample size, and *J*-number of repeat visits. *True* N^{tot} is the actual number of occupied sites out of 400 while N^{tot}-hat is the estimated number of occupied sites averaged over the 500 synthetic datasets. *Range* represents the minimum and maximum estimated number of occupied sites out of the 500 synthetic datasets for each set of design criteria. *p-hat* is the estimated detection probability while *RBIAS* and *RMSE* refer to relative bias and relative root mean-squared error, respectively. Note that one simulation trial encountered fatal errors and is not represented (n = 100, p = 0.75, J = 5, Habitat 2).

Scenario	n	p	J	Habitat	True Ntot	Ntot-hat	Range		Ntot RBias	Ntot RMSE	p-hat	p RBias	p RMSE
1	20	0.25	3	1	143	185.52	92.58	276.63	0.30	0.36	0.19	-0.23	0.37
min		0.25			143	83.60	10.00	166.00	-0.42	0.00	0.06	-0.75	0.00
max		0.25			143	251.65	186.00	360.00	0.76	0.76	0.48	0.92	0.92
2	50	0.25	3	1	162	170.75	105.22	248.02	0.05	0.19	0.23	-0.09	0.24
min		0.25			162	73.06	39.00	127.00	-0.55	0.00	0.07	-0.71	0.00
max		0.25			162	251.45	165.00	356.00	0.55	0.55	0.47	0.86	0.86
3	100	0.25	3	1	162	171.75	121.15	231.79	0.06	0.17	0.23	-0.06	0.19
min		0.25			162	103.53	68.00	140.00	-0.36	0.00	0.13	-0.47	0.00
max		0.25			162	253.24	196.00	314.00	0.56	0.56	0.41	0.62	0.62
4	150	0.25	3	1	150	159.48	116.63	210.81	0.06	0.15	0.24	-0.05	0.16
min		0.25			150	105.37	74.00	143.00	-0.30	0.00	0.13	-0.48	0.00
max		0.25			150	207.64	160.00	270.00	0.38	0.38	0.36	0.42	0.48
5	20	0.5	3	1	170	167.89	128.14	213.68	-0.01	0.10	0.49	-0.02	0.09
min		0.5			170	107.37	49.00	151.00	-0.37	0.00	0.24	-0.52	0.00

max		0.5			170	206.83	172.00	276.00	0.22	0.37	0.62	0.24	0.52
6	50	0.5	3	1	159	158.57	135.02	185.67	0.00	0.07	0.49	-0.01	0.08
min		0.5			159	102.98	68.00	141.00	-0.35	0.00	0.37	-0.26	0.00
max		0.5			159	192.58	166.00	226.00	0.21	0.35	0.58	0.17	0.26
7	100	0.5	3	1	150	149.46	130.93	171.86	0.00	0.06	0.49	-0.02	0.08
min		0.5			150	118.93	98.00	141.00	-0.21	0.00	0.37	-0.26	0.00
max		0.5			150	173.75	153.75	207.00	0.16	0.21	0.61	0.22	0.26
8	150	0.5	3	1	151	151.24	136.31	169.24	0.00	0.05	0.50	-0.01	0.07
min		0.5			151	126.70	114.00	142.00	-0.16	0.00	0.40	-0.19	0.00
max		0.5			151	171.69	154.00	195.00	0.14	0.16	0.61	0.22	0.22
9	20	0.75	3	1	148	138.10	110.42	169.53	-0.07	0.11	0.74	-0.01	0.05
min		0.75			148	64.62	20.00	134.00	-0.56	0.00	0.46	-0.38	0.00
max		0.75			148	161.36	138.00	227.00	0.09	0.56	0.83	0.11	0.38
10	50	0.75	3	1	170	160.46	143.06	179.38	-0.06	0.07	0.75	0.00	0.03
min		0.75			170	124.24	104.00	150.00	-0.27	0.00	0.67	-0.10	0.00
max		0.75			170	180.00	172.38	199.00	0.06	0.27	0.80	0.07	0.10

11	100	0.75	3	1	157	157.77	146.74	169.84	0.00	0.03	0.75	0.00	0.03
min		0.75			157	130.98	118.00	145.00	-0.17	0.00	0.67	-0.10	0.00
max		0.75			157	168.36	158.00	181.00	0.07	0.17	0.82	0.09	0.10
12	150	0.75	3	1	155	151.22	142.57	160.96	-0.02	0.04	0.75	0.00	0.03
min		0.75			155	137.08	127.00	148.00	-0.12	0.00	0.67	-0.11	0.00
max		0.75			155	160.18	152.00	170.00	0.03	0.12	0.81	0.08	0.11
13	20	0.25	5	1	148	155.94	98.47	224.87	0.05	0.19	0.23	-0.08	0.23
min		0.25			148	69.38	6.00	132.15	-0.53	0.00	0.04	-0.84	0.00
max		0.25			148	226.17	159.00	329.00	0.53	0.53	0.44	0.76	0.84
14	50	0.25	5	1	153	156.69	122.64	197.62	0.02	0.12	0.24	-0.03	0.13
min		0.25			153	90.88	59.00	134.00	-0.41	0.00	0.10	-0.59	0.00
max		0.25			153	218.48	173.00	263.00	0.43	0.43	0.33	0.33	0.59
15	100	0.25	5	1	154	159.00	131.94	192.21	0.03	0.10	0.24	-0.02	0.11
min		0.25			154	105.48	84.00	132.00	-0.32	0.00	0.17	-0.33	0.00
max		0.25			154	210.48	174.00	251.00	0.37	0.37	0.31	0.25	0.33
16	150	0.25	5	1	149	154.05	129.92	184.31	0.03	0.10	0.25	-0.02	0.11

min		0.25			149	105.98	87.00	130.00	-0.29	0.00	0.16	-0.37	0.00
max		0.25			149	190.67	160.00	234.00	0.28	0.29	0.34	0.35	0.37
17	20	0.5	5	1	160	149.38	121.59	181.25	-0.07	0.10	0.50	-0.01	0.06
min		0.5			160	103.54	20.00	134.78	-0.35	0.00	0.23	-0.55	0.00
max		0.5			160	179.22	160.35	236.00	0.12	0.35	0.60	0.19	0.55
18	50	0.5	5	1	141	142.95	125.08	163.19	0.01	0.06	0.50	0.00	0.05
min		0.5			141	106.33	82.00	131.00	-0.25	0.00	0.43	-0.14	0.00
max		0.5			141	160.25	144.00	179.00	0.14	0.25	0.57	0.15	0.15
19	100	0.5	5	1	152	150.36	138.42	163.50	-0.01	0.04	0.50	0.00	0.05
min		0.5			152	132.17	117.00	145.00	-0.13	0.00	0.43	-0.13	0.00
max		0.5			152	165.32	154.00	178.00	0.09	0.13	0.58	0.16	0.16
20	150	0.5	5	1	157	156.79	147.19	167.51	0.00	0.03	0.50	0.00	0.04
min		0.5			157	142.75	131.00	155.00	-0.09	0.00	0.44	-0.11	0.00
max		0.5			157	167.87	159.00	179.00	0.07	0.09	0.56	0.11	0.11
21	20	0.75	5	1	162	163.07	134.24	194.92	0.01	0.05	0.75	0.00	0.03
min		0.75			162	132.64	104.00	172.00	-0.18	0.00	0.69	-0.08	0.00

max		0.75			162	182.86	163.98	220.00	0.13	0.18	0.81	0.08	0.08
22	50	0.75	5	1	159	158.54	140.96	177.80	0.00	0.04	0.75	0.00	0.02
min		0.75			159	137.27	118.00	161.00	-0.14	0.00	0.70	-0.06	0.00
max		0.75			159	174.22	163.68	195.00	0.10	0.14	0.80	0.07	0.07
23	100	0.75	5	1	140	135.90	125.05	148.18	-0.03	0.05	0.75	0.00	0.02
min		0.75			140	117.43	105.00	130.00	-0.16	0.00	0.67	-0.10	0.00
max		0.75			140	149.49	140.00	160.00	0.07	0.16	0.81	0.08	0.10
24	150	0.75	5	1	167	163.51	156.44	171.40	-0.02	0.03	0.75	0.00	0.02
min		0.75			167	149.80	140.00	160.00	-0.10	0.00	0.71	-0.06	0.00
max		0.75			167	172.25	165.00	180.00	0.03	0.10	0.80	0.07	0.07
25	20	0.25	3	2	141	172.50	82.41	273.02	0.22	0.34	0.20	-0.18	0.39
min		0.25			141	53.43	6.00	146.00	-0.62	0.00	0.07	-0.74	0.00
max		0.25			141	253.52	160.00	375.00	0.80	0.80	0.71	1.85	1.85
26	50	0.25	3	2	141	174.40	104.91	254.62	0.24	0.34	0.21	-0.16	0.30
min		0.25			141	54.86	21.00	110.00	-0.61	0.00	0.06	-0.78	0.00
max		0.25			141	256.19	196.00	361.00	0.82	0.82	0.46	0.83	0.83

27	100	0.25	3	2	142	147.26	101.98	205.26	0.04	0.18	0.24	-0.03	0.20
min		0.25			142	74.66	44.00	120.18	-0.47	0.00	0.09	-0.64	0.00
max		0.25			142	224.32	155.00	335.00	0.58	0.58	0.40	0.61	0.64
28	150	0.25	3	2	149	159.89	120.70	208.45	0.07	0.17	0.24	-0.02	0.17
min		0.25			149	96.26	70.00	131.00	-0.35	0.00	0.14	-0.45	0.00
max		0.25			149	221.01	161.00	303.00	0.48	0.48	0.37	0.49	0.49
29	20	0.5	3	2	157	145.21	105.69	190.00	-0.08	0.13	0.49	-0.02	0.13
min		0.5			157	70.89	41.00	121.00	-0.55	0.00	0.15	-0.71	0.00
max		0.5			157	203.49	151.00	301.00	0.30	0.55	0.70	0.39	0.71
30	50	0.5	3	2	134	124.69	96.40	157.52	-0.07	0.13	0.50	-0.01	0.10
min		0.5			134	74.30	46.00	103.00	-0.45	0.00	0.36	-0.27	0.00
max		0.5			134	158.83	134.00	205.00	0.19	0.45	0.67	0.35	0.35
31	100	0.5	3	2	154	154.64	136.83	175.08	0.00	0.06	0.50	0.00	0.07
min		0.5			154	118.54	97.00	143.00	-0.23	0.00	0.40	-0.20	0.00
max		0.5			154	177.45	159.00	199.00	0.15	0.23	0.59	0.19	0.20
32	150	0.5	3	2	140	143.43	127.71	161.97	0.02	0.07	0.49	-0.02	0.07

min		0.5			140	109.71	94.00	129.00	-0.22	0.00	0.39	-0.23	0.00
max		0.5			140	171.20	149.00	198.00	0.22	0.22	0.58	0.17	0.23
33	20	0.75	3	2	148	131.47	104.20	164.67	-0.11	0.14	0.74	-0.01	0.04
min		0.75			148	50.27	6.00	122.00	-0.66	0.00	0.61	-0.19	0.00
max		0.75			148	159.60	141.90	194.00	0.08	0.66	0.86	0.15	0.19
34	50	0.75	3	2	144	138.53	120.87	157.09	-0.04	0.07	0.75	0.00	0.04
min		0.75			144	108.77	80.00	135.00	-0.24	0.00	0.64	-0.14	0.00
max		0.75			144	155.33	139.00	174.00	0.08	0.24	0.83	0.10	0.14
35	100	0.75	3	2	166	163.28	152.10	175.05	-0.02	0.04	0.75	0.00	0.03
min		0.75			166	143.81	128.00	159.00	-0.13	0.00	0.69	-0.09	0.00
max		0.75			166	175.37	168.68	187.00	0.06	0.13	0.83	0.11	0.11
36	150	0.75	3	2	140	139.63	129.66	150.66	0.00	0.03	0.74	-0.01	0.04
min		0.75			140	123.70	112.00	136.00	-0.12	0.00	0.66	-0.12	0.00
max		0.75			140	149.74	141.00	161.00	0.07	0.12	0.82	0.10	0.12
37	20	0.25	5	2	143	138.99	83.84	209.32	-0.03	0.19	0.24	-0.05	0.25
min		0.25			143	56.27	6.00	98.00	-0.61	0.00	0.05	-0.81	0.00

max		0.25			143	236.20	156.00	350.00	0.65	0.65	0.46	0.84	0.84
38	50	0.25	5	2	139	132.93	101.11	171.54	-0.04	0.12	0.25	-0.01	0.14
min		0.25			139	81.97	49.00	121.00	-0.41	0.00	0.14	-0.46	0.00
max		0.25			139	182.13	139.00	255.00	0.31	0.41	0.35	0.41	0.46
39	100	0.25	5	2	144	148.26	123.25	177.48	0.03	0.09	0.25	-0.02	0.11
min		0.25			144	113.23	79.00	139.00	-0.21	0.00	0.17	-0.34	0.00
max		0.25			144	197.55	155.00	235.00	0.37	0.37	0.34	0.34	0.34
40	150	0.25	5	2	129	133.80	108.74	165.75	0.04	0.12	0.24	-0.04	0.13
min		0.25			129	93.62	73.00	115.00	-0.27	0.00	0.13	-0.47	0.00
max		0.25			129	177.48	143.00	220.00	0.38	0.38	0.33	0.31	0.47
41	20	0.5	5	2	138	129.26	95.63	165.76	-0.06	0.11	0.49	-0.01	0.10
min		0.5			138	80.48	6.00	132.00	-0.42	0.00	0.09	-0.82	0.00
max		0.5			138	165.44	138.75	276.00	0.20	0.42	0.64	0.27	0.82
42	50	0.5	5	2	143	144.78	127.14	163.44	0.01	0.05	0.50	0.00	0.05
min		0.5			143	106.88	84.00	133.00	-0.25	0.00	0.41	-0.17	0.00
max		0.5			143	163.50	148.00	181.00	0.14	0.25	0.57	0.14	0.17

43	100	0.5	5	2	164	163.04	151.70	175.32	-0.01	0.03	0.50	0.00	0.04
min		0.5			164	147.16	133.00	161.00	-0.10	0.00	0.43	-0.14	0.00
max		0.5			164	176.20	166.00	188.00	0.07	0.10	0.56	0.11	0.14
44	150	0.5	5	2	138	131.49	121.11	142.91	-0.05	0.06	0.50	0.00	0.05
min		0.5			138	116.49	106.00	128.00	-0.16	0.00	0.44	-0.13	0.00
max		0.5			138	146.61	136.00	158.00	0.06	0.16	0.57	0.14	0.14
45	20	0.75	5	2	148	136.56	105.66	168.63	-0.08	0.11	0.75	0.00	0.04
min		0.75			148	76.78	29.00	122.00	-0.48	0.00	0.66	-0.13	0.00
max		0.75			148	161.70	151.45	190.00	0.09	0.48	0.85	0.13	0.13
46	50	0.75	5	2	143	136.82	117.43	156.61	-0.04	0.08	0.75	0.00	0.03
min		0.75			143	92.33	64.00	120.00	-0.35	0.00	0.68	-0.09	0.00
max		0.75			143	162.88	148.00	178.00	0.14	0.35	0.82	0.09	0.09
47	100	0.75	5	2									
min		0.75											
max		0.75											
48	150	0.75	5	2	140	138.43	129.77	147.80	-0.01	0.03	0.75	0.00	0.03

min		0.75			140	122.37	114.00	132.00	-0.13	0.00	0.69	-0.08	0.00
max		0.75			140	148.90	141.23	158.00	0.06	0.13	0.80	0.07	0.08
49	20	0.25	3	3	150	188.00	94.17	281.27	0.25	0.32	0.20	-0.20	0.36
min		0.25			150	89.07	9.00	162.00	-0.41	0.00	0.07	-0.72	0.00
max		0.25			150	250.51	171.00	362.00	0.67	0.67	0.57	1.28	1.28
50	50	0.25	3	3	155	178.96	106.74	258.33	0.15	0.26	0.22	-0.13	0.29
min		0.25			155	65.73	30.00	120.00	-0.58	0.00	0.07	-0.73	0.00
max		0.25			155	260.40	184.00	355.00	0.68	0.68	0.47	0.89	0.89
51	100	0.25	3	3	145	158.84	107.09	219.90	0.10	0.21	0.24	-0.05	0.21
min		0.25			145	83.99	39.00	120.00	-0.42	0.00	0.11	-0.55	0.00
max		0.25			145	228.60	180.00	310.00	0.58	0.58	0.44	0.78	0.78
52	150	0.25	3	3	149	157.58	112.57	212.66	0.06	0.17	0.24	-0.05	0.19
min		0.25			149	88.47	64.00	124.00	-0.41	0.00	0.13	-0.50	0.00
max		0.25			149	229.43	173.00	298.00	0.54	0.54	0.41	0.64	0.64
53	20	0.5	3	3	158	151.11	103.39	204.79	-0.04	0.13	0.48	-0.05	0.17
min		0.5			158	78.13	25.00	135.75	-0.51	0.00	0.08	-0.85	0.00

max		0.5			158	209.93	156.00	301.00	0.33	0.51	0.72	0.44	0.85
54	50	0.5	3	3	156	149.89	123.73	180.04	-0.04	0.09	0.49	-0.02	0.10
min		0.5			156	100.55	67.98	139.00	-0.36	0.00	0.33	-0.35	0.00
max		0.5			156	195.51	154.00	236.00	0.25	0.36	0.61	0.22	0.35
55	100	0.5	3	3	152	148.61	129.30	170.35	-0.02	0.07	0.50	0.00	0.08
min		0.5			152	118.63	96.00	139.00	-0.22	0.00	0.37	-0.26	0.00
max		0.5			152	173.53	154.00	200.00	0.14	0.22	0.62	0.23	0.26
56	150	0.5	3	3	143	139.97	123.84	159.07	-0.02	0.06	0.50	-0.01	0.08
min		0.5			143	114.32	99.00	129.00	-0.20	0.00	0.35	-0.30	0.00
max		0.5			143	167.40	146.00	199.00	0.17	0.20	0.63	0.26	0.30
57	20	0.75	3	3	146	135.04	97.78	172.10	-0.08	0.13	0.73	-0.02	0.08
min		0.75			146	60.68	23.00	113.00	-0.58	0.00	0.52	-0.31	0.00
max		0.75			146	167.72	139.00	203.00	0.15	0.58	0.88	0.17	0.31
58	50	0.75	3	3	157	159.78	138.50	179.81	0.02	0.06	0.74	-0.01	0.04
min		0.75			157	114.03	88.00	142.00	-0.27	0.00	0.64	-0.15	0.00
max		0.75			157	182.57	165.00	200.00	0.16	0.27	0.84	0.12	0.15

59	100	0.75	3	3	150	140.75	128.12	153.68	-0.06	0.07	0.75	0.00	0.04
min		0.75			150	120.32	106.00	135.00	-0.20	0.00	0.67	-0.11	0.00
max		0.75			150	155.24	145.00	167.00	0.03	0.20	0.83	0.11	0.11
60	150	0.75	3	3	166	165.91	157.13	175.04	0.00	0.02	0.75	0.00	0.03
min		0.75			166	153.49	143.00	163.00	-0.08	0.00	0.68	-0.10	0.00
max		0.75			166	175.86	167.00	187.00	0.06	0.08	0.82	0.09	0.10
61	20	0.25	5	3	153	158.02	92.64	228.96	0.03	0.18	0.23	-0.06	0.25
min		0.25			153	74.86	6.00	135.00	-0.51	0.00	0.06	-0.76	0.00
max		0.25			153	223.62	163.00	327.00	0.46	0.51	0.48	0.93	0.93
62	50	0.25	5	3	160	167.48	125.05	216.84	0.05	0.16	0.23	-0.07	0.17
min		0.25			160	99.36	65.00	141.00	-0.38	0.00	0.10	-0.58	0.00
max		0.25			160	234.21	187.00	287.00	0.46	0.46	0.36	0.43	0.58
63	100	0.25	5	3	158	159.23	128.39	196.76	0.01	0.12	0.24	-0.05	0.14
min		0.25			158	81.27	60.00	109.00	-0.49	0.00	0.15	-0.41	0.00
max		0.25			158	219.61	185.00	255.00	0.39	0.49	0.33	0.33	0.41
64	150	0.25	5	3	159	160.63	136.60	188.98	0.01	0.08	0.24	-0.02	0.10

min		0.25			159	124.84	106.00	145.00	-0.21	0.00	0.18	-0.29	0.00
max		0.25			159	202.82	173.00	234.00	0.28	0.28	0.33	0.31	0.31
65	20	0.5	5	3	158	153.38	122.06	182.55	-0.03	0.09	0.50	-0.01	0.08
min		0.5			158	79.48	22.00	151.00	-0.50	0.00	0.28	-0.43	0.00
max		0.5			158	177.51	158.00	223.00	0.12	0.50	0.62	0.23	0.43
66	50	0.5	5	3	152	150.50	132.54	167.82	-0.01	0.05	0.50	0.00	0.05
min		0.5			152	117.21	90.00	145.00	-0.23	0.00	0.43	-0.14	0.00
max		0.5			152	170.82	157.00	185.00	0.12	0.23	0.58	0.15	0.15
67	100	0.5	5	3	169	171.32	160.82	182.21	0.01	0.03	0.50	0.00	0.04
min		0.5			169	156.37	143.00	169.00	-0.07	0.00	0.44	-0.12	0.00
max		0.5			169	184.60	174.00	196.00	0.09	0.09	0.56	0.12	0.12
68	150	0.5	5	3	166	163.51	154.00	173.66	-0.02	0.03	0.50	0.00	0.04
min		0.5			166	146.50	136.00	157.00	-0.12	0.00	0.44	-0.12	0.00
max		0.5			166	175.23	166.00	186.00	0.06	0.12	0.57	0.13	0.13
69	20	0.75	5	3	163	159.37	123.53	191.32	-0.02	0.11	0.75	0.00	0.04
min		0.75			163	78.92	31.00	138.00	-0.52	0.00	0.61	-0.19	0.00

max		0.75			163	189.96	169.00	215.00	0.17	0.52	0.85	0.13	0.19
70	50	0.75	5	3	149	143.73	125.86	161.24	-0.04	0.06	0.75	0.00	0.03
min		0.75			149	103.90	75.00	133.00	-0.30	0.00	0.67	-0.10	0.00
max		0.75			149	157.46	144.50	173.00	0.06	0.30	0.81	0.09	0.10
71	100	0.75	5	3	159	157.97	147.65	168.30	-0.01	0.02	0.75	0.00	0.02
min		0.75			159	144.75	129.00	158.00	-0.09	0.00	0.70	-0.07	0.00
max		0.75			159	167.07	157.00	177.00	0.05	0.09	0.80	0.06	0.07
72	150	0.75	5	3	161	164.32	156.50	172.15	0.02	0.03	0.75	0.00	0.02
min		0.75			161	153.71	143.00	163.00	-0.05	0.00	0.70	-0.07	0.00
max		0.75			161	173.83	169.00	182.00	0.08	0.08	0.80	0.07	0.07

Table C.2. Simulation results from 500 synthetic datasets with different design criteria for the SSOCC model (Single-season occupancy). Habitat refers to the amount of generated spatial correlation where Habitat 1 has the most spatial correlation and Habitat 3 has no spatial correlation (see text for more details). Parameters are *p*-detection probability, *n*- initial sample size, and *J*-number of repeat visits. *Sample size* is the realized number of sites sampled from the adaptive sampling models and is the true sample size for this model. *True* N^{tot} is the actual number of occupied sites out of 400 while N^{tot}-hat is the estimated number of occupied sites averaged over the 500 synthetic datasets. *Var(Ntot)* represents the estimated variance of the estimated number of occupied sites out of the. *p-hat* is the estimated detection probability while *RBIAS* and *RMSE* refer to relative bias and relative root mean-squared error, respectively. Note that one simulation trial encountered fatal errors and is not represented (n = 100, p = 0.75, J = 5, Habitat 2).

Scenario	*n*	*Sample Size*	*p*	*J*	*Habitat*	*True Ntot*	*Ntot-hat*	*Var(Ntot)*	*Ntot RBias*	*Ntot RMSE*	*p-hat*	*p RBias*	*p RMSE*
1	20	57.61	0.25	3	1	143	167.77	44.71	0.17	0.43	0.24	-0.04	0.41
min		20.00	0.25			143	34.64	0.16	-0.76	0.00	0.02	-0.91	0.00
max		116.00	0.25			143	397.40	93.96	1.78	1.78	0.64	1.54	1.54
2	50	122.83	0.25	3	1	162	162.36	59.82	0.00	0.22	0.26	0.03	0.24
min		67.00	0.25			162	78.17	14.81	-0.52	0.00	0.12	-0.53	0.00
max		183.00	0.25			162	313.50	85.48	0.94	0.94	0.42	0.69	0.69
3	100	199.56	0.25	3	1	162	169.46	51.43	0.05	0.20	0.25	-0.01	0.21
min		156.00	0.25			162	88.37	12.11	-0.45	0.00	0.12	-0.51	0.00
max		238.00	0.25			162	291.88	69.44	0.80	0.80	0.46	0.82	0.82
4	150	243.75	0.25	3	1	150	153.03	45.89	0.02	0.17	0.25	0.00	0.18
min		211.00	0.25			150	93.92	27.70	-0.37	0.00	0.14	-0.43	0.00
max		282.00	0.25			150	235.31	63.92	0.57	0.57	0.39	0.55	0.55
5	20	151.33	0.5	3	1	170	170.16	52.11	0.00	0.09	0.50	0.00	0.09
min		44.00	0.5			170	116.93	37.91	-0.31	0.00	0.35	-0.31	0.00

max		208.00	0.5			170	224.96	76.58	0.32	0.32	0.63	0.27	0.31
6	50	206.93	0.5	3	1	159	158.03	40.28	-0.01	0.07	0.50	0.01	0.08
min		115.00	0.5			159	122.58	26.05	-0.23	0.00	0.39	-0.22	0.00
max		249.00	0.5			159	193.71	55.04	0.22	0.23	0.66	0.32	0.32
7	100	254.70	0.5	3	1	150	152.21	35.44	0.01	0.08	0.49	-0.01	0.08
min		201.00	0.5			150	117.85	26.08	-0.21	0.00	0.35	-0.30	0.00
max		288.00	0.5			150	188.27	46.66	0.26	0.26	0.63	0.27	0.30
8	150	291.10	0.5	3	1	151	151.38	27.90	0.00	0.06	0.50	0.00	0.07
min		253.00	0.5			151	124.09	20.61	-0.18	0.00	0.41	-0.19	0.00
max		329.00	0.5			151	180.07	37.59	0.19	0.19	0.61	0.21	0.21
9	20	151.14	0.75	3	1	148	149.43	44.60	0.01	0.08	0.75	0.00	0.05
min		29.00	0.75			148	116.05	30.48	-0.22	0.00	0.63	-0.17	0.00
max		220.00	0.75			148	204.77	78.48	0.38	0.38	0.87	0.16	0.17
10	50	222.27	0.75	3	1	170	169.95	33.24	0.00	0.05	0.75	0.00	0.04
min		172.00	0.75			170	144.31	25.65	-0.15	0.00	0.64	-0.14	0.00
max		261.00	0.75			170	192.72	45.54	0.13	0.15	0.85	0.13	0.14

11	100	287.90	0.75	3	1	157	157.23	22.02	0.00	0.04	0.75	0.00	0.03
min		250.00	0.75			157	137.47	15.79	-0.12	0.00	0.65	-0.13	0.00
max		318.00	0.75			157	175.77	29.75	0.12	0.12	0.82	0.09	0.13
12	150	309.25	0.75	3	1	155	155.03	18.14	0.00	0.03	0.75	0.00	0.03
min		270.00	0.75			155	136.10	13.10	-0.12	0.00	0.67	-0.10	0.00
max		337.00	0.75			155	169.43	26.71	0.09	0.12	0.83	0.11	0.11
13	20	95.27	0.25	5	1	148	154.04	53.71	0.04	0.20	0.25	-0.02	0.21
min		20.00	0.25			148	62.29	3.83	-0.58	0.00	0.09	-0.64	0.00
max		166.00	0.25			148	252.59	83.35	0.71	0.71	0.46	0.83	0.83
14	50	171.48	0.25	5	1	153	156.00	44.71	0.02	0.13	0.25	-0.01	0.15
min		83.00	0.25			153	93.96	18.30	-0.39	0.00	0.12	-0.51	0.00
max		233.00	0.25			153	213.87	64.91	0.40	0.40	0.35	0.42	0.51
15	100	237.51	0.25	5	1	154	159.19	41.83	0.03	0.11	0.25	-0.02	0.12
min		185.00	0.25			154	114.67	29.69	-0.26	0.00	0.16	-0.36	0.00
max		275.00	0.25			154	216.28	53.25	0.40	0.40	0.33	0.31	0.36
16	150	273.50	0.25	5	1	149	152.64	41.10	0.02	0.10	0.25	-0.02	0.12

min		231.00	0.25			149	111.86	30.47	-0.25	0.00	0.17	-0.33	0.00
max		305.00	0.25			149	205.71	52.56	0.38	0.38	0.33	0.32	0.33
17	20	159.33	0.5	5	1	160	160.31	41.87	0.00	0.07	0.50	0.00	0.06
min		28.00	0.5			160	125.28	28.68	-0.22	0.00	0.40	-0.21	0.00
max		226.00	0.5			160	198.04	59.55	0.24	0.24	0.59	0.19	0.21
18	50	217.26	0.5	5	1	141	140.99	35.11	0.00	0.06	0.50	0.00	0.06
min		164.00	0.5			141	112.01	26.68	-0.21	0.00	0.41	-0.17	0.00
max		252.00	0.5			141	164.11	46.15	0.16	0.21	0.58	0.16	0.17
19	100	266.27	0.5	5	1	152	152.25	25.71	0.00	0.04	0.50	0.00	0.05
min		217.00	0.5			152	133.11	19.51	-0.12	0.00	0.42	-0.16	0.00
max		299.00	0.5			152	171.40	33.20	0.13	0.13	0.56	0.12	0.16
20	150	306.93	0.5	5	1	157	156.94	18.74	0.00	0.03	0.50	0.00	0.04
min		270.00	0.5			157	135.01	13.66	-0.14	0.00	0.43	-0.13	0.00
max		335.00	0.5			157	171.62	23.48	0.09	0.14	0.56	0.12	0.13
21	20	182.45	0.75	5	1	162	162.65	38.64	0.00	0.06	0.75	0.00	0.03
min		131.00	0.75			162	138.54	26.97	-0.14	0.00	0.67	-0.10	0.00

max		217.00	0.75			162	185.75	48.00	0.15	0.15	0.81	0.08	0.10
22	50	233.36	0.75	5	1	159	158.86	28.15	0.00	0.04	0.75	0.00	0.03
min		205.00	0.75			159	127.26	20.69	-0.20	0.00	0.69	-0.07	0.00
max		273.00	0.75			159	179.51	36.85	0.13	0.20	0.82	0.10	0.10
23	100	260.57	0.75	5	1	140	140.24	21.95	0.00	0.04	0.75	0.00	0.03
min		220.00	0.75			140	124.18	16.03	-0.11	0.00	0.69	-0.07	0.00
max		300.00	0.75			140	158.07	29.96	0.13	0.13	0.82	0.09	0.09
24	150	323.01	0.75	5	1	167	167.29	13.69	0.00	0.02	0.75	0.00	0.02
min		285.00	0.75			167	155.55	8.92	-0.07	0.00	0.70	-0.06	0.00
max		349.00	0.75			167	178.71	20.28	0.07	0.07	0.81	0.07	0.07
25	20	51.18	0.25	3	2	141	149.33	60.87	0.06	0.41	0.27	0.06	0.41
min		20.00	0.25			141	25.74	0.13	-0.82	0.00	0.01	-0.97	0.00
max		115.00	0.25			141	399.15	96.94	1.83	1.83	0.64	1.56	1.56
26	50	112.83	0.25	3	2	141	155.70	58.06	0.10	0.30	0.24	-0.04	0.28
min		64.00	0.25			141	60.24	17.44	-0.57	0.00	0.05	-0.80	0.00
max		171.00	0.25			141	292.32	90.13	1.07	1.07	0.46	0.83	0.83

27	100	180.82	0.25	3	2	142	146.68	52.68	0.03	0.21	0.25	0.00	0.22
min		135.00	0.25			142	75.98	0.01	-0.46	0.00	0.10	-0.61	0.00
max		219.00	0.25			142	399.99	82.08	1.82	1.82	0.42	0.67	0.67
28	150	248.18	0.25	3	2	149	154.28	47.17	0.04	0.19	0.25	-0.01	0.19
min		205.00	0.25			149	98.21	0.00	-0.34	0.00	0.08	-0.69	0.00
max		283.00	0.25			149	400.00	67.38	1.68	1.68	0.41	0.62	0.69
29	20	103.78	0.5	3	2	157	158.33	54.94	0.01	0.12	0.50	0.00	0.12
min		40.00	0.5			157	78.88	26.25	-0.50	0.00	0.31	-0.39	0.00
max		189.00	0.5			157	228.02	79.19	0.45	0.50	0.66	0.32	0.39
30	50	147.04	0.5	3	2	134	134.25	50.77	0.00	0.11	0.50	0.00	0.11
min		87.00	0.5			134	93.49	36.54	-0.30	0.00	0.36	-0.28	0.00
max		208.00	0.5			134	212.30	70.84	0.58	0.58	0.67	0.33	0.33
31	100	261.93	0.5	3	2	154	153.00	33.72	-0.01	0.06	0.50	0.01	0.07
min		196.00	0.5			154	124.28	24.90	-0.19	0.00	0.39	-0.22	0.00
max		302.00	0.5			154	175.13	43.65	0.14	0.19	0.61	0.22	0.22
32	150	293.54	0.5	3	2	140	140.75	28.67	0.01	0.06	0.50	0.00	0.07

min		260.00	0.5			140	117.68	19.82	-0.16	0.00	0.35	-0.30	0.00
max		318.00	0.5			140	167.64	39.92	0.20	0.20	0.59	0.19	0.30
33	20	148.37	0.75	3	2	148	148.63	43.88	0.00	0.08	0.75	0.00	0.05
min		24.00	0.75			148	112.26	31.95	-0.24	0.00	0.63	-0.16	0.00
max		207.00	0.75			148	206.15	83.93	0.39	0.39	0.85	0.13	0.16
34	50	207.84	0.75	3	2	144	143.73	35.33	0.00	0.06	0.75	0.00	0.04
min		127.00	0.75			144	116.83	25.35	-0.19	0.00	0.63	-0.17	0.00
max		257.00	0.75			144	174.66	56.32	0.21	0.21	0.84	0.12	0.17
35	100	296.91	0.75	3	2	166	165.65	21.63	0.00	0.03	0.75	0.00	0.03
min		246.00	0.75			166	144.44	15.25	-0.13	0.00	0.67	-0.11	0.00
max		331.00	0.75			166	182.65	31.37	0.10	0.13	0.83	0.11	0.11
36	150	302.54	0.75	3	2	140	140.93	17.62	0.01	0.04	0.75	-0.01	0.04
min		269.00	0.75			140	126.92	12.46	-0.09	0.00	0.65	-0.13	0.00
max		338.00	0.75			140	161.20	24.94	0.15	0.15	0.83	0.11	0.13
37	20	75.42	0.25	5	2	143	148.19	60.26	0.04	0.21	0.25	0.00	0.22
min		20.00	0.25			143	60.19	11.33	-0.58	0.00	0.06	-0.77	0.00

max		148.00	0.25			143	336.90	90.21	1.36	1.36	0.44	0.76	0.77
38	50	144.49	0.25	5	2	139	139.21	47.71	0.00	0.13	0.25	0.01	0.16
min		80.00	0.25			139	90.15	30.29	-0.35	0.00	0.14	-0.43	0.00
max		226.00	0.25			139	207.60	74.49	0.49	0.49	0.36	0.45	0.45
39	100	226.56	0.25	5	2	144	144.59	39.11	0.00	0.10	0.25	0.00	0.12
min		166.00	0.25			144	110.36	27.41	-0.23	0.00	0.16	-0.36	0.00
max		271.00	0.25			144	195.98	54.48	0.36	0.36	0.33	0.32	0.36
40	150	262.08	0.25	5	2	129	131.35	43.06	0.02	0.11	0.25	0.00	0.12
min		226.00	0.25			129	95.12	31.18	-0.26	0.00	0.16	-0.37	0.00
max		295.00	0.25			129	172.54	60.17	0.34	0.34	0.34	0.36	0.37
41	20	108.73	0.5	5	2	138	137.81	50.47	0.00	0.11	0.50	0.00	0.08
min		20.00	0.5			138	97.31	30.81	-0.29	0.00	0.38	-0.25	0.00
max		190.00	0.5			138	187.43	77.62	0.36	0.36	0.66	0.33	0.33
42	50	216.47	0.5	5	2	143	143.28	33.37	0.00	0.06	0.50	0.00	0.05
min		123.00	0.5			143	120.83	23.20	-0.16	0.00	0.43	-0.14	0.00
max		269.00	0.5			143	170.89	49.74	0.20	0.20	0.58	0.15	0.15

43	100	297.07	0.5	5	2	164	164.74	20.56	0.00	0.03	0.50	0.00	0.04
min		252.00	0.5			164	147.97	13.38	-0.10	0.00	0.43	-0.14	0.00
max		328.00	0.5			164	184.36	29.55	0.12	0.12	0.55	0.11	0.14
44	150	297.47	0.5	5	2	138	138.22	21.34	0.00	0.04	0.50	0.00	0.05
min		264.00	0.5			138	120.84	16.46	-0.12	0.00	0.43	-0.15	0.00
max		325.00	0.5			138	154.86	30.16	0.12	0.12	0.58	0.15	0.15
45	20	118.26	0.75	5	2	148	148.78	48.15	0.01	0.10	0.75	0.00	0.04
min		37.00	0.75			148	70.23	32.83	-0.53	0.00	0.62	-0.18	0.00
max		185.00	0.75			148	192.61	78.77	0.30	0.53	0.84	0.12	0.18
46	50	185.33	0.75	5	2	143	143.46	37.23	0.00	0.06	0.75	0.00	0.03
min		103.00	0.75			143	113.77	25.79	-0.20	0.00	0.68	-0.09	0.00
max		247.00	0.75			143	175.14	56.17	0.22	0.22	0.82	0.09	0.09
47	100		0.75	5	2								
min			0.75										
max			0.75										
48	150	308.05	0.75	5	2	140	140.26	15.64	0.00	0.03	0.75	0.00	0.03

min		273.00	0.75			140	122.87	10.59	-0.12	0.00	0.69	-0.08	0.00
max		337.00	0.75			140	153.59	20.34	0.10	0.12	0.81	0.08	0.08
49	20	50.98	0.25	3	3	150	169.41	49.79	0.13	0.44	0.25	0.01	0.43
min		20.00	0.25			150	43.89	1.11	-0.71	0.00	0.04	-0.84	0.00
max		105.00	0.25			150	353.88	96.57	1.36	1.36	0.59	1.35	1.35
50	50	114.05	0.25	3	3	155	170.20	58.87	0.10	0.31	0.25	-0.02	0.30
min		61.00	0.25			155	79.16	1.42	-0.49	0.00	0.06	-0.74	0.00
max		164.00	0.25			155	333.00	91.36	1.15	1.15	0.50	0.99	0.99
51	100	184.12	0.25	3	3	145	157.35	46.61	0.09	0.25	0.24	-0.04	0.24
min		126.00	0.25			145	89.02	2.38	-0.39	0.00	0.10	-0.62	0.00
max		243.00	0.25			145	309.56	70.09	1.13	1.13	0.42	0.69	0.69
52	150	242.50	0.25	3	3	149	150.71	47.66	0.01	0.16	0.25	0.02	0.17
min		207.00	0.25			149	95.12	14.26	-0.36	0.00	0.13	-0.47	0.00
max		283.00	0.25			149	246.83	67.71	0.66	0.66	0.43	0.73	0.73
53	20	90.48	0.5	3	3	158	159.68	59.11	0.01	0.15	0.50	0.00	0.13
min		23.00	0.5			158	93.55	4.98	-0.41	0.00	0.26	-0.49	0.00

max		151.00	0.5			158	266.90	92.16	0.69	0.69	0.69	0.37	0.49
54	50	189.56	0.5	3	3	156	158.40	46.61	0.02	0.08	0.49	-0.01	0.09
min		115.00	0.5			156	121.96	34.97	-0.22	0.00	0.37	-0.27	0.00
max		262.00	0.5			156	200.43	66.07	0.28	0.28	0.62	0.25	0.27
55	100	250.63	0.5	3	3	152	151.02	34.45	-0.01	0.06	0.50	0.01	0.07
min		195.00	0.5			152	125.13	25.39	-0.18	0.00	0.40	-0.19	0.00
max		296.00	0.5			152	182.88	45.45	0.20	0.20	0.60	0.21	0.21
56	150	289.41	0.5	3	3	143	142.42	29.62	0.00	0.06	0.50	0.01	0.07
min		248.00	0.5			143	119.67	21.75	-0.16	0.00	0.41	-0.18	0.00
max		324.00	0.5			143	164.73	39.03	0.15	0.16	0.63	0.26	0.26
57	20	95.52	0.75	3	3	146	145.78	52.47	0.00	0.11	0.75	0.00	0.07
min		31.00	0.75			146	98.82	36.20	-0.32	0.00	0.55	-0.26	0.00
max		178.00	0.75			146	199.13	73.41	0.36	0.36	0.93	0.24	0.26
58	50	212.04	0.75	3	3	157	157.90	35.79	0.01	0.05	0.75	0.00	0.04
min		141.00	0.75			157	131.95	25.08	-0.16	0.00	0.66	-0.12	0.00
max		271.00	0.75			157	182.80	55.53	0.16	0.16	0.83	0.10	0.12

59	100	272.15	0.75	3	3	150	150.40	25.15	0.00	0.04	0.75	0.00	0.04
min		223.00	0.75			150	135.26	16.55	-0.10	0.00	0.64	-0.14	0.00
max		313.00	0.75			150	168.91	34.67	0.13	0.13	0.83	0.11	0.14
60	150	333.11	0.75	3	3	166	166.00	14.60	0.00	0.03	0.75	0.00	0.03
min		296.00	0.75			166	153.41	9.35	-0.08	0.00	0.67	-0.10	0.00
max		356.00	0.75			166	179.74	20.99	0.08	0.08	0.82	0.09	0.10
61	20	65.52	0.25	5	3	153	158.12	53.23	0.03	0.24	0.25	-0.01	0.24
min		20.00	0.25			153	74.48	0.69	-0.51	0.00	0.08	-0.68	0.00
max		116.00	0.25			153	305.49	92.90	1.00	1.00	0.44	0.76	0.76
62	50	158.08	0.25	5	3	160	163.21	52.59	0.02	0.13	0.25	0.00	0.15
min		93.00	0.25			160	104.65	0.96	-0.35	0.00	0.12	-0.50	0.00
max		226.00	0.25			160	269.54	70.00	0.68	0.68	0.36	0.43	0.50
63	100	231.05	0.25	5	3	158	161.00	44.49	0.02	0.10	0.25	0.00	0.12
min		162.00	0.25			158	118.12	15.14	-0.25	0.00	0.15	-0.39	0.00
max		295.00	0.25			158	209.29	57.33	0.32	0.32	0.34	0.37	0.39
64	150	285.32	0.25	5	3	159	159.44	36.53	0.00	0.08	0.25	0.00	0.10

min		249.00	0.25			159	123.84	26.41	-0.22	0.00	0.18	-0.26	0.00
max		322.00	0.25			159	204.24	47.55	0.28	0.28	0.33	0.32	0.32
65	20	130.14	0.5	5	3	158	158.99	41.78	0.01	0.08	0.50	0.00	0.07
min		30.00	0.5			158	117.76	25.21	-0.25	0.00	0.38	-0.25	0.00
max		219.00	0.5			158	213.54	70.83	0.35	0.35	0.60	0.21	0.25
66	50	211.44	0.5	5	3	152	151.93	30.61	0.00	0.05	0.50	0.00	0.06
min		132.00	0.5			152	120.54	20.02	-0.21	0.00	0.42	-0.17	0.00
max		266.00	0.5			152	173.32	44.46	0.14	0.21	0.58	0.17	0.17
67	100	313.08	0.5	5	3	169	168.48	18.67	0.00	0.03	0.50	0.00	0.04
min		265.00	0.5			169	152.74	11.26	-0.10	0.00	0.44	-0.11	0.00
max		354.00	0.5			169	186.07	27.22	0.10	0.10	0.55	0.11	0.11
68	150	333.82	0.5	5	3	166	166.06	17.27	0.00	0.03	0.50	0.00	0.04
min		299.00	0.5			166	150.87	12.37	-0.09	0.00	0.43	-0.13	0.00
max		362.00	0.5			166	179.22	23.47	0.08	0.09	0.57	0.13	0.13
69	20	140.66	0.75	5	3	163	162.44	48.18	0.00	0.08	0.75	0.00	0.04
min		42.00	0.75			163	120.11	33.79	-0.26	0.00	0.63	-0.15	0.00

max		217.00	0.75			163	212.32	74.67	0.30	0.30	0.85	0.14	0.15
70	50	206.38	0.75	5	3	149	148.40	34.69	0.00	0.06	0.75	0.00	0.03
min		117.00	0.75			149	124.90	24.63	-0.16	0.00	0.68	-0.09	0.00
max		260.00	0.75			149	173.02	53.79	0.16	0.16	0.81	0.09	0.09
71	100	293.59	0.75	5	3	159	159.15	17.82	0.00	0.03	0.75	0.00	0.02
min		235.00	0.75			159	145.40	11.99	-0.09	0.00	0.70	-0.07	0.00
max		328.00	0.75			159	175.38	28.85	0.10	0.10	0.80	0.06	0.07
72	150	338.94	0.75	5	3	161	161.06	11.25	0.00	0.02	0.75	0.00	0.02
min		311.00	0.75			161	148.59	7.36	-0.08	0.00	0.70	-0.06	0.00
max		361.00	0.75			161	173.21	15.99	0.08	0.08	0.79	0.05	0.06

Table C.3. Simulation results from 500 synthetic datasets with different design criteria for the ACS model (adaptive-cluster sampling using the modified Horvitz-Thompson estimator). Habitat refers to the amount of generated spatial correlation where Habitat 1 has the most spatial correlation and Habitat 3 has no spatial correlation (see text for more details). Parameters are *p*-detection probability, *n*- initial sample size, and *J*-number of repeat visits. *True* N^{tot} is the actual number of occupied sites out of 400 while N^{tot}-hat is the estimated number of occupied sites averaged over the 500 synthetic datasets. *Var(Ntot)* represents the estimated variance of the estimated number of occupied sites out of the. *Coverage* represents the percent of confidence intervals that contained the true occupancy rate out of the 500 synthetic data sets. *RBIAS* and *RMSE* refer to relative bias and relative root mean-squared error, respectively. Note that one simulation trial encountered fatal errors and is not represented (n = 100, p = 0.75, J = 5, Habitat 2).

Scenario	*n*	*p*	*J*	*Habitat*	*True Ntot*	*Ntot-hat.HT*	*Var(Ntot.Ht)*	*Ntot.HT RBias*	*Ntot.HT RMSE*	*Coverage HT*
1	20	0.25	3	1	143	37.01	625.89	0.74	0.76	0.09
min		0.25			143	0.00	0.00	-0.43	0.01	0.00
max		0.25			143	204.92	1994.72	1.00	1.00	1.00
2	50	0.25	3	1	162	41.26	251.09	0.75	0.75	0.00
min		0.25			162	0.00	0.00	0.39	0.39	0.00
max		0.25			162	99.18	519.44	1.00	1.00	0.00
3	100	0.25	3	1	162	43.57	107.55	0.73	0.74	0.00
min		0.25			162	12.00	35.00	0.51	0.51	0.00
max		0.25			162	80.06	185.06	0.93	0.93	0.00
4	150	0.25	3	1	150	40.95	54.01	0.73	0.73	0.00
min		0.25			150	16.60	18.14	0.51	0.51	0.00
max		0.25			150	73.00	87.24	0.89	0.89	0.00
5	20	0.5	3	1	170	86.02	1207.47	0.49	0.53	0.36
min		0.5			170	0.00	0.00	-0.22	0.00	0.00

max		0.5			170	207.95	1995.20	1.00	1.00	1.00
6	50	0.5	3	1	159	85.59	413.73	0.46	0.48	0.09
min		0.5			159	29.19	154.69	-0.11	0.02	0.00
max		0.5			159	176.90	662.25	0.82	0.82	1.00
7	100	0.5	3	1	150	82.93	160.40	0.45	0.46	0.01
min		0.5			150	40.10	75.17	0.10	0.10	0.00
max		0.5			150	134.61	243.49	0.73	0.73	1.00
8	150	0.5	3	1	151	86.39	82.32	0.43	0.43	0.00
min		0.5			151	52.42	42.89	0.16	0.16	0.00
max		0.5			151	127.33	112.31	0.65	0.65	0.00
9	20	0.75	3	1	148	112.61	1405.34	0.24	0.35	0.83
min		0.75			148	0.00	0.00	-0.62	0.00	0.00
max		0.75			148	240.32	1990.26	1.00	1.00	1.00
10	50	0.75	3	1	170	138.42	471.05	0.19	0.23	0.67
min		0.75			170	78.01	174.18	-0.21	0.00	0.00
max		0.75			170	205.12	681.82	0.54	0.54	1.00

11	100	0.75	3	1	157	130.61	165.61	0.17	0.19	0.48
min		0.75			157	83.38	87.07	-0.09	0.00	0.00
max		0.75			157	170.84	238.51	0.47	0.47	1.00
12	150	0.75	3	1	155	129.08	90.79	0.17	0.18	0.26
min		0.75			155	96.31	60.02	0.00	0.00	0.00
max		0.75			155	155.40	119.96	0.38	0.38	1.00
13	20	0.25	5	1	148	38.88	660.02	0.74	0.76	0.10
min		0.25			148	0.00	0.00	0.02	0.02	0.00
max		0.25			148	145.44	1811.12	1.00	1.00	1.00
14	50	0.25	5	1	153	38.59	234.39	0.75	0.76	0.00
min		0.25			153	0.00	0.00	0.37	0.37	0.00
max		0.25			153	97.05	520.64	1.00	1.00	0.00
15	100	0.25	5	1	154	41.86	102.74	0.73	0.73	0.00
min		0.25			154	8.00	23.48	0.43	0.43	0.00
max		0.25			154	87.38	176.63	0.95	0.95	0.00
16	150	0.25	5	1	149	41.36	54.10	0.72	0.73	0.00

min		0.25			149	5.95	8.58	0.50	0.50	0.00
max		0.25			149	74.55	88.21	0.96	0.96	0.00
17	20	0.5	5	1	160	83.42	1197.19	0.48	0.53	0.37
min		0.5			160	0.00	0.00	-0.23	0.01	0.00
max		0.5			160	196.40	1975.70	1.00	1.00	1.00
18	50	0.5	5	1	141	74.36	374.85	0.47	0.49	0.11
min		0.5			141	19.27	103.55	0.09	0.09	0.00
max		0.5			141	128.15	570.94	0.86	0.86	1.00
19	100	0.5	5	1	152	83.85	162.13	0.45	0.46	0.01
min		0.5			152	39.49	77.09	0.08	0.08	0.00
max		0.5			152	140.24	246.13	0.74	0.74	1.00
20	150	0.5	5	1	157	91.17	85.98	0.42	0.43	0.00
min		0.5			157	57.39	50.84	0.15	0.15	0.00
max		0.5			157	133.05	121.57	0.63	0.63	0.00
21	20	0.75	5	1	162	128.83	1399.91	0.20	0.32	0.79
min		0.75			162	20.00	224.89	-0.70	0.00	0.00

max		0.75			162	276.01	1974.04	0.88	0.88	1.00
22	50	0.75	5	1	159	128.18	442.21	0.19	0.24	0.64
min		0.75			159	60.14	169.06	-0.14	0.00	0.00
max		0.75			159	182.03	637.82	0.62	0.62	1.00
23	100	0.75	5	1	140	115.67	159.46	0.17	0.20	0.53
min		0.75			140	65.84	76.80	-0.11	0.00	0.00
max		0.75			140	155.30	232.60	0.53	0.53	1.00
24	150	0.75	5	1	167	140.36	90.33	0.16	0.17	0.26
min		0.75			167	102.59	45.36	-0.08	0.00	0.00
max		0.75			167	181.18	128.90	0.39	0.39	1.00
25	20	0.25	3	2	141	36.78	626.43	0.74	0.76	0.09
min		0.25			141	0.00	0.00	0.12	0.12	0.00
max		0.25			141	123.46	1680.00	1.00	1.00	1.00
26	50	0.25	3	2	141	35.92	220.94	0.75	0.75	0.00
min		0.25			141	0.00	0.00	0.42	0.42	0.00
max		0.25			141	82.14	455.82	1.00	1.00	0.00

27	100	0.25	3	2	142	37.19	95.56	0.74	0.74	0.00
min		0.25			142	8.57	23.06	0.39	0.39	0.00
max		0.25			142	86.31	200.31	0.94	0.94	0.00
28	150	0.25	3	2	149	40.65	54.07	0.73	0.73	0.00
min		0.25			149	11.89	16.94	0.55	0.55	0.00
max		0.25			149	66.45	89.50	0.92	0.92	0.00
29	20	0.5	3	2	157	79.51	1173.42	0.49	0.54	0.33
min		0.5			157	0.00	0.00	-0.33	0.03	0.00
max		0.5			157	208.59	1990.71	1.00	1.00	1.00
30	50	0.5	3	2	134	68.72	374.97	0.49	0.51	0.15
min		0.5			134	24.00	153.41	-0.05	0.05	0.00
max		0.5			134	141.32	639.05	0.82	0.82	1.00
31	100	0.5	3	2	154	84.66	161.65	0.45	0.46	0.01
min		0.5			154	37.54	72.04	0.10	0.10	0.00
max		0.5			154	138.13	237.69	0.76	0.76	1.00
32	150	0.5	3	2	140	79.40	81.28	0.43	0.44	0.00

min		0.5			140	44.16	45.71	0.16	0.16	0.00
max		0.5			140	116.99	112.13	0.68	0.68	0.00
33	20	0.75	3	2	148	112.99	1388.93	0.24	0.36	0.80
min		0.75			148	20.00	308.96	-0.75	0.00	0.00
max		0.75			148	258.52	1973.73	0.86	0.86	1.00
34	50	0.75	3	2	144	112.90	492.01	0.22	0.27	0.68
min		0.75			144	44.54	237.22	-0.26	0.00	0.00
max		0.75			144	181.52	664.27	0.69	0.69	1.00
35	100	0.75	3	2	166	134.26	201.01	0.19	0.21	0.42
min		0.75			166	90.37	134.38	-0.07	0.00	0.00
max		0.75			166	177.05	266.86	0.46	0.46	1.00
36	150	0.75	3	2	140	116.58	85.66	0.17	0.18	0.30
min		0.75			140	71.35	55.24	-0.06	0.00	0.00
max		0.75			140	148.74	120.66	0.49	0.49	1.00
37	20	0.25	5	2	143	35.85	614.52	0.75	0.77	0.10
min		0.25			143	0.00	0.00	0.14	0.14	0.00

max		0.25			143	122.94	1680.00	1.00	1.00	1.00
38	50	0.25	5	2	139	35.86	220.75	0.74	0.75	0.00
min		0.25			139	0.00	0.00	0.40	0.40	0.00
max		0.25			139	83.23	456.86	1.00	1.00	0.00
39	100	0.25	5	2	144	37.33	93.67	0.74	0.74	0.00
min		0.25			144	8.00	21.86	0.50	0.50	0.00
max		0.25			144	71.43	161.91	0.94	0.94	0.00
40	150	0.25	5	2	129	34.28	47.97	0.73	0.74	0.00
min		0.25			129	12.58	15.89	0.50	0.50	0.00
max		0.25			129	65.08	85.97	0.90	0.90	0.00
41	20	0.5	5	2	138	71.03	1076.96	0.49	0.54	0.46
min		0.5			138	0.00	0.00	-0.22	0.03	0.00
max		0.5			138	168.50	1909.64	1.00	1.00	1.00
42	50	0.5	5	2	143	75.94	392.20	0.47	0.49	0.13
min		0.5			143	16.52	104.79	-0.01	0.01	0.00
max		0.5			143	144.66	637.42	0.88	0.88	1.00

43	100	0.5	5	2	164	88.45	170.18	0.46	0.47	0.01
min		0.5			164	39.05	70.98	0.17	0.17	0.00
max		0.5			164	136.84	243.24	0.76	0.76	1.00
44	150	0.5	5	2	138	74.81	83.26	0.46	0.47	0.00
min		0.5			138	43.74	48.52	0.19	0.19	0.00
max		0.5			138	111.43	118.31	0.68	0.68	0.00
45	20	0.75	5	2	148	116.48	1479.77	0.21	0.35	0.82
min		0.75			148	0.00	0.00	-0.57	0.00	0.00
max		0.75			148	232.94	1988.78	1.00	1.00	1.00
46	50	0.75	5	2	143	111.36	501.73	0.22	0.27	0.68
min		0.75			143	34.12	208.71	-0.21	0.00	0.00
max		0.75			143	172.49	665.08	0.76	0.76	1.00
47	100	0.75	5	2						
min		0.75								
max		0.75								
48	150	0.75	5	2	140	114.45	86.15	0.18	0.20	0.28

min		0.75			140	83.82	46.65	-0.06	0.00	0.00
max		0.75			140	148.45	123.13	0.40	0.40	1.00
49	20	0.25	3	3	150	36.76	625.78	0.75	0.78	0.10
min		0.25			150	0.00	0.00	0.18	0.18	0.00
max		0.25			150	122.49	1680.00	1.00	1.00	1.00
50	50	0.25	3	3	155	40.91	249.70	0.74	0.74	0.00
min		0.25			155	0.00	0.00	0.25	0.25	0.00
max		0.25			155	115.80	573.95	1.00	1.00	1.00
51	100	0.25	3	3	145	38.13	97.98	0.74	0.74	0.00
min		0.25			145	8.00	23.21	0.46	0.46	0.00
max		0.25			145	78.31	185.71	0.94	0.94	0.00
52	150	0.25	3	3	149	39.41	54.55	0.74	0.74	0.00
min		0.25			149	11.28	17.18	0.52	0.52	0.00
max		0.25			149	71.85	95.41	0.92	0.92	0.00
53	20	0.5	3	3	158	81.59	1205.85	0.48	0.53	0.35
min		0.5			158	0.00	0.00	-0.29	0.02	0.00

max		0.5			158	204.46	1995.19	1.00	1.00	1.00
54	50	0.5	3	3	156	81.84	425.58	0.48	0.50	0.10
min		0.5			156	25.57	157.46	0.01	0.01	0.00
max		0.5			156	153.95	653.99	0.84	0.84	1.00
55	100	0.5	3	3	152	82.11	173.01	0.46	0.47	0.00
min		0.5			152	32.98	75.79	0.15	0.15	0.00
max		0.5			152	128.56	246.46	0.78	0.78	1.00
56	150	0.5	3	3	143	76.88	88.69	0.46	0.47	0.00
min		0.5			143	39.17	50.10	0.22	0.22	0.00
max		0.5			143	112.03	120.91	0.73	0.73	0.00
57	20	0.75	3	3	146	110.30	1455.22	0.24	0.37	0.79
min		0.75			146	0.00	0.00	-0.86	0.00	0.00
max		0.75			146	271.46	1992.04	1.00	1.00	1.00
58	50	0.75	3	3	157	123.14	515.30	0.22	0.27	0.65
min		0.75			157	53.17	232.13	-0.25	0.00	0.00
max		0.75			157	196.50	685.68	0.66	0.66	1.00

59	100	0.75	3	3	150	119.53	208.70	0.20	0.23	0.46
min		0.75			150	75.04	123.75	-0.10	0.00	0.00
max		0.75			150	164.65	261.58	0.50	0.50	1.00
60	150	0.75	3	3	166	136.86	101.64	0.18	0.19	0.22
min		0.75			166	105.39	69.61	-0.09	0.00	0.00
max		0.75			166	180.72	133.14	0.37	0.37	1.00
61	20	0.25	5	3	153	37.99	645.87	0.75	0.77	0.04
min		0.25			153	0.00	0.00	0.21	0.21	0.00
max		0.25			153	121.47	1679.16	1.00	1.00	1.00
62	50	0.25	5	3	160	39.89	243.46	0.75	0.76	0.00
min		0.25			160	0.00	0.00	0.43	0.43	0.00
max		0.25			160	91.75	488.04	1.00	1.00	0.00
63	100	0.25	5	3	158	40.93	104.49	0.74	0.74	0.00
min		0.25			158	4.57	11.72	0.53	0.53	0.00
max		0.25			158	73.94	178.34	0.97	0.97	0.00
64	150	0.25	5	3	159	42.66	57.65	0.73	0.73	0.00

min		0.25			159	13.95	21.39	0.53	0.53	0.00
max		0.25			159	74.16	90.97	0.91	0.91	0.00
65	20	0.5	5	3	158	80.15	1187.20	0.49	0.54	0.32
min		0.5			158	0.00	0.00	-0.32	0.03	0.00
max		0.5			158	208.05	1996.28	1.00	1.00	1.00
66	50	0.5	5	3	152	80.11	419.22	0.47	0.49	0.10
min		0.5			152	25.07	159.58	-0.16	0.11	0.00
max		0.5			152	176.49	679.34	0.84	0.84	1.00
67	100	0.5	5	3	169	92.17	182.52	0.45	0.46	0.00
min		0.5			169	27.67	66.11	0.23	0.23	0.00
max		0.5			169	130.26	243.07	0.84	0.84	0.00
68	150	0.5	5	3	166	92.24	96.66	0.44	0.45	0.00
min		0.5			166	53.79	61.16	0.20	0.20	0.00
max		0.5			166	132.83	125.89	0.68	0.68	0.00
69	20	0.75	5	3	163	125.87	1551.92	0.23	0.33	0.77
min		0.75			163	20.00	367.97	-0.55	0.00	0.00

max		0.75			163	251.94	1994.16	0.88	0.88	1.00
70	50	0.75	5	3	149	113.75	505.45	0.24	0.29	0.63
min		0.75			149	50.97	244.54	-0.22	0.00	0.00
max		0.75			149	182.06	681.32	0.66	0.66	1.00
71	100	0.75	5	3	159	128.34	200.04	0.19	0.22	0.42
min		0.75			159	80.93	125.40	-0.11	0.00	0.00
max		0.75			159	177.02	265.59	0.49	0.49	1.00
72	150	0.75	5	3	161	134.61	89.27	0.16	0.18	0.25
min		0.75			161	104.05	46.84	-0.03	0.00	0.00
max		0.75			161	165.89	121.19	0.35	0.35	1.00

Table C.4. Simulation results from 500 synthetic datasets with different design criteria for the ACS model (adaptive-cluster sampling using the modified Hansen-Hurwitz estimator). Habitat refers to the amount of generated spatial correlation where Habitat 1 has the most spatial correlation and Habitat 3 has no spatial correlation (see text for more details). Parameters are *p*-detection probability, *n*- initial sample size, and *J*-number of repeat visits. *True* N^{tot} is the actual number of occupied sites out of 400 while N^{tot}-hat is the estimated number of occupied sites averaged over the 500 synthetic datasets. *Var(Ntot)* represents the estimated variance of the estimated number of occupied sites out of the. *Coverage* represents the percent of confidence intervals that contained the true occupancy rate out of the 500 synthetic data sets. *RBIAS* and *RMSE* refer to relative bias and relative root mean-squared error, respectively. Note that one simulation trial encountered fatal errors and is not represented (n = 100, p = 0.75, J = 5, Habitat 2).

Scenario	*n*	*p*	*J*	*Habitat*	*True Ntot*	*Ntot-hat.HH*	*Var(Ntot.HH)*	*Ntot.HH RBias*	*Ntot.HH RMSE*	*Coverage HH*
1	20	0.25	3	1	143	36.80	632.48	0.74	0.77	0.09
min		0.25			143	0.00	0.00	-0.54	0.02	0.00
max		0.25			143	220.00	1980.00	1.00	1.00	1.00
2	50	0.25	3	1	162	41.31	259.51	0.74	0.75	0.00
min		0.25			162	0.00	0.00	0.41	0.41	0.00
max		0.25			162	96.00	521.14	1.00	1.00	0.00
3	100	0.25	3	1	162	43.57	116.45	0.73	0.74	0.00
min		0.25			162	12.00	35.27	0.48	0.48	0.00
max		0.25			162	84.00	201.09	0.93	0.93	0.00
4	150	0.25	3	1	150	40.95	61.22	0.73	0.73	0.00
min		0.25			150	16.00	25.77	0.50	0.50	0.00
max		0.25			150	74.67	101.89	0.89	0.89	0.00
5	20	0.5	3	1	170	86.08	1287.44	0.49	0.54	0.43
min		0.5			170	0.00	0.00	-0.29	0.06	0.00

max		0.5			170	220.00	2000.00	1.00	1.00	1.00
6	50	0.5	3	1	159	85.60	470.92	0.46	0.48	0.17
min		0.5			159	32.00	210.29	-0.06	0.04	0.00
max		0.5			159	168.00	696.00	0.80	0.80	1.00
7	100	0.5	3	1	150	82.63	196.60	0.45	0.46	0.02
min		0.5			150	36.00	99.27	0.12	0.12	0.00
max		0.5			150	132.00	268.00	0.76	0.76	1.00
8	150	0.5	3	1	151	86.36	112.95	0.43	0.44	0.00
min		0.5			151	50.67	74.24	0.19	0.19	0.00
max		0.5			151	122.67	142.70	0.66	0.66	0.00
9	20	0.75	3	1	148	112.04	1533.32	0.24	0.36	0.87
min		0.75			148	0.00	0.00	-0.62	0.05	0.00
max		0.75			148	240.00	2000.00	1.00	1.00	1.00
10	50	0.75	3	1	170	138.75	634.63	0.18	0.24	0.69
min		0.75			170	72.00	421.71	-0.32	0.01	0.00
max		0.75			170	224.00	714.29	0.58	0.58	1.00

11	100	0.75	3	1	157	130.74	264.50	0.17	0.20	0.62
min		0.75			157	68.00	171.03	-0.10	0.01	0.00
max		0.75			157	172.00	297.09	0.57	0.57	1.00
12	150	0.75	3	1	155	128.79	145.83	0.17	0.19	0.41
min		0.75			155	90.67	117.64	-0.08	0.00	0.00
max		0.75			155	168.00	163.49	0.42	0.42	1.00
13	20	0.25	5	1	148	38.64	666.48	0.74	0.76	0.10
min		0.25			148	0.00	0.00	-0.22	0.19	0.00
max		0.25			148	180.00	1980.00	1.00	1.00	1.00
14	50	0.25	5	1	153	38.59	243.86	0.75	0.76	0.00
min		0.25			153	0.00	0.00	0.37	0.37	0.00
max		0.25			153	96.00	521.14	1.00	1.00	0.00
15	100	0.25	5	1	154	41.70	111.91	0.73	0.73	0.00
min		0.25			154	8.00	23.76	0.43	0.43	0.00
max		0.25			154	88.00	208.00	0.95	0.95	0.00
16	150	0.25	5	1	149	41.55	61.98	0.72	0.72	0.00

min		0.25			149	8.00	13.15	0.43	0.43	0.00
max		0.25			149	85.33	112.63	0.95	0.95	0.00
17	20	0.5	5	1	160	83.36	1250.40	0.48	0.53	0.41
min		0.5			160	0.00	0.00	-0.25	0.00	0.00
max		0.5			160	200.00	2000.00	1.00	1.00	1.00
18	50	0.5	5	1	141	74.54	424.48	0.47	0.50	0.13
min		0.5			141	16.00	109.71	-0.13	0.04	0.00
max		0.5			141	160.00	685.71	0.89	0.89	1.00
19	100	0.5	5	1	152	83.82	198.59	0.45	0.46	0.01
min		0.5			152	40.00	109.09	0.05	0.05	0.00
max		0.5			152	144.00	279.27	0.74	0.74	1.00
20	150	0.5	5	1	157	90.55	116.76	0.42	0.43	0.00
min		0.5			157	56.00	80.81	0.18	0.18	0.00
max		0.5			157	128.00	146.04	0.64	0.64	0.00
21	20	0.75	5	1	162	130.96	1668.00	0.19	0.33	0.83
min		0.75			162	20.00	380.00	-0.73	0.01	0.00

max		0.75			162	280.00	2000.00	0.88	0.88	1.00
22	50	0.75	5	1	159	127.76	609.57	0.20	0.25	0.79
min		0.75			159	56.00	344.00	-0.31	0.01	0.00
max		0.75			159	208.00	714.29	0.65	0.65	1.00
23	100	0.75	5	1	140	114.48	245.48	0.18	0.22	0.60
min		0.75			140	68.00	171.03	-0.26	0.00	0.00
max		0.75			140	176.00	298.67	0.51	0.51	1.00
24	150	0.75	5	1	167	140.46	152.11	0.16	0.18	0.45
min		0.75			167	101.33	126.95	-0.09	0.01	0.00
max		0.75			167	181.33	166.32	0.39	0.39	1.00
25	20	0.25	3	2	141	36.64	632.24	0.74	0.76	0.10
min		0.25			141	0.00	0.00	0.15	0.15	0.00
max		0.25			141	120.00	1680.00	1.00	1.00	1.00
26	50	0.25	3	2	141	35.82	228.54	0.75	0.75	0.00
min		0.25			141	0.00	0.00	0.43	0.43	0.00
max		0.25			141	80.00	457.14	1.00	1.00	0.00

27	100	0.25	3	2	142	37.07	100.86	0.74	0.74	0.00
min		0.25			142	8.00	23.76	0.41	0.41	0.00
max		0.25			142	84.00	201.09	0.94	0.94	0.00
28	150	0.25	3	2	149	40.68	60.91	0.73	0.73	0.00
min		0.25			149	10.67	17.42	0.55	0.55	0.00
max		0.25			149	66.67	93.21	0.93	0.93	0.00
29	20	0.5	3	2	157	79.56	1212.68	0.49	0.54	0.36
min		0.5			157	0.00	0.00	-0.27	0.02	0.00
max		0.5			157	200.00	2000.00	1.00	1.00	1.00
30	50	0.5	3	2	134	68.40	397.24	0.49	0.51	0.13
min		0.5			134	24.00	161.14	-0.07	0.04	0.00
max		0.5			134	144.00	658.29	0.82	0.82	1.00
31	100	0.5	3	2	154	84.80	200.36	0.45	0.46	0.02
min		0.5			154	36.00	99.27	0.09	0.09	0.00
max		0.5			154	140.00	275.76	0.77	0.77	1.00
32	150	0.5	3	2	140	78.98	105.74	0.44	0.44	0.00

min		0.5			140	37.33	56.79	0.14	0.14	0.00
max		0.5			140	120.00	140.94	0.73	0.73	1.00
33	20	0.75	3	2	148	112.04	1523.80	0.24	0.37	0.84
min		0.75			148	20.00	380.00	-0.76	0.05	0.00
max		0.75			148	260.00	2000.00	0.86	0.86	1.00
34	50	0.75	3	2	144	113.02	568.12	0.22	0.28	0.69
min		0.75			144	40.00	257.14	-0.28	0.00	0.00
max		0.75			144	184.00	709.71	0.72	0.72	1.00
35	100	0.75	3	2	166	133.73	267.38	0.19	0.22	0.49
min		0.75			166	88.00	208.00	-0.11	0.01	0.00
max		0.75			166	184.00	301.09	0.47	0.47	1.00
36	150	0.75	3	2	140	116.70	137.91	0.17	0.19	0.53
min		0.75			140	69.33	96.17	-0.09	0.01	0.00
max		0.75			140	152.00	158.12	0.50	0.50	1.00
37	20	0.25	5	2	143	35.76	621.60	0.75	0.77	0.10
min		0.25			143	0.00	0.00	0.16	0.16	0.00

max		0.25			143	120.00	1680.00	1.00	1.00	1.00
38	50	0.25	5	2	139	35.86	228.23	0.74	0.75	0.00
min		0.25			139	0.00	0.00	0.37	0.37	0.00
max		0.25			139	88.00	490.29	1.00	1.00	0.00
39	100	0.25	5	2	144	37.54	102.04	0.74	0.74	0.00
min		0.25			144	8.00	23.76	0.47	0.47	0.00
max		0.25			144	76.00	186.55	0.94	0.94	0.00
40	150	0.25	5	2	129	34.12	52.01	0.74	0.74	0.00
min		0.25			129	10.67	17.42	0.48	0.48	0.00
max		0.25			129	66.67	93.21	0.92	0.92	0.00
41	20	0.5	5	2	138	71.04	1116.08	0.49	0.54	0.49
min		0.5			138	0.00	0.00	-0.30	0.01	0.00
max		0.5			138	180.00	1980.00	1.00	1.00	1.00
42	50	0.5	5	2	143	75.82	430.16	0.47	0.49	0.15
min		0.5			143	16.00	109.71	-0.01	0.01	0.00
max		0.5			143	144.00	658.29	0.89	0.89	1.00

43	100	0.5	5	2	164	88.21	205.91	0.46	0.47	0.02
min		0.5			164	40.00	109.09	0.07	0.07	0.00
max		0.5			164	152.00	285.58	0.76	0.76	1.00
44	150	0.5	5	2	138	74.56	101.13	0.46	0.47	0.00
min		0.5			138	45.33	67.44	0.15	0.15	0.00
max		0.5			138	117.33	139.12	0.67	0.67	1.00
45	20	0.75	5	2	148	116.16	1564.24	0.22	0.35	0.85
min		0.75			148	0.00	0.00	-0.76	0.05	0.00
max		0.75			148	260.00	2000.00	1.00	1.00	1.00
46	50	0.75	5	2	143	110.64	561.89	0.23	0.28	0.68
min		0.75			143	32.00	210.29	-0.29	0.01	0.00
max		0.75			143	184.00	709.71	0.78	0.78	1.00
47	100	0.75	5	2						
min		0.75								
max		0.75								
48	150	0.75	5	2	140	113.84	135.90	0.19	0.21	0.44

min		0.75			140	77.33	104.67	-0.16	0.01	0.00
max		0.75			140	162.67	161.94	0.45	0.45	1.00
49	20	0.25	3	3	150	36.64	631.28	0.76	0.78	0.11
min		0.25			150	0.00	0.00	0.20	0.20	0.00
max		0.25			150	120.00	1680.00	1.00	1.00	1.00
50	50	0.25	3	3	155	40.94	257.07	0.74	0.74	0.00
min		0.25			155	0.00	0.00	0.28	0.28	0.00
max		0.25			155	112.00	576.00	1.00	1.00	1.00
51	100	0.25	3	3	145	38.18	103.64	0.74	0.74	0.00
min		0.25			145	8.00	23.76	0.48	0.48	0.00
max		0.25			145	76.00	186.55	0.94	0.94	0.00
52	150	0.25	3	3	149	39.19	58.88	0.74	0.74	0.00
min		0.25			149	10.67	17.42	0.50	0.50	0.00
max		0.25			149	74.67	101.89	0.93	0.93	0.00
53	20	0.5	3	3	158	81.72	1236.12	0.48	0.53	0.38
min		0.5			158	0.00	0.00	-0.27	0.01	0.00

max		0.5			158	200.00	2000.00	1.00	1.00	1.00
54	50	0.5	3	3	156	81.98	456.62	0.47	0.50	0.12
min		0.5			156	24.00	161.14	-0.13	0.08	0.00
max		0.5			156	176.00	704.00	0.85	0.85	1.00
55	100	0.5	3	3	152	81.74	195.04	0.46	0.47	0.00
min		0.5			152	28.00	78.91	0.18	0.18	0.00
max		0.5			152	124.00	259.27	0.82	0.82	1.00
56	150	0.5	3	3	143	76.42	103.01	0.47	0.47	0.00
min		0.5			143	37.33	56.79	0.20	0.20	0.00
max		0.5			143	114.67	137.24	0.74	0.74	0.00
57	20	0.75	3	3	146	110.44	1513.88	0.24	0.37	0.84
min		0.75			146	0.00	0.00	-0.78	0.04	0.00
max		0.75			146	260.00	2000.00	1.00	1.00	1.00
58	50	0.75	3	3	157	122.48	594.27	0.22	0.28	0.68
min		0.75			157	56.00	344.00	-0.38	0.02	0.00
max		0.75			157	216.00	713.14	0.64	0.64	1.00

59	100	0.75	3	3	150	118.55	250.48	0.21	0.24	0.50
min		0.75			150	72.00	178.91	-0.12	0.01	0.00
max		0.75			150	168.00	295.27	0.52	0.52	1.00
60	150	0.75	3	3	166	136.68	150.11	0.18	0.20	0.33
min		0.75			166	98.67	124.71	-0.09	0.00	0.00
max		0.75			166	181.33	166.32	0.41	0.41	1.00
61	20	0.25	5	3	153	37.88	650.44	0.75	0.77	0.04
min		0.25			153	0.00	0.00	0.22	0.22	0.00
max		0.25			153	120.00	1680.00	1.00	1.00	1.00
62	50	0.25	5	3	160	39.95	251.37	0.75	0.76	0.00
min		0.25			160	0.00	0.00	0.40	0.40	0.00
max		0.25			160	96.00	521.14	1.00	1.00	0.00
63	100	0.25	5	3	158	40.99	110.35	0.74	0.74	0.00
min		0.25			158	4.00	12.00	0.52	0.52	0.00
max		0.25			158	76.00	186.55	0.97	0.97	0.00
64	150	0.25	5	3	159	42.60	63.42	0.73	0.73	0.00

min		0.25			159	13.33	21.63	0.55	0.55	0.00
max		0.25			159	72.00	99.06	0.92	0.92	0.00
65	20	0.5	5	3	158	79.68	1215.84	0.50	0.54	0.36
min		0.5			158	0.00	0.00	-0.27	0.01	0.00
max		0.5			158	200.00	2000.00	1.00	1.00	1.00
66	50	0.5	5	3	152	80.18	449.34	0.47	0.49	0.11
min		0.5			152	24.00	161.14	-0.16	0.11	0.00
max		0.5			152	176.00	704.00	0.84	0.84	1.00
67	100	0.5	5	3	169	91.62	211.94	0.46	0.47	0.01
min		0.5			169	24.00	68.36	0.17	0.17	0.00
max		0.5			169	140.00	275.76	0.86	0.86	1.00
68	150	0.5	5	3	166	92.63	118.64	0.44	0.45	0.00
min		0.5			166	50.67	74.24	0.12	0.12	0.00
max		0.5			166	146.67	155.85	0.69	0.69	1.00
69	20	0.75	5	3	163	126.44	1646.52	0.22	0.34	0.81
min		0.75			163	20.00	380.00	-0.72	0.02	0.00

max		0.75			163	280.00	2000.00	0.88	0.88	1.00
70	50	0.75	5	3	149	113.89	570.99	0.24	0.29	0.71
min		0.75			149	56.00	344.00	-0.23	0.02	0.00
max		0.75			149	184.00	709.71	0.62	0.62	1.00
71	100	0.75	5	3	159	127.88	261.27	0.20	0.23	0.52
min		0.75			159	72.00	178.91	-0.16	0.01	0.00
max		0.75			159	184.00	301.09	0.55	0.55	1.00
72	150	0.75	5	3	161	134.51	149.05	0.16	0.18	0.42
min		0.75			161	90.67	117.64	-0.08	0.01	0.00
max		0.75			161	173.33	164.80	0.44	0.44	1.00