

The Influence of Probability of Detection When Modeling Species
Occurrence Using GIS and Survey Data

By
Alison K. Williams

Dissertation submitted to the Faculty of the
Virginia Polytechnic Institute and State University
In partial fulfillment of the requirements for the degree of

Doctorate of Philosophy
in
Wildlife Science

Jim Berkson, Chair
Paul Angermeier
Carlyle Brewster
Marcella Kelly
Dean Stauffer
Christopher Zobel

December 2003
Blacksburg, Virginia

Keywords: habitat suitability model, probability of detection, salamander, Bayesian logistic regression, multivariate statistics, ecological niche factor analysis

Copyright 2003, Alison K. Williams

The Influence of Probability of Detection When Modeling Species Occurrence Using GIS and Survey Data

Alison K Williams

(ABSTRACT)

I compared the performance of habitat models created from data of differing reliability. Because the reliability is dependent on the probability of detecting the species, I experimented to estimate detectability for a salamander species. Based on these estimates, I investigated the sensitivity of habitat models to varying detectability.

Models were created using a database of amphibian and reptile observations at Fort A.P. Hill, Virginia, USA. Performance was compared among modeling methods, taxa, life histories, and sample sizes. Model performance was poor for all methods and species, except for the carpenter frog (*Rana virgatipes*). Discriminant function analysis and ecological niche factor analysis (ENFA) predicted presence better than logistic regression and Bayesian logistic regression models. Database collections of observations have limited value as input for modeling because of the lack of absence data. Without knowledge of detectability, it is unknown whether non-detection represents absence.

To estimate detectability, I experimented with red-backed salamanders (*Plethodon cinereus*) using daytime, cover-object searches and nighttime, visual surveys. Salamanders were maintained in enclosures (n = 124) assigned to four treatments, daytime–low density, daytime–high density, nighttime–low density, and nighttime–high density. Multiple observations of each enclosure were made. Detectability was higher using daytime, cover-object searches (64%) than nighttime, visual surveys (20%). Detection was also higher in high-density (49%) versus low-density enclosures (35%).

Because of variation in detectability, I tested model sensitivity to the probability of detection. A simulated distribution was created using functions relating habitat suitability to environmental variables from a landscape. Surveys were replicated by randomly selecting locations (n = 50,

100, 200, or 500) and determining whether the species was observed, based on the probability of detection ($p = 40\%$, 60% , 80% , or 100%). Bayesian logistic regression and ENFA models were created for each sample. When detection was 80 – 100%, Bayesian predictions were more correlated with the known suitability and identified presence more accurately than ENFA.

Probability of detection was variable among sampling methods and effort. Models created from presence/absence data were sensitive to the probability of detection in the input data. This stresses the importance of quantifying detectability and using presence-only modeling methods when detectability is low. If planning for sampling as an input for suitability modeling, it is important to choose sampling methods to ensure that detection is 80% or higher.

ACKNOWLEDGMENTS

I would like to thank my advisor, Dr. Jim Berkson, and committee members, Dr. Paul Angermeier, Dr. Carlyle Brewster, Dr. Dean Stauffer, and Dr. Chris Zobel. They have always been available, ready to help, and have provided insightful reviews and ideas for this project. A special thanks also goes to Dr. Marcella Kelly, who graciously agreed to take part in the exam committee for my defense. I appreciate the time and effort that went into reading and critiquing this dissertation during a busy end of the semester. Thanks also go to all of the individuals at the Conservation Management Institute that helped me along the way, including Jeff Waldon, Scott Klopfer, Stacy McNulty, Verl Emrick, and Lola Roghair. If it wasn't for my experiences with CMI (formerly the Fish and Wildlife Information Exchange) while I was an undergraduate at Virginia Tech, it is likely that I would not be here as a graduate student.

There are also many people and organizations vital to the success of this project. Dr. Joe Mitchell at the University of Richmond provided his database of amphibian and reptile observations at Fort A.P. Hill that made Chapter 2 possible. Joe's experience with these species far surpassed mine and he was a great source of ideas and information at the beginning of this project. Dr. Henry Wilbur at the University of Virginia provided valuable advice and assistance with conducting the experiment in Chapter 3 at the Mountain Lake Biological Research Station. I also owe a large thank you to the other previous and current graduate students and members of the Berkson lab, Beth Walls, Jodi Dew, Michelle Davis, Lenka Hurton, Jay McGhee, Mary Tilton, Whitney Grogan, and David Hata, who have provided wonderful reviews of manuscripts and presentations and plenty of other helpful advice.

And finally, I would like to specially thank all of my friends and family that have been much needed source of support and sanity during my graduate school adventure. Thanks to my friends Mandy and Till Rosenberger who provided a welcome break from work every week and constant friendship and previous officemate Tim Copeland for always being able to make me laugh and reminding me of all of the important things in life other than grad school. A huge thanks to the Blacksburg Rescue Squad and all of my friends and fellow workers, I am proud to be a member. Volunteering with the rescue squad has given me new insight on the world, taught me skills I would have otherwise never learned, and been a source of support and purpose when graduate school seemed daunting. Special thanks also goes out to my tres amigos, Eric Stanley, Brian Ekey, and Chris Rourke. And especially to Chris, who has always been an encouraging source of motivation and welcome escape from work. And finally, none of this would have been possible without the support and love from my parents and family who have always encouraged me to succeed at whatever I desired.

TABLE OF CONTENTS

TITLE PAGE	i
ABSTRACT	ii
ACKNOWLEDGMENTS	iv
TABLE OF CONTENTS	vi
LIST OF TABLES	ix
LIST OF FIGURES	x
CHAPTER 1: LITERATURE REVIEW	1
Goals of Habitat Modeling	1
Habitat Modeling Philosophy	2
Assumptions of Habitat Modeling	2
Methods of Habitat Modeling	4
Logistic Regression.....	5
Discriminant Function Analysis	5
Artificial Neural Networks.....	6
Classification And Regression Trees	6
Ecological Niche Factor Analysis.....	7
Bayesian Statistical Methods	8
Applications of Bayesian Statistics in the Wildlife Field	9
Geographic Information Systems and Habitat Modeling	10
Factors Influencing Probability of Occurrence Prediction	11
Model Testing	14
Model Testing Using Data Simulation	16
Habitat Modeling for Amphibians and Reptiles	16
LITERATURE CITED	17
CHAPTER 2: A COMPARISON OF MODELS USING OCCURRENCE DATA TO PREDICT THE PRESENCE OR ABSENCE OF REPTILE AND AMPHIBIAN SPECIES	26
ABSTRACT	26
INTRODUCTION	27
Species Distribution Data	27
Problems with Presence-Only Data	28
Ecological Niche Factor Analysis	28
Models Using Presence and Absence Data	29
Bayesian Models	30
Model Comparisons	31
Application to Amphibian and Reptile Taxa	31
Study Objectives	32
METHODS	32
Survey Data	32
Predictor Variables	33
Model Creation	34

RESULTS	36
Ecological Niche Factor Analyses	36
Discriminant Function Analysis	38
Logistic Regression	40
Bayesian Logistic Regression	40
Model Comparison	41
DISCUSSION	42
Comparison of Model Accuracy	42
Taxonomic Differences in Model Performance	43
Correlation Between Sample Size and Model Accuracy	44
Appropriateness of Pseudo-Absence Data	45
Appropriateness of Observation Data for Modeling	46
Conclusions	47
LITERATURE CITED	47
CHAPTER 3: REDUCING FALSE ABSENCES IN SURVEY DATA: DETECTION PROBABILITIES OF RED-BACKED SALAMANDERS	70
ABSTRACT	70
INTRODUCTION	70
METHODS	72
Experimental Enclosures	73
Data Collection	74
Statistical Analyses	75
RESULTS	76
DISCUSSION	80
Sampling Method Effects on Detection Probability	80
Sampling Method Effects on Salamander Condition	81
Salamander Density Effects on Detection Probability	81
Weather Influences on Salamander Detection	82
Impact of Sampling Effort on Detection Probabilities	83
Experiment Assumptions	84
Recommendations for Sampling	85
Management Implications	86
LITERATURE CITED	86
CHAPTER 4: EFFECTS OF SPECIES DETECTABILITY AND SAMPLE SIZE ON PREDICTING HABITAT SUITABILITY: AN ANALYSIS OF BAYESIAN AND ECOLOGICAL NICHE FACTOR ANALYSIS MODELS USING SIMULATED DATA 97	
ABSTRACT	97
INTRODUCTION	98
METHODS	101
Landscape Data	101
Generating the Distribution of a Virtual Species	102
Generating Sampling Data	103
Bayesian Logistic Regression Model	104
Ecological Niche Factor Analysis	105

Model Evaluation	106
RESULTS	107
Habitat Suitability	107
Environmental Predictor Variables	107
Bayesian Logistic Regression Models	108
Ecological Niche Factor Analysis Models	109
Model Comparison	110
DISCUSSION	112
Relationship Between Probability of Detection and Bayesian Model Accuracy	112
Bayesian Logistic Regression versus ENFA	112
Influence of Sample Size on Model Accuracy	114
Assumptions of Simulation	114
Conclusions	115
LITERATURE CITED	116

LIST OF TABLES

Table 2. 1. Amphibian and reptile species selected for modeling probability of occurrence and the number of presence and absence locations recorded at Fort A.P. Hill during 1992-1998.	52
Table 2. 2. Habitat variables measured for each location in the presence/absence data set.	53
Table 2. 3. Eigenvalues and percent of variation explained for the first marginality factor, marginality scores, specialization scores, and the % of information explained by the factors included in the habitat suitability model for each species calculated using ecological niche factor analysis.	54
Table 2. 4. Kappa statistics for ecological niche factor (ENFA), discriminant analysis (DFA), logistic regression (LR), and Bayesian models for amphibian and reptiles species at Fort A.P. Hill.	55
Table 2. 5. Comparison of mean sensitivity, specificity, and Kappa statistics for each of the four modeling methods.	56
Table 3. 1. Number of red-backed salamanders, by treatment group, unable to be located at the end of the experiment in July 2001. These salamanders are assumed to have either died or escaped. The percentages are based on the total number of salamanders, however high density treatments contained twice the number of salamanders than low density treatments.	90
Table 4. 1. Mean and standard deviation of the number of presence locations in the data sets used to build Bayesian logistic regression and ENFA models, listed in order from greatest to fewest presence locations.	119
Table 4. 2. Analysis of variance of environmental predictor variables as sampled by the “virtual ecologist” for 50 repetitions. F statistics and p values listed for each factor, with factors significant at the 95% level in bold.	120
Table 4. 3. Comparison of model accuracy using least square means and 95% confidence intervals among Bayesian logistic regression and ENFA models built with data from differing probability of detection scenarios.	121
Table 4. 4. Analysis of variance of model accuracy, as measured by correlation of predicted and true suitability, model sensitivity (proportion of presences correctly predicted), and specificity (proportion of absences correctly predicted). Analysis of fifteen replications (n = 480) of each combination of model type (Bayesian logistic regression and ENFA), sample sizes (n = 50, 100, 200, and 500) and probabilities of detection (p = 40%, 60%, 80%, and 100%) as sampled by the “virtual ecologist.” F statistics and p values listed for each factor, with factors significant at the 95% level in bold.	122

LIST OF FIGURES

FIGURE LEGENDS	57
Figure 2. 1.	58
Figure 2. 2.	59
Figure 2. 3.	60
Figure 2. 4.	61
Figure 2. 5.	62
Figure 2. 6.	63
Figure 2. 7.	64
Figure 2. 8.	65
Figure 2. 9.	66
Figure 2. 10.	67
Figure 2. 11.	68
Figure 2. 12.	69
FIGURE LEGENDS	91
Figure 3. 1.	92
Figure 3. 2.	93
Figure 3. 3.	94
Figure 3. 4.	95
Figure 3. 5.	96
FIGURE LEGENDS	123
Figure 4. 1.	124
Figure 4. 2.	125
Figure 4. 3.	126
Figure 4. 4.	127
Figure 4. 4 Continued	128
Figure 4. 5.	129
Figure 4. 6.	130
Figure 4. 7.	131
Figure 4. 8.	132
Figure 4. 9.	133

CHAPTER 1: LITERATURE REVIEW

Knowledge of the presence or absence of wildlife species and their distribution across a landscape is critical for making wildlife management decisions. However, direct inventories and surveys of wildlife are expensive and time-consuming (Mack et al. 1997) and are therefore impractical to conduct over extensive landscapes (Cardillo et al. 1999). In addition, landscapes vary temporally and are manipulated by management actions, which can alter wildlife species composition or abundance across a landscape. Because of the expense and impracticality of sampling across an entire landscape, wildlife-habitat relationship models have been frequently used in wildlife management since the 1970's (Anderson and Gutzwiller 1984). Habitat models have been developed to predict the presence or absence, geographical distribution, or population size of wildlife species by associating characteristics of the species to features of their environments (Morrison et al. 1992). The variety of the types of habitat models reflects the wide range of goals for which habitat models can be useful.

Goals of Habitat Modeling

Habitat models are useful tools for a variety of different wildlife management objectives. Distributions of wildlife species can be predicted for geographical areas that have not been extensively surveyed. Predictions of areas of high species diversity (Butterfield et al. 1994) or locations of species of concern (Sperduto and Congalton 1996) can be used to identify geographic locations for more intensive study. Habitat models are also useful for predicting areas of suitable habitat that may not be currently used by wildlife species (Lawton and Woodroffe 1991), serving as an aid to species re-introduction or prediction of the spread of an introduced species (Fielding and Haworth 1995). A primary interest of managers and planners is the prediction of the impacts of habitat manipulations and management decisions on wildlife species (Anderson and Gutzwiller 1994, Austin et al. 1996). In many cases, conservation and management decisions must be made in a relatively short period of time with limited information (Palma et al. 1999). Habitat modeling techniques that can create large-scale predictions of habitat suitability for wildlife species, without detailed knowledge of their physiology and behavior, can be a cost-effective and productive endeavor.

Habitat Modeling Philosophy

Although the goals of wildlife-habitat models may differ, all models are derived under one of two philosophies (Tobalske and Tobalske 1999). Deductive habitat models are created using existing knowledge of species habitat preferences, physiology, and/or behavior (Guisan and Zimmerman 2000). From these known relationships, a model can be constructed to deduce the locations across a landscape where suitable sites for the species will occur. The second, more commonly observed, modeling philosophy involves induction and is often used to define and quantify our understanding of wildlife-habitat relationships (Tobalske and Tobalske 1999). Inductive habitat models are developed from known locations of wildlife species and inferences about quality habitat are derived from habitat measures surrounding these locations.

Following one of the two philosophies, habitat models have been created for either individual species or for guilds, consisting of groups of species with similar habitat requirements. Guild or community habitat models were developed to meet mandates of maintaining sustainable populations of all species, where it would be impractical to individually model all species present (Verner 1984). Single species models have been more commonly used for particular species of interest, such as game species and threatened or endangered species. The state variable and output of habitat models differs with the goals and objectives for each model. However, the relationship between the state variable and output of the habitat model is generally static and deterministic (Guisan and Zimmerman 2000). Habitat models have been developed to predict species richness (Jorgensen and Demarais 1999), presence or absence of a species (Butterfield et al. 1994), probability of a species occurrence (Austin et al. 1996, Carroll et al. 1999), or an index of habitat suitability for a species (Hepinstall et al. 1996).

Assumptions of Habitat Modeling

Regardless of the state variable(s) predicted, wildlife-habitat models rely on a number of usually implicit assumptions. The success of habitat models is ultimately related to the existence of strong and predictable associations between species and habitat variables (Cardillo et al. 1999). In reality however, habitat is not the only factor that determines the distribution of species. Interspecific interactions, such as predation and competition, may have a significant effect on the

distribution and abundance of certain species (Morin 1981). Habitat models are based on key ecological concepts, such as the niche (Capen et al. 1986) and carrying capacity (Schamberger and O'Neil 1986). Habitat models based on empirical observations predict the realized (or ecological) niche because of the inclusion of biotic interactions and competitive inclusion in the observed data. On the other hand, the broader fundamental niche, which may be of more interest when predicting the species' response to habitat alteration, is a function of physiology and environmental constraints (Guisan and Zimmerman 2000). Habitat models also implicitly involve the concept of carrying capacity by assuming equilibrium between the species and their pattern of occurrence in the habitat.

Habitat models are simplified representations of complex ecological processes and cannot include every factor that influences a species occurrence or abundance (Reichert and Omlin 1997). Habitat models make the assumption that particular habitat variables can explain significant patterns and variability in species occurrence or abundance. In areas with rapidly changing environmental conditions, caused by severe disturbance, human influence, or successional dynamics, static habitat models may not be well suited for prediction (Guisan and Zimmerman 2000). When locations with particular habitat characteristics are associated with the occurrence of a species, it is often assumed that the locations are quality habitat. The probability of occupancy is thus used as a surrogate for habitat quality, when in fact the occupied habitat may not be "good" habitat (Tyre et al. 2001, VanHorne 1983). In landscapes where species exhibit source-sink population dynamics (Pulliam 1998), species occur both in "good" or source habitat (where birth rates exceed death rates) and in sink habitat (where death rates exceed birth rates). Demographic stochasticity in fecundity or survival rates and local dispersal also limit the ability of habitat models to predict occurrence as a function of habitat quality (Tyre et al. 2001). The choice of spatial scale (Wiens 1989), environmental predictor variables (Guisan and Zimmerman 2000), and the representation of habitat relationships within a model also reflect assumptions by the modeler. The choice of modeling method from simple indices to multivariate analyses reflects an assumption by the model-builder and can impact the resulting model accuracy and generality. Regardless of the conceptual assumptions, the cost efficiency and rapid development of habitat models still render them useful for large-scale prediction.

Each habitat model also has assumptions relevant to the particular data and methods used to build the model. The quality of the empirical data used to construct the model will affect the predictive accuracy of the model and should ideally result from a systematic, stratified, or random sampling design that adequately represents the study area and range of habitat variables. Observation of the species during sampling is not without error. Even if a species is present at a location, it may not be detected in the survey. If the species is difficult to detect, the population is small, or the survey methodology is insufficient or inappropriate, the species may not be detected by the observer when it does occur (Tyre et al. 2001). Failing to detect a species when it does occur leads to false absences in the survey data and unreliable distribution information. Error in distribution data can also occur from mistakes in species identification and data recording.

Methods of Habitat Modeling

Many methods have been used to model habitat suitability (see review by Guisan and Zimmerman 2000), but these methods can be classified into two groups: those requiring presence-absence data and those requiring presence-only data. Most of the commonly used statistical modeling techniques are based on multiple regression methods and require binary data (presence/absence data in this case) for model construction. Generalized linear modeling (GLM; McCullagh and Nelder 1989), discriminant function analysis (Davis 1986), generalized additive modeling (GAM; Hastie and Tibshirani 1986), artificial neural networks, and classification and regression trees (CART), are examples of methods that have been used to model presence-absence data. More recently, methods of habitat modeling that use presence-only data have been developed, such as ecological niche factor analysis (ENFA; Hirzel et al. 2002) and genetic algorithm for rule set prediction (GARP, Stockwell and Noble 1992). Presence-only data are more difficult to model statistically and have several drawbacks for habitat modeling. Other than the lack of absence data, there is unknown sampling bias when using ad-hoc or opportunistic observations. Often the locations of the samples are correlated with other factors, such as accessibility or habitat type (Zaniewski et al. 2002). In addition, the sampling bias for different species is unknown. For example, herbarium data sets often have a large number of occurrences of rare species in comparison to more common species (Zaniewski et al. 2002). Although the

performance of presence-only data models has been less accurate than models constructed from presence-absence data, the presence-only methods are robust to the quality and quantity of data and often presence-only data are the only available data because of the expense and difficulty of conducting field surveys.

Logistic Regression

Logistic regression is a statistical tool for the analysis of binary data, such as presence and absence data. A specific example of a general linear model (GLM), a logistic regression model relates a linear combination of the environmental variables to the predicted variable by means of a logistic link function which constrains the predicted variable to a probability from 0 to 1 (Guisan and Zimmerman 2000). Using presence/absence data, logistic regression has been used to model probability of occurrence for a multitude of different species including: fisher (*Martes pennanti*) (Carroll et al. 1999), Iberian lynx (*Lynx pardinus*) (Palma et al. 1999), buzzard (Austin et al. 1996), American woodcock (*Scolopax minor*) (Storm et al. 1995), great bustards (*Otis tarda*) (Osborne et al. 2001), Siberian flying squirrels (*Pteromys volans*) (Reunanen 2002), groups of woodpecker species (Tobalske and Tobalske 1999) and Himalayan river birds (Manel et al. 1999b). In comparisons between logistic regression and discriminant function analysis models, logistic regression has been found to have slightly higher classification accuracy (Storm et al. 1995; Manel et al. 1999b) or similar accuracy depending on the method of data partitioning.

Discriminant Function Analysis

Linear discriminant function analysis is a multivariate classification technique that can be used to describe how two or more groups differ in relation to selected habitat characteristics and classify additional observations into one of the groups. Discriminant function analysis is the second most common multivariate technique used in fisheries research (Paukert and Wittig 2002) and is commonly used in wildlife studies (Johnson 1981). To test the significance of the separation of the groups, discriminant analysis relies on several assumptions, including: a random, independent sample, the probability of an unknown point belonging to either group is equal, equality of the variance-covariance matrix, multivariate normality, and that all of the data used to

create the discriminant function were classified correctly (Davis 1986). However, provided that sample sizes are large, discriminant function analysis is robust to violations of these assumptions. It is recommended that the number of observations used to create the discriminant function is at least three times the number of environmental variables included in the model (Williams and Titus 1988).

Artificial Neural Networks

More recently, artificial neural networks have been applied to problems in ecological fields since the early 1990's (Lek and Guegan 1999). Artificial neural networks are patterned on the function of the human brain (Manel et al. 1999a) and represent a black box approach to ecological modeling. Although the underlying ecological relationships may be unknown, an artificial neural network can learn the patterns of correlation between input data and resulting output values. By using training data, the artificial neural network can learn complex and non-linear relationships and then can predict the output for a second, independent dataset (Lek and Guegan 1999). Neural networks have been found to perform with better accuracy than logistic regression when the habitat selection process is non-linear and incapable of being sufficiently explained using generalized linear models (Ozesmi and Ozesmi 1999). However, another model comparison found that logistic regression always outperformed artificial neural networks while more efficiently using computer time (Manel et al. 1999b). Artificial neural networks are appropriate when correlation is present (Zobel et al. 2004) and provides advantages over more standard multivariate statistical analyses with assumptions of independent variables.

Classification And Regression Trees

Classification and regression trees are multivariate techniques that result in a tree structure that branches at nodes based on the value of environmental predictors to describe the observations. Classification trees are built with presence-absence data and the splits in the tree are based on the proportion of presences and absences in the groups (De'ath and Fabricius 2000). This approach has been used to predict the presence and absence of nesting habitat for cavity-nesting birds based on landscape patterns (Lawler and Edwards 2002). Regression trees are similar, but the splits in the tree are based on minimizing the sum of squares within each group (De'ath and

Fabricius 2000). Regression trees have been used to relate relative density of threatened gopher tortoises (*Gopherus agassizii*) to habitat characteristics (Andersen et al. 2000). To summarize a regression tree, the proportion of the total sum of squares explained by the tree is useful for validation, while the misclassification rate of the number of observations misclassified by the tree is used for classification trees (De'ath and Fabricius 2000).

Ecological Niche Factor Analysis

Ecological niche factor analysis (ENFA) is a relatively new multivariate approach developed to predict habitat suitability when absence data for the species are not available (Hirzel et al. 2002). Other presence-only modeling methods, such as GARP (Stockwell and Peters 1999), were not specifically reviewed here, given recent focus on ENFA methods and comparisons with other multivariate techniques. ENFA compares the distribution of the presence observations in the multidimensional space of the environmental variables to the entire study area (Guisan and Zimmerman 2000). The suitability is based on functions that define the marginality of the species, i.e. how the species mean differs from the mean of the entire area, and the specialization of the species, i.e. ratio of the overall variance to the species variance. The factor analysis extracts linear combinations of the environmental variables against which the species shows the greatest degree of marginality and specialization (Hirzel et al. 2002). The key advantage to ENFA is that absence data, which is often not available or difficult to obtain, is not necessary for the analysis. In a simulation study, ENFA was robust to the quality and quantity of data for species with differing distributions: spreading, equilibrium, and overabundant. In comparison with logistic regression, ENFA was superior in predicting suitability for the species with a spreading distribution, equal for the equilibrium distribution, and slightly outperformed when the species was abundant (Hirzel 2001).

Overall, numerous methods have been used to model the relationships among species occurrences and environmental variables. Although some comparisons have been conducted between the methods (Hirzel 2001, Manel 1999b), there is an overall lack of comparative studies that could be useful when choosing to begin a modeling study (Guisan and Zimmerman 2000). Little guidance, such as the comparison of generalized additive models (GAM) and ENFA by

Zaniewski et al. (2002), is available for deciding when to use a presence-only data method versus a presence/absence data method. Presence-only data methods can be used when absence data does not exist or when the absence data are unreliable. However, absence data can also be simulated in different ways, including random selection from the entire landscape, allowing the use of presence/absence data methods (Zaniewski et al. 2002) with presence-only data sets. The choice of modeling methods is further complicated by a choice between frequentist versus Bayesian methods. Although all of the previously discussed multivariate methods are based on frequentist statistics, Bayesian statistical theory is also well suited for problems of probabilistic modeling.

Bayesian Statistical Methods

Bayesian statistical theory represents a philosophical departure from the commonly used frequentist statistics in the wildlife field. While frequentist statistical theory treats an unknown model parameter as a fixed value, Bayesian statistical theory treats the parameter as a random variable where the uncertainty in the parameter value is represented by a probability distribution (Reichert and Omlin 1997). Bayesian statistics also allow for the quantification and inclusion of prior knowledge into the statistical procedure. Model parameters are estimated by using Bayes theorem that combines the prior probability with a likelihood function to result in a posterior probability distribution of the model parameters. The posterior probability represents a compromise between the prior knowledge and the new observed data with each weighted based on their variance (Gelman et al. 1995). Overall, Bayesian statistical inference incorporates prior knowledge and data to produce estimates of parameters and their uncertainty in an easily understandable framework (Ellison 1996).

Bayesian statistics have several advantages that would promote their use in the wildlife field. A useful feature of Bayesian statistics is the ability to easily update models based on new information. Because the posterior probability is a distribution, the posterior probability can then become the prior probability when new information is acquired and a new posterior probability distribution can be produced. In addition, Bayesian statistics have advantages for the analysis of non-replicated landscape studies because frequentist statistics are based on a long-range

frequency view of probability (Reckhow 1990). However, Bayesian statistics allow a more meaningful and straightforward interpretation of probabilistic inferences when a large number of repeated samples are not available. For hypothesis testing in a Bayesian framework, the tests are only conditioned on the existing data sample, unlike frequentist methods where the test is also conditioned on data that were not observed (i.e. data more extreme than the observed values) (Reckhow 1990).

Despite the advantages of Bayesian statistics, their use in the wildlife field has been minimal for several reasons. The specification of prior information and its incorporation into the data analysis can be a subjective step in an otherwise objective scientific process. The choice of a prior distribution is a contentious issue and researchers can choose from a range of objective reference or non-informative prior distributions to subjective distributions of personal beliefs (Dennis 1996). In addition, to obtain exact probabilities using Bayes theorem, a normalization constant (the marginal distribution) must be calculated, which can be an unwieldy integral function. In many cases, analytical solutions are not available and posterior distributions must be estimated via sampling algorithms. Until recently, computer technology was insufficient for the quantity and complexity of the necessary computations (Carlin and Louis 1998). User-friendly software packages, such as BUGS (Bayesian inference using Gibbs sampling, developed by the Medical Research Council Biostatistics Unit, Institute of Public Health, Cambridge, England), are increasing in availability and use (Meyer and Millar 1999).

Applications of Bayesian Statistics in the Wildlife Field

Although not as commonly used as frequentist statistics, Bayesian statistics have been used for a variety of analyses. In both wildlife and entomology studies, Bayesian statistics have been used to combine remote sensing data with other diverse types of data such as topography (Hutchinson 1982), distance to environmental features (Wood et al. 1992), and bird survey data (Aspinall and Veitch 1993, Hepinstall and Sader 1997, Tucker et al. 1997). Bayesian statistics have been used to estimate population parameters, such as survival and movement probabilities, based on capture-recapture data (Dupuis 1995, Ver Hoef 1996, Chavez-Demoulin 1999). Methods for updating Bayesian models were developed for analyzing recurring waterfowl surveys, leading to

improved precision of population size estimators (Johnson 1989). For estimating parameters for groups of species, hierarchical Bayesian models have been used for analysis of population trends of groups of bird species observed in Breeding Bird Survey data (Sauer and Link 2002). Bayesian statistics have also aided in the analysis of presence-absence data by inferring estimates for the probability of occurrence of unobserved species (Nicholson and Barry 1995) and estimates of the number of species in a community (Lewins and Joanes 1984, Solow 1994). Program MARK (White 2004) has incorporated occupancy models based on McKenzie et al. (2002) that predict the proportion of sites occupied and a Bayesian analysis of the probability of detection at a site (MacKenzie et al. 2002). Pereira and Itami (1991) used a combination of multivariate and Bayesian methods to predict the distribution of the Mt. Graham red squirrel. Trend surface modeling created a surface of the location of squirrel habitat use and the surface was subsequently used as the prior probability in the Bayesian model (Pereira and Itami 1991). This model and the advances in Bayesian modeling for avian species have shown the promise and utility of Bayesian statistics for habitat modeling, although the implementation of these methods has been minimal.

Geographic Information Systems and Habitat Modeling

A common use for probability of occurrence models is to predict the likely distribution of the species across a heterogeneous landscape (Austin et al. 1996, Tucker et al. 1997, Carroll et al. 1999, Tobalske and Tobalske 1999, Boone and Krohn 2000). Although the modeling methods previously discussed are not explicitly spatial, the models can be applied to an entire landscape to result in a spatially explicit model by treating the landscape as a collection of smaller areas that each has an associated probability of occurrence (Boroski et al. 1996, Hepinstall et al. 1996, van Horssen et al. 1999). To accomplish this task, geographic information systems (GIS) are frequently used in concert with habitat modeling, both as a source of environmental predictor variables and as a method of displaying model results.

A GIS is essentially an integration of a computerized mapping system with a digital database. A GIS is used for the input, storage, analysis, manipulation, and display of geographic and associated attribute data (Coulson 1992). With the advances in computer technology and

increasing interest in spatial relationships and landscape ecology, GIS technology has become increasingly used in wildlife management and research. Environmental data stored in a GIS, such as land cover and topography, are frequently used to derive habitat variables for modeling (Burrough and McDonnell 1998). GIS also has the capability to easily store, reference, and relate the spatial locations and attributes of species survey data with other environmental data (Burrough and McDonnell 1998). Analytical GIS techniques for manipulating spatial data are increasingly used to measure landscape metrics, and the interspersion and juxtaposition of resources for habitat analyses (Donovan et al. 1987). The obvious utility of GIS has resulted in both the creation of spatially-explicit habitat models and the modification of non-spatial habitat models for implementation within a GIS (Hepinstall et al. 1996).

Factors Influencing Probability of Occurrence Prediction

With concentration on GIS technology and implementation, it can be easy to overlook a variety of input data and biological factors that can influence the ability of a model to accurately predict the probability of occurrence of a wildlife species. The quality of the presence-absence data used to define the species-habitat relationships directly impacts the prediction accuracy of the resulting model. The sample size of the observation data, i.e. the number of occurrences, affects prediction accuracy (Stockwell and Peterson 2002, Pearce and Ferrier 2000a). Large numbers of occurrences may be necessary to maximize model accuracy (Pearce and Ferrier 2000a), although the exact number of observations may depend on the scale of the model (Karl et al. 2000). The sampling design for collecting the presence-absence data is an important consideration, although currently, too many habitat models are created using data from observational studies lacking a sampling design (Conroy and Noon 1995, Guisan and Zimmerman 2000). To maximize prediction accuracy, an appropriate sampling design that covers the range of habitats and accounts for species-specific factors, such as home range size and degree of spatial aggregation (Boone and Krohn 1999), is necessary.

Although the quantity and quality of occurrence data is important, often a limiting factor in habitat modeling is the accuracy of the absence data (Hirzel et al. 2002). The accuracy of absence data is difficult to assess because the failure to observe a species at a location may be

due to several reasons: the failure of the survey to detect the species when it was present (McArdle 1990, Solow 1994), historical reasons cause the species to be absent even in suitable habitat (Hirzel et al. 2002), or the true absence of the species because of unsuitable habitat at that location. False absences in presence/absence data occur when the species is actually present at a site, but the species is not detected in the survey. Many factors including sampling methods, environmental conditions, species-specific characteristics, and population density of the species can influence the detectability of a species during a survey. Species-specific characteristics related to the appearance and behavior of the species, as well as the habitats used by the species, can influence the species detection probability (Bayley and Peterson 2001). Selecting sampling methods with high probabilities of detection and ensuring adequate sampling effort (Angermeier and Smogor 1995), in terms of both the size of the sampling area and number of repeat samples, are important for obtaining accurate absence data. Advances in methods to estimate probability of presence, given a lack of detection at a location, have been made using Bayesian models in fisheries research (Bayley and Peterson 2001), although no evidence of similar models could be found in the wildlife field.

In addition to species distribution data uncertainty, factors involved in the construction of habitat models also influence the accuracy of probability of occurrence. The spatial scale or resolution of the model can determine the patterns of association that are detected between the habitat variables and species presence/absence (see Wiens 1989). The effect of spatial scale on the accuracy of the model is influenced by environmental variability, for example, a finer spatial scale can improve accuracy in a heterogeneous environment (Karl et al. 2000). The choice of spatial scale also affects the influence of spatial autocorrelation on the model. Spatial autocorrelation reduces the independence of the observations used in the creation of the model, possibly leading to overestimating the degrees of freedom in statistical tests of model parameters (Guisan and Zimmerman 2000). When known to be a factor, spatial autocorrelation can be included in the modeling process through an autocorrelative model (Augustin et al. 1996). The generalization of habitat patterns and species predictions of occurrence into summary statistics leads to the loss of information (Conroy and Noon 1996).

The choice of which habitat factors to include in the model, which algorithm to use in the selection of habitat factors (i.e. stepwise procedures) (Pearce and Ferrier 2000b), and the level of complexity of the model also contribute to uncertainty in the model output. Biological meaning behind the selection of habitat variables when using step-wise selection procedures may be difficult to interpret or non-existent. Because the statistical techniques are correlational, statistical significance may not reflect causal relationships and true biological significance (Noon 1986). A trade-off exists between the complexity of the model and the quality of the model fit to the data (Reichert and Omlin 1997). A comparison of habitat data sources for coarse resolution prediction of distributional patterns showed little difference in prediction accuracy between vegetation, climate, and spatial autocorrelation models (Beard et al. 1999). When more than one of the habitat variables was included, the models only slightly increased in accuracy. Error and uncertainty in the habitat variables used as predictors propagates through the modeling process and influences the resulting model output. Several sources of uncertainty have been identified when dealing with GIS-derived habitat variables: the loss of information from spatial generalization, loss of information due to the precision of the selected classification scheme, positional errors and misclassification, and selecting an unrepresentative area of the entire landscape (Stoms et al. 1992). In addition, habitat models created using statistical methods may lack generality. Habitat models created in one area can underestimate prediction uncertainty when extrapolated to other areas (Reichert and Omlin 1997). Habitat models created for the same species in different locations may consist of entirely different sets of habitat variables (Manel et al. 1999a). Although some sources of error and uncertainty can be controlled during the modeling process, it is necessary to understand the sources and extent of other uncertainty to derive realistic estimates for the model output.

Overall, habitat modeling attempts have been most successful for species that are common, range-restricted, and specialist species in comparison to species that are rare, wide-ranging, or generalists (Lancia et al. 1986). Differences in the ability to accurately model particular species can be related to factors such as: extent of geographic distribution (Stockwell and Peterson 2002), niche width (number of habitats used by the species), home range size, detectability of the species, degree of spatial aggregation, degree of density dependence (Boone and Krohn 1999)

and colonization history (Hirzel et al. 2001). The relationships between these factors and the accuracy of different types of habitat models are often not known.

Model Testing

Because of the numerous sources of uncertainty in the habitat modeling process, it is vitally important to test the output predictions of the model. Model testing provides information about the performance and reliability of the model, as well as providing additional data for model improvement. For example, lack of validation is a key argument of critics of Habitat Suitability Index models and can impede the development and implementation of these models (Cole and Smith 1983). Model testing can be conducted on many levels and in a variety of ways. The most common level of model testing is the validation of the final output of the model, but testing the assumptions, variables, and components of the model are valuable levels of validation as well (Schamberger and O'Neil 1986).

Several methods and data sources for model validation are frequently used. Ideally, an independent data set should be used to test model performance, but often this type of data set does not exist. In this case, the original data for model creation can be separated into a training set and testing set (Guisan and Zimmerman 2000). Jackknife sampling is a special case of data separation where one case is sequentially held out with the remaining observations forming the training set (Capen et al. 1986). A general rule was proposed by Huberty (1994) for selecting the ratio of the size of the training data set to the testing data set that is dependent on the number of predictor variables in the model. Finally, resubstitution of the data used for model building into the model as a test data set is a test of the model, but tends to produce optimistic measures of accuracy (Fielding and Bell 1997, Olden et al. 2002).

Several options exist to quantify the accuracy of probability of occurrence models. Two components are important in the assessment of the model: sensitivity, i.e. agreement between the predicted probability of occurrence and observed presence, and specificity, i.e. the ability of the model to distinguish absence locations (Fielding and Bell 1997). Models with high sensitivity are good at predicting species presence, while models with high specificity are good at predicting

absence. Probability of occurrence models are generally measured by setting a threshold probability value and transforming the probabilities into binary presence/absence data. The predicted presence/absence can then be compared with the test data in a 2 x 2 table. From this table, also known as a confusion matrix, measures such as percent correctly classified, omission and commission errors (Fielding and Bell 1997), and kappa (Cohen 1960) can be obtained. Kappa uses all of the information in the confusion matrix and assesses the improvement in classification over chance (Fielding and Bell 1997). The choice of a threshold probability to transform the predictions to presence/absence can be arbitrary (Pearce and Ferrier 2000a) and statistics based on the confusion matrix are influenced by the frequency of occurrence of the species (Manel et al. 2001). The receiver operating characteristic (ROC) plot methodology was developed to provide a threshold independent measure of accuracy and results in a plot of the relative proportions of correctly classified sites over a range of threshold levels. The area under the ROC plot can then be compared among competing models (Cumming 2000). The ROC method of model evaluation is becoming more widespread, but Manel et al. (2001) found the method to be highly correlated with the Cohen's kappa statistic, which is simpler to compute. The choice of statistic to compare model accuracy is important, but identifying the sources of the errors is the ultimate goal.

Errors in the prediction of species' occurrence can be viewed as falling into two categories: algorithmic errors, which are derived from model construction and data uncertainties, and biotic errors (Fielding and Bell 1997). Biotic errors are inevitable because the model is a simplification of reality and cannot include all of the habitat factors that influence the species distribution. Misclassification of locations by the model can be related to several aspects of the species' ecology. Most presence/absence habitat models assume that the suitable area of the landscape is saturated by the species of interest (Capen et al. 1986); however, if this assumption is not met the model will result in a large percentage of commission errors, where the species is predicted to occur but is not found at the location. In addition, intraspecific and interspecific interactions can influence the performance of the model. Other problems related to judging the accuracy of habitat models are spatial autocorrelation of species occurrences and the influence of historical events on the species' distribution (Fielding and Bell 1997). Understanding both the biotic and

model construction reasons for errors in habitat model performance is a necessary step towards identifying aspects of the model that can be improved.

Model Testing Using Data Simulation

When testing different modeling methods, it is beneficial to have the species distribution and environmental data known with certainty. Simulation can produce a data set that can be fully controlled and known with certainty by the researcher (Hirzel et al. 2001). Simulation can be used for many of the steps in the modeling process. Environmental variables or species' distributions can be simulated across a landscape to have complete control over the input data used for model creation. Hirzel et al. (2001) created a "virtual species" by simulating the ecological niche of the species and applied it to a real landscape, resulting in a known habitat suitability map for the virtual species. In another case (Tyre et al. 2001), a simulated landscape was generated using a fractal algorithm. Tyre et al. (2001) also introduced the concept of a "virtual ecologist," which simulates an ecologist conducting a survey by sampling within a known landscape. Multiple samples can be simulated from the same landscape and modeling methods can be tested by how closely they predict the "truth" data set. Accuracy assessment of the models is simple and certain because the "truth" landscape and species distribution is perfectly known (Hirzel et al. 2001).

Habitat Modeling for Amphibians and Reptiles

Although habitat models have been widely used in the wildlife field, there is a relative lack of predictive habitat models for amphibian and reptile species. Even though recent concern over declines of amphibian populations (Fisher and Shaffer 1996, Halliday 1998) has led to increased research and policy (*ex.* the Declining Amphibian Task Force created by the International Union for the Conservation of Nature and Natural Resources and the U.S. Department of Interior's Taskforce on Amphibian Declines and Deformities) interest (Beebee 1996, Baker 1999), few habitat suitability models for these species exist (although see Guerry and Hunter 2002, Lowe and Bolger 2002). Amphibian and reptile taxa present a challenge to habitat modelers because of their wide range of aquatic and terrestrial life histories (Conant and Collins 1998) and relatively few presence-absence data sets based on long-term studies. Although standard

techniques exist for the inventory and monitoring of amphibian (Heyer et al. 1994) and reptile species, detectability varies among these species because of the wide range of life histories and is often unknown. Many species tend to be secretive and spend time hidden underground or underneath objects (Conant and Collins 1998). Given the recent emphasis to develop monitoring programs and management plans for amphibian and reptile species, inventory (presence-only) data sets are becoming increasingly common. The ability to develop accurate habitat suitability models for a range of amphibian and reptile species from these inventory data sets would be useful for researchers and managers.

LITERATURE CITED

- Andersen, M. C., J. M. Watts, J. E. Freilich, S. R. Yool, G. I. Wakefield, J. F. McCauley, and P. B. Fahnestock. 2000. Regression-tree modeling of desert tortoise habitat in the central Mojave desert. *Ecological Applications* 10:890-897.
- Anderson, S. H., and K. J. Gutzwiller. 1994. Habitat evaluation methods. Pages 592-606 in T.A. Bookhout, editor. *Research and management techniques for wildlife and habitats*. Fifth edition, The Wildlife Society, Bethesda, Maryland.
- Angermeier, P. L. and R. A. Smogor. 1995. Estimating number of species and relative abundances in stream-fish communities: effects of sampling effort and discontinuous spatial distributions. *Canadian Journal of Fisheries and Aquatic Sciences* 52:936-949.
- Aspinall, R. and N. Veitch. 1993. Habitat Mapping From Satellite Imagery and Wildlife Survey Data Using a Bayesian Modeling Procedure in a Gis. *Photogrammetric Engineering and Remote Sensing* 59:537-543.
- Augustin, N. H., M. A. Muggleston, S. T. Buckland. 1996. An autologistic model for the spatial distribution of wildlife. *Journal of Applied Ecology* 33:339-347.
- Austin, G. E., C. J. Thomas, D. C. Houston, and D. B. A. Thompson. 1996. Predicting the spatial distribution of buzzard *Buteo Buteo* nesting areas using a geographical information system and remote sensing. *Journal of Applied Ecology* 33:1541-1550.
- Baker, B. 1999. New federal task force tackles amphibian troubles. *Bioscience*, 49:366.
- Bayley, P. B. and J. T. Peterson. 2001. An approach to estimate probability of presence and richness of fish species. *Transactions of the American Fisheries Society* 130:620-633.
- Beard, K. H., N. Hengartner, and D. K. Skelly. 1999. Effectiveness of predicting breeding bird distributions using probabilistic models. *Conservation Biology* 13:1108-1116.
- Beebee, T. J. C. 1996. *Ecology and conservation of amphibians*. Chapman and Hall, London.

- Bender, L. C., G. J. Roloff, and J. B. Haufler. 1996. Evaluating confidence intervals for habitat suitability models. *Wildlife Society Bulletin* 24:347-352.
- Boone, R. B. and W. B. Krohn. 1999. Modeling the occurrence of bird species: are the errors predictable? *Ecological Applications* 9:835-848.
- Boone, R. B. and W. B. Krohn. 2000. Predicting broad-scale occurrences of vertebrates in patchy landscapes. *Landscape Ecology* 15:63-74.
- Boroski, B. B., R. H. Barrett, I. C. Timossi, and J. G. Kie. 1996. Modelling habitat suitability for black-tailed deer (*Odocoileus hemionus columbianus*) in heterogenous landscapes. *Forest Ecology and Management* 88:157-165.
- Breining, D. R., V. L. Larson, B. W. Duncan, and R. B. Smith. 1998. Linking habitat suitability to demographic success in Florida scrub-jays. *Wildlife Society Bulletin* 26:118-128.
- Brooks, R. P. 1997. Improving habitat suitability index models. *Wildlife Society Bulletin* 25: 163-167.
- Burgman, M. A., D. R. Breining, B. W. Duncan, and S. Ferson. 2001. Setting reliability bounds on habitat suitability indices. *Ecological Applications* 11:70-78.
- Burrough, P. A. and R. A. McDonnell, R.A. 1998. *Principles of Geographical Information Systems*. Oxford University Press, Oxford.
- Butterfield, B. R., B. Csuti, J. M. and Scott. 1994. Modeling vertebrate distributions for Gap analysis. Pages 53-68 in R. I. Miller, editor. *Mapping the diversity of nature*. Chapman and Hall, London.
- Capen, D. E., J. W. Fenwick, D. B. Inkley, and A. C. Boynton. 1986. Multivariate models of songbird habitat in New England forests. Pages 171-175 in J. Verner, M. L. Morrison, and C. J. Ralph, editors. *Wildlife 2000: Modeling habitat relationships of terrestrial vertebrates*. University of Wisconsin Press, Madison, Wisconsin, USA.
- Cardillo, M., D. W. Macdonald, and S. P. Rushton. 1999. Predicting mammal species richness and distributions: testing the effectiveness of satellite-derived land cover data. *Landscape Ecology* 14:423-435.
- Carlin, B. P. and T. A. Louis. 1998. *Bayes and empirical Bayes methods for data analysis*. Chapman and Hall/CRC, Boca Raton, Florida.
- Carroll, C., W. J. Zielinski, and R. F. Noss. 1999. Using presence-absence data to build and test spatial habitat models for the fisher in the Klamath Region, USA. *Conservation Biology* 13:1344-1359.
- Chavez-Demoulin, V. 1999. Bayesian inference for small-sample capture-recapture data. *Biometrics* 55:727-731.

- Cohen, J. 1960. A coefficient of agreement for nominal scales. *Education and Psychological Measurements* 41:687-699.
- Cole, C. A. and R. L. Smith. 1983. Habitat suitability indices for monitoring wildlife populations. *Transactions of the North American Wildlife and Natural Resources Conference* 48:367-375.
- Conant, R. and J. T. Collins. 1998. *A field guide to reptiles and amphibians of eastern and central North America*. Third edition, expanded. Houghton Mifflin Company, Boston, Massachusetts, USA. 616 pp.
- Conroy, M. J. and B. R. Noon. 1996. Mapping of species richness for conservation of biological diversity: conceptual and methodological issues. *Ecological Applications* 6: 763-773.
- Coulson, R. N. 1992. Intelligent geographic information systems and integrated pest management. *Crop Protection* 11:507-516.
- Cumming, G. S. 2000. Using between-model comparisons to fine-tune linear models of species ranges. *Journal of Biogeography* 27:441-455.
- Davis, F. W., D. M. Stoms, J. E. Estes, J. Scepan, and J. M. Scott. 1990. An information system approach to the preservation of biological diversity. *International Journal of Geographical Information Systems* 4:55-78.
- Davis, J. C. 1986. *Statistics and data analysis in geology*. Second edition. John Wiley and Sons, New York, New York, USA.
- De'ath, G. and K. E. Fabricius. 2000. Classification and regression trees: a powerful yet simple technique for ecological data analysis. *Ecology* 81:3178-3192.
- Dennis, B. 1996. Discussion: Should ecologists become Bayesians? *Ecological Applications* 6: 1095-1103.
- Donovan, M. L., D. L. Rabe, and C. E. Olson, Jr. 1987. Use of geographic information systems to develop habitat suitability models. *Wildlife Society Bulletin* 15:574-579.
- Dupuis, J. A. 1995. Bayesian estimation of movement and survival probabilities from capture-recapture data. *Biometrika* 82:761-772.
- Ellison, A. M. 1996. An introduction to Bayesian inference for ecological research and environmental decision-making. *Ecological Applications* 6:1036-1046.
- Fielding, A. H. and J. F. Bell. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation* 24:38-49.
- Fielding, A. H. and P. F. Haworth. 1995. Testing the generality of bird-habitat models. *Conservation Biology* 9:1466-1481.

- Fisher, R. N. and H. B. Shaffer. 1996. The decline of amphibians in California's Great Central Valley. *Conservation Biology* 10:1387-1397.
- Gaudette, M. T. 1986. Modeling winter habitat for white-tailed deer in southwestern Virginia. M.S. thesis, Virginia Polytechnic Institute and State University, Blacksburg, Virginia. 68 p.
- Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin. 1995. Bayesian data analysis. Chapman and Hall/CRC, Boca Raton, Florida, USA
- Grubb, T. G. 1988. Pattern recognition- a simple tool for evaluating wildlife habitat. USDA Forest Service Rocky Mountain Forest and Range Experiment Station, No. RM-487, Fort Collins, Colorado.
- Guerry, A. D. and M. L. Hunter, Jr. 2002. Amphibian distributions in a landscape of forests and agriculture: an examination of landscape composition and configuration. *Conservation Biology* 16:745-754.
- Guisan, A. and N. E. Zimmerman. 2000. Predictive habitat distribution models in ecology. *Ecological Modelling* 135:147-186.
- Halliday, T. 1998. A declining amphibian conundrum. *Nature* 394:418-419.
- Hastie, T. and R. Tibshirani. 1986. Generalized additive models. *Statistical Science* 3:297-318.
- Hecnar, S. J. and R. T. M'Closkey. 1997. Changes in the composition of a ranid frog community following bullfrog extinction. *American Midland Naturalist* 137:145-150.
- Hepinstall, J. A., L. P. Queen, and P. A. Jordan. 1996. Application of a modified habitat suitability index model for moose. *Photogrammetric Engineering and Remote Sensing* 62:1281-1286.
- Hepinstall, J. A. and S. A. Sader. 1997. Using Bayesian statistics, Thematic Mapper satellite imagery, and Breeding Bird Survey data to model bird species probability of occurrence in Maine. *Photogrammetric Engineering and Remote Sensing* 63:1231-1237.
- Heyer, W. R., Donnelly, M. A., McDiarmid, R. W., Hayek, L. C., and Foster, M. S., editors. 1994. Measuring and monitoring biological diversity: Standard methods for amphibians. Smithsonian Institution Press, Washington D.C., USA.
- Hirzel, A. H., J. Hausser, D. Chessel, and N. Perrin. 2002. Ecological-niche factor analysis: how to compute habitat-suitability maps without absence data? *Ecological Modelling* 83:2027-2036.
- Hirzel, A. H., V. Helfer, and F. Metral. 2001. Assessing habitat-suitability models with a virtual species. *Ecological Modelling* 145:111-121.
- Huberty, C. J. 1994. Applied discriminant analysis. Wiley, New York, New York, USA.

- Hutchinson, C. 1982. Techniques for combining Landsat and ancillary data for digital classification improvement. *Photogrammetric Engineering and Remote Sensing* 48:123-130.
- Johnson, D. H. 1989. An empirical Bayes approach to analyzing recurring animal surveys. *Ecology* 70:945-952.
- Johnson, D. H. 1981. The use and misuse of statistics in wildlife habitat studies. Pages 11-19 in D. E. Capen, ed. *The use of multivariate statistics in studies on wildlife habitat*. U.S. Forest Service General Technical Report RM-87, Burlington, Vermont, USA.
- Karl, J. W., P. J. Heglund, E. O. Garton, J. M. Scott, N. M. Wright, and R. L. Hutto. 2000. Sensitivity of species habitat-relationship model performance to factors of scale. *Ecological Applications* 10:1690-1705.
- Kurzejeski, E. W. and J. B. Lewis. 1985. Application of PATREC modeling to wild turkey management in Missouri in J. Earl and M. C. Kennamer, editors. *Proceedings of the Fifth National Wild Turkey Symposium*.
- Lancia, R. A., D. A. Adams, and B. W. Anderson. 1986. Temporal and spatial aspects of species-habitat models. Pages 65-69 in J. Verner, M. L. Morrison, and C. J. Ralph, editors. *Wildlife 2000: Modeling habitat relationships of terrestrial vertebrates*. University of Wisconsin Press, Madison, Wisconsin, USA.
- Lancia, R. A., S. D. Miller, D. A. Adams, and D. W. Hazel. 1982. Validating habitat quality assessment: An example. *Transactions of the North American Wildlife and Natural Resources Conference* 47:96-110.
- Lawler, J. J. and T. C. Edwards, Jr. 2002. Landscape patterns as habitat predictors: building and testing models for cavity-nesting birds in the Uinta Mountains of Utah, USA. *Landscape Ecology* 17:233-245.
- Lawton, J. H. and G. L. Woodroffe. 1991. Habitat and the distribution of water voles: why are there gaps in a species range? *Journal of Animal Ecology* 60:79-91.
- Lek, S. L. and J. F. Guegan. 1999. Artificial neural networks as a tool in ecological modeling, an introduction. *Ecological Modelling* 120:65-73.
- Lewins, W. A. and D. N. Joanes. 1984. Bayesian estimation of the number of species. *Biometrics* 40:323-328.
- Lowe, W. H. and D. T. Bolger. 2002. Local and landscape-scale predictors of salamander abundance in New Hampshire headwater streams. *Conservation Biology* 16:183-193.
- Mack, E. L., L. G. Firbank, P. E. Bellamy, S. A. Hinsley, and N. Veitch. 1997. The comparison of remotely sensed and ground-based habitat area data using species-area models. *Journal of Applied Ecology* 34:1222-1228.

- Manel, S., H. C. Williams, and S. J. Ormerod. 2001. Evaluating presence-absence models in ecology: the need to account for prevalence. *Journal of Applied Ecology* 38:921-931.
- Manel, S., J. M. Dias, S. T. Buckton, and S. J. Ormerod. 1999a. Alternative methods for predicting species distribution: an illustration with Himalayan river birds. *Journal of Applied Ecology* 36:734-747.
- Manel, S., J. M. Dias, and S. J. Ormerod. 1999b. Comparing discriminant analysis, neural networks and logistic regression for predicting species distributions: a case study with a Himalayan river bird. *Ecological Modelling* 120:337-347.
- McArdle, B. H. 1990. When are rare species not there? *Oikos* 57:276-277.
- McCullagh, P., and J. A. Nelder. 1989. *Generalized linear models*. Second ed., Chapman and Hall.
- MacKenzie, D. I., J. D. Nichols, G. B. Lachman, S. Droege, J. A. Royle, and C. A. Langtimm. 2002. Estimating site occupancy when detection probabilities are less than one. *Ecology* 83:2248-2255.
- Meyer, R. and R. B. Millar. 1999. BUGS in Bayesian stock assessments. *Canadian Journal of Fisheries and Aquatic Sciences* 56:1078-1086.
- Morin, P. J. 1981. Predatory salamanders reverse the outcome of competition among three species of anuran tadpoles. *Science* 212:1284-1286.
- Morrison, M. L., B. G. Marcot, and R. W. Mannan. 1992. *Wildlife-habitat relationships: Concepts and applications*. University of Wisconsin Press, Madison, Wisconsin, USA.
- Nicholson, M. and J. Barry. 1995. Inferences from spatial surveys about the presence of an unobserved species. *Oikos* 72:74-78.
- Noon, B.R. 1986. Summary: biometric approaches to modeling – the researcher’s viewpoint. Pages 197-201 in J. Verner, M. L. Morrison, and C. J. Ralph, editors. *Wildlife 2000: Modeling habitat relationships of terrestrial vertebrates*. University of Wisconsin Press, Madison, Wisconsin, USA.
- Olden, J. D., D. A. Jackson, and P. R. Peres-Neto. 2002. Predictive models of fish species distributions: a note on proper validation and chance predictions. *Transactions of the American Fisheries Society* 131:329-336.
- Osborne, P. E., J. C. Alonso, and R. G. Bryant. 2001. Modelling landscape-scale habitat use using GIS and remote sensing: a case study with great bustards. *Journal of Applied Ecology* 38:458-471.
- Ozesmi, S. L. and U. Ozesmi. 1999. An artificial neural network approach to spatial habitat modelling with interspecific interaction. *Ecological Modelling* 116:15-31.

- Palma, L., P. Beja, and M. Rodrigues. 1999. The use of sighting data to analyse Iberian Lynx habitat and distribution. *Journal of Applied Ecology* 36:812-824.
- Paukert, C. P. and T. A. Wittig. 2002. Applications of multivariate statistical methods in fisheries. *Fisheries* 27:16-22.
- Pearce, J. and S. Ferrier. 2000. Evaluating the predictive performance of habitat models developed using logistic regression. *Ecological Modelling* 133:225-245.
- Pearce, J. and S. Ferrier. 2000a. An evaluation of alternative algorithms for fitting species distribution models using logistic regression. *Ecological Modelling* 128:127-147.
- Pereira, J. M. C. and R. M. Itami. 1991. GIS-based habitat modeling using logistic multiple-regression - a study of the Mt Graham red squirrel. *Photogrammetric Engineering and Remote Sensing* 57:1475-1486.
- Pulliam, H. R. 1988. Sources, sinks, and population regulation. *American Naturalist* 132:652-661.
- Reckhow, K. H. 1990. Bayesian inference in non-replicated ecological studies. *Ecology* 71:2053-2059.
- Reichert, P. and M. Omlin. 1997. On the usefulness of overparameterized ecological models. *Ecological Modelling* 95:289-299.
- Reunanen, P., A. Nikula, M. Monkkonen, E. Hurme, and V. Nivala. 2002. Predicting occupancy for the Siberian flying squirrel in old-growth forest patches. *Ecological Applications* 12:1188-1198.
- Roloff, G. J. and B. J. Kernohan. 1999. Evaluating reliability of habitat suitability index models. *Wildlife Society Bulletin* 27:973-985.
- Roseberry, J. L. and S. D. Sudkamp. 1998. Assessing the suitability of landscapes for northern bobwhite. *Journal of Wildlife Management* 62:895-902.
- Sauer, J. R. and W. A. Link. 2002. Hierarchical modeling of population stability and species group attributes from survey data. *Ecology* 83:1743-1751.
- Schairer, G. 1999. Landscape Level Evaluation of Northern Bobwhite Habitats in Eastern Virginia using Landsat TM Imagery. M.S. thesis, Virginia Polytechnic Institute and State University, Blacksburg, Virginia. 168 p.
- Schamberger, M. L. and L. J. O'Neil. 1986. Concepts and constraints of habitat-model testing. Pages 5-10 *in* J. Verner, M. L. Morrison, and C. J. Ralph, editors. *Wildlife 2000: Modeling habitat relationships of terrestrial vertebrates*. University of Wisconsin Press, Madison, Wisconsin, USA.
- Solow, A. R. 1994. On the Bayesian estimation of the number of species in a community.

- Ecology 75:2139-2142.
- Sperduto, M. B. and R. G. Congalton. 1996. Predicting rare orchid (small whorled pogonia) habitat using GIS. *Photogrammetric Engineering and Remote Sensing*, 62:1269-1279.
- Stockwell, D. R. B. and D. Peters. 1999. The GARP modeling system: problems and solutions to automated spatial prediction. *International Journal of Geographic Information Science*, 13:143-158.
- Stockwell, D. R. B. and A. T. Peterson. 2002. Effects of sample size on accuracy of species distribution models. *Ecological Modelling* 148:1-13.
- Stoms, D. M., F. W. Davis, and C. B. Cogan. 1992. Sensitivity of wildlife habitat models to uncertainties in GIS data. *Photogrammetric Engineering and Remote Sensing* 58:843-850.
- Storm, G. L., M. J. Lovallo, D. S. Klute, W. M. Tzilkowski, and T. Delong. 1995. Predicting American woodcock presence in Pennsylvania from local and landscape-scale habitat variables. *Northeast Wildlife* 52:39-48.
- Swimley, T. J., T. L. Serfass, R. P. Brooks, and W. M. Tzilkowski. 1998. Predicting river otter latrine sites in Pennsylvania. *Wildlife Society Bulletin* 26:836-845.
- Tobalske, C. and B. W. Tobalske. 1999. Using atlas data to model the distribution of woodpecker Species in the Jura, France. *Condor* 101:472-483.
- Tucker, K., S. P. Rushton, R. A. Sanderson, E. B. Martin, and J. Blaiklock. 1997. Modelling bird distributions - a combined GIS and Bayesian rule-based approach. *Landscape Ecology* 12:77-93.
- Tyre, A. J., H. P. Possingham, and D. B. Lindenmayer. 2001. Inferring process from pattern: can territory occupancy provide information about life history parameters? *Ecological Applications* 11:1722-1737.
- U.S. Fish and Wildlife Service 1981. Standards for the development of habitat suitability index models, *Ecological Services Manual* 103, Release No. 1-81, U.S. Fish and Wildlife Service, Division of Ecological Services, Washington, D.C., USA.
- Van Horne, B. 1983. Density as a misleading indicator of habitat quality. *Journal of Wildlife Management* 47:893-901.
- Van Horssen, P. W., P. P. Schot, and A. Barendregt. 1999. A GIS-based plant prediction model for wetland ecosystems. *Landscape Ecology* 14:253-265.
- Ver Hoef, J. M. 1996. Parametric empirical Bayes methods for ecological applications. *Ecological Applications* 6:1047-1055.
- Verner, J. 1984. The guild concept applied to management of bird populations. *Environmental*

Management 8:1-14.

White, G. Accessed January 2004. Program MARK.

<http://www.cnr.colostate.edu/~gwhite/mark/mark.htm>.

Wiens, J. A. 1989. Spatial scaling in ecology. *Functional Ecology* 3:385-397.

Williams, B. K. and K. Titus. 1988. Assessment of sampling stability in ecological applications of discriminant analysis. *Ecology* 69:1275-1285.

Williams, G. L., K. R. Russell, and W. K. Seitz. 1978. Pattern recognition as a tool in the ecological analysis of habitat. Classification, inventory, and analysis of fish and wildlife habitat: Proceedings of a national symposium. FWS/OBS-78-76. 78. U.S. Fish and Wildlife Service Biological Services Program.

Wood, B. L., L. R. Beck, R. K. Washino, K. A. Hibbard and J. S. Salute. 1992. Estimating high mosquito-producing rice fields using spectral and spatial data. *International Journal of Remote Sensing* 13:2813-2826.

Zaniewski, A. E., Lehmann, A., and J. McC. Overton. 2002. Predicting species spatial distributions using presence-only data: a case study of native New Zealand ferns. *Ecological Modelling* 157:261-280.

Zobel, C. W., Cook, D. F., and Q. J. Nottingham. 2004. An augmented neural network classification approach to detecting mean shifts in correlated manufacturing process parameters. *International Journal of Production Research* 42:741-758.

CHAPTER 2: A COMPARISON OF MODELS USING OCCURRENCE DATA TO PREDICT THE PRESENCE OR ABSENCE OF REPTILE AND AMPHIBIAN SPECIES

ABSTRACT

Databases of species observations from surveys and ad hoc observations are frequently maintained by managers for a particular area. Obtaining information from this type of survey data about the habitat associations of species can be an efficient method of predicting habitat suitability across a landscape. Many multivariate statistical methods have been used to develop models of habitat associations. The objective of this study was to compare the effectiveness of discriminant function analysis, logistic regression, ecological niche factor analysis (ENFA), and Bayesian logistic regression methods to predict species occurrences using a large database of species occurrences. The four methods were used with the same data set to create models of habitat suitability for 29 amphibian and reptile species. Sensitivity, specificity, and Kappa statistics were used to assess model discrimination (the ability to correctly classify locations as present or absent) and compare the four modeling methods. Overall, model discrimination was poor and did not differ among modeling methods. *Rana virgatipes*, the carpenter frog, was the only species that was modeled with a greater than 0.4 Kappa statistic using all four methods. Discriminant function analysis and ENFA predicted presence locations with greater success than logistic regression and Bayesian logistic regression methods. Logistic regression and Bayesian logistic regression correctly predicted absence locations more frequently than discriminant function analysis. All three methods predicted absences more successfully than ENFA. The large number of absence locations and method of defining absence locations influenced the model results. Without knowledge of the detection probability of amphibian and reptile species, it is unknown if the absence locations truly represent absences instead of non-detections. ENFA methods, which only use presence data, are better suited to the survey data set. If information about the detection probability of species exists, the Bayesian modeling methods can be modified to account for detectability. Overall, database collections of species occurrences appear to have limited use as data sources for habitat suitability modeling.

INTRODUCTION

Knowledge of the presence or absence of wildlife species and their distribution across a landscape is critical for wildlife management and conservation planning purposes. However, direct inventories and surveys of wildlife are expensive and time-consuming (Mack et al. 1997) and can be impractical to conduct over extensive landscapes (Cardillo et al. 1999). With an increase in the availability of environmental data from geographic information systems (GIS), modeling habitat suitability has become an increasingly common practice. Habitat suitability modeling provides consistency and allows relatively rapid prediction of distributions for large numbers of species (Fleishman et al. 2001). Modeling habitat suitability is often done using correlative methods that rely on links between species distribution records and environmental predictors (Robertson et al. 2003). Methods of extracting information about species-habitat relationships from existing survey and environmental data to predict the potential distribution of wildlife species across landscapes are valuable tools to managers and decision makers.

Species Distribution Data

One of the most important factors to the success of habitat suitability modeling is the species distribution data that is correlated to the environmental predictors. Most multivariate statistical techniques used for habitat suitability modeling, such as logistic regression and discriminant function analysis, require both presence and absence data. However, absence data are frequently not available or are difficult to reliably obtain (Hirzel et al. 2001). Particularly for elusive and difficult to detect species, failing to detect a species during a survey does not result in a reliable absence classification (Bayley and Peterson 2001). Species with a low probability of detection may have a large number of false absences in survey data. In particular, many amphibian and reptile species may have low probabilities of detection because of their secretive habits. Some amphibian and reptile species can live underground, within the leaf litter, or under cover objects with little activity on the ground surface (Conant and Collins 1998). Knowledge of the probability of detecting these species during a survey is important for determining the reliability of presence/absence survey data.

Problems with Presence-Only Data

Large datasets of species observations (presence-only data) exist from surveys, opportunistic sightings, and museum specimens. In many areas, presence-only data is readily available, but data from systematic surveys to determine presence/absence is too cost and/or time prohibitive to obtain (Zaniewski et al. 2002). In addition to not being amenable to many multivariate statistical modeling techniques, presence-only data present disadvantages for habitat suitability modeling. First, due to the opportunistic nature of species observations, there is an unknown sampling bias (Reutter et al. 2003). All habitats may not be equally or systematically sampled, and the probability of a location being sampled may be correlated with environmental factors (Zaniewski et al. 2003). Despite disadvantages to presence-only data, it does represent information about a species' distribution and methods for modeling habitat suitability from this data could aid conservation and management decisions. Recent interest in developing methods for modeling presence-only data have led to the creation of Ecological Niche Factor Analysis (ENFA, Hirzel et al. 2002), Genetic Algorithm for Rule-Set Prediction (GARP, Stockwell and Peters 1999), and other techniques.

Ecological Niche Factor Analysis

Using presence-only data, ENFA compares the locations of species occurrence in a multidimensional space of ecological variables to the entire study area (Hirzel et al. 2002). Built on the concept of the fundamental niche (Hutchinson 1957), a species is not expected to be randomly distributed in relation to all habitat variables, but is expected to occur within a smaller range of values of each of the habitat variables. The distribution of species locations along habitat variables may differ from the entire study area in terms of the mean (marginality) and/or variance (specialization). A species with high marginality prefers habitat that is different from the overall mean of the study area. A species with high specialization occurs at a very restricted range of a habitat variable in relation to the distribution of the variable for the entire study area. ENFA transforms multiple correlated habitat variables into uncorrelated factors that are linear combinations of the habitat variables (Hirzel et al. 2002). Habitat variables that are key to the

marginality and specialization of the species can be determined and it is possible to compute habitat suitability across a gridded study area to produce a habitat suitability map. Provided that the same study area is used, comparisons among species can be made in terms of their marginality, specialization, and predicted habitat suitability. ENFA methods were selected for analysis because of the recent increase in their application to the wildlife field (Reutter et al. 2003, Zaniewski et al. 2002) and the user-friendly software, Biomapper (Hirzel 2000) created for the implementation of ENFA.

Models Using Presence and Absence Data

Many different multivariate statistical techniques have been used to predict habitat suitability using presence/absence data. These techniques include discriminant analysis, logistic regression (a specific case of a generalized linear model, GLM), generalized additive models (GAM), artificial neural networks, and decision-tree methods (see review by Guisan and Zimmerman 2000). Discriminant function analysis and logistic regression were selected for analysis in this study because these methods appeared to be the most frequently used in the wildlife field and previous comparisons of the methods had been conducted (Manel 1999). These methods develop functions to best discriminate between the presence and absence data based on the values of environmental predictor variables. After the functions are created using the presence/absence data, they can be applied to any location on the landscape and result in a habitat suitability map. In comparisons of a presence-only method (ENFA) with a presence/absence method (GLM) under simulated conditions, the presence/absence method produced a model with better predictive power when the absences truly represented a location of low habitat quality (Hirzel et al. 2001). When the absence data was fallacious, i.e. the absence was due to historical reasons and not habitat suitability, the overall predictive power of the GLM model was decreased (Hirzel et al. 2001). Given that reliable absence data are difficult to obtain or are unavailable, pseudo-absences can be generated by randomly selecting locations where the species was not observed (Stockwell and Peterson 2002b) these multivariate techniques have also been used with presence data and generated pseudo-absences, resulting in models that predicted species distributions with greater accuracy than ENFA (Zaniewski et al. 2002).

Bayesian Models

Bayesian statistical theory represents a philosophical departure from the frequentist statistics commonly used in probability of occurrence models. Because of the difference in philosophy, Bayesian and frequentist methods are not directly comparable. However, both types of methods have been used to address similar problems in the wildlife field. Recently gaining increased acceptance in the wildlife field, Bayesian statistical inference incorporates prior knowledge and data to produce probability estimates of model parameters that can be updated when new data is available (Wade 2000, Ellison 1996). Bayesian philosophy is well suited for questions about probability and situations where data are collected on a recurring basis. While frequentist statistical theory treats an unknown model parameter as a fixed value and the data as random, Bayesian statistical theory treats the data as fixed and parameters as random variables where uncertainty in the parameters is represented by a probability distribution (Reichert and Omlin 1997). Bayesian statistics also allow for the quantification and inclusion of prior knowledge into the statistical procedure, which is not appropriate in frequentist analysis. Model parameters are estimated by using Bayes theorem which combines the prior probability with a likelihood function to result in a posterior probability distribution of the model parameters. The posterior probability incorporates both the prior knowledge and the new observed data with each weighted based on their variance (Gelman et al. 1995).

Using presence/absence data, Bayesian statistics have been used for habitat modeling, as well as to infer estimates of the number of species in a community (Lewins and Joanes 1984, Solow 1994) and the probability of occurrence of unobserved species (Nicholson and Barry 1995). With recent interest in Bayesian statistics, habitat suitability modeling has successfully used Bayesian logistic regression and resulted in probability of occurrence distributions that are more informative than a binary presence/absence classification (Fleishman et al. 2001). Combining information about species detectability derived through sampling or prior knowledge, with models of habitat suitability has the potential to produce cost-effective, accurate models that require less sampling effort (Peterson and Dunham 2003).

Model Comparisons

With many different modeling techniques available, comparisons of the predictive power and efficiency of these techniques are useful (Manel et al. 1999, Hirzel et al. 2001, Zaniewski et al. 2002, Robertson et al. 2003). However, performance of the modeling techniques can be sensitive to different situations, such as differing taxa, sample size (Stockwell and Peterson 2002a), sampling strategy (Hirzel and Guisan 2002), colonization history, and interspecific competition. In particular, the relationship between sample size and model accuracy varied between GARP (a modeling method for presence and pseudo-absence data) and logistic regression models (Stockwell and Peterson 2002a). GARP models reached maximum accuracy with approximately 50 samples, while logistic regression models did not reach maximum accuracy until approximately 100 samples (Stockwell and Peterson 2002a). In another comparison, logistic regression was sensitive to the quantity of data, while ENFA was robust to sample sizes of 300 and 1200 samples (Hirzel et al. 2001). Comparative studies are important for assessing the suitability of particular methods under certain conditions (Pearce and Ferrier 2000, Guisan and Zimmerman 2000) and provide guidance for others to choose an efficient modeling technique for their data. Further study is needed to investigate appropriate sample sizes and whether modeling methods designed for presence/absence data can be used with pseudo-absence data to produce models with greater accuracy than models created from presence-only data (Robertson et al. 2003).

Application to Amphibian and Reptile Taxa

Although habitat modeling has become increasingly frequent, there remains a lack of predictive habitat models for amphibians and reptiles. Amphibian and reptile taxa present a challenge to habitat modelers because of their wide range of aquatic and terrestrial life histories and differences in detectability during surveys (ex. subterranean salamander during dry conditions vs. vocalizing male frogs at a pond). Modeling methods may have varying effectiveness for different taxa, but many methods (e.g. ENFA) have not been used with amphibian and reptile taxa. Given the recent emphasis on developing monitoring programs and management plans for

herpetological species and the difficulty in obtaining survey data, habitat models that could predict their distribution across a landscape using documented occurrence locations would be useful and cost efficient for managers. Because of the lack of existing habitat models for reptiles and amphibians, understanding which modeling methods are best suited for these species can lead to more accurate predictions of distribution as data from the recently implemented monitoring becomes available.

Study Objectives

The goal of this study was to evaluate the effectiveness of using a modeling method for presence-only data versus methods for presence/absence data to predict the occurrence of reptile and amphibian species, given a collection of occurrence locations. The appropriateness of using a collection of occurrence locations and pseudo-absence data to extract information on habitat suitability relationships was examined. My first objective was to compare ENFA, discriminant function, logistic regression, and Bayesian analyses to determine which modeling method best predicts species occurrence when using existing survey data. A second objective was to examine model performance in relation to taxonomic classification to determine if taxa differ in sensitivity to selection of modeling methods. Lastly, I examined the relationship between the sample size of survey data and model performance to determine the sensitivity of modeling methods to the quantity of data.

METHODS

Survey Data

Known presence locations for 29 species (Table 2. 1) were obtained from an amphibian and reptile database for Fort A.P. Hill Military Reservation, Virginia. Fort A.P. Hill consists of 30,329 ha in Caroline and Essex counties in eastern Virginia. The military base is approximately 80% forested and has forestry management programs for loblolly pine and hardwood timber harvest. Elevation ranges from 6.1 to 76 m. Numerous wetlands, including vernal pools, streams, beaver ponds and impoundments, exist within the military base (Mitchell and McNulty 1999).

The database was compiled from an intensive herpetological survey supervised by Dr. Joseph Mitchell (University of Richmond) in 1996-1998 and opportunistic observations of species recorded since 1992 (Mitchell and McNulty 1999). Because surveys in some areas of Fort A.P. Hill were infrequent, records from these areas were excluded. Only geographically-referenced records from the database were used resulting in observations of 2136 individuals of 53 amphibian and reptile species at 1053 locations (Figure 2. 1).

Because all amphibian and reptile species that were observed at a location were recorded, I classified all locations where the species of interest was not observed as non-detections. The locations that were classified as non-detections were used as the basis for generating pseudo-absences. Pseudo-absences (henceforth referred to as absences) were created for each species by selecting all non-detection locations that were farther than 100 m from a presence location of that species. The 100-m buffer was an arbitrary distance to prevent locations near presence locations from being classified as absences. This resulted in 655 to 1023 absence locations for the 29 species (Table 2. 1).

Predictor Variables

Habitat data about land cover, wetlands, and human disturbance were selected as potentially important variables for predicting habitat suitability for amphibians and reptiles. Variables were selected that could be calculated from existing GIS data and were potentially correlated with a wide range of amphibian and reptile species. Existing GIS data for Fort A.P. Hill consisted of a vector road coverage and a land cover polygon coverage obtained from Fort A.P. Hill (1999). In addition, a National Wetlands Inventory wetlands coverage (US Fish and Wildlife Service 1999), a 7.5-minute digital elevation model (DEM, US Geological Survey 2000), and a grid coverage of satellite-derived classified land cover used in the Virginia Gap analysis (Waldon et al. 2001). All grid data had a 30 x 30-m cell size resolution. Spatial data analysis was conducted using ArcView 3.2 (Environmental Systems Research Institute 1999).

Fourteen variables (Table 2. 2) were calculated from the existing GIS data. Distances to the nearest road, forest, and maintained or non-maintained field were calculated from the Fort A.P. Hill land cover polygon coverage and distance to the nearest wetland from the National Wetlands Inventory wetlands coverage. Slope and aspect were calculated from the DEM. Because of the circular nature of aspect data, aspects were transformed using a cosine transformation of the angle in radians for analyses. To determine land cover around the immediate vicinity of each presence and absence location, a 50-m buffer was established around each location and the percentage of each land cover classification in each buffer area was summarized. The percentages of forest, wetland, and open field area in the 50-m buffer area were used as variables in the analysis. Because of other land cover classes, the sum of the percentage variables did not equal 100%. To examine land cover on a larger scale, a neighborhood analysis using 10 x 10 cell (300 x 300-m) area around each cell was conducted to determine the mean area of forest, wetland, non-forest, and disturbed land cover as classified from satellite imagery.

Model Creation

Binary logistic regression and discriminant function analyses were conducted using SPSS, version 11.0 (SPSS 2002) software. A forward stepwise selection procedure was used to build the logistic regression models with an entry alpha value of 0.05 and removal value of 0.10. Stepwise model building procedures were also used for the creation of the discriminant functions for each species. Variables significant at an alpha level of 0.05 were added to the discriminant function and variables significant at an alpha level of 0.10 were retained in the function. Logistic regression and discriminant function models were created for each of the 29 species.

Ecological niche factor analysis was conducted using BioMapper 2.1 software (Hirzel et al. 2000). Because BioMapper uses a grid (raster) maps, all ArcView grid coverages were converted to Idrisi format (using tools provided in the BioMapper 2.1 software) and verified to be the same extent. A raster map of presence locations was created for each species. The same set of predictor variables was used for each species that included all of the variables (Table 2. 2) with the exception of the three land cover variables based on the 50-m buffer of the land cover polygon coverage. Converting the polygon coverage to a grid coverage matching in extent and

grid cell size of 30 x 30-m produced an undesirable level of error at the 50-m scale used for the buffer. All predictor variables were normalized using a Box-Cox transformation. The marginality (measure of the difference between means for the species of interest and the global distribution) and specialization (measure of the difference in variation between the species of interest and global distribution) were calculated for each species. For prediction of habitat suitability, not all factors were used. Because the first few factors incorporate most of the variance, only factors that explained a significant portion of the total variance in comparison to MacArthur's broken-stick distribution were used for prediction (Hirzel et al. 2002). The broken stick distribution represents an expected distribution if the total variance were broken at random into a number of partitions equivalent to the number of eigenvectors. Therefore, comparing the distribution of the ENFA factors with the random, broken-stick distribution allows for the selection of ENFA factors that explain significantly more of the variance in comparison to the expected, random distribution. The factors that were significant were used for the creation of the habitat suitability map.

The specification of the Bayesian model follows the logistic regression format (Neter et al. 1996) by assuming the presence/absence data ($y[i]$) are from a Bernoulli distribution with a probability of presence ($p[i]$). A logit transformation ($p[i]/[1-p[i]]$) of the probability of presence is then related to a regression of environmental covariates. The environmental covariates were individually selected for each species based on the significant variables in the previously calculated logistic regression and discriminant function models. An example of the Bayesian logistic regression model in WinBugs code using two environmental covariates (x_1 and x_2) is depicted below:

```

For (i in 1:N){ y[i] ~ dbern(p[i])
                logit(p[i]) <- alpha0 + alpha1 * x1[i] + alpha2 * x2[i]
            }

```

The model then specifies non-informative priors for each of the logistic regression coefficients ($\alpha_0 - \alpha_2$ in the example above) using a normal distribution. Bayesian analyses were conducted using WinBugs (Version 1.4, Imperial College & MRC, UK) software. All simulations were checked to ensure that convergence was reached. To reduce autocorrelation within the simulation process, the model updates during simulation were thinned by keeping 1 of

every 10 updates. After thinning, the first 5,000 updates of each simulation were discarded (burn in) and the following 50,000 updates were used to produce the posterior distribution. The mean of the posterior distribution for $p[i]$ was used as the predicted probability of occurrence for comparison with the other models.

Predicted probability of occurrence was transformed into predicted presence/absence through use of a threshold value. Locations with predicted probability of occurrence > 0.5 were classified as present. Probabilities less than 0.5 were classified as absent. To compare results among the models, sensitivity (conditional probability that the location is predicted present when it is actually present) and specificity (conditional probability that the location is predicted absent when it is actually absent) were calculated for each model (Fielding and Bell 1997) using a resubstitution of the data (Olden et al. 2002). Kappa statistics (Cohen 1960) have been frequently used to quantify the level of agreement between observed and predicted data. Kappa statistics were calculated for each model to assess discrimination and compare among modeling methods (Manel et al. 2001, Pearce and Ferrier 2000). I followed Monserud and Leemans (1992) suggestions for interpretation of Kappa statistics: no agreement, <0.05 ; very poor, 0.05-0.2; poor, 0.20-0.40; fair, 0.40-0.55; good, 0.55-0.70; very good, 0.70-0.85; excellent, 0.85-0.99; and perfect, 0.99-1.0. Statistical comparisons among modeling methods were made using analysis of variance (ANOVA) with modeling method as a categorical factor. ANOVA was also used to examine differences in model performance among taxa, with Tukey's pairwise comparisons to examine differences among means when differences existed among taxa. Relationships between model accuracy and sample size were quantified using simple linear regression.

RESULTS

Ecological Niche Factor Analyses

Ecological niche factor analysis was performed for all species except *Kinosternon subrubrum*, which had too few occurrence observations. Marginality ranged between 0.349 and 0.916 (Table 2. 3). Marginality values are generally between 0 and 1, with high marginality values denoting that the species was found in a different habitat in relation to the average of the entire military

base. Specialization ranged between 1.066 and 2.722 (Table 2. 3). A larger specialization value denotes a species with a narrower niche (specialist) in relation to the all of the available habitats at the military base. A smaller specialization value indicates a more generalist species. For example, a species with a high marginality score, prefers habitats that are very different from the average available habitat. If the species also had a high specialization score, it would occur only on a restricted range of those habitats. If the species had a lower specialization score it would occur over a wider range of habitats that are different from the overall landscape. A species with a low marginality score would occur in habitats that are common throughout the landscape and represent the average habitat in the landscape. No relationship was noted between marginality and specialization scores (Figure 2. 2). The first factor of ENFA accounts for 100% of the marginality of the species. For the first factor, the percentage of variation explained is a measure of the specialization on the habitat conditions that are the most different from the average for the military base. For example, *Plethodon cinereus* had 52.4% of its specialization accounted for by the marginality factor, meaning that the species shows a very restricted range on the habitat variables (distance to nearest forest, neighboring disturbance, and neighboring forest) from which it differs the most from the overall mean of all areas on the military base. Overall, the marginality factor accounted for between 6.5 – 52.4% of the specialization (Table 2. 3). Eigenvalues ranged from 0.818 to 28.929 for the marginality factor, with higher values denoting greater sensitivity to shifts along this optimal habitat conditions axis (Reutter et al. 2003). The second eigenvector accounted for an additional 18 to 45% of the specialization for each species. When the habitat suitability maps were calculated for each species, the overall amount of information explained by the retained factors (as recommended by McArthur's broken-stick model) ranged from 64 to 93%. The amount of information explained by the model was low (below 75%) for 12 of the 28 species (Table 2. 3).

Habitat suitability maps were created for each species and translated to presence/absence using a 0.5 suitability threshold value. Mean sensitivity (% presence locations correctly predicted) for all species was 68.7% and ranged from 58.3% to 81.1% (Figure 2. 3). Mean specificity (% absence locations correctly predicted) for all species was 56.1% and ranged from 31.4% to 90.7%. Sensitivity was not related to the sample size of presence locations (d.f. = 1,25, F = 2.04, P = 0.166) or absence locations (d.f. = 1,25, F = 2.04, P = 0.166). Similarly, specificity was not

related to the sample size of presence (d.f. = 1,25, F = 3.32, P = 0.08) or absence locations (d.f. = 1,25, F = 0.19, P = 0.848).

The mean Kappa statistic for all ENFA models was 0.083; the Kappa statistic ranged from 0 to 0.426 (Table 2. 4). The model for *Rana virgatipes* had the highest Kappa statistic (0.426) and was the only species model to be classified as having fair agreement between the predicted and observed data. Three species, *Sternotherus odoratus*, *Pseudemys rubiventris* and *Chrysemys picta*, had models with Kappa statistics between 0.2 and 0.4 representing poor agreement among the model predictions with the presence and pseudo-absence data. The remaining 24 species showed very poor or no agreement among the model predictions and observed data. Kappa statistics were not related to the sample size of presence locations (d.f. = 1,25, F = 0.01, P = 0.910) or absence locations (d.f. = 1,25, F = 0.001, P = 0.948).

Examining the taxa groups, one-way ANOVA analyses showed that mean marginality (d.f. = 4, 23, F = 4.56, P = 0.007) and specialization (d.f. = 4, 23, F = 3.11, P = 0.035) scores differed among taxa (Figure 2. 4). Turtle species had higher mean marginality than salamander and lizard taxa using Tukey's pairwise comparisons. Snake species were more specialized than frog/toad species. Sensitivity (d.f. = 4,22, F = 0.49, P = 0.743), specificity (d.f. = 4,22, F = 2.75, P = 0.054), and Kappa statistics (d.f. = 4,22, F = 1.36, P = 0.279) for ENFA models did not differ among taxa groups (Figure 2. 5).

Discriminant Function Analysis

Discriminant function models were built to predict habitat suitability for 28 species, with the exception of *Bufo fowleri* for which no significant predictors were found. Mean sensitivity (% presence locations correctly predicted) for all species was 65.2% and ranged from 33.3% to 93.8% (Figure 2. 6). Models for five species (*Coluber constrictor*, *Elaphe obsoleta*, *Notophthalmus viridescens*, and *Terrapene carolina*) correctly predicted 50% or less of the known presence locations. Models for most species correctly predicted between 50 – 70% of the presence locations correctly. However, models for eight species (*Nerodia sipedon*, *Ambystoma maculatum*, *Kinosternon subrubrum*, *Chrysemys picta*, *Pseudemys rubiventris*, *Rana virgatipes*,

Sternotherus odoratus, and *Cnemidophorus sexlineatus*) showed greater success, predicting at least 70% of the presence locations correctly. Mean specificity (% of absence locations correctly predicted) for all species was 72.3% and ranged from 36.6% to 87.6% (Figure 2. 6). Models for two species (*Cnemidophorus sexlineatus* and *Sceloporus undulatus*) correctly predicted fewer than 50% of the absence locations used in the model creation. Greater than 70% of absence locations were correctly predicted for 19 of the 28 species. Models for six species (*Rana virgatipes*, *Nerodia sipedon*, *Sternotherus odoratus*, *Kinosternon subrubrum*, *Chrysemys picta*, and *Pseudemys rubiventris*) predicted at least 70% of both presence and absence locations correctly. However, discriminant function models for 22 of the 29 species had an unacceptable level of error in either predicting presence and/or absence locations. The sample sizes of neither presence nor absence locations were correlated with sensitivity and specificity rates.

The mean Kappa statistic for all models was 0.153 (range 0.014 – 0.449), showing very poor agreement between model predictions and data used in model creation (Table 2. 4). Models for 20 of the 29 species showed poor agreement with Kappa statistics lower than 0.2. Seven species (*Pseudemys rubiventris*, *Ambystoma maculatum*, *Sternotherus odoratus*, *Rana clamitans*, *Acris crepitans*, *Chrysemys picta*, and *Rana virgatipes*) had Kappa statistics between 0.2 and 0.4, representing poor agreement. Only the model for *Rana virgatipes* (Kappa statistic = 0.449) showed fair agreement between model predictions and observed data. The Kappa statistic was positively related to the sample size of presence locations (d.f. = 1,26, F = 9.24, P = 0.005) and negatively related to the sample size of absence locations (d.f. = 1,26, F = 10.13, P = 0.004). The sample size of presence locations accounted for 26.2% of the variation in Kappa and the sample size of absence locations accounted for 28.0%.

Mean sensitivity for models was not different (d.f. = 4,23, F = 1.88, P = 0.148) among taxa (Figure 2. 7). Mean specificity differed among taxa (d.f. = 4,23, F = 7.94, P < 0.001), with models for lizard species having the least accuracy when predicting absence locations. The Kappa statistic for the discriminant function models did not differ among taxa (d.f. = 4,23, F = 1.42, P = 0.258).

Logistic Regression

Logistic regression models were built for 28 species, with the exception of *Bufo fowleri* for which no significant predictor variables were found. Mean sensitivity for all logistic regression models was 4.6% and ranged from 0 to 51.4% (Figure 2. 8). Models for 27 species predicted fewer than 50% of the known presence locations, with 18 of the models failing to predict any presence locations. The model for *Rana virgatipes* was the only logistic regression model that predicted over 50% of the presence locations correctly. In contrast, mean specificity for all models was 99.7% and ranged from 95.9 to 100% (Figure 2. 8) because models predicted low habitat suitability across the landscape. Model specificity was negatively related to the sample size of presence locations (d.f. = 1,26, F = 13.54, P = 0.001) and positively related to the sample size of absence locations (d.f. = 1,26, F = 9.24, P = 0.005).

The mean Kappa statistic for all models was very poor (0.062) and ranged from 0 to 0.622 (Table 2. 4). For 24 species models, Kappa statistics revealed very poor agreement between model predictions and the data used to create the models. Models for three species, *Chelydra serpentina*, *Rana clamitans*, and *Rana virgatipes*, had poor agreement with Kappa statistics between 0.20 and 0.40. The Kappa statistic was highest for the *Rana virgatipes* logistic regression model (0.62) showing good agreement between the model predictions and observed data. The Kappa statistic was not related to the sample size of presence locations (d.f. = 1,26, F = 0.55, P = 0.464) or the sample size of absence locations (d.f. = 1,26, F = 0.61, P = 0.443).

In a comparison of model performance among taxa, no differences were found (Figure 2. 9). Model sensitivity (d.f. = 4,23, F = 0.84, P = 0.513) and specificity (d.f. = 4,23, F = 0.58, P = 0.679) did not differ among taxa. The mean Kappa statistic also did not vary among taxa (d.f. = 4,23, F = 0.83, P = 0.521).

Bayesian Logistic Regression

Mean sensitivity for all Bayesian logistic regression models was 5.1% and ranged from 0 to 48.6% (Figure 2. 10). All models predicted fewer than half of the known presence locations correctly. The models performed better at predicting absences, with a mean specificity for all models of 99.7% and ranged from 95.7 to 100% (Figure 2. 10). Model specificity was negatively related to the sample size of presence locations (d.f. = 1,26, F = 12.94, P = 0.001) and positively related to the sample size of absence locations (d.f. = 1,26, F = 8.89, P = 0.006).

The mean Kappa statistic for Bayesian logistic regression models for all species was 0.069 and ranged from 0 to 0.591 (Table 2. 4). Models for 24 species showed very poor or no agreement between model predictions and observed data, with Kappa statistics less than 0.20. Models for three species, *Chelydra serpentina*, *Nerodia sipedon*, and *Rana clamitans*, showed poor agreement with Kappa statistics between 0.20 to 0.40. The model for *Rana virgatipes* was the only Bayesian logistic regression model with a Kappa statistic greater than 0.4, demonstrating good agreement between the predicted and observed data. The Kappa statistic was not related to the sample size of presence (d.f. = 1,26, F = 0.26, P = 0.617) or absence locations (d.f. = 1,26, F = 0.35, P = 0.561).

No differences in model performance were found among taxa (Figure 2. 11). Model sensitivity (d.f. = 4,23, F = 0.48, P = 0.749), specificity (d.f. = 4,23, F = 0.52, P = 0.723), and mean Kappa statistic (d.f. = 4,23, F = 0.48, P = 0.747) did not vary among taxa.

Model Comparison

Twenty-seven of the 29 species were suitable for modeling by all four modeling techniques. A comparison of the outcomes of these four modeling techniques revealed differences in the mean sensitivity and specificity. Discriminant function analysis and ENFA methods predicted presence locations (sensitivity) with greater success than logistic regression and Bayesian logistic regression methods (d.f. = 3,100, F = 266.05, P = 0.013). Mean sensitivity of discriminant function and ENFA models was correlated (Pearson correlation coefficient = 0.48, P < 0.001). Mean sensitivity was also correlated between logistic regression and Bayesian logistic regression (Pearson correlation coefficient = 0.98, P < 0.001). Prediction of absence

locations (specificity) also differed among methods with logistic regression and Bayesian logistic regression successfully predicting greater numbers of absences than discriminant function analysis. All three methods had higher accuracy in predicting absences (d.f. = 3,100, $F = 95.65$, $P < 0.001$) than ENFA. Mean specificity for logistic regression and Bayesian logistic regression models were correlated (Pearson correlation coefficient = 0.999, $P < 0.001$). Discriminant function analysis and ENFA specificity were also correlated (Pearson correlation coefficient = 0.422, $P = 0.032$). The mean Kappa for all methods was low (Table 2. 5) and did not significantly differ among methods (d.f. = 3,100, $F = 2.40$, $P = 0.073$). Kappa values among all comparisons of modeling methods were significantly correlated and Pearson correlation coefficients ranged from 0.643 to 0.971.

Rana virgatipes was the only species modeled that showed fair to good agreement between the model predictions and observed data, with Kappa statistics greater than 0.4 for all modeling methods. Nine species were modeled by at least one method with a resulting Kappa statistic greater than 0.2. Primarily aquatic species ($n = 12$) were modeled with greater accuracy than terrestrial species ($n = 14$) by all modeling methods (Figure 2.12). Kappa statistics were higher for models of aquatic species than terrestrial species when using logistic regression ($F = 6.07$, $P = 0.021$), discriminant function analysis ($F = 10.2$, $P = 0.004$), ENFA ($F = 9.69$, $P = 0.005$), and Bayesian models ($F = 6.49$, $P = 0.017$).

DISCUSSION

Comparison of Model Accuracy

Overall, model performance was poor for all modeling methods when used with the existing survey data and pseudo-absences. However, differences in model sensitivity and specificity were found among modeling methods. Discriminant function analysis and ENFA predicted presence locations with greater success than logistic regression or Bayesian logistic regression methods. ENFA was expected to predict a greater number of presence locations because the model is scaled between 0-1 and constrained to have at least one cell on the predicted habitat suitability map equal to one (Hirzel et al. 2002). No other modeling methods had this constraint. Logistic regression and Bayesian logistic regression tended to have greater success predicting

absences, by predicting low habitat suitability for all locations. For the carpenter frog (*Rana virgatipes*), the one species adequately modeled using the existing data, logistic regression produced the model with the best discrimination of presence and absence locations. This was similar to previous comparisons of logistic regression and discriminant function analysis which found that logistic regression had higher classification accuracy (Manel et al. 1999). However, for the majority of species, ENFA appears to be the method best suited for this data set because it produced models with acceptable and consistent accuracy in predicting species' presence in comparison to other modeling methods.

Taxonomic Differences in Model Performance

Differences among taxa were also examined to identify any general trends when modeling amphibian and reptile species. One advantage of ENFA is the ability to compare marginality and specialization among species, which has ecological interpretations related to niche of the species. Mean marginality of turtle species was found to be higher than that of salamander and lizard species. Turtle species were associated with wetland related variables, which varied from the mean of the landscape, which is primarily forested. Snake species were also found to have higher specialization than frog/toad species. The locations of snake species occurred along a more restricted range of habitat variables than the frog/toads species in relation to the entire landscape. These taxonomic differences however, did not translate into differences in model performance. Dividing species into general taxonomic groups may not be the most effective classification when modeling, given the wide range of life histories and habitat preferences within each taxa. For example, both eastern box turtles (*Terrapene carolina*) and eastern musk turtles (*Sternotherus odoratus*) are members of the Chelonia order, but eastern box turtles are primarily terrestrial unlike the aquatic eastern musk turtle (Martof et al. 1980). Examining taxa from a family level of classification may result in patterns of model accuracy that can be generalized by taxa. This was not examined in this study because the limited number of species modeled resulted in few (sometimes only one) species from a family modeled.

Mean sensitivity, specificity, and Kappa values did not differ among taxa for any of the modeling methods, with the exception of discriminant function analysis. Discriminant function

models had less success predicting absence locations for lizard species in comparison to other taxonomic groups. This may be an anomalous result due to the small sample size ($n = 3$) and little variation of habitat preferences and life histories of lizard species. All three lizard species, the six-lined racerunner, five-lined skink, and northern fence lizard, are terrestrial that prefer open areas with basking locations such as rocks, fences, and logs (Martof et al. 1980). Habitat preferences are not completely similar, for example, five-lined skinks prefer moist areas and six-lined racerunners well-drained or sandy areas (Conant and Collins 1998). The inability of discriminant function models to predict absence locations may also be a result of low probability of detection for lizard species. The discriminant function model predicts high suitability across the landscape, which may accurately represent the actual habitat suitability. But, if the probability of detection is low, it is likely that the species was not detected and the model results are compared against an incorrectly classified location. Therefore, the error may be in the data used for the assessment of accuracy and not the model predictions.

Although model accuracy was low for most models, differences in accuracy were noted between aquatic and terrestrial species. Models for aquatic species may have been more accurate than those for terrestrial species for several reasons. The presence data used to create the models may have been more representative and accurate for aquatic species. Methods of sampling for aquatic species may have been associated with higher probabilities of detection. However, sampling effort in observation databases is not quantified and differences in the input data between aquatic and terrestrial species cannot be determined. A second reason for the increased accuracy for aquatic species may be related to the environmental data used as predictors. If identification of wetland-related features is more accurate than other land cover types, the resulting models for aquatic species would be more accurate. The greater accuracy of aquatic species models cannot be attributed to a specific cause because of the post-hoc, correlative nature of the analysis. However, the result does present an area of further research

Correlation Between Sample Size and Model Accuracy

The discrimination capacity of ENFA models was robust to the range of sample sizes examined in this study. This was not the case for discriminant function analysis, where model performance

improved with larger number of presence locations within the data set. With more presence locations in the data set, there were generally fewer absence locations and greater distinction in the measured environmental variables between the presence and absence data. However, because ENFA compares the presence locations to all locations within the landscape, additional presence locations do not result in greater distinction in the habitat variables between presence vs. all locations. When the sample sizes of data sets are small or vary among different species being modeled, ENFA would be the better choice of model.

Although logistic regression has been found to be one of the least accurate modeling methods at low sample sizes (Stockwell and Peterson 2002a), I found no differences in the overall discrimination capacity of logistic regression and Bayesian logistic regression models in relation to sample size. The proportion of absence locations that were correctly predicted did increase when more absence locations were used to construct the models. As more absence locations are used to build the models, the models predict lower habitat suitability for all locations, which translates into a greater number of locations predicted as absences. A logistic regression model that predicts the entire landscape as absences will have a specificity equal to one, but that does not necessarily mean that the model is useful for identifying absence locations.

Appropriateness of Pseudo-Absence Data

The use of pseudo-absence data (Zaniewski 2002) affected all models and the resulting comparison of accuracy. Because the data used to represent absence locations was created from non-detection locations and not a specifically designed study, it is unknown whether these pseudo-absences reflect true absence at a location. It is likely that the pseudo-absence data contains a large number of false absence locations, particularly for difficult to detect species. In addition to obscuring discrimination between presence and absence locations during model building, the false absences are also a problem when determining the accuracy of the model. When the model predicts a location as a presence and the same location is classified as an absence in the data set, it is possible that the model prediction of high suitability and species presence is correct and the location is actually a false absence in the data set where the species was present but not detected. Therefore, measures of specificity (accuracy of absence

predictions) may be misleading if the data used for judging model accuracy do not represent true absences. Because of the large number of absence locations and the inclusion of specificity in the calculation of the statistic, the Kappa statistic may be biased when the data sets do not reflect true absence data (Anderson et al. 2003). However, measures of sensitivity are more reliable because an observation of the species at a site is considered to be proof that the species is present. Therefore, model predictions of presence are compared against known, verified presence locations.

Appropriateness of Observation Data for Modeling

The observation locations from Fort A.P. Hill do not represent the ideal data set for habitat suitability modeling, but it does appear to be representative of the available data at many military and other managed lands. The biggest drawback to this and similar data sets for modeling is the lack of a randomized or systematic survey design. The ad-hoc surveys result in a collection of survey locations that may not represent all available habitats and the locations may be correlated with other factors (Zaniewski et al. 2002). For example, observations of snakes in this data set may be correlated with roads because they were found on the road. Often, convenience is a factor in ad-hoc surveys and survey locations are located within shorter distances from the road than would be expected if random locations in the landscape were selected. In addition, sampling effort may be undeterminable from collections of observation locations and repeat sampling at a location is usually not conducted. In the data set used in this study, numerous sampling methods were used by several different observers resulting in an unknown and probably highly variable probability of detection for each species. Because the database is maintained over multiple years, changes in habitat may have occurred over the time period of data collection, which leads to difficulties when selecting environmental predictor variables for modeling. Overall, there are many valid reasons to not use observational data for habitat suitability modeling and it is obviously not the first choice for a model-building data set. However, this data set is representative of typical data sets available and when better data sets are not available, the question becomes whether or not a model built from the existing data is useful.

The usefulness of a habitat model depends upon the objective of the modeler. Models may be useful for screening, research, or planning, monitoring, and assessment (Guisan and Zimmerman 2000). A habitat model built from a collection of species' observations may be useful for examining potential relationships between habitat variables and species occurrence or identifying unsurveyed areas where the species is likely to occur. If models from the data work to predict patterns in species occurrences, the models are of potential use to managers even if the models do not explain the process behind the patterns. In this study, the models were successful for the carpenter frog, but not for other species. Information was gained by modeling the other species, even though an accurate presence/absence map was not an outcome. Information about the relationship between species' occurrences and the selected environmental variables was gathered and future modeling efforts can look at different aspects of the species' habitat. Additionally, the models can direct future field work towards species whose occurrence is not well predicted by current habitat models.

Conclusions

Overall, model performance was less than ideal for all modeling methods and amphibian and reptile species. ENFA analysis appears to be the method best suited for the occurrence data set, given the lack of reliable absence data and relatively high model sensitivities. ENFA does allow for information on species distributions to be extracted from a large collection of opportunistic and ad hoc occurrence observations. Bayesian models approach the modeling problem with a different philosophy, so the choice between using a Bayesian logistic regression model and a frequentist method should be based on more than just the resulting accuracy of the model. Bayesian models have the capability to incorporate the probability of detection of a species, but information about detectability cannot be extracted from the presence-only data. Given a priori knowledge about detectability or a small-scale field survey, Bayesian models have potential for being a cost-effective approach to modeling (Peterson and Dunham 2003). However, the most successful approach to habitat suitability modeling is still specifically designed surveys that can collect reliable presence and absence data.

LITERATURE CITED

- Bayley, P.B. and J.T. Peterson. 2001. An approach to estimate probability of presence and richness of fish species. *Transactions of the American Fisheries Society*, 130:620-633.
- Cardillo, M., MacDonald, D.W., and Rushton, S.P. 1999. Predicting mammal species richness and distributions: testing the effectiveness of satellite-derived land cover data. *Landscape Ecology*, 14:423-435.
- Cohen, J. 1960. A coefficient of agreement for nominal scales. *Educational and Psychological Measurement*, 20:37-46.
- Conant, R. and J. T. Collins. 1998. *A field guide to reptiles and amphibians of eastern and central North America*. Third edition, expanded. Houghton Mifflin Company, Boston, Massachusetts, USA. 616 pp.
- Dillon, W. R. and M. Goldstein. 1984. *Multivariate analysis methods and applications*. John Wiley and Sons, New York, NY. 587 pp.
- Ellison, A. M. 1996. An introduction to Bayesian inference for ecological research and environmental decision-making. *Ecological Applications* 6:1036-1046.
- Environmental Systems Research Institute. 1999. *ArcView 3.2*. Environmental Systems Research Institute, Inc., Redlands, California, USA.
- Fielding, A.H. and J.F. Bell. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, 24:38-49.
- Fleishman, E., MacNally, R., Fay, J.P., and D.D. Murphy. 2001. Modeling and predicting species occurrence using broad-scale environmental variables: an example with butterflies of the Great Basin. *Conservation Biology*, 15:1674-1685.
- Gelman, A., J. B. Carlin, H. S. Stern, and D. B. Rubin. 1995. *Bayesian data analysis*. Chapman and Hall/CRC, Boca Raton, Florida, USA.
- Guisan, A., and N.E. Zimmerman. 2000. Predictive habitat distribution models in ecology. *Ecological Modelling*, 135:147-186.
- Hirzel, A. and A. Guisan. 2002. Which is the optimal sampling strategy for habitat suitability modeling. *Ecological Modelling*, 157:331-341.
- Hirzel, A. H., J. Hausser, D. Chessel, and N. Perrin. 2002. Ecological-niche factor analysis: how to compute habitat-suitability maps without absence data? *Ecology*, 2027-2036.
- Hirzel, A.H., Helfer, V., and F. Metral. 2001. Assessing habitat-suitability models with a virtual species. *Ecological Modelling*, 145:111-121.
- Hirzel, A. H., Hausser, J., and N. Perrin. 2000. *Biomapper 2.1*. Laboratory for Conservation Biology, University of Lausanne, Switzerland.

- Hutchinson, G. E. 1957. Concluding remarks. Cold Spring Harbor Symposium on Quantitative Biology, 22, Population Studies: Animal Ecology and Demography, Cold Spring Harbor, pp. 415-427.
- Lewins, W. A. and D. N. Joanes. 1984. Bayesian estimation of the number of species. *Biometrics* 40:323-328.
- Mack, E.L., Firbank, L.G., Bellamy, P.E., Hinsley, S.A., and Veitch, N. 1997. The comparison of remotely sensed and ground-based habitat area data using species-area models. *Journal of Applied Ecology*, 34:1222-1228.
- Manel, S., Dias, J.M., and Ormerod, S.J. 1999. Comparing discriminant analysis, neural networks, and logistic regression for predicting species distributions: a case study with a Himalayan river bird. *Ecological Modelling*, 120:337-347.
- Manel, S., Williams, H.C., and S.J. Ormerod. 2001. Evaluating presence-absence models in ecology: the need to account for prevalence. *Journal of Applied Ecology*, 38:921-931.
- Martof, B.S., Palmer, W.M., Bailey, J.R., and J.R. Harrison III. 1980. Amphibians and reptiles of the Carolinas and Virginia. The University of North Carolina Press, Chapel Hill, North Carolina, USA.
- Mitchell, J.C. and McNulty, S. 1999. Distribution and habitat affinities of amphibians and reptiles on Fort A.P. Hill, Virginia: A Gap analysis study. Final Report submitted to Environmental and Natural Resources Division, Fort A.P. Hill, Bowling Green, Virginia. Fish and Wildlife Information Exchange, Virginia Polytechnic Institute and State University. 112 pp.
- Monserud, R. A. and R. Leemans. 1992. Comparing global vegetation maps with the Kappa statistic. *Ecological Modelling*, 62:275-293.
- Neter, J., Kutner, M. H., Nachtsheim, C. J., and W. Wasserman. 1996. Logistic regression, Poisson regression, and generalized linear models. Chapter 14 *In Applied Linear Regression Models*. Third edition. Irwin Book Team, Times Mirror Higher Education Group, Chicago, Illinois, USA.
- Nicholson, M. and J. Barry. 1995. Inferences from spatial surveys about the presence of an unobserved species. *Oikos* 72:74-78.
- Olden, J.D., Jackson, D.A., and P.R. Peres-Neto. 2002. Predictive models of fish species distributions: a note on proper validation and chance predictions. *Transactions of the American Fisheries Society*, 131:329-336.
- Pearce, J. and S. Ferrier. 2000. Evaluating the predictive performance of habitat models developed using logistic regression. *Ecological Modelling*, 133:225-245.

- Peterson, J. T. and J. Dunham. 2003. Combining inferences from models of capture efficiency, detectability, and suitable habitat to classify landscapes for conservation of threatened bull trout. *Conservation Biology*, 17(4):1070-1077.
- Reichert, P. and M. Omlin. 1997. On the usefulness of overparameterized ecological models. *Ecological Modelling* 95:289-299.
- Reutter, B. A., Helfer, V., Hirzel, A. H., and P. Vogel. 2003. Modelling habitat-suitability using museum collections: an example with three sympatric *Apodemus* species from the Alps. *Journal of Biogeography*, 30:581-590.
- Robertson, M. P., Peter, C. I., Villet, M. H., and B. S. Ripley. 2003. Comparing models for predicting species' potential distributions: a case study using correlative and mechanistic predictive modeling techniques. *Ecological Modelling*, 164:153-167.
- Solow, A. R. 1994. On the Bayesian estimation of the number of species in a community. *Ecology* 75:2139-2142.
- SPSS, Inc. 2002. SPSS. Version 11.0. SPSS, Inc., Chicago, Illinois, USA.
- Stockwell, D. R. B. and D. Peters. 1999. The GARP modeling system: problems and solutions to automated spatial prediction. *International Journal of Geographic Information Science*, 13:143-158.
- Stockwell, D. R. B. and A. T. Peterson. 2002a. Effects of sample size on accuracy of species distribution models. *Ecological Modelling*, 148:1-13.
- Stockwell, D. R. B. and A. T. Peterson. 2002b. Controlling bias in biodiversity data. Pages 537-546 In *Predicting species occurrences: issues of accuracy and scale*. Scott, J. M., Heglund, P. J., Morrison, M. L., Haufler, J. B., Raphael, M. G., Wall, W. A., and F. B. Samson, editors. Island Press, Washington, DC, USA.
- US Fish and Wildlife Service. 1999. National Wetlands Inventory. Download data webpage, <http://www.nwi.fws.gov/downloads.htm>.
- US Geological Survey. 2000. US GeoData Digital Elevation Model. Fact Sheet 040-00, <http://mac.usgs.gov/mac/isb/pubs/factsheets/fs04000.html>.
- Wade, P. R. 2000. Bayesian methods in conservation biology. *Conservation Biology*, 14:1308-1316.
- Waldon, J. L., Vaughn, M., Kbpfer, S. D., McNulty, S., Morton, D., Schairer, G., Clark, N., and J. McClafferty. 2001. The Virginia gap analysis project: final report. Submitted to the US Geological Survey, Biological Resources Division, Gap Analysis Program. <http://www.cmiweb.org/anonymous/vagap/vareport.pdf>

Williams, A. K. and J. Berkson. 2003. Reducing false absences in survey data: detection probabilities of red-backed salamanders. *Journal of Wildlife Management*, in review.

Zaniewski, A. E., Lehmann, A., and J. McC. Overton. 2002. Predicting species spatial distributions using presence-only data: a case study of native New Zealand ferns. *Ecological Modelling*, 157:261-280.

Table 2. 1. Amphibian and reptile species selected for modeling probability of occurrence and the number of presence and absence locations recorded at Fort A.P. Hill during 1992-1998.

Scientific Name	Common Name	Taxa	Presence Locations	Absence Locations
<i>Acris crepitans</i>	Northern cricket frog	Frog/Toad	160	694
<i>Ambystoma maculatum</i>	Spotted salamander	Salamander	199	678
<i>Ambystoma opacum</i>	Marbled salamander	Salamander	41	949
<i>Bufo americanus</i>	American toad	Frog/Toad	98	772
<i>Bufo fowleri</i>	Fowler's toad	Frog/Toad	49	879
<i>Carphophis amoenus</i>	Eastern wormsneak	Snake	41	972
<i>Chelydra serpentina</i>	Snapping turtle	Turtle	29	956
<i>Chrysemys picta</i>	Eastern painted turtle	Turtle	96	801
<i>Cnemidophorus sexlineatus</i>	Six-lined racerunner	Lizard	16	1023
<i>Coluber constrictor</i>	Northern black racer	Snake	37	937
<i>Diadophis punctatus</i>	Northern ring-necked snake	Snake	21	1000
<i>Elaphe obsoleta</i>	Black rat snake	Snake	25	985
<i>Eumeces fasciatus</i>	Five-lined skink	Lizard	62	869
<i>Hyla chrysocelis</i>	Cope's gray treefrog	Frog/Toad	163	682
<i>Kinosternon subrubrum</i>	Eastern mud turtle	Turtle	9	1008
<i>Nerodia sipedon</i>	Northern water snake	Snake	30	903
<i>Notophthalmus viridescens</i>	Red spotted newt	Salamander	119	725
<i>Plethodon cinereus</i>	Red-backed salamander	Salamander	63	937
<i>Plethodon cylindraceus</i>	White-spotted slimy salamander	Salamander	38	940
<i>Pseudacris crucifer</i>	Northern spring peeper	Frog/Toad	101	789
<i>Pseudemys rubiventris</i>	Northern red-bellied cooter	Turtle	44	902
<i>Rana catesbeiana</i>	American bullfrog	Frog/Toad	66	808
<i>Rana clamitans</i>	Northern green frog	Frog/Toad	176	655
<i>Rana palustris</i>	Pickerel frog	Frog/Toad	37	900
<i>Rana sylvatica</i>	Wood frog	Frog/Toad	65	890
<i>Rana virgatipes</i>	Carpenter frog	Frog/Toad	72	865
<i>Sceloporus undulatus</i>	Northern fence lizard	Lizard	41	912
<i>Sternotherus odoratus</i>	Stinkpot	Turtle	46	900
<i>Terrapene carolina</i>	Eastern box turtle	Turtle	46	937

Table 2. 2. Habitat variables measured for each location in the presence/absence data set.

Variable Name	Variable Description
Distance to Nearest Road	Distance (m) from the location to the nearest asphalt, gravel, or dirt road
Distance to Nearest Wetland	Distance (m) from the location to nearest area classified as wetland
Distance to Nearest Forest	Distance (m) from the location to nearest area classified as forest
Distance to Nearest Non-Forest	Distance (m) from the location to nearest area classified as a maintained or non-maintained field
Percentage of Forest in 50 m Buffer	Percent of forested area in a surrounding 50 m buffer around each location from the Fort A.P. Hill landcover map
Percentage of Wetland in 50 m Buffer	Percent of wetland area in a surrounding 50 m buffer around each location from the National Wetlands Inventory map
Percentage of Field in 50 m Buffer	Percent of field area in a surrounding 50 m buffer around each location from the Fort A.P. Hill landcover map
Neighboring Disturbance	Percent of grid cells classified as disturbed in a 10x10 window (300x300m) from classified satellite imagery
Neighboring Forest	Percent of grid cells classified as forest in a 10x10 window (300x300m) from classified satellite imagery
Neighboring Wetland	Percent of grid cells classified as wetland in a 10x10 window (300x300m) from classified satellite imagery
Neighboring Field	Percent of grid cells classified as field in a 10x10 window (300x300m) from classified satellite imagery
Elevation	Elevation from the digital elevation model (m)
Slope	Slope derived from the digital elevation model
Aspect	Aspect derived from the digital elevation model and cosine transformed

Table 2. 3. Eigenvalues and percent of variation explained for the first marginality factor, marginality scores, specialization scores, and the % of information explained by the factors included in the habitat suitability model for each species calculated using ecological niche factor analysis.

Species	First Eigenvalue	% Variation Explained	Marginality	Specialization	% Information Explained
<i>Chelydra serpentina</i>	4.044	21.8	0.808	1.361	93
<i>Plethodon cinereus</i>	24.676	52.4	0.481	2.169	91
<i>Diadophis punctatus</i>	28.929	39.0	0.524	2.722	88
<i>Carphophis amoenus</i>	8.803	33.0	0.403	1.632	87
<i>Cnemidophorus sexlineatus</i>	18.181	37.6	0.554	2.200	86
<i>Nerodia sipedon</i>	7.815	25.8	0.906	1.740	86
<i>Rana virgatipes</i>	4.536	17.3	0.916	1.619	86
<i>Eumeces fasciatus</i>	2.401	13.2	0.455	1.348	82
<i>Pseudemys rubiventris</i>	3.241	15.2	0.891	1.460	82
<i>Sceloporus undulatus</i>	2.902	13.8	0.513	1.449	82
<i>Plethodon cylindraceus</i>	4.833	23.0	0.349	1.449	80
<i>Sternotherus odoratus</i>	6.934	22.3	0.911	1.765	80
<i>Elaphe obsoleta</i>	3.711	14.2	0.673	1.619	79
<i>Ambystoma opacum</i>	1.926	8.9	0.505	1.475	78
<i>Rana palustris</i>	2.007	9.0	0.729	1.492	76
<i>Rana catesbeiana</i>	1.974	15.2	0.710	1.141	75
<i>Bufo fowleri</i>	1.441	11.1	0.645	1.140	73
<i>Chrysemys picta</i>	3.599	19.7	0.855	1.351	73
<i>Acris crepitans</i>	0.818	6.5	0.699	1.124	72
<i>Notophthalmus viridescens</i>	1.030	8.2	0.588	1.118	71
<i>Pseudacris crucifer</i>	1.180	10.3	0.652	1.071	71
<i>Coluber constrictor</i>	2.052	16.4	0.661	1.119	69
<i>Rana clamitans</i>	1.255	10.2	0.671	1.109	67
<i>Rana sylvatica</i>	1.111	7.7	0.472	1.198	66
<i>Terrapene carolina</i>	1.467	9.9	0.604	1.216	66
<i>Ambystoma maculatum</i>	1.313	11.3	0.580	1.076	65
<i>Hyla chrysoscelis</i>	1.454	12.8	0.576	1.066	65
<i>Bufo americanus</i>	1.042	9.0	0.527	1.075	64

Table 2. 4. Kappa statistics for ecological niche factor (ENFA), discriminant analysis (DFA), logistic regression (LR), and Bayesian models for amphibian and reptiles species at Fort A.P. Hill.

Species	ENFA Kappa	DFA Kappa	LR Kappa	Bayesian Kappa
<i>Acris crepitans</i>	*	0.352	0	0
<i>Ambystoma maculatum</i>	0.109	0.227	0.076	0.09
<i>Ambystoma opacum</i>	0.023	0.037	0	0
<i>Bufo americanus</i>	0.012	0.125	0	0
<i>Bufo fowleri</i>	0.024	**	**	**
<i>Carphophis amoenus</i>	0.060	0.107	0	0
<i>Chelydra serpentina</i>	0.068	0.096	0.237	0.237
<i>Chrysemys picta</i>	0.202	0.379	0.162	0.162
<i>Cnemidophorus sexlineatus</i>	0.035	0.014	0	0
<i>Coluber constrictor</i>	0.0	0.086	0	0
<i>Diadophis punctatus</i>	0.064	0.071	0	0.084
<i>Elaphe obsoleta</i>	0.015	0.081	0	0
<i>Eumeces fasciatus</i>	0.008	0.084	0	0
<i>Hyla chrysoscelis</i>	0.028	0.172	0.007	0.008
<i>Kinosternon subrubrum</i>	***	0.053	0	0
<i>Nerodia sipedon</i>	0.163	0.173	0.121	0.270
<i>Notophthalmus viridescens</i>	0.0	0.128	0	0
<i>Plethodon cinereus</i>	0.086	0.182	0.027	0.025
<i>Plethodon cylindraceus</i>	0.053	0.160	0	0
<i>Pseudacris crucifer</i>	0.085	0.086	0	0
<i>Pseudemys rubiventris</i>	0.261	0.217	0	0
<i>Rana catesbeiana</i>	0.094	0.183	0	0
<i>Rana clamitans</i>	0.047	0.307	0.273	0.271
<i>Rana palustris</i>	0.082	0.089	0	0
<i>Rana sylvatica</i>	0.016	0.033	0.028	0.028
<i>Rana virgatipes</i>	0.426	0.449	0.622	0.591
<i>Sceloporus undulatus</i>	0.010	0.021	0	0
<i>Sternotherus odoratus</i>	0.307	0.279	0.178	0.178
<i>Terrapene carolina</i>	0.0	0.099	0	0

* No eigenvalues were greater than 1, therefore no habitat suitability map was constructed.

** No significant relationships found with any environmental predictor variables.

*** Too few presence observations to perform ENFA.

Table 2. 5. Comparison of mean sensitivity, specificity, and Kappa statistics for each of the four modeling methods.

Modeling Method	Mean Sensitivity ^a	Mean Specificity ^b	Mean Kappa
Ecological Niche Factor Analysis	0.68	0.57	0.087
Discriminant Function Analysis	0.65	0.72	0.15
Logistic Regression	0.050	0.99	0.067
Bayesian Logistic Regression	0.055	0.99	0.075

^a Proportion of presence locations correctly predicted.

^b Proportion of absence locations correctly predicted.

FIGURE LEGENDS

Figure 2. 1. Map of Fort A.P. Hill showing locations of amphibian and reptile occurrence observations (dots), boundary of the military base, and roads within the base. The large area without species observations had few surveys, therefore presence locations from the area were not used in modeling.

Figure 2. 2. Marginality and specialization scores calculated from ecological niche factor analysis for each amphibian and reptile species.

Figure 2. 3. Sensitivity and specificity for ecological niche factor analysis models calculated for each species.

Figure 2. 4. Boxplots of marginality and specialization for each taxa calculated using ecological niche factor analysis. Boxplots represent the median value and first and third quartiles. Whiskers on the boxplot represent the lowest and highest values within the limits set by 1.5 times the interquartile range. Dots represent outliers.

Figure 2. 5. Boxplots of sensitivity, specificity, and Kappa statistics by taxa calculated using ecological niche factor analysis.

Figure 2. 6. Sensitivity and specificity for discriminant function models developed for each amphibian and reptile species.

Figure 2. 7. Boxplots of sensitivity, specificity, and Kappa statistics by taxa for discriminant function models.

Figure 2. 8. Sensitivity and specificity for logistic regression models developed for each amphibian and reptile species. A limited number of points are visible on the graph, because multiple species had similar values of sensitivity = 0 and specificity = 1, resulting in overlapping points.

Figure 2. 9. Boxplots of sensitivity, specificity, and Kappa statistics by taxa for logistic regression models.

Figure 2. 10. Sensitivity and specificity for Bayesian logistic regression models developed for each amphibian and reptile species. A limited number of points are visible on the graph, because multiple species had similar values of sensitivity = 0 and specificity = 1, resulting in overlapping points.

Figure 2. 11. Boxplots of sensitivity, specificity, and Kappa statistics by taxa for Bayesian logistic regression models.

Figure 2. 12. Boxplots of Kappa statistics for aquatic and terrestrial species using logistic regression (LR), discriminant function analysis (DFA), ecological niche factor analysis (ENFA), and Bayesian logistic regression (BAYES) models. A = Aquatic species, T = Terrestrial species.



Figure 2. 1.

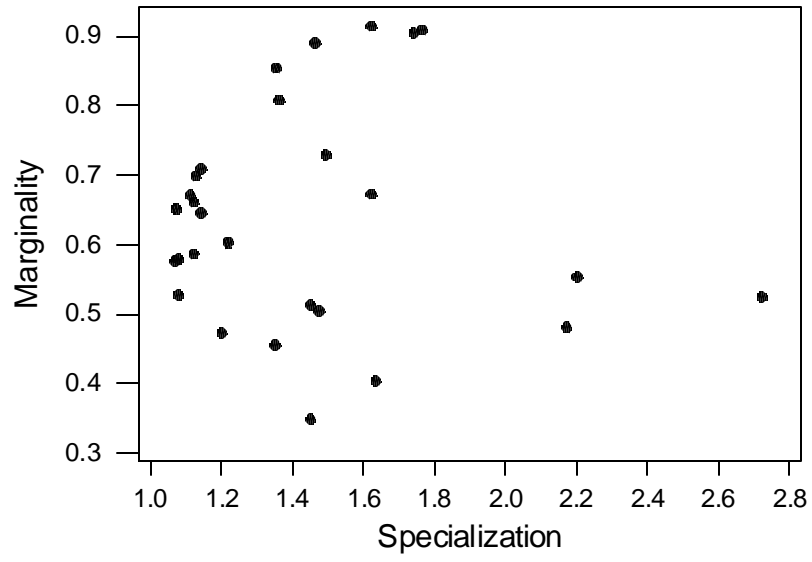


Figure 2. 2.

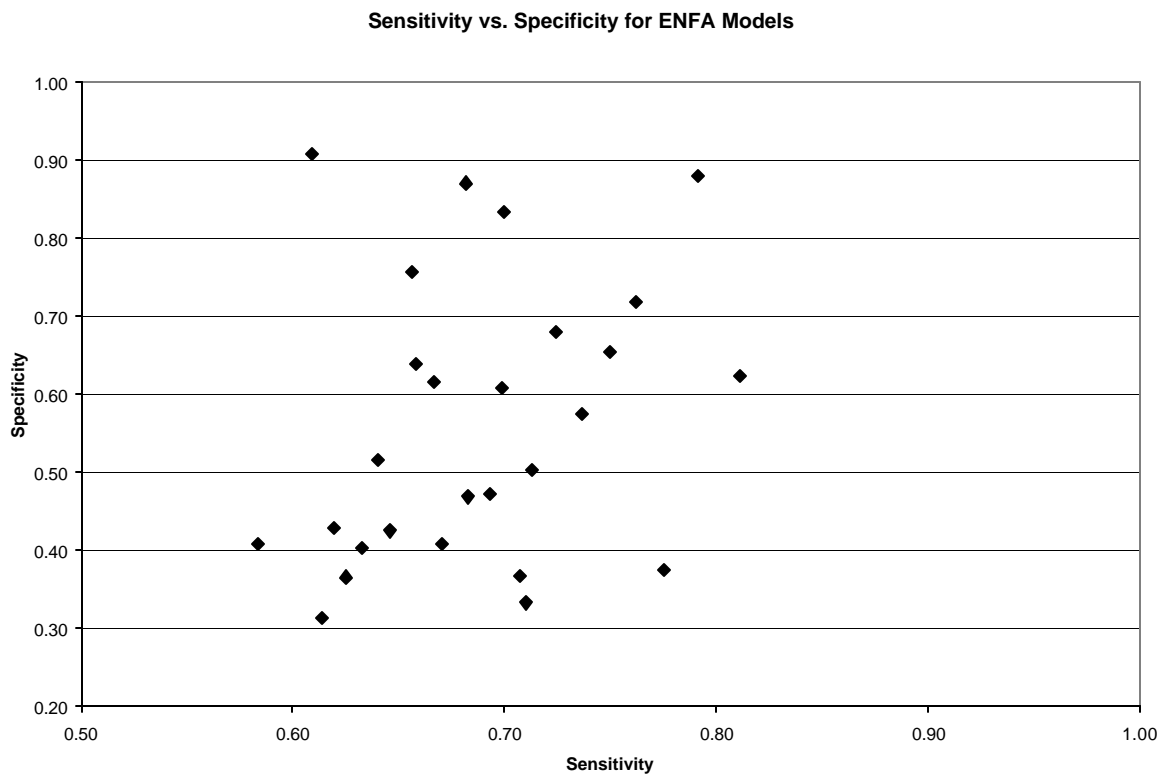


Figure 2. 3.

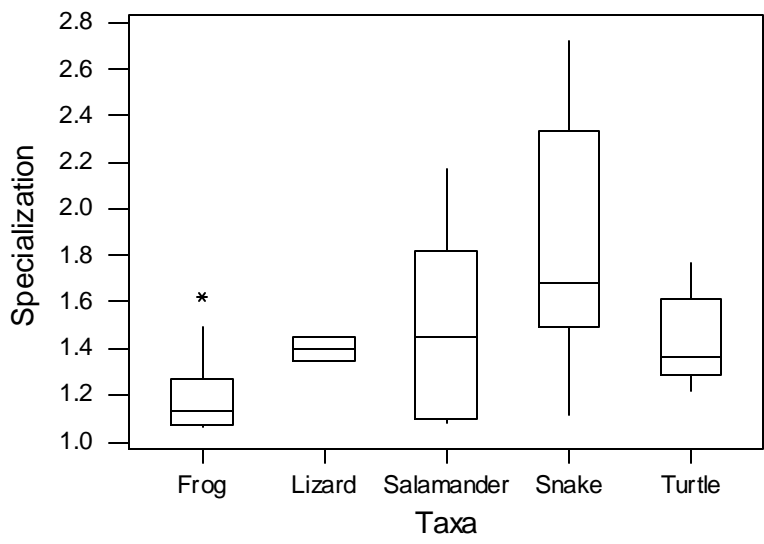
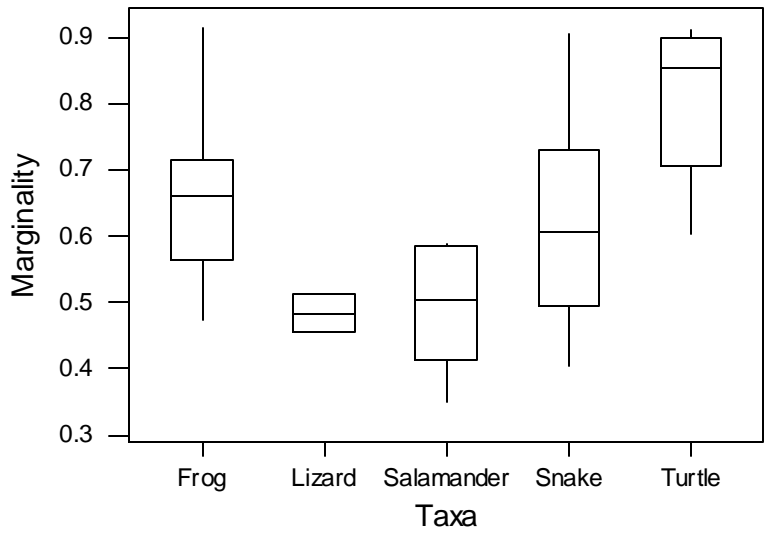
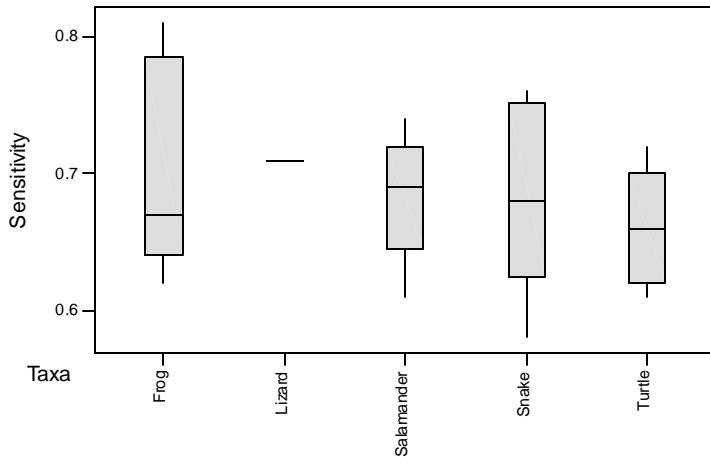
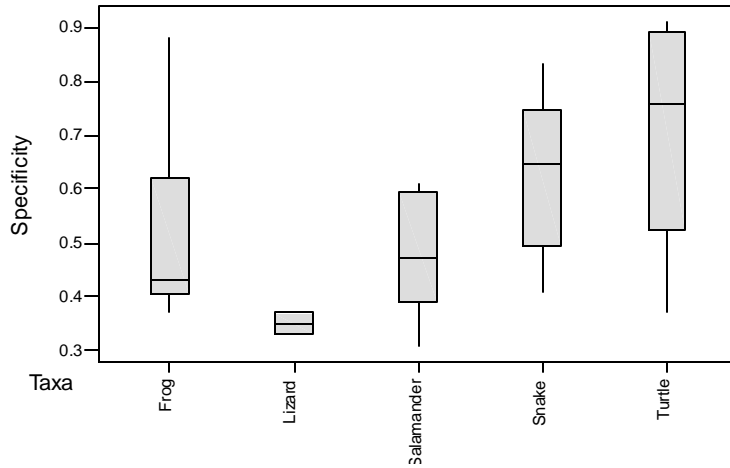


Figure 2. 4.

Boxplots of Sensitivity by Taxa



Boxplots of Specificity by Taxa



Boxplots of Kappa by Taxa

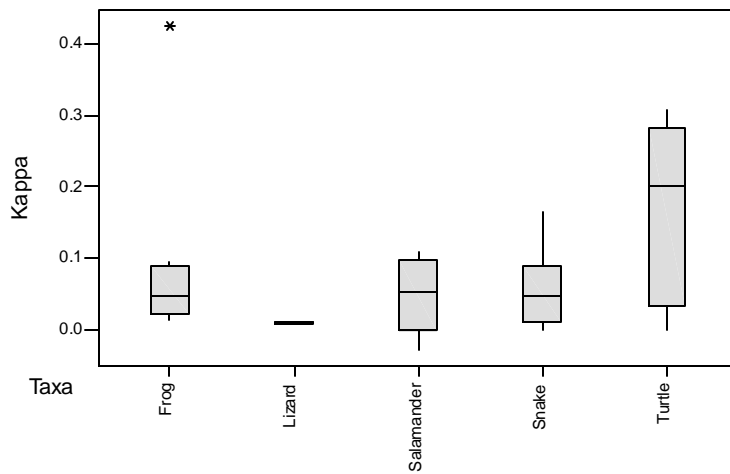


Figure 2. 5.

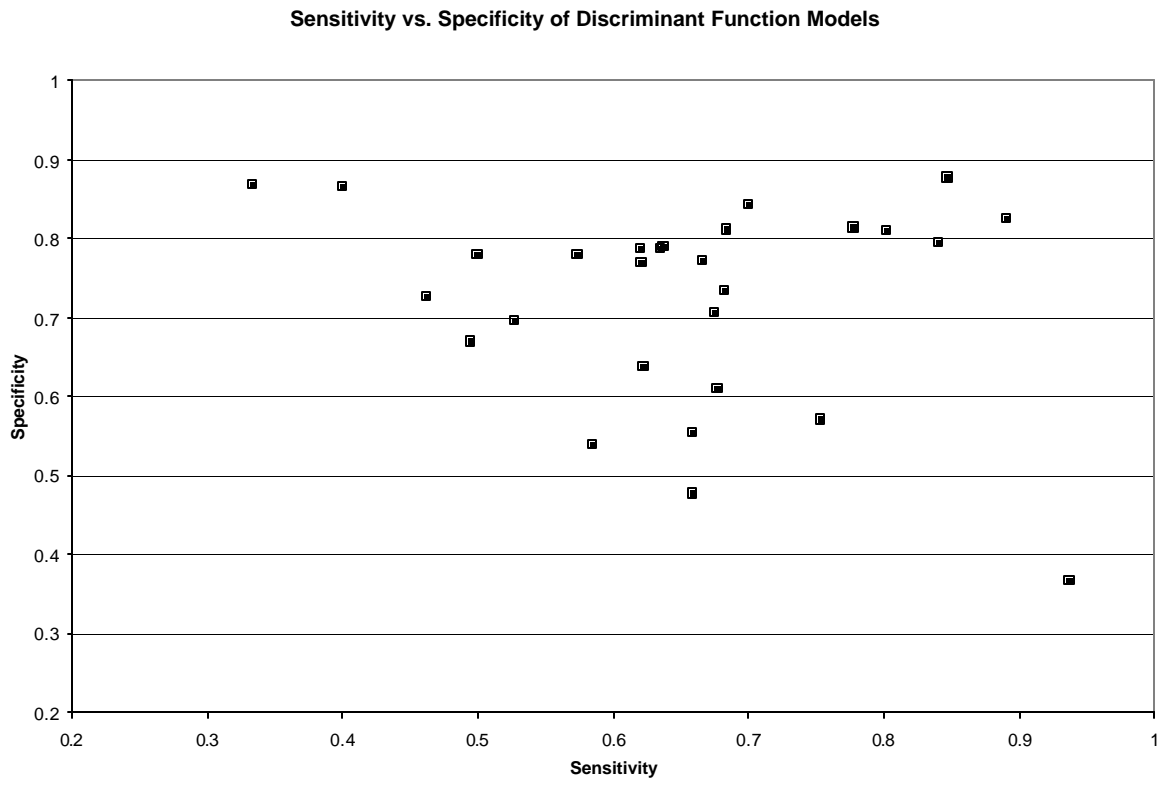
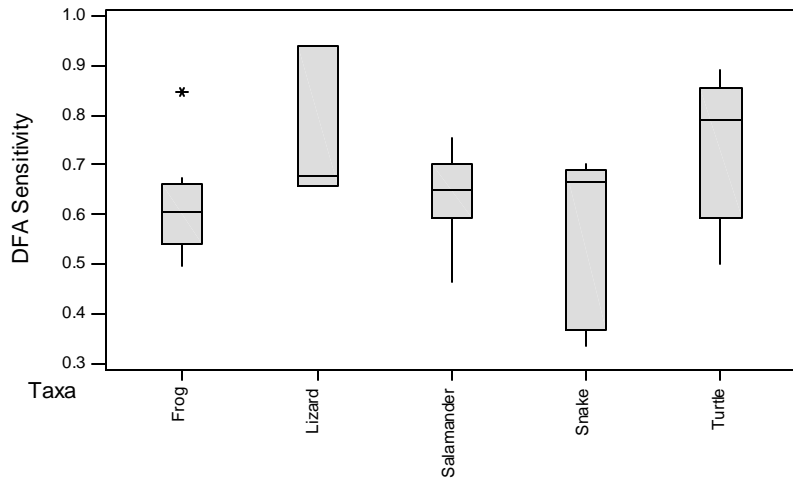
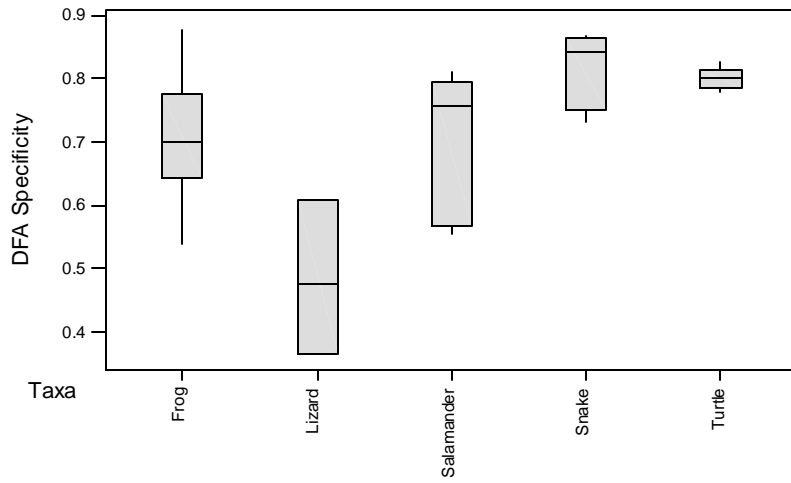


Figure 2. 6.

Boxplots of DFA Sensitivity by Taxa



Boxplots of DFA Specificity by Taxa



Boxplots of DFA Kappa by Taxa

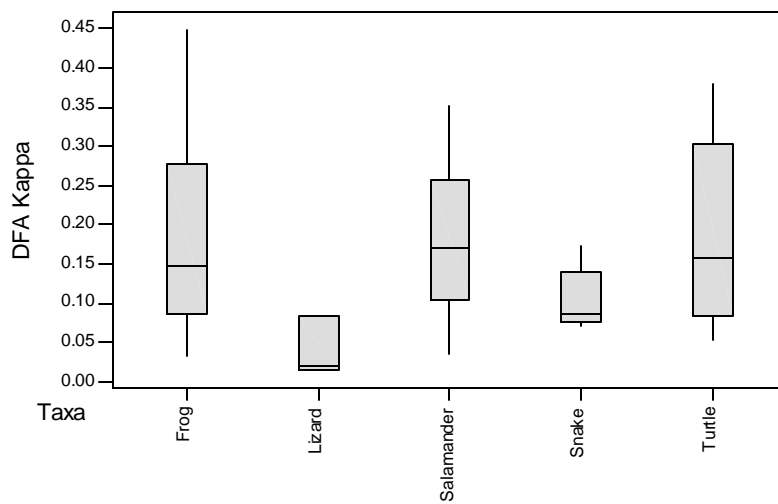


Figure 2. 7.

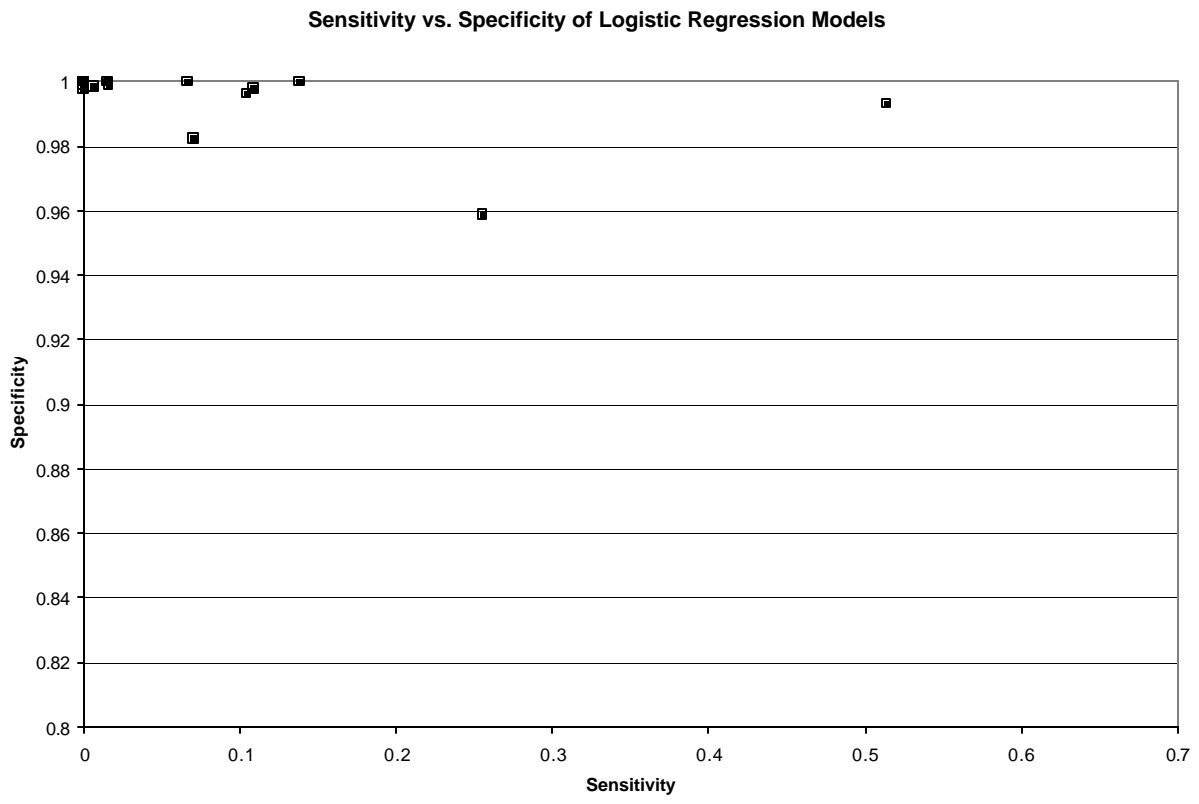
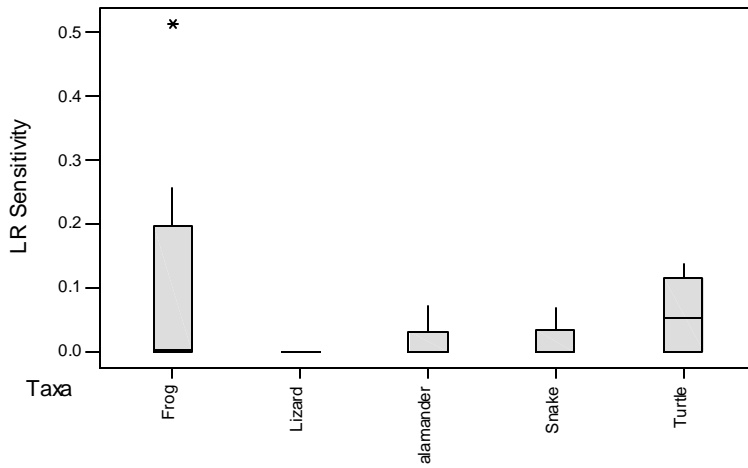
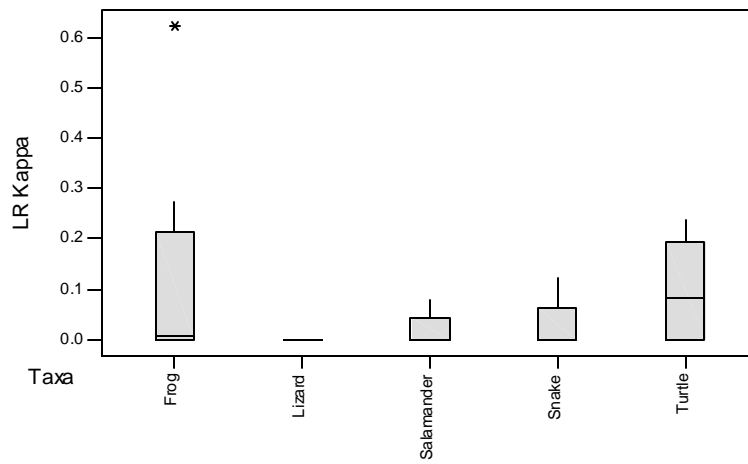


Figure 2. 8.

Boxplots of LR Sensitivity by Taxa



Boxplots of LR Kappa by Taxa



Boxplots of LR Specificity by Taxa

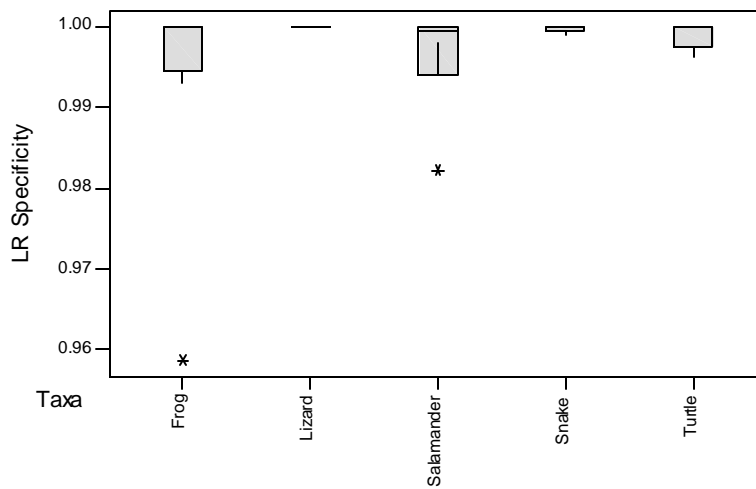


Figure 2. 9.

Sensitivity vs. Specificity of Bayesian Logistic Regression Models

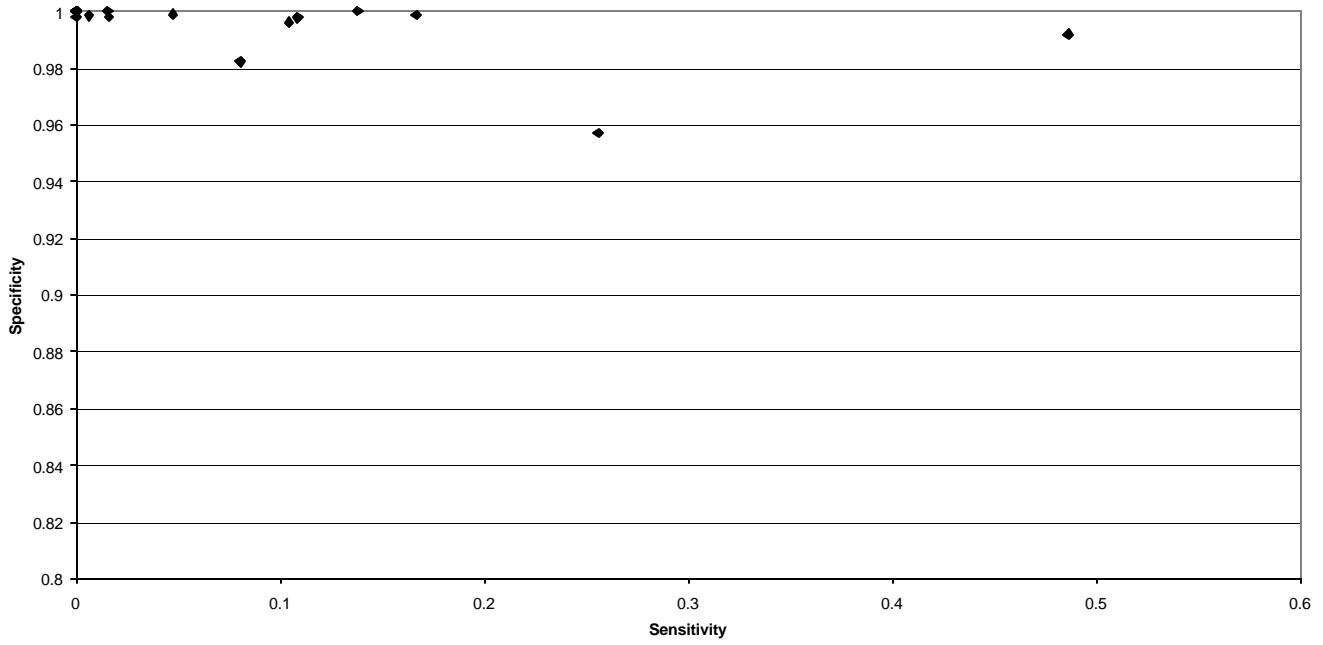


Figure 2. 10.

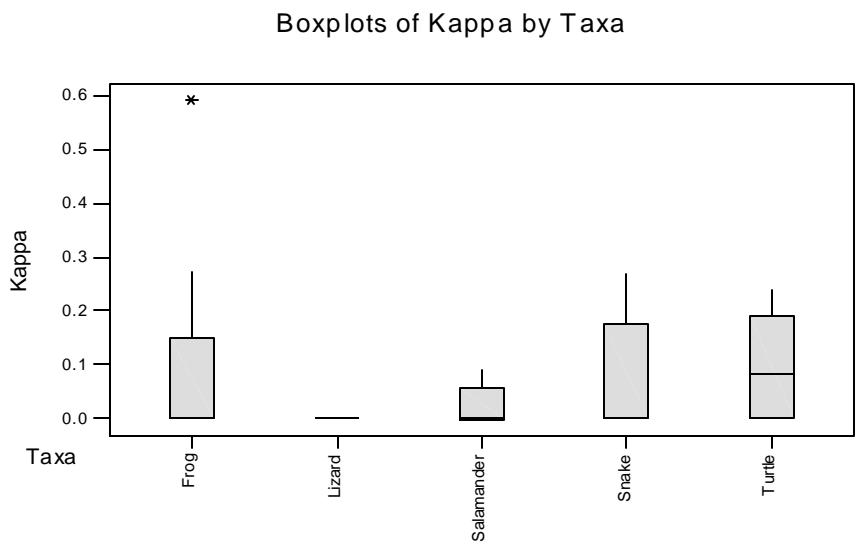
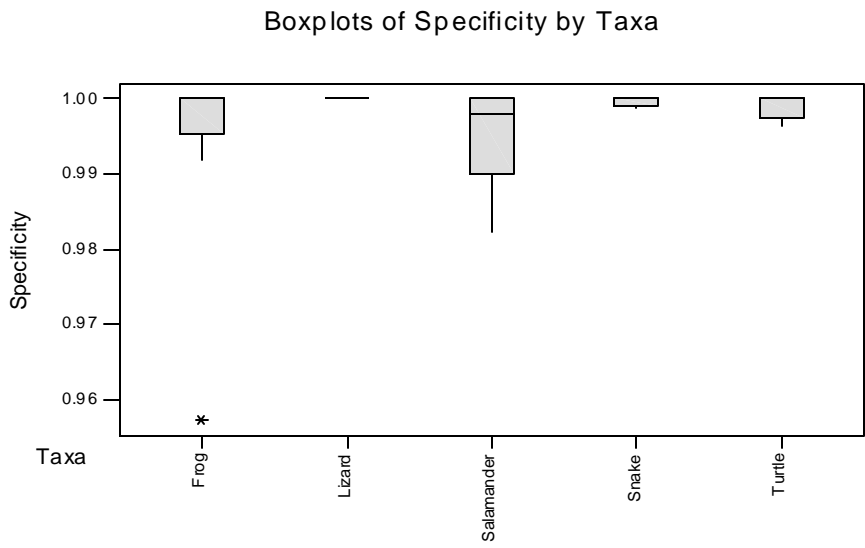
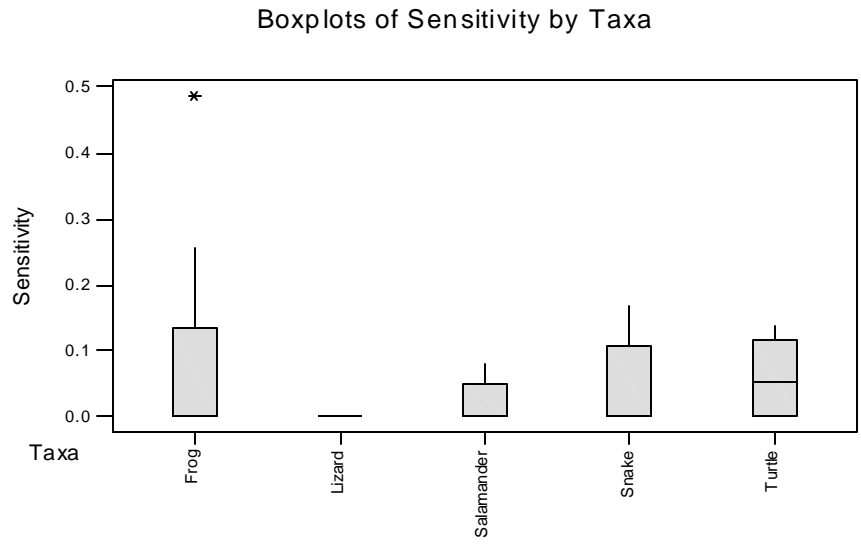


Figure 2. 11.

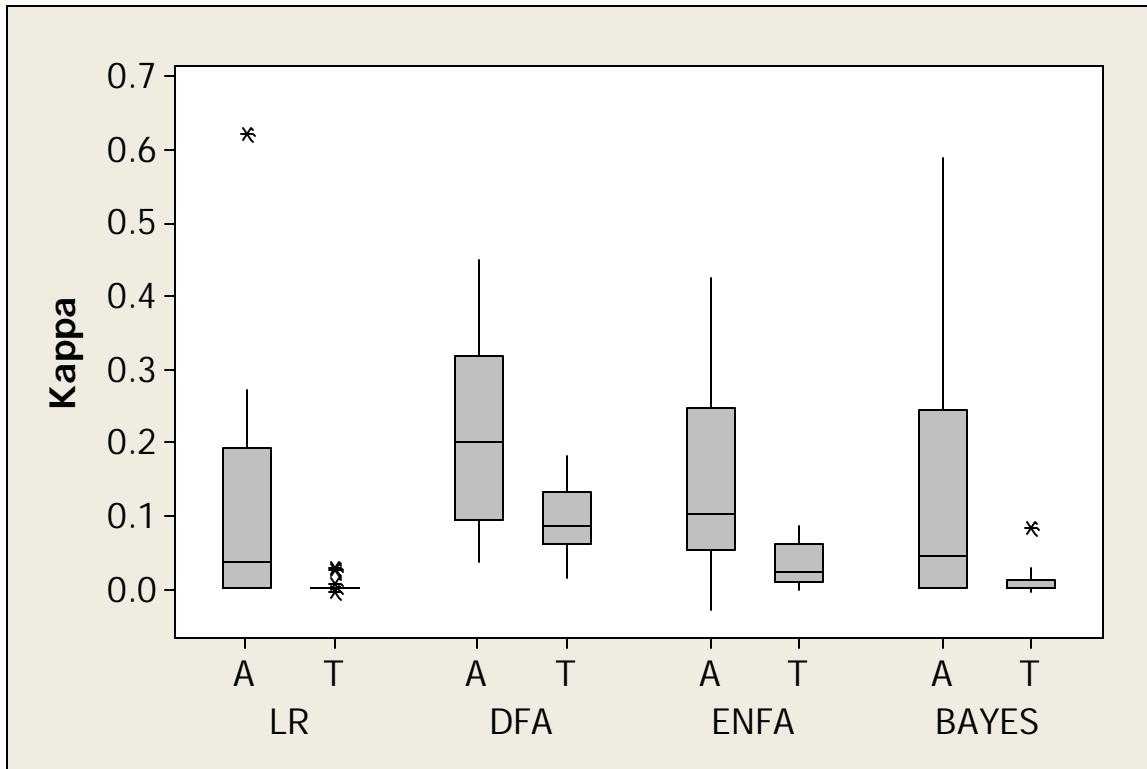


Figure 2. 12.

CHAPTER 3: REDUCING FALSE ABSENCES IN SURVEY DATA: DETECTION PROBABILITIES OF RED-BACKED SALAMANDERS

ABSTRACT

Surveys to determine the presence of a species may be used as the basis for monitoring programs, wildlife-habitat models, or management decisions. Errors in survey data can result from non-detections, where the species is actually present at the site, but not detected in the survey. We designed an experiment to estimate the probability of detection of red-backed salamanders (*Plethodon cinereus*) using different sampling methods and salamander densities. Salamanders were introduced into 124 enclosures, which were randomly assigned to a sampling treatment, either daytime cover-object searches or nighttime visual searches, and a density treatment, either low or high densities of red-backed salamanders. Probability of detection was significantly higher using daytime cover-object searches and in enclosures with higher salamander density. Weather conditions influenced the probability of detection. Increases in temperature correlated with decreases in detection during daytime cover-object sampling. Low humidity limited detection during nighttime visual surveys. Recent rainfall was positively related to detection during nighttime visual surveys. Probability of detection estimates are important for future design and analysis of survey data. Because of the high mean probability of detection, daytime surveys of cover-objects are recommended as the survey method of choice if only a small number of samples can be obtained. However, given their lower variability, nighttime visual surveys are the recommended sampling method when multiple samples can be obtained. Daytime cover-object surveys may be most effective at determining presence or absence at a site, but nighttime visual surveys may be more useful for indices of population trends. Non-detections can be reduced by ensuring large sampling areas for cover-object searches and multiple samples at the same location for nighttime searches.

INTRODUCTION

Surveys for wildlife species are frequently conducted to determine the presence of a species at a particular site. Although observing a species in a survey leads to a presence classification, non-detection of a species may be caused by either the absence of the species at that location or a

failure of the survey to detect the species when it is actually present (Fellers 1997). Survey data forms the basis for many types of models and inferences about the probability of occurrence of the species at a location (Munger et al. 1998, Carroll et al. 1999, Angermeier et al. 2002) and is often used to assess the occurrence and status of threatened and endangered species (Gerrard et al. 2001, Odom et al. 2001). Non-detections in survey data can underestimate occurrence at a location and lead to inaccurate inferences about site occupancy.

Terrestrial salamanders use habitats both above and below the ground, with the majority of individuals believed to be underground at any given time (Taub 1961, Fraser 1976). For this reason, the probability of detection during field surveys is likely to be lower than the probability of occurrence. Information about the probability of detection of terrestrial salamanders using different sampling methods is lacking, but vital for the design and analysis of inventory and monitoring surveys (Bury and Major 1997). Our study develops methodology to estimate the detection rates expected for red-backed salamanders under specific survey methods and illustrates how these rates can be estimated for other species and surveys.

Many factors influence detectability of terrestrial salamanders during a survey, including sampling methods, environmental conditions, and population density. Environmental conditions such as weather, temperature, and climate can influence the detectability to varying degrees (Zug 1993). The ability of some sampling methods to detect a species may be sensitive to weather conditions, although other sampling methods may be unaffected. The population size or density of a species in a location also may correlate with the probability that the species is detected (Nichols 1998). An increasing population density should lead to an increasing probability of detection, but the form of this relationship is unknown and may not be linear. Given the numerous factors that can influence the probability of detection, estimating this probability can be difficult in uncontrolled, observational studies.

In recent years, salamander populations have received greater attention for inventory and monitoring efforts. Changes in site occupancy over time are important to researchers and managers and, given the sensitivity of amphibian populations to environmental conditions, may be an early warning of habitat degradation (Welsh and Droege 2001). In particular, *Plethodon*

spp. have been well-studied (Petranka 1998) because of their terrestrial habitats and potential sensitivity to forest management practices (Petranka 1994). Comparisons of sampling methods occur frequently in the literature, but usually compare the relative change in an index value among the methods, which cannot be translated into a detection probability without accurate knowledge of population size (Adams and Freedman 1999, Crosswhite et al. 1999, Jung et al. 2000, Paszkowski et al. 2002). On populations of known size, we examined the performance of two sampling techniques used for the inventory and monitoring of *Plethodon* salamanders in the eastern United States: nighttime, visual encounter surveys constrained by area and daytime cover-object searches (DeGraaf and Yamasaki 1992, Grant et al. 1992, Heyer et al. 1994, Monti et al. 2000). Estimating the probability of detection of red-backed salamanders using these sampling methods is important for the design of inventory and monitoring survey protocols and analysis of survey data.

The objective of our study was to estimate the probability of detection of red-backed salamanders using daytime cover-object searches and nighttime visual searches constrained by area, under high and low salamander densities. An experimental approach, using known salamander density, permitted rigorous estimation and comparison of the probability of detection. In addition, we examined how weather conditions, the number of sampling occasions, and sample area size influenced the probability of detection of red-backed salamanders when using the two sampling methods. The results of this study can be used for planning future salamander studies to reduce non-detections and as a factor in the analysis of salamander survey data to account for detection probability.

METHODS

Salamanders were housed in enclosures (described below) during the experiment, which served as the experimental units. The enclosures were located in deciduous forest at the Mountain Lake Biological Research Station (MLBS) in Giles County, Virginia. Red-backed salamanders are common at MLBS, with the home range size of adult salamanders estimated to be less than 1 m² (Mathis 1991). We used a 2x2 factorial approach that was further expanded during analysis by the addition of a factor related to time. Experimental enclosures were randomly assigned to

either a daytime cover-object search or nighttime visual search sampling method, and to either a high density or low density of salamanders. The four treatments (high density-daytime samples, low density-daytime samples, high density-nighttime samples, and low density-nighttime samples) each had 31 replicates (total sample size, $n=124$).

Experimental Enclosures

The enclosures were constructed from 1.4 x 0.84 x 0.3 m (area of 1.2 m²) black plastic tubs (MacCourt, Crystal Lake, IL) with 5-cm-diameter holes cut in the bottom and sides. The holes were covered with mesh screening attached to the enclosure tub with silicone caulking. The top of the enclosure was secured with mesh screening caulked to the tub to prevent salamander escape. In enclosures assigned to daytime sampling treatments that needed to be accessed to search cover objects, one corner of the screening was loosened and then secured with clothespins. All enclosures had approximately 10-12 cm of soil and had been used in previous years as salamander enclosures. Prior to the experiment, the soil in each enclosure was sifted by hand to ensure that no salamanders were present. Leaf litter was collected, homogenized in a large tub, and distributed evenly over the soil in each of the enclosures. A split log was placed in each enclosure to serve as a natural cover object.

Salamanders were caught by hand in areas near the enclosures. Captured salamanders were confined separately and transported to the lab, where they were weighed to the nearest 0.1 mg and snout-vent length (SVL) was measured to the nearest 0.1 mm. Individuals were preliminarily sexed in the field by their head shape (Quinn and Graves 1999), and this was confirmed in the lab using a fiber optic light. Only adult male salamanders were selected for the experiment to eliminate confounding of results with age and sex-related differences. Groups of six salamanders were randomly chosen from all of the collected salamanders (verified to be adult and male) and the six salamanders were randomly divided into the four treatment assignments, with 1 salamander in the low density treatments and 2 in the high density treatments. Each treatment was randomly assigned to an enclosure. Although treatments were not initiated simultaneously, there were equal sample sizes of treatments at all times.

Data Collection

Treatments were initiated between May 19, 2001 and June 6, 2001, with sampling beginning 5 - 6 days after salamander introduction (May 24, 2001 – June 12, 2001 depending on the date the treatment was initiated). Sampling of all enclosures was concluded on July 24, 2001. Day samples were made in the afternoon, prior to sunset. Cover-object searches were performed by lifting a corner of the cover object, counting the number of salamanders visible, and returning the cover object to its original location with the least amount of disturbance possible. Night samples were initiated between one-half hour after sunset and midnight. Visual searches were conducted using a headlamp through the mesh screen and the cover object was not disturbed (Heyer 1994).

Sampling was conducted 3 times per week (approximately every other day), but night samples on occasion were collected on consecutive nights to coincide with rain events. Daytime samples were never taken on consecutive days (regardless of rain events) to minimize and maintain a consistent level of disturbance to the cover object microhabitat. For both daytime and nighttime sampling, several samples for all enclosures were excluded, so that all samples included in the analyses were made by the same observer to minimize observer bias (Bury and Major 1997). This resulted in a total of 1250 samples collected for the 62 daytime sampling enclosures, with the number of observations for each enclosure ranging from 13-25 depending on the date the treatment was initiated. A total of 1496 samples were collected for the 62 nighttime sampling enclosures, with the number of observations for each enclosure ranging from 17-29 depending on the date the treatment was initiated. Because the date of introduction varied and influenced the number of sampling nights, a time factor based on the introduction date was included in the probability of detection analyses.

At the conclusion of the experiment, salamanders were recovered from each enclosure and its location was recorded. Salamanders were weighed to the nearest 0.1 mg and measured SVL to the nearest 0.1 mm, then released at their approximate location of initial capture at MLBS.

Weather data including temperature (degrees Celsius), humidity, and rainfall were recorded every half hour at a weather station maintained at MLBS. Weather data were summarized to a

daily (0000 hrs to 2330 hrs) average for temperature and humidity. Rainfall totals were summarized for day samples from 1200 hrs the previous day to 1130 hrs of the sample day. For night samples, rainfall totals were summarized from 0000 hrs to 2330 hrs of the sample day.

Statistical Analyses

For each sample day or night, the number of salamanders observed in the enclosure was recorded. For analysis, only enclosures where all salamanders were recovered at the completion of the experiment were used. Sixteen enclosures were excluded from analysis because the salamanders were not recovered and presumed to have either died or escaped at some point during the experiment. Each enclosure was classified as either a presence (regardless of whether 1 or 2 salamanders were observed) or no detection for each sample. The proportion of times the enclosure was classified as a presence (probability of detection) was calculated for each enclosure. The probability of detection was used for data analysis and estimation of treatment effects using analysis of variance for a 2x2 factorial and calculation of least squared means.

The percent of enclosures classified as present on a given sampling day was calculated to test for weather-related influences on detection probabilities. Step-wise multiple linear regression (SAS, Version 8, Cary, NC) was used with rainfall, humidity, and temperature as dependent variables to determine the relationship between weather and detectability.

A bootstrap analysis was used to examine the influence of the number of sampling occasions and size of the sampling area on the probability of detection. Separate bootstrap analyses were conducted for each density-sampling treatment. For the analysis of the number of sampling occasions, an enclosure was randomly selected. Next, 1 to 15 random samples were drawn with replacement from all of the sampling occasions for that enclosure. If at least 1 salamander was observed during at least one of the randomly selected sampling occasions, the enclosure was classified as present. A new enclosure was then randomly selected with replacement, with this process repeated for 500 replications. This resulted in 500 presence or non-detection classifications for each level of sampling effort from 1 to 15 sampling occasions. The mean

proportion of enclosures correctly classified as present was then calculated for each level of sampling effort.

A similar bootstrapping procedure was used to examine the effect of the size of the sampling area on the probability of detection. For this analysis, 1 sampling occasion was randomly selected. One to 10 enclosures were then randomly selected with replacement from all of the enclosures sampled on that sampling occasion. A presence or absence classification was then determined based on whether or not at least 1 salamander was observed in at least 1 of the enclosures. For example, 1 enclosure represents a small sampling area but the 10 randomly selected enclosures were treated as 1 large sampling area. If at least 1 salamander was seen in 1 of the 10 enclosures, the large sampling area was classified as present. A new sampling occasion was then randomly drawn with replacement, and this process was repeated for 500 replications. The mean proportion of times a sampling area, consisting of 1 to 10 enclosures, was correctly classified as present was then calculated for each area size.

RESULTS

A total of 186 adult male red-backed salamanders were caught by hand and placed in enclosures. Prior to introducing the salamanders into the enclosures, mean salamander mass was 1.12 g (range 0.73 - 1.65 g). Snout-vent length (SVL) averaged 44.0 mm (range 35.9 - 51 mm).

A total of 20 salamanders (11%) were not recovered from 16 enclosures and were assumed to have died or escaped from the enclosure (Table 3. 1). Holes, where the screening had become unfastened from the plastic enclosure, were found on several of the enclosures where salamanders were not recovered. These 16 enclosures were not included in the analysis. This exclusion resulted in a sample size of 108 experimental enclosures for analyses (n = 28 for low density daytime treatment, n = 24 for high density daytime treatment, n = 29 for low density nighttime treatment, and n = 27 for high density nighttime treatment).

The location of the salamander at the conclusion of the experiment showed a consistent pattern for all treatments except the high density, cover-object sampled treatment. For all night

sampling treatments and the low density, cover-object sampled treatment, approximately 70-75% of the salamanders were found underneath the cover object, 18-21% were found in the leaf litter, and 7-11% were recovered from within the soil. The high density, cover-object sampled treatment appeared to vary from this pattern, with a tendency to find fewer salamanders (54%) under cover objects and more salamanders recovered from the leaf litter (27%) and soil (19%).

On average, salamanders in all treatment enclosures gained mass during the experiment (mean mass gain: low density, daytime = 0.19 g, high density, daytime = 0.12 g, low density, nighttime = 0.26 g, high density, nighttime = 0.26 g). Sampling method was a significant factor in predicting mean mass gain ($F = 7.63$, $df = 1,87$, $P = 0.007$). A comparison of least square means determined salamanders in enclosures sampled using nighttime visual surveys gained more mass than those in enclosures sampled using daytime cover-object searches.

Daytime cover-object sampling detected salamanders more frequently than nighttime visual searches ($F = 118.62$, $df = 1,101$, $P < 0.0001$). An examination of least square means shows that daytime cover-object sampled enclosures were classified as present 64.3% of the time, while only 19.8% of the nighttime visual surveys resulted in a correct classification. Although detection probabilities were greater for daytime cover-object searches, variation among replicate enclosures was greater than nighttime visual surveys (Figure 3. 1). Daytime cover object searches resulted in a full range of detection probabilities from 0 (a salamander was never detected in the enclosure) to 1 (at least one salamander was detected on each sampling occasion). Replicate enclosures for nighttime visual surveys showed less variability in the proportion of samples correctly classified as present, but nighttime visual surveys lacked the high detection probabilities seen with cover object searches.

Salamander density affected the probability of detection ($F = 10.94$, $df = 1,101$, $P = 0.0013$). High density enclosures were correctly classified as present more frequently (least square mean = 48.8%) than low density enclosures (least square mean = 35.3%), regardless of sampling method.

No evidence of interaction between sampling method and density was found ($F = 1.06$, $df = 1,101$, $P = 0.3063$). Because salamanders were introduced into enclosures on different dates, a time factor also was examined. The date of introduction did not influence the resulting probability of detection ($F = 1.75$, $df = 3,101$, $P = 0.1626$).

Total rainfall during the previous 24 h showed positive relationships with the percent present in low and high density, nighttime sampled enclosures. An interaction between rainfall and temperature was selected as a predictor variable in the stepwise multiple linear regressions for both low and high density, nighttime sampled enclosures. For low density, nighttime samples, the interaction between total daily rainfall and temperature ($F = 5.58$, $df = 1,27$, $P = 0.0256$) explained 17.1% of the variation in the percent of enclosures classified as present. For high density, nighttime samples, total daily rainfall ($F = 9.49$, $df = 1,27$, $P = 0.0047$) explained slightly more, 26%, of the variation in the percent of enclosures classified as present. Average daily humidity was not selected as a factor in the multiple linear regressions for the nighttime sampling treatments. However, average daily humidity did show evidence of a limiting factor relationship (Figure 3. 2) with both of the treatments.

For the observed range of temperatures, average daily temperature was the only predictor variable selected for inclusion in the stepwise multiple linear regression models for low and high density, cover object sampled treatments. Average daily temperature was negatively related to the percent of enclosures classified as present for low density, cover object sampled treatments ($F = 6.58$, $df = 1,23$, $P = 0.0173$) and explained 22.2% of the variation. For the high-density, cover object sampled treatment, a similar negative relationship was observed with average daily temperature explaining approximately 25.6% of the variation in percent presence ($F = 7.92$, $df = 1,23$, $P = 0.0098$). Rainfall during the previous 24 h, average humidity, and all interactions between predictor variables were not selected for inclusion in either of the multiple linear regressions for low or high density cover object treatments.

We checked to see if nighttime, visual surveys that occurred on wet nights were more effective than our more general results. Given the positive relationship between rainfall in the previous 24 h and the proportion of visually sampled enclosures classified as present, nighttime samples

collected when the total rainfall for the previous 12 h was greater than 0 were selected. A noticeable improvement in the probability of detection by nighttime visual surveys was observed when only samples from rainy nights (Figure 3. 3) were examined, in comparison to all sample nights (Figure 3. 1). Analysis of least squared means showed that the mean probability of detection for nighttime visual surveys (for both low and high density treatments) improved from 0.20 (0.14 - 0.25, 95% CI) to 0.34 (0.27 - 0.40, 95% CI) when surveys were conducted on wet nights. In comparison to cover-object surveys conducted in all weather conditions, no difference in detection probability existed between high density, nighttime visual surveys conducted during wet weather ($x = 0.41$, 95% CI 0.31 - 0.50) and low density, cover-object surveys ($x = 0.56$, 95% CI 0.48 - 0.63).

Sampling effort consists of both the number of sampling occasions and the size of the sampling area. The bootstrap analysis allowed the number of sampling occasions to be varied from 1 to 15, while keeping the size of the sampling area constant. For all density and sampling treatments, increasing the number of sampling occasions led to an increase in the percent of enclosures that were correctly classified (Figure 3. 4). Regardless of the number of sampling occasions, daytime cover object sampling correctly classified a larger proportion of the enclosures than nighttime sampling. The percentage of enclosures correctly predicted decreased substantially after 3 samples and reached an asymptote around 5 samples for cover object sampling treatments. For nighttime sampling treatments, the percentage correctly predicted increased until approximately 8 to 10 samples were taken. After 10 nighttime samples are taken, additional samples only slightly increase the percentage correctly predicted.

A similar bootstrapping procedure was used to vary the size of the sampling area (as measured by the number of enclosures) when only sampling on one occasion. The percent of sampling areas correctly classified increased for all treatments as the size of the sampling area increased (Figure 3. 5). The percentage of cover object sampled treatments correctly classified was close to maximum when the sampling area was at least equal to the size of three enclosures. For night sampled treatments, the percent of sampling areas correctly classified was lower than that of cover object sampled treatments regardless of the size of the sampling effort. Increasing the size

of the sampling area did not increase the percentage of correctly classified areas as greatly as increasing the number of sampling occasions for nighttime samples.

DISCUSSION

Sampling Method Effects on Detection Probability

In a controlled, experimental setting, daytime cover-object searches were more effective at detecting red-backed salamanders than nighttime visual searches. However, in the high-density enclosures, approximately 34% of the cover-object searches resulted in no detections. For nighttime visual surveys of these enclosures, no detections should be expected for 81% of the surveys. Adjusting the protocol to include only nighttime surveys on rainy nights improves the probability of detection, but still resulted in no detection for 66% of the surveys. Typical visual survey protocols involve sampling only on ideal, wet nights to maximize detection probabilities. However, under ideal weather conditions, no detections were common occurrences and should be incorporated into the calculation of probabilities of occurrence for both sampling methods that were examined.

Night surveys have previously been described as more variable than daytime cover-object searches (Petranka et al. 1994), but our study found that mean detectability over multiple nights was less variable among replicate enclosures. Using daytime cover-object searches, results ranged from replicate enclosures where a salamander was detected on every sampling occasion, to replicate enclosures where a salamander was never detected on a single occasion. A consistent probability of detection with low variability is a desired attribute of a sampling method. A low bias in the probability of detection (fewer no detections) is also a desired attribute of a sampling method. For red-backed salamanders, researchers apparently have the choice between a survey method with a lower bias and higher variability (daytime cover-object searches) or a method with a higher bias and lower variability (nighttime visual searches).

If the survey objective is to inventory the species, the daytime cover-object search method is more effective. If researchers are attempting to quantify the probability of occurrence,

particularly for comparison across time or space, the nighttime visual search method would likely be preferred due to the lower variability.

Sampling Method Effects on Salamander Condition

Repeated sampling of cover objects appeared to have an effect on salamander condition and distribution, especially in the high density enclosures. Individuals in the high-density, cover-object sampling treatment gained significantly less mass than those in the low-density, visually sampled enclosures. The greatest number of dead or missing salamanders occurred in the high-density, cover-object search treatment and distribution of the salamanders recovered from these enclosures varied from the other treatments. Fewer salamanders were found under cover objects in the high density enclosures where cover objects had been repeatedly disturbed. Repeated sampling of cover objects may alter microhabitat conditions under the cover object and place additional stress on salamanders in comparison to nighttime visual surveys. The repeated disturbance of cover object sampling may be a factor in the increased variability of the detection probability for daytime cover object sampled treatments.

Salamander Density Effects on Detection Probability

Enclosures with a higher density of salamanders were more likely to be correctly classified as present than were lower density enclosures. However, the sampling method had a larger effect on the probability of detection than did the salamander density. When the survey goal is to inventory the species, fewer samples will be necessary when salamander density is high, regardless of the survey method. In surveys designed to monitor population abundance where density is unknown, more samples will be necessary to ensure that the probability of detection is high. In our experiment, because of the small size of the enclosures, only two density treatments (1 salamander vs. 2 salamanders) were used. At higher densities, the influence of salamander density on the probability of detection may differ. However, given the territorial nature of red-backed salamanders at cover objects (Jaeger et al. 1982, Smith and Pough 1994), a difference between the two density treatments was expected if one of the salamanders was excluded from the cover object.

Weather Influences on Salamander Detection

Weather conditions and moisture previously have been found to influence the activity levels of salamanders (Grover 1998). Due to their lungless nature, red-backed salamanders are physiologically constrained to maintain moist skin (Heatwole 1960). Our results confirm that red-backed salamanders are more likely to be detected on wet, cool nights during visual surveys. Humidity also seems to be a limiting factor. When average daily humidity was low, the proportion of correctly classified enclosures was always low. Limiting factor relationships frequently occur in ecology and linear regression cannot appropriately quantify the relationship (Thomson et al. 1996). Given the constraint of maintaining moisture for respiration, low humidity may limit the surface activity and therefore be responsible for the low proportion of times a red-backed salamander was detected. Higher humidity may allow for more surface activity (Heatwole 1960) and greater proportion of salamander detections, but other factors may limit the activity of salamanders on these days. However, over half of the salamanders were either not active on the surface or not detected by the observer on nights with ideal weather conditions.

The success of daytime cover-object searches was negatively correlated with mean daily temperature during May - July. Given the observed range of temperatures from 11 to 21 degrees Celsius, fewer of the enclosures were correctly classified as present using cover-object searches at higher temperatures. Because of the correlative nature of the analysis, cause and effect cannot be determined. At higher temperatures, salamanders may select microhabitats other than underneath the available cover object. However, the same negative correlation would be observed if the probability of detecting a salamander under a cover object decreased with repeated disturbance of the cover object. Because of the increasing trend in mean daily temperature over the time period of the experiment, lower temperatures and higher detection probabilities would occur together at the beginning of the experiment without implying causation.

Sampling protocols should be designed to maximize the probability of detection and should account for weather conditions. Conducting nighttime visual surveys during cool, rainy and

humid weather or daytime surveys during cool weather is likely to improve the probability of detection of red-backed salamanders. If weather conditions are not included in the sampling protocol, the differences in detection probability should be recognized and accounted for when comparing samples that were taken under different weather conditions.

Impact of Sampling Effort on Detection Probabilities

Quantifying sampling effort involves specifying both the number of sampling occasions and the size of the sampling area. Examining these two aspects of sampling effort separately, shows that cover-object sampling is more sensitive to the size of the sampling area. As the size of the sampling area is increased, the percent of correctly classified areas using cover-object searches quickly reaches an asymptote near 100. Provided that at least 3 enclosures (3.6-m²) in the high density treatment and 5 (6-m²) enclosures in the low density treatment are sampled on any given occasion, it is highly likely that at least one salamander will be observed and the area classified as a presence.

Conversely, nighttime sampling is more influenced by the number of sampling occasions than by the size of the sampling area. By maintaining the size of the sampling area constant at a size equal to 1 enclosure, increasing the number of sampling occasions to 8 to 10 led to the majority of nighttime sampled enclosures being correctly predicted as present. Adding sampling occasions increases the percent of enclosures correctly classified by nighttime sampling at a greater rate and to a higher asymptote than adding to the size of the sampling area. Given that salamander activity at night is influenced by weather conditions, salamander activity may be similar among enclosures on a given night. Because of this correlation, adding sampling occasions from different dates with varied weather conditions was more effective for detecting salamanders than increasing the size of the sampled area on one particular night.

Overall, the percent of sampling areas where at least one salamander was detected increased as the number of sampling occasions and the size of the sampling area increased. In general, more samples and larger sample area are desired for both inventory and monitoring of red-backed salamanders. However, the classification accuracy for all density and sampling treatments was

asymptotic, where additional sampling occasions and larger areas did not significantly increase the accuracy and would not be cost-effective. Because survey design represents a trade-off between accuracy and limited resources to conduct the survey, determining the level of sampling effort that provides the maximum information gain for the least amount of survey effort is valuable. For cover object sampling conducted during one sampling occasion, a sampling area with a size equivalent to 3 to 5 enclosures (3.6 to 6-m²) provided a high accuracy predictions. For nighttime sampling, a sample size of 8 to 10 occasions of one sampling area (1.2-m²) represented a trade-off providing the desired high percent of correctly classified enclosures and least amount of effort.

Experiment Assumptions

The experimental approach used to examine the questions of detection probability of red-backed salamanders provides rigor for the estimates of treatment effects. However, many assumptions must be made in the experimental design and some natural conditions and variability are sacrificed to achieve the increased experimental rigor. Application of the experimental results to natural conditions is dependent on whether salamanders exhibit typical behavior in the enclosures. The enclosures were successful in supporting the survival needs of red-backed salamanders for an extended period of time. Although 11% of the salamanders were not recovered from the enclosures, the recovered salamanders on average gained mass and were in good condition. In addition, salamanders in all treatments were recovered from all locations within the enclosure with no evidence of avoidance of any portion of the enclosure.

Probability of detection of red-backed salamanders under natural conditions may vary from the experiment due to differences in salamander and/or cover-object density. The small size of the enclosures limited the number of salamanders that could be enclosed together, resulting in only two density treatments. Typical home ranges at MLBS are small (less than 1 m², Mathis 1991), and the low density treatment resulted in 1.2 m² per salamander in comparison to the high density treatment of 0.6 m² per salamander. In reality, a range of densities is possible and it may not be accurate to extrapolate the relationship of probability of detection to densities greater or lesser than those included in the experiment. The availability of cover objects may also play a

role in the probability of detection of red-backed salamanders (Grover 1998). In the low density treatments, no competition for the single cover object was possible, but competition between the two salamanders in the high density treatment was possible for the single cover object. The ratio of salamanders to cover objects may influence the probability of detection, with further experimentation required to quantify the relationship. In addition, the probability of detection of female or juvenile salamanders may differ from the adult, male salamanders used in the experiment because of possible differences in activity patterns, microhabitat selection, or physiological constraints.

Recommendations for Sampling

Although the probability of detection of red-backed salamanders under natural conditions may vary from the experimental estimates, the experimental estimates provide an initial estimate that is better than assuming 100% detectability. Based on the estimates of detectability, cover-object searches appear to be the most effective method for determining presence of red-backed salamanders, if only a small, limited number of samples can be taken. If few samples can be obtained, the chances of obtaining a no detection when using nighttime visual searches is significantly higher than cover-object searches. However, if multiple samples (particularly over a short period of time) can be obtained, nighttime visual surveys would be preferable because of the consistent probability of detection among replicate sample areas and the lesser impacts on the condition and distribution of the salamanders. Nighttime visual surveys are less disturbing to the cover objects and given a large enough collection of repeat samples, the problem of incorrectly classifying the sample location will be low. Sampling protocols should also incorporate weather conditions that are correlated with the probability of detection, in an effort to maximize this probability. Sampling areas for cover-object searches should be at least 3 to 5 times larger than the size of the enclosures in this experiment (at least 3.2 to 6-m²) or contain at least 3 to 5 cover objects. Sampling areas for nighttime surveys can be smaller than those for cover-object searches, but multiple samples should be obtained for the sampling area. The exact number of repeat samples to take at a sampling area depends on the desired accuracy of presence classification.

Understanding the probability of detection of red-backed salamanders and how it varies based on sampling method, weather, and density is important both for future study design and analysis of data that has already been collected. Using the estimated probability of detection and number of repeat samples at a location, the probability that a no detection is actually a presence location can be estimated for existing data collected using one of the sampling methods. Areas of no detection can be ranked based on their likelihood of actually being an absence, and areas can be prioritized for future sampling or excluded from further analysis or modeling. Future comparisons of how the experimental estimates correlate with those obtained under natural conditions and estimates for additional amphibian species would be valuable.

Management Implications

It is clearly not appropriate when using survey data to assume that the observed probability of detection is equal to the probability of occurrence. In most cases, it is likely that a no detection location is actually a presence that was not detected. No detections lead to underestimating the probability of occurrence as well as raising its associated variance. This can have important implications for wildlife conservation and management. Estimating the probability of detection of a species using a particular sampling method is necessary for effective study design and meaningful data analysis. This experiment provides estimates of detection probabilities for one species, but also serves as an example of how small experiments can be used to estimate these probabilities under controlled conditions.

LITERATURE CITED

- Adams, J. D., and B. Freedman. 1999. Comparative catch efficiency of amphibian sampling methods in terrestrial habitats in southern New Brunswick. *Canadian Field-Naturalist* 113:493-496.
- Angermeier, P. L., K. L. Krueger, and C. A. Dolloff. 2002. Discontinuity in stream-fish distributions: Implications for assessing and predicting species occurrence. Pages 519-527 in J. M. Scott, P. J. Heglund, M. L. Morrison, J. B. Haufler, M. G. Raphael, W. A. Wall, and F. B. Samson, editors. *Predicting species occurrences: issues of accuracy and scale*. Island Press, Covelo, California, USA.
- Bury, R. B., and D. J. Major. 1997. Integrated sampling for amphibian communities in montane

- habitats. Northwest Fauna 4:75-82.
- Carroll, C., W. J. Zielinski, and R. F. Noss. 1999. Using presence-absence data to build and test spatial habitat models for the fisher in the Klamath Region, USA. Conservation Biology 13:1344-1359.
- Crosswhite, D. L., S. F. Fox, and R. E. Hill. 1999. Comparison of methods for monitoring reptiles and amphibians in upland forests of the Ouachita Mountains. Proceedings of the Oklahoma Academy of Science 79:45-50.
- DeGraaf, R. M., and M. Yamasaki. 1992. A nondestructive technique to monitor the relative abundance of terrestrial salamanders. Wildlife Society Bulletin 20:260-264.
- Fellers, G. M. 1997. Design of amphibian surveys. Northwest Fauna 4:23-34.
- Fraser, D. F. 1976. Empirical evaluation of the hypothesis of food competition in salamanders of the genus *Plethodon*. Ecology 57:459-471.
- Friend, G. R. 1984. Relative efficiency of two pitfall-drift fence systems for sampling small vertebrates. Australian Zoologist 22:423-433.
- Gerrard, R., P. Stine, R. Church, and M. Gilpin. 2001. Habitat evaluation using GIS - a case study applied to the San Joaquin kit fox. Landscape and Urban Planning 52:239-255.
- Grant, B. W., A. D. Tucker, J. E. Lovich, A. M. Mills, P. M. Dixon, and J. W. Gibbons. 1992. The use of coverboards in estimating patterns of reptile and amphibian biodiversity. Pages 379-403 in D. R. McCullough, and R. H. Barrett, editors. Wildlife 2001: Populations. Elsevier Applied Science, London, United Kingdom.
- Grover, M. C. 1998. Influence of cover and moisture on abundances of the terrestrial salamanders *Plethodon cinereus* and *Plethodon glutinosus*. Journal of Herpetology 32:489-497.
- Heatwole, H. 1960. Burrowing ability and behavioral responses to desiccation of the salamander, *Plethodon cinereus*. Ecology 41:661-668.
- Heyer, W. R., M. A. Donnelly, R. W. McDiarmid, L. C. Hayek, and M. S. Foster, editors. 1994. Measuring and Monitoring Biological Diversity: Standard Methods for Amphibians. Smithsonian Institution, Washington D.C., USA.
- Jaeger, R. G., D. Kalvarsky, and N. Shimizu. 1982. Territorial behaviour of the red-backed salamander: expulsion of intruders. Animal Behaviour 30:490-496.
- Jung, R. E., S. Droege, J. R. Sauer, and R. B. Landy. 2000. Evaluation of terrestrial and streamside salamander monitoring techniques at Shenandoah National Park. Environmental Monitoring and Assessment 63:65-79.

- Mathis, A. 1991. Territories of male and female terrestrial salamanders: costs, benefits, and intersexual spatial associations. *Oecologia* 86:433-440.
- Monti, L., M. Hunter, Jr., and J. Whitham. 2000. An evaluation of the artificial cover object (ACO) method for monitoring populations of the redback salamander, *Plethodon cinereus*. *Journal of Herpetology* 34:624-629.
- Munger, J. C., M. Gerber, K. Madrid, M.-A. Carroll, W. Petersen, and L. Herberger. 1998. U.S. National Wetlands Inventory classifications as predictors of the occurrence of Columbia spotted frogs (*Rana luteiventris*) and Pacific treefrogs (*Hyla regilla*). *Conservation Biology* 12:320-330.
- Nichols, J. D., T. Boulinier, J. E. Hines, K. H. Pollock, and J. R. Sauer. 1998. Estimating rates of local species extinction, colonization, and turnover in animal communities. *Ecological Applications* 8:1213-1225.
- Odom, R. H., W. M. Ford, J. W. Edwards, C. W. Stihler, and J. M. Menzel. 2001. Developing a habitat model for the endangered Virginia northern flying squirrel (*Glaucomys sabrinus fuscus*) in the Allegheny mountains of west Virginia. *Biological Conservation* 99:245-252.
- Paszkowski, C. A., G. Scrimgeour, B. A. Gringas, and S. Kendall. 2002. A comparison of techniques for assessing amphibian assemblages on streams in the western boreal forest. *Canadian Field-Naturalist* 116:116-119.
- Petranka, J. W. 1994. Response to impact of timber harvesting on salamanders. *Conservation Biology* 8:302-304.
- Petranka, J. W. 1998. Salamanders of the United States and Canada. Smithsonian Institution Press, Washington D.C., USA.
- Quinn, V. S., and B. M. Graves. 1999. A technique for sexing red-backed salamanders (*Plethodon cinereus*). *Herpetological Review* 30:32.
- Smith, E. M., and F. H. Pough. 1994. Intergeneric aggression among salamanders. *Journal of Herpetology* 29:198-210.
- Taub, F. B. 1961. The distribution of red-backed salamanders, *Plethodon c. cinereus*, within the soil. *Ecology* 42:681-698.
- Thomson, J. D., G. Weiblen, B. A. Thomson, S. Alfaro, and P. Legendre. 1996. Untangling multiple factors in spatial distributions: Lilies, gophers, and rocks. *Ecology* 77:1698-1715.
- Welsh, Jr., H. H., and S. Droege. 2001. A case for using Plethodontid salamanders for monitoring biodiversity and ecosystem integrity of North American forests.

Conservation Biology 15:558-569.

Zug, G. R. 1993. Herpetology: An introductory biology of amphibians and reptiles. Academic Press, San Diego, California, USA.

Table 3. 1. Number of red-backed salamanders, by treatment group, unable to be located at the end of the experiment in July 2001. These salamanders are assumed to have either died or escaped. The percentages are based on the total number of salamanders, however high density treatments contained twice the number of salamanders than low density treatments.

Salamander Density	Sampling Method	No. Salamanders Not Recovered	% Salamanders Not Recovered
Low	Cover Object	3	1.6 %
High	Cover Object	10	5.4 %
Low	Visual	2	1.1 %
High	Visual	5	2.7 %
	Total	20	10.8 %

FIGURE LEGENDS

Figure 3. 1. Mean detection probability of red-backed salamanders for each treatment at Mountain Lake Biological Research Station. A = Low density, daytime samples, B = High density, daytime samples, C = Low density, nighttime samples, and D = High density, nighttime samples.

Figure 3. 2. Relationship between mean daily humidity and the percent of enclosures classified as present for the high density (HN) and low density (LN) nighttime visual search treatments during May – July, 2001.

Figure 3. 3. Mean probability of detection for low density, nighttime (A) and high density, nighttime (B) sampled enclosures calculated using only wet sample nights between May – July, 2001.

Figure 3. 4. Percent of enclosures correctly classified as present for each treatment when varying the number of sampling occasions using a bootstrap analysis with 500 replications. LD = low density, daytime cover object search, HD = high density, daytime cover object search, LN = low density, nighttime visual search, HN = high density, nighttime visual search.

Figure 3. 5. Percent of sampling areas, consisting of 1 – 10 enclosures, correctly predicted when the size of the sampling area was varied using a bootstrap analysis with 500 replications. LD = low density, daytime cover object search, HD = high density, daytime cover object search, LN = low density, nighttime visual search, HN = high density, nighttime visual search.

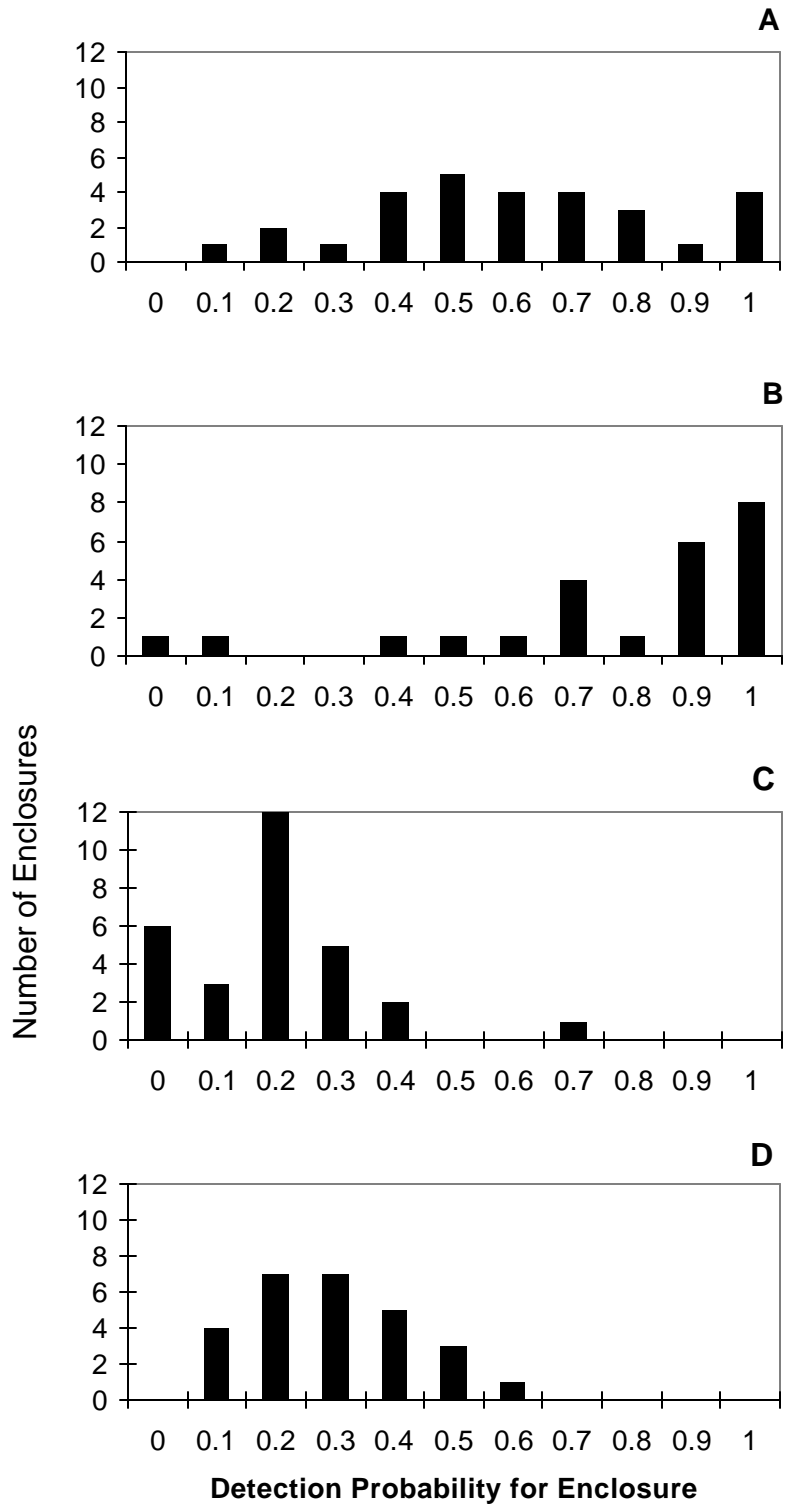


Figure 3. 1.

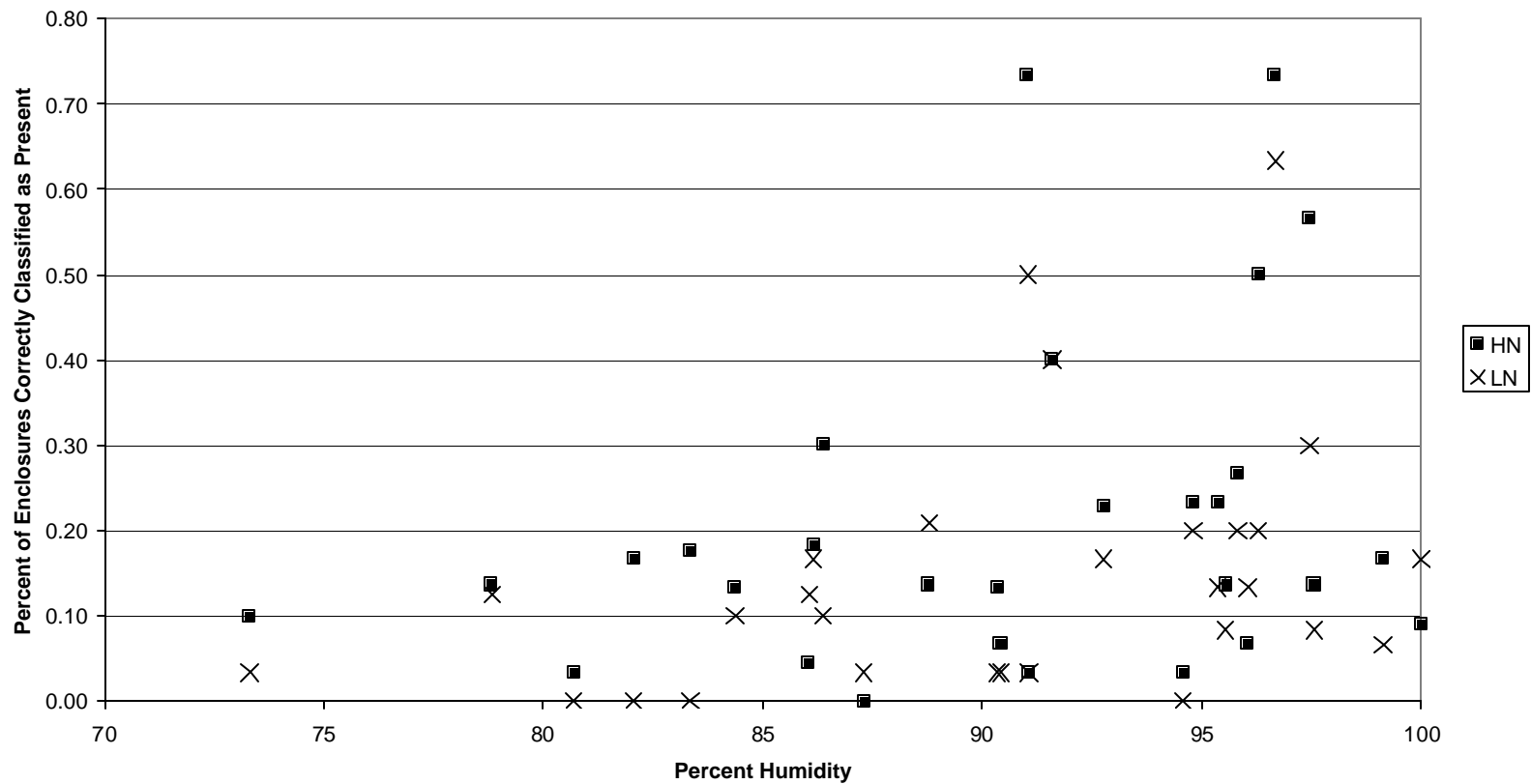


Figure 3. 2.

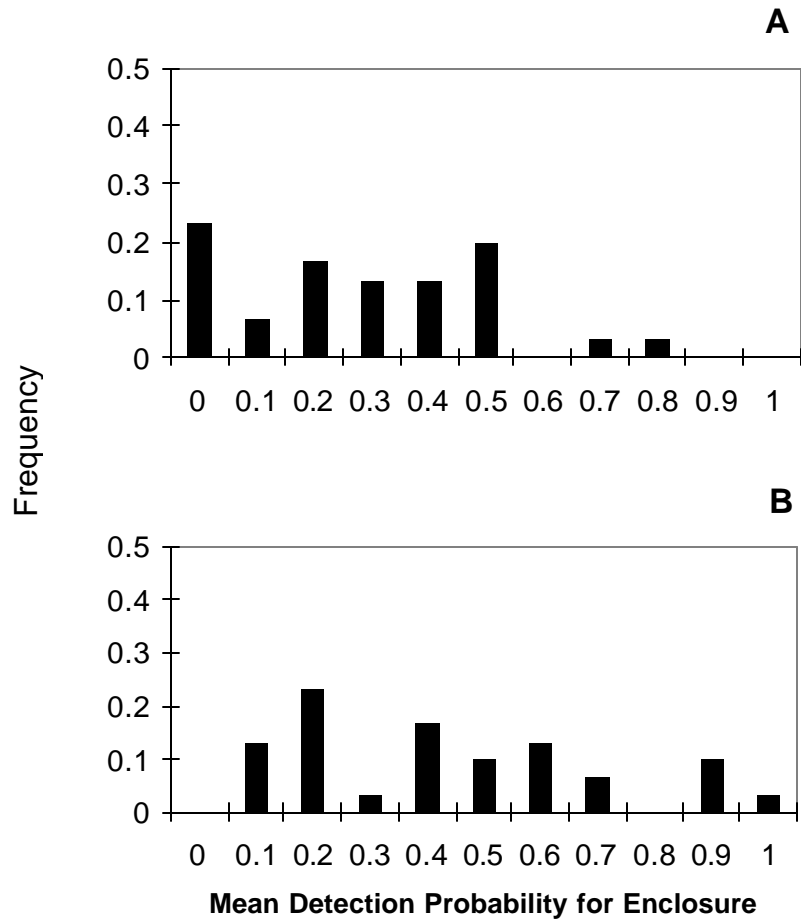


Figure 3. 3.

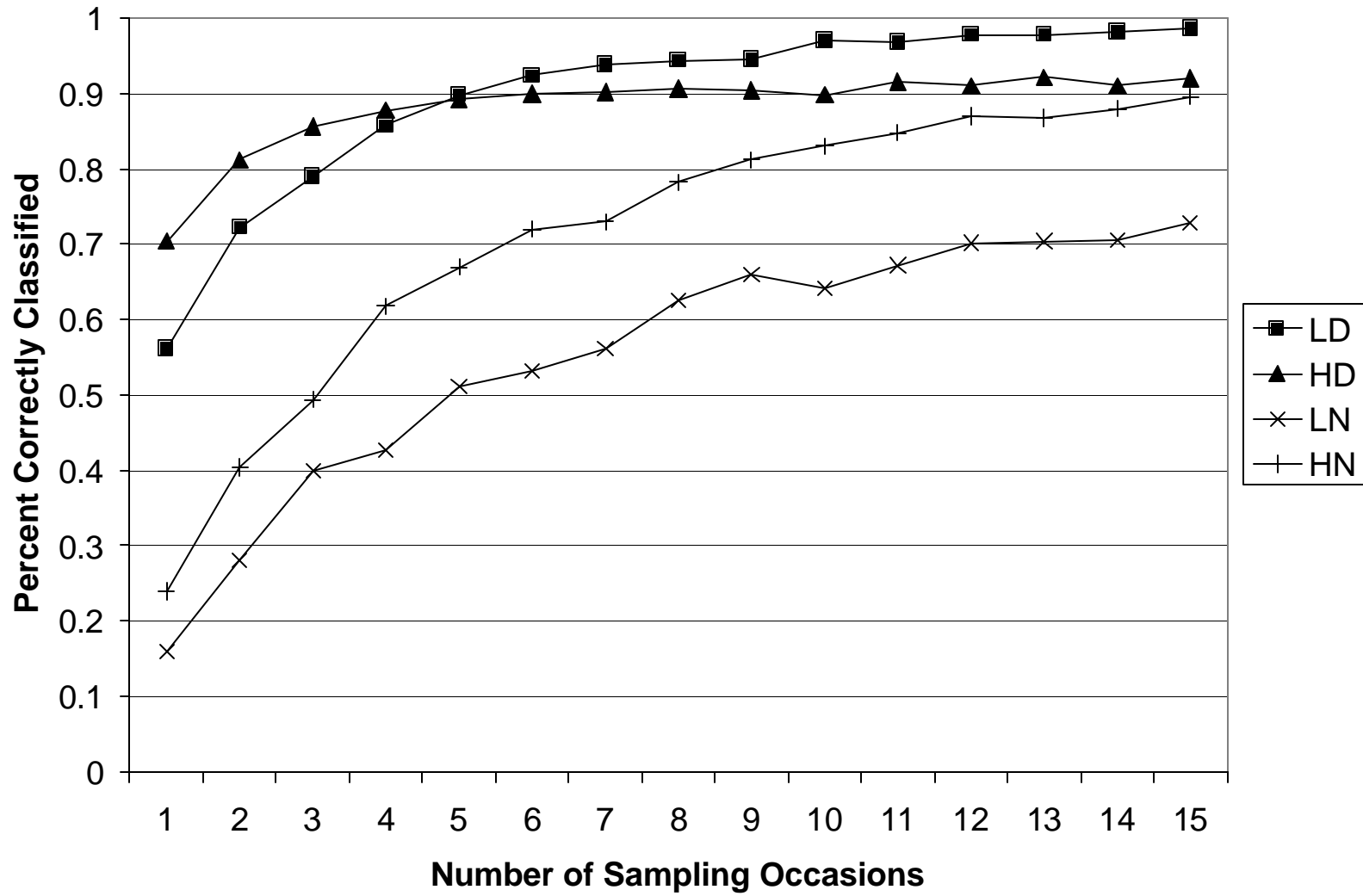


Figure 3. 4.

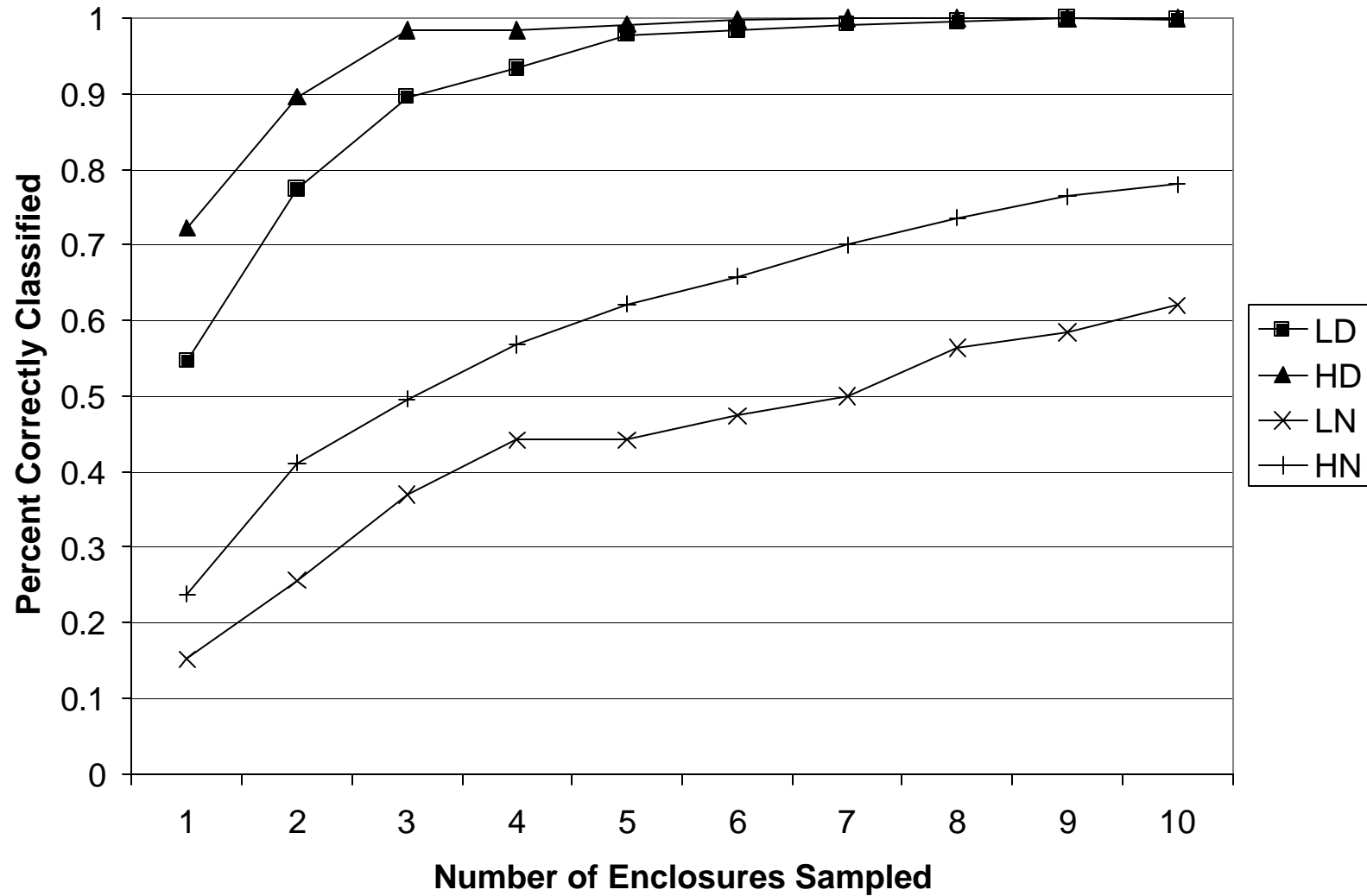


Figure 3. 5.

CHAPTER 4: EFFECTS OF SPECIES DETECTABILITY AND SAMPLE SIZE ON PREDICTING HABITAT SUITABILITY: AN ANALYSIS OF BAYESIAN AND ECOLOGICAL NICHE FACTOR ANALYSIS MODELS USING SIMULATED DATA

ABSTRACT

Survey data consisting of presence and absence locations is often unreliable because of difficulty in classifying nondetection locations as absences. When the probability of detecting the species during a survey is low, locations where the species is not detected may be incorrectly classified as absence locations (false absences). The robustness of habitat suitability modeling methods that use presence/absence data to varying levels of species' detectability has yet to be examined. Additionally, methods of habitat suitability modeling that rely only on presence data have been used with increasing frequency. In this study, I examine which type of model (presence/absence or presence-only modeling method) performs better under varying conditions of species detectability and sample size. Using a real landscape, I generated a species distribution so the true habitat suitability and presence/absence of the species was known without error. The survey process was replicated through a random selection of $n = 50, 100, 200, \text{ or } 500$ locations and the survey data obtained are dependent on both the true presence or absence of the species and the probability of detecting the species in the survey. Four probability of detection distributions were examined (100%, 80%, 60%, and 40% with a $SD = 0.05$). Bayesian logistic regression (presence/absence method) and ecological niche factor analysis (ENFA, presence-only method) models were constructed with each of the simulated surveys and were compared in terms of their ability to accurately predict probability of occurrence, presence locations and absence locations. At low levels of species' detectability, the mean of the environmental predictor variables measured at the absence locations during the survey approaches the mean for the presence locations; however, the means are more separated at high levels of detectability. Model performance was more sensitive to the probability of detection than to sample size. When species detection is $\geq 80\%$, Bayesian logistic regression predicts habitat suitability and presence locations more accurately than ENFA. When detection was $\leq 60\%$, ENFA predicted presence locations more accurately. ENFA models predicted habitat suitability equally as well as Bayesian logistic regression when the probability of detection was 40%. Both methods appear to

be robust to sample sizes > 50 locations. The results have implications for designing surveys and provide guidance for the choice of modeling method for analysis of existing data.

INTRODUCTION

Species distribution maps are important tools for wildlife research, management, and conservation. Developing maps of species distributions across a landscape is often done using habitat suitability modeling because of the expense and length of time required for extensive field surveys. Models that combine multivariate statistical techniques with geographical information systems (GIS) are frequently used to predict a probability of occurrence at a location based on environmental attributes (Guisan and Zimmerman 2000, Hirzel et al. 2002). This allows prediction of species distributions over a large landscape area in a relatively short time-frame, conducive for management decisions. For most habitat suitability models, information about habitat relationships is extracted from a data set of species presence/absence at a set of locations. The majority of multivariate statistical techniques require binary data, i.e. presence and absence, for habitat suitability modeling (Franklin 1998, Guisan and Zimmerman 2000). It is important to have a sufficient sample of accurate, representative presence/absence data because this data set is the foundation of the habitat suitability model.

However, absence data are typically difficult to obtain reliably. A location may be incorrectly classified as an absence during a survey because the species was not detected, even though it was present at the location (Nicholson and Barry 1995). A false absence (McArdle 1990, Solow 1993), i.e. not detecting a species at a location where it occurs, can occur for many reasons. Some species are particularly cryptic and secretive, making them difficult to detect (Hirzel et al. 2002). Sampling methods may be inappropriate or have varying effectiveness for detecting the species (Angermeier et al. 2002). For example, red-backed salamanders (*Plethodon cinereus*) were detected more often (64% of samples) using cover-object searches than with nighttime visual encounter searches (19% of samples) in enclosures where the salamanders were known to be present (Chapter 3, Williams 2003). Insufficient sampling effort at a location can also result in a false absence. A large number of repeat visits or large sampling area are often needed to determine absence of a species with acceptable accuracy (Angermeier and Smogor 1995).

Understanding the probability of detection of the species is necessary for determining the reliability of absence data. If the probability of detection of a species is high, given the sampling methods, the resulting absence data should reflect true absences. However, if the probability of detection is low, the absence data is less reliable and likely contains many false absences. False absences obscure the relationships between species occurrence and habitat characteristics (Hirzel et al. 2001), leading to inaccuracies in the resulting habitat suitability models.

Accurate presence/absence data sets are also needed to assess the accuracy of habitat suitability models. Assessments of model error are dependent on qualities of the data set used for testing (Boone and Krohn 1999). Judging the model results against a data set with false absences provides an unreliable estimate of model accuracy. For example, the model may correctly predict a location as a presence, but the location is incorrectly classified as an absence in the data set because the species was not detected. In this example, model accuracy would be underestimated.

Recently, methods of habitat suitability modeling based on presence-only data, such as ecological niche factor analysis (ENFA, Hirzel et al. 2002), environmental envelopes (Kadmon et al. 2003), genetic algorithms (Anderson et al. 2003), and others (Dettmers and Bart 1999), have been developed for situations where absence data is unavailable or unreliable. Presence-only data is more difficult to model and is not suited for many of the commonly-used statistical modeling methods. Several drawbacks exist to using presence-only data, with the most obvious disadvantage being the lack of accurate absence locations. There is often an unknown sampling bias in most presence-only data sets because the data are collected in an ad-hoc, opportunistic, or non-systematic design (Zaniewski et al. 2002). Often, the location of the samples is correlated with other factors related to convenience of sampling or particular habitat characteristics (Kadmon et al. 2003). Another drawback to presence-only data sets is a potential sampling bias among species. For example, herbarium data sets typically have a disproportionately greater sample size of rare species in comparison to common species (Zaniewski et al. 2002). Despite the drawbacks to modeling with presence-only data, these modeling methods are becoming more widely used because of the abundance of presence-only data (Reutter et al. 2003, Peterson and Robins 2003).

However, the choice of whether to use a presence-only or a presence/absence modeling technique is not necessarily straightforward. For data sets where absence data are not available, selecting random locations from the landscape has been used to represent absence data (Zaniewski et al. 2002, Palma et al. 1999). In comparisons of presence/absence (generalized linear models and generalized additive models) and presence-only (ENFA) modeling methods, presence/absence methods produced slightly better results when the absence data truly reflected low habitat quality (Hirzel et al. 2001, Zaniewski et al. 2002), but performed poorer when the absence data did not correspond with low habitat quality (Hirzel et al. 2001). Presence-only methods appeared to be robust to both data quality and quantity (Hirzel et al. 2001). However, the quality of the absence data is often unknown in many presence/absence data sets. The question of how reliable the absence data must be to gain advantage from using presence/absence modeling methods has yet to be investigated. Because the reliability of the absence data is dependent on the probability of detection of the species, understanding the impact of the probability of detection on model results can lead to guidelines for model selection.

The goal of this paper was to test how modeling methods based on presence/absence data and presence-only data perform when species detectability varies. Bayesian logistic regression was selected as the presence/absence data method because of the increased prevalence of Bayesian statistics in the wildlife field and several desirable characteristics of Bayesian models. Bayesian models can be updated when new presence/absence data becomes available. In addition, the models predict parameter distributions, instead of point estimates, which reflect the level of uncertainty in the parameter estimate. Ecological niche factor analysis was selected as the presence-only modeling method. Previous comparisons of ENFA and other multivariate techniques have been conducted, but have not included detection probability. Also, both ENFA and Bayesian logistic regression were methods used in Chapter 2 to predict the distribution of amphibian and reptile species.

The study had three objectives. The first objective was to determine the relationship between probability of detection and predictive accuracy of Bayesian logistic regression models. ENFA methods are not directly affected by probability of detection because only presence data is used,

excluding potential absence and false absence data. Secondly, I evaluated which modeling method (Bayesian logistic regression or ENFA) performed better under differing scenarios of probability of detection. The third objective was to evaluate the relationship between sample size and model accuracy for Bayesian logistic regression and ENFA models.

To test the influence of probability of detection and sample sizes on the resulting habitat suitability model, I created simulated data. Using a real landscape, I created a “virtual species” distribution (Hirzel et al. 2001) and replicated the field survey process (“virtual ecologist,” Tyre et al. 2001). Creation of simulated data allowed the true habitat suitability and presence/absence at a location to be known with certainty, providing a straightforward assessment of model accuracy. When using simulated data sets, the input data for the model can be controlled and the model results can then be tested against the known species distribution. In this paper, simulation of the field survey process allowed the habitat suitability modeling process to be tested under different probability of detection scenarios and with more samples and replications than could be collected in the field.

METHODS

Landscape Data

Field data used for the creation of the Bayesian logistic regression model came from a 7.53 x 7.53 km area within Fort A.P. Hill Military Reservation, Virginia. Fort A.P. Hill consists of 30,329 ha in Caroline and Essex Counties in eastern Virginia. The military base is approximately 80% forested, with elevation ranging from 6.1 – 76 m. Numerous wetlands and bodies of water, including vernal pools, streams, beaver ponds and impoundments, exist within the military base (Mitchell and McNulty 1999). Available environmental data used for model creation included a digital elevation model (DEM), a wetlands map from the National Wetlands Inventory, a land cover map developed at Fort A.P. Hill, and a VA Gap Analysis land cover map derived from satellite imagery. Four environmental variables were selected and mapped (Figure 4. 1) across the entire landscape: distance to the nearest road, distance to the nearest wetland or water, percentage of forest in the surrounding 300 x 300-m area, and elevation. Further details

on these variables and the source of the data can be found in Chapter 2. Maps were created using tools for IDRISI included in the Biomapper 2.1 software.

Generating the Distribution of a Virtual Species

To simulate a species distribution across the landscape, a “virtual species” (Hirzel et al. 2001) was created with its occurrence dependent on the four environmental variables. The 7.53 x 7.53-km area was divided into a grid of 251 x 251 cells, with a resulting resolution of 30-m cell size. The virtual species was created by simulating an ecological niche (Hutchinson 1957) using the methods developed by Hirzel et al. (2001). Functions relating habitat suitability for the species with each of the environmental variables were developed (Figure 4. 2). Habitat suitability was related to the percentage of surrounding forest by an increasing linear relationship. Truncated linear functions were used to related habitat suitability with distance to the nearest road (increasing truncated linear) and distance to nearest wetland (decreasing truncated linear). A Gaussian function was used to represent habitat suitability based on elevation. Using these habitat suitability functions, each of the maps of the environmental variables was transformed into a partial habitat suitability map (Figure 4. 1). A weighted average of the four partial habitat suitability coefficients for each location was calculated to combine the partial suitability maps into a final “truth” habitat suitability map. Variability was also included by the addition of a random value generated from a uniform distribution between –0.05 and 0.05. Variability could have been included using a normal distribution; however, I chose to follow Hirzel et al.’s (2001) methodology. The “truth” habitat suitability (H) was calculated for each grid cell in the landscape using the following equation (Hirzel et al. 2001):

$$H = \frac{1}{\sum w_i} \sum w_i H_i + e$$

In this equation, w_i is the weight assigned to the i^{th} partial habitat suitability, H_i is the value of the i^{th} partial habitat suitability and e is the random variation. The weights assigned to the partial habitat suitabilities were arbitrarily selected to produce a final landscape with a wide range of habitat suitabilities (Figure 4. 3). Following calculations of habitat suitability for each grid cell, the resulting “truth” habitat suitability was mapped (Figure 4. 1).

The “truth” habitat suitability map served as the basis for generating the distribution of the virtual species. The presence (denoted by a 1) or absence (denoted by a 0) of the virtual species at a location in the landscape was calculated using a threshold value applied to the habitat suitability map. All locations with habitat suitability of greater than or equal to 0.50 were classified as present and all locations with suitability values less than 0.50 were absences. Different choices of threshold values can be used to represent differing scenarios of species abundance. For example, choosing lower threshold values can simulate an overabundant species that has spread to less suitable habitats (Hirzel 2001). Only one threshold value of 0.5 was selected for this study to limit the number of variables and concentrate specifically on the effects of the surveyor’s probability of detecting the species. Applying the 0.5 threshold to all grid cells in the landscape, a “truth” presence/absence map of the species distribution was produced. Both the “truth” habitat suitability and presence/absence maps are the source for the sample data used in model creation and provide a known reference to judge the accuracy of the resulting models.

Generating Sampling Data

To obtain sample data to build the Bayesian habitat suitability models, we followed the concept of a “virtual ecologist” (Tyre et al. 2001) who samples at locations within the landscape. The virtual ecologist randomly sampled for the presence or absence of the species at a specified number of locations. We tested 4 sample sizes, with the virtual ecologist sampling 50, 100, 200, or 500 randomly selected locations. When the probability of detection of the species was 100%, the virtual ecologist sampled the presence or absence of the species from the distribution map without error. However, we examined the impact of three additional scenarios of detection probability. We simulated surveys with 80%, 60%, and 40% mean probability of detection. The standard deviation of the probability of detection was 0.05 for each of the simulated surveys. For every presence location that was sampled by the virtual ecologist, each of the four detection probability distributions (100%, 80%, 60%, and 40% means) were used to determine if the virtual ecologist observed the species. To determine whether the virtual ecologist detected the species, a probability of detection was randomly selected from the distribution being tested. A Bernoulli distribution with the selected probability of detection was used to determine presence or absence at the survey location. A uniform random variate [0-1] was then generated and the

location was recorded as present if the uniform random variate was less than or equal to the probability of detection at that location. Otherwise, the location was recorded as an absence in the survey data set (representing a “false absence”). For each “truth” presence location sampled by the virtual ecologist, this process of determining whether or not the species was detected was repeated for each of the four probability of detection distributions. For sampled locations where the virtual species was absent on the “truth” distribution map, the virtual ecologist never detected the species. This resulted in four detection/non-detection data sets (100%, 80%, 60%, and 40% mean detectability) for each of the 4 specified sample sizes (N = 50, 100, 200, and 500). The entire sampling process was then repeated to produce 15 replications for model-building. An additional 35 replications (total = 50) were used to compare the means and standard deviations of the environmental variables measured at presence and absence locations. Landscape data were stored in Excel software and the “virtual ecologist” process was implemented using Visual Basic for Applications (VBA).

Bayesian Logistic Regression Model

The model specification for this study replicates the model that was used in Chapter 2. The specification of the Bayesian model begins by assuming the presence/absence data ($y[i]$) are from a Bernoulli distribution with a probability of presence ($p[i]$). The probability that the species occurs was determined from a logistic regression of environmental variables. All four environmental variables used to generate the species’ distribution were included as independent variables in the models. The Bayesian logistic regression model, written in WinBugs code, is depicted below:

```

For (i in 1:N){
  y[i] ~ dbern(p[i])
  logit(p[i]) <- alpha0 + alpha1 * x1[i] + alpha2 * x2[i] + alpha3 * x3[i] + alpha4 * x4[i]
}

```

Where N = size of the sample taken by the virtual ecologist

$y[i]$ = binary presence or absence
 $p[i]$ = probability of habitat suitability
 α_0 - α_4 = regression coefficients

$x1[i] - x4[i]$ = values of environmental covariates

The model then specifies non-informative priors for each of the logistic regression coefficients ($\alpha_0 - \alpha_4$) using a normal distribution. The models were initialized with values of 0 for $\alpha_0 - \alpha_4$. Bayesian analyses were conducted using WinBugs (Version 1.4, Imperial College & MRC, UK) software. All simulations were checked to ensure that convergence was reached. To reduce autocorrelation, the samples were thinned by keeping 1 update from every 20 updates produced. After thinning, the first 5,000 updates of each simulation were discarded (burn in) and the following 15,000 updates were used for analysis.

Ecological Niche Factor Analysis

The simulated data sets were also used for ecological niche factor analysis (ENFA). The locations that were sampled and classified as absences by the virtual ecologist were eliminated from the data sets and only the presence locations were used for ENFA. All data sets ($n = 240$, 4 sample sizes \times 4 detection probabilities \times 15 repetitions) were analyzed using both Bayesian logistic regression and ENFA.

ENFA was performed using Biomapper 2.1 software (Hirzel et al. 2000) that includes a collection of tools for statistical analysis and mapping using IDRISI raster format. The four environmental variables were normalized using a Box-Cox transformation (Sokal and Rohlf 1981). Ecological niche factors were computed for each data set, summarizing the transformed environmental variables into four uncorrelated factors. The first factor accounted for the marginality of the virtual species (the difference between the species' mean and the mean of the entire simulated landscape). The other three factors explained the specialization of the virtual species, by maximizing the ratio of the variance of the entire landscape to the variance of the species' distribution (Hirzel et al. 2001). Not all factors were retained for habitat suitability prediction. Factors that explained a significant amount of variation in comparison to McArthur's broken stick distribution were kept and a minimum of 2 factors was always retained. In all models, the retained factors explained greater than 80% of the total variance. A habitat suitability index was calculated for each cell in the landscape ($n = 63,001$) by comparing the values of the factors at each location with the medians of the species distribution on the factors

(Hirzel et al 2002). Using each of the 240 data sets, ENFA was conducted and a predicted habitat suitability map for the 251 x 251 grid cell area was produced.

Model Evaluation

Performance of Bayesian logistic regression and ENFA models was evaluated in relation to both calibration (agreement between predicted habitat suitability and “true” habitat suitability) and discrimination capacity (ability to correctly predict presence and absence locations) (Pearce and Ferrier 2000). Predicted habitat suitability for Bayesian and ENFA models was compared to the “truth” habitat suitability map at the sampled locations within each data set. Because Bayesian models produce a posterior distribution of predicted habitat suitability, the mean of this distribution was used as the estimate for predicted habitat suitability during model evaluation. Both presence and absence locations were used for model evaluation, although the absence data were not used for building the ENFA models. A Pearson correlation statistic was calculated using Minitab statistical software (Release 14, Minitab, Inc.). To assess discrimination capacity, mean predicted habitat suitability was also transformed into predicted presence/absence through use of a threshold value. Locations with a greater than 0.5 predicted habitat suitability were classified as present. Probabilities less than 0.5 were classified as absent. To compare results between the model and “truth” distribution, sensitivity (conditional probability that the location is predicted present when it is actually present) and specificity (conditional probability that the location is predicted absent when it is actually absent) were calculated for each model (Fielding and Bell 1997). The correlation coefficient, sensitivity, and specificity statistics were calculated for each model ($n = 480$). Analysis of variance was conducted using model type, detection rates, and sample sizes as factors. All possible interactions among factors were also included in the ANOVA model. For significant factors, comparisons of least squared means were used to evaluate effects.

To investigate the influence of the choice of threshold value on the discrimination capacity of the model, the threshold value was varied for a subset of the models. For the first 3 repetitions of each treatment (model type x sample size x probability of detection), the threshold value was

varied between 0.1 and 0.9 in increments of 0.1. Sensitivity and specificity statistics were then calculated for the model at each threshold value and averaged over the 3 repetitions.

RESULTS

Habitat Suitability

Construction of the “truth” habitat suitability map resulted in a landscape with a mean suitability of 0.547 that ranged from 0 – 1 (Figure 4. 3). When a 0.50 threshold value was applied, 40,194 out of 63,001 landscape cells were classified as present. The number of presence locations included in a data set (Table 4. 1) was a function of the number of presence locations in the landscape, size of the sample ($n = 50, 100, 200, 500$), and probability of the virtual ecologist detecting the species ($p = 40\%, 60\%, 80\%, 100\%$). Data sets produced from sampling by the virtual ecologist contained between 24.67% - 64.02% presence locations out of the total number of samples (Table 4. 1).

Environmental Predictor Variables

The means of environmental predictor variables were calculated for each sample taken by the virtual ecologist. Mean values of all environmental variables differed based on the classification of the site (presence or absence) and probability of detection (Table 4. 2). There was a significant interaction between the site classification and probability of detection for all environmental variables. For locations classified as “present” by the virtual ecologist, the mean values of the environmental variables were similar regardless of the probability of detection (Figure 4. 4). For locations classified as “absent,” mean values of the environmental variables were different at each probability of detection level (Figure 4. 4). As probability of detection decreased, the mean of the absence locations became closer to the mean of the presence locations for all environmental variables. However, even at the lowest detection probability level (40%), the means of the environmental variables differed between presence and absence locations. Although a decreasing probability of detection did not impact the means for presence locations, the variation around the mean values of the environmental variables increased (Figure 4. 4). Sample size influenced the mean values of distance to wetland, distance to road, and percentage

of forest variables, but was not significant for elevation. Comparisons of least squared means revealed differences between sample sizes of $n=50$ and $n=100$, but neither sample size was significantly different from sample sizes $n=200$ and $n=500$.

Bayesian Logistic Regression Models

Problems with model convergence were found when constructing the Bayesian logistic regression models with small sample sizes of $n = 50$. Because errors were dependent on the sample used to build the model, the data set was discarded and a new sample drawn when a model was unable to be constructed with WinBugs. To reach a total of 15 successful model repetitions, a total of 19 data sets were unable to be used because of errors related to convergence. Convergence-related issues were not a problem for any data sets containing sample sizes of $n=100$, 200, or 500.

Mean correlation between the predicted habitat suitability and true habitat suitability was 0.73, $SE = 0.0044$. Both the probability of detection and sample size influenced the predictive accuracy of the models. Correlation between the predicted habitat suitability and true habitat suitability decreased as the probability of detection decreased (Figure 4. 5A). Bayesian logistic regression models constructed from data sets with 100% detection had greater correlation with the true suitability than models from data sets with 60% and 40% detection (Table 4. 3). Suitability correlations for models from data sets with 80% detection were not significantly different from models with either 100% or 60% probability of detection, but were more strongly correlated with the true suitability than models with 40% detection (Table 4. 3). The size of the sample data set used to build the Bayesian model also affected the correlation of habitat suitability. Models based on sample data sets of 50 locations were poorer at predicting habitat suitability than models from data sets of $n = 100$, 200, and 500 locations (Figure 4. 5A). For sample sizes of $n = 100$, 200, and 500, differences in the correlation of habitat suitability were not significant and the models performed similarly.

Model sensitivity, i.e. the ability of the model to correctly predict presence locations, averaged 0.58 ($SE = 0.0044$) for all Bayesian logistic regression models, meaning that 58% of the true

presence locations were correctly classified by the model. Mean sensitivity declined from 0.90 when the virtual ecologist sampled without error to 0.10 when the probability of the virtual ecologist detecting the species was 40% (Figure 4. 6A). The probability of detection and the interaction between model type and probability of detection accounted for the greatest proportion of variance in model sensitivity (Table 4. 4). Sample size was less influential than probability of detection, but models from data sets of $n = 100$ predicted presence locations better (l.s. mean = 0.606, SE = 0.0089) on average than models from $n = 200$ (l.s. mean = 0.569) and 500 (l.s. mean = 0.563) locations. Models from data sets of $n = 50$ locations (l.s. mean = 0.571) were not significantly different from any of the other models based on greater sample sizes.

Specificity, or the ability of Bayesian logistic regression models to correctly predict absence locations, averaged 0.90 (SE = 0.0042) for all models. Specificity was correlated with probability of detection, with specificity increasing from 0.77 to 0.99 as the probability of detection decreased (Figure 4. 7A). Specificity was statistically similar between models from the 40% and 60% probability of detection scenarios (Table 4. 3). Both the 40% and 60% detection scenarios produced models with higher specificity than models based on data sets from the 80% (l.s. mean = 0.87) and 100% (l.s. mean = 0.77) probability of detection scenarios. Model specificity was not dependent on the size of the data set used for model construction when the sample contained 100 or more locations. Sample sizes of 50 locations produced models with slightly lower specificity at all of the examined detection probabilities (Figure 4. 7A).

Ecological Niche Factor Analysis Models

For ENFA models, predicted habitat suitability was less correlated with true habitat suitability (l.s. mean = 0.65, SE = 0.004506) than Bayesian logistic regression models (l.s. mean = 0.73) on average for all models. ENFA models performed consistently for all probabilities of detection that were tested, with no significant differences in the correlation of suitability among probability of detection levels of 40% to 100% (Figure 4. 5B). Prediction of habitat suitability was influenced by low sample size. Predictions from models constructed from data sets of $n = 50$ observations (l.s. mean = 0.61, SE = 0.0090) were less correlated with true suitability than model predictions based on sample sizes of $n = 100$ (l.s. mean = 0.65), $n = 200$ (l.s. mean = 0.68), $n =$

500 (l.s. mean = 0.67). There were no significant differences in the correlation of suitability among models constructed from $n = 100$, 200 , and 500 .

Average sensitivity for ENFA models was 0.63 , meaning that the models predicted 63% of the known presence locations correctly. Probability of detection did not influence model sensitivity (Figure 4. 6B), with no significant differences among 100% detection (l.s. mean = 0.65 , SE = 0.0089), 80% detection (l.s. mean = 0.63), 60% detection (l.s. mean = 0.64), and 40% detection (l.s. mean = 0.62). Model sensitivity was not dependent on the sample size of observations used to build the model, with no significant differences among $n = 50$ (l.s. mean = 0.65 , SE = 0.0089), $n = 100$ (l.s. mean = 0.63), $n = 200$ (l.s. mean = 0.62), and $n = 500$ (l.s. mean = 0.63).

Average specificity for ENFA models was 0.84 and was not influenced by the probability of detection (Figure 4. 7A). All levels of probability of detection produced models that correctly predicted 84% of the absence locations (although the absence locations were not used in the construction of the model). Models constructed from data sets based on $n = 50$ locations predicted fewer absence locations correctly (l.s. mean = 0.79 , SE = 0.0085) than models based on data sets of $n = 100$ (l.s. mean = 0.85), $n = 200$ (l.s. mean = 0.86), and $n = 500$ (l.s. mean = 0.86) locations.

Model Comparison

Predictive accuracy, quantified as correlation of predicted and true suitability and model sensitivity and specificity, differed based on the choice of modeling method (Table 4. 4). However, the relationships between model type and measurements of predictive accuracy were not consistent for all levels of probability of detection and sample size, with significant interactions among factors (Table 4. 4). As probability of detection or sample size change, the choice of which modeling method produces the highest accuracy changes.

The relationships among model accuracy and probability of detection differed based on the type of model, evidenced by a significant interaction of model type and probability of detection in the analysis of variance (Table 4. 4). When probability of detection was 100% , Bayesian logistic

regression produced models with greater correlation to the true habitat suitability, higher accuracy of classifying presence locations, and lower accuracy of classifying absence locations, than ENFA models (Table 4. 3). A similar pattern was found when probability of detection was 80%, with the exception that the accuracy in classifying absence locations was equivalent between Bayesian and ENFA models. When probability of detection was 60%, Bayesian methods produced models with higher suitability correlation and accuracy predicting absence locations, but ENFA models had higher accuracy in predicting presence locations. When probability of detection was 40%, Bayesian and ENFA methods produced models with similar correlations to true habitat suitability, ENFA models had higher accuracy predicting presence locations, and Bayesian models had higher accuracy predicting absence locations.

The relationships of sample size with correlations of habitat suitability and model specificity were similar for both Bayesian logistic regression and ENFA models, evidenced by the non-significance of the interaction between model type and sample size (Table 4. 4). Both modeling methods produced models with lower correlations of habitat suitability and specificities when using sample sizes of $n = 50$ locations. The relationship of model sensitivity and sample size differed between Bayesian and ENFA models. ENFA produced models of similar sensitivity regardless of sample size, while Bayesian logistic regression models produced from data sets of $n = 100$ locations had greater sensitivity than models from $n = 50, 200,$ and 500 locations.

The choice of threshold value did influence model results. When detectability was 40%, a threshold value of 0.21 produced the highest model sensitivity and specificity for Bayesian logistic regression models (Figure 4. 8). When detectability was 100%, Bayesian models reached a similar sensitivity and specificity when the threshold value was 0.63. ENFA models performed consistently and reached similar levels of model sensitivity and specificity when the threshold value was low (0.33), regardless of the probability of detection.

DISCUSSION

Relationship Between Probability of Detection and Bayesian Model Accuracy

A strong relationship was found between the probability of detection of the species in the sample data set with the resulting accuracy of the Bayesian model built from the data. When species detection was high, Bayesian logistic regression models predicted both the true habitat suitability and presence/absence of the species well. Decreasing the probability of detection from 100% (i.e. always detecting and classifying the location as present when the species occurs), created more false absences within the sample data sets and altered the mean values of habitat variables for absence locations in the sample data sets. Addition of false absences to the sample data, shifted the mean values of environmental variables at absence locations towards the mean of the variables at presence locations. In effect, this shift makes it increasingly difficult to distinguish between presence and absence locations on the basis of values of environmental variables. The error in the sample data as a result of reduced detection of the species led to lower correlation of the predicted habitat suitability with the true habitat suitability and decreased accuracy in predicting presence locations. The increase in accuracy of predicting absence locations at low levels of detection resulted from the larger number of absence locations in the sample data set. As probability of detection decreases, the proportion of absence locations within the sample data increases. As a result, the Bayesian logistic regression models predict lower suitability and absence for a larger portion of the landscape. The proportion of presence versus absence locations within the data set is known to influence logistic regression models. At low probabilities of detection, Bayesian logistic regression model performance is poor because low habitat suitability is predicted for almost all locations and fewer than 10% of the locations where the species occurs are predicted as present.

Bayesian Logistic Regression versus ENFA

Choosing between modeling habitat suitability with a method that uses presence/absence data, such as Bayesian logistic regression, or with a method that uses presence-only data, such as ENFA, depends on more than the probability of detection. The decision is influenced by the source of the sampling data and whether the data for model creation have already been collected

or are still in the survey design phase (Figure 4. 9). In addition, prior knowledge or an estimate of the probability of detection involved in the collection of the survey data may or may not be available. For example, if the probability of detection is unknown, the choice of modeling technique to achieve the best result is different from situations where the probability of detection is known to be low. In modeling situations where the probability of detection is not known, the choice of which aspect of the model, accuracy of suitability prediction or discrimination ability (distinguishing between presence and absence locations), is more important becomes a factor.

Starting with the simplest case where sample data already exists and the probability of detection is known, the choice of modeling technique is straightforward. When the probability of detection is high ($\geq 80\%$), Bayesian logistic regression is a better method for predicting the actual habitat suitability and classifying presence and absence locations. When the sample data has already been collected and the probability of detection is known to be less than 80%, ENFA becomes the better choice as a method for modeling. ENFA is robust to the probability of detection because it is not influenced by false absences. When the probability of detection is low, fewer presence locations are found and included within the sample data set. So in this study, when probability of detection was low (40%) and 100 samples were collected, Bayesian logistic regression models used all 100 locations (including 60% of the presence locations incorrectly classified as absence locations) while ENFA only used the 25 locations where the species was correctly detected (100 locations sampled \times 62% of landscape where species actually occurs \times 40% detection where it actually occurs).

In modeling situations where the probability of detection of the species is unknown and unable to be estimated, it becomes important to examine the objectives for building the model. If it is more important to predict the true habitat suitability at specific locations, Bayesian logistic regression produced models that were more highly correlated (or at least equally correlated) with the true suitability than ENFA models at all levels of probability of detection. However, if the modeling objective is to discriminate between presence and absence locations with the highest accuracy, ENFA methods performed more consistently at predicting both presence and absence locations with relatively high accuracy.

In modeling situations where the sample data have not yet been collected and the survey is still in the design phases, it is possible to ensure that an estimate of species' detection is obtained and the methods are designed to produce the highest rate of detection possible. If detection is estimated at less than 80%, it is advisable to alter the survey design to increase detection by increasing the number of repeat visits to the site, identifying and using a more efficient sampling technique or combination of techniques, or increasing the size of the sample location (i.e. the resolution or minimum cell size of the model). Based on this study, it would be more beneficial to increase the number of repeat visits to a site to ensure detection of at least 80% than to increase the number of locations sampled.

Influence of Sample Size on Model Accuracy

The size of the sample data set did not have a large effect on the accuracy of habitat suitability models produced by either Bayesian logistic regression or ENFA. Model accuracy of Bayesian methods was robust to samples containing 100 to 500 presence/absence locations. ENFA methods were robust to data sets based on 100 to 500 samples that contained an average of 25 to 316 presence locations. Models constructed from data sets based on $n = 50$ locations had several drawbacks. For Bayesian logistic regression, many of these data sets produced errors when running under WinBugs. Although the model is appropriate for the data, the models run into computational errors within the software that are dependent on the specific data set. When Bayesian models of $n = 50$ locations were successfully constructed, they had lower correlations with the true habitat suitability and decreased accuracy in predicting absence locations. ENFA models based on $n = 50$ locations, actually were built using 12 to 31 presence locations depending on the probability of detection. These ENFA models also produced predictions that were less correlated with true suitability and had lower specificity. Habitat suitability modeling will be most successful using both Bayesian logistic regression and ENFA methods if the sample data sets contain greater than or equal to 100 locations or 30 presence locations.

Assumptions of Simulation

The results of this study are dependent on the assumptions used to create the "virtual species" distribution and replicate the sampling process with the "virtual ecologist." Simulated data have

been used frequently in other types of studies (Mesple et al. 1996), but have only recently been applied to habitat suitability modeling (Hirzel et al. 2001, Tyre et al. 2001). Previous simulation studies have found that the performance of different types of habitat suitability modeling methods are sensitive to species' prevalence across the landscape (Hirzel et al. 2001) and the sampling strategy for obtaining sample data (Hirzel and Guisan 2002). Bias could have been introduced if the simulated species distribution does not correctly depict a typical reality. Only one species distribution was examined that had a prevalence of 62%. The relationship between probability of detection and model accuracy may be different when species' prevalence is more or less common than the simulated distribution. Following Hirzel et al.'s (2001) example, I reduced the risk of bias by trying to keep the simulations as realistic as possible through the use of real environmental data from an existing landscape. I also used both linear and non-linear components within the niche function and added stochasticity to the resulting "true" habitat suitability. So, although simulated data were used, prediction accuracy for any method was not 100% because of variability introduced in the calculation of the true habitat suitability, non-linearity in the relationships between suitability and environmental variables, and sampling error.

In addition the choice of 0.50 as the threshold value for determining presence or absence, influenced the measures of model accuracy. With Bayesian logistic regression, model sensitivity and specificity were improved when a lower threshold value was selected when detectability was low. The threshold value of 0.50 appears to work best for Bayesian logistic regression when detectability is between 80-100%. For ENFA models, a lower threshold value, closer to 0.3, yielded the best results when discriminating between presence and absence locations. Suitability predicted by ENFA models was less correlated with the true suitability and predicted lower suitability on average. Regardless of detectability, using a lower threshold value for ENFA models resulted in increased success of model discrimination between presence and absence locations.

Conclusions

A strong relationship exists between the probability of species' detection and the accuracy of Bayesian logistic regression models. ENFA models are insensitive to this factor because of their

exclusion of absence data, including the false absence data that is prevalent when species' detection is low. ENFA models were also robust to sample size and produced reliable models that were able to discriminate between presence and absence locations with 63% and 84% accuracy, respectively. Bayesian methods produced models of higher accuracy when the probability of species' detection was high, but much poorer models when detection was 60% or less. Knowledge of model performance under differing scenarios of species' detectability and sample size is important for selecting the appropriate modeling technique under differing conditions and designing appropriate surveys to obtain the most effective data for habitat suitability modeling.

LITERATURE CITED

- Anderson, R. P., Lew, D., and R. T. Peterson. 2003. Evaluating predictive models of species' distributions: criteria for selecting optimal models. *Ecological Modelling*, 162:211-232.
- Angermeier, A. L. and R. A. Smogor. 1995. Estimating number of species and relative abundances in stream-fish communities: effects of sampling effort and discontinuous spatial distributions. *Canadian Journal of Fisheries and Aquatic Sciences*, 52:936-949.
- Angermeier, P. L., Krueger, K. L., and C. A. Dolloff. 2002. Discontinuity in stream-fish distributions: Implications for assessing and predicting species occurrence. Pp. 519-527 In J. M. Scott, P. J. Heglund, M. L. Morrison, J. B. Haufler, M. G. Raphael, W. A. Wall, and F. B. Samson, eds. *Predicting species occurrences: issues of accuracy and scale*. Island Press, Covelo, CA.
- Bayley, P. B. and J. T. Peterson. 2001. An approach to estimate probability of presence and richness of fish species. *Transactions of the American Fisheries Society*, 130:620-633.
- Boone, R. B. and W. B. Krohn. 1999. Modeling the occurrences of bird species: are the errors predictable. *Ecological Applications*, 9:835-848.
- Ellison, A. M. 1996. An introduction to Bayesian inference for ecological research and environmental decision-making. *Ecological Applications*, 6:1036-1046.
- Dettmers, R. and J. Bart. 1999. A GIS modeling method applied to predicting forest songbird habitat. *Ecological Applications*, 9:152-163.
- Fielding, A. H. and J. F. Bell. 1997. A review of methods for the assessment of prediction errors in conservation presence/absence models. *Environmental Conservation*, 24:38-49.
- Fleishman, E., MacNally, R., Fay, J. P., and D. D. Murphy. 2001. Modeling and predicting

- species occurrence using broad-scale environmental variables: an example with butterflies of the Great Basin. *Conservation Biology*, 15:1674-1685.
- Franklin, J. 1988. Predicting the distribution of shrub species in southern California from climate and terrain-derived variables. *Journal of Vegetation Science*, 9:733-748.
- Guisan, A. and N. E. Zimmerman. 2000. Predictive habitat distribution models in ecology. *Ecological Modelling*, 135:147-186.
- Hirzel, A. H., and A. Guisan. 2002. Which is the optimal strategy for habitat suitability modeling? *Ecological Modelling*, 157:331-341.
- Hirzel, A. H., Hausser, J., Chessel, D., and N. Perrin. 2002. Ecological-niche factor analysis: how to compute habitat suitability maps without absence data? *Ecology*, 83:2027-2036.
- Hirzel, A. H., Helfer, V., and F. Metral. 2001. Assessing habitat-suitability models with a virtual species. *Ecological Modelling*, 145:111-121.
- Hutchinson, G. E. 1957. Concluding remarks. *Cold Spring Harbor Symposium on Quantitative Biology*, 22:415-427.
- Kadmon, R., Farber, O., and A. Danin. 2003. A systematic analysis of factors affecting the performance of climatic envelope models. *Ecological Applications*, 13:853-867.
- McArdle, B. H. 1990. When are rare species not there? *Oikos*, 57:276-277.
- Meyer, R. and R. B. Millar. 1999. BUGS in Bayesian stock assessments. *Canadian Journal of Fisheries and Aquatic Sciences*, 56:1078-1086.
- Monserud, R. A. and R. Leemans. 1992. Comparing global vegetation maps with the Kappa statistic. *Ecological Modelling*, 62:275-293.
- Nicholson, M. and J. Barry. 1995. Inferences from spatial surveys about the presence of an unobserved species. *Oikos*, 72:74-78.
- Olden, J. D., Jackson, D. A., and P. R. Peres-Neto. 2002. Predictive models of fish species distributions: a note on proper validation and chance predictions. *Transactions of the American Fisheries Society*, 131:329-336.
- Palma, L., Beja, P., and M. Rodrigues. 1999. The use of sighting data to analyse Iberian lynx habitat and distribution. *Journal of Applied Ecology*, 36:812-824.
- Pearce, J. and S. Ferrier. 2000. Evaluating the predictive performance of habitat models developed using logistic regression. *Ecological Modelling*, 133:225-245.
- Peterson, A. T. and C. R. Robins. 2003. Using ecological-niche modeling to predict barred owl invasions with implications for spotted owl conservation. *Conservation Biology*,

17:1161-1165.

Peterson, J. T. and J. Dunham. 2003. Combining inferences from models of capture efficiency, detectability, and suitable habitat to classify landscapes for conservation of threatened bull trout. *Conservation Biology*, 17:1070-1077.

Reutter, B. A., Helfer, V., Hirzel, A. H., and P. Vogel. 2003. Modelling habitat-suitability using museum collections: an example with three sympatric *Apodemus* species from the Alps. *Journal of Biogeography*, 30:581-590.

Solow, A. R. 1993. Inferring extinction from sighting data. *Ecology*, 74:962-964.

Tyre, A. J., Possingham, H. P., and D. B. Lindenmayer. 2001. Inferring process from pattern: can territory occupancy provide information about life history parameters. *Ecological Applications* 11:1722-1737.

Wade, P. R. 2000. Bayesian methods in conservation biology. *Conservation Biology*, 14:1308-1316.

Zaniewski, A. E., Lehmann, A., and J. McC. Overton. 2002. Predicting species spatial distributions using presence-only data: a case study of native New Zealand ferns. *Ecological Modelling*, 157:261-280.

Table 4. 1. Mean and standard deviation of the number of presence locations in the data sets used to build Bayesian logistic regression and ENFA models, listed in order from greatest to fewest presence locations.

Sample Size	Detection Probability	Mean Number of Presence Locations ^a	Standard Deviation	Mean % Presence Locations ^b
500	100	316.18	12.18	63.24
500	80	252.86	11.67	50.57
500	60	191.16	10.41	38.23
500	40	127.26	10.19	25.45
200	100	125.96	6.50	62.98
200	80	100.6	6.92	50.30
200	60	76.78	5.12	38.39
100	100	64.02	4.63	64.02
100	80	51.66	5.56	51.66
200	40	49.34	5.47	24.67
100	60	38.42	4.43	38.42
50	100	31.92	3.29	63.84
100	40	25.62	4.50	25.62
50	80	25.36	3.56	50.72
50	60	19.52	3.25	39.04
50	40	12.76	2.63	25.52

^aMean number of presence locations identified by the “virtual ecologist” from 15 replications of the specified sample size and detection probability.

^bPercentage of the samples in the data set that are classified by the “virtual ecologist” as present.

Table 4. 2. Analysis of variance of environmental predictor variables as sampled by the “virtual ecologist” for 50 repetitions. F statistics and p values listed for each factor, with factors significant at the 95% level in bold.

Source of Variance	d.f.	Distance to Wetland	Distance to Road	Elevation	% Forest
Site Classification (Presence or Absence)	1	15843.25 P < 0.001	7078.68 P < 0.001	8012.35 P < 0.001	3332.18 P < 0.001
Sample Size	3	4.12 P = 0.006	4.21 P = 0.006	0.79 P = 0.497	3.01 P = 0.029
Detection Probability	3	395.08 P < 0.001	171.37 P < 0.001	201.98 P < 0.001	83.0 P < 0.001
Classification*Sample Size	3	1.51 P = 0.210	1.33 P = 0.262	0.39 P = 0.758	0.55 P = 0.648
Classification*Detection Probability	3	386.93 P < 0.001	180.82 P < 0.001	180.33 P < 0.001	85.8 P < 0.001
Sample Size*Detection Probability	9	0.07 P = 1.0	0.13 P = 0.999	0.14 P = 0.999	0.07 P = 1.0
Classification*Sample Size*Detection Probability	9	0.28 P = 0.981	0.50 P = 0.875	0.50 P = 0.875	0.30 P = 0.976

Table 4. 3. Comparison of model accuracy using least square means and 95% confidence intervals among Bayesian logistic regression and ENFA models built with data from differing probability of detection scenarios.

Model Type	Probability of Detection	Mean	95% CI	Mean Comparison ^a
Correlation of Predicted and True Habitat Suitability				
Bayesian	100%	0.7896	0.7719 – 0.8072	A
Bayesian	80%	0.7595	0.7418 – 0.7771	AB
Bayesian	60%	0.7330	0.7153 – 0.7506	B
ENFA	100%	0.6694	0.6517 – 0.6870	C
ENFA	80%	0.6634	0.6457 – 0.6810	C
ENFA	60%	0.6539	0.6363 – 0.6716	C
Bayesian	40%	0.6370	0.6194 – 0.6547	CD
ENFA	40%	0.6183	0.6006 – 0.6359	D
Model Sensitivity (Proportion of True Presence Locations Correctly Predicted)				
Bayesian	100%	0.9042	0.8867 – 0.9216	A
Bayesian	80%	0.7881	0.7706 – 0.8055	B
ENFA	100%	0.6497	0.6322 – 0.6671	C
ENFA	60%	0.6373	0.6199 – 0.6547	C
ENFA	80%	0.6332	0.6157 – 0.6506	C
ENFA	40%	0.6174	0.6000 – 0.6348	C
Bayesian	60%	0.5211	0.5036 – 0.5385	D
Bayesian	40%	0.0961	0.0787 – 0.1136	E
Model Specificity (Proportion of True Absence Locations Correctly Predicted)				
Bayesian	40%	0.9931	0.9764 – 1.0000	A
Bayesian	60%	0.9633	0.9466 – 0.9799	A
Bayesian	80%	0.8658	0.8491 – 0.8824	B
ENFA	40%	0.8448	0.8282 – 0.8615	B
ENFA	80%	0.8431	0.8265 – 0.8598	B
ENFA	100%	0.8403	0.8236 – 0.8569	B
ENFA	60%	0.8365	0.8198 – 0.8531	B
Bayesian	100%	0.7695	0.7529 – 0.7862	C

^aMeans with the same letter are not significantly different at alpha = 0.05.

Table 4. 4. Analysis of variance of model accuracy, as measured by correlation of predicted and true suitability, model sensitivity (proportion of presences correctly predicted), and specificity (proportion of absences correctly predicted). Analysis of fifteen replications (n = 480) of each combination of model type (Bayesian logistic regression and ENFA), sample sizes (n = 50, 100, 200, and 500) and probabilities of detection (p = 40%, 60%, 80%, and 100%) as sampled by the “virtual ecologist.” F statistics and p values listed for each factor, with factors significant at the 95% level in bold.

Source of Variance	d.f.	Habitat Suitability Correlation	Sensitivity	Specificity
Model Type (Bayesian/ENFA)	1	151.88 P <0.001	82.30 P <0.001	89.19 P <0.001
Sample Size	3	27.51 P <0.001	2.66 P = 0.048	16.70 P <0.001
Detection Probability	3	48.56 P <0.001	869.05 P < 0.001	71.73 P <0.001
Model Type*Sample Size	3	1.08 P = 0.358	3.28 P = 0.021	1.99 P = 0.114
Model Type*Detection Probability	3	11.54 P <0.001	761.55 P < 0.001	70.85 P <0.001
Sample Size*Detection Probability	9	5.11 P <0.001	0.75 P = 0.666	0.18 P = 0.996
Model Type*Sample Size*Detection Probability	9	1.59 P = 0.115	0.52 P = 0.863	0.21 P = 0.993

FIGURE LEGENDS

Figure 4. 1. Process of creating “virtual species” distribution from maps of environmental factors (distance to nearest road, distance to wetland, elevation, and percentage of forest in surrounding area), to maps of partial suitability for each environmental factor, to the final suitability map and the presence/absence map. Each map is composed of 251×251 grid cells with each grid cell measuring 30×30-m. Light areas on the raw data maps represent areas with a high measurement of the specific environmental variable, dark areas represent low values. Suitability maps range from 0-1, with light areas representing high suitability and dark areas representing low suitability.

Figure 4. 2. Partial habitat suitability functions used to relate the environmental factors (distance to nearest road, distance to wetland, elevation, and percentage of forest in surrounding area) to a quantified suitability for the “virtual species.”

Figure 4. 3. Distribution of final habitat suitability values from the “truth” habitat suitability map.

Figure 4. 4. Distributions of sampled environmental variables (distance to nearest road, distance to wetland, elevation, and percentage of forest in surrounding area) at four levels of species detection (40%, 60%, 80%, and 100%) when sample size = 500 presence and absence locations as sampled by the virtual ecologist during 50 repetitions.

Figure 4. 5. Mean correlation of predicted vs. actual habitat suitability for Bayesian logistic regression and ENFA models constructed under four different scenarios of species’ detection (40%, 60%, 80%, and 100%) and four sample sizes (n = 50, 100, 200, and 500) with 15 replications of each treatment combination.

Figure 4. 6. Mean sensitivity for Bayesian logistic regression and ENFA models constructed under four different scenarios of species’ detection (40%, 60%, 80%, and 100%) and four sample sizes (n = 50, 100, 200, and 500) with 15 replications of each treatment combination.

Figure 4. 7. Mean specificity for Bayesian logistic regression and ENFA models constructed under four different scenarios of species’ detection (40%, 60%, 80%, and 100%) and four sample sizes (n = 50, 100, 200, and 500) with 15 replications of each treatment combination.

Figure 4. 8. Mean sensitivity and specificity of Bayesian logistic regression and ENFA models at differing threshold values from 0.1 to 0.9. Models were created from n=500 observations.

Figure 4. 9. Questions relative to detection probability that should be considered when choosing between Bayesian logistic regression and ENFA as modeling methods for presence/absence survey data.

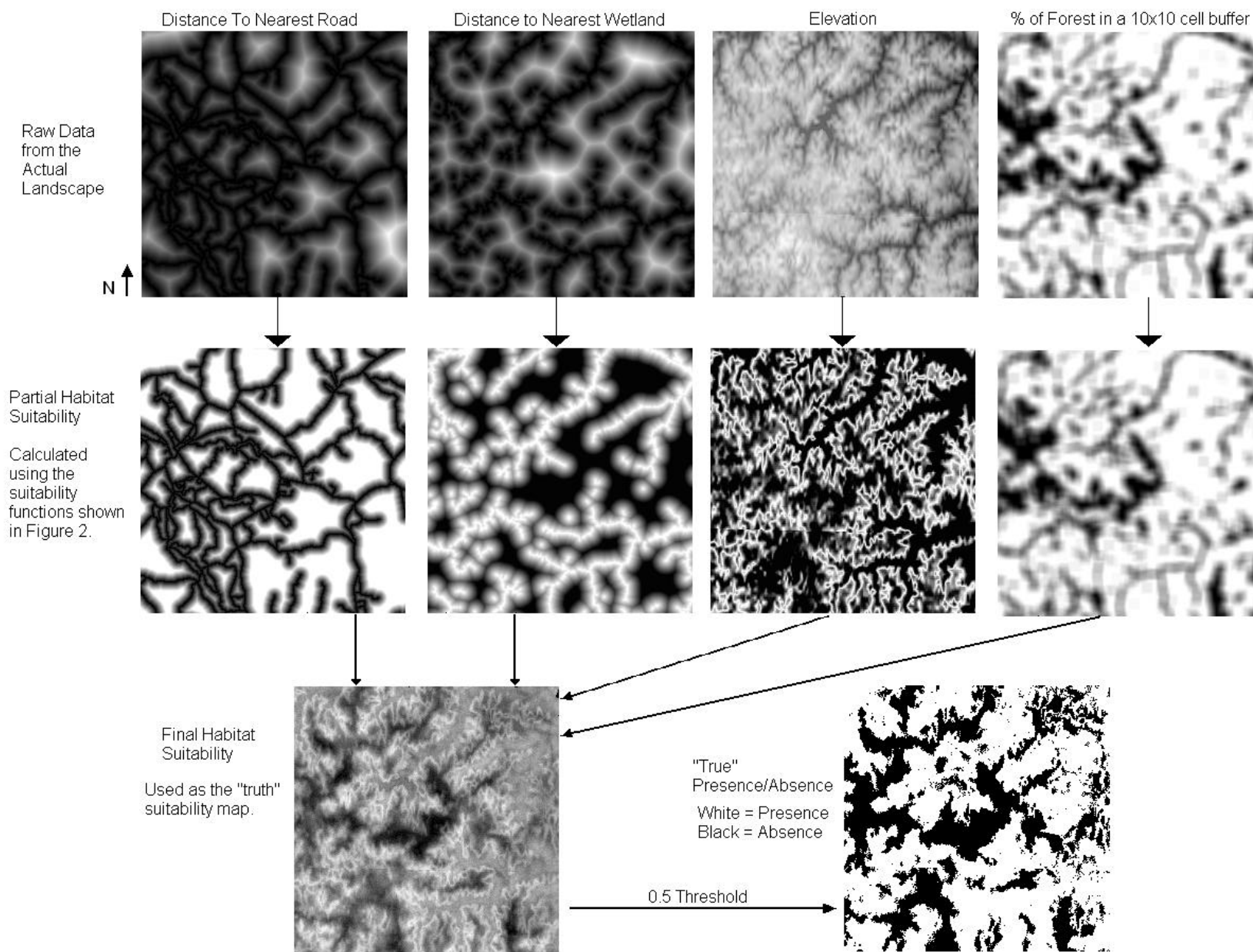


Figure 4. 1.

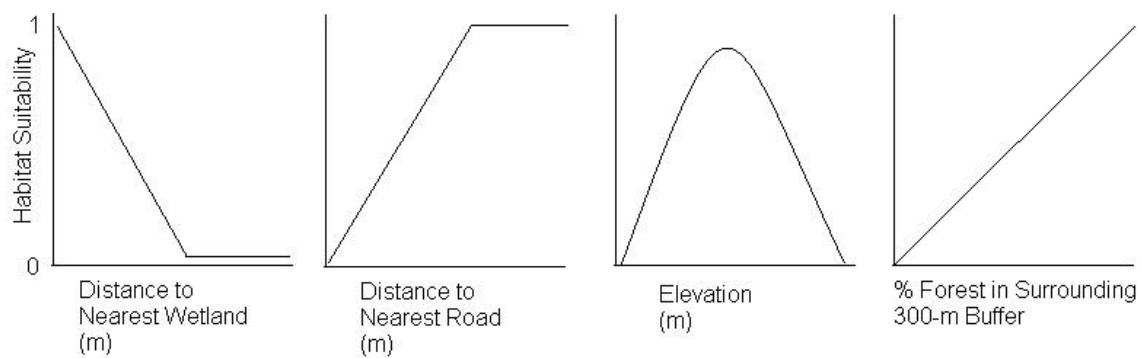


Figure 4. 2.

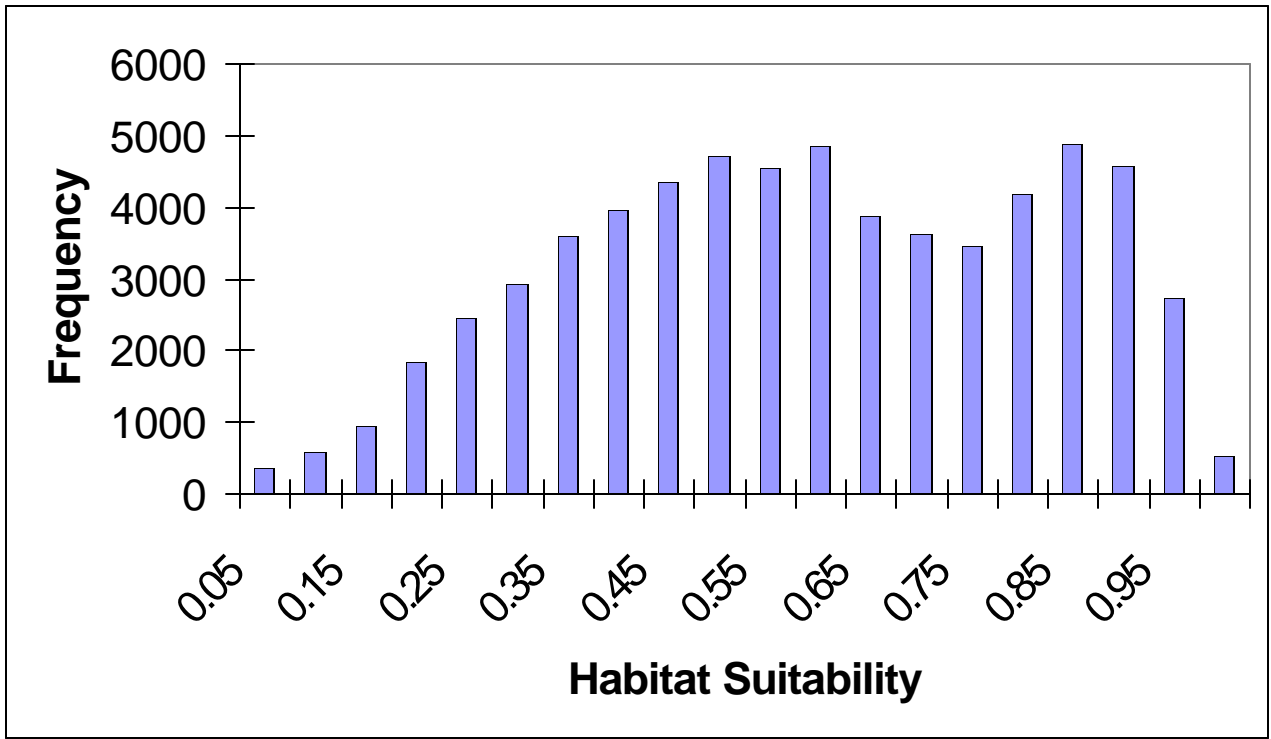


Figure 4. 3.

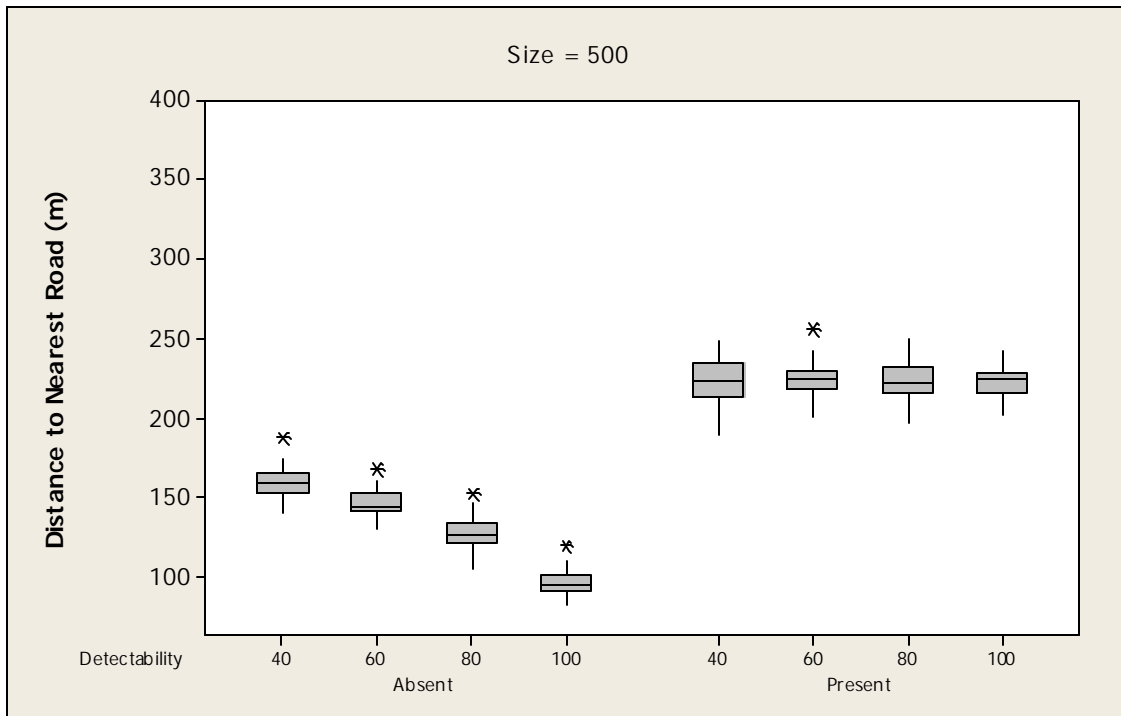
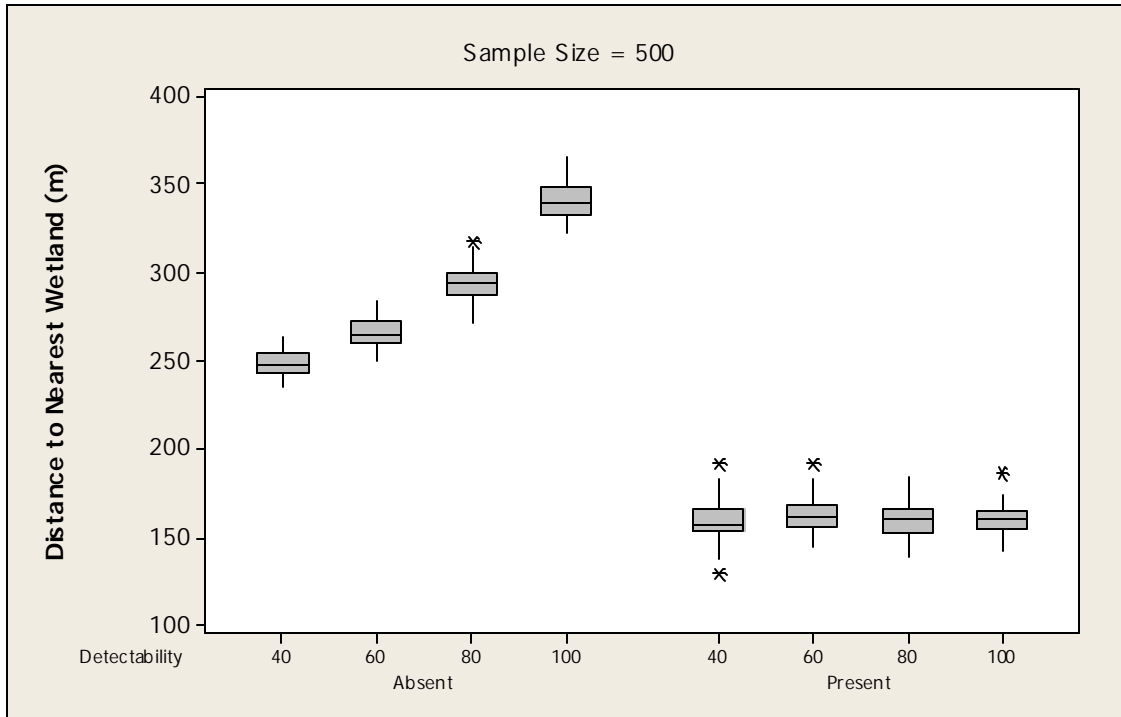


Figure 4. 4.

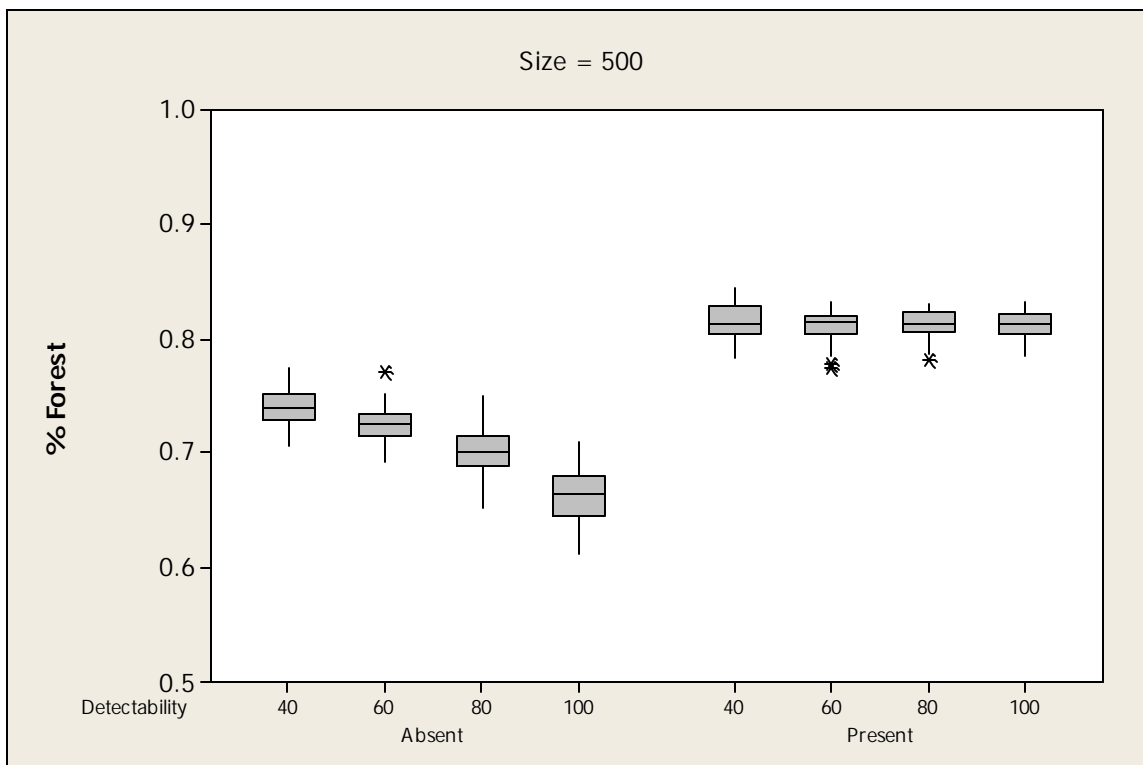
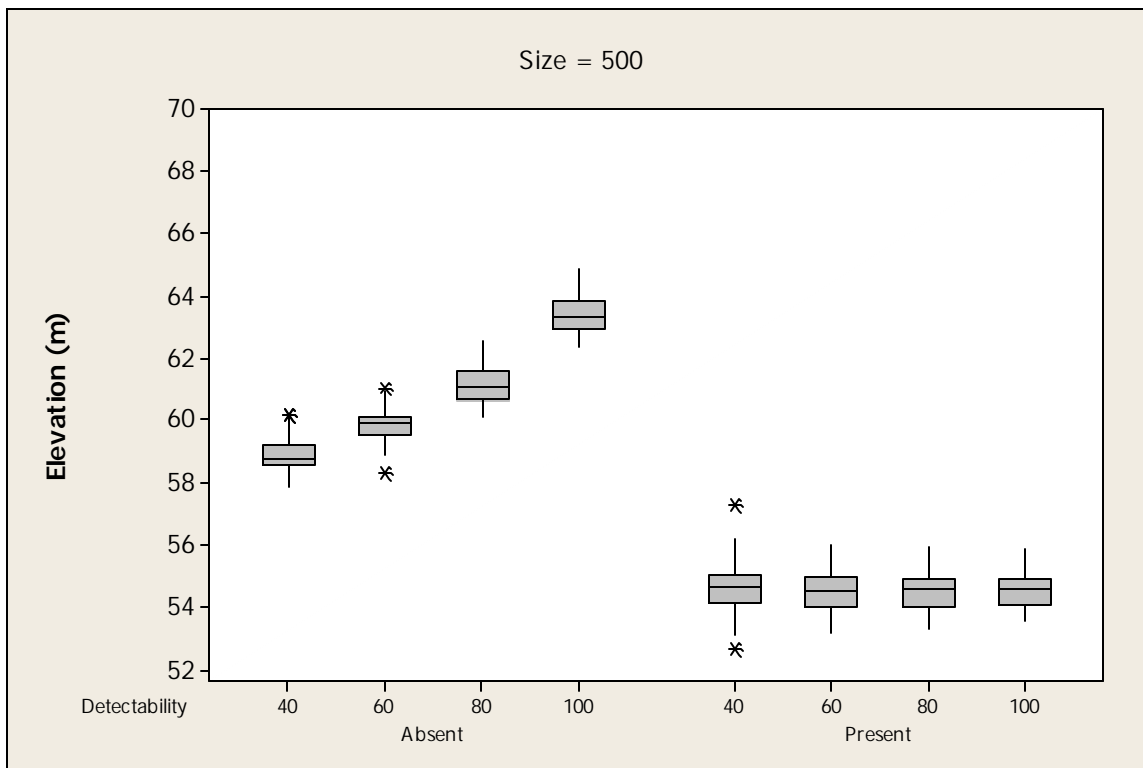
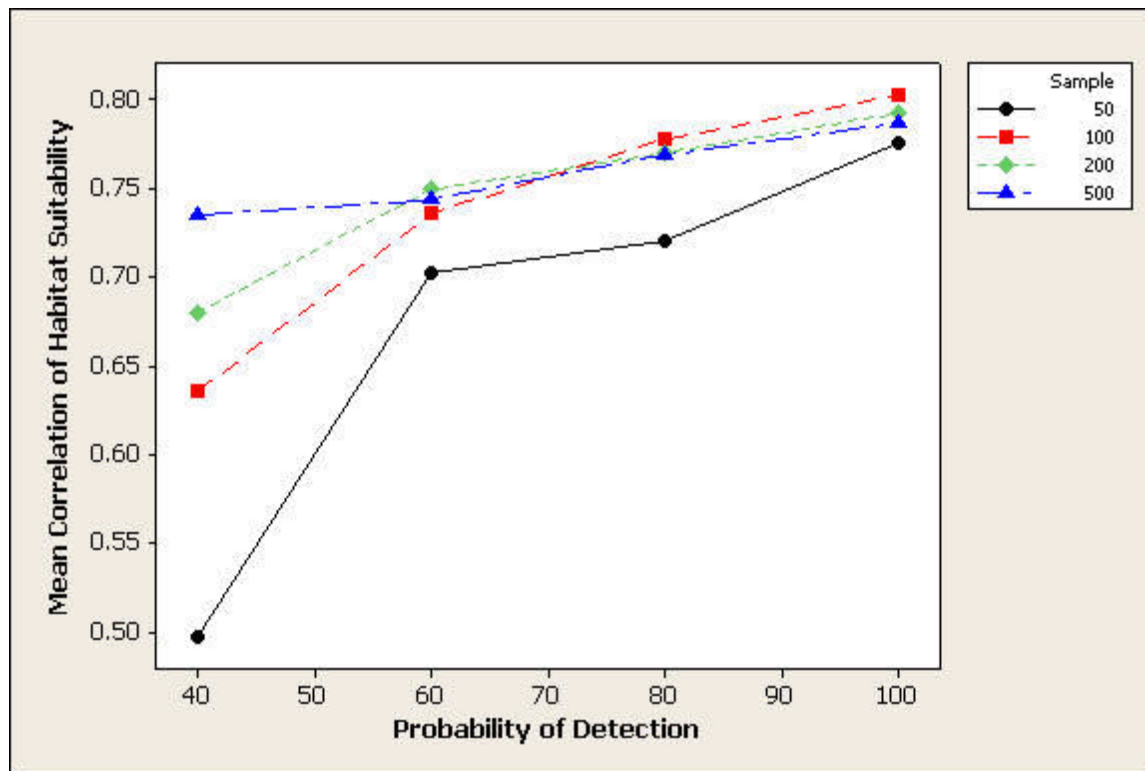


Figure 4. 4 Continued.

A. Bayesian Logistic Regression



B. ENFA

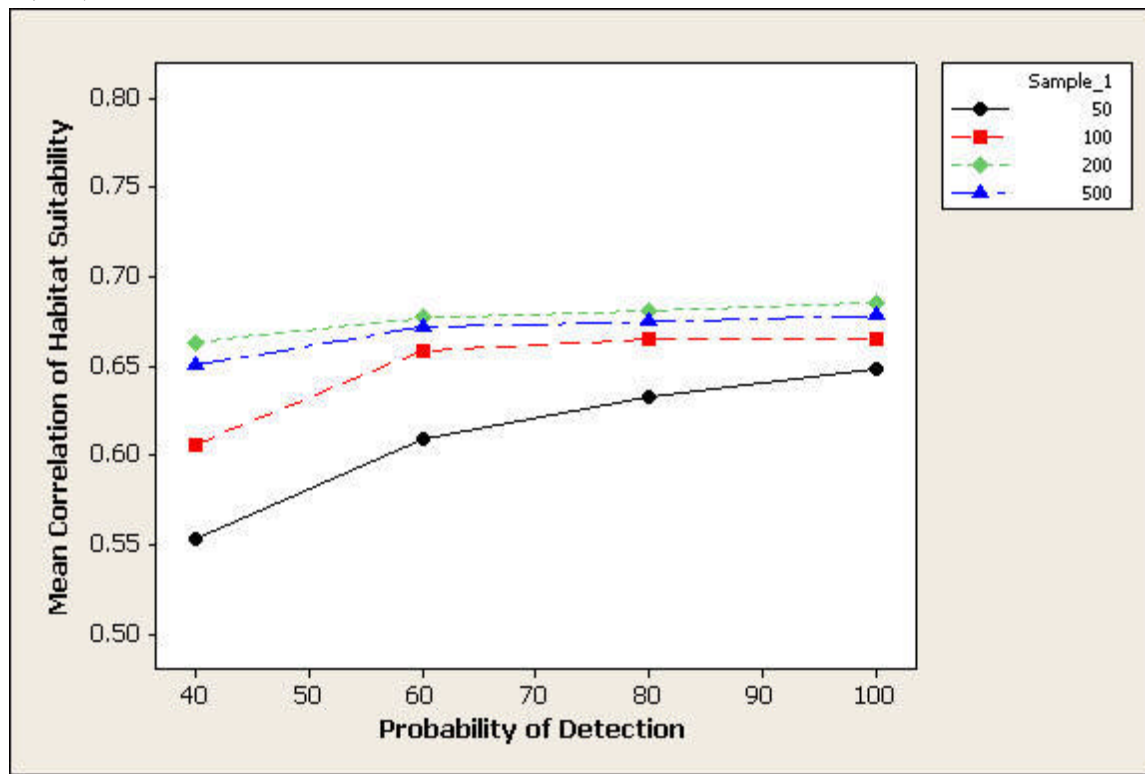
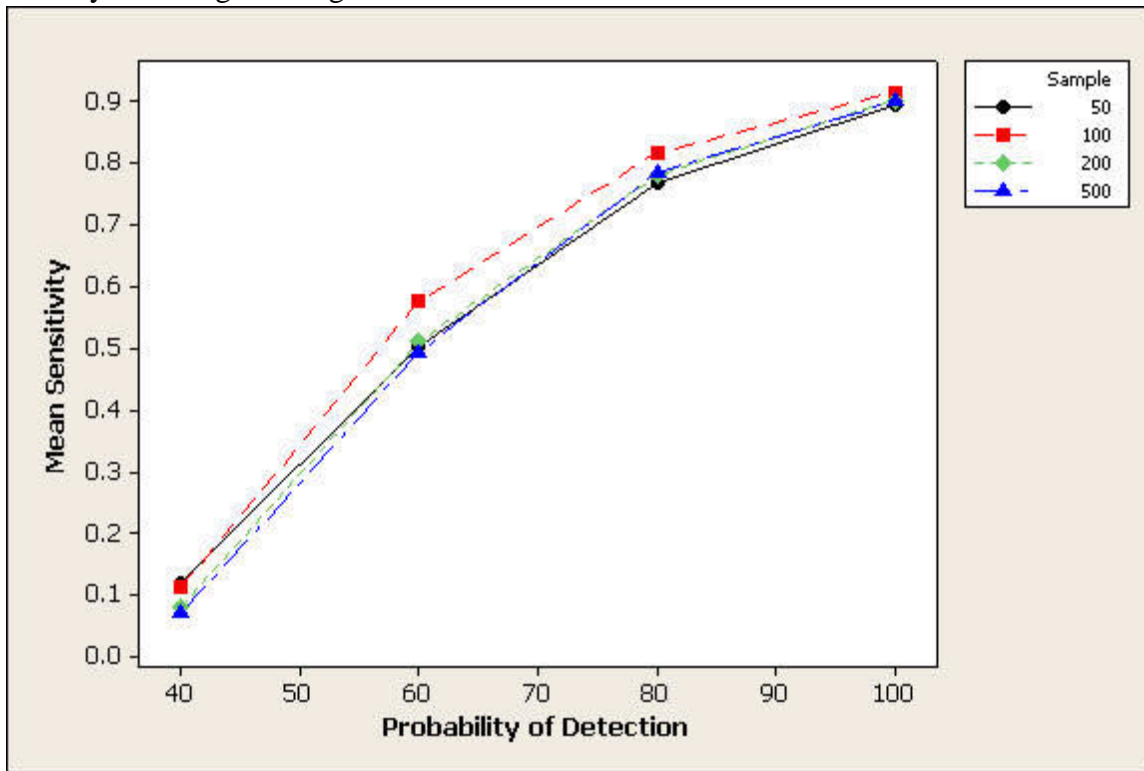


Figure 4. 5.

A. Bayesian Logistic Regression



B. ENFA

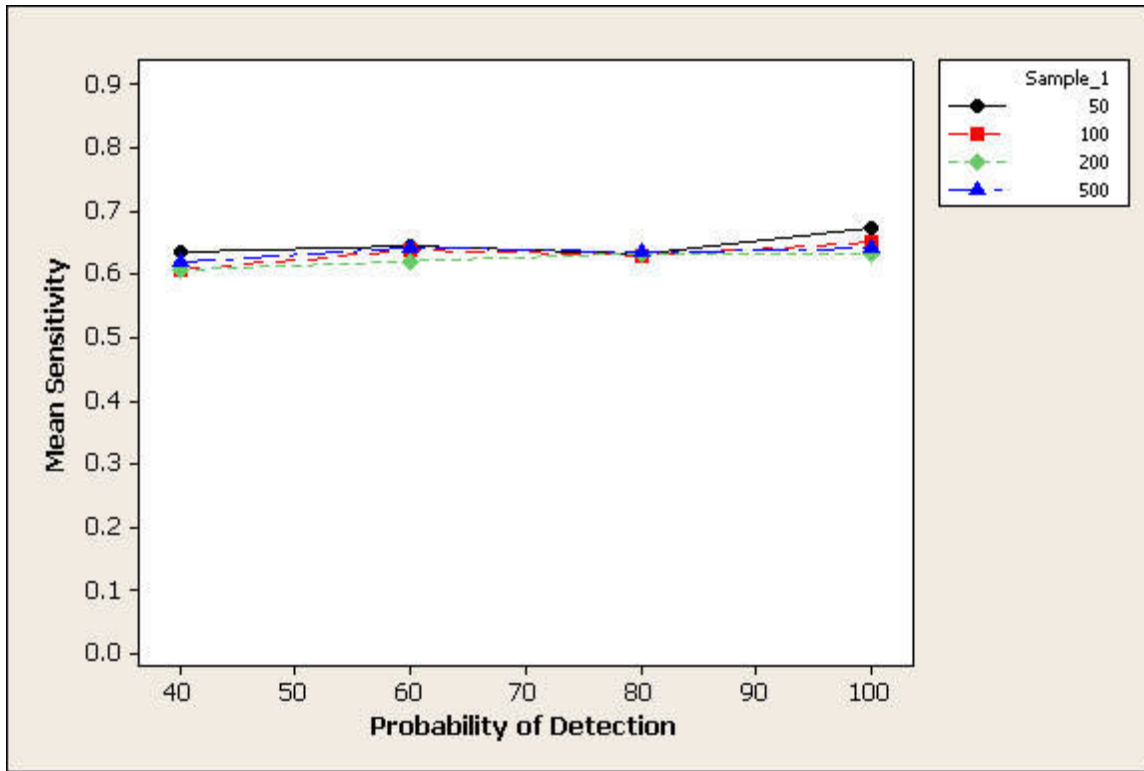
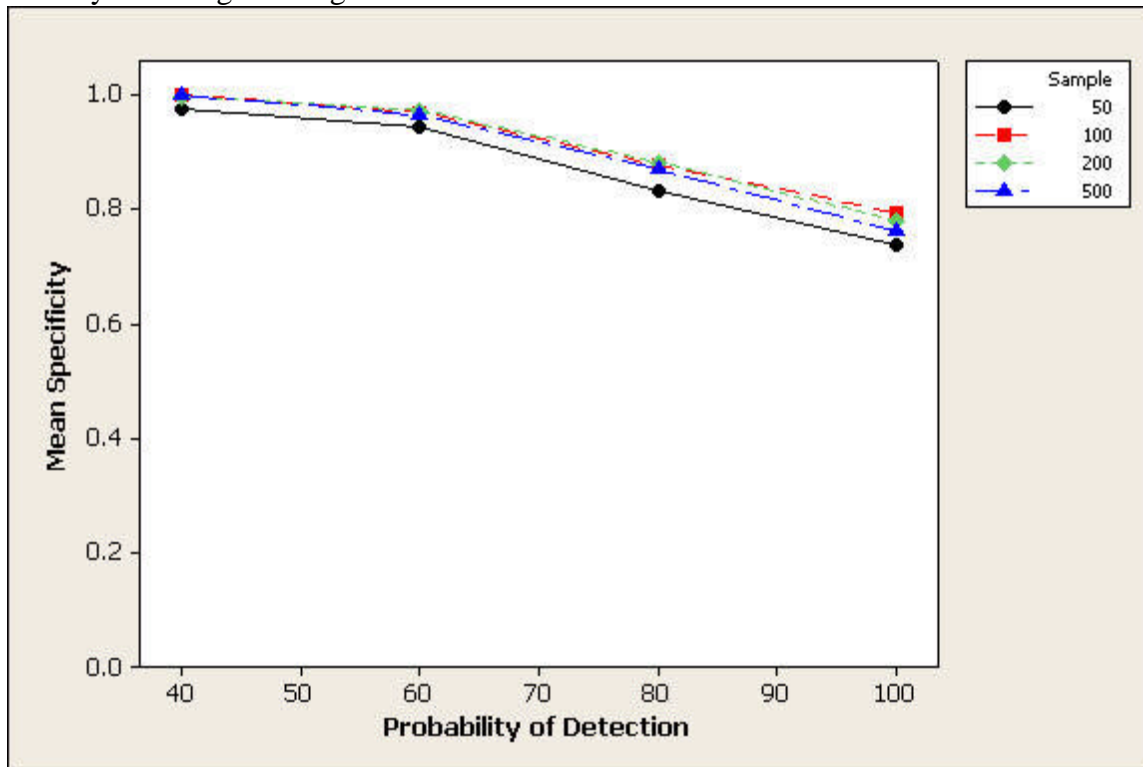


Figure 4. 6.

A. Bayesian Logistic Regression



B. ENFA

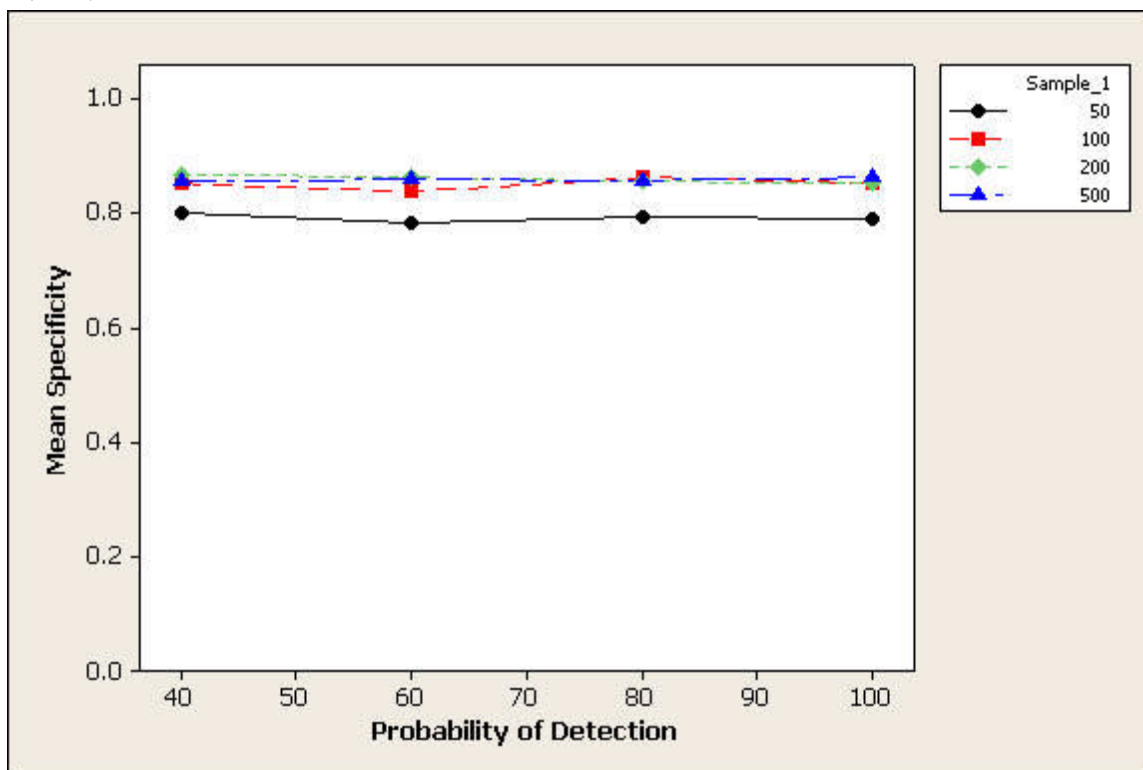


Figure 4. 7.

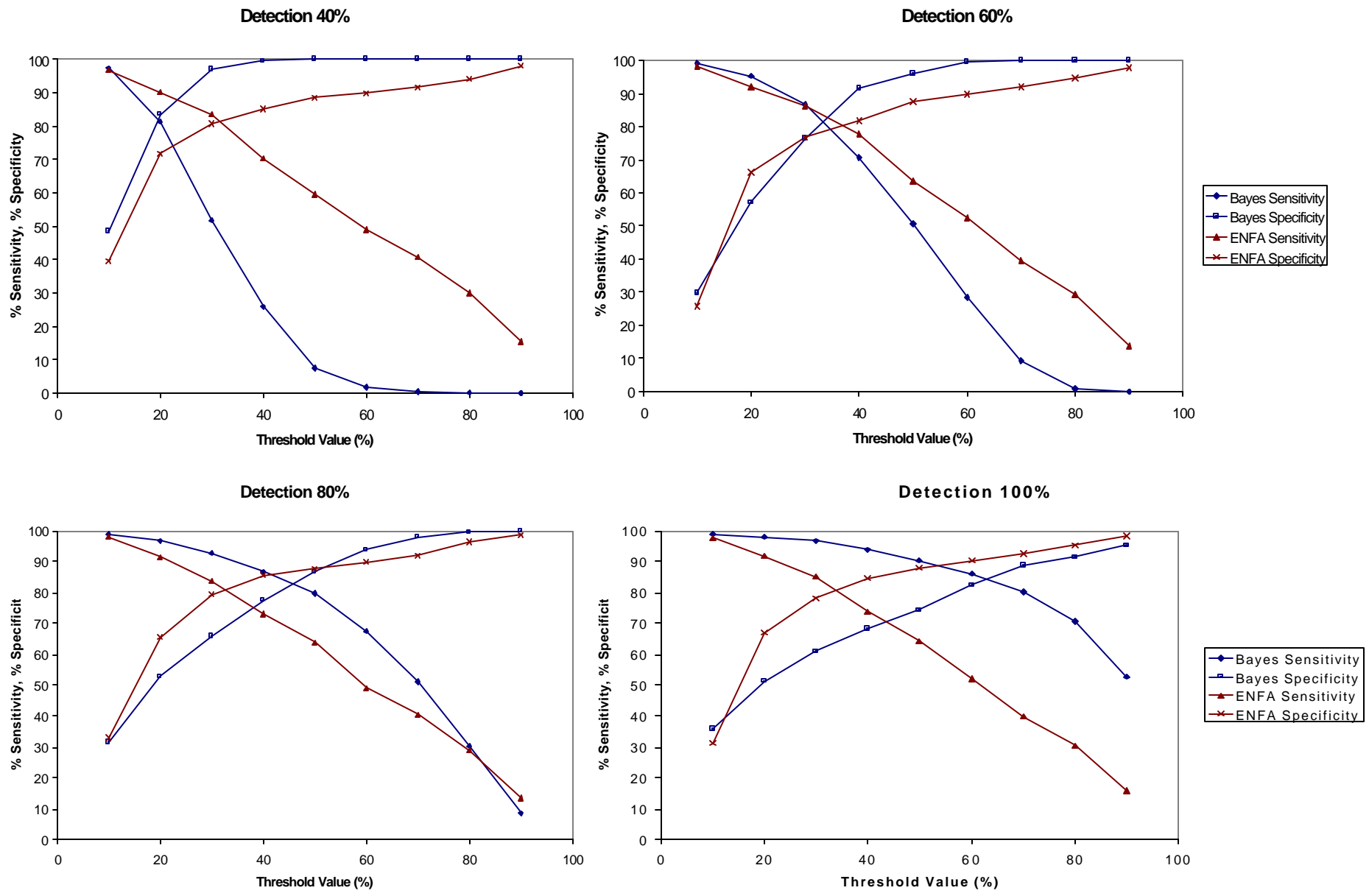


Figure 4. 8.

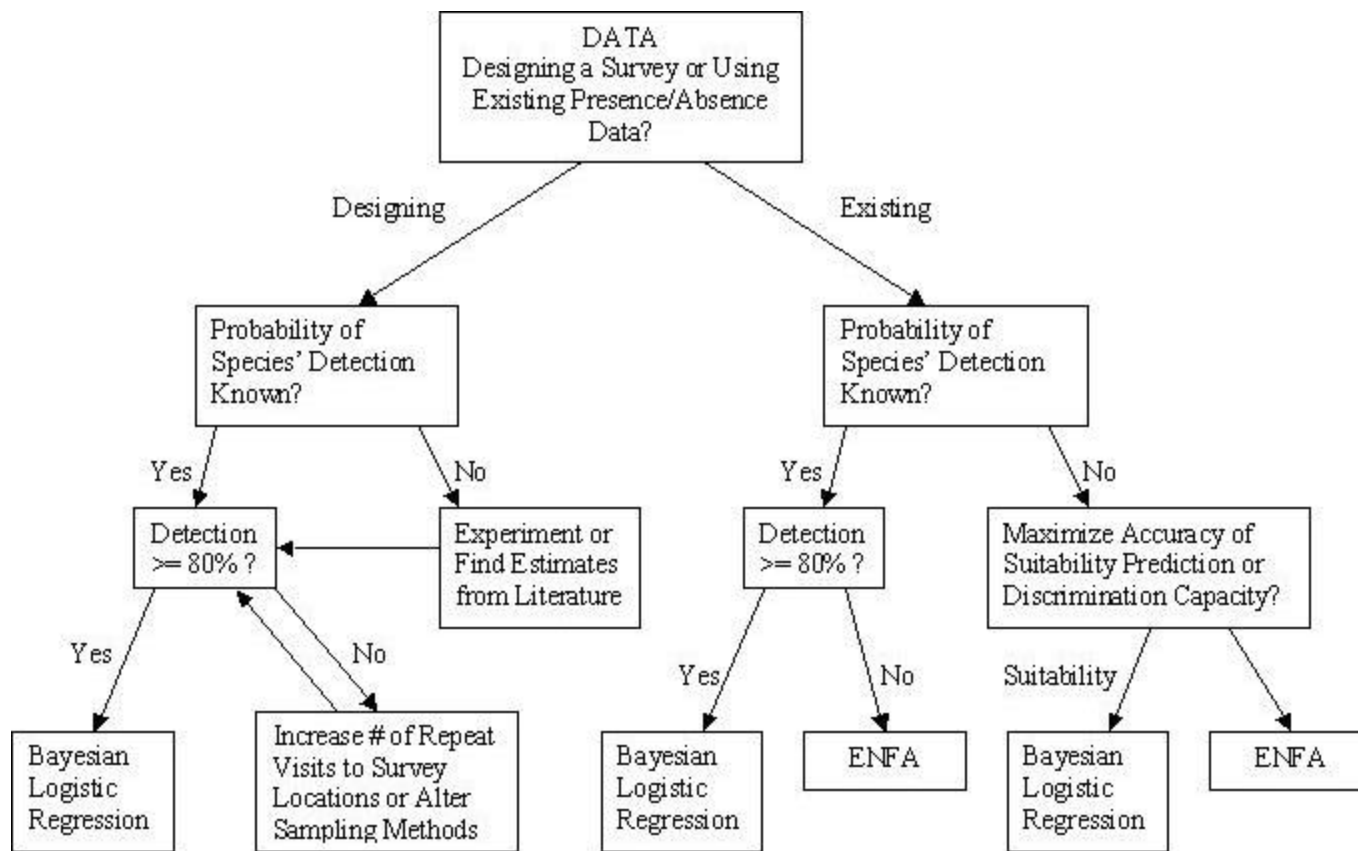


Figure 4. 9.

SUMMARY

Presence-absence data is often more available than abundance data for many wildlife species and wildlife scientists and managers frequently use presence-absence data to make conclusions and management decisions about wildlife populations. Habitat suitability models are often based on presence-absence data by using habitat data to best discriminate between known presence and absence locations. Frequently however, little consideration is given to the accuracy of the absence data. As a result of the enclosure experiment in this work, I have shown that the accuracy of absence data from wildlife surveys can vary based on the sampling method and density of the species of interest. Specific actions, such as adjusting sampling protocols based on weather, number of repeat visits, and size of the sampling area, have varying effectiveness at improving the probability of detecting the species depending on the sampling method.

Comparisons of modeling methods using the same species occurrence data set also verified the importance of having reliable absence data. Models built from large databases of species observations, where the probability of detection and sampling methods were unknown or inconsistent, had poor accuracy in discriminating between presence and absence locations regardless of the modeling method selected. The importance of knowing the probability of detection was illustrated with a simulation experiment. Models that were based on presence-absence data predicted suitability better and were more able to accurately discriminate between presence and absence locations when the probability of detecting the species was between 80-100%. However, when detection probabilities were lower, models that were based on presence-only data had the advantage of higher accuracy. This stresses the need for designing wildlife surveys protocols to obtain high levels of probability of detection. However, in the increasingly frequent position where previously collected data is the only option available, it is vital to know or estimate the probability of detection to be able to choose the best type of model.

Overall, this work finds that probability of detection is not an issue that can be overlooked when dealing with presence-absence data. Probability of detection should be considered when designing survey protocols and estimated by direct experiments or repeat sampling if unknown. If previously collected data is used for modeling purposes, modeling methods based on presence-absence data should be used if the probability of detection for the survey was high. If the

probability of detection was low or unknown, modeling methods based on presence-only data provide the most reliable results. However, databases of previously collected species occurrences are not the ideal data source for habitat modeling projects and specifically designed surveys with high probability of detection are encouraged for all future modeling projects.

VITA

Alison K. Williams was born on October 11, 1974 in San Antonio, Texas. She lived in Midlothian, Virginia until graduating from Midlothian High School in 1992. Alison attended college at Virginia Polytechnic Institute and State University from 1992 to 1996 and received a B.S. in Forestry and Wildlife Management. In 1996, Alison moved to Columbia, Missouri to attend graduate school at the University of Missouri. She completed her M.S. degree in Wildlife Ecology during 1999. During the summers between semesters, Alison worked as a field technician on many different projects in southwest Virginia and southeast and central Missouri. Alison returned to Virginia Polytechnic Institute and State University in 1999 to work with Dr. Jim Berkson and the Conservation Management Institute. In 2004, Alison completed her Ph.D. in Wildlife Science at Virginia Polytechnic Institute and State University.