HIERARCHICAL MODELS OF AVIAN DISTRIBUTIONS FOR THE STATE OF GEORGIA

by

JAY EARL HOWELL

(Under the Direction of Michael Conroy)

ABSTRACT

In order to predict the distributions of nine breeding birds in the state of Georgia, I built models consisting of four hierarchical levels built on nested mapping units of decreasing area. The models were fit to the hierarchical logistic regression model using MCMC through the program WinBugs. Results of three fold cross validation showed an average overall correct classification rate of 72%. I tested the models in two management scenarios, choosing between three possible reserves (coarse) and recommending an optimal size for wildlife openings (fine). I evaluated model performance by their ability to differentiate between alternatives despite model uncertainty. We tested the impact of using AIC or DIC and of estimation based on model averaging. The models were able to choose a clear alternative in the coarse scenario, but not the fine one. Using AIC versus DIC or model averaging had no impact on the ranking of alternatives.

INDEX WORDS: Hierarchical Models, Wildlife Habitat Relationships, MCMC, Model Uncertainty, Georgia, Birds, AIC, DIC
HIERARCHICAL MODELS OF AVIAN DISTRIBUTIONS FOR THE STATE OF GEORGIA

by

JAY EARL HOWELL

B.S., The University of Tennessee, Knoxville, 1999

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

MASTER OF SCIENCE

ATHENS, GEORGIA

2003
HIERARCHICAL MODELS OF AVIAN DISTRIBUTIONS FOR THE STATE OF GEORGIA

by

JAY EARL HOWELL

Major Professor: Michael Conroy
Committee: James Peterson
            Robert Cooper

Electronic Version Approved:

Maureen Grasso
Dean of the Graduate School
The University of Georgia
December 2003
ACKNOWLEDGMENTS

I acknowledge the financial support of the Georgia Department of Natural Resources, Wildlife Resources Division, The National GAP Analysis Project, The Georgia GAP Analysis Project, and the United States Geological Survey. I thank the Georgia Partners in Flight cooperators for the use of their data, particularly Nathan Klauss who assisted me in obtaining the latest point counts. I thank the University of Georgia Institute of Ecology, Natural Resources and Spatial Analysis Laboratory for the use of their 1998 landcover and facilities for GIS work. I would especially like to thank Matt Elliot for all of his help with using GIS to measure my habitat variables. The Warnell School of Forest Resources and the USGS Cooperative Fish and Wildlife Unit at the University of Georgia provided valuable resources as well.

I would like to thank Dr. Michael Conroy for serving as my major professor and for providing guidance and support throughout the long process of developing my thesis. I also thank Dr. James Peterson for serving on my committee, providing technical support on the hierarchical models, and last minute financial support. Thanks to Dr. Robert Cooper for serving on my committee. Thanks to Clint Moore for help developing my SAS programs. Thanks to the other students in the Conroy lab who provided support and advice and didn’t complain when my model runs tied up computers for months at a time.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>ACKNOWLEDGMENTS</th>
<th>iv</th>
</tr>
</thead>
<tbody>
<tr>
<td>CHAPTER</td>
<td></td>
</tr>
<tr>
<td>1 INTRODUCTION AND LITERATURE REVIEW</td>
<td>1</td>
</tr>
<tr>
<td>2 HIERARCHICAL MODELS OF AVIAN DISTRIBUTIONS FOR THE STATE OF GEORGIA: MODEL BUILDING AND TESTING</td>
<td>25</td>
</tr>
<tr>
<td>3 HIERARCHICAL MODELS OF AVIAN DISTRIBUTIONS FOR THE STATE OF GEORGIA: MODEL APPLICATION</td>
<td>62</td>
</tr>
<tr>
<td>4 SUMMARY</td>
<td>90</td>
</tr>
</tbody>
</table>
CHAPTER 1

INTRODUCTION AND LITERATURE REVIEW

Biologists have long sought to understand the interactions between wildlife and the habitats in which they occur. From the earliest naturalists to modern day wildlife studies, we have sought to describe the locations where animals occur based on a wide range of possible descriptors (Morrison et al. 1998). In the mid seventies and early eighties, wildlife biologists began building models of these relationships with an intention to use these relationships to guide land management (Verner et al. 1984). Since then, while techniques have advanced, the goal has remained the same; that is, advancing understanding of these relationships and using that understanding to guide conservation.

Wildlife habitat relationship models generally fall into two classes, qualitative descriptions of suitable habitat based on the knowledge of the model builders and empirical models which relate the occurrence of a given species in a location to some vector of measurements taken on that location through a variety of statistical means. The Habitat Suitability Index (HSI) models (U.S. Fish and Wildlife Service 1981), the habitat matrices developed by Hamel (1992), and the pattern recognition (PATREC) models (Williams et al. 1978) are examples of the first sort of models. Some of these sorts of models, while useful in some ways, have been shown to perform poorly in several tests (Giles 1978, Lancia et al. 1982, Cole and Smith 1983). That is why most recent efforts have focused on developing and testing the techniques used to build the empirical models.
There are a wide variety of techniques used to estimate the relationships between species and habitat vectors. The most common empirical approaches involve relating presence and absence using multivariate techniques such as logistic regression, discriminant function analysis, or principal components analysis (Capen et al. 1984, Pereira and Itami 1991, Austin et al. 1996, Oeszemi and Mitsch 1997, Carroll et al. 1999, Manel et al. 1999). Other studies have developed techniques like Mahalanobis Distance (Clark et al. 1993) and Cumulative Distribution Functions (Detmers and Bart 1999) that are specific to predicting relationships when only presence data is available. Finally, cutting edge techniques that have imported from the artificial intelligence field are being used including Neural Networks (Manel et al. 1999) and Genetic Algorithms (Peterson and Kluza 2003).

GAP ANALYSIS

The application of wildlife habitat relationship modeling to landscape ecology was initially limited due to the logistic difficulties involved in collecting information on a landscape scale, which made such studies cost prohibitive (Urban 2000). With the advances in GIS and remote sensing, it became possible for ecologists to craft cost effective landscape approaches to WHR modeling. Scott et al. (1993) outlined a program to take advantage of this new capability on a nationwide basis called the GAP Analysis Project.

GAP Analysis identifies the gaps in representation of biodiversity in areas managed exclusively or primarily for the long-term conservation of native species and natural ecosystems (Scott et al. 1993). The national GAP office is funded from several sources including the United States Geological Survey and the Nature Conservancy (National GAP 2000). Below the national level, each state has its own individual project that completes the GAP analysis for that state.
(National GAP 2000). GAP analysis produces a map that consists of three layers (Scott et al. 1993). The first layer is a map of the vegetation communities derived from satellite imagery (Stoms 1998). The second layer uses the first layer in conjunction with a wildlife habitat relationship matrix to predict the distribution of every native vertebrate that occurs in the region (Csuti and Crist 1998). The third layer is a map of lands that have been designated as being set aside specifically for conservation (Edwards et al. 1998). These layers are then compared to determine which areas that have high species diversity are not protected for conservation (Scott et al. 1993, Jennings 2000). Based on these so-called gaps in what is called a “coarse filter” approach, managers can make optimal planning decisions in terms of land acquisition and finer scale studies (Scott et al. 1993, Kiester et al. 1996). One of the critical links in this process is the vertebrate distribution layer.

The vertebrate distribution layer is derived in two main steps (Scott et al. 1993, Csuti and Crist 1998). First the range of each species is mapped using occurrence records (Scott et al. 1993, Csuti and Crist 1998). Second, the wildlife habitat matrix is created through an extensive literature review that delineates habitats within the vegetation map where the species is considered present (Scott et al. 1993, Csuti and Crist 1998). The range maps and the matrix may then be reviewed by experts in each state (Csuti and Crist 1998). In some cases individual projects may use additional constraints to further refine their models such as minimum critical area (Pearlstine et al. 2002). The matrix and any additional constraints are then applied throughout the species’ predicted range (Scott et al. 1993, Csuti and Crist 1998).

Confidence in the vertebrate distribution map may be limited in that the relationships used to predict the final distributions are not tied to any specific data (Mannan et al. 1996). Although some GAP models have performed relatively well (Edwards et al. 1996), expert
opinion models similar to those employed by GAP historically did not perform well. Further, these models are unable to quantitatively reflect the relationship between variables or provide statistical measures of uncertainty.

One of the goals of landscape WHR modeling is to drive decision-making (Verner et al. 1984, Buckland and Elston 1993, Bradbury et al. 2000). When GAP is discussed in terms of applications, it is usually couched in “coarse filter” terms (Scott et al. 1993, Root et al. 2003). That is, it is assumed that GAP will operate at coarse spatial scales, but not at fine scales (O’Neil et al. 1997, Hansen et al. 1999). Use of GAP predictions as a coarse filter assumes predictive continuity among all levels. That is, because GAP models predict at scales that differ from the problems to which they are applied, wildlife relationships to habitat patterns must be assumed to hold between scales (Conroy and Noon 1996, Bolger et al. 1997). This is potentially troubling because patterns in landscape processes have been shown to vary across scales (Levin 1993). Further, it remains unclear why applications at the finer scales are untenable, given that the models predict at those scales and should therefore be able to reflect the impacts of habitat change at this level.

Therefore a set of empirical models that predict presence and absence based on the same types of habitat inputs available to GAP Analysis and that explicitly include hierarchical scales of information may provide an alternative approach to vertebrate distribution mapping which addresses these concerns. Further, I feel it would be beneficial to test these models using management applications at various scales to test their ability to function at and across various levels.
Habitat selection has been regularly characterized based on local effects; however, many have proposed that surrounding habitats should also be considered (Wiens 1989, Pearson 1993, Saab 1999, Penhollow and Stauffer 2000). Indeed, the strength of biogeographical patterns or priority of conservation selection may be related to the spatial scale of the investigation as well as processes within the landscape (Stoms 1994). It has further been suggested that habitat selection may be viewed as a hierarchical process (Schaeffer and Messier 1995). How an organism responds to an individual patch requires explicit recognition of the hierarchical structure of nested patches (Kotliar and Wiens 1990). Specification of a reliable spatial model requires selection of a model structure appropriate with regard to spatial scale, level of resolution, and the specific form of functional relationships (Conroy et al. 1995). It is therefore necessary to choose a model structure that implicitly models these hierarchical relationships.

A hierarchical data structure is one that has data describing individuals that may be logically grouped as well as data that describes those groups (Bryk and Raudenbush 1998). This structure can be created through the use of nested habitat blocks of decreasing area. Hierarchical regression models are useful whenever there is covariate information at multiple levels of variation (Gelman et al. 1995). In addition to explicitly modeling multiple scales, the hierarchical model improves estimation of individual effects by borrowing from the ensemble of data (Bryk and Raudenbush 1998). Hierarchical logistic regression seems the most appropriate tool when dealing with binary outcomes and multiple covariates (Wong and Mason 1985). The parameters of the hierarchical logistic regression model can be difficult to estimate using traditional techniques (Wong and Mason 1985).
BAYESIAN ESTIMATION USING MCMC

Bayesian Estimation has many advantages that contribute to our goals for this project. The Bayesian paradigm implicitly recognizes uncertainty about model parameters and outputs, allowing us to explore the impact of uncertainty on decision-making. Further, MCMC can be a useful tool for estimating models with complex structures and parameters resistant to standard analysis (Link et al. 2002). Hierarchical logistic regression includes fixed effects as well as random group level effects leading to the so-called mixed model (Wong and Mason 1985). MCMC provides a straightforward method for dealing with these random effects (Gelman et al. 1995, Congdon 2001). Under a Bayesian analysis, any parameter \( \theta \) is regarded as a random variable (Gelman et al. 1995). Inference about \( \theta \) referenced by data \( Y \) is then based on the posterior distribution \( f(\theta|Y) \), obtained by applying Bayes’ theorem to the prior \( \pi(\theta) \), or what is known about \( \theta \) without reference to the data, and the sampling distribution \( f(Y|\theta) \) (Link et al. 2002), using the equation:

\[
f(\theta | Y) = f(Y | \theta) \pi(\theta) / \int f(Y | \theta) \pi(\theta) d\theta .
\]  

However, for many models calculation of the posterior distribution is prohibitively difficult due the intractability of the integral in (1) (Link et al. 2002). Monte Carlo simulation is often used to calculate posterior distributions in place of intractable mathematical calculation by supposing some function \( g(\theta) \) of model parameters that can be expressed as the expected value of a function of the data \( h(Y) \) (Link et al. 2002) through the equation:

\[
g(\theta) = \frac{1}{N} \sum_{i=1}^{N} h(Y_i) .
\]
Under a Bayesian analysis one would need to take independent samples from the posterior
distribution \( f(\theta | Y) \), which is once again difficult due to the integral in equation 1. A first order
Markov Chain is a sequence of dependent observations with the property that the distribution of
any observation, given all previous observations, is dependent only upon the observation directly
before it (Link et al. 2002). One can approximate the posterior distribution by generating a first
order Markov chain of dependent draws \( \theta^*_i \) from the posterior using the Metropolis-Hasting
algorithm and thereby approximate the features of that distribution using the equation:

\[
G(Y) = 1/N \sum_{i=M+1}^{N+M} H(\theta^*_i),
\]

where the first \( M \) values of the Markov chain are discarded (Link et al. 2002).

DATA SOURCES

Data for this project came from two principle sources. Habitat measures were derived
from an 18-class land cover map generated by the Georgia GAP Analysis Project (Natural
Resources Spatial Analysis Laboratory, 2001). This map consisted of classified 30x30 m Landsat
TM imagery from 1996-1998 delineated into 18 major land cover classes. For a more complete
description of classification methods see Natural Resources Spatial Analysis Laboratory 18 class
land cover Map of Georgia metadata. Additional GIS data sources were used, including
coverages of streams and ecoregions also generated by the Georgia GAP Analysis Project.
Occurrence records were taken from the Georgia Partners in Flight (PIF) Point Count Database.
PIF is a cooperative effort among federal, state and local government agencies, philanthropic
organizations, professional organizations, conservation groups, industry, the academic
community and private individuals created in response to concerns about decline among many
land bird species (PIF 2003). The goal of PIF is to focus resources from multiple agencies on the
improvement of monitoring and inventory, research, management, and education involving birds and their habitats (PIF 2003). The database is a combination of point counts taken by PIF cooperators through out Georgia from 1995-2000.

STUDY SPECIES

Nine avian species were chosen for this modeling effort. They were chosen to represent a cross section of preferred habitat types, relative frequency, migratory patterns, and PIF concern. These species were the Acadian Flycatcher (*Empidonax virescens*), Brown-Headed Nuthatch (*Sitta pusilla*), Carolina Wren (*Thryothorus ludovicianus*), Indigo Bunting (*Passerina cyanea*), Northern Cardinal (*Cardinalis cardinalis*), Prairie Warbler (*Dendroica discolor*), White-eyed Vireo (*Vireo griseus*), Wood Thrush (*Hylocichla mustelina*), and Yellow-billed Cuckoo (*Coccyzus americanus*).

Acadian Flycatchers (ACFL) are common breeders throughout Georgia in the summer (Greene et al. 1945). Most Acadians depart for eastern Colombia and Ecuador between August-September and then return April-May (Bent 1942, Mumford 1964, Walkinshaw 1966). Acadians feed mostly on insects though hawking or by gleaning from leaves and trunks while in flight (Bent 1942, Mumford 1964). Acadian flycatcher may be found in deciduous and mixed forests including heavily wooded bottomlands, wooded ravines, cypress swamps, and forested waterways (Bent 1942, Mumford 1964, Walkinshaw 1966, Hespenheide 1971).

The Brown-Headed Nuthatch (BHNU) is a common permanent resident throughout the state of Georgia except in northeastern counties (Greene et al. 1945). It is one of the few cooperative breeders native to North America (Withgott and Smith 1998). Males generally choose cavities for nesting in Feb-May (Withgott and Smith 1998). The nuthatch is primarily
insectivorous and forages through bark picking; however, in winter it may switch to pine seeds (Bent 1948). They are also tool users, using chips of bark to pry off other bark chips while foraging (Morse 1968, Pranty 1995, Withgott and Smith 1998). The nuthatch is declining throughout its range, yet it remains common in pine habitats (Withgott and Smith 1998). It occurs primarily in association with pine trees, inhabiting a wide variety of southeastern pine forest habitats (Withgott and Smith 1998). It prefers open mature pinewoods, including burned-over areas (Bent 1948, Hamel 1992, Harrop and Quin 1995, Withgott and Smith 1998). It can also be found in cypress swamps and pine forest edges (Bent 1948, Harrop and Quin 1995). It also may occur in parks and residential areas where large pines are available (Withgott and Smith 1998). It is more abundant in older pine stands than younger ones (Johnston and Odum 1956, Johnson and Landers 1982).

The Carolina Wren (CARW) is a common permanent resident distributed throughout the state of Georgia (Greene et al. 1945). The wren forages on or near the ground, in leaf litter, around downed trees, upturned roots and vegetation tangles primarily on insects and seeds (Bent 1948, Strain and Mumme 1988, Haggerty and Morton 1995). BBS data show the Carolina Wren significantly increasing in abundance in the state of Georgia (Haggerty and Morton 1995). It inhabits a wide range of habitats so long as they include moderate to dense shrub or bushy cover (Bent 1948, James 1971, Conner et al. 1983, Haggerty and Morton 1995). It is also found in wooded residential areas with shrubs (Beissinger and Osbourne 1982). It prefers moist bottomland woods to dry upland ones (Bent 1948, James and Neal 1986).

The Indigo Bunting (INBU) breeds throughout the state in the summer, but is uncommon on the coast (Greene et al. 1945). Buntings leave the North American breeding grounds in late September and October for the Neotropics from which they return in late April and May (Taber
and Johnston 1968, Payne 1991). It feeds on small insects, spiders, seeds, and berries on the
ground or in low scrub (Payne 1992). The Indigo Bunting population has been increasing or
remaining the same throughout its range (Payne 1992). The Indigo Bunting prefers brushy and
weedy habitats along the edges of woods and cultivated lands, including roads, utility right of
ways, railway siding, and riparian habitats (Payne 1992). It is generally absent from closed
canopy forest, residential areas and intensively cultivated or grazed areas (Payne 1992).

The Northern Cardinal (NOCA) is a very common year round resident found throughout
Georgia (Greene et al. 1945). The cardinal is an opportunistic feeder foraging on the ground for
seeds an insects in open areas while in closed canopies it eats buds and insect larvae on trees and
shrubs (Dow 1969, Halkin and Linville 1999). Population density has increased throughout the
United States due to natural and human induced changes in habitat (Halkin and Linville 1999).
The cardinal occurs in habitats throughout its range in areas with shrubs or small trees including
forest edges, interior shrubby areas within second growth forest, marsh edges, riparian forest,
grasslands with shrubs, old fields, agricultural hedgerows, and landscaping around buildings

The Prairie Warbler (PRAW) is fairly common in Georgia during the summer, especially
in coastal counties (Greene et al. 1945). They are medium distance migrants that head south for
the Caribbean and Florida in mid to late summer and return beginning in March (Raffaele et al.
1998, Nolan et al. 1999). It feeds primarily on insects, spiders, small soft-bodied arthropods, and
mollusks (Nolan et al. 1999). This species is often cited as an example of Neotropical Migrant
decline, with decreases in a majority of the physiographic regions in the Breeding Bird Survey
between 1966 and 1996 (Nolan et al. 1999). The Prairie Warbler occurs in various shrubby areas
lacking closed canopies (Nolan et al. 1999). It can be found most often in open southern pine
forest, dunes, abandoned field and orchards, and grassland forest edges (Griscom 1979, Schorre 1998, Nolan et al. 1999).

The White-eyed Vireo (WEVI) is common throughout Georgia in the summer and found rarely in southern parts during the winter (Greene et al. 1945). It migrates in October to wintering grounds in Central America and the southeastern U.S. and returns in mid March (Andrele 1966, Hopp et al. 1995). It feeds primarily on insects in deciduous medium height growth (Bent 1950, Noland and Wooldridge 1962, Hopp et al. 1995). Population trends for the White-eyed vireo based on the Breeding Bird Survey show declines between 1966 and 1988 (Sauer and Droege 1992). It can be found in secondary deciduous scrub, overgrown pastures and abandoned farmland, wood margins, late to middle stage succession, streamside thickets, and low swamps (Bent 1950, Graber et al. 1985).

The Wood Thrush (WOTH) is common throughout Georgia during the breeding season (Greene et al. 1945). In September, they fly south to southeastern Mexico and Panama and then return to breed in early April (Roth et al. 1996). Wood Thrushes feed mostly on soil invertebrates in leaf litter or on semibare ground where herbaceous cover is open (Holmes and Robinson 1988, Roth et al. 1996). Continent wide data from the BBS show significant declines in Wood Thrush densities (Sauer and Droege 1992, Sauer et al. 1997). Wood Thrushes may be found in areas with moderate sub canopy and shrub density, decaying leaf litter, moist soil and fairly open forest floor, including interior and edges of deciduous or mixed forests and small forest fragments in residential areas (Bent 1949, Bertin 1977, Noon 1981, James et al. 1984, Robbins et al. 1989, Roth et al. 1996).

The Yellow-billed Cuckoo (YBCU) is fairly common throughout Georgia during the breeding season (Greene 1945). It migrates to South America in late August and returns in late
spring (Hughes 1999). This cuckoo forages primarily for large insects by gleaning from leaves and stems but is also known to eat small lizards and frogs (Nolan and Thompson 1975, Hamilton and Hamilton 1965). Although populations can vary greatly locally (Eaton 1988), overall, eastern U.S. populations have shown highly significant declines between 1980-1994 (Hughes 1999). The Yellow-billed Cuckoo prefers open woodland with clearings and low dense scrubby vegetation and is often found in riparian areas while avoiding dense forest or large urban areas (Bent 1940, Eaton 1988, Hughes 1999).

OBJECTIVES

The purpose of my study was to build empirical hierarchical models of the distributions of nine avian species throughout the state of Georgia and then test those models through application to management scenarios. I wished to answer the following questions:

1. Can an empirical model be built that predicts distributions on the same scale and uses the same inputs as standard GAP Analysis?
2. What resolution of spatial information is useful for modeling each species?
3. How will these models perform when used in the ways suggested by GAP Analysis, including coarse scale applications like land purchasing and fine applications like habitat manipulation?
4. What effect will model uncertainty have on potential optimal decisions based on these models?

My results will show the efficacy and validity of an empirical approach to statewide modeling. My results will also explore the incorporation of model uncertainty into management.
LITERATURE CITED


National Gap Analysis Program. 2000. National Gap Analysis Program Homepage:
http://www.gap.uidaho.edu


Scott, J. M., Davis, F., Csuti, B., Noss, R., Butterfield, B., Groves, C., Anderson, H., Caicco, S.,

46:346-358

Scot, D. M. and M. D. Jennings eds, Moscow, ID: National GAP Analysis Project. 1.1-1.6

Strain, J. G. and R. L. Mumme. 1988. Effects of Food Supplementation, Song Playback, and
Temperature on Vocal Territorial Behavior of Carolina Wrens. *Auk*. 105:11-16


CHAPTER 2

HIERARCHICAL MODELS OF AVIAN DISTRIBUTIONS FOR THE STATE OF GEORGIA:
MODEL BUILDING AND TESTING

---

**Abstract:** To predict the distributions of nine breeding birds in the state of Georgia, we built models consisting of four hierarchical levels built on nested mapping units of decreasing area: 90,000 hectares (level 4), 3600 hectares (level 3), 144 hectares (level 20, and 5.76 hectares (level 1). We used the Partners in Flight (PIF) database of point counts to generate presence vs. absence at locations across the state of Georgia for nine avian species: Acadian Flycatcher (*Empidonax virescens*), Brown-Headed Nuthatch (*Sitta pusilla*), Carolina Wren (*Thryothorus ludovicianus*), Indigo Bunting (*Passerina cyanea*), Northern Cardinal (*Cardinalis cardinalis*), Prairie Warbler (*Dendroica discolor*), Yellow-Billed Cuckoo (*Coccyzus americanus*), White-eyed Vireo (*Vireo griseus*), and Wood Thrush (*Hylocichla mustelina*). At each location, vectors of habitat measurements were taken from the Georgia GAP landcover and other GIS sources for each of the hierarchical levels. The models were then fit to a hierarchical logistic regression model using MCMC and Gibbs sampling through the program WinBugs. Model selection was performed using AIC and DIC. For all species, models from the final top three set utilized variables from multiple hierarchical levels. For these models, we tested out of sample prediction using three fold cross validation. Results showed an average overall correct classification rate of 72%. Finally, we found that many variable relationships at coarse spatial scales deviated from what we expected, based local habitat preferences from our literature review. This led us to conclude that microhabitat preference may not be an accurate predictor for coarse scale habitat interactions.

**Key Words:** GAP Analysis, Hierarchical Models, MCMC, logistic regression, Bayesian, AIC, DIC, Georgia
GAP analysis is an effort to address conservation of biodiversity on a national scale through identification of the gaps in representation of biodiversity in areas managed exclusively or primarily for the long-term conservation of native species and natural ecosystems (Scott et al. 1993). GAP analysis is performed by producing a map with three layers (Scott et al. 1993); a map of the vegetation communities derived from satellite imagery (Stoms 1998), a predicted distribution for every native vertebrate that occurs in the region (Csuti and Crist 1998), and a map of lands that have been designated as being set aside specifically for conservation (Edwards et al. 1998). These layers are then compared to determine which areas have high species diversity, but are not protected for conservation (Scott et al. 1993, Jennings 2000). Based on these so-called gaps, managers can make optimal planning decisions in terms of land acquisition and finer-scale studies (Scott et al. 1993, Kiester et al. 1996).

To create the vertebrate distribution layer, the range of each species is mapped using occurrence records (Scott et al. 1993, Csuti and Crist 1998). Then, a wildlife habitat matrix, created through an extensive literature review that delineates habitats within the vegetation map where the species is considered present, is applied throughout the species’ range (Scott et al. 1993, Csuti and Crist 1998).

This form of modeling raises two concerns. First, confidence in the vertebrate distribution map may be limited in that the relationships used to predict the final distributions are not tied to any specific data (Mannan et al, 1996). Although tests of some GAP models have performed relatively well (Edwards et al. 1996), expert opinion models similar to those employed by GAP historically have not performed well (Giles 1978, Lancia et al. 1982, Cole and Smith 1983). Further, these models are unable to quantitatively reflect the relationship between variables or provide statistical measures of uncertainty, which has implications for the potential application...
of GAP species richness maps for conservation area choice and design as the decision making process is heavily dependent on the level of confidence in the quality of the models used (Conroy and Noon 1996).

Second, habitat selection in GAP models is characterized based on local effects; however, many have proposed that surrounding habitats should also be considered (Wiens 1989, Pearson 1993, Saab 1999, Penhollow and Stauffer 2000). Indeed, the strength of biogeographical patterns or priority of conservation selection may be related to the spatial scale of the investigation as well as processes within the landscape (Stoms 1994). How an organism responds to an individual patch requires explicit recognition of the hierarchical structure of nested patches (Kotliar and Wiens 1990). Because specification of a reliable spatial model requires selection of a model structure appropriate with regard to spatial scale, level of resolution, and the specific form of functional relationships (Conroy et al. 1995), a model of habitat selection should integrate local and surrounding effects.

In order to address these concerns, we built hierarchical empirical models that predict the probability of occurrence for nine avian species throughout the state of Georgia based on spatially nested habitat inputs from the Georgia GAP Analysis 18-class land cover map.

METHODS

In order to build our models, we used records taken from the Georgia Partners in Flight (PIF) Point Count Database. The database is a combination of point counts taken by PIF cooperators through out Georgia from 1995-2000. We had coordinates for 787 stations distributed throughout the state (Figure 2.1). Stations varied in the years that they were surveyed resulting in a total of 3873 point counts that we could use for our modeling. From records in the
database, we recorded presence or absence at each point for nine avian species, the Acadian Flycatcher (*Empidonax virescens*), Brown-Headed Nuthatch (*Sitta pusilla*), Carolina Wren (*Thryothorus ludovicianus*), Indigo Bunting (*Passerina cyanea*), Northern Cardinal (*Cardinalis cardinalis*), Prairie Warbler (*Dendroica discolor*), White-eyed vireo (*Vireo griseus*), Wood Thrush (*Hylocichla mustelina*), and Yellow-billed Cuckoo (*Coccyzus americanus*), for each year that the station was surveyed. Species were chosen in order to get a cross section of preferred habitat types, frequency based on relative numbers of detections, and migration patterns (Table 2.1).

We modeled hierarchical spatial habitat patches through four levels of nested blocks of decreasing area. We maintained a 1:25 ratio between each level (Figure 2.2). Level four blocks, intended to be about the size of a Georgia county, were 90,000 ha. Level three blocks, intended to be about the size of a habitat reserve, were 3,600 ha. Level two blocks, intended to represent a community, were 144 ha. Finally, level one blocks at 5.76 ha represented an inflated individual territory. This inflated value was used because exploratory results showed smaller blocks failed to adequately capture enough heterogeneity for our purposes.

Habitat measures within each block were derived from an 18-class land cover map generated by the Georgia GAP Analysis Project (Natural Resources Spatial Analysis Laboratory, 2001). This map consisted of classified 30x30 m Landsat TM imagery from 1996-1998 delineated into 18 major land cover classes (Table 2.2). Habitat variables were measured in four ways (Table 2.3). First, we measured the percentage of each cover type within the hierarchical blocks. Then we measured percentage of certain land cover types pooled into less specific categories. We took measures of forest fragmentation by creating a map of forested blocks with a 100 ha minimum area. This minimum area was used to insure that blocks were not comprised
primarily of edge habitat and to eliminate tiny isolated forest patches. We then measured how each hierarchical mapping unit intersected with the forest blocks. Finally, we measured riparian area by buffering a coverage of Georgia streams with an area 100m on each side. We then measured the area of this buffer within each mapping unit.

For each species we conducted a review of the literature to determine variables likely to affect habitat choice. Based on this review, we selected variables that we expected would have positive associations as well as those that we believed would affect species occurrence in a negative way. We chose four variables at level one and three variables at each of the upper levels (Table 4). We deliberately repeated variables between levels in some cases in order to investigate effects of scale. We also chose variables we believed are useful to management applications.

Acadian Flycatchers (ACFL) were relatively uncommon in our data set with 857 detections at our PIF stations out of a possible 3873 compared to other birds in our species set. They may be found in deciduous and mixed forests including heavily wooded bottomlands, wooded ravines, cypress swamps, and forested waterways (Bent 1942, Mumford 1964, Walkinshaw 1966, Hespenheide 1971). At level one we chose the variables dfor, mfor, riparian, and lfba (Table 2.3). At level two we chose the variables forwet, riparian, and tlifba. At level three we chose the variables forest, wetland, and ifba. At level four we chose the variables wetland, urban, and forest.

Brown-Headed Nuthatches (BHNU) were relatively rare in our data set with 228 of 3873 possible detections. They occur primarily in association with pine trees, inhabiting a wide variety of southeastern pine forest habitats (Withgott and Smith 1998). It prefers open mature pinewoods, including burned-over areas (Bent 1948, Hamel 1992, Harrop and Quin 1995, Withgott and Smith 1998). It can also be found in cypress swamps and pine forest edges (Bent
At level one we chose the variables forwet, efor, mfor, and clearcut (Table 2.3). At level two we chose the variables efor, forwet, and clearcut. At level three we chose the variables efor, natopen, and riparian. At level four we chose the variables efor, wetland, and urban.

Carolina Wrens (CARW) were relatively common within our data set with 1691 of 3873 possible detections. It inhabits a wide range of habitats so long as they include moderate to dense shrub or bushy cover (Bent 1948, James 1971, Conner et al. 1983, Haggerty and Morton 1995). It is also found in wooded residential areas with shrubs (Beissinger and Osbourne 1982). It prefers moist bottomland woods to dry upland ones (Bent 1948, James and Neal 1986). At level one we chose the variables lowurb, forest, clearcut, and utility (Table 2.3). At level two we chose the variables forest, natopen, and lowurb. At level three we chose the variables forest, natopen, and unnatopen. At level four we chose the variables ifba, open, and urban.

Indigo Buntings (INBU) were relatively uncommon in our data set with 576 of 3873 possible detections. The Indigo bunting prefers brushy and weedy habitats along the edges of woods and cultivated lands, including roads, utility right of ways, railway siding, and riparian habitats (Payne 1992). It is generally absent from closed canopy forest, residential areas and intensively cultivated or grazed areas (Payne 1992). At level one we chose the variables natopen, ag, riparian, and forest (Table 2.3). At level two we chose the variables forest, natopen, and unnatopen. At level three we chose the variables natopen, unatopen, and lowurb. At level four we chose the variables natopen, unatopen, and urban.

Northern Cardinals (NOCA) were the most common bird in our data set with 2010 of 3873 possible detections. The cardinal occurs in habitats throughout its range in areas with shrubs or small trees including forest edges, interior shrubby areas within second growth forest,
marsh edges, riparian forest, grasslands with shrubs, old fields, agricultural hedgerows, and landscaping around buildings (Bent 1968, Dow 1969, Emlen 1972). At level one we chose the variables natopen, riparian, lowurb, and ifba (Table 2.3). At level two we chose the variables riparian, lowurb, and open. At level three we chose the variables lowurb, open, and ifba. At level four we chose the variables lowurb, open, and ifba.

Prairie Warblers (PRAW) occurred rarely in our data set with 304 of 3873 possible detections. They occur in various shrubby areas lacking closed canopies (Nolan et al. 1999). They can be found most often in open southern pine forest, dunes, abandoned field and orchards, and grassland forest edges (Griscom 1979, Schorre 1998, Nolan et al. 1999). At level one we chose the variables beaches, efor, clearcut, and pasture (Table 2.3). At level two we chose the variables pasture, clearcut, and beaches. At level three we chose the variables efor, natopen, and unnatopen. At level four we chose the variables unnatopen, natopen, and urban.

The White-eyed Vireo (WEVI) was relatively uncommon in our data set with 799 of 3873 possible detections. It can be found in secondary deciduous scrub, overgrown pastures and abandoned farmland, wood margins, late to middle stage succession, streamside thickets, and low swamps (Bent 1950, Graber et al. 1985). At level one we chose the variables cleacut, pasture, dfor, and ifba (Table 2.3). At level two we chose the variables dfor, clearcut, and riparian. At level three we chose the variables forest, natopen, and ag. At level four we chose the variables ifba, open, and urban.

Wood Thrushes (WOTH) were uncommon in our data set with 435 detections. Wood thrushes may be found in areas with moderate sub canopy and shrub density, decaying leaf litter, moist soil and fairly open forest floor, including interior and edges of deciduous or mixed forests and small forest fragments in residential areas (Bent 1949, Bertin 1977, Noon 1981, James et al.
1984, Robbins et al. 1989, Roth et al. 1996). At level one we chose the variables dfor, mfor, lfb, and forwet (Table 2.3). At level two we chose the variables forwet, riparian, and lfb. At level three we chose the variables forest, wetland, and ifba. At level four we chose the variables ifba, open, and urban.

The Yellow-billed Cuckoo (YBCU) was fairly uncommon in our data set with 751 of 3873 possible detections. They prefer open woodland with clearings and low dense scrubby vegetation and are often found in riparian areas while avoiding dense forest or large urban areas (Bent 1940, Eaton 1988, Hughes 1999). At level one we chose the variables dfor, natopen, riparian, and ifba (Table 2.3). At level two we chose the variables unnatopen, natopen, and riparian. At level three we chose the variables wetland, open, and ifba. At level four we chose the variables ifba, open, and urban.

For computational purposes, all habitat data were standardized. That is, the mean was set to zero and the standard deviation set to 1 by subtracting the sample mean from each observation and dividing by the sample’s standard error. Standardization has implications primarily for interpretation in that probability of occurrence is at the intercept value when all habitat variables are at their means.

We related presence or absence to our hierarchical habitat data based on a hierarchical logit model (Wong and Mason 1985, Bryk and Raudenbush 1993). We begin with a single level linear model:

$$Y_i = \alpha_0 + \alpha_1 X_{i1} + \ldots + \alpha_P X_{Pi} + r_i,$$

where $Y_i$ is the response and $X_{i1} \ldots X_{Pi}$ are the habitat variables at level one for observation $i$, $\beta_0$ is the intercept, $\alpha_1 \ldots \alpha_P$ are the coefficients, and $r$ is the error assumed to be normally distributed with a mean of zero and variance $\sigma^2$. However, our observations belong to a spatially
explicit hierarchy of levels, and we assume that observations that occur in the within the same
level two block are more similar to each other than observations which occur outside of that
block. Therefore, we can express the relationship between response and predictor variables in
level two block \( j \) can be expressed as:

\[
Y_{ij} = \alpha_{0j} + \alpha_{1j}X_{1ij} + \ldots + \alpha_{Pj}X_{Pij} + r_{ij},
\]

where the variables are as described in eq. (1). Next, we assume that the intercept and
coefficients in eq. (2) vary between level two blocks based on a set of habitat variables \( W_{1j} - W_{Sj} \)
measured for each block. The influence of these variables may be modeled by expressing the
level one coefficients (\( \alpha \)'s) as linear functions of these habitat measures:

\[
\alpha_{0j} = \beta_{00} + \beta_{01}W_{1j} + \ldots \beta_{0S}W_{Sj} + \mu_{0j},
\]

\[
\alpha_{1j} = \beta_{10} + \beta_{11}W_{1j} + \ldots \beta_{1S}W_{Sj} + \mu_{1j},
\]

\[
\alpha_{Pj} = \beta_{P0} + \beta_{P1}W_{1j} + \ldots \beta_{PS}W_{Sj} + \mu_{PJ},
\]

where \( \beta_{Q0} \ldots \beta_{PS} \) are the level 2 coefficients, \( \mu_{0j} \ldots \mu_{PJ} \) are the random effects, and \( W_{1S} \ldots W_{Sj} \)
are the level 2 habitat measures for group \( j \). The random effects \( \mu_{0j} \ldots \mu_{PJ} \) represent the effect
associated with each level 2 unit that is not explained by the model. The random effects are
assumed to be normally distributed with mean 0 and variance \( \tau \). The complete two level linear
model can be expressed by substituting the level 2 components in eq. (3) for \( \alpha_{Pj} \) in eq. (2):

\[
Y_{ij} = \beta_{00} + \beta_{01}W_{1j} + \ldots \beta_{0S}W_{Sj} + \mu_{0j} + \beta_{10}X_{1ij} + \beta_{11}W_{1j}X_{1ij} + \ldots \beta_{1S}W_{Sj}X_{1ij} + \mu_{1j}X_{1ij} + \ldots \beta_{P0}X_{Pij} + \beta_{P1}W_{1j}X_{Pij} + \ldots \beta_{PS}W_{Sj}X_{Pij} + \mu_{PJ}X_{Pij} + r_{ij}.
\]

Finally, the hierarchical logit model may be gained by assuming a binomial distribution and a
logit link as:

\[
\eta_{ij} = \beta_{00} + \beta_{01}W_{1j} + \ldots \beta_{0S}W_{Sj} + \mu_{0j} + \beta_{10}X_{1ij} + \beta_{11}X_{1ij}W_{1j}X_{1ij} + \ldots \beta_{1S}W_{Sj}X_{1ij} + \mu_{1j}X_{1ij} + \ldots \beta_{P0}X_{Pij} + \beta_{P1}W_{1j}X_{Qij} + \ldots \beta_{PS}W_{Sj}X_{Pij} + \mu_{PJ}X_{Pij},
\]
where $\eta_{ij}$ is the log odds of occurrence for level one block $i$ in level two block $j$.

Our model involved four hierarchical levels. Unfortunately a fully realized four level hierarchical model was deemed impractical for our project because, due to the nature of the model, complexity increases geometrically with each hierarchical level. We were forced to simplify the model by making the potentially unrealistic assumption of non-interaction within and between hierarchical levels. This was accomplished through intercept only substitution:

$$\alpha_{00} = \beta_{00} + \beta_{01} W_{1jkl} + \ldots \beta_{0S} W_{Sjkl} + \mu_{0jkl}, \quad (6a)$$

$$\beta_{00} = \chi_{00} + \chi_{01} V_{1kl} + \ldots \chi_{0R} V_{Rkl} + \mu_{0kl}, \quad (6b)$$

$$\chi_{00} = \delta_{00} + \delta_{01} U_{11} + \ldots \delta_{0Q} U_{Q1} + \mu_{0l}, \quad (6c)$$

where $\chi_{00} \ldots \chi_{0R}$ and $V_{1kl} \ldots V_{Rkl}$ are the level three coefficients and predictors respectively and $\delta_{00} \ldots \delta_{0Q}$ and $U_{11} \ldots U_{Q1}$ are the level four coefficients and predictors respectively. Substitution results in a model with the form:

$$\eta_{ijkl} = \delta_{00} + \delta_{01} U_{11} + \ldots \delta_{0Q} U_{Q1} + \mu_{0l} + \chi_{00} + \chi_{01} V_{1kl} + \ldots \chi_{0R} V_{Rkl} + \mu_{0kl} + \beta_{00} + \beta_{01} W_{1jkl} + \ldots \beta_{0S} W_{Sjkl} + \mu_{0jkl} + \alpha_{1} X_{1ijkl} + \ldots + \alpha_{P} X_{Pijkl}.$$

$$\eta_{ijkl} = \delta_{00} + \delta_{01} U_{11} + \ldots \delta_{0Q} U_{Q1} + \mu_{0l} + \chi_{00} + \chi_{01} V_{1kl} + \ldots \chi_{0R} V_{Rkl} + \mu_{0kl} + \beta_{00} + \beta_{01} W_{1jkl} + \ldots \beta_{0S} W_{Sjkl} + \mu_{0jkl} + \alpha_{1} X_{1ijkl} + \ldots + \alpha_{P} X_{Pijkl} + \gamma_{m} + r_{ijklm}, \quad (7)$$

Our data was taken from a database that contained seven years worth of data. The nature of our habitat data was such that we could only record it in one year, 1998, and we assumed it changed minimally. We decided to model the effect of year as an additional random effect. This leaves us with the final form of the model:

$$\eta_{ijklm} = \delta_{00} + \delta_{01} U_{11} + \ldots \delta_{0Q} U_{Q1} + \mu_{0l} + \chi_{00} + \chi_{01} V_{1kl} + \ldots \chi_{0R} V_{Rkl} + \mu_{0kl} + \beta_{00} + \beta_{01} W_{1jkl} + \ldots \beta_{0S} W_{Sjkl} + \mu_{0jkl} + \alpha_{1} X_{1ijkl} + \ldots + \alpha_{P} X_{Pijkl} + \gamma_{m} + r_{ijklm}, \quad (8)$$

where $\gamma_{m}$ is the random effect of time period $m$.

The random error associated with each observation in this model is $\mu_{0l} + \mu_{0kl} + \mu_{0jkl} + r_{ijklm}$. Although the random effect of each observation ($r_{ijklm}$) is captured by binomial sampling
error, the random effects of location within our hierarchical blocks must be estimated. In addition, we have the random effect of time. These random effects are difficult to estimate using standard maximum likelihood techniques (Wong and Mason 1985). For this reason we chose to conduct our modeling in a Bayesian framework using MCMC (Gelman et al. 1995, Congdon 2001, Link et al. 2002), specifically the program WinBugs (Lunn 2003). For a more thorough analysis of MCMC see Link et al. (2003).

Winbugs uses Gibbs’ sampling to estimate the posterior distribution of a given set of parameters \( \theta \), based on prior distributions \( f(\theta) \), a set of data \( y \) that informs on these priors, and a model likelihood eq. (8). We chose to use diffuse prior distributions in order to reflect a lack of prior knowledge about model parameters. Coefficients \( (\alpha, \beta, \chi, \delta) \) were assigned normal distributions with a mean of 0 and precision of .000001. Priors for the random effects \( (\mu_{0l}, \mu_{0kl}, \mu_{0jkl}) \) were specified by \( \mu_n \sim \text{norm}(0, \tau_n) \) with \( \tau \) as hyperparameters for each level \( n \) expressed in the form \( \tau \sim \text{gamma}(.001, 001) \). Finally the effect of time \( \gamma \) was specified as \( \gamma_m \sim \text{norm}(0, \iota) \) with \( \iota \) as a hyperparameter specified by \( \iota \sim \text{gamma}(.001, 001) \).

We chose a candidate set of ten models for each species. To create the candidate set, we first fit all possible models without random effects using standard maximum likelihood techniques in SAS (SAS Institute 2001). The candidate set contained the top 3 models from this run based on AIC (Burnham and Anderson 2002). We also chose two models randomly from the remaining possible models. Finally we fit the full model and models with one hierarchical level removed. Each model in the candidate set was run through 11,000 iterations from which we discarded the first 1000 runs as a burn in period.

Due to the nature of MCMC, we calculated AIC in an unorthodox way. That is, AIC was estimated at each MCMC iteration. This results in a distribution of AIC. For model selection, we
took the mean of the distribution. It is unclear what affect this had on our value of AIC.
Additionally, AIC is calculated by adding twice the number of parameters to the deviance.
Unfortunately, in a situation where one has random group effects it is unclear how to include
them in the count of parameters. In model fitting, there is a random effect estimated for each
block in each level. This would be 384 variables for level 2, 119 variables for level 3, and 40
variables for level four. These parameters do not seem to have the same meaning as the
parameters that serve as coefficients for our habitat variables and would likely distort the effects
of adding and removing levels from the model on AIC. In our modeling, we decided, therefore,
not to include random effects in the parameter count for AIC. This means there was no
discounting of additional levels beyond cost of additional habitat coefficients.

Initial runs showed that our candidate sets likely did not contain the best model; the
models chosen at random often had the lowest AIC scores of the models we tested. Also, several
parameters had posterior distributions with means at or near zero. Based on the results from
those runs, we conducted a stepwise approach of adding and removing variables from the models
for each species based on the posterior distributions from the initial run until we were satisfied
with the resulting model set. Note that this stepwise procedure does not guarantee that the final
model set contains the “best” model, however we believed it to be a reasonable compromise
between achieving this “best” model and expediency.

In order to insure that the MCMC results from the top three models were close to the
posterior distributions of the parameters, we conducted a new run for each of them. These runs
were conducted using two chains for 22,000 iterations, with the first 2000 discarded as burn in.
We used the two chains to test for convergence using the Gelman-Rubin convergence statistic
(Brooks and Gelman 1998). Our tests showed that the parameters for environmental effects
showed good convergence. However, hyperparameters for the unique effects of group membership often failed to converge. The initial runs also showed significant autocorrelation for some parameters, that is the values for individual samples were dependent upon preceding observations. This has an impact on the ability of the estimates to approximate the posterior distribution. To address this concern, we ran the top three models for each species through 20,000 iterations and thinned to every 10th sample. In most cases, this eliminated autocorrelation. Finally, in addition to AIC, we calculated DIC for each of these models.

Ultimately, we were interested in the ability of these models to make predictions. Therefore, we tested out-of-sample prediction performance for the top three models from each species. First, we partitioned the data into thirds, using two thirds of the data to fit our models and estimate parameters. The fitted models were then used to predict the remaining one third of the data and calculate prediction error. For these runs, error rates were calculated by applying a specific threshold for each species calculated as the prevalence of that species within the dataset, that is the number of detections for each species divided by the total number of records. If the predicted probability of occurrence was greater than the threshold, then the species was predicted to be present, otherwise an absence was predicted. We then averaged the classification rates from each third.

RESULTS

The final model set including scores for AIC and DIC can be seen in Table 2.5. Model rankings often varied between AIC and DIC. The final set contained models with multiple levels for every species.
In the course of recalculating AIC for the third stepwise model for the Prairie Warbler, we discovered that our previous value was erroneous and, therefore, that model did not belong in the actual top 3 of the models we tested for this species. We decided to leave the model in the set because it demonstrates an important trend in the hierarchical nature of the models. That is, that in all cases where level 2 was not considered in the modeling, the values for our model selection criteria were much larger than for models which included said level. This jump could have been the result of the failure of our method of calculating AIC to discount for the addition of the random effects, however a similar jump was observed in the values for DIC which presumably does not have that problem.

The relationships predicted by the models are summarized in Table 2.6. In most cases the relationships were in line with what we expected based on our literature review. There were some models, however, which returned unexpected results.

We considered the Brown-headed Nuthatch to be a forest bird, so the positive relationship to natural open areas was unexpected. The positive effect of natural open areas may stem from the fact that this nuthatch prefers open forested habitats, a distinction our models do not make due to the limitations of the land cover classes available as habitat inputs. That is, a dense evergreen forest is classified the same as an open one. Also of note is the fact that there is a negative relationship for clear cuts at level 1 in the third stepwise model, whereas the relationship is positive in the remaining two models. A review of the posterior distributions for clearcut_l1, however, shows that they are distributed fairly closely about zero (Figure 2.3). The result is that the mean effect listed in the table does not fully represent the actual impact. Therefore the flip-flop may reflect slight distributional changes rather than a change in the relationship.
The models for the Indigo Bunting and the Wood Thrush showed positive relationships to urban area at the coarsest level. We believe this may be explained two ways. First, they may be statistical artifacts caused of the general lack of urban areas sampled by our point counts. Second, it may be that the birds are not actually responding positively to urban areas, rather some other factor associated with urban development may be exerting the positive influence.

The results of our cross validation may be seen in Table 2.7. Overall correct classification rates varied from 53% for the Carolina Wren models to as high as 86% for the Prairie Warbler with a species mean of 72%. Individual classification rates show that this is mostly due to very high rates for correctly classifying absence. The models generally performed poorly in classifying presence with values ranging from 29% of presences correctly classified for the Brown-headed Nuthatch to 64% for the Northern Cardinal with an overall species mean of 45%.

DISCUSSION

After our modeling was complete we had three potential methods for model selection, AIC, DIC, and classification rates from the cross validation (Table 2.8). As the table shows, under each species there were several cases where each method produced varying results. For instance, in the Acadian Flycatcher models, each method returned a different top model among the three possibilities. Even in cases where the top model may be agreed upon, the orders of subsequent models were rarely identical among methods. Indeed, such agreement was achieved only in the ranking of the models for the Carolina Wren. We will discuss the strengths and weaknesses of each in our application.

AIC is suspect because of the method we used to calculate it. We do not know what effect not including random effects in the parameter count had, but it may have led to a bias
towards including additional levels. Additionally, we do not know whether the median of the posterior distribution of AIC may have been a more appropriate representative value than the mean.

Using the results of the cross validation for model selection can be problematic as well. The actual rates themselves are based on applying a threshold value to change probability of occurrence to predictions of presence or absence. This means that the eventual rates are highly dependent upon the values selected for those thresholds.

DIC is an alternative information theoretic criterion to AIC derived specifically with modeling problems such as this one in mind (Spiegelhalter et al. 2002). Unfortunately, we only became aware of it after the bulk of our modeling was already complete. For this reason, much of our early work in reducing the model set was conducted exclusively based on AIC. Given the differences in model order, it may be instructive to go back and redo the modeling based entirely on DIC to determine how different the final top-three model set would be.

Our modeling showed the usefulness of considering hierarchical levels. Most species models had one to two variables in the first and second level. The White-eyed Vireo was the only species modeled where the top model had no level one (small scale) variables. We believe this reflects the inability of the land cover classes to reflect the shrubby edge habitats that this species prefers. For all of our species, failure to consider level two led to extreme increases in both AIC and DIC as demonstrated by the third stepwise model for the Prairie Warbler. We believe this reflects both the importance of this level, and the fact that this is the level with the most individual group effects. That is, the 384 level two unique group effects account for a great deal of the variation between observations.
Model complexity did not seem related to habitat preference. We expected that specialist species, like the Brown-headed Nuthatch, would use mainly the finer scale levels, while the generalists would be more likely to favor the coarse scales. While both of our generalist species, the Carolina Wren and the Northern Cardinal, made use of all four levels, they were not the only species to do so, and our specialists seemed equally inclined to including the coarser scales. This may reflect the fact that our finest scale is still somewhat coarse compared to the habitat characteristics usually used to define a species’ requirements. Also, forest birds versus open space birds varied equally in their use of spatial scales.

In many cases, the eventual habitat relationships predicted by our models were at odds with what we expected based on our literature review. For instance, the models for the Prairie Warbler predicted negative relationships for natural and unnatural open areas at level four. This is contrary to the bird’s documented preference for shrubby open habitats. However, these discrepancies all occurred at the coarse spatial scales. This leads us to conclude that in considering hierarchical spatial relationships, the sort of localized habitat preferences commonly studied for most species may not be accurate guides to how the species relates to those same factors at coarser spatial scales. This is shown several times in our model set. For instance, the indigo bunting has a positive relationship to natural open areas at level one but then shows a negative relationship to these same areas when dealing with the coarser scale of level four. A similar relationship only in reverse is shown in the models for the Northern Cardinal. In those models, the Cardinal shows a positive relationship to open areas at the coarse levels and a negative relationship to natural open areas at finer scales.

A central weakness to using the spatial habitat data in our application was the lack of independence between habitat measures; that is, increases in one cover type must lead to
decreases in other cover types. This means that correlations with variables could reflect reactions to the way changes in one cover type generally impact another cover type. For instance, the Indigo Bunting shows positive relationships to increases in urban areas. Indigo Buntings do not generally occur in urban areas, therefore the relationship predicted by the model is suspect. However, the positive relationship may not reflect a positive reaction to urban itself, rather it may show a reaction to other types of habitat which generally are found along with urban areas. We recommend understanding of the biology involved before application to management.

In the next chapter, we will apply these models to simulated management scenarios to evaluate the impacts of statistical uncertainty on decision-making. We will also look at the impact of applying our different model selection criteria on those scenarios.
LITERATURE CITED


SAS Institute Inc. 2001, SAS: ver 8.02, Cary, NC


Figure 2.1: Partners in Flight (PIF) stations in Georgia (1995 – 2001).
Figure 2.2: Hierarchical mapping units created through spatial nesting for hierarchical logistic regression modeling for the state of Georgia based on the 1998 Georgia GAP Analysis 18-class landcover map.
Figure 2.3 Posterior Distributions for the parameters associated with clearcut_11 for the Brown-Headed Nuthatch models.
Table 2.1: Avian species chosen for hierarchical logistic regression modeling for the state of Georgia based on the 1998 Georgia GAP Analysis 18-class landcover map and Georgia Partners in Flight (PIF) point counts (1995-2001). References for habitat and migration categories located within text. Frequency reflects occurrence within the PIF point counts.

<table>
<thead>
<tr>
<th>Common Name</th>
<th>Scientific Name</th>
<th>Code</th>
<th>Frequency</th>
<th>Habitat</th>
<th>Migration</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acadian Flycatcher</td>
<td><em>Empidonax Virescens</em></td>
<td>ACFL</td>
<td>Uncommon</td>
<td>Forest</td>
<td>Migrant</td>
</tr>
<tr>
<td>Brown-Headed Nuthatch</td>
<td><em>Sitta pusilla</em></td>
<td>BHNU</td>
<td>Rare</td>
<td>Forest</td>
<td>Resident</td>
</tr>
<tr>
<td>Carolina Wren</td>
<td><em>Thryothorus ludovicianus</em></td>
<td>CARW</td>
<td>Common</td>
<td>General</td>
<td>Resident</td>
</tr>
<tr>
<td>Indigo Bunting</td>
<td><em>Passerina cyanea</em></td>
<td>INBU</td>
<td>Uncommon</td>
<td>Open</td>
<td>Migrant</td>
</tr>
<tr>
<td>Northern Cardinal</td>
<td><em>Cardinalis cardinalis</em></td>
<td>NOCA</td>
<td>Common</td>
<td>General</td>
<td>Resident</td>
</tr>
<tr>
<td>Prairie Warbler</td>
<td><em>Dendroica discolor</em></td>
<td>PRAW</td>
<td>Rare</td>
<td>Open</td>
<td>Migrant</td>
</tr>
<tr>
<td>White-Eyed Vireo</td>
<td><em>Coccyzus Americanus</em></td>
<td>WEVI</td>
<td>Uncommon</td>
<td>Open</td>
<td>Migrant w/Resident</td>
</tr>
<tr>
<td>Wood Thrush</td>
<td><em>Vireo griseus</em></td>
<td>WOTH</td>
<td>Uncommon</td>
<td>Forest</td>
<td>Migrant</td>
</tr>
<tr>
<td>Yellow-Billed Cuckoo</td>
<td><em>Hylocichla mustelina</em></td>
<td>YBCU</td>
<td>Uncommon</td>
<td>Forest</td>
<td>Migrant</td>
</tr>
</tbody>
</table>
Table 2.2: Landcover Classification System used by the Georgia GAP Analysis Project (2003) for the 1998 18-class Statewide Landcover Map based on Landsat Thematic Mapping imagery using a 30 x 30 meter pixel.

<table>
<thead>
<tr>
<th>Class Name</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>Beaches/Dunes/Mud</td>
<td>Open sand, sandbars, sand dunes, mud - natural environments as well as exposed sand from dredging and other activities.</td>
</tr>
<tr>
<td>Open Water</td>
<td>Lakes, rivers, ponds, ocean, industrial water, aquaculture.</td>
</tr>
<tr>
<td>Transportation</td>
<td>Roads, railroads, airports, and runways.</td>
</tr>
<tr>
<td>Utility Swaths</td>
<td>Swaths maintained for transmission lines.</td>
</tr>
<tr>
<td>Low Intensity Urban</td>
<td>Single family dwellings, recreation, cemeteries, playing fields, campus-like institutions, parks, and schools.</td>
</tr>
<tr>
<td>High Intensity Urban</td>
<td>Multi-family dwellings, commercial/industrial, prisons and At’l Int’l Speedway, junkyards, confined animal operations.</td>
</tr>
<tr>
<td>Clearcut/Sparse</td>
<td>Recent clearcuts, sparse vegetation, clearcut wetlands.</td>
</tr>
<tr>
<td>Quarries/Strip Mines</td>
<td>Exposed rock and soil from industrial uses, gravel pits, landfills.</td>
</tr>
<tr>
<td>Rock Outcrop</td>
<td>Rock outcrops, mountaintops, and barren land.</td>
</tr>
<tr>
<td>Deciduous Forest</td>
<td>Forest composed of at least 75% deciduous trees in the canopy, mountain shrub/scrub, and deciduous woodland.</td>
</tr>
<tr>
<td>Evergreen Forest</td>
<td>Evergreen forest, at least 75% evergreen trees, managed pine plantations, evergreen woodland.</td>
</tr>
<tr>
<td>Mixed Forest</td>
<td>Mixed deciduous/coniferous, fall line and coastal plain shrub/scrub, mixed woodland.</td>
</tr>
<tr>
<td>Golf Courses</td>
<td>Golf courses.</td>
</tr>
<tr>
<td>Pasture</td>
<td>Pasture, non-tilled grasses.</td>
</tr>
<tr>
<td>Row Crop</td>
<td>Row crops, orchards, vineyards, groves, horticultural businesses.</td>
</tr>
<tr>
<td>Forested Wetland</td>
<td>Cypress gum, evergreen wetland, deciduous wetland, depressional wetlands, and shrub wetlands.</td>
</tr>
<tr>
<td>Coastal Marsh</td>
<td>Coastal freshwater and brackish marsh.</td>
</tr>
<tr>
<td>Non-forested Wetland</td>
<td>Inland freshwater marsh.</td>
</tr>
</tbody>
</table>
Table 2.3: Habitat variables used hierarchical logistic regression modeling for the state of Georgia based on the 1998 Georgia GAP Analysis 18-class landcover map (2003) and Georgia Partners in Flight (PIF) point counts (1995-2001). Values are based on the 1998 Georgia GAP 18-class Statewide Landcover Map classification system.

<table>
<thead>
<tr>
<th>Variable</th>
<th>Description</th>
</tr>
</thead>
<tbody>
<tr>
<td>beaches</td>
<td>% of the beaches cover type</td>
</tr>
<tr>
<td>water</td>
<td>% of the water cover type</td>
</tr>
<tr>
<td>roads</td>
<td>% of the roads cover type</td>
</tr>
<tr>
<td>utility</td>
<td>% of the utility cover type</td>
</tr>
<tr>
<td>lowurb</td>
<td>% of the low density urban cover type</td>
</tr>
<tr>
<td>highurb</td>
<td>% of the high density urban cover type</td>
</tr>
<tr>
<td>clearcut</td>
<td>% of the clearcut cover type</td>
</tr>
<tr>
<td>quarry</td>
<td>% of the quarry cover type</td>
</tr>
<tr>
<td>Rock</td>
<td>% of the bare rock cover type</td>
</tr>
<tr>
<td>Dfor</td>
<td>% of the deciduous forest cover type</td>
</tr>
<tr>
<td>Efor</td>
<td>% of the evergreen forest cover type</td>
</tr>
<tr>
<td>Mfor</td>
<td>% of the mixed forest cover type</td>
</tr>
<tr>
<td>Golf</td>
<td>% of the golf course cover type</td>
</tr>
<tr>
<td>Past</td>
<td>% of the pasture cover type</td>
</tr>
<tr>
<td>rcrop</td>
<td>% of the row crop cover type</td>
</tr>
<tr>
<td>forwet</td>
<td>% of the forested wetland cover type</td>
</tr>
<tr>
<td>marsh</td>
<td>% of the salt water wetland cover type</td>
</tr>
<tr>
<td>Wet</td>
<td>% of the fresh water wetland cover type</td>
</tr>
<tr>
<td>forest</td>
<td>Sum of dfor, mfor, efor, and forwet</td>
</tr>
<tr>
<td>urban</td>
<td>Sum of lowurb, highurb, and rock</td>
</tr>
<tr>
<td>open</td>
<td>Sum of utility, clearcut, golf, past, rcrop, marsh, and wet</td>
</tr>
<tr>
<td>natopen</td>
<td>Sum of clearcut, marsh, and wet</td>
</tr>
<tr>
<td>unnatopen</td>
<td>Sum of utility, golf, past, rcrop</td>
</tr>
<tr>
<td>Ag</td>
<td>Sum of past and rcrop</td>
</tr>
<tr>
<td>wetland</td>
<td>Sum of forwet, marsh, and wet</td>
</tr>
<tr>
<td>Lfba</td>
<td>Area of the largest forested block with area within the unit</td>
</tr>
<tr>
<td>Ifba</td>
<td>Area within the unit consisting of forested blocks</td>
</tr>
<tr>
<td>Llifba</td>
<td>Largest area within the unit from a single forested block</td>
</tr>
<tr>
<td>tafba</td>
<td>Total area of forested blocks which intersect the unit</td>
</tr>
<tr>
<td>Tlifba</td>
<td>Total area of the forested block which has the greatest area within the unit</td>
</tr>
<tr>
<td>riparian</td>
<td>Area within 100 meters of a stream within the unit</td>
</tr>
</tbody>
</table>
Table 2.4: Candidate habitat variables selected for each species and hierarchical level based on life histories of each species (citation within text) for hierarchical logistic regression modeling for the state of Georgia based on the 1998 Georgia GAP Analysis 18-class landcover map (2003) and Georgia Partners in Flight (PIF) point counts (1995-2001). Habitat variable abbreviations are described in table 2.3.

<table>
<thead>
<tr>
<th>Species</th>
<th>Level1</th>
<th>Level2</th>
<th>Level3</th>
<th>Level4</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acadian Flycatcher</td>
<td>defor, mfor, riparian, lfba</td>
<td>forwet, riparian, tlifba</td>
<td>forest, wetland, ifba</td>
<td>wetland, urban, forest</td>
</tr>
<tr>
<td>Brown-Headed Nuthatch</td>
<td>forwet, efor, mfor, clearcut</td>
<td>efor, forwet, clearcut</td>
<td>efor, natopen, riparian</td>
<td>efor, wetland, urban</td>
</tr>
<tr>
<td>Carolina Wren</td>
<td>lowurb, forest, clearcut, utility</td>
<td>forest, natopen, lowurb</td>
<td>forest, natopen, unnatopen</td>
<td>natopen, unnatopen, lowurb</td>
</tr>
<tr>
<td>Indigo Bunting</td>
<td>natopen, ag, riparian, forest</td>
<td>riparian, lowurb, open</td>
<td>riparian, lowurb, open</td>
<td>riparian, lowurb, open</td>
</tr>
<tr>
<td>Northern Cardinal</td>
<td>beaches, efor, clearcut, pasture</td>
<td>efor, natopen, unnatopen</td>
<td>lowurb, open, ifba</td>
<td>lowurb, open, ifba</td>
</tr>
<tr>
<td>Prairie Warbler</td>
<td>dfor, natopen, ifba, riparian</td>
<td>dfor, clearcut, riparian</td>
<td>forest, natopen, ag</td>
<td>ifba, open, urban</td>
</tr>
<tr>
<td>White-Eyed Vireo</td>
<td>clearcut, pasture, dfor, ifba</td>
<td>forwet, riparian, tlifba</td>
<td>forest, wetland, ifba</td>
<td>ifba, open, urban</td>
</tr>
<tr>
<td>Wood Thrush</td>
<td>defor, mfor, lfba, forwet</td>
<td>unnatopen, natopen, riparian</td>
<td>wetland, open, ifba</td>
<td>ifba, open, urban</td>
</tr>
<tr>
<td>Yellow-Billed Cuckoo</td>
<td>defor, natopen, ifba, riparian</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>
Table 2.5: Top three model set for each species from hierarchical logistic regression modeling for the state of Georgia based on the 1998 Georgia GAP Analysis 18-class landcover map (2003) and Georgia Partners in Flight (PIF) point counts (1995-2001). Models Ranked by AIC (Burnham and Anderson 2002) and DIC (Spiegelhalter et al. 2003). Model name reflects American Ornithologists Union species code plus the step in the model selection process in which the model was fitted.

<table>
<thead>
<tr>
<th>Species</th>
<th>Model</th>
<th>AIC</th>
<th>ΔAIC&lt;sup&gt;a&lt;/sup&gt;</th>
<th>AIC wt&lt;sup&gt;b&lt;/sup&gt;</th>
<th>DIC</th>
<th>ΔDIC&lt;sup&gt;c&lt;/sup&gt;</th>
<th>DIC wt&lt;sup&gt;d&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acadian Flycatcher</td>
<td>ACFLStep3</td>
<td>2647</td>
<td>0</td>
<td>0.628532</td>
<td>2863.96</td>
<td>40.88</td>
<td>1.33E-09</td>
</tr>
<tr>
<td></td>
<td>ACFLStep2</td>
<td>2649</td>
<td>2</td>
<td>0.231224</td>
<td>2863.64</td>
<td>40.56</td>
<td>1.56E-09</td>
</tr>
<tr>
<td></td>
<td>ACFLStep1</td>
<td>2650</td>
<td>3</td>
<td>0.140244</td>
<td>2823.08</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Brown-Headed Nuthatch</td>
<td>BHUNStep3</td>
<td>1214</td>
<td>0</td>
<td>0.574097</td>
<td>1303.8</td>
<td>1.35</td>
<td>0.531673</td>
</tr>
<tr>
<td></td>
<td>BHUNStep4</td>
<td>1215</td>
<td>1</td>
<td>0.348207</td>
<td>1305.25</td>
<td>1.45</td>
<td>0.257502</td>
</tr>
<tr>
<td></td>
<td>BHUNStep2</td>
<td>1218</td>
<td>4</td>
<td>0.077696</td>
<td>1305.65</td>
<td>1.85</td>
<td>0.210825</td>
</tr>
<tr>
<td>Carolina Wren</td>
<td>CARWStep3</td>
<td>4440</td>
<td>0</td>
<td>0.50648</td>
<td>4594.21</td>
<td>0</td>
<td>0.714337</td>
</tr>
<tr>
<td></td>
<td>CARWStep2</td>
<td>4441</td>
<td>1</td>
<td>0.307196</td>
<td>4596.5</td>
<td>2.29</td>
<td>0.227319</td>
</tr>
<tr>
<td></td>
<td>CARWRand1</td>
<td>4442</td>
<td>2</td>
<td>0.186324</td>
<td>4599.22</td>
<td>5.01</td>
<td>0.058344</td>
</tr>
<tr>
<td>Indigo Bunting</td>
<td>INBUSStep4</td>
<td>2194</td>
<td>0</td>
<td>0.50648</td>
<td>2332.07</td>
<td>0</td>
<td>0.507838</td>
</tr>
<tr>
<td></td>
<td>INBUSStep1</td>
<td>2195</td>
<td>1</td>
<td>0.307196</td>
<td>2333.68</td>
<td>1.61</td>
<td>0.227048</td>
</tr>
<tr>
<td></td>
<td>INBUSStep3</td>
<td>2196</td>
<td>2</td>
<td>0.186324</td>
<td>2333.37</td>
<td>1.3</td>
<td>0.265114</td>
</tr>
<tr>
<td>Northern Cardinal</td>
<td>NOCASStep3</td>
<td>4310</td>
<td>0</td>
<td>0.805512</td>
<td>4480.72</td>
<td>2.35</td>
<td>0.166978</td>
</tr>
<tr>
<td></td>
<td>NOCASStep2</td>
<td>4313</td>
<td>3</td>
<td>0.179734</td>
<td>4478.37</td>
<td>0</td>
<td>0.540698</td>
</tr>
<tr>
<td></td>
<td>NOCASStep1</td>
<td>4318</td>
<td>8</td>
<td>0.014753</td>
<td>4479.6</td>
<td>1.23</td>
<td>0.292324</td>
</tr>
<tr>
<td>Prairie Warbler</td>
<td>PRAWStep2</td>
<td>1249</td>
<td>0</td>
<td>0.731059</td>
<td>1366.21</td>
<td>1.11</td>
<td>0.364705</td>
</tr>
<tr>
<td></td>
<td>PRAWRand1</td>
<td>1251</td>
<td>2</td>
<td>0.268941</td>
<td>1365.1</td>
<td>0</td>
<td>0.635295</td>
</tr>
<tr>
<td></td>
<td>PRAWStep3</td>
<td>1537</td>
<td>288</td>
<td>2.12E-63</td>
<td>1562.2</td>
<td>197.1</td>
<td>1.01E-43</td>
</tr>
<tr>
<td>White-eyed Vireo</td>
<td>WEVISStep4</td>
<td>2731</td>
<td>0</td>
<td>0.422319</td>
<td>2895.28</td>
<td>0</td>
<td>0.610798</td>
</tr>
<tr>
<td></td>
<td>WEVISStep3</td>
<td>2731</td>
<td>0</td>
<td>0.422319</td>
<td>2900.4</td>
<td>5.12</td>
<td>0.047218</td>
</tr>
<tr>
<td></td>
<td>WEVISStep1</td>
<td>2733</td>
<td>2</td>
<td>0.155362</td>
<td>2896.44</td>
<td>1.16</td>
<td>0.341985</td>
</tr>
<tr>
<td>Wood Thrush</td>
<td>WOTHRand1</td>
<td>2127</td>
<td>0</td>
<td>0.736125</td>
<td>2213.78</td>
<td>1.31</td>
<td>0.244331</td>
</tr>
<tr>
<td></td>
<td>WOTHStep2</td>
<td>2130</td>
<td>3</td>
<td>0.164252</td>
<td>2212.47</td>
<td>0</td>
<td>0.470373</td>
</tr>
<tr>
<td></td>
<td>WOTHL3Out</td>
<td>2131</td>
<td>4</td>
<td>0.099624</td>
<td>2213.47</td>
<td>1</td>
<td>0.285296</td>
</tr>
<tr>
<td>Yellow-billed Cuckoo</td>
<td>YBCUSStep4</td>
<td>3112</td>
<td>0</td>
<td>0.451863</td>
<td>3174.64</td>
<td>0</td>
<td>0.426043</td>
</tr>
<tr>
<td></td>
<td>YBCUSStep3</td>
<td>3113</td>
<td>1</td>
<td>0.274069</td>
<td>3175.07</td>
<td>0.43</td>
<td>0.343621</td>
</tr>
<tr>
<td></td>
<td>YBCUSStep5</td>
<td>3113</td>
<td>1</td>
<td>0.274069</td>
<td>3175.87</td>
<td>1.23</td>
<td>0.230336</td>
</tr>
</tbody>
</table>

<sup>a</sup> Difference between AIC for that model and the model with the lowest AIC.
<sup>b</sup> Model weight calculated based on AIC.
<sup>c</sup> Difference between DIC for that model and the model with the lowest DIC.
<sup>d</sup> Model weight calculated based on DIC.
Table 2.6: Positive and negative habitat variable associations generated for each model from hierarchical logistic regression modeling for the state of Georgia based on the 1998 Georgia GAP Analysis 18-class landcover map (2003) and Georgia Partners in Flight (PIF) point counts (1995-2001). Model name reflects American Ornithologists Union species code plus the step in the model selection process in which the model was fitted. Variable abbreviations described in table 2.3 with an additional tag denoting the hierarchical level that it was taken from.

<table>
<thead>
<tr>
<th>Model</th>
<th>Positive</th>
<th>Negative</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACFLStep2</td>
<td>riparian_l2, forwet_l2, dfor_l1, riparian_l1</td>
<td></td>
</tr>
<tr>
<td>ACFLStep3</td>
<td>forwet_l2, dfor_l1, riparian_l1</td>
<td></td>
</tr>
<tr>
<td>ACFLStep1</td>
<td>wetland_l3, riparian_l2, forwet_l2, dfor_l1, riparian_l1</td>
<td></td>
</tr>
<tr>
<td>BHNUStep4</td>
<td>natopen_l3, efor_l2, efor_l1, clearcut_l1</td>
<td>riparian_l3, clearcut_l1</td>
</tr>
<tr>
<td>BHUNStep3</td>
<td>natopen_l3, efor_l2, efor_l1</td>
<td>riparian_l3, clearcut_l1</td>
</tr>
<tr>
<td>BHUNStep2</td>
<td>natopen_l3, efor_l2, efor_l1, clearcut_l1</td>
<td>riparian_l3, forwet_l2</td>
</tr>
<tr>
<td>CARWStep2</td>
<td>ifba_l4, open_l4, unnatopen_l3, natopen_l2, forest_l1</td>
<td></td>
</tr>
<tr>
<td>CARWStep3</td>
<td>open_l4, unnatopen_l3, natopen_l2, forest_l1, lowurb_l1</td>
<td></td>
</tr>
<tr>
<td>CARWRand1</td>
<td>lowurb_l4, unnatopen_l3, forest_l2, forest_l1</td>
<td></td>
</tr>
<tr>
<td>INBUStep3</td>
<td>urban_l4, natopen_l1, ag_l1</td>
<td>natopen_l4, forest_l2, forest_l1</td>
</tr>
<tr>
<td>INBUStep4</td>
<td>urban_l4, natopen_l1</td>
<td>natopen_l4, forest_l2, forest_l1</td>
</tr>
<tr>
<td>INBUStep1</td>
<td>urban_l4, unnatopen_l2, natopen_l1, ag_l1</td>
<td>natopen_l4, forest_l1</td>
</tr>
<tr>
<td>NOCAStep2</td>
<td>open_l4, open_l2, riparian_l2</td>
<td>ifba_l3, natopen_l1</td>
</tr>
<tr>
<td>NOCAStep1</td>
<td>lowurb_l4, open_l4, open_l2, riparian_l2, lowurb_l2</td>
<td>ifba_l3, lowurb_l3, natopen_l1</td>
</tr>
<tr>
<td>NOCAStep3</td>
<td>open_l4, open_l2, riparian_l2</td>
<td>natopen_l1</td>
</tr>
<tr>
<td>PRAWStep2</td>
<td>beaches_l2, efor_l1, clearcut_l1</td>
<td>natopen_l4, unnatopen_l4</td>
</tr>
<tr>
<td>PRAWRand1</td>
<td>efor_l3, natopen_l3, beaches_l2, efor_l1, clearcut_l1</td>
<td>natopen_l4, unnatopen_l4</td>
</tr>
<tr>
<td>PRAWStep3</td>
<td>efor_l1, clearcut_l1</td>
<td>natopen_l4, unnatopen_l4</td>
</tr>
<tr>
<td>WEVIStep4</td>
<td>clearcut_l2</td>
<td>ifba_l4, forest_l3, ag_l3</td>
</tr>
<tr>
<td>WEVIStep3</td>
<td>clearcut_l2</td>
<td>forest_l3, ag_l3</td>
</tr>
<tr>
<td>WEVIStep1</td>
<td>clearcut_l2, dfor_l1</td>
<td>ifba_l4, forest_l3, ag_l3</td>
</tr>
<tr>
<td>WOTHStep2</td>
<td>ifba_l4, forest_l3, Ifba_l1</td>
<td>wetland_l3, riparian_l2</td>
</tr>
<tr>
<td>WOTHRand1</td>
<td>ifba_l4, urban_l4, forwet_l2, dfor_l1, forwet_l1, Ifba_l1</td>
<td>wetland_l3, riparian_l2</td>
</tr>
<tr>
<td>WOTHL3Out</td>
<td>ifba_l4, open_l4, urban_l4, dfor_l1, forwet_l1, Ifba_l1</td>
<td>forwet_l2, riparian_l2, tlfba_l2, mfor_l1</td>
</tr>
<tr>
<td>YBCUStep3</td>
<td>ifba_l4, ifba_l3, unnatopen_l2, riparian_l1</td>
<td>urban_l4</td>
</tr>
<tr>
<td>YBCUStep5</td>
<td>ifba_l4, ifba_l3, unnatopen_l2, riparian_l1, dfor_l1</td>
<td></td>
</tr>
<tr>
<td>YBCUStep4</td>
<td>ifba_l4, ifba_l3, unnatopen_l2, riparian_l1</td>
<td></td>
</tr>
</tbody>
</table>
Table 2.7: Mean classification rates from three fold cross validation for the top three models for each species from hierarchical logistic regression modeling for the state of Georgia based on the 1998 Georgia GAP Analysis 18-class landcover map (2003) and Georgia Partners in Flight (PIF) point counts (1995-2001). Model name reflects American Ornithologists Union species code plus the step in the model selection process in which the model was fitted.

<table>
<thead>
<tr>
<th>Model</th>
<th>Threshold(^a)</th>
<th>pc(^b)</th>
<th>pe(^c)</th>
<th>cc0(^d)</th>
<th>cc1(^e)</th>
<th>ec(^f)</th>
<th>eo(^g)</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACFLStep2</td>
<td>0.221275</td>
<td>0.790769</td>
<td>0.209321</td>
<td>0.895259</td>
<td>0.422637</td>
<td>0.104741</td>
<td>0.577363</td>
</tr>
<tr>
<td>ACFLStep3</td>
<td>0.221275</td>
<td>0.790421</td>
<td>0.209579</td>
<td>0.895292</td>
<td>0.421354</td>
<td>0.104708</td>
<td>0.578646</td>
</tr>
<tr>
<td>ACFLStep1</td>
<td>0.221275</td>
<td>0.785088</td>
<td>0.214912</td>
<td>0.909498</td>
<td>0.347281</td>
<td>0.090502</td>
<td>0.652719</td>
</tr>
<tr>
<td>BHNUStep4</td>
<td>0.058869</td>
<td>0.855544</td>
<td>0.144456</td>
<td>0.890704</td>
<td>0.293377</td>
<td>0.109296</td>
<td>0.706623</td>
</tr>
<tr>
<td>BHNUStep3</td>
<td>0.058869</td>
<td>0.852697</td>
<td>0.147303</td>
<td>0.888156</td>
<td>0.285877</td>
<td>0.111844</td>
<td>0.714123</td>
</tr>
<tr>
<td>BHNUStep2</td>
<td>0.058869</td>
<td>0.849881</td>
<td>0.150119</td>
<td>0.884451</td>
<td>0.297193</td>
<td>0.115549</td>
<td>0.702807</td>
</tr>
<tr>
<td>CARWStep3</td>
<td>0.436612</td>
<td>0.583527</td>
<td>0.416473</td>
<td>0.617094</td>
<td>0.437672</td>
<td>0.204847</td>
<td>0.562328</td>
</tr>
<tr>
<td>CARWStep2</td>
<td>0.436612</td>
<td>0.558198</td>
<td>0.441802</td>
<td>0.586939</td>
<td>0.497105</td>
<td>0.214867</td>
<td>0.531915</td>
</tr>
<tr>
<td>CARWRand1</td>
<td>0.436612</td>
<td>0.524684</td>
<td>0.475316</td>
<td>0.601054</td>
<td>0.426138</td>
<td>0.398946</td>
<td>0.573862</td>
</tr>
<tr>
<td>INBUStep4</td>
<td>0.148722</td>
<td>0.746311</td>
<td>0.253689</td>
<td>0.783227</td>
<td>0.535008</td>
<td>0.216773</td>
<td>0.464992</td>
</tr>
<tr>
<td>INBUStep3</td>
<td>0.148722</td>
<td>0.745419</td>
<td>0.254581</td>
<td>0.781074</td>
<td>0.541308</td>
<td>0.218926</td>
<td>0.458692</td>
</tr>
<tr>
<td>INBUStep1</td>
<td>0.148722</td>
<td>0.743979</td>
<td>0.256021</td>
<td>0.779497</td>
<td>0.540677</td>
<td>0.220503</td>
<td>0.459323</td>
</tr>
<tr>
<td>NOCASStep3</td>
<td>0.518978</td>
<td>0.64867</td>
<td>0.35133</td>
<td>0.655019</td>
<td>0.642786</td>
<td>0.344981</td>
<td>0.357214</td>
</tr>
<tr>
<td>NOCASStep2</td>
<td>0.518978</td>
<td>0.641028</td>
<td>0.358972</td>
<td>0.657955</td>
<td>0.625672</td>
<td>0.342405</td>
<td>0.374328</td>
</tr>
<tr>
<td>NOCASStep1</td>
<td>0.518978</td>
<td>0.627885</td>
<td>0.372115</td>
<td>0.639399</td>
<td>0.617214</td>
<td>0.360601</td>
<td>0.382786</td>
</tr>
<tr>
<td>PRAWRand1</td>
<td>0.078492</td>
<td>0.863525</td>
<td>0.136475</td>
<td>0.905367</td>
<td>0.372336</td>
<td>0.094633</td>
<td>0.627664</td>
</tr>
<tr>
<td>PRAWStep2</td>
<td>0.078492</td>
<td>0.857899</td>
<td>0.142101</td>
<td>0.904119</td>
<td>0.31523</td>
<td>0.095881</td>
<td>0.68477</td>
</tr>
<tr>
<td>PRAWStep1</td>
<td>0.078492</td>
<td>0.76724</td>
<td>0.23726</td>
<td>0.790249</td>
<td>0.497105</td>
<td>0.209751</td>
<td>0.502895</td>
</tr>
<tr>
<td>WEVIS4step4</td>
<td>0.2063</td>
<td>0.721405</td>
<td>0.278595</td>
<td>0.795153</td>
<td>0.437672</td>
<td>0.204847</td>
<td>0.562328</td>
</tr>
<tr>
<td>WEVIS4step3</td>
<td>0.2063</td>
<td>0.719726</td>
<td>0.280274</td>
<td>0.785133</td>
<td>0.468085</td>
<td>0.214867</td>
<td>0.531915</td>
</tr>
<tr>
<td>WEVIS4step1</td>
<td>0.2063</td>
<td>0.718384</td>
<td>0.281616</td>
<td>0.792713</td>
<td>0.432416</td>
<td>0.207287</td>
<td>0.567584</td>
</tr>
<tr>
<td>WOTHStep2</td>
<td>0.112316</td>
<td>0.69968</td>
<td>0.30032</td>
<td>0.710762</td>
<td>0.612092</td>
<td>0.289238</td>
<td>0.387908</td>
</tr>
<tr>
<td>WOTHRand1</td>
<td>0.112316</td>
<td>0.697586</td>
<td>0.302414</td>
<td>0.710442</td>
<td>0.595977</td>
<td>0.289558</td>
<td>0.404023</td>
</tr>
<tr>
<td>WOTHL3out</td>
<td>0.112316</td>
<td>0.68834</td>
<td>0.31166</td>
<td>0.711896</td>
<td>0.502161</td>
<td>0.288104</td>
<td>0.497839</td>
</tr>
<tr>
<td>YBCUS4step5</td>
<td>0.193907</td>
<td>0.658245</td>
<td>0.341755</td>
<td>0.732864</td>
<td>0.348026</td>
<td>0.267136</td>
<td>0.651974</td>
</tr>
<tr>
<td>YBCUS4step4</td>
<td>0.193907</td>
<td>0.657727</td>
<td>0.342273</td>
<td>0.731038</td>
<td>0.352936</td>
<td>0.268962</td>
<td>0.647064</td>
</tr>
<tr>
<td>YBCUS4step3</td>
<td>0.193907</td>
<td>0.640592</td>
<td>0.359408</td>
<td>0.693978</td>
<td>0.418673</td>
<td>0.306022</td>
<td>0.581327</td>
</tr>
<tr>
<td>Mean</td>
<td>0.719791</td>
<td>0.280209</td>
<td>0.767703</td>
<td>0.45965</td>
<td>0.232297</td>
<td>0.54035</td>
<td></td>
</tr>
</tbody>
</table>

\(^a\) Probability of occurrence above which presence is predicted.
\(^b\) Rate that observations were correctly classified.
\(^c\) Rate that observations were incorrectly classified.
\(^d\) Rate that observations that were absences were correctly classified.
\(^e\) Rate that observations that were presences were correctly classified.
\(^f\) Rate that observations that were absences were incorrectly classified.
\(^g\) Rate that observations that were presences were incorrectly classified.
Table 2.8: Model ranking by AIC (Burnham and Anderson 2003), DIC (Spiegelhalter et al. 2003), and percent of observations correctly classified in three fold cross validation for the top three models from hierarchical logistic regression modeling for the state of Georgia based on the 1998 Georgia GAP Analysis 18-class landcover map (2003) and Georgia Partners in Flight (PIF) point counts (1995-2001). Model name reflects American Ornithologists Union species code plus the step in the model selection process in which the model was fitted.

<table>
<thead>
<tr>
<th>Model</th>
<th>AIC</th>
<th>DIC</th>
<th>Cross Validation</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACFLStep1</td>
<td>3</td>
<td>1</td>
<td>3</td>
</tr>
<tr>
<td>ACFLStep2</td>
<td>2</td>
<td>2</td>
<td>1</td>
</tr>
<tr>
<td>ACFLStep3</td>
<td>1</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>BHNUSStep2</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>BHNUSStep3</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>BHNUSStep4</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>CARWRand1</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>CARWStep2</td>
<td>2</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>CARWStep3</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>INBUSStep1</td>
<td>2</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>INBUSStep3</td>
<td>3</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>INBUSStep4</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>NOCASStep1</td>
<td>3</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>NOCASStep2</td>
<td>2</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>NOCASStep3</td>
<td>1</td>
<td>3</td>
<td>1</td>
</tr>
<tr>
<td>PRAWRand1</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>PRAWStep2</td>
<td>1</td>
<td>2</td>
<td>2</td>
</tr>
<tr>
<td>PRAWStep3</td>
<td>3</td>
<td>3</td>
<td>3</td>
</tr>
<tr>
<td>WEVISStep1</td>
<td>3</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>WEVISStep3</td>
<td>2</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>WEVISStep4</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>WOTHL2Out</td>
<td>3</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>WOTHRand1</td>
<td>1</td>
<td>3</td>
<td>2</td>
</tr>
<tr>
<td>WOTHStep2</td>
<td>2</td>
<td>1</td>
<td>1</td>
</tr>
<tr>
<td>YBCUSStep3</td>
<td>2</td>
<td>2</td>
<td>3</td>
</tr>
<tr>
<td>YBCUSStep4</td>
<td>1</td>
<td>1</td>
<td>2</td>
</tr>
<tr>
<td>YBCUSStep5</td>
<td>3</td>
<td>3</td>
<td>1</td>
</tr>
</tbody>
</table>
CHAPTER 3

HIERARCHICAL MODELS OF AVIAN DISTRIBUTIONS FOR THE STATE OF GEORGIA:

MODEL APPLICATION\textsuperscript{2}

\textsuperscript{2} Howell, J. E., M. J. Conroy, and J. T. Peterson. To be submitted to The Journal of Wildlife Management.
Abstract: In a previous paper to predict the distributions of nine breeding birds in the state of Georgia, we built models consisting of four hierarchical levels based on nested mapping units of decreasing area: 90,000 hectares (level 4), 3600 hectares (level 3), 144 hectares (level 2), and 5.76 hectares (level 1). In this paper we tested the ability of these models to choose between alternatives in management scenarios at two different scales. The coarse scale scenario involved choosing between three possible reserves in the Georgia Piedmont. The fine scale scenario involved recommending an optimal size for wildlife openings on a reserve. We evaluated the performance of these models at each scale by their ability to clearly differentiate between alternatives in the face of model uncertainty. In addition we tested the impact of using AIC or DIC as the model selection criterion and of estimation based on model averaging versus best model selection. The models were able to choose a clear alternative in the coarse scenario. The models failed to clearly differentiate among alternative opening sizes in the fine scale scenario. In both scenarios, model selection using AIC or DIC had no impact on the ranking of decision alternatives. Also in both scenarios, model averaging versus top model only selection had no impact on the rankings of management alternatives using AIC and DIC. We concluded that model uncertainty was such that the models will perform well in coarse applications, but fine applications are unlikely to allow discrimination among alternatives. We also concluded that although model selection method did not affect outcomes in our scenarios, the potential for influence exists.

Key Words: GAP Analysis, Hierarchical Models, MCMC, logistic regression, Bayesian, AIC, DIC, Georgia
The application of wildlife habitat relationship modeling to landscape ecology was initially limited due to the logistic difficulties involved in collecting information on a landscape scale which made such studies cost prohibitive (Urban 2000). With the advances in GIS and remote sensing, it became possible for ecologists to craft cost effective landscape approaches to WHR modeling. In their 1993 monograph, Scott et al. outlined a program to take advantage of this new capability on a nationwide basis called the GAP Analysis Project.

GAP Analysis identifies the gaps in representation of biodiversity in areas managed exclusively or primarily for the long-term conservation of native species and natural ecosystems (Scott et al. 1993). GAP analysis produces a map that consists of three layers (Scott et al. 1993). The first layer is a map of the vegetation communities derived from satellite imagery (Stoms 1998). The second layer uses the first layer in conjunction with a wildlife habitat relationship matrix to predict the distribution of every native vertebrate that occurs in the region (Csuti and Crist 1998). The third layer is a map of lands that have been designated as being set aside specifically for conservation (Edwards et al. 1998). These layers are then compared to determine which areas that have high species diversity are not protected for conservation (Scott et al. 1993, Jennings 2000). Based on these so-called gaps in what is called a “coarse filter” approach, managers can make optimal planning decisions in terms of land acquisition and finer scale studies (Scott et al. 1993, Kiester et al. 1996).

In previous work (Chapter 2), we introduced an alternative modeling method to the standard GAP wildlife habitat matrices. In that chapter, we explored the modeling process and some of the standard methods of model validation. In essence, we created a set of species models that predicted presence or absence based on a hierarchy of habitat information.
We modeled hierarchical spatial habitat patches through four levels of nested blocks of decreasing area. Level four blocks, intended to be about the size of a Georgia county, were 90,000 hectares. Level three blocks, intended to be about the size of a habitat reserve, were 3,600 hectares. Level two blocks, intended to represent a community, were 144 hectares. Finally, level one blocks at 5.76 hectares represented an inflated individual territory. At each level we took measures of habitat based on an 18-class land cover map generated by the Georgia GAP Analysis Project (Natural Resources Spatial Analysis Laboratory, 2001). These values took the form of percentages of land cover types, measures of forest fragmentation, and riparian area.

We then modeled the effects of habitat on the presence or absence of nine avian species within these areas using hierarchical logistic regression (Wong and Mason 1985, Bryk and Raudenbush 1993). Our models took the form:

\[
p_{ijklm} = \delta_{00} + \delta_{01}U_{1l} + \ldots + \delta_{0Q}U_{Ql} + \mu_{0l} + \chi_{01}V_{1kl} + \ldots + \chi_{0R}V_{Rkl} + \mu_{0kl} + \beta_{01}W_{1jkl} + \ldots + \beta_{0S}W_{Sjkl} + \mu_{0jkl} + \alpha_{1}X_{1ijkl} + \ldots + \alpha_{P}X_{Pijkl} + \gamma_{m},
\]

where \( p_{ijklm} \) is the log odds of presence in level 1 unit \( i \) from level 2 unit \( j \) in level 3 unit \( k \) in level 4 unit \( l \) in year \( m \). \( \delta_{01}, \ldots, \delta_{0Q} \) and \( U_{1l}, \ldots, U_{Ql} \) are the level 1 coefficients and habitat variables respectively. \( \beta_{00}, \ldots, \beta_{0S} \) and \( W_{1jkl}, \ldots, W_{Sjkl} \) are the level 2 coefficients and habitat variables respectively. \( \chi_{00}, \ldots, \chi_{0R} \) and \( V_{1kl}, \ldots, V_{Rkl} \) are the level 3 coefficients and habitat variables respectively. \( \delta_{00}, \ldots, \delta_{0Q} \) and \( U_{1l}, \ldots, U_{Ql} \) are the level 4 coefficients and habitat variables respectively. \( \mu_{0jkl}, \mu_{0jk}, \text{ and } \mu_{0ij} \) are the random effects of level membership. \( \gamma_{m} \) is the random effect of time period \( m \).

In their monograph, Scott et al. 1993 described the GAP analysis process as a “coarse filter”. That is, it is assumed that GAP will operate at coarse spatial scales, but not at fine scales (Edwards 1996, O’Neil et al. 1997, Hansen et al. 1999). Use of GAP predictions as a coarse filter
assumes predictive continuity among all levels. That is, because GAP models predict at scales that differ from the problems to which they are applied, the relation between habitat patterns and wildlife response must be assumed to hold between scales (Bolger et al. 1997). This is potentially troubling because patterns in landscape processes have been shown to vary across scales (Levin 1993). Further it remains unclear why applications of GAP at the finer scales are untenable. The predictive unit of the Georgia GAP Analysis project is a 30 by 30 m pixel, and it is possible in theory for GAP models to be used to predict responses to habitat change at this level. Based on this dichotomy, we sought to explore the ability of our models which we proposed as an alternative to GAP to perform in two management situations where decision-making was occurring at a coarse scale in one application and at a fine scale in the other.

Studies have investigated the impacts of uncertainty on wildlife habitat relationship models (Fielding and Bell 1997, Flather et al. 1997). Here, we evaluate the ability of our models to perform in our two applications, given uncertainty in the models. We ask whether the differences in the objective values between management alternatives are great enough in each of our applications to delineate a clearly superior alternative when model uncertainty is considered.

In our previous work (Chapter 2), we explored the use of two different information theoretic approaches to model selection. We conducted the bulk of model selection using Akaike’s Information Criterion (AIC) (Burnham and Anderson 2002). Afterwards, we revisited the model set and re-ranked the models using the Deviance Information Criterion (DIC) (Spiegelhalter et al. 2002). We found that in many cases the model rankings did not agree between the different approaches, and therefore sought to investigate the potential impacts of the application of either method in light of our management simulations.
METHODS

In order to test our models, we devised two hypothetical management scenarios. The first scenario involved choosing between three potential reserves. The second scenario involved choosing an optimum size for the creation of wildlife openings. The first situation tested the ability of the models to perform on coarse scale applications while the second tested their ability to perform at fine scales.

For the first scenario we selected three potential reserves in the Piedmont of Georgia (Figure 3.1). We selected areas already under some sort of protection in order to avoid obviously unsuitable areas such as those that included major metropolitan areas. The first area is the Hard Labor Creek State Park. The second and third areas are separate parts of the Oconee National Forest. The areas were chosen for their relative proximity to each other while still occurring in separate level 4 blocks. Habitat values within each block within the three areas were then measured from the land cover map used in the model fitting and a buffered coverage of streams for riparian area (Natural Resources Spatial Analysis Laboratory, 2001).

In our second scenario, we investigated determining an optimal size for wildlife openings. We used area three from our first scenario as the test reserve. We selected 20 locations within the area as locations for management (Figure 3.2). We then took habitat measures from the land-cover map after converting 57600 m² (opening1), 32400 m² (opening2), 14400 m² (opening3), and 3600 m² (opening 4) of forest to the clear-cut cover type at each of those locations. We also had the values calculated in scenario one for area three to use as a base line or a fifth option of no management.

We investigated the impacts of our model selection criteria by varying the way probability of occurrence was calculated for each species at each mapping unit. First we
calculated probability of occurrence based on the top models as ranked by AIC and DIC. Because we were also interested in exploring model averaging, we calculated model weights based on AIC and DIC (Table 3.1) with the model weights for DIC calculated using the same method as for AIC (Burnham and Anderson 2002). DIC, like AIC, is based on deviance plus a discounting factor for number of parameters, except for DIC it is an estimated effective number of parameters in place of a fixed parameter number (Spiegelhalter et al. 2002). We then produced model-averaged estimates for probability of occurrence based on these weights. This left us with four values for probability of occurrence for each species and mapping unit.

The next step in our management simulations was to develop an objective function that combined probability of occurrence for each species from the four model selection regimes in order to evaluate our management alternatives. Our function took the form:

\[ O_i = \sum_{j=1}^{q} p_{ij} \cdot v_j, \quad (2) \]

where \( O_i \) is the objective value at habitat block \( i \), \( p_{ij} \) is the probability of occurrence of species \( j \) at block \( i \), and \( v_j \) is the value assigned to species \( j \). We developed a set of values for each of our nine species based on Georgia Partners in Flight species of concern (Partners in Flight 2003) and consultations with Georgia Department of Natural Resources staff (Table 2). These values have implications for our results and applications as we discuss in the following sections.

To evaluate the performance of the models at different scales, it was necessary to consider uncertainty about model parameters and the ability of the models to recommend decisions in the face of model uncertainty. The Bayesian methods we used to estimate model parameters explicitly recognizes this uncertainty by defining parameters as random variables with a mean and variance that are then estimated (Gelman et al. 1995, Congdon 2002). Essentially we are interested in the distributions of \( O_i \) given the distributions of \( p_{ij} \) under each
alternative and how much separation occurs among alternatives. It was computationally too difficult to simultaneously calculate the distributions of $p_{ij}$ for all species and combine them in our objective function under our various model selection regimes. Instead we approximated this by obtaining .025, .05, .1, .25, .5, .75, .9, .95, and .975 quantiles of the posterior distributions of our model parameters and then used these to calculate probability of occurrence under our model selection regimes for each management alternative. The values at each quantile were then combined using our objective function. This left us with a cumulative distribution function (CDF) for our objective values at each mapping unit. Finally, an overall CDF for each management alternative was calculated by averaging the CDF’s for the mapping units associated with each alternative.

RESULTS

In Management Scenario One, Area Three was the preferred alternative followed by Area One and then Area Two. Area Three clearly differentiated itself from the alternative reserves (Figure 3.3-3.6). While Areas One and Two show little separation in their final CDF’s, Area Three is shows clear stochastic dominance over the other alternatives. The mean for Area Three is above topmost edge of the CDF’s for the other areas resulting in very little overlap between them.

Model selection criterion did not have an effect on the preferred alternative. Objective values from calculating probability of occurrence based on a single top model from AIC, a single top model from DIC, a model averaged estimate based on AIC weights, and model averaged estimates based on DIC resulted in the same preferred management alternative and the same ranking of alternatives in all cases. Actual values for the objective function CDF’s did vary
among regimes, but decision order did not vary. The values for the objective function were higher when using the models and weights based on AIC for model averaging and top model only selection than the values computed using DIC weights and selection. Model averaging versus top model only selection did not show any clear trends.

In management scenario two the management alternatives were ranked from most preferred to least as Opening1, Opening2, Opening3, Opening4, and No Management (Figure 3.7-3.10). Unlike in scenario one, however, there was not enough separation among alternatives to overcome uncertainty. The CDF’s for each alternative had little separation between them. There is almost 100% percent overlap among all the alternatives. Even at the extremes of the rankings, that is, between no management and opening level 1, there is nearly complete overlap.

As in scenario one, model selection criterion did not affect the ranking of the alternatives. Objective values from calculating probability of occurrence based on a single top model from AIC, a single top model from DIC, a model averaged estimate based on AIC weights, and model averaged estimates based on DIC weights again resulted in the same preferred management alternative and the same ranking of alternatives in all cases. Scenario two also showed similar trends in the magnitude of the objective values as seen in scenario one. That is using AIC selection and weights tended to produce larger objective values than DIC and model averaging did not produce and clear trends.

DISCUSSION

In our testing, model selection criterion had no impact on the eventual ranking of the decision in either of our scenarios. Intuitively, we know that this would not be the case in all possible applications. In any situation where models that predict based on differing inputs are
used in the same application their eventual agreement is based on how that application affects the inputs that are different between models. Indeed, the fact that the magnitude of the objective values was different when using AIC as the criterion versus using DIC shows that each approach would yield different objective values, although in this case that did not change the ordering of the alternatives.

Model averaging in some ways alleviates this problem. By including additional models weighted by their scores, one can reduce the impact of a few wildly differing models. In our scenarios, we found no difference in the ranking of alternatives when using model averaging.

Another way to influence the importance of selection criterion lies in the objective function itself. Each species was assigned a value based on our conservation goals. By increasing the values of species where AIC and DIC differed most starkly in their model rankings, one could increase the impact of the model selection criterion on the eventual decision.

The models succeeded in choosing an alternative in our coarse scale application. It is important to note however that success or failure actually depended greatly on our alternative reserves. If we had been trying to choose between area one and area two we would not have been able to make a decision. This boils down to how different are the alternative areas in terms of the model, in essence how different is one largely forested area on the Piedmont from another. In the case of our alternatives, the answer can be seen in the level four hierarchical habitat blocks. When we mapped the objective values for every mapping unit within the level four blocks that contained our alternative reserves, we found that the units in the same block as area three averaged higher scores than the other two level four blocks. Clearly there is something in that block that increases the objective values for any area within that unit. Whether it is the presence of Area Three within that unit that makes it desirable or something outside of the potential
reserve is immaterial, because the two cannot be separated. The fact is that only by considering more than the area within the preserve itself were we able to delineate the desirability of that alternative. This clearly demonstrates the potential and the power of our hierarchical approach. However, this sensitivity does come at a cost.

In Scenario Two, we were not able to choose a single best alternative in light of model uncertainty. The application in this case was occurring at a scale too fine for our models to clearly detect a difference. Consider that the model includes input from habitat blocks that cover 90,000 ha; the difference between creating 115.2 ha of clear cut and 64.8 ha is going to be difficult to detect. It is important to note that, because the model includes input from that scale and indeed predicts probability of occurrence at that scale, the model will return an answer to the problem. It is only in light of model uncertainty that we find the answer unsatisfactory.

This leaves us with the question of the significance of these results in terms of appropriate applications. We believe that our results demonstrate that the common coarse scale applications for which GAP models are most often proposed are appropriate scales for our model. By contrast, finer-scale applications are inappropriate because model uncertainty will be too great for the small changes in objective value that result from local habitat manipulation.

In Scenario Two, the management alternatives ended up being ranked consistently in favor of increasing clear cut area. Although we concluded that model uncertainty was such that we were unable to differentiate between adding the clear cuts and no management, we believe that this trend is nonetheless interesting. Essentially, the increase in objective values is most likely due to increasing the probability of open space birds into an area where that probability was near zero beforehand. One should note, however, that there is likely some point at which increasing clear cut area begins to negatively affect the forest birds causing a decrease in
objective value. Unfortunately, our alternatives did not include this possible switch point. An interesting follow up study would be to determine the switch point, that is the point at which increasing clear cut area leads to a decrease in objective value, assuming there is one.

These results illustrate an important facet of any multi-species conservation effort, namely that what is good for one species is not necessarily good and possibly even bad for another. Striking a balance between these conflicting goals is difficult at best, and needs to be taken into account when constructing an objective function. In our objective function we clearly valued certain species over others, for instance the Northern Cardinal was assigned a value of 10 while an Acadian Flycatcher was valued at 55. We believe this was consistent with the goals of conservation managers, however it necessarily means the function is more sensitive to changes in Flycatcher occurrence. These objective values also interact with probability of occurrence in important ways. For instance, a species may occur so infrequently that its probability of occurrence in a particular area will always be lower than a more common bird irrespective of the quality of the habitat. While this does not matter when comparing areas for a single species, it becomes important for multiple species applications. Consider our objective function where the value of an area for a given bird is simply its probability of occurrence multiplied by the value for that species. This means that a species, which may have a high relative value, may occur so infrequently that the resulting product could be lower than a more commonly occurring bird with a smaller assigned value. It is important in any final application, to make sure the eventual objective function accurately reflects the goals of the manager using it. We recommend assessing the models in terms of individual species first, then assessing the situation from a multi species standpoint.
Our findings have some important consequences for GAP applications. The main reason our models were able to perform in the coarse application was the implicit recognition of the impact of surrounding area on the potential reserves. GAP models do not include this recognition. Conversely, the failure at the finer scales may have been due to recognition of wide area effects reducing the ability to interpret effects at local scales. This means that GAP models, which deal only with local effects, may actually perform worse at coarse scales and better at finer scales than our models. Unfortunately this could not be tested in the same way our models were, because there is no good way to estimate statistical uncertainty for the GAP models. However, if true, this type of performance would further challenge the premise that GAP models can legitimately be used as “coarse filters” for conservation decision making (Conroy and Noon 1996).

We have shown that our models were able to make decisions in the coarse scale application, but not in the fine scale. We showed the importance of including information surrounding an area when evaluating it for a potential reserve. We showed that, while choice of model selection criterion did not have an impact on decision making in our scenarios, under other scenarios this might not be the case. Finally, we showed that model uncertainty is the principle limiting factor in the ability of our models to choose between management alternatives.
LITERATURE CITED


O’Neill, R. V., Hunsaker, C. T., Jones, K. B., Riitters, K. H., Wickham, J. D., Schwartz, P. M.,
Goodman, I. A., Jackson, B. L., and Baillargeon, W. S. 1997. Monitoring Environmental
Quality at the Landscape Scale. *Bioscience.* 47:513-519


Scott, J. M., Davis, F., Cutsi, B., Noss, R., Butterfield, B., Groves, C., Anderson, H., Caicco, S.,
Geographic Approach to Protection of Biological Diversity. *Wildlife Monographs.* 123:1-
41

of Model Complexity and Fit. *Journal of the Royal Statistical Society Series B-Statistical
Methodology.* 64:583-616

Scot, D. M. and M. D. Jennings eds, Moscow, ID: National GAP Analysis Project. 1.1-
1.6


Wong, G. Y. and Mason, W. M. 1985. The Hierarchical Logistic-Regression Model for
Figure 3.1: Management Scenario 1, three potential reserves, Hard Labor Creek State Park (Area 1) and the Oconee National Forest (Area 2, Area 3).
Figure 3.2: Locations of Management Patches on Area3 for Management Scenario 2.
Figure 3.3: Management Scenario One: CDF’s of the mean objective values for Area1, Area2 and Area 3 using model averaging based on AIC weights and habitat values from the Georgia GAP 1998 18-class Landcover Map.
Figure 3.4: Management Scenario One: CDF’s of the mean objective values for Area1, Area2, and Area3 based on model averaging using DIC weights and habitat values from the Georgia GAP 1998 18-Class Landcover Map.
Figure 3.5: Management Scenario One: CDF’s of the mean objective values for Area1, Area2, and Area3 based on the top model selected by AIC and habitat values from the Georgia GAP 1998 18-Class Landcover Map.
Figure 3.6: Management Scenario One: CDF’s of the mean objective values for Area1, Area2, and Area3 based on the top model selected by DIC and habitat values from the Georgia GAP 1998 18-Class Landcover Map.
Figure 3.7: Management Scenario Two: CDF CDF’s of the mean objective values for the management alternatives using model averaging based on AIC weights and habitat values from the Georgia GAP 1998 18-class Landcover Map.
Figure 3.8: Management Scenario Two: CDF’s of the mean objective values for the management alternatives using model averaging based on DIC weights and habitat values from the Georgia GAP 1998 18-class Landcover Map.
Figure 3.9: Management Scenario Two: CDF’s of the mean objective values for the management alternatives based on the top model selected by AIC and habitat values from the Georgia GAP 1998 18-class Landcover Map.
Figure 3.10: Management Scenario Two: CDF’s of the mean objective values for the management alternatives based on the top model selected by DIC and habitat values from the Georgia GAP 1998 18-class Landcover Map.
Table 3.1: Top three model set for each species for each model from hierarchical logistic regression modeling for the state of Georgia based on the 1998 Georgia GAP Analysis 18-class landcover map (2003) and Georgia Partners in Flight (PIF) point counts (1995-2001). Model name reflects American Ornithologists Union species code plus the step in the model selection process in which the model was fitted (example: ACFLStep3 is the third model fitted using stepwise selection for the Acadian Flycatcher). Models Ranked by AIC (Burnham and Anderson 2002) and DIC (Spiegelhalter et al. 2003).

<table>
<thead>
<tr>
<th>Species</th>
<th>Model</th>
<th>AIC</th>
<th>ΔAIC&lt;sup&gt;a&lt;/sup&gt;</th>
<th>AIC wt&lt;sup&gt;b&lt;/sup&gt;</th>
<th>DIC</th>
<th>ΔDIC&lt;sup&gt;c&lt;/sup&gt;</th>
<th>DIC wt&lt;sup&gt;d&lt;/sup&gt;</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACFL</td>
<td>ACFLStep3</td>
<td>2647</td>
<td>0</td>
<td>0.628532</td>
<td>2863.96</td>
<td>40.88</td>
<td>1.33E-09</td>
</tr>
<tr>
<td>ACFL</td>
<td>ACFLStep2</td>
<td>2649</td>
<td>2</td>
<td>0.231224</td>
<td>2863.64</td>
<td>40.56</td>
<td>1.56E-09</td>
</tr>
<tr>
<td>ACFL</td>
<td>ACFLStep1</td>
<td>2650</td>
<td>3</td>
<td>0.140244</td>
<td>2823.08</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>BHNU</td>
<td>BHUNStep3</td>
<td>1214</td>
<td>0</td>
<td>0.574097</td>
<td>1303.8</td>
<td>0</td>
<td>0.531673</td>
</tr>
<tr>
<td>BHNU</td>
<td>BHNUStep4</td>
<td>1215</td>
<td>1</td>
<td>0.348207</td>
<td>1305.25</td>
<td>1.45</td>
<td>0.257502</td>
</tr>
<tr>
<td>BHNU</td>
<td>BHNUStep2</td>
<td>1218</td>
<td>4</td>
<td>0.077696</td>
<td>1305.65</td>
<td>1.85</td>
<td>0.210825</td>
</tr>
<tr>
<td>CARW</td>
<td>CARWStep3</td>
<td>4440</td>
<td>0</td>
<td>0.50648</td>
<td>4594.21</td>
<td>0</td>
<td>0.714337</td>
</tr>
<tr>
<td>CARW</td>
<td>CARWStep2</td>
<td>4441</td>
<td>1</td>
<td>0.307196</td>
<td>4596.5</td>
<td>2.29</td>
<td>0.227319</td>
</tr>
<tr>
<td>CARW</td>
<td>CARWRand1</td>
<td>4442</td>
<td>2</td>
<td>0.186324</td>
<td>4599.22</td>
<td>5.01</td>
<td>0.058344</td>
</tr>
<tr>
<td>INBU</td>
<td>INBUSStep4</td>
<td>2194</td>
<td>0</td>
<td>0.50648</td>
<td>2332.07</td>
<td>0</td>
<td>0.507838</td>
</tr>
<tr>
<td>INBU</td>
<td>INBUSStep1</td>
<td>2195</td>
<td>1</td>
<td>0.307196</td>
<td>2333.68</td>
<td>1.61</td>
<td>0.227048</td>
</tr>
<tr>
<td>INBU</td>
<td>INBUSStep3</td>
<td>2196</td>
<td>2</td>
<td>0.186324</td>
<td>2333.37</td>
<td>1.3</td>
<td>0.265114</td>
</tr>
<tr>
<td>NOCA</td>
<td>NOCASStep3</td>
<td>4310</td>
<td>0</td>
<td>0.805512</td>
<td>4480.72</td>
<td>2.35</td>
<td>0.166978</td>
</tr>
<tr>
<td>NOCA</td>
<td>NOCASStep2</td>
<td>4313</td>
<td>3</td>
<td>0.179734</td>
<td>4478.37</td>
<td>0</td>
<td>0.540698</td>
</tr>
<tr>
<td>NOCA</td>
<td>NOCASStep1</td>
<td>4318</td>
<td>8</td>
<td>0.014753</td>
<td>4479.6</td>
<td>1.23</td>
<td>0.292324</td>
</tr>
<tr>
<td>PRAW</td>
<td>PRAWStep2</td>
<td>1249</td>
<td>0</td>
<td>0.731059</td>
<td>1366.21</td>
<td>1.11</td>
<td>0.364705</td>
</tr>
<tr>
<td>PRAW</td>
<td>PRAWRand1</td>
<td>1251</td>
<td>2</td>
<td>0.268941</td>
<td>1365.1</td>
<td>0</td>
<td>0.635295</td>
</tr>
<tr>
<td>PRAW</td>
<td>PRAWStep3</td>
<td>1537</td>
<td>288</td>
<td>2.12E-63</td>
<td>1562.2</td>
<td>197.1</td>
<td>1.01E-43</td>
</tr>
<tr>
<td>WEVI</td>
<td>WEVISStep4</td>
<td>2731</td>
<td>0</td>
<td>0.422319</td>
<td>2895.28</td>
<td>0</td>
<td>0.610798</td>
</tr>
<tr>
<td>WEVI</td>
<td>WEVISStep3</td>
<td>2731</td>
<td>0</td>
<td>0.422319</td>
<td>2900.4</td>
<td>5.12</td>
<td>0.047218</td>
</tr>
<tr>
<td>WEVI</td>
<td>WEVISStep1</td>
<td>2733</td>
<td>2</td>
<td>0.155362</td>
<td>2896.44</td>
<td>1.16</td>
<td>0.341985</td>
</tr>
<tr>
<td>WOTH</td>
<td>WOTHRand1</td>
<td>2127</td>
<td>0</td>
<td>0.736125</td>
<td>2213.78</td>
<td>1.31</td>
<td>0.244331</td>
</tr>
<tr>
<td>WOTH</td>
<td>WOTHStep2</td>
<td>2130</td>
<td>3</td>
<td>0.164252</td>
<td>2212.47</td>
<td>0</td>
<td>0.470373</td>
</tr>
<tr>
<td>WOTH</td>
<td>WOTHL3Out</td>
<td>2131</td>
<td>4</td>
<td>0.099624</td>
<td>2213.47</td>
<td>1</td>
<td>0.285296</td>
</tr>
<tr>
<td>YBCU</td>
<td>YBCUSStep4</td>
<td>3112</td>
<td>0</td>
<td>0.451863</td>
<td>3174.64</td>
<td>0</td>
<td>0.426043</td>
</tr>
<tr>
<td>YBCU</td>
<td>YBCUSStep3</td>
<td>3113</td>
<td>1</td>
<td>0.274069</td>
<td>3175.07</td>
<td>0.43</td>
<td>0.343621</td>
</tr>
<tr>
<td>YBCU</td>
<td>YBCUSStep5</td>
<td>3113</td>
<td>1</td>
<td>0.274069</td>
<td>3175.87</td>
<td>1.23</td>
<td>0.230336</td>
</tr>
</tbody>
</table>

<sup>a</sup> Difference between AIC for that model and the model with the lowest AIC.
<sup>b</sup> Model weight calculated based on AIC.
<sup>c</sup> Difference between DIC for that model and the model with the lowest DIC.
<sup>d</sup> Model weight calculated based on DIC.
Table 3.2: Species Values ($v_j$) for the objective function used to evaluate predictions in management scenarios from hierarchical logistic regression models for the state of Georgia based on the 1998 Georgia GAP Analysis 18-class landcover map (2003) and Georgia Partners in Flight (PIF) point counts (1995-2001). Values assigned based on PIF species of concern and priorities of the Georgia Department of Natural Resources.

<table>
<thead>
<tr>
<th>Species</th>
<th>Scientific Name</th>
<th>Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Acadian Flycatcher</td>
<td><em>Empidonax Virescens</em></td>
<td>55</td>
</tr>
<tr>
<td>Brown-headed Nuthatch</td>
<td><em>Sitta pusilla</em></td>
<td>45</td>
</tr>
<tr>
<td>Carolina Wren</td>
<td><em>Thryothorus ludovicianus</em></td>
<td>25</td>
</tr>
<tr>
<td>Indigo Bunting</td>
<td><em>Passerina cyanea</em></td>
<td>45</td>
</tr>
<tr>
<td>Northern Cardinal</td>
<td><em>Cardinalis cardinalis</em></td>
<td>10</td>
</tr>
<tr>
<td>Prairie Warbler</td>
<td><em>Dendroica discolor</em></td>
<td>45</td>
</tr>
<tr>
<td>White-eyed Vireo</td>
<td><em>Coccyzus Americanus</em></td>
<td>35</td>
</tr>
<tr>
<td>Wood Thrush</td>
<td><em>Vireo griseus</em></td>
<td>50</td>
</tr>
<tr>
<td>Yellow-billed Cuckoo</td>
<td><em>Hylocichla mustelina</em></td>
<td>45</td>
</tr>
</tbody>
</table>
CHAPTER 4

SUMMARY

The goals of my project were to build empirical models that predicted distributions for multiple avian species using data sources available to the Georgia Gap Analysis Project and to test the application of those models to decision making. We were interested in testing the “coarse filter” assumptions of GAP and the impacts of model uncertainty (Conroy and Noon 1996).

We built our empirical models nine avian species based on a hierarchy of habitat information. Habitat was measured at four hierarchical scales using nested blocks; level one (5.76 ha), level two (144 ha), level three (3600 ha), and level four (90,000 ha). These habitat measures were calculated using the composition of the 1998 18-class landcover map from the Georgia GAP analysis project (Natural Resource Spatial Analysis Laboratory 2001). We used the Georgia Partners in Flight point count database to measure presence or absence at stations distributed throughout the state for Acadian Flycatcher (*Empidonax virescens*), Brown-Headed Nuthatch (*Sitta pusilla*), Carolina Wren (*Thryothorus ludovicianus*), Indigo Bunting (*Passerina cyanea*), Northern Cardinal (*Cardinalis cardinalis*), Prairie Warbler (*Dendroica discolor*), Yellow-Billed Cuckoo (*Coccyzus americanus*), White-eyed Vireo (*Vireo griseus*), and Wood Thrush (*Hylocichla mustelina*). We then fit our models using hierarchical logistic regression (Wong and mason 1985, Bryk and Raudenbush 1992).

Hierarchical logistic regression works by taking variables from one hierarchical level and substituting linear combinations of variables from the next higher hierarchical level. In the case of our modeling, we only substituted hierarchical information for intercept values. This meant
that our models did not reflect interaction between variables from across levels. Our model took the form:
\[ p_{ijklm} = \delta_{00} + \delta_{01}U_{1l} + \ldots + \delta_{0Q}U_{Qi} + \mu_{0j} + \chi_{01}V_{1kl} + \ldots + \chi_{0R}V_{Rkl} + \mu_{0k} + \beta_{01}W_{ijkl} + \ldots + \beta_{0S}W_{Sijkl} + \mu_{0jkl} + \alpha_{1}X_{1ijkl} + \ldots + \alpha_{P}X_{Pijkl} + \gamma_{m}, \]  

(1)
where \( p_{ijklm} \) is the log odds of presence in level 1 unit \( i \) from level 2 unit \( j \) in level 3 unit \( k \) in level 4 unit \( l \) in year \( m \). \( \delta_{01} \ldots \delta_{0Q} \) and \( U_{1l} \ldots U_{Qi} \) are the level 1 coefficients and habitat variables respectively. \( \beta_{00} \ldots \beta_{0S} \) and \( W_{ijkl} \ldots W_{Sijkl} \) are the level 2 coefficients and habitat variables respectively. \( \chi_{01} \ldots \chi_{0R} \) and \( V_{1kl} \ldots V_{Rkl} \) are the level three coefficients and habitat variables respectively. \( \delta_{00} \ldots \delta_{0Q} \) and \( U_{1l} \ldots U_{Qi} \) are the level four coefficients and habitat variables respectively. \( \mu_{0jkl}, \mu_{0jk}, \) and \( \mu_{0j} \) are the random effects of level membership. \( \gamma_{m} \) is the random effect of time period \( m \).

The random effects \( \mu_{0jkl}, \mu_{0jk}, \gamma_{m}, \) and \( \mu_{0j} \) in our models would have been difficult to estimate using standard maximum likelihood techniques (Wong and Mason 1985). For this reason, we chose to conduct our modeling in a Bayesian framework using MCMC (Gelman et al. 1995, Congdon 2001, Link et al. 2002), specifically the program WinBugs (Lunn 2003).

We created a set of candidate models for each species based on the results of all possible models fit without random effects. We took the top three models from this run based on AIC (Burnham and Anderson 2002), two models chosen at random, models in which one hierarchical level, and the full model. After initial results showed the best model was likely not included in this set, we conducted stepwise selection until we were satisfied with the final set.

We then took the top three models for each species based on AIC and calculated DIC for them (Spiegelhalter et al. 2002). We found that DIC usually disagreed with the rankings produced by AIC. We estimated out of sample prediction by conducting three fold cross
validation on each model in this final set. Performance was generally good with a mean overall correct classification rate for all models of 70%. This is somewhat misleading however, because correct classification was heavily skewed toward predicting absence, which averaged 76%. Classification rates of presence were not as high averaging only 45%. We attribute this discrepancy to the fact that none of the species we modeled occurred more frequently than at 50 percent of our stations and most, seven of nine, occurred at less than 25 percent of our stations.

In many cases, the habitat relationships predicted by our models were different from what we expected, based on our literature review. For instance, the models for the Prairie Warbler predicted negative relationships for natural and unnatural open areas at level four. This is contrary to the bird’s documented preference for shrubby open habitats. However, these discrepancies all occurred at the coarse spatial scales. This leads us to conclude that in considering hierarchical spatial relationships, the sort of localized habitat preferences commonly studied for most species may not be accurate guides to how the species relates to those same factors at coarser spatial scales. This is explicitly shown by the indigo bunting, which has a positive relationship to natural open areas at level one, but then shows a negative relationship to these same areas at the coarser scale of level four.

After our models were built, we applied them to two hypothetical management situations. Management Scenario One involved choosing between three potential reserves on the Georgia Piedmont; Hard Labor Creek State Park (Area One) and two parts of the Oconee national Forest (Area Two and Area Three). Management Scenario Two involved choosing and optimum size for wildlife openings created in a forested reserve. Scenario one tested performance at coarse scales and Scenario two tested fine scales. We evaluated model performance for each scenario in terms of model uncertainty. That is, was the difference returned by our objective function
between decision alternatives great enough for us to choose a clearly superior alternative in light of model uncertainty. Finally, due to the lack of agreement between AIC and DIC about model selection, we sought to test the impact of using either criteria as well as model averaging based on them.

In Scenario One, Area Three was the clearly superior alternative. While Areas One and Two were too similar to distinguish between them, our results for Area three showed clear separation. When we mapped the objective values for every mapping unit within the level four blocks that contained our alternative reserves, we found that the units in the same block as area three averaged higher scores than the other two level four blocks. We concluded that it was information derived from the level four scale that enabled the separation in our management alternatives. We concluded that application at this scale was appropriate for our models, but that GAP models which do not consider area beyond the preserve itself may perform worse challenging the “coarse filter” assumptions of GAP.

In Scenario Two, there was no clearly superior alternative. Although objective value increased with larger clearcuts, separation between alternatives was not great enough to overcome model uncertainty. We concluded that while our models will demonstrate differences when changes are made at fine scales, the inclusion of wide area effects makes these differences too small in light of model uncertainty, and therefore our models should not be applied to decision-making at these scales. GAP models, which do not include wide area effects, however, may perform better at these scales.

In both scenarios, model selection criteria did not have an effect on the ordering of alternative decisions. We also found no difference in order when using model averaging. Actual
objective values were different when each method was used. This led us to conclude that while no effect was observed in our Scenarios, the potential for impact exists.

The results of our study speak to the efficacy of our approach. For all of our species, a hierarchy of levels was involved in the final best model. The ability of our models to perform in the coarse scale scenario was based on the explicit recognition of the importance of multi-scale information. We acknowledge that for many species in the GAP Analysis process this type of modeling would not be possible due to constraints on available information. We believe our results show the value in doing so in cases where it is possible.
LITERATURE CITED


