Appendix AL

Joy, S.M. 2002. Northern Goshawk Habitat on the Kaibab National Forest Arizona: Factors Affecting Nest Locations and Territory Quality, Ph.d. Dissertation, University of Colorado, 223pp.

DISSERTATION

NORTHERN GOSHAWK HABITAT ON THE KAIBAB NATIONAL FOREST IN ARIZONA: FACTORS AFFECTING NEST LOCATIONS AND TERRITORY QUALITY

Submitted by

Suzanne Merideth Joy

Graduate Degree Program in Ecology

In partial fulfillment of the requirements

For the Degree of Doctor of Philosophy

Colorado State University

Fort Collins, CO

Fall 2002

COLORADO STATE UNIVERSITY

18 December 2001

WE HEREBY RECOMMEND THAT THE DISSERTATION PREPARED UNDER OUR SUPERVISION BY SUZANNE MERIDETH JOY ENTITILED NORTHERN GOSHAWK HABITAT ON THE KAIBAB NATIONAL FOREST IN ARIZONA: FACTORS AFFECTING NEST LOCATIONS AND TERRITORY QUALITY BE ACCEPTED AS FULFILLING IN PART REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY.

Committee on Graduate Work

Advisor

Co-Advisor

Department Head

ABSTRACT OF DISSERTATION

NORTHERN GOSHAWK HABITAT

ON THE KAIBAB NATIONAL FOREST IN ARIZONA:

FACTORS AFFECTING NEST LOCATIONS AND TERRITORY QUALITY

The northern goshawk has been at the center of controversy over its habitat management in the southwestern United States for the past decade. Fire suppression, grazing, and especially logging in coniferous forests are thought to be responsible for altering the composition and structure of forest habitat used by goshawks and their prey, thereby resulting in population declines. Conservation efforts for goshawks have focused on restoring forest structure to pre-settlement conditions using a combination of mechanical (logging) and natural (fire) methods.

Many studies of goshawk habitat requirements have focused on the hawks' habitat use, food habits, movements, distribution, demographics, and diets; however, no studies have attempted to use spatially explicit models to describe the spatial dynamics between goshawks and their environment, which is of critical importance to their survival. In this study, I develop a dynamic spatial simulation model to assess the spatial relationships between goshawk habitat composition and structure and the location of active nests (i.e., in which eggs were laid), and I examine the relationships between the amount and arrangement of habitat elements surrounding nests in high quality territories. Prior to assessing these relationships, I developed models of vegetation composition and

structure for the study area and assess the models' performance. The location of goshawk nests and assessments of territory quality estimated from reproductive performance are based on a population of banded goshawks on 101 territories studied from 1991 through 2000 on the Kaibab National Forest (North Kaibab Ranger District) in northern Arizona.

Suzanne Merideth Joy
Graduate Degree Program in Ecology
Colorado State University
Fort Collins, CO 80523
Fall 2002

DEDICATION

I dedicate this research to my parents, Sherry and Fred, for sharing their love of knowledge and writing with me, and to my husband, Vern, and son, Quinn, for their unconditional love and infinite patience throughout my graduate experience.

ACKNOWLEDGEMENTS

It takes a village to support a doctoral candidate, and I needed no less than that to complete this dissertation research. I am especially grateful to my family for their love and encouragement in my pursuit of this degree. My parents, Sherry and Frederick Joy, sisters, Stephanie, Michele and Ann Joy, and in-laws, Richard and Micki Dick, continually believed in my abilities and encouraged me to pursue my dreams. Vernon Thomas, my husband and friend, has been a constant source of support, love, inspiration, and understanding. I have learned much from his resiliency. Over the years, Vern has also provided many hours of GIS technical advice and assisted with customized ArcView (Avenue) applications. And, finally, I am profoundly grateful to my son, Quinn, for reminding me that family life is as, if not more, important to my well-being than life outside the home. Quinn's unconditional love, enthusiasm, and inquisitive nature make every day a joyous occasion.

My committee members have been a valuable source of inspiration and direction throughout my study, each playing a role in my education and professional growth as well. I thank Phil Chapman for challenging and improving my knowledge of multivariate statistical techniques, and providing other statistical advice on numerous occasions. Melinda Laituri, my academic advisor, has been a mentor and friend, encouraging me when I was frustrated and providing me with objectivity when sorely needed. Robin Reich, my research advisor, not only introduced me to spatial modeling, but has also been a friend. Robin guided my spatial modeling efforts throughout this project and was

frequently available during non-working hours to discuss spatial modeling concepts and analytical techniques. I am grateful to Robin for teaching me how to see research problems in a spatially explicit manner. Richard Reynolds, Research Wildlife Biologist with the USDA Forest Service (USFS) Rocky Mountain Research Station (RMRS), continually challenged my knowledge of goshawk biology and forest ecology. I am especially grateful to Richard for having the courage, fifteen years ago, to hire a laboratory microbiologist to conduct a field-intensive study of *Accipiter* hawks, forever changing the direction of my life. During the course of this dissertation work, Richard not only kept me employed, but also provided me with the opportunity and resources to pursue my doctoral degree.

Rudy King, Biometrician for the RMRS, furthered my understanding of clustering techniques, the use of mixed model analysis of variance, and logistic regression. My appreciation for Rudy's statistical advice and reviews of the Chapters 2-4 of this dissertation is immense. Brian Cade, Statistician with the USGS Fish and Wildlife Service, provided me with much useful insight into habitat modeling and multiresponse permutation procedures, as well as reviewing an earlier draft of Chapter 3 of the dissertation.

A field-based project of this scope and duration cannot be carried out alone. Many individuals contributed to the daily logistics, data collection, entry and verification, and otherwise day-to-day activities of this long-term research effort. I am particularly grateful to Shelley Bayard de Volo, Donna Laing, Jeffrey Lambert, Susan Salafsky, John Seyfried, and David Wiens for their technical and personal support, as well as their friendship. Many others employee and volunteers also devoted several years to

monitoring the goshawk population on the Kaibab Plateau and their contributions cannot be understated.

The support, encouragement, and feedback I received from my friends made each day's triumphs more rewarding and setbacks more bearable. I would especially like to thank Brian Linkhart, whose camaraderie and positive outlook buoyed my morale on many occasions, and whose discussions on demography and GIS inspired me to greater levels of problem solving. I am grateful to Michael Williams for reviewing earlier drafts of Chapters 1-3, as well as for endless discussions of statistical techniques and parenting, all of which I enjoyed greatly. Others who were a constant source cheer, advice, and support were Amy Cade, Robert (Rocky) Coleman, Michelle and Jerry Davis, Sandra Haire, Charles Johnson, Pamela Kaval, Mary Jane McCool, and Davol Tedder.

This project was funded by the USFS RMRS, USFS Region 3, and the Arizona Department of Game and Fish. Housing and logistical support during the field portion of this study were provided by the USFS North Kaibab Ranger District (Fredonia, Arizona).

TABLE OF CONTENTS

Page

ABSTRACT OF DISSERTATION....iii

DEDICATION....v

ACKNOWLEDGEMENTS....vi

LIST OF FIGURES....xii

LIST OF TABLES....xvi

INTRODUCTION TO DISSERATION....1
- WILDLIFE HABITAT RELATIONSHIPS....1
- THE NORTHERN GOSHAWK....2
- MANAGEMENT ON THE KAIBAB NATIONAL FOREST....3
- OVERVIEW OF DISSERTATION....4
- LITERATURE CITED....8

CHAPTER ONE: A NON-PARAMETRIC, SUPERVISED CLASSIFICATION OF VEGETATION TYPES ON THE KAIBAB NATIONAL FOREST USING DECISION TREES....10
- ABSTRACT....10
- INTRODUCTION....11
- STUDY AREA....13
- METHODS....16
  - Data Collection....16
    - Field Data....16
      - Spectrally-derived plots....16
      - Goshawk nest plots....17
      - Randomly located plots....17
    - Landsat TM and GIS Data....19
    - Photographic Data....20
  - Image Classification....20
  - Accuracy Assessment....23
- RESULTS....25
  - Image Classification....25
  - Accuracy Assessment....30
- DISCUSSION AND CONCLUSIONS....36

Page

ACKNOWLEDGEMENTS....38
LITERATURE CITED....38
APPENDIX – Application of the multivariate composite estimator used to assess classification accuracy....44

CHAPTER TWO: MODELING FOREST STRUCTURE ON THE KAIBAB NATIONAL FOREST IN ARIZONA....46
ABSTRACT....46
INTRODUCTION....47
STUDY AREA....49
METHODS....51
Data Collection....51
Field Data....51
GIS and Landsat TM Data....54
Modeling Forest Structure....55
Model Evaluation....59
Estimating Means and Variances....62
RESULTS....64
Modeling Forest Structure....64
Model Evaluation....75
Estimating Means and Variances....82
DISCUSSION AND CONCLUSIONS....85
ACKNOWLEDGEMENTS....89
LITERATURE CITED....90

CHAPTER THREE: PREDICTING THE LOCATION OF NORTHERN GOSHAWK NESTS: MODELING THE SPATIAL DEPENDENCY BETWEEN NEST LOCATIONS AND FOREST STRUCTURE....96
ABSTRACT....96
INTRODUCTION....97
STUDY AREA....99
METHODS....102
The Data....102
Goshawk Nest Locations....102
Field Data....103
Goshawk nest plots....103
Randomly located plots....103
GIS and Landsat TM Data....103
Field Measurements....104
Spatial Distribution of Active Goshawk Nest....105
Gibbsian Pairwise Potential Model....108
Potential Energy of Goshawk Nests....109
Model Parameter Estimation....110

Page

Potential Energy Between Nests and Forest Structure.....112
Modeling Nest Site Suitability.....113
Simulating the Spatial Distribution of Goshawk Nests.....115
RESULTS.....118
Modeling Nest Site Suitability.....118
Trends in Nest Habitat Use.....121
Nest Habitat.....121
Simulating the Spatial Distribution of Nests.....129
DISCUSSION AND CONCLUSIONS.....135
ACKNOWLEDGEMENTS.....144
LITERATURE CITED.....145

CHAPTER FOUR: CORRELATES OF NORTHERN GOSHAWK TERRITORY QUALITY.....153
ABSTRACT.....153
INTRODUCTION.....154
STUDY AREA.....157
METHODS.....160
Terminology.....160
Population Sampling.....161
Qualitative Ranking of Territories.....162
Goshawk Age and Experience.....165
Habitat Modeling.....167
Correlating Habitat Quality with Environmental Variables.....168
Predicting the Location and Quality of Territories.....172
RESULTS.....174
Population Sampling.....174
Qualitative Ranking of Territories.....175
Goshawk Age and Experience.....175
Correlating Habitat Quality with Environmental Variables.....179
Predicting the Location and Quality of Territories.....200
DISCUSSION AND CONCLUSIONS.....203
MANAGEMENT IMPLICATIONS.....208
ACKNOWLEDGEMENTS.....209
LITERATURE CITED.....210

DISSERTATION SUMMARY.....220
OVERVIEW.....220
FUTURE WORK.....222

LIST OF FIGURES

Figure Page

I.1 Overview of the process leading to the dynamic spatial simulation used to predict the location of active northern goshawk nests and the assessment of territory quality.....5

1.1 1997 Landsat TM image (bands 4, 3, and 2) of the Kaibab National Forest (North Kaibab Ranger District), Arizona, above 2182 m in elevation, showing spectrally-derived, random, and northern goshawk nest plots.....15

1.2 Plot layout used to sample vegetative types on the Kaibab National Forest (North Kaibab Ranger District), Arizona, and the relationship of the plot to Landsat TM imagery.....18

1.3 Decision tree classifier for ponderous pine and other dominant vegetative types on the Kaibab National Forest (North Kaibab Ranger District), Arizona.....26

1.4 Decision tree classifier for dominant vegetative types other than pure ponderosa pine on the Kaibab National Forest (North Kaibab Ranger District), Arizona.....28

1.5 Dominant vegetative types on the study area in Arizona.....31

2.1 Location and map of the study area on the Kaibab National Forest (North Kaibab Ranger District) in Arizona, showing the distribution and spatial arrangement of sample plots.....50

2.2 Layout of the field plots used on the Kaibab National Forest (North Kaibab Ranger District), Arizona, showing the arrangement of sample sub-plots.....53

2.3 Predictions of forest structure for percent canopy closure, total basal area ($m^2 ha^{-1}$), proportion of ponderosa pine, and proportion of spruce and fir basal areas on the Kaibab National Forest (North Kaibab Ranger District) in Arizona.....67

Figure Page

2.4 Predictions of forest structure for the proportion of aspen basal area, maximum understory height (m), and presence of seedling and/or sapling trees on the Kaibab National Forest (North Kaibab Ranger District) in Arizona....68

2.5 Prediction error from cross-validation of the multiple linear regression models for percent canopy closure, total basal area, proportion of ponderosa pine and spruce-fir basal areas, and maximum height of the understory vegetation....73

2.6 Semi-variogram plots for residuals from the multiple linear regression models of percent canopy closure, total basal area, proportion of ponderosa pine and spruce-fir basal areas, and maximum height of the understory vegetation, and the logistic regression of seedling and/or sapling presence....76

2.7 Histograms showing the distribution of errors from cross-validation of the models of percent canopy closure, total basal area, proportion of ponderosa pine and spruce-fir basal areas, and maximum height of the understory vegetation....81

2.8 Receiver operated characteristic curve used to discriminate between the presence and absence of seedlings and/or sapling trees....83

3.1 Distribution and arrangement of nest plots and random plots used to model forest structure displayed among dominant vegetation classes on the North Kaibab Ranger District, Kaibab National Forest, Arizona....100

3.2 Bounded region (B) showing the relative location of 27 active northern goshawk nests from 1998 used to model the spatial relationship between active nests and forest structure....106

3.3 The location of active northern goshawk nests between 1991 and 1998 on the North Kaibab Ranger District, Kaibab National Forest, Arizona....107

3.4 Spatial distribution of estimated "good" and "poor" locations for goshawk nests on the North Kaibab Ranger District, Kaibab National Forest, Arizona, and all nests active between 1991 and 1998....119

3.5 Plot of the transformed K-function, $L(h) = \{K(h)/\pi\}^{1/2}$, against distance h, used to model the spatial arrangement of individual northern goshawk nests on the bounded region (B) on the North Kaibab Ranger District, Kaibab National Forest, Arizona, for 200 realizations of a spatial Poisson process....130

Figure Page

3.6 Plot of the fitted pairwise potential model for individual northern goshawk nests on the bounded region (B) on the North Kaibab Ranger District, Kaibab National Forest, Arizona.....132

3.7 Plot of the transformed K-function, $L(h) = \{K(h)/\pi\}^{1/2}$, against distance h, used to model the spatial arrangement of individual northern goshawk nests on the study area on the North Kaibab Ranger District, Kaibab National Forest, Arizona, for 200 realizations of the (a) nest component of the point process model, (b) forest component of the point process model, (c) point process model that takes into consideration the territoriality between individual active nests and forest structure.....134

3.8 Realization of the point process model within the bounded region (B) that takes into consideration the territoriality between individual northern goshawk nests and forest structure on the North Kaibab Ranger District, Kaibab National Forest, Arizona, shown along with the locations of the 27 active northern goshawk nests used in fitting the model.....136

3.9 Realization of the point process model that takes into consideration the territoriality between individual northern goshawk nests and forest structure on the North Kaibab Ranger District, Kaibab National Forest, Arizona.....137

4.1 Location of the study area ≥ 2,182 m above mean sea level on the North Kaibab Ranger District (hatched region) of the Kaibab National Forest in Arizona, showing random plots and goshawk territory centroids.....158

4.2 Configuration of circle and ring plots used to assess the amount and spatial arrangement of habitat components, respectively, around random plots and northern goshawk territories demonstrating higher and lower reproductive performance on the Kaibab National Forest (North Kaibab Ranger District), Arizona.....169

4.3 Clusters distinguishing higher from lower quality northern goshawk territories on Kaibab National Forest (North Kaibab Ranger District) in Arizona.....176

4.4 Distribution of higher and lower quality northern goshawk territories on the Kaibab National Forest (North Kaibab Ranger District), Arizona.....177

Figure Page

4.5 Significant relationships from analysis of variance tests of the amount (proportion) of habitat components around nests on northern goshawks territories of higher and lower quality and random plots on the Kaibab National Forest (North Kaibab Ranger District) in Arizona.....184

4.6 Significant relationships in analysis of variance tests of the spatial arrangement of habitat components around nests on northern goshawks territories of higher and lower quality and random plots on the Kaibab National Forest (North Kaibab Ranger District) in Arizona.....185

S.1 Overview of the dynamic spatial simulation model that includes predictions of northern goshawk habitat quality and changes in forest structure due to the implementation of management alternatives on the Kaibab National Forest (North Kaibab Ranger District) in Arizona.....223

LIST OF TABLES

Table Page

1.1 Modified land cover classification system (Anderson et al., 1976) for vegetation types on the Kaibab National Forest (North Kaibab Ranger District), Arizona....22

1.2 Resubstitution classification error rates of the decision tree for the vegetative types on the Kaibab National Forest (North Kaibab Ranger District), Arizona....29

1.3 Distribution of sample plots by vegetative type on the Kaibab National Forest (North Kaibab Ranger District), Arizona....32

1.4 Joint probability error matrix for the vegetative types and photo-interpretation of 767 sample plots on the Kaibab National Forest (North Kaibab Ranger District), Arizona....33

1.5 Joint probability error matrix for the decision tree for vegetative types and field data on the Kaibab National Forest (North Kaibab Ranger District), Arizona....35

2.1 Summary statistics showing the range of observed topographic and remotely sensed data from 1556 sample subplots on the Kaibab National Forest (North Kaibab Ranger District), Arizona....65

2.2. Topographic characteristics, Landsat TM data, and vegetative classes used to develop the multiple linear regression models of forest structure and a logistic regression model of the presence of seedlings and/or saplings on the Kaibab National Forest (North Kaibab Ranger District), Arizona....70

2.3 Topographic characteristics, Landsat TM data, and vegetation classes used in regression trees to describe the residuals in the multiple linear regression models of forest structure on the Kaibab National Forest (North Kaibab Ranger District), Arizona....72

2.4 Overall effectiveness of the multiple linear regression models and regression trees used to describe continuous variables of forest spatial structure on the Kaibab National Forest (North Kaibab Ranger District), Arizona....74

Table Page

2.5 Summary statistics and prediction bias showing the level of agreement between observed and estimated models of forest structure on the Kaibab National Forest (North Kaibab Ranger District), Arizona.....77

2.6 Summary statistics for the estimation errors of the models of forest structure on the Kaibab National Forest (North Kaibab Ranger District), Arizona, based on 10-fold cross-validation of the models.....80

2.7 Accuracy assessment for the logistic regression model and cross-validation predicting the presence of seedling and/or sapling trees on the Kaibab National Forest (North Kaibab Ranger District), Arizona.....84

2.8 Comparison of sample statistics for estimating average canopy closure on the Kaibab National Forest (North Kaibab Ranger District), Arizona, using field data and the canopy model.....86

3.1 Distribution of estimated good and poor northern goshawk nest habitat by vegetation class on the North Kaibab Ranger District, Kaibab National Forest, Arizona.....120

3.2 Total number of territories and active northern goshawk nests between 1991 and 1998 above 2,182 m in elevation on the North Kaibab Ranger District, Kaibab National Forest, Arizona.....122

3.3 Number of active nests between 1991 and 1998 by estimated suitability (good, poor) of nest locations and vegetative class on the North Kaibab Ranger District, Kaibab National Forest, Arizona.....123

3.4 Standardized regression coefficients for variables that maximize the likelihood of a northern goshawk nest occurring in a vegetation class on the North Kaibab Ranger District, Kaibab National Forest, Arizona.....124

3.5 Means for variables that maximize the likelihood of a northern goshawk nest occurring in a vegetation class on the North Kaibab Ranger District, Kaibab National Forest, Arizona.....126

3.6 Results of the Cramér von-Mises goodness-of-fit test used to test the null hypothesis that northern goshawk nests in 1998 were randomly distributed on the North Kaibab Ranger District, Kaibab National Forest, Arizona.....131

3.7 Distribution of probabilities of finding a northern goshawk nest associated with predicted and observed nest points on the North Kaibab Ranger District, Kaibab National Forest, Arizona.....138

Table Page

4.1 Fisher's Exact Tests of the number of young northern goshawks (banded recruits and unbanded hawks) that joined the breeding population on 101 territories on the Kaibab National Forest (North Kaibab Ranger District), Arizona, evaluated between levels of territory quality (higher, lower).....178

4.2 Results of ANOVA tests for differences in the amount of vegetative types and diversity of vegetation around plots at northern goshawk nets and random locations on the Kaibab National Forest (North Kaibab Ranger District) in Arizona.....180

4.3 Results of ANOVA tests for differences in the spatial arrangement of vegetative types and diversity of vegetation around plots at northern goshawk nets and random locations on the Kaibab National Forest (North Kaibab Ranger District) in Arizona.....182

4.4 Proportions of vegetative types and diversity for higher and lower quality northern goshawk territories and random locations on the Kaibab National Forest (North Kaibab Ranger District) in Arizona.....187

4.5 Proportion of vegetative types and diversity for higher and lower quality northern goshawk territories and random locations by plot size on the Kaibab National Forest (North Kaibab Ranger District) in Arizona.....189

4.6 Results of MRPP tests of the distribution in the proportion of vegetative types and vegetative diversity for higher and lower quality goshawk territories and random locations for all circle or ring plot sizes considered simultaneously, on the Kaibab National Forest (North Kaibab Ranger District), Arizona.....194

4.7 Results of univariate MRPP tests of the distribution in the proportion of vegetative types and vegetative diversity for higher and lower quality goshawk territories and random locations for each circle and ring plot size on the Kaibab National Forest (North Kaibab Ranger District), Arizona.....197

4.8 Significant vegetative coefficients for predicting the location of northern goshawk territories vs. random sites on the Kaibab National Forest (North Kaibab Ranger District) in Arizona fitted by logistic regression.....201

4.9 Significant habitat coefficients for predicting higher (vs. lower) quality nests for northern goshawks on the Kaibab National Forest (North Kaibab Ranger District) in Arizona fitted by logistic regression.....202

INTRODUCTION TO DISSERTATION

WILDLIFE-HABITAT RELATIONSHIPS

Integral to any study of wildlife species are the relationships that exist between the species and its habitat. This relationship provides insights into the distribution and abundance of individuals, and can be used to predict species occurrence and responses to habitat changes on the landscape. On a broader level, the understanding of wildlife-habitat relationships and the functional role a species plays in its environment provide a knowledge platform for maintaining biologically diverse and viable ecosystems. This understanding is also essential to managing landscapes for population persistence. Most wildlife species occupy a range of habitats that differ in the resources necessary to reproduce and survive. Typically, animals settle into high quality habitats first and, as these fill, settle into progressively lower quality habitats. Whereas high quality habitats confer sufficient survival and reproduction for population persistence or growth, low quality habitats may not. In the latter, populations may not produce sufficient young to balance mortality. Measures of habitat quality should therefore be based on the vital rates (reproduction and/or survival) of individuals in the population expressed over their lifetimes. In this dissertation, I used spatial modeling techniques, to characterize a species' environment and explore the relationships between the species and its habitat in order to answer fundamental questions about factors affecting the species' distribution, abundance, population performance. These issues, and their management implications, are outlined in the following chapters of the dissertation.

THE NORTHERN GOSHAWK

The northern goshawk (*Accipiter gentilis atricapillus*; hereafter referred to as goshawk) is a forest-dwelling hawk that occurs throughout North America, Asia, and Europe (Johnsgard 1990, Squires and Reynolds 1997). In North America, it breeds in boreal and temperate forests, using a variety of forest types and structures for nesting and foraging. Goshawks prey on small to medium-sized birds and mammals, such as song birds woodpeckers, jays, game birds, chipmunks, squirrels, and rabbits. Its long tail, short rounded wings, and short perch/short flight foraging tactics are morphological and behavioral adaptations for living in vegetation-filled environments.

During the past decade, the goshawk has become the focus of research in forests of North America due to apparent conflicts between population performance and forest management practices (Reynolds 1983, 1989, Crocker-Bedford 1990, Reynolds et al. 1992, Kennedy 1997, Peck 2000). Habitat loss due to changes in forest structure and composition (Reynolds 1983, 1989, Speiser and Basakowski 1984, Crocker-Bedford 1990, Reynolds et al. 1992) are thought cause population declines. As a result, goshawks are considered a "sensitive" species by most National Forests, "special status species" by most National Parks, and "Category II" species by the Fish and Wildlife Service (FWS). Since 1992, the FWS has been petitioned three times to list the goshawk as "threatened."

Because goshawks occupy a variety of forest types that differ in the types and availabilities of resources, one would expect differential vital rates among pairs occupying the range of habitat conditions. This association has particular import for goshawks because they display high site fidelity (Reynolds et al., in review); once they choose a breeding location, they rarely change sites. This behavioral attribute has large ecological and

management implications because individual pairs may not respond to habitat alterations by moving to new breeding sites. Population persistence should be evaluated on the demographic performance of individual pairs, rather than on presence alone.

MANAGEMENT ON THE KAIBAB NATIONAL FOREST

The study area is the Kaibab National Forest's (KNF) North Kaibab Ranger District (NKRD) on the Kaibab Plateau in Arizona. The forest on the study area is composed of ponderosa pine, mixed conifer, spruce, fir, and aspen trees, all of which are used by the goshawk and its prey. Single-tree selection cutting began on the NKRD in the 1920s. This cutting practice continued, along with small (0.1 km^2) clearcuts in the central part of the Kaibab Plateau, until the late-1970s when intensive stand-level management began. Shelterwood, seed, salvage, removal, and thinning cuts persisted until 1991 when the NKRD implemented forest treatments designed to restore the habitat of goshawks and their prey (Reynolds et al. 1992). Livestock grazing was common on the NKRD between the late 1800s and the mid-1920s. Reduction in the frequency of natural ground fires since the early 1900s through road building, grazing, and fire suppression has affected the composition and structure of all forests on the NKRD by allowing the regeneration of pine and shade-tolerant tree species; therefore, much of the forest on the NKRD has been altered.

Some management-related changes in the forest appear to have affected the distribution and abundance of goshawks by decreasing the amount of available habitat or by altering the composition and abundance of prey resources. For example, the Kaibab squirrel (*Sciurus aberti kaibabensis*), a primary goshawk prey species, depends on ponderosa pine for food (cambium, apical buds, cones) and nest sites (Hall 1981). Lack of sufficient ponderosa

pine regeneration in mixed-conifer forests resulted in a gradual decline of this tree, loss of the squirrel in these forests, and a reduction of the carrying capacity for goshawks.

The management guidelines (Management Recommendations for the Northern Goshawk in the Southwestern United States, Reynolds et al. 1992) by which forests on the NKRD are currently governed describe desired forest conditions within three spatially explicit areas of the goshawk's nesting home range (24.3 km^2): the nest (0.12 km^2), post-fledging family (1.70 km^2), and foraging (21.9 km^2) areas. These areas encompass the movements and behavior of goshawks associated with nesting, pre-independence of fledglings, and foraging, respectively. The guidelines advocate thinning and prescribed fire to achieve landscape-scale changes in the vegetation that simulate natural ecological processes that support the goshawk and its prey.

OVERVIEW OF DISSERTATION

The overall objective of my dissertation was to assess the relationship between the abundance, distribution and demographic performance of goshawks, and the range of habitat conditions supporting the population (Fig. I.1). I accomplish this by first developing fine-scale (10-m x10-m spatial resolution) habitat models for the study area. The spatial arrangement of vegetative components (forest species composition and structure) that characterize the goshawk's habitat and the distribution of nests on the study area are described using a Gibbsian pairwise potential model. The vegetation is then linked to the spatial arrangement of paired individuals using a dynamic spatial simulation model to assess the influence of habitat and territoriality on goshawk distribution. Finally, I evaluate the habitat correlates of territory quality. The goshawk population of interest occurs in northern

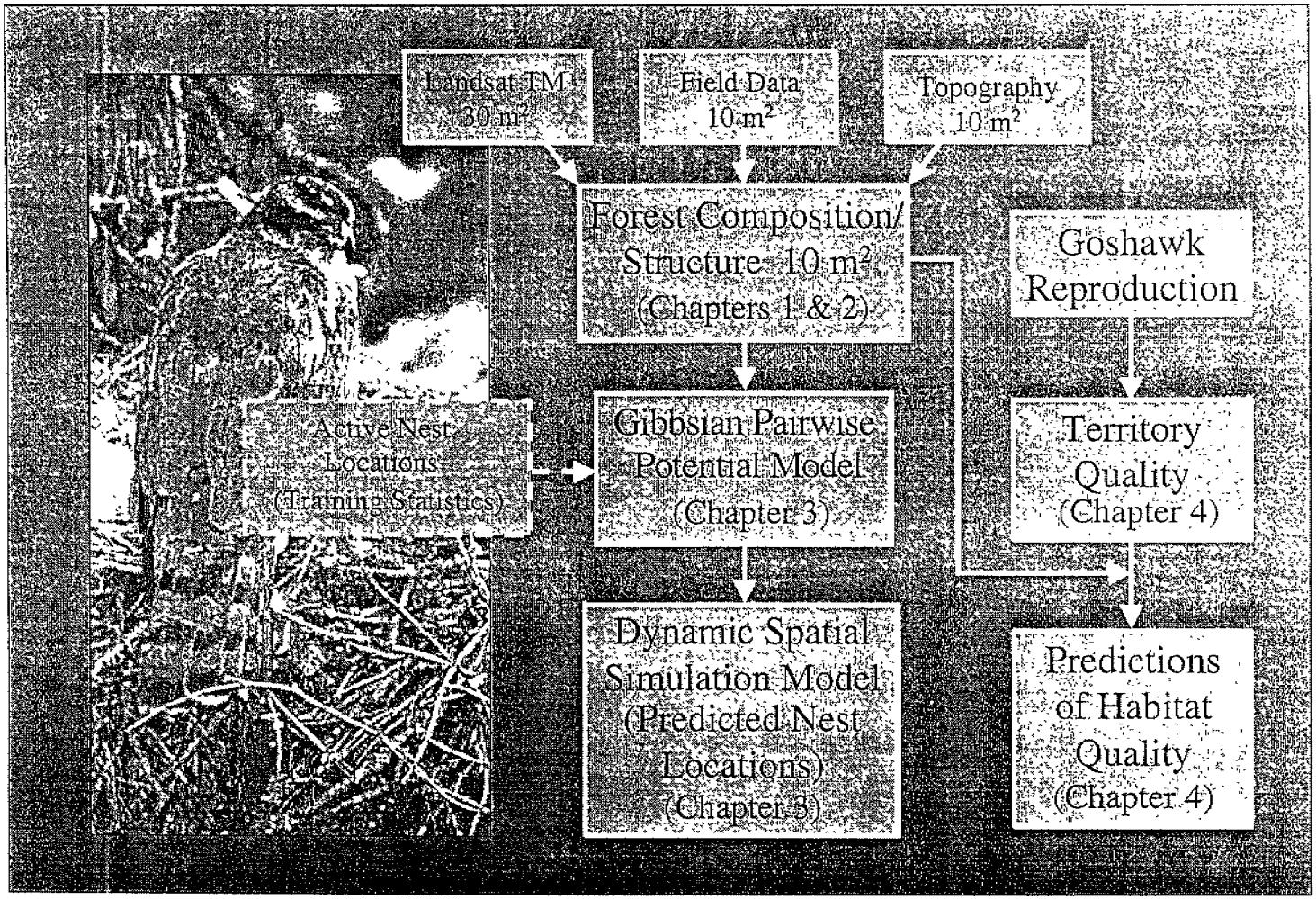


Figure I.1. Overview of the process leading to the dynamic spatial simulation used to predict the location of active northern goshawk nests and the assessment of territory quality.

BEST AVAILABLE

Arizona on 1,281 km^2 of the NKRD above 2,182 m in elevation. This population is considered relatively closed, with little emigration or immigration of adult, banded goshawks (Reynolds et al., in review).

Characterizing goshawk habitat (Chapters 1, 2) required developing habitat models for the study area at a spatial resolution meaningful to the grain of interest, i.e., that of nest sites/areas. Chapters 1 and 2 are devoted to modeling forest species composition and structure, respectively, to a 10-m spatial resolution, reflecting nearly a century of management history on the NKRD. In Chapter 1, I used decision trees to model the vegetative types. The models are based on field data, remotely sensed imagery, and measures of topography. In Chapter 2, I used a combination of trend surface models, regression trees, and logistic regression to model forest structure from field data, remotely sensed imagery, measures of topography, and the map of vegetation types developed in Chapter 1. Both modeling procedures (composition, structure) used fine-grain field data to improve the resolution of the remotely sensed data on which the classifications are based, a novel and cost-saving approach for classifying large study areas. All models are also evaluated extensively for their performance. The accuracy each model of vegetative type was assessed using independent aerial photo interpretation and a post-stratified, multivariate composite estimator. Models of forest structure were validated using a stratified 11-fold cross-validation method.

In Chapter 3, I simulate the spatial interaction of goshawks with conspecifics and their forest environment using a Gibbsian pairwise potential model (Ogata and Tanemura 1989). The model describes the spatial dependency (1) among goshawk nest locations on the study area and (2) between nest locations and forest composition/structure within 10-

m x 10-m areas containing nests. Although spatial modeling techniques have been used to investigate wildlife-habitat relationships, no studies have examined multi-way, dynamic relationships that occur among individuals in a population, and between the individuals and their environment. Nonetheless, models based on multi-way, dynamic relationships are integral to understanding how populations respond to changes in habitat, such as those resulting from forest management. The dynamic modeling effort presented here, referred to as a dynamic spatial simulation, is unique in its approach in that it uses both the hawk's territorial behavior and habitat conditions interactively to simulate the spatial distribution of nests and to investigate whether the availability of nest habitat limits, or otherwise affects the distribution and abundance of, goshawks. Using logistic regression, I then model the probability of an occurrence of an active nest within a 10-m x 10-m area. By correlating the location of known nests with environmental variables that account for the large-grain variability in the landscape, I identified habitat that is more likely to contain active nests.

In Chapter 4, I address an important ecological question about goshawk population performance; namely, whether habitat differences occur among goshawk territories with high and low reproductive performance and random sites. Habitat correlates are measured as the proportion of vegetation types (results of Chapter 1) and diversity around active nests within territories. Logistic regression models are then developed to predict the probability that a site contains a goshawk territory and that the territory will have high reproductive potential. To date, no goshawk studies have been conducted at sufficiently large temporal (8-10 years) or spatial (landscapes encompassing >50 territories) scales to infer goshawk territory quality from demographic performance.

This study represents 10 years of goshawk population monitoring, including 101 territories, on a relatively isolated landform, the Kaibab Plateau.

Each of the following chapters in my dissertation was written as a manuscript for peer-reviewed journals. Chapter 1 has been accepted by the International Journal of Remote Sensing. Chapter 2 is in review with Forest Science. Chapter 3 is in review with Ecological Modeling. I am currently revising Chapter 4 for publication in an ornithological or wildlife journal. Chapters 1-3 reflect the inclusion of co-authors (Robin M. Reich, Richard T. Reynolds).

LITERATURE CITED

Crocker-Bedford, C. 1990. Goshawk reproduction and forest management. Wildl. Soc. Bull. 18:262-269.

Hall, J. G. 1981. A field study of the Kaibab squirrel in Grand Canyon National Park. Wildl. Monogr. 75:6-54.

Johnsgard. P. A. 1990. Hawks, eagles, and Falcons of North America. Smithsonian Institution Press, Washington, D.C. 403 pp.

Kennedy, P. L. 1997. The northern goshawk (Accipiter gentilis atricapillus): is there evidence of a population decline? J. Raptor Res. 31:95-106.

Ogata, Y., and Tanemura, M. 1989. Likelihood estimation of soft-core interaction potentials for Gibbsian point patterns. Ann. Inst. Stat. Math. 41:583-600.

Peck, J. 2000. Seeing the forest through the eyes of a hawk: an evaluation of recent efforts to protect northern goshawk populations in southwestern forests. Nat. Res. J. 40:125-156.

Reynolds, R.T. 1983. Management of western coniferous forest habitat for nesting Accipiter hawks. Gen. Tech. Rep. RM-102. Fort Collins, CO: U.S.D.A., Forest Service, Rocky Mount. For. Range Exp. Stat. 7 p.

Reynolds, R.T. 1989. Accipiters. Pages 92-102 *in* B. G. Pendleton, G. E. Ruibal, D. L. Krahe, K. Steenhof, M. N. Kochert, M. N. LeFranc, Jr., eds. Proc. of the Western Raptor Management Symposium and Workshop. Natl. Wildl. Fed. Sci. Tech. Ser. No. 12, Washington, DC. 317 p.

Reynolds, R.T., R. T. Graham, M. H. Reiser, R. L. Bassett, P. L. Kennedy, D. A. Boyce, Jr., G. Goodwin, R. Smith, and E. L. Fisher. 1992. Management recommendations for the northern goshawk in the southwestern United States. U.S.D.A., For. Serv., Gen. Tech. Rep. RM-217, Rocky Mount. For. Range Exp. Stat., Fort Collins, CO. 90 p.

Reynolds, R. T. S. M. Joy, J. D. Wiens, S. R. Salafsky, and J. C. Seyfried. In review. Territory and mate fidelity and breeding dispersal in the northern goshawk. Condor.

Squires, J. R., and R. T. Reynolds. 1997. Northern goshawk (*Accipiter gentilis*). *In* The Birds of North America, No. 298, A. Poole and F. Gill, eds. Acad. Nat. Sci., Philadelphia, PA, and The Amer. Ornithol. Union, Washington, D.C. 32 p.

Speiser, R., and T. Bosakowski. 1984. History, status, and future management of goshawk nesting in New Jersey. Rec. New Jersey Birds. 10:29-33.

CHAPTER ONE

A NON-PARAMETRIC, SUPERVISED CLASSIFICATION OF VEGETATIVE TYPES ON THE KAIBAB NATIONAL FOREST USING DECISION TREES

ABSTRACT

Traditional land classification techniques that use remotely sensed imagery are typically limited to the fixed spatial resolution of the sensors (e.g., 30 m for Landsat Thematic Mapper (TM) imagery). However, the study of some ecological processes requires land cover classifications at finer spatial resolutions. In this paper, we model forest vegetation types on the Kaibab National Forest (KNF) in northern Arizona to a 10-m spatial resolution with field data, using topographical information (elevation, slope, aspect, landform) and Landsat 5 TM imagery as auxiliary variables. Vegetation types were identified by clustering the field variables total basal area and proportion of basal area by species, and then using a decision tree based on the auxiliary variables to predict the vegetation types. Use of additional variables (canopy closure, height of the herbaceous and shrub understory vegetation, presence of small, young trees in the understory, and proportion of ground covered by vegetation) did not improve the model. Vegetation types modeled included pinyon-juniper, ponderosa pine, mixed conifer, spruce- and deciduous-dominated mixes, and openings. Results of the decision tree had a resubstitution accuracy of 87%. To independently assess the accuracy of the final vegetation maps, we used a post-stratified, multivariate composite estimator. Overall

accuracy was 74.5% with a Kappa statistic of 49.9%. Sources of error included differentiating between mixed conifer and spruce-dominated forest types and between openings in the forest and deciduous-dominated mixes. Inclusion of a larger sample of field plots in our study design might have mitigated these classification dilemmas. Overall, our non-parametric classification method successfully identified dominant vegetation types on the study area at a finer spatial resolution than can typically be achieved using traditional classification techniques.

INTRODUCTION

Habitat changes due to forest management (e.g., tree harvests, fire suppression) are thought to be responsible for declining populations of northern goshawks (*Accipiter gentilis atricapillus*; hereafter referred to as goshawk) in the North America (Reynolds 1983, 1989, Speiser and Basakowski 1984, Crocker-Bedford 1990, Reynolds et al. 1992). An accurate description of habitat is, therefore, paramount to understanding the relationship between goshawk demographic performance and its changing habitat. Modeling habitat characteristics is often difficult when a study area is large and diverse and complete sampling of environmental variables is unrealistic. Remote sensing technology, however, may be used as an indirect means of obtaining information on the biophysical characteristics of large areas.

Traditional land classification techniques incorporate information derived from remotely sensed data, such as satellites (e.g., Landsat TM sensors), to develop models of land cover and are limited to the fixed spatial resolution of those data (e.g., 30 m). In addition, land cover classifications are frequently driven by (supervised classification) or

defined by (unsupervised classification) the interpretation of photography or by incidental knowledge of the study area. The former requires a high level of skill to reduce interpretation error, while the latter may not provide comprehensive coverage of the study area. Ground sampling at select points, assumed to be error free, is often used as auxiliary information to validate the classification, increase the precision of the classification parameters, and/or compute a classification error rate. Disadvantages of traditional classification techniques are the expertise required to carry out and interpret the initial classification and the inability to interpolate the classified data to a scale finer than the spatial resolution of the pixel associated with the imagery.

Herein, we describe a method for modeling the composition of goshawk habitat on the Kaibab National Forest (KNF) in northern Arizona using non-traditional classification techniques. Vegetative classes are first derived from a cluster analysis of the field data. Decision trees are then used to model the vegetation throughout the study area using independent variables such as Landsat TM bands, slope, aspect, elevation, and landform. Advantages of using a decision tree over traditional remote sensing classification techniques (Friedl and Brodley 1997, De'Ath and Fabricius 2000) include the non-parametric nature of the test, the test's invariance to transformations of the independent variables, ease of interpretation, and the robustness of results. Unlike many supervised and unsupervised parametric classification algorithms, decision trees can be used to describe non-linear interactions. Individual splits in a decision tree may also be checked for physical justification. In comparative analyses with equivalent linear models (analysis of variance, linear regression) (De'Ath and Fabricius 2000) and maximum likelihood estimators (Hansen et al. 1996, Friedl and Brodley 1997), decision trees

outperformed both classification techniques. The major disadvantage associated with decision trees, and one that may hinder their widespread application for developing vegetation maps, is the need for large sample sizes (200 or more sample plots; S.M. Joy, pers. obs.) when dealing with complex data sets (i.e., those with non-linear or high-order interactions). Despite this limitation, decision trees have been used effectively to classify remotely sensed imagery (Michaelson et al. 1994, Friedl and Brodley 1997), validate cluster-driven classification algorithms (Iñiguez 2000), and explore and interpret ecological data (Baker 1993, De'Ath and Fabricius 2000).

In our study, we used more than one source of data to assess the accuracy of the estimated vegetation map. Field data, considered 'error-free', were used with photo-interpreted data to independently assess model accuracy. However, combining data from several such 'sources' increased the complexity of the accuracy assessment as different sampling protocols were used. Consequently, we used a composite estimator (Green and Strawderman 1990) that combines information from our various sampling sources to estimate accuracy. We discuss the application of our multivariate composite estimator to assess the accuracy of the classification procedure used to identify the major vegetation types on our study area.

STUDY AREA

The Kaibab Plateau, in northern Arizona, is an oval-shaped (95 km x 55 km), limestone plateau that rises from a shrub-steppe plain at 1750 m above mean sea level (asl) to its highest point at 2800 m. Surface weathering has produced gentle drainages and moderately sloping valleys on the landform. The Plateau is bounded by escarpments

of the Grand Canyon of the Colorado River on its south side, by steep slopes on the east, and gentle slopes on the north and west sides that descend to the shrub-steppe plain. The study area, which occupies the northern two-thirds of the Plateau, includes all of the Kaibab National Forest (KNF), above 2182 m asl (Fig. 1.1), or about 128,500 ha [estimated from digital elevation models (DEM) using ARC/INFO® (ESRI, 1995)]. The KNF is managed by the North Kaibab Ranger District (NKRD; Fredonia, AZ).

Pinyon (*Pinus edulis*)-juniper (*Juniperus* spp.) woodlands occur at elevations below 2075 m asl. Ponderosa pine (*P. ponderosa*) forests occur between 2075 and 2450 m asl, mixed conifer (*P. ponderosa, Pseudotsuga mensiesii, Abies concolor,* and *Picea pungens)* forests between 2450 and 2650 m asl, and spruce (*Picea engelmannii*)-fir (*A. lasiocarpa*) forests between 2650 and 2800 m asl (Rasmussen, 1941, White and Vankat, 1993). At transition zones of forest types, associated tree species typically intermix because of differences in slope and aspect. A series of narrow meadows containing grasses and herbaceous vegetation occur on the Plateau.

Each forest type has been altered by some form of management: livestock grazing, fire suppression, thinning, shelterwood, seed-tree and sanitation cuts, and clearcuts. Prior to the introduction of livestock grazing (late-1800s), fire suppression (beginning in the early 1900s), and intensive logging (beginning in the 1980s), many trees were in mature size classes, and occurred singly or in groups (Rasmussen 1941). Forest understories were dominated by grasses (*Poa*, *Sitanion*, and *Muhlengergia* spp.) and typically free from smaller trees (Rasmussen 1941, Merkle 1962). Currently much of the ponderosa pine type understory is dense with pine reproduction and, at higher elevations, with white fir reproduction.

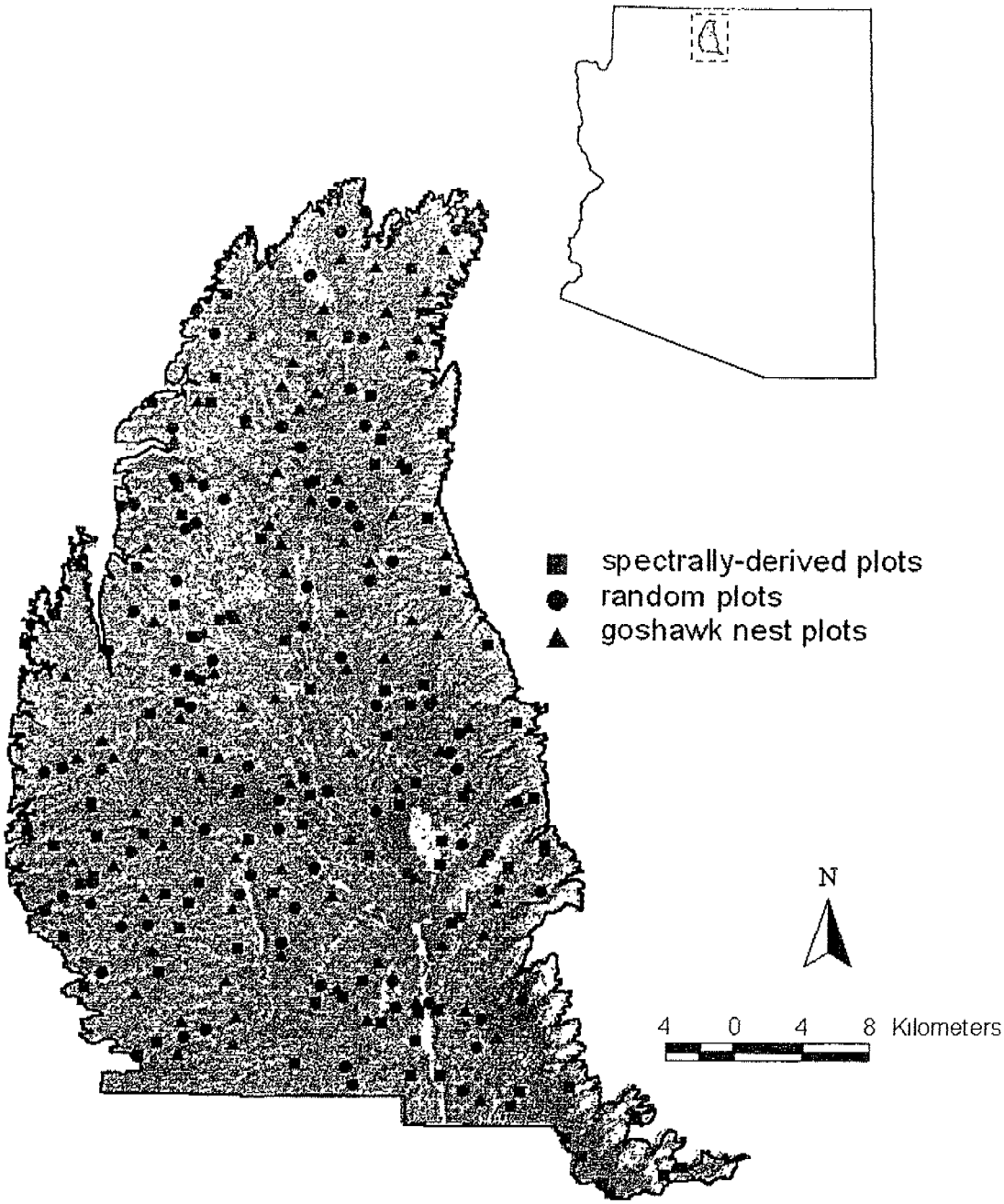


Figure 1.1. 1997 Landsat TM image (bands 4, 3, and 2) of the Kaibab National Forest (North Kaibab Ranger District), Arizona, above 2182 m in elevation. Points indicate spectrally-derived (square), random (circle), and northern goshawk nest (triangles) plots.

METHODS

Data Collection

Field Data

Field sampling occurred between July and September 1997. A total of 272 field plots (Fig. 1.1) obtained from three sources were used to develop a vegetation map of the study area:

Spectrally-derived plots - To capture the spectral variability on the study area, we used an unsupervised classification procedure [(CLASSIFICATION; IMAGINE® version 8.3, ERDAS 1997) (ISODATA algorithm, Tou and Gonzalez 1974)] on a 1994 Landsat TM scene of the study area (16 July, centered on Path 37and Rows 34 and 35) to identify 50 spectral classes. The ISODATA algorithm computes the minimum distance between spectral signatures to identify clusters of pixels with similar spectral characteristics. Use of additional spectral classes might have captured more of the spectral variability in the image; however, we believed that 50 classes sampled a reasonable spectral range and that the remaining plot sources (below) would capture any residual spectral variability. To sample the corresponding vegetative characteristics in the field, we generated 100 random coordinates, two per spectral class (EVALUATE CLASSIFICATION; IMAGINE®, ERDAS 1997). Each coordinate represented the center of a 3 x 3 pixel (90-m x90-m) 'window' of like spectral class. Due to steep terrain, we sampled 98 of the 100 plots.

Goshawk nest plots - Goshawks typically lay eggs in two or more 'alternate' nests within their territories among breeding years (Reynolds et al. 1994). To estimate the range of vegetative characteristics at goshawk nest areas, a sample plot, centered on the nest tree, was established at one randomly chosen alternate nest in 95 goshawk territories that had been identified on the KNF between 1991 and 1997. Due to data collection errors during field sampling, 92 of the 95 nest tree plots were used in analyses.

Randomly located plots – One hundred sample plots were located randomly to capture the remaining vegetative and topographic variability on the study area. The sample plots were established irrespective of territories and nests. Due to steep terrain or the remoteness of plots and time constraints, 85 of the 100 sample plots were measured. The majority (13 of 15) of plots not measured were at the edge of the study area where access was difficult due to steep slopes or cliffs.

Each plot was oriented in a north-south/east-west direction and divided into a cluster of nine 10-m x 10-m subplots corresponded to a 30-m x 30-m pixel on a Landsat TM image. In developing this vegetation map, we used only data from the central 10-m x 10-m subplot on each plot (Fig. 1.2). The remaining eight subplots were sampled to model forest structure (canopy closure, total basal area, proportion of pine, spruce-fir, and aspen basal area, maximum height of the understory vegetation, and presence or absence of small trees in the understory) on the study area (see Chapter 2).

Coordinates for the spectrally-derived and nest plots were assigned to the center of the plot because of the geographic specificity of the unit (Landsat TM pixel or nest tree) on which the plots were based. Coordinates for random plots were systematically assigned to the lower left-hand corner of each plot. A Trimble Navigation Pathfinder™

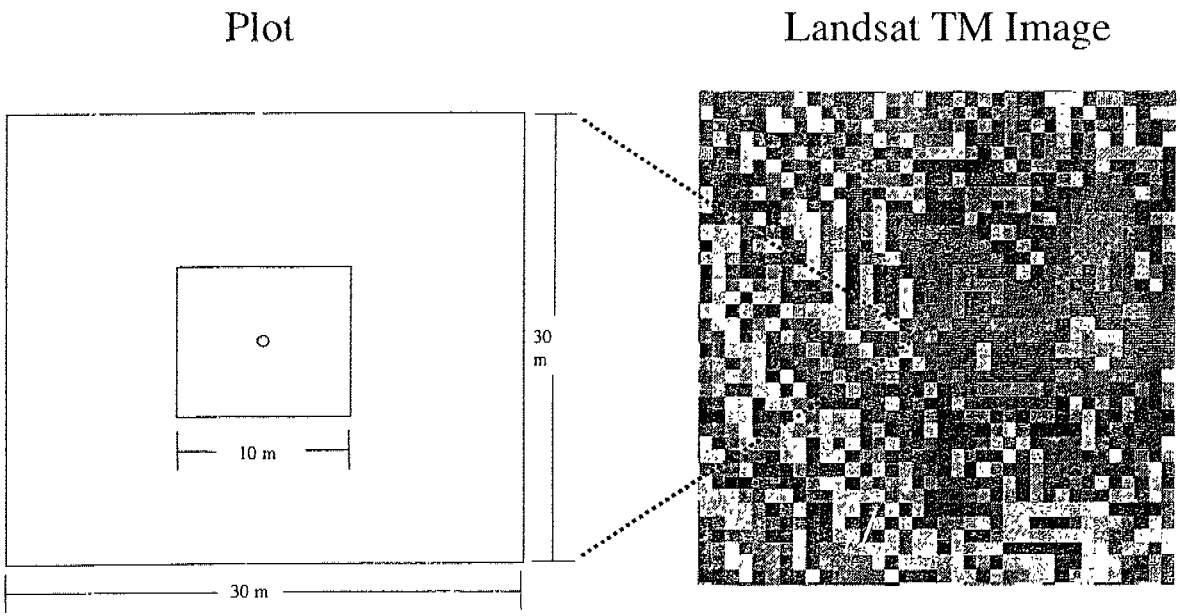


Figure 1.2. Plot layout used to sample vegetative types on the Kaibab National Forest (North Kaibab Ranger District), Arizona, to a 10-m spatial resolution and the relationship of the plot to Landsat TM imagery.

Asset Surveyor Global Positioning System (GPS) with an estimated accuracy of 1-3 m was used to ascertain actual plot locations and account for orienteering errors. Measurements taken on plots included canopy closure [estimated with a concave, spherical densiometer (Lemmon 1956, 1957)], overstory species, total basal area by species [estimated with a 4.6 (metric) factor prism], height of the understory vegetation, % ground cover, and presence of seedling and sapling trees.

Digital maps were generated to assist with the orienteering of sample plots on the ground. Each digital map was composed of a digital elevation model (DEM) (USGS, 1:24000, 30-m spatial resolution) overlaid with digital line graphs of NKRD roads (R. Crawford, pers. comm.) and the sample plots.

Landsat TM and GIS Data

We obtained a cloud-free, 22 June 1997 TM image (Path 37, Row 35) of the study area that corresponded with the timing of field data collection. Band layers 1-5 and 7 were exported as ARC/INFO® (ESRI 1995) grid coverages and resampled (RESAMPLE, nearest neighbor, GRID Module) to 10 m, corresponding to the spatial resolution of the field data. The value of each pixel was averaged by passing a 3 x 3 moving window (FOCALMEAN, GRID Module) over the resampled grids. Thus, averaging occurred every 10 m to smooth the spectral transition between features. This resulted in a grid with a piece-wise, continuous surface where every 10-m x 10-m pixel represented the average of the surrounding 30-m x 30-m pixels, including the central 10-m pixel of each original Landsat TM pixel whose value did not change. This was done to reduce potential registration errors and to reflect changes in forest structure and vegetative classes among the nine10-m x 10-m subplots measured on the ground. Resampling was

also important because not all of the forested plots fell within spectrally distinct areas; some plots were in transition zones between spectral classes. Average DN (digital number) values associated with the 272 sample plots were extracted from each of the six Landsat TM grids using a customized ArcView® (ESRI 1998) application.

Elevation, slope, aspect, and landform (McNab 1989) were also determined for each sample plot from DEMs. Landform (McNab 1989) is an index that expresses surface shape as a measure of surface concavity or convexity (computed as the mean slope gradient from the original cell to adjacent cells in 4 directions) creating a continuous variable. Prior to extracting the cell values, each grid was resampled to 10 m as described above.

Photographic Data

True color aerial photography (1:12,000, 1991, NKRD), infrared National High Altitude Photography (NHAP) (1:58,000, 1980, USGS), Digital Orthophoto Quadrangles (DOQs) (1:24,000, 1992, USDA Forest Service Geometric Service Center, USGS), and photographs taken during field sampling were used in the assessment of classification accuracy. Information on forest management activities [Resource Information System (RIS) data] aided in identifying management treatments that occurred subsequent to the acquisition of the photography and was provided by the NKRD (K. Fuelling, D. Steffensen, pers. comm.).

Image Classification.

A hierarchical clustering algorithm (HCLUST; S-PLUS©, Statistical Sciences 1995), based on group averaging, was used to group the field data into clusters with similar vegetative characteristics (i.e., species composition, basal area, canopy closure,

understory vegetation, etc.). The algorithm produces a two-dimensional dendrogram representing the fusion of clusters at each successive stage of the analysis. At each stage, the two 'nearest' clusters are combined to form one bigger cluster (initially each cluster contains a single sample plot). This process continues until only one group containing all of the observations remains. Empirical studies have shown that when there are an unequal number of observations per cluster, group averaging is more successful at identifying clusters than other hierarchical clustering methods (Everitt 1993). Group averaging also has superior performance when the data contain evidence of a distinct cluster structure (Everitt 1993).

The resulting field data clusters were assigned to modified Anderson Level II ('opening' class only) or III (all other forest types) (Anderson et al. 1976) land cover classes (Table 1.1). Anderson Levels II and III were assumed to be sufficient for modeling dominant goshawk habitat. S-PLUS© (TREE; Statistical Sciences 1995) was used to generate a stepwise decision tree (Breiman et al. 1984, Friedl and Brodley 1997, De'Ath and Fabricius 2000) that identified independent variables (Landsat TM bands, elevation, slope, aspect, or landform) that were important in discriminating among vegetation types. The decision tree uses a binary partitioning algorithm that maximizes the dissimilarities among groups to compare all possible splits among the independent variables and splits within each independent variable to partition the data into new subsets. Once the algorithm partitions the data into new subsets, new relationships are developed to split the new subsets. The algorithm recursively splits the data in each subset until either the subset is homogeneous or the subset contains too few observations (< 5) to be split further. To prevent over fitting the data, a pruning algorithm (Friedl and

Table 1.1. Modified land cover classification system (Anderson et al., 1976) for vegetation types on the Kaibab National Forest (North Kaibab Ranger District), Arizona. Characteristics at higher levels are nested within the lower level.

Land Cover		
Level I	Level II	Level III
Forest	Coniferous	Ponderosa pine
		Mixed Conifer
	Deciduous	Deciduous-dominated mix
	Mixed	Pinyon-Juniper
		Spruce-dominated Mix
Non-Forested	Opening	

Brodley 1997) was used to eliminate subsets that were fit to noise in the data. Decision tree criteria were used as 'training' statistics to classifying the 1997 Landsat image. The classification was achieved using Arc Macro Language [AML (CON; ARC/INFO®, ESRI 1995)] to produce the final vegetation map.

Accuracy Assessment

A sample-based assessment of accuracy (i.e., that based on the same data used to generate the classification) for the decision trees was calculated by weighting the classification error associated with a given vegetation type proportional to its area, a process referred to as post-stratification (Cochran 1977:134-135). To assess a model's classification accuracy independently, field data alone may be used if the data were not involved in the classification procedure or came from a sufficiently large sample. However, if the sample of field data is small or the data were used in developing the classification procedure (such as here), independent auxiliary information is needed in conjunction with the field data to estimate accuracy. When data from different sources are combined in this manner, the complexity of the accuracy assessment increases due to different sampling protocols. A composite estimator (Green and Stawderman 1990) may be used to combine the data from the different sources and improve estimates of model accuracy (APPENDIX). Czaplewski (2000) provides detailed discussion on the use of auxiliary information to assess the accuracy of vegetation maps derived from remotely sensed data (see also Kalkhan et al. 1996, 1998).

To independently assess the accuracy of our classification using the multivariate composite estimator, we generated 498 random points as auxiliary data using simple random sampling. The location of the 498 random points and 269 of the 272 field plots

(photographic data were unavailable for three of the field plots) were transferred to aerial photographs, NHAP, and/or DOQs. An independent photo interpreter (i.e., one with no knowledge of the classification) identified the dominant vegetation type corresponding to each 10-m x 10-m sample location (random photo points, field plots). Grid values were then extracted from the vegetation map for each location using a customized ArcView® (ESRI 1998) application.

In this study, because we used three selection criteria (spectrally-derived, nest-based, and random plots) to sample the field data, increasing the complexity of our analysis as formulas used in summarizing these data depended on sampling design (Stehman 1999). When vegetation types associated with the sample data are proportional to the true values, the sample data should provide good estimates of the overall accuracy, as a simple random sample distributes itself approximately proportionally among strata (Cochran 1977:134-135). However, because of the nesting requirements of the goshawk, the vegetation types associated with nest-tree plots are not likely to be in proportion to those observed on the study area outside of nest areas. Also, because spectrally-derived plots were designated based on reflectance values, they may not be in proportion to the vegetation types observed on the study are. As a result, the combination of these data may lead to poor estimates of the overall model accuracy. Stehman (1996) points out, however, that if simple random sampling is combined with a post-stratified estimator, the estimator is nearly as efficient as a proportionally-allocated stratified design. A post-stratified estimator combined with simple random sampling should also perform nearly as efficiently as an optimally-allocated stratified design for estimating overall accuracy (Stehman 1999), assuming a large enough sample, because an optimally-allocated

stratified design is usually only slightly more precise than a proportionally-allocated stratified design (Cochran 1977:327-335).

To compensate for differences in the selection probabilities associated with our three sets of sample data, we adjusted the joint probability matrix associated with the three sampling procedures using post-stratification. This adjusted error matrix was used to calculate a Kappa statistic to estimate the difference between the classified and ground-verified themes, and the agreement contributed by chance. Because of the complexity of our sampling design, the usual formula for calculating the variance of the Kappa is not appropriate, as the data do not follow a multinomial distribution. Under these conditions, bootstrapping techniques can be used to construct confidence bands around the estimates of Kappa (Kalkhan et al. 1997). As we were not testing hypotheses or making comparisons, this was not done.

RESULTS

Image Classification

Preliminary analysis indicated that the accuracy of the classification could be increased by modeling pure ponderosa pine forests separately from the other vegetation types. Sample plots with 100 % of basal area in ponderosa pine were assigned a value of 1 and all other plots were assigned a value of 0 when fitting the decision tree (Fig. 1.3). The final decision tree had 25 terminal nodes. Discriminating variables included elevation, slope, aspect, landform, and Landsat TM bands 1-5 and 7. The initial split, which maximized the distance between the response variables (pine/other), was based on elevation (2579 m) and separated the higher elevation spruce-fir forests from the pine and

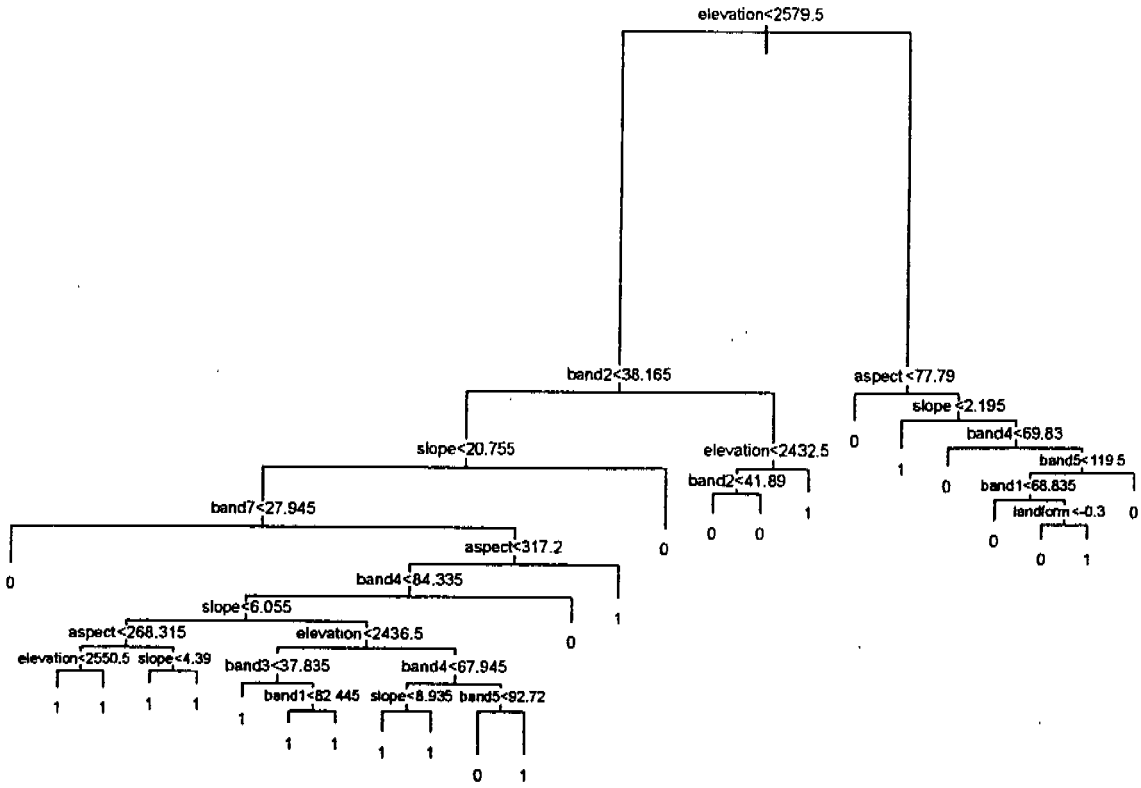


Figure 1.3. Decision tree classifier for ponderous pine (1) and other dominant vegetative types (0) on the Kaibab National Forest (North Kaibab Ranger District), Arizona. Independent variables (and units) used in the classification procedure include elevation (m), slope (%), aspect (degrees), landform (index), and Landsat TM bands 1-5 and 7 (digital number).

pinyon-juniper forests that occurred at lower elevations. Ninety-eight % (148/151) of the pure pine plots and 86 % (104/121) of "other" forest type plots (i.e., ponderosa pine mixed with other species, spruce, fir and deciduous vegetation, openings) were correctly classified for an overall resubstitution model accuracy of 92 %.

After we developed the decision tree for pure ponderosa pine, the pure pine sample plots were removed from the data set. We identified six additional vegetation classes from the remaining data by clustering the proportion of basal area associated with the remaining species. Vegetation classes included pinyon-juniper, mixed conifer, fir-dominated mix, spruce-dominated mix, deciduous-dominated mix, and openings. We later consolidated mixed conifer and fir-dominated mix into the mixed conifer class to improve the model.

The final decision tree (Fig. 1.4) for the "other" (i.e., non-pure pine) vegetation classes had 20 terminal nodes. Important variables in the partitioning criteria included elevation, slope, aspect, and Landsat TM bands 1, 2, 4, 5, and 7. The initial split was again based on elevation (2418 m) and separated pinyon-juniper and some deciduous-dominated classes and openings from higher elevational forest mixes. Landsat TM band 7, representing the mid-infrared (2.08-2.35 μM) spectral range, further distinguished openings from the majority of higher elevational forest mixes. Eighty-seven % (14/16) of the pinyon-juniper sites were correctly classified (Table 1.2). Similarly, 88 % (29/33) of the mixed conifer sites, 87 % (27/31) of the spruce-dominated sites, 90 % (18/20) of the deciduous-dominated sites, and 76 % of the openings (16/21) were correctly classified. The overall resubstitution accuracy of this decision tree was 86 %. The combined model (pure ponderosa pine and other vegetation classes) correctly classified 93 % (20/272) of

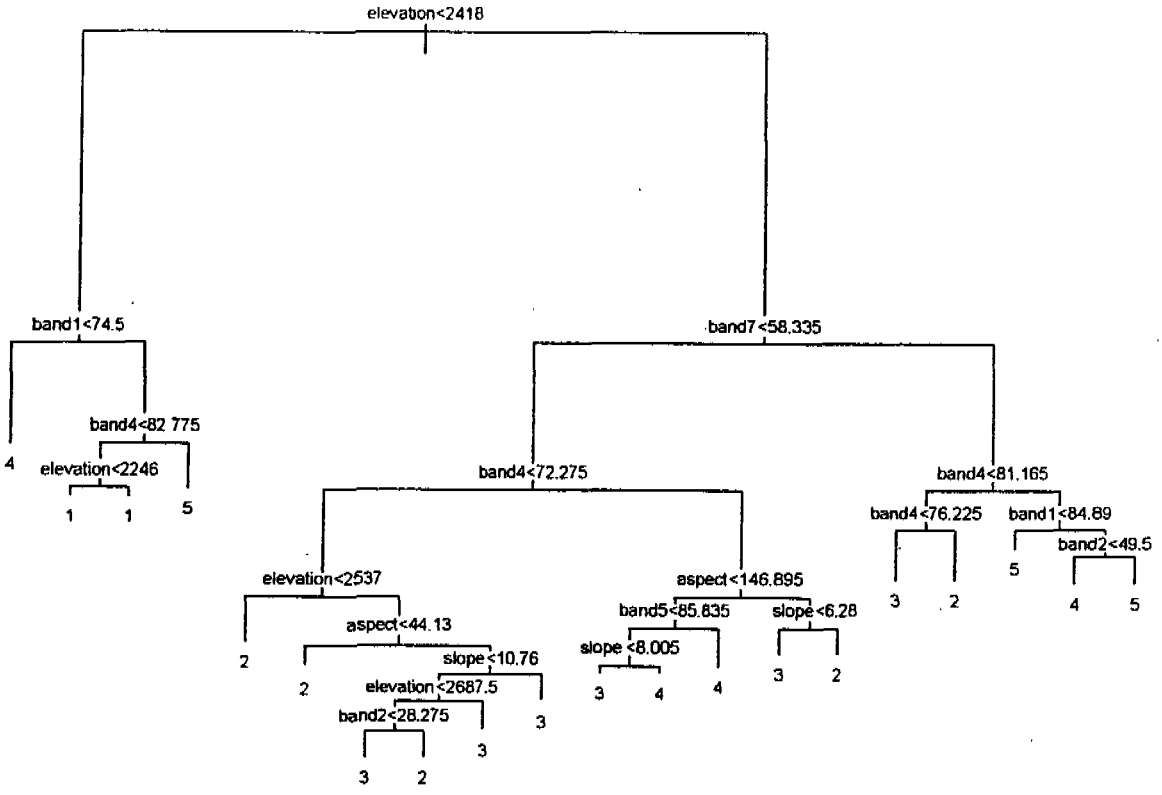


Figure 1.4. Decision tree classifier for dominant vegetative types other than pure ponderosa pine on the Kaibab National Forest (North Kaibab Ranger District), Arizona. Vegetation types include pinyon-juniper (1), mixed conifer (2), spruce-dominated mixes (3), deciduous-dominated mixes (4), and openings (5). Independent variables (and units) used in the classification procedure include elevation (m), slope (%), aspect (degrees), and Landsat TM bands 1, 2, 4, 5 and 7 (digital number).

Table 1.2. Resubstitution classification error rates of the decision tree for the vegetation types on the Kaibab National Forest (North Kaibab Ranger District), Arizona. Error rates were adjusted (post-stratified) for the proportions of the vegetation classes on the study area.

Vegetation Type	Misclassified Plots	Total Plots	Classification Error Rate	Post-stratified Classification Error Rate
Pinyon-Juniper	2	16	0.125	0.010
Ponderosa Pine	3	151	0.020	0.011
Mixed Conifer	4	33	0.121	0.014
Spruce-dominated mix	4	31	0.129	0.013
Deciduous-dominated Mix	2	20	0.100	0.009
Openings	5	21	0.238	0.014
Column Totals	20	272		
Overall Error			0.125	0.071

the sample plots. Overall post-stratified resubstitution accuracy for the model was also 93% (Table 1.2). In both models (pure ponderosa pine and all other vegetation types), the use of additional variables (e.g., canopy closure, height of the understory vegetation, presence of seedling and sapling, and proportion of ground covered) tended to obscure the cluster structure and did not improve our ability to identify major vegetation types.

On the final vegetation map (Fig. 1.5), ponderosa pine occupied the majority (55.5%) of the study area (Table 1.3) and occurred at lower elevations, with mixed conifer (11.3%) and spruce-dominated mixes (10.1%) occurring at higher elevations. Pinyon-juniper (8.2%) was found predominantly along lower elevational edges of the study area and where crown-destroying fire and intensive management (e.g., shelterwood and seed-tree cuts) had occurred at lower elevations. Deciduous-dominated mixes (8.8%) and openings (6.1%) occurred throughout the study area. All field plots, except goshawk nests, were sampled in relatively close proportion to the occurrence of vegetation type in the model (Table 1.3).

Accuracy Assessment

The overall accuracy of the classification model when compared to the photo interpreted data was 62.8%, with a Kappa of 32.9% (Table 1.4). The low Kappa statistic suggests a poor relationship between the classification derived from the decision tree and the photo interpreted classification. Low accuracy estimates associated with mixed conifer, spruce-dominated mixes, and openings contributed largely to the low overall accuracy (Table 1.4). Errors associated with coniferous species were attributed to difficulties in distinguishing among mixed stands of pine, fir, and spruce on the photographs, especially in small ($\leq 100\ m^2$) patches of individual species. Photo

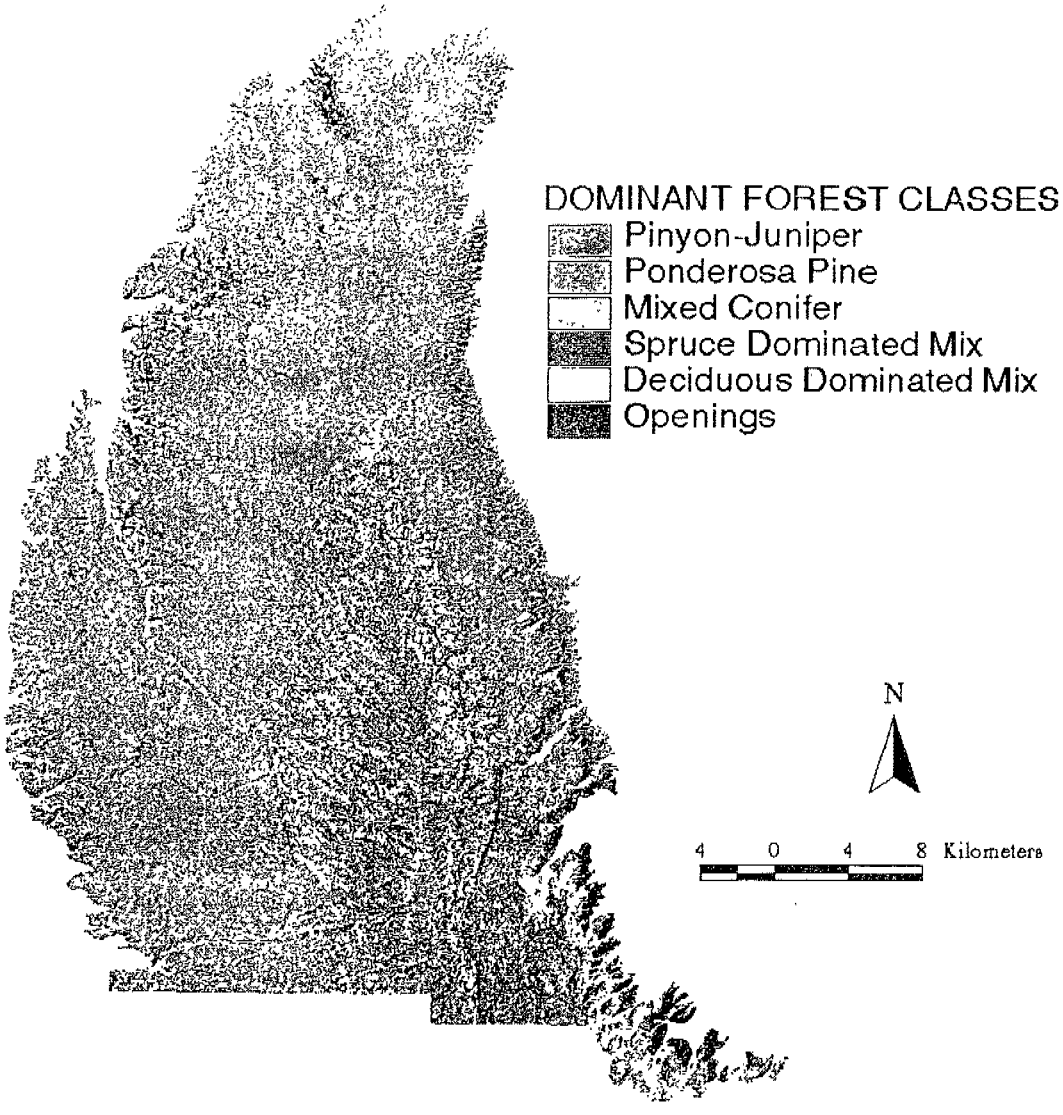


Figure 1.5. Dominant vegetative types on the study area in Arizona.

Table 1.3. Distribution of sample plots by vegetation type on the Kaibab National Forest (North Kaibab Ranger District), Arizona.

Vegetation Type	Number of Pixels (10 m x 10 m)	Proportion of Total Area	Spectrally-Derived Plots	Goshawk Nest Plots	Randomly Located Plots	All Field Plots	Photo Interpretation Plots
Ponderosa Pine	7,137,947	0.56	0.45	0.70	0.59	0.58	0.58
Mixed Conifer	1,454,207	0.11	0.14	0.12	0.10	0.12	0.11
Spruce-dominated Mix	1,300,085	0.10	0.11	0.12	0.09	0.11	0.11
Deciduous-dominated Mix	1,124,891	0.09	0.09	0.06	0.07	0.07	0.10
Pinyon-Juniper	1,055,726	0.08	0.10	0.00	0.06	0.05	0.07
Openings	778,865	0.06	0.11	0.00	0.09	0.07	0.03
Column Totals	112,851,721	1.00	1.00	1.00	1.00	1.00	1.00

Table 1.4. Joint probability error matrix for the vegetation types and photo-interpretation of 767 sample plots on the Kaibab National Forest (North Kaibab Ranger District), Arizona.

	Vegetation Types (Classified Map)							
Vegetation Type (Photographs)	PJ	PP	MC	SD	D	OP	Row Totals	Accuracy
Pinyon-Juniper (PJ)	0.0313	0.0222	0.0000	0.0000	0.0000	0.0078	0.0613	0.5106
Ponderosa Pine (PP)	0.0222	0.4003	0.0117	0.0104	0.0261	0.0078	0.4785	0.8365
Mixed Conifer (MC)	0.0026	0.0561	0.0469	0.0130	0.0143	0.0039	0.1368	0.3429
Spruce-dominated Mix (SD)	0.0000	0.0209	0.0326	0.0717	0.0169	0.0026	0.1447	0.2687
Deciduous-dominated Mix (D)	0.0013	0.0378	0.0104	0.0052	0.0235	0.0091	0.0873	0.4955
Openings (OP)	0.0052	0.0430	0.0091	0.0078	0.0117	0.0143	0.0911	0.1571
Column Totals	0.0626	0.5803	0.1107	0.1081	0.0925	0.0455	0.9997	

Overall Accuracy = 62.82% Standard Error = 0.99%

Kappa = 32.90%

interpretation of these forest types using mid-summer, true color aerial photography and black and white DOQs was difficult. The large error associated with openings was a result of forest regeneration following tree harvests or fire subsequent to the acquisition of photography and to the misclassification of openings containing a small arboreal component. Spectral similarities between deciduous and open forest vegetation also contributed to the error estimate for openings. Narrow (< 800 m) meadows comprised of grasses and forbs were often classified as deciduous-dominated vegetation, rather than as openings.

After using the multivariate composite estimator (APPENDIX) to adjust for differences in the photo-interpretation of random points and field plots, and adjusting for the proportion of vegetation types (i.e., post-stratification) on the study area, the performance of the classification procedure improved to 74.5% (Table 1.5). All vegetation types except openings had an estimated accuracy greater than 50%. This level of accuracy is comparable to results achieved with traditional classification techniques that use a larger (typically 900 m^2) per-pixel area. Our Kappa statistic for the classification model was approximately 50%, which indicated good agreement between the classification model and observations on the ground.

The upward adjustment in the overall accuracy and Kappa statistic after correcting for classification errors in photographic interpretation indicated a larger classification error associated with the independent photo-plots compared to the error of classifying field plots. There was no indication of any subjective bias associated with the photo-interpretation of the sample plots to suggest a loss of efficiency in estimating the weights of the composite estimator. If a bias were to exist, an alternative composite estimator

Table 1.5. Joint probability error matrix for the decision tree for vegetation types and the field data. Probabilities were corrected for differences in the photo-interpretation of random points and field plots using the multivariate composite estimator and adjusted to the proportions of the vegetation classes on the Kaibab National Forest (North Kaibab Ranger District), Arizona.

	Vegetation Types (Classified Map)							
Vegetation Type (Field Data)	PJ	PP	MC	SD	D	OP	Row Totals	Accuracy
Pinyon-Juniper (PJ)	0.0374	0.0133	0.0000	0.0000	0.0000	0.0115	0.0622	0.6014
Ponderosa Pine (PP)	0.0149	0.4956	0.0100	0.0035	0.0118	0.0102	0.5460	0.9076
Mixed Conifer (MC)	0.0000	0.0308	0.0667	0.0192	0.0046	0.0000	0.1213	0.5499
Spruce-dominated Mix (SD)	0.0000	0.0179	0.0197	0.0736	0.0139	0.0010	0.1261	0.5017
Deciduous-dominated Mix (D)	0.0000	0.0261	0.0100	0.0038	0.0402	0.0000	0.0801	0.5838
Opening (OP)	0.0000	0.0163	0.0055	0.0089	0.0057	0.0279	0.0643	0.4333
Column Totals	0.0523	0.6000	0.1119	0.109	0.0762	0.0506	1.0000	

Overall Accuracy = 74.47% (Overall Standard Error = 1.56%)

Kappa = 49.87%

proposed by Green and Strawderman (1990) could be employed. This estimator behaves as well as the usual precision-weighted multivariate composite estimator when auxiliary data are unbiased, yet is superior when the auxiliary information is severely biased.

DISCUSSION AND CONCLUSIONS

Goshawks occur in a wide range of forest types and structures throughout their geographic range in North America. Qualitative differences among the habitats occupied may be measured in terms of the fitness of individuals who occupy the range of habitat conditions (Fretwell and Lucas 1970, Van Horne 1983). To understand how goshawks are influenced by their environment, the relationships between their demographic performance on territories and the composition and spatial arrangement of vegetation within their territories need to be determined. Thus, we need to accurately model the vegetation to adequately characterize goshawk habitat use on the study area.

Dominant vegetation types on the study area included pinyon-juniper, ponderosa pine, mixed conifer, spruce-dominated mixes, deciduous-dominated mixes, and openings. Stands of pure ponderosa pine were identified with high (91%) resubstitution accuracy. Differentiating between mixed conifer and spruce-dominated forest types and between deciduous-dominated mixes and openings was difficult due to their spectral and/or physical similarities. Auxiliary variables (canopy closure, understory vegetative species and height, proportion of ground covered, and presence of seedling/sapling) did not improve the accuracy of the model because these variables tended to reflect differences in stocking levels, and not differences in species composition. Modifying our sampling to include additional field plots in these vegetation classes might have mitigated these

classification dilemmas. Our efforts to model vegetation type, however, were incidental to the primary sampling objective, which was to model forest structure on the study area to a 10-m spatial resolution (Chapter 2) and so we were limited by the protocol thereby established.

Use of independent data (i.e., those not used in the classification procedure) to assess the accuracy of a classification of remotely sensed data is now common practice. The assessment procedure should take into consideration the sampling design, sample size, and the classification scheme (Congalton 1991). Our method of accuracy assessment uses photo-interpretation as a relatively inexpensive source of independent reference data to assess the accuracy of our vegetation map. Use of the multivariate composite estimator also takes into account our multiple sampling schemes, by combining the independent set of photo-interpreted sample data with our field data to obtain a more efficient estimate of the overall classification accuracy (74.5%). This method corrected for photo-interpretation errors such as misclassification and incorrect plot locations, which can confound the assessment (Czaplewski 2000). However, the overall accuracy of our map may be slightly higher or lower than this because of registration errors between the Landsat TM image and the field data (Reimann et al. 2000). We feel that registration errors were negligible in our case.

Overall, the non-parametric classification method presented here successfully identified dominant vegetation types at a smaller spatial resolution than is typically achieved using traditional classification techniques. We used this model to aid in the description of forest structure on the study area (Chapter 2). We are also currently correlating our model of vegetative composition with a ranking of goshawk territories

based on the hawk's long-term demographic performance to identify forest characteristics of goshawk habitat quality (Chapter 4).

ACKNOWLEDGEMENTS

We gratefully acknowledge Carrie M. Erickson, Matthew A. Gavin, Luke J. H. Hunt, Amy M. Iñiguez, Jose M. Iñiguez, Mohammed A. Kalkhan, Donna C. Laing, Jeffrey S. Lambert, Joanna L. Nelson, Susan S. Salafsky, John C. Seyfried, Geroge S. Stamatellos, Rebecca A. Steffensen, Vernon L. Thomas, J. David Wiens, and Laura E. Williams for their assistance collecting field data. Vernon L. Thomas also provided customized ArcView® code and guidance on GIS manipulations. We thank Donna C. Laing for the interpretation of photographic images. Rudy M. King, William J. Krausmann, and Michael S. Williams provided helpful reviews of the manuscript. The North Kaibab Ranger District (NKRD, Fredonia, AZ) kindly provided housing and logistical support. We also thank Rick Crawford, Karl Fuelling, and Dave Steffensen of the NKRD for support in acquiring RIS data, DOQs, and road coverages of the KNF. This study was funded by Region 3 (Albuquerque, NM) and the Rocky Mountain Research Station (Fort Collins, CO), USDA Forest Service.

LITERATURE CITED

Anderson, J. R., E. Hardy, J. Roach, and R. Witmer. 1976. A land use and land cover classification system for use with remote sensor data. U.S.D.I. Geolog. Surv., Prof. Paper 964 (Washington, D.C.: US Government Printing Office). 28 p.

Baker, F. A. 1993. Classification and regression tree analysis for assessing hazard of pine mortality caused by *Heterobasidion annosum*. Plant Dis. 77:136-139.

Breiman, L., J. H. Friedman, R. A. Olshen, and C. J. Stone. 1984. Classification and Regression trees. Wadsworth Ind. Group, Belmont, CA. 358 p.

Cochran, W. G. 1977. Sampling Techniques. John Wiley & Sons, NY. 428 p.

Congalton, R. G. 1991. A review of assessing the accuracy of classifications of remotely sensed data. Rem. Sens. Environ. 37:35-46.

Crocker-Bedford, C. 1990. Goshawk reproduction and forest management. Wildl. Soc. Bull. 18:262-269.

Czaplewski, R. L. 2000. Accuracy assessment and areal estimates using two-phase stratified random sampling, cluster plots, and the multivariate composite estimator. Pages 79-100 *in* Quantifying Spatial Uncertainty in Natural Resources: Theory and Applications for GIS and Remote Sensing, H. T. Mowrer and R. G. Congalton, eds. Ann Arbor Press, Chelsea, Michigan.

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.

ERDAS. 1997. IMAGINE® 8.3 Software and On-line Help Manual. ERDAS, Inc., Atlanta, GA.

ESRI. 1995. ARC/INFO® Software and on-line help manual. Environ. Res. Inst., Inc., Redlands, CA.

ESRI. 1998. ArcView® 3.1. Environ. Res. Inst., Inc., Redlands, CA.

Everitt, B. S. 1993. Cluster Analysis. Edward Arnold, London. 170 p.

Fretwell, S. D., and H. L. Lucas. 1970. On territorial behavior and other factors influencing habitat distribution in birds. I. Theoretical development. Acta Biotheor. 19:16-36.

Friedl, M. A., and C. E. Brodley. 1997. Decision tree classification of land cover from remotely sensed data. Rem. Sens.Environ. 61:399-409.

Green, E. J., and W. E. Strawderman. 1990. A multivariate composite estimator. Pages 181-188 *in* Proc. State-of-the-art Methodology of Forest Inventory. U.S.D.A. For. Serv., Gen. Tech. Rep. PNW-GTR-263. Pacific Northwest Res. Stat., Portland, Oregon.

Hansen, M., R. Dubayah, and R. Defries. 1996. Classification trees: an alternative to traditional land cover classifiers. Int. J. Rem. Sens. 17:1075-1081.

Iñiguez, J. M. 2000. Cover type and structural classification in the Sky Islands of southeastern Arizona. M.Sc. thesis, N. Ariz. Univ., Flagstaff. 75 p.

Kalkhan, M. A., R. M. Reich, and R. L. Czaplewski. 1996. Statistical properties of measures of association and the Kappa statistic for assessing the accuracy of remotely sensed data using double sampling. Pages 467-476 *in* Spatial Accuracy Assessment in Natural Resources and Environmental Sciences: Second International Symposium. U.S.D.A., For. Serv., Gen. Tech. Rep. RM-GTR-277, Rocky Mount. For. Range Exp. Stat., Fort Collins, CO.

Kalkhan, M. A., R. M. Reich, and R. L. Czaplewski. 1997. Variance estimates and confidence intervals for the Kappa measure of classification. Can. J. Rem. Sens. 23:246-252.

Kalkhan, M. A., R. M. Reich, and T. J. Stohlgren. 1998. Assessing the accuracy of Landsat Thematic Mapper classification using double sampling. Int. J. Rem. Sens. 19:2049-2060.

Lemmon, P. E. 1956. A spherical densiometer for estimating forest overstory density. For. Sci. 2:314-320.

Lemmon, P. E. 1957. A new instrument for measuring forest overstory density. J. For. 55:667-668.

McNab, W. H. 1989. Terrain shape index: quantifying effect of minor landforms on tree height. For. Sci. 35:91-104.

Merkle, J. 1962. Plant communities of the Grand Canyon area, Arizona. Ecology 43:698-711.

Michaelson, J., D. S. Schmiel, M. A. Friedl, F. W. Davis, and R. O. Dubayah. 1994. Regression tree analysis of satellite and terrain data to guide vegetation sampling and surveys. J. Veget. Sci. 5:673-686.

Rasmussen, D. I. 1941. Biotic communities of the Kaibab Plateau, Arizona. Ecol. Monogr. 11:230-274.

Reimann, R., M. Hoppus, and A. Lister. 2000. Using arrays of small ground sample plots to assess the accuracy of Landsat TM-derived forest cover maps. Proc., Accuracy 2000, 4th International Symposium on Spatial Accuracy Assessment in Natural Resources and Environmental Sciences, 12-14 July 2000. Amsterdam, The Netherlands.

Reynolds, R.T. 1983. Management of western coniferous forest habitat for nesting Accipiter hawks. U.S.D.A. For. Serv., Gen. Tech. Rep. RM-102, Rocky Mount. For. Range Exp. Stat., Fort Collins, CO. 7 p.

Reynolds, R.T. 1989. Accipiters. Pages 92-102 *in* B. G. Pendleton, G. E. Ruibal, D. L. Krahe, K. Steenhof, M. N. Kochert, M. N. LeFranc, Jr., eds. Proc. of the Western Raptor Management Symposium and Workshop. Natl. Wildl. Fed. Sci. Tech. Ser. No. 12, Washington, DC. 317 p.

Reynolds, R.T., R. T. Graham, M. H. Reiser, R. L. Bassett, P. L. Kennedy, D. A. Boyce, Jr., G. Goodwin, R. Smith, and E. L. Fisher. 1992. Management recommendations for the northern goshawk in the southwestern United States. U.S.D.A. For. Serv., Gen. Tech. Rep. RM-217, Rocky Mount. For. Range Exp. Stat., Fort Collins, CO. 90 p.

Reynolds, R. T., S. M. Joy, and D. G. Leslie. 1994. Nest productivity, fidelity, and spacing of northern goshawks in northern Arizona. Stud. Avian Biol. 16:106-113.

Speiser, R., and T. Bosakowski. 1984. History, status, and future management of goshawk nesting in New Jersey. Rec. New Jersey Birds 10:29-33.

Statistical Sciences. 1995. S-PLUS© 3.3Statistical Software Package for Personal Computers. StatSci Division, MathSoft, Inc., Seattle, WA.

Stehman, S. U. 1996. Estimating the kappa coefficient and its variance under stratified random sampling. Am. Soc. Photogr. Rem. Sens. 62:401-402.

Stehman, S. U. 1999. Basic probability sampling designs for thematic map accuracy assessment. Int. J. Rem. Sens. 20:2423-2441.

Tou, J. T., and R. C. Gonzalez. 1974. Pattern Recognition Principles. Addison-Wesley Publ. Co., Reading, MA. 348 p.

Van Horne, B. 1983. Density as a misleading indicator of habitat quality. J. Wildl. Manage. 43:893-901.

White, M. A., and J. T. Vankat. 1993. Middle and high elevation coniferous forest communities of the North Rim region of the Grand Canyon National Park, Arizona, USA. Vegetat. 109:161-174.

APPENDIX:

Applicaion of the Multivariate Composite Estimator Used to Assess Classification Accuracy

Estimators that combine information from various sources are referred to as composite estimators (Green and Strawderman 1990). To apply the multivariate composite estimator to assess the accuracy of our classification procedure, we collected the data in two phases. In phase one, n' (296 field and 498 random) plots were located on aerial photographs. Each of the n' photo plots were classified by photo interpretation into one of k classes. This information was used to construct a k x k matrix of joint probabilities of all possible classification outcomes associated with the interpretation of plots on the photographs and the classification procedure under evaluation. To implement the composite estimator, we stacked the columns of the k x k error matrix of joint probabilities on top of each other to create a k^2 x 1 vector (x'), along with a k^2 x k^2 variance-covariance matrix of V(x').

In phase two, the field plots (n, a sub-sample of n') were classified into one of k classes associated with the classification outcome. This information was combined with the photo interpretation and classification outcome on our map to construct a k x k x k matrix of joint probabilities associated with the three classification procedures. This matrix of joint probabilities was re-arranged to create the k^3 x 1 vector, y_1, and k^3 x k^3 variance-covariance matrix V(y_1). Given an appropriately structured k^2 x k^3 matrix of zeros and ones (H_x), the vector y_1 can be collapsed to provide the k^2 x 1 vector, $x = H_x y_1$,

of joint probabilities associated with the photo interpretation of field plots and the classification procedure being evaluated. Next, the composite estimator was used to combine these two sources of information to obtain a more efficient k^3 x 1 vector:

$$y_c = y_1 + K(x' - x) \qquad [A.1.1]$$

of joint probabilities associated with the three classification procedure, where

$$K = V(y_1)H_y'[H_y V(y_1)H_y' + V(x')]^{-1} \qquad [A.1.2]$$

is the k^3 x k^2 weighting factor, H_y is a k^2 x k^3 matrix of zeros and ones used to sum the joint probabilities within the appropriate groups to form the desired k^2 x k^1 vector $y = H_y y_c$ of joint probabilities associated with the classification of the field plots and the classification procedure, and H_y' is the transpose of H_y. In equation A.1.1, $(x'-x)$ is an estimate of the difference in the classification errors association with the photo interpretation in the first and second phases of sampling. If there are no differences (i.e. x' = x), the most efficient estimation of the classification error is based on the data collected in the second phase of sampling. If there are differences, the classification errors are adjusted accordingly (equation A.1.1) to take these differences into consideration.

CHAPTER TWO

MODELING FOREST STRUCTURE ON THE KAIBAB NATIONAL FOREST IN ARIZONA

ABSTRACT

Models based on the classification of Landsat Thematic Mapper (TM) imagery are typically limited by the 28.5-m spatial resolution of the satellite's sensors. Using a combination of multiple linear regression models and binary regression trees, or logistic regression, we modeled forest structure to a 10-m spatial resolution to identify habitat structural components associated with northern goshawks on the Kaibab National Forest in northern Arizona. We spatially interpolated field data from our 1,285-km^2 study area using Landsat TM imagery, elevation, slope, aspect, landform, and vegetation class as auxiliary information. Spatial autoregressive models were used to adjust for the presence of spatially autocorrelated residuals. Our models accounted for 45-83% of the variability in spatial structure of canopy closure, total basal area, proportion of ponderosa pine, spruce/fir and aspen basal areas, height of the understory vegetation, and presence of tree regeneration. Cross-validation showed that each model had nominal prediction bias and that the estimation error variance could be used to assess uncertainty for new observations. Finally, we provide an example of how the models may be used to obtain unbiased estimates of forest structure where no sample plots are located.

INTRODUCTION

The way in which an animal perceives its environment, and thus selects habitat, ranges from general landscape features that define its broad surroundings to fine-scale habitat characteristics that result in specific use of a site (Svardson 1949, Hilden 1965). An understanding of a species' habitat requirements, therefore, entails knowing the extensive and intensive features responsible for habitat choice. Models of wildlife-habitat relations attempt to mimic these decisions at appropriate scales. However, it is typically site-specific knowledge of an animal's relationship to its habitat that leads to useful, predictive models about habitat use (Morrison et al. 1992, p. 119). Although generally sensitive to local conditions, such models provide the foundation for more general, theoretical models that may include fine, as well as broad scale, habitat information (Van Horne 2002).

Detailed habitat models are typically limited by the spatial resolution of the data on which the models are based. For wide-ranging animals, obtaining sufficient spatial coverage further complicates the modeling process, as time and resource limitations preclude detailed sampling over large areas. Use of remotely sensed information, such as multi-spectral satellite imagery, allows one to derive large amounts of habitat-based information over large areas; however, these sensors gather information at a fixed spatial resolution (e.g., 28.5 x 28.5-m's for Landsat satellite imagery) and some habitat choices (such as nest tree selection) may occur at scales finer than the spatial resolution of those data. Both Metzger (1997) and Joy et al. (2001, Chapter 1) improved the spatial resolution of satellite-based classifications of forest structure and composition, respectively, using fine-scale field data to drive the classification procedure.

Here, we describe the variability associated with forest structural components on the Kaibab National Forest (KNF) in Arizona, where a population of northern goshawks (*Accipiter gentilis*; hereafter referred to as goshawk) has been studied for 12 years (Reynolds et al. 1994, Joy et al. 1994, Reynolds and Joy 1998). Nearly all of the KNF has been altered by some form of management during the past 100 years (Crocker-Bedford 1990, White and Vankat 1993, Burnett 1991, Anonymous 2000). Our models of forest structure were based on field measurements at random locations and goshawk nest sites, and included Landsat TM bands, topographic measures such as slope, aspect, elevation and landform, and vegetation classes as auxiliary independent variables.

We first modeled the variability in forest structure using multiple linear regression models. We then modeled the residuals from the regression models using binary regression trees. The residuals for each model were assessed for spatial correlations. The final models predicted forest structural components thought to be important in goshawk-habitat decisions across the study area. To evaluate the predictive performance of the models, we used 11-fold cross-validation (Breiman et al. 1984, Efron and Tibshirani 1993:237-255, Guisan and Zimmermann 2000, Steele 2000). We then quantified and evaluated the feasibility of using the estimation error variance to explain estimation uncertainty. Binary data (presence/absence) of forest regeneration were modeled using logistic regression and evaluated using receiver operated characteristic curves (Egan 1975:142-149). Finally, we demonstrate the application of our models for estimating forest structure where no sample plots are located. The models we developed formed the foundation for further study of the spatial relationships between goshawk habitat use and the hawk's changing environment.

STUDY AREA

Our study was located on the Kaibab Plateau, an isolated and forested plateau in north-central Arizona. The study area (1,285 km^2) includes all KNF lands on the North Kaibab Ranger District (NKRD) above 2,182 m in elevation (Fig. 2.1). The study area was bounded on the south by the Grand Canyon National Park and escarpments of the Grand Canyon of the Colorado River, to the west and east by slopes that descend to tributary canyons of the Colorado River and shrub-steppe plains, and to the north by gentle slopes that descend to the plains. Five vegetation classes dominate the study area (Joy et al. 2001, Chapter 1): (1) pinyon (*Pinus edulis*)-juniper (*Juniperus* spp.) woodlands (106 km^2) occur at lower elevations (2,182-2,250 m) and mix with ponderosa pine (*P. ponderosa*) at higher, transitional zones; (2) ponderosa pine comprises over 50% (714 km^2) of the forested area and occurs between 2,250 and 2,550 m; (3) mixed conifer (145 km^2), which occurs between 2,550 and 2,650 m, is comprised of ponderosa pine mixed with white fir (*Abies concolor*), Douglas-fir (*Pseudotsuga mensiesii*) and/or aspen (*Populus tremuloides*), and spruce (*Picea pungens, P. englemannii*) at higher elevations; (4) spruce-dominated stands occur with sub-alpine fir (*A. lasiocarpa*) (130 km^2) above 2,650 m; (5) deciduous (aspen, *Quercus gambeli*)-dominated mixes (112 km^2) occur throughout the forest and are common where extensive disturbance (fire, logging) has occurred. A series of long, narrow meadows containing grasses and herbaceous vegetation occur within the forest.

The KNF receives about 64 cm of precipitation annually (White and Vankat 1993), with winter snowpacks of 2-3 m. A drought period typically occurs in May and June, followed by mid- to late-summer thunderstorms with frequent showers.

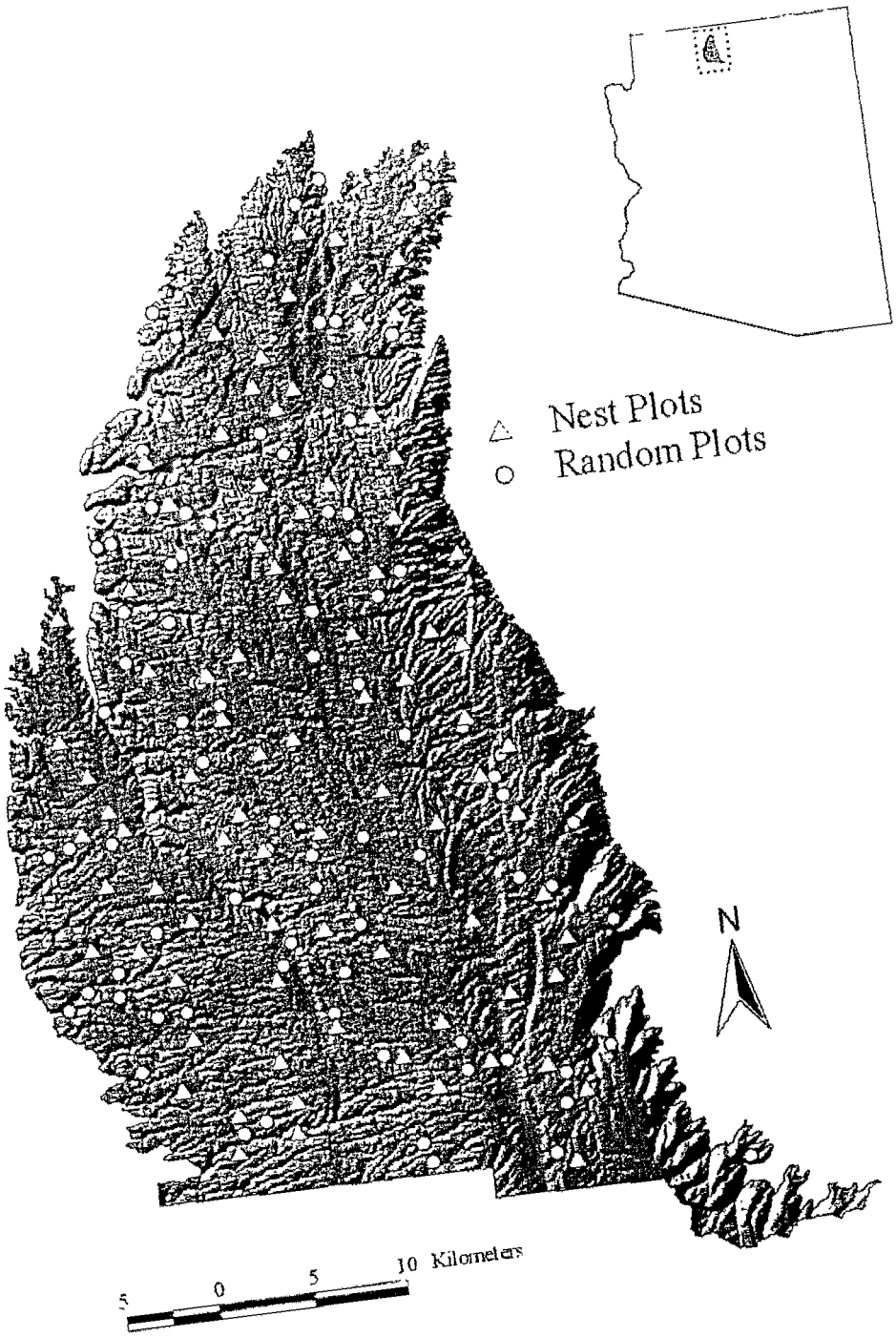


Figure 2.1. Location and map of the study area on the Kaibab National Forest (North Kaibab Ranger District) in Arizona, showing the distribution and spatial arrangement of sample plots. Triangles are plots at northern goshawk nests (n = 92) and circles are randomly located plots (n = 85).

Forest management began on the NKRD in the 1920s in the form of sanitation cuts and single-tree selection. These continued, along with occasional small (0.1 km^2) clearcuts, until the late-1970s when intensive stand-level management began. Shelterwood, seed, salvage, removal, and thinning cuts continued until 1991 when the NKRD implemented forest treatments (Reynolds et al. 1992) designed to enhance the habitat of goshawks and their prey. Livestock grazing was common on the NKRD between the late 1800s and the mid-1920s. The suppression of naturally-occurring fires started in the early 1900s and continues today.

METHODS

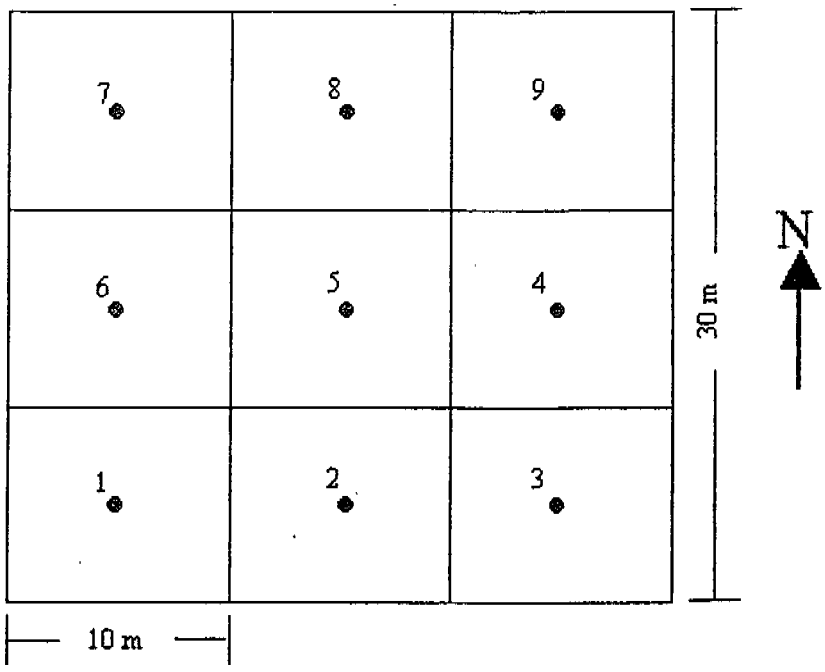
Data Collection

Field Data

During August and September 1997, we sampled (Fig. 2.1) the forest vegetation throughout the study area. We located 100 random plots, irrespective of goshawk territories. Despite our inability to sample 15 of the 100 plots along the edge of the study area due to steep terrain, plots at random sites were expected to capture the range of variability in forest structure on the study area. To sample structural components within goshawk nests sites, 95 plots at nest sites were also measured. Because goshawks use 2-7 alternate nests within territories (Reynolds et al. 1994, Reynolds and Joy 1998), we randomly selected one alternate nest within each of the 95 goshawk territories identified between 1991 and 1998 (Reynolds et al., in review) to measure the vegetative characteristics at nest sites. Due to data collection errors, only 92 of the 95 nest plots were used in the analyses. Although goshawk nest sites have been described as relatively

homogeneous (Reynolds et al. 1982, Moore and Henny 1983, Spieser and Bosakowski 1987, Hayward and Escano 1989), goshawks on the KNF appeared to nest in a variety of forest conditions (e.g., pure pine, mixed conifer, aspen, spruce-fir, etc.), reflective of the conditions available to them, reducing the potential for sampling bias towards homogeneous plots.

Random and nest-site plots (Fig. 2.1) were comprised of nine 10-m x 10-m sub-plots (Fig. 2.2) that, together, corresponded to a resampled 30-m x 30-m pixel on a Landsat TM image. Sub-plots within plots were considered correlated and dependent. Departures from independence are common in spatial data, and are accounted for in robust terms in the modeling process (Cressie 1991:3-4). Plots were established systematically in a north-south, east-west manner and their locations verified with a Trimble Navigation Pathfinder™ Asset Surveyor Global Positioning System with an estimated accuracy of 1-3 m. Vegetative characteristics measured on the sub-plots included basal area by tree species ($m^2 ha^{-1}$, measured with a 4.6 factor prism), maximum height of the understory vegetation (i.e., grasses, forbs, shrubs; m), the presence or absence of seedling and/or sapling trees (hereafter, referred to as "regeneration"), and percent canopy closure measured using a spherical densiometer (Lemmon 1956, 1957). Canopy closure was estimated by averaging four readings, one from each cardinal direction. The decision to use the densiometer's wide angle of view to estimate canopy closure was based on the desire to include the depth of canopy over a large area that corresponded somewhat with the remote sensing imagery (Jennings et al. 1999), while minimizing the influence of the height to the base of the live crown on the estimates



Plot and Sub-plots

Figure 2.2. Layout of the field plots used on the Kaibab National Forest (North Kaibab Ranger District), Arizona, showing the arrangement of sample sub-plots..

(Bunnell and Vales 1990). Total basal area was calculated for each sub-plot, and basal area of each tree species was expressed as a proportion of the total basal area.

GIS and Landsat TM Data

For each sub-plot we derived GIS grids of elevation, slope, aspect, and landform from digital elevation models (1:24,000, 30-m x 30-m spatial resolution; US Geological Survey) using ArcView® (ESRI 1998). Landform (McNab 1989) is an index that expresses surface shape as a measure of surface concavity (negative values) or convexity (positive values) creating a continuous variable. Surface shape was computed as the mean slope gradient from the original cell to adjacent cells in 4 directions using Avenue (ESRI 1998) code. Prior to generating grid coverages for each topographic variable, the elevation grid was resampled (RESAMPLE, nearest neighbor, Grid Module; ARC/INFO®, ESRI 1995) to a 10-m spatial resolution, corresponding to the resolution of the field data. Grids of spectral bands 1-5 and 7 from a cloud-free, 1997 Landsat TM image (22 June, Path 37, Row 35) were georectified to sub-pixel accuracy and resampled to a 30-m pixel resolution using cubic convolution (USDA Forest Service, Remote Sensing Lab., Albuquerque, NM; pers. comm.). We then resampled (nearest neighbor) each grid to a 10-m spatial resolution and averaged (FOCALMEAN, Grid Module; ARC/INFO®, ESRI 1995) the data values by moving a 3 x 3 pixel window over the resampled grids. Each 10-m x 10-m pixel of resampled Landsat data therefore represented an average of the surrounding 30-m x 30-m pixel, including the central 10-m pixel of the original 30-m Landsat pixel, whose value did not change. It was necessary to average the Landsat information because some sample plots fell within transition zones between vegetation classes and structural composition. The FOCALMEAN function

(ARC/INFO®, ESRI 1995) was preferred over other averaging techniques (e.g., bilinear interpolation, cubic convolution) because the averaging window did not alter the data values to the extent that weighted averaging (bilinear interpolation) or curve fitting (cubic convolution) did, resulting in loss of information. The chosen averaging method was intended to improve the relationship between the Landsat bands and dependent variables.

A grid of six vegetation classes, modeled to a 10-m spatial resolution, was derived from cluster analyses and decision trees (Joy et al. 2001, Chapter 1). Values for all grid layers of information were derived for each 10-m sub-plot using Avenue (ESRI 1998) code.

Modeling Forest Structure

Modeling of forest structure followed Metzger (1997), where the classification of remotely sensed imagery was based primarily on field data, and the corresponding spatial resolution at which those data were collected, and not on the resolution of the spectral scanners. Multiple linear regression analysis (OLS; Reich and Davis 1998; S-PLUS©, Statistical Sciences 2000) was used to explore the coarse-scale variability in continuous measures of forest structure (elevation, slope, aspect, landform, Landsat TM bands, and vegetation class). The base model to which all other models were compared was for ponderosa pine. To account for differences in the models for ponderosa pine and other vegetation classes with respect to the independent variables (elevation, slope, aspect, landform, and the six Landsat bands), dummy variables were included in the models (Neter et al. 1985:329-330). The inclusion of dummy variables allowed us to quantify classes of vegetation, model the different vegetation classes simultaneously, and test for significance among vegetation classes. For each component of forest structure modeled,

we used a stepwise procedure (elimination was based on a p-value ≤ 0.10) to identify a set of independent variables to include in the regression models.

We modeled the error (i.e., residuals) associated with the multiple linear regression models using binary regression trees (RT) (TREE; S-PLUS©, Statistical Sciences 2000). The RT is a non-parametric approach to regression that compares all possible splits among the independent (continuous) variables using a binary partitioning algorithm that maximizes the dissimilarities among groups (Hansen et al. 1996, De'Ath and Fabricius 2000). Once the algorithm partitions the data into new subsets, new relationships are developed, assessed, and split into new subsets. The algorithm recursively splits the data in each subset until either the subset is homogeneous or the subset contains too few observations (e.g., < 5) to be split further. Interpolation using RTs is relatively insensitive to missing data (Friedl and Brodley 1997, De'Ath and Fabricius 2000). Independent variables considered in the RT included elevation, slope, aspect, landform, Landsat TM bands, and vegetation class, the latter being treated as a categorical variable. Classification trees have been used successfully to interpret Landsat imagery in the past (Michaelson et al. 1994, Hansen et al. 1996, Friedl and Brodley 1997). To avoid over-fitting the models, a 10-fold cross-validation procedure (De'Ath and Fabricius 2000) was used to estimate the prediction error for a given tree size. Estimates of the prediction error were then plotted against the sequence of tree sizes. The tree was then pruned to the size that minimized the deviance (prediction error) (De'Ath and Fabricius 2000). Although, we recognize that cross-validation does not identify the absolute minimum deviance associated with the trees, we used cross-validation as a guideline to select the best tree size that would approximate the minimum deviance.

Semi-variograms were used to evaluate spatial dependencies among the residuals from the various models. If the residuals exhibited spatial dependencies, a spatial autoregressive (SAR) model was used to obtain generalized least squares estimates of the regression coefficients (Upton and Fingleton 1985:349-373). The model residuals were then reevaluated to ensure the removal of the spatial dependencies. In fitting the SAR models, a spatial weight matrix (i.e., a block diagonal matrix) was used in which diagonal elements represented the spatial arrangements of sub-plots within each cluster plot. We adapted the queen's definition of contiguity to define the spatial dependency among the sub-plots (Upton and Fingleton 1985:349-373). Off-diagonal elements were assigned a value of zero, which assumed no spatial dependency between cluster plots. This assumption was reasonable as all but one pair of cluster plots were more than 100 m apart (interquartile range = 1006-1925 m).

Grids representing the various components of forest structure were generated (CON; Grid Module, ARC/INFO®, ESRI 1995) using the estimated model parameters. Similarly, grids representing the error in each multiple linear regression model were generated (CON; Grid Module, ARC/INFO®, ESRI 1995) by passing each grid for the appropriate independent variable through the RTs. Our final, predicted surfaces for each continuous variable of forest structure were obtained from the sum of the grids associated with the multiple linear regression models and RTs.

The effectiveness of the final models was evaluated using a goodness-of-prediction statistic (G) (Agterberg 1984, Kravchenka and Bullock 1999, Guisan and Zimmermann 2000, Schloeder et al. 2001). The G-value, measures how effective a

prediction might be relative to that which could have been derived using the sample mean (Agterberg 1984):

$$G = \left(1 - \left\{ \sum_{i=1}^{n} \left[Z_i - \hat{Z}_i\right]^2 \Big/ \sum_{i=1}^{n} \left[Z_i - \bar{Z}\right]^2 \right\}\right), \qquad [2.1]$$

where Z_i is the observed value of the ith observation, $\hat{Z}_i$ is the predicted value of the ith observation, and $\bar{Z}$ is the sample mean. A G-value equal to 1 indicates perfect prediction, a positive value indicates a more reliable model than if one had used the sample mean, a negative value indicates a less reliable model than if one had used the sample mean, and a value of zero indicates that the sample mean should be used to estimate Z.

We used a logistic regression model (GLM; S-PLUS© Statistical Sciences 2000) to predict the probability of observing tree regeneration. A categorical variogram (Soares 1992) was used to test for spatial dependencies in regeneration on the sample plots, prior to fitting the logistic regression model, and to account for spatial dependency that might exist among the sample plots when fitting the logistic regression model. We used a combination of forward and backward elimination to identify the independent variables that were important in discriminating between the presence and absence of regeneration. A receiver operated characteristic curve (ROC) (Egan 1975:142-149), which shows the true-positive rate (TPR) versus the false-positive rate (FPR), was used to evaluate the performance of the model for different thresholds of prediction:

$$TPR = \frac{\sum_{i=1}^{n} I_{\hat{x}_i > b} I_{x_i = 1}}{\sum_{i=1}^{n} x}$$

$$FPR = \frac{\sum_{i=1}^{n} I_{\hat{x}_i > b} I_{x_i = 0}}{\sum_{i=1}^{n} (1 - x_i)} \qquad [2.2]$$

In the above equations, TPR is equal to the number of sub-plots where the estimated probability of regeneration presence is greater than the threshold, b, ($0 \leq b \leq 1$) for sites where regeneration was truly present ($x = 1$), divided by the total number of sub-plots where regeneration was truly present. FPR is equal to the number of sub-plots where the estimated probability of regeneration presence is greater than the threshold, b, for sites where regeneration was truly absent ($x = 0$), divided by the total number of sub-plots where regeneration was truly absent. The greater the height of the ROC curve above the TPR=FPR line, the better the model discriminates between true presence and true absence. The threshold value, b, that maximized the distance between the ROC curve and the TPR=FPR line was used to generate a surface showing the presence or absence of regeneration on the KNF. The classification of this surface was performed in ARC/INFO® (ESRI 1995) using AML (CON; Grid Module).

Model Evaluation

Stratified 11-fold cross-validation (Efron and Tibshirani 1993:237-255) was used to estimate the prediction error for each component of forest structure, except regeneration. The data were sorted *apriori* by sampling method (random and fixed) and cluster number to mimic the sampling protocol used in the field (Kish and Frankel 1974). The data were split into K=11 parts consisting of approximately 17 clusters each. For

each kth part, the regression and RT models were fitted to the remaining K-1=10 parts of the data. The fitted model was used to predict the kth (i.e., removed) part of the data. This process was repeated 11 times so that each observation was excluded from the model construction step and its response predicted.

To evaluate the effectiveness of the models, we computed various measures of the prediction error. Prediction bias (Williams 1997) was calculated for each validation data set as a percentage of the true value:

$$B = \frac{100\sum_{i=1}^{n}\left(Z_i - \hat{Z}_i^{-k(i)}\right)}{\sum_{i=1}^{n} Z_i}, \qquad [2.3]$$

where Z_i is the observed value, and $\hat{Z}_i^{-k(i)}$ denotes the estimated value for observation i (i = 1, 2, …, n) computed with the k(i)th part of the data removed. Accuracy (Kravchenko and Bullock 1999, Schloeder et al. 2001) was measured by the mean absolute error (MAE), which is a measure of the sum of residuals (i.e., actual minus predicted):

$$MAE = \frac{1}{n}\sum_{i=1}^{n}\left[\left|Z_i - \hat{Z}_i^{-k(i)}\right|\right], \qquad [2.4]$$

and the root mean squared error (RMSE), which measured of the square root of the sum of squared residuals:

$$RMSE = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left[Z_i - \hat{Z}_i^{-k(i)}\right]^2}\ . \qquad [2.5]$$

Small MAE values indicate a model with few errors, while small values of RMSE indicate more accurate predictions on a point-by-point basis (Schloeder et al. 2001). To assess the estimation uncertainty in the models (Isaaks and Srivastava 1989:231-234) we

calculated the estimation error variance (EEV), $\hat{\sigma}_i^{2(-k(i))}$, for each observation in the kth part of the data:

$$\hat{\sigma}_i^{2(-k(i))} = MSE^*\left[\left(X_i^{-k(i)}\right)\left(X^{*\prime}X^*\right)\left(X_i^{-k(i)}\right)+1\right]+MSE(RT)+2Cov\left(\hat{Y},\hat{\eta}\right), \qquad [2.6]$$

where MSE^* is the regression mean squared error for the TS model fitted using K-1 parts of the data, X^* is a matrix of independent variables associated with the K-1 parts of the data, $X_i^{-k(i)}$ is a vector of independent variables associated with the ith observation in the kth part of the data, *MSE(RT)* is the mean squared error of the RT used to describe the error in the multiple linear regression model, and $Cov\left(\hat{Y},\hat{\eta}\right)$ is the covariance between the estimated values, $\left(\hat{Y}\right)$, from the TS model and the predicted residuals, $\left(\hat{\eta}\right)$, from the RT for the K-1 parts of the data. The consistency between the EEV ($\hat{\sigma}_i^{2(-k(i))}$), and the observed estimation errors (i.e., true errors), $e_i^{-k(i)} = \left(Z_i - Z_i^{-k(i)}\right)$, was calculated using the standard mean squared error (SMSE) (Hevesi et al. 1992):

$$\text{SMSE} = \frac{1}{n}\sum_{i=1}^{n}\frac{\left(e_i^{-k(i)}\right)^2}{\hat{\sigma}_i^{2(-k(i))}}. \qquad [2.7]$$

EEVs were assumed consistent with true errors if the SMSE fell within the interval $\left[1 \pm 2(2/n)^{1/2}\right]$ or [0.928 - 1.072] (Hevesi et al. 1992). Paired t-tests ($\alpha = 0.05$) were used to test for differences between the mean estimation errors and zero. The EEVs were also used to construct 95% confidence intervals around individual estimates. Coverage rates were calculated as the proportion of individual confidence intervals that contained the true value.

Ten-fold cross-validation was performed to evaluate the predictive performance of the logistic regression model in discriminating between the presence and absence of

tree regeneration. A Kappa statistic, which estimates the difference between the observed and estimated values, and the agreement contributed by chance, was calculated for both the fitted model and prediction based on cross-validation.

Estimating Means and Variances

In addition to being able to assess the level of uncertainty associated with our models, it is also important that the models are capable of providing unbiased estimates at any spatial scale or level of support. It is also important that we are able to place bounds on the error of estimation. To demonstrate this concept, we will assume that one is interested in estimating a mean (e.g., canopy closure, basal area, height understory vegetation, etc.) per sub-plot and placing a bound on the error of estimation. A sample of n 30-m x 30-m cluster plots consisting of $m = 9$ 10-m x 10-m sub-plots are randomly selected from one of the modeled surfaces. The modeled surfaces are used to provide an estimate ($\hat{z}$) on each of the nm sub-plots, along with the model prediction variance ($\hat{\sigma}^2$) using Eq. 2.6. An estimate of the mean value per sub-plot ($\hat{\bar{z}}_{sp}$) is given by

$$\hat{\bar{z}}_{sp} = \frac{1}{nm}\sum_{i=1}^{n}\sum_{j=1}^{m}\hat{z}_{ij} = \frac{1}{n}\sum_{i=1}^{n}\hat{\bar{z}}_i\,, \qquad [2.8]$$

where $\hat{z}_{ij}$ is the estimated value on the jth sub-plot from cluster i, and $\hat{\bar{z}}_i$ is the average for the ith cluster. If clusters of the same size are sampled, the total sum of squares associated with estimating the mean can be partitioned into the within-cluster sum of squares (SSW) and between-cluster sum of squares (SSB) (Scheaffer et al. 1996:299-301). With appropriate divisors, these sum of squares become the usual mean squares of an analysis of variance. The within-cluster mean square (MSW) is given by

$$MSW = \frac{SSW}{n(m-1)} = \frac{1}{n(m-1)}\sum_{i=1}^{n}\sum_{j=1}^{m}\left(z_{ij} - \bar{z}_i\right)^2 \approx \frac{1}{nm^2}\sum_{i=1}^{n}\sum_{j=1}^{m}\hat{\sigma}_{ij}^2, \qquad [2.9]$$

where $\frac{1}{n(m-1)}\sum_{i=1}^{n}\sum_{j=1}^{m}\left(z_{ij} - \bar{z}_i\right)^2$ is the MSW one would typically use in cluster sampling and $\frac{1}{nm^2}\sum_{i=1}^{n}\sum_{j=1}^{m}\hat{\sigma}_{ij}^2$ is its equivalent using the EEV formula. The between-cluster mean square (MSB) is given by

$$MSB = \frac{SSB}{n-1} = \frac{m}{n-1}\sum_{i=1}^{n}\left(\bar{z}_i - \bar{z}_{sp}\right)^2 \approx \frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{m}\hat{\sigma}_{ij}^2, \qquad [2.10]$$

where $\frac{m}{n-1}\sum_{i=1}^{n}\left(\bar{z}_i - \bar{z}_{sp}\right)^2$ is the general formula for calculating the MSB and $\frac{1}{n}\sum_{i=1}^{n}\sum_{j=1}^{m}\hat{\sigma}_{ij}^2$ is its equivalent using the EEV formula. The MSB can be used to calculate the variance of $\hat{\bar{z}}_{sp}$ as follows:

$$\hat{V}\left(\hat{\bar{z}}_{sp}\right) = \frac{MSB}{nm}. \qquad [2.11]$$

The estimates of the population mean possess special properties when all of the cluster plots are of equal size (i.e., 30 m x 30 m). First, the estimator $\hat{\bar{z}}_{sp}$ is an unbiased estimator of the population mean; and second, $\hat{V}\left(\hat{\bar{z}}_{sp}\right)$ is an unbiased estimator the variance of $\hat{\bar{z}}_{sp}$. If cluster plots of unequal sizes are used, alternative formulas need to be applied to estimate the mean and variance, and these estimates may be biased due to the unequal cluster sizes (Scheaffer et al. 1996:299-301).

RESULTS

Modeling Forest Structure

Summary statistics associated with the topographic variables used in modeling forest structure varied as expected (Table 2.1) based on digital elevation models (U.S. Geological Survey) for the study area. Slopes ranged from 0.5% to 32%, with steeper areas found along the edges of the study area, and in canyons and drainages. Elevations ranged from 2,182 m to 2,805 m at subplots, with higher elevations represented in the south-central portion of the KNF, near the border with Grand Canyon National Park. Landform indices at subplots indicated greater concavity (5.6) than convexity (-2.8), and were highly variable (CV%=603.4).

The Landsat TM scanners simultaneously collected radiance values [recorded as digital numbers (DN)] from the study area surface in each of the 6 bands of interest (Table 2.1), which included the blue-green (band 1), green (band 2), red (band 3), near-infrared (band 4), and mid-infrared (bands 5 and 7) wavelengths (Avery and Berlin 1992:141). Reflectance values measured by each band had a maximum range of 22-214 digital numbers (DN): band 1 – 65-148 DN; band 2 – 25-83 DN; band 3 – 23-129 DN; band 4 – 54-124 DN; band 5 – 50-214 DN; band 7 – 22-124 DN.

The final models used to describe forest structure on the KNF (Fig. 2.3, 2.4) included the continuous variables: percent canopy closure (CC), total tree basal area (TBA), proportion of ponderosa pine (PPB) and spruce/fir (PSB) basal areas, and maximum height of the understory vegetation (HUV), and the binary variable for the presence/absence of regeneration. Due to the sparseness of deciduous forest types (i.e., aspen or oak) on the sample plots (aspen occurred on 25 % and oak on 1 % of the

Table 2.1. Summary statistics [minimum, mean, maximum, standard deviation (SD), and percentage coefficient of variation (CV%)] showing the range of observed topographic and remotely sensed data from 1556 sample subplots on the Kaibab National Forest (North Kaibab Ranger District), Arizona.

Variable	Minimum	Mean	Maximum	SD	CV%
Topography					
Slope (%)	0.480	7.202	32.53	4.930	68.45
Aspect (degrees)	0	186.7	358.4	116.5	62.39
Elevation (m)	2182	2482	2805	156.2	6.290
Landform	-2.800	0.137	5.600	0.826	603.4
Landsat TM (Digital Numbers)					
Band 1 (0.45-0.52 μm)	65.22	75.25	148.0	9.029	12.00
Band 2 (0.52-0.60 μm)	24.67	31.77	83.00	5.919	18.63
Band 3 (0.63-0.69 μm)	23.44	36.54	129.3	11.09	30.35
Band 4 (0.76-0.90 μm)	53.67	71.23	123.7	9.623	13.51

Table 2.1 (continued)

Band 5 (1.55-1.75 μm)	50.00	92.94	214.0	25.57	27.52
Band 7 (2.08-2.35 μm)	22.00	43.06	124.3	16.07	36.47

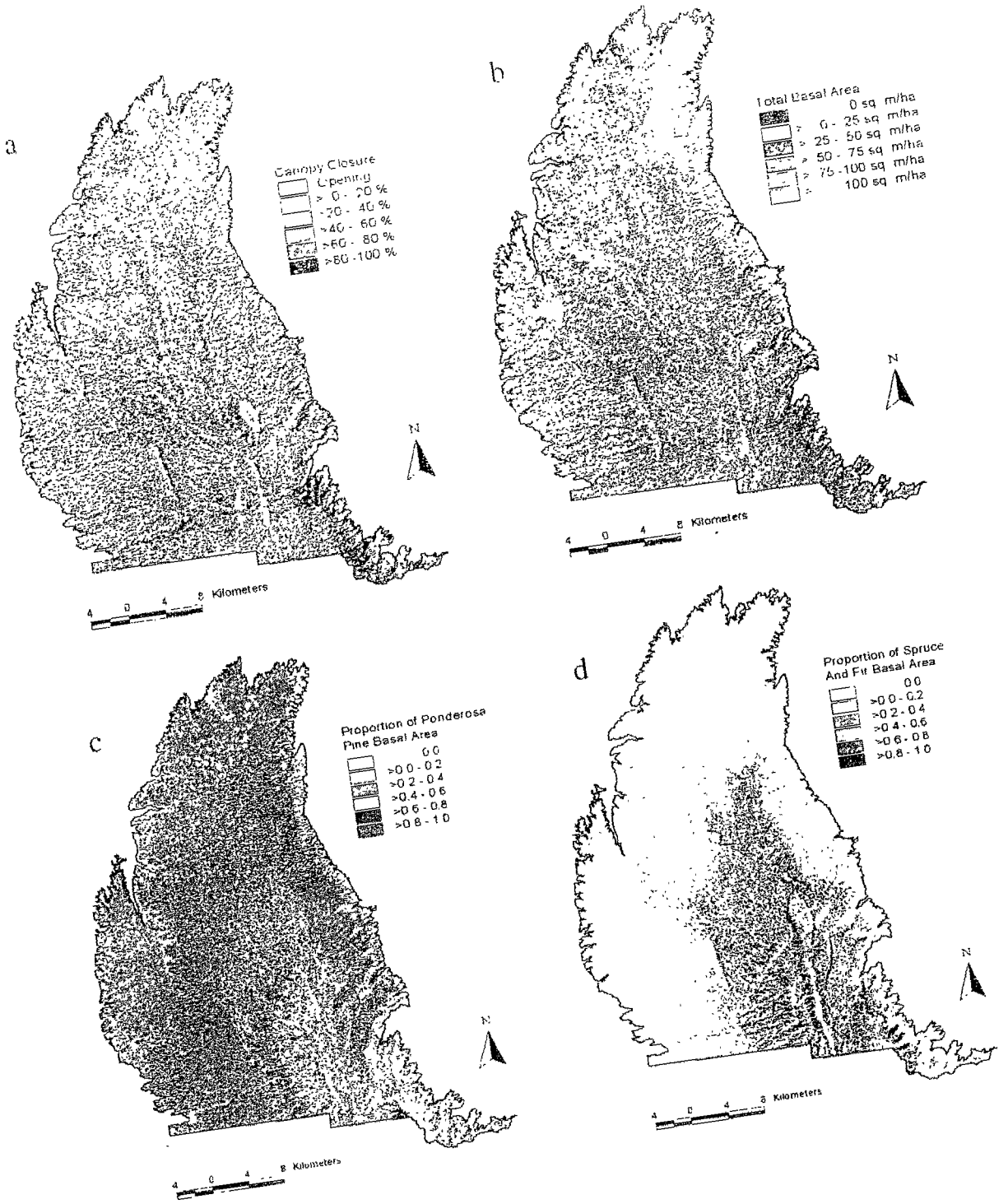


Figure 2.3. Predictions of forest structure for (a) percent canopy closure, (b) total basal area (m^2 ha^{-1}), (c) proportion of ponderosa pine, and (d) proportion of spruce and fir basal area on the Kaibab National Forest (North Kaibab Ranger District), Arizona.

Figure 2.4. Predictions of forest structure for (a) proportion of aspen basal area, (b) maximum understory height (m), and (c) presence of seedling and/or sapling trees on the Kaibab National Forest (North Kaibab Ranger District), Arizona.

subplots), we were not able to model the proportion of deciduous basal area directly. However, because aspen was the only other dominant vegetation class encountered on the study area, other than the classes we modeled (note that components of the mixed conifer class were accounted for in the PPB and PSB structural classes), we estimated the proportion of aspen basal area as 1-(PPB+PSB). Slope, elevation, landform, and Landsat TM Bands 3 and 7, and mixed conifer and spruce-dominated vegetation classes were the most frequently occurring variables (Table 2.2) in the multiple linear regression models. All independent variables were used in one or more of the TS models. All of the topographic, Landsat TM, and vegetation class variables (Table 2.3) were also used in RTs to describe the error in one or more of the multiple linear regression models. Notably, however, none of the vegetation classes contributed to classifying residuals in the models for the proportion of spruce/fir basal area and maximum understory height, indicating that the regression tree models for these components could not account for additional structural variability due to species composition. Vegetation classes were important in nearly all other models of forest structure, however. The tree sizes selected to minimize the deviance (prediction error) in the RTs ranged from 23 to 49 splits (Fig. 2.5).

The overall contribution of the linear regression and binary regression tree models (Table 2.4) in describing forest structure varied with the model. The multiple linear regression models alone explained 36% (HUV) to 76% (PPB) of the observed variability in forest structural components. The RT models accounted for an additional 4% (PSB) to 18% (CC) of the unexplained variability in the multiple linear regression models, based on the proportion to which the RT contributed to the final effectiveness of the models.

Table 2.2. Topographic characteristics, Landsat TM data, and vegetation classes used to develop the multiple linear regression models of forest structure and a logistic regression model of the presence of seedlings and/or saplings on the Kaibab National Forest (North Kaibab Ranger District), Arizona. The reference model for vegetative types is that for ponderosa pine. Dots indicate where intercepts differed from the reference model, light gray cells indicate where slopes differed from the reference model, and dark gray cells indicate where slopes and intercepts differed from the reference model. Significant indicator relationships (dummy variables) between vegetation types and main variables are provided. Independent variables with a p-value ≤ 0.10 were retained in the model.

Regression Model	Topography[1]				Landsat TM Data[2]						Vegetation Class[3]				
	S	A	E	L	B1	B2	B3	B4	B5	B7	PJ	MC	SD	DD	OP
Canopy Closure[4]	PJ, SD, OP	OP		PJ, OP	SD	SD	SD, OP		SD	DD					
Total Basal Area (BA) (ft^2/acre)	MC	PJ, MC, SD	SD, OP	•	MC, SD	MC	SD	MC	MC, OP	•					
Ponderosa Pine BA[4]	•	MC	PJ, MC, OP	PJ, DD, OP			PJ	PJ, DD	MC, OP	MC			•		
Spruce/Fir BA[4]	MC	MC, SD, DD	MC, DD	•	MC	MC, SD, DD	SD	DD	MC, SD	MC, DD					
Understory Height	•		•	MC, OP		•	SD, DD	MC, OP	MC, DD	PJ, MC					
Seedlings/Saplings[5]		•	•		•		•		•	•	•	•	•	•	•

[1] S – slope, A – aspect, E – elevation, L – landform.

Table 2.2. (continued)

[2] B1 – band 1, B2 – band 2, B3 – band 3, B4 – band 4, B5 – band 5, B7 – band 7.

[3] PJ – pinyon-juniper, MC – mixed conifer, SD – spruce-dominated mix, DD – deciduous-dominated mix, OP – opening.

[4] Proportion.

[5] Presence/absence.

Table 2.3. Topographic characteristics, Landsat TM data, and vegetation classes used in regression trees to describe the residuals in the multiple linear regression models of forest structure on the Kaibab National Forest (North Kaibab Ranger District), Arizona.

Model Residuals	Topography[1]				Landsat TM Data[2]						Vegetation Class[3]					
	S	A	E	L	B1	B2	B3	B4	B5	B7	PJ	PP	MC	SD	DD	OP
Canopy Closure[4]	•	•	•	•	•	•	•			•		•	•		•	•
Total Basal Area (BA) (ft^2/ac)	•	•	•	•		•	•	•		•			•	•	•	
Ponderosa Pine BA[4]	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•	•
Spruce/Fir BA[4]	•	•	•	•	•	•		•								
Understory Height	•	•	•	•	•	•	•	•		•						

[1] S – slope, A – aspect, E – elevation, L – landform.

[2] B1 – band 1, B2 – band 2, B3 – band 3, B4 – band 4, B5 – band 5, B7 – band 7.

[3] PJ – pinyon-juniper, PP – ponderosa pine, MC – mixed conifer, SD - spruce-dominated mix, DD – deciduous-dominated mix, OP – opening.

[4] Proportion.

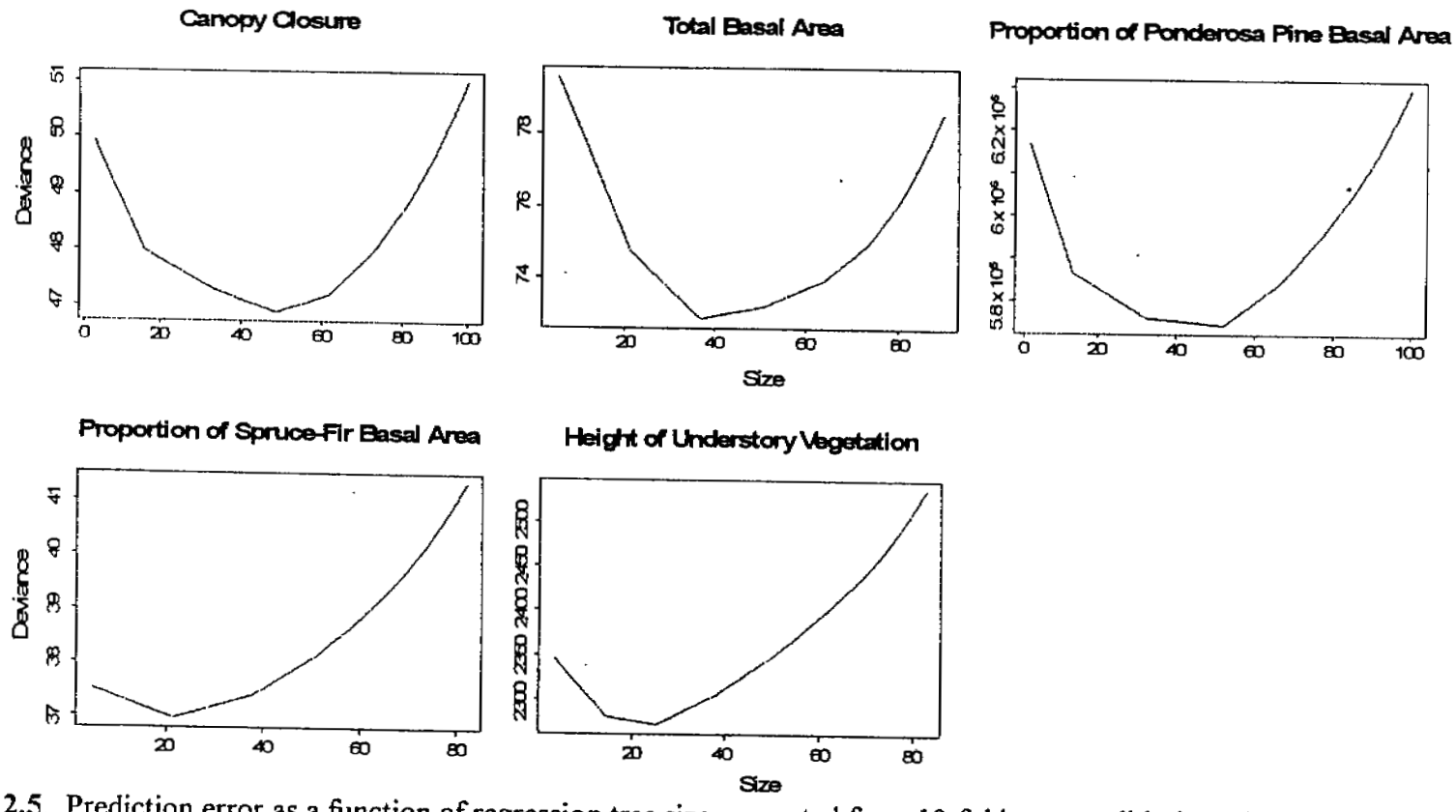


Figure 2.5. Prediction error as a function of regression tree size generated from 10-fold cross-validation of the residuals from the multiple linear regression models of percent canopy closure, total basal area, proportion of ponderosa pine and spruce/fir basal areas, and maximum height of the understory vegetation on the Kaibab National Forest (North Kaibab Ranger District), Arizona. Regression tree size was selected to minimize the deviance (i.e., prediction error).

Table 2.4. Overall effectiveness (G-value) of the multiple linear regression models (LS) and regression trees (RT) used to describe continuous variables of forest spatial structure on the Kaibab National Forest (North Kaibab Ranger District), Arizona. LS models were fit using either ordinary least squares or spatial autoregressive (SAR) models. The small standard errors $s(\hat{\lambda})$ associated with the spatial autocorrelation coefficients from the SAR models suggested that use of the SAR models was appropriate.

Model	G-statistic		SAR[2]	
	LS	(LS + RT)	$\hat{\lambda}$	$s(\hat{\lambda})$
Canopy Closure[1]	0.534	0.712	—	—
Total Basal Area (BA) (m^2/ha)	0.397	0.565	—	—
Ponderosa Pine BA[1]	0.763	0.830	0.407	0.025
Spruce/fir BA[1]	0.736	0.772	0.354	0.026
Understory Height (m)	0.361	0.452	0.452	0.025

[1] Proportion.

[2] $\hat{\lambda}$ = spatial autocorrelation coefficient $(-1 \le \hat{\lambda} \le 1)$, $s(\hat{\lambda})$ = estimated standard error.

Overall model performance ranged from a low of 0.45 for HUV to a high of 0.83 for PPB. The remaining three models had G-values ranging from 0.56 to 0.77.

The analysis of residuals from the individual models data indicated that three of the models (PPB, PSB, HUV) had spatially autocorrelated residuals. The regression models for these components were refitted using a SAR model (Table 2.4). The small standard errors, $s(\hat{\lambda})$, associated with the estimates of the spatial autocorrelation coefficient ($\hat{\lambda}$) for the PPB, PSB, and HUV models suggested that use of the SAR model was appropriate. All model residuals were approximately normally distributed. We identified no significant spatial correlations in the sample data for the presence of seedling/sapling trees (Fig. 2.6).

Model Evaluation

Summary statistics for the various components of forest structure modeled in this study (Table 2.5) were consistent with what was expected based on forest inventory data on the KNF (Resource Information System, NKRD; K. Fuelling, D. Steffensen, pers. comm.). Mean observed canopy closure was 62%, and ranged from no canopy (openings) to 100% closure. Total observed tree basal area ranged from 0 to 96 m^2 ha^{-1} on the sample subplots, while the mean observed proportions of total basal area by species ranged from a high of 0.62 (PPB) to a low of 0.17 (PSB). On individual sample plots, the proportions of the total basal area by species ranged from 0.0 to 1.0. The understory vegetation reached a mean maximum height of 0.56 m. The coefficient of variation was also reasonably small for all classes of forest structure, except PSB (observed CV% = 349), which was large. Spruce and fir generally made up only a small proportion of the total basal area on a given sample plot, but there were some sample

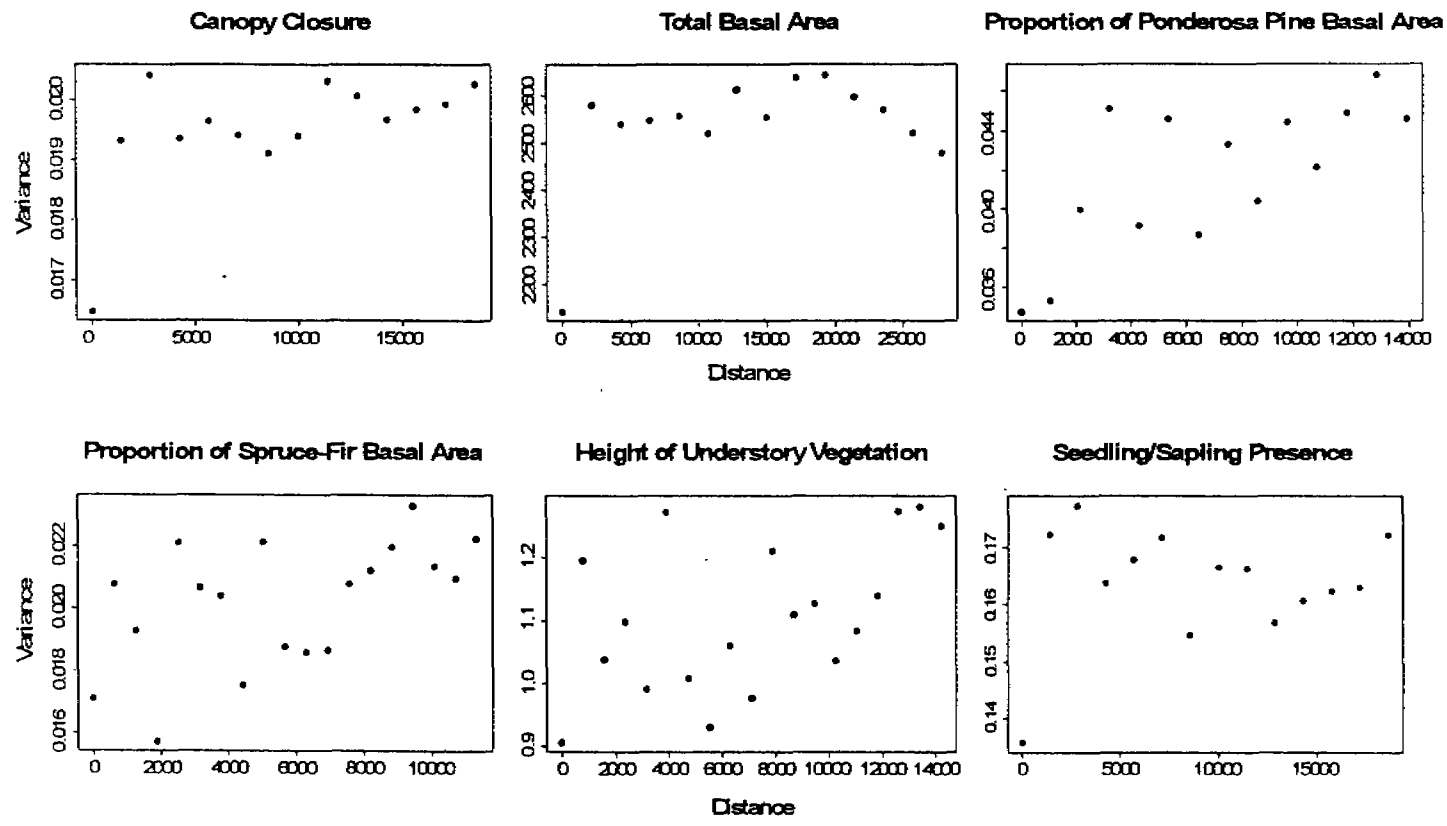


Figure 2.6. Semi-variogram plots for residuals from the multiple linear regression models of percent canopy closure, total basal area, proportion of ponderosa pine and spruce/fir basal areas, and maximum height of the understory vegetation, and the logistic regression of seedling and/or sapling presence on the Kaibab National Forest (North Kaibab Ranger District), Arizona, after correcting for spatial correlations in the residuals..

Table 2.5. Summary statistics and prediction bias showing the level of agreement between observed (obs) and estimated (est) models of forest structure on the Kaibab National Forest (North Kaibab Ranger District), Arizona. Prediction bias, based on estimated values from 10-fold cross-validation procedures, was nominal for all models.

	CC[1]		TBA[1]		PPB[1,2]		PSB[1,2]		HUV[1,2]	
Statistic	Obs	Est	Obs	Est	Obs	Est	Obs	Est	Obs	Est
Number	1556	1556	1556	1556	1556	1556	1556	1556	1556	1556
Mean	0.625	0.625	29.27	28.88	0.623	0.643	0.166	0.168	0.565	0.564
SD	0.264	0.214	18.67	15.04	0.421	0.352	0.286	0.247	0.425	0.376
CV(%)	42.20	34.30	14.61	11.93	67.60	54.70	349.1	147.2	22.93	66.70
Minimum	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
First quartile	0.489	0.498	13.74	19.88	0.133	0.306	0.0	0.0	0.305	0.382
Median	0.692	0.687	27.48	27.59	0.857	0.828	0.0	0.039	0.381	0.475

Table 2.5. (continued)

Third quartile	0.827	0.772	41.22	37.88	1.0	0.945	0.250	0.294	0.686	0.626
Maximum	0.998	1.0	96.18	80.01	1.0	1.0	1.0	1.0	3.658	3.485
Bias (%)		0.00		1.33		-3.21		-1.20		0.18

[1] CC – proportion of canopy closure, TBA – total basal area (m^2/ha), PPB – proportion ponderosa pine basal area, PSB – proportion spruce/fir basal area, HUV – height of understory vegetation (m).

[2] Models account for the presence of spatially autocorrelated errors.

plots on which spruce or fir dominated. resulting in distributions highly skewed to the right. The mean predicted proportion of aspen basal on the study area was 2.0 (range 0.0-1.0, SD = 0.32).

Prediction bias was nominal (Table 2.5) for all models. Minimum, maximum, and quartile values showed that estimated and observed value distributions were similar for all models. Because only three (CC, PPB, and PSB) of the five models of continuous data had similar units, further comparisons of performance were possible only for those three models. Estimation errors (Table 2.6) for the proportional variable CC had the greatest spread compared to the estimation errors for the models of PSB and PPB. Except for the PPB model, the mean estimation errors for the models did not differ significantly from zero (p-value ≥ 0.05). MAE values for the models (Table 2.6) suggested that PPB was the least accurate model of the three compared, while the model for PSB was most accurate. RMSE results showed that the model for CC resulted in more accurate predictions on a point-by-point basis than the other models. The distribution of errors from the cross-validation were approximately normally distributed (Fig. 2.7).

Except for the PSB model, the EEV overestimated the true errors (SMSE < 1) by as much as 14% (Table 2.6). The 0.95 confidence interval coverage rates for these models ranged from 93% to 96%. The lowest coverage rate (90%) was associated with the PSB model in which the EEV underestimated the true errors by 13% (SMSE > 1). These results suggest that the EEV could be used to assess estimates of uncertainty for new observations and to insure a 95% confidence interval around our estimates.

Table 2.6. Summary statistics for the estimation errors of the models of forest structure[1] on the Kaibab National Forest (North Kaibab Ranger District), Arizona, based on 10-fold cross-validation of the models. Mean estimation errors for all models, except the proportion of ponderosa pine basal area, did not differ (p-value ≥ 0.05) from zero.

Statistic	CC	TBA	PPB[2]	PSB[2]	HUV[2]
Number	1556	1556	1556	1556	1556
Mean	0.002	0.333	-0.020	-0.004	0.001
Interquartile range	0.248	22.02	0.236	0.049	0.364
MAE[3]	0.163	13.99	0.192	0.128	0.319
RMSE[4]	0.213	18.25	0.292	0.232	0.506
SMSE[5]	0.856	0.870	0.867	1.131	0.907
0.95 confidence interval coverage rate	0.954	0.958	0.926	0.902	0.939

[1] CC – proportion of canopy closure, TBA – total basal area m^2/ha), PPB – proportion ponderosa pine basal area, PAB – proportion aspen basal area, PSB – proportion spruce-fir basal area, HUV – height of understory vegetation (m).

[2] Models accounting for the presence of spatially autocorrelated errors.

[3] Mean absolute error.

[4] Root mean squared error.

[5] Standardized mean squared error.

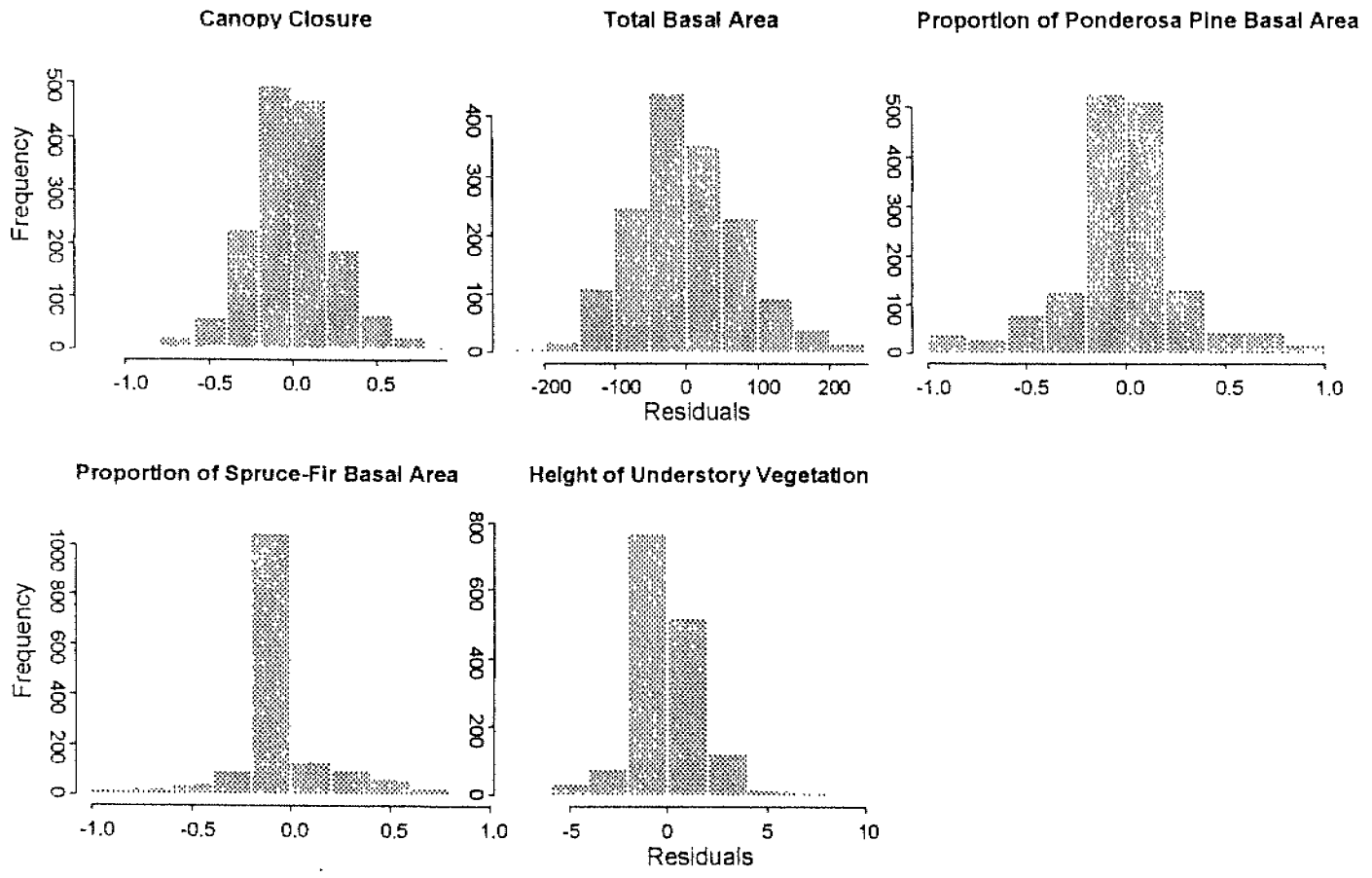


Figure 2.7. Histograms showing the distribution of errors from the 11-fold cross-validation of the models of percent canopy closure, total basal area, proportion of ponderosa pine and spruce/fir basal areas, and maximum height of the understory vegetation on the Kaibab National Forest (North Kaibab Ranger District), Arizona, after correcting for spatial correlations in the residuals.

The ROC curve (Fig. 2.8) used to evaluate the logistic regression model of tree understory presence showed that a high true-positive rate of prediction could be achieved, but only at the expense of an increased false-positive rate. The threshold value, *b*, that maximized the difference between the ROC curve and the TPR = FPR line was 0.485. Using this threshold value, the overall accuracy of the logistic regression model in discriminating tree understory presence from absence was 74% (Table 2.7). The Kappa statistic was 0.30. Accuracies ranged from 35% (absence) to 92% (presence) for the individual classes. The greatest confusion in fitting the model was associated with the absence of regeneration, which was confused with the presence of regeneration. Ten-fold cross-validation (Table 2.7) yielded similar results. Overall cross-validation accuracy was 72% (Kappa = 0.22), with accuracies for individual classes ranging from a low of 24% to a high of 94%. Again, the greatest confusion was associated with predicting the absence of regeneration.

Estimating Means and Variances

The models developed in this paper can be used to obtain unbiased estimates of forest structure at a 10-m spatial resolution anywhere on the study area, and place a bound on the error of estimation. We demonstrate the process using the model for canopy closure, but any component of forest structure could be used. Grid values from our canopy model were extracted at n plots. For comparative purposes, we randomly select the same number of plots (n = 174 30-m x 30-m cluster plots consisting of m = 9 10-m x 10-m sub-plots) as in our original sampling scheme, but any number of plots could be used.

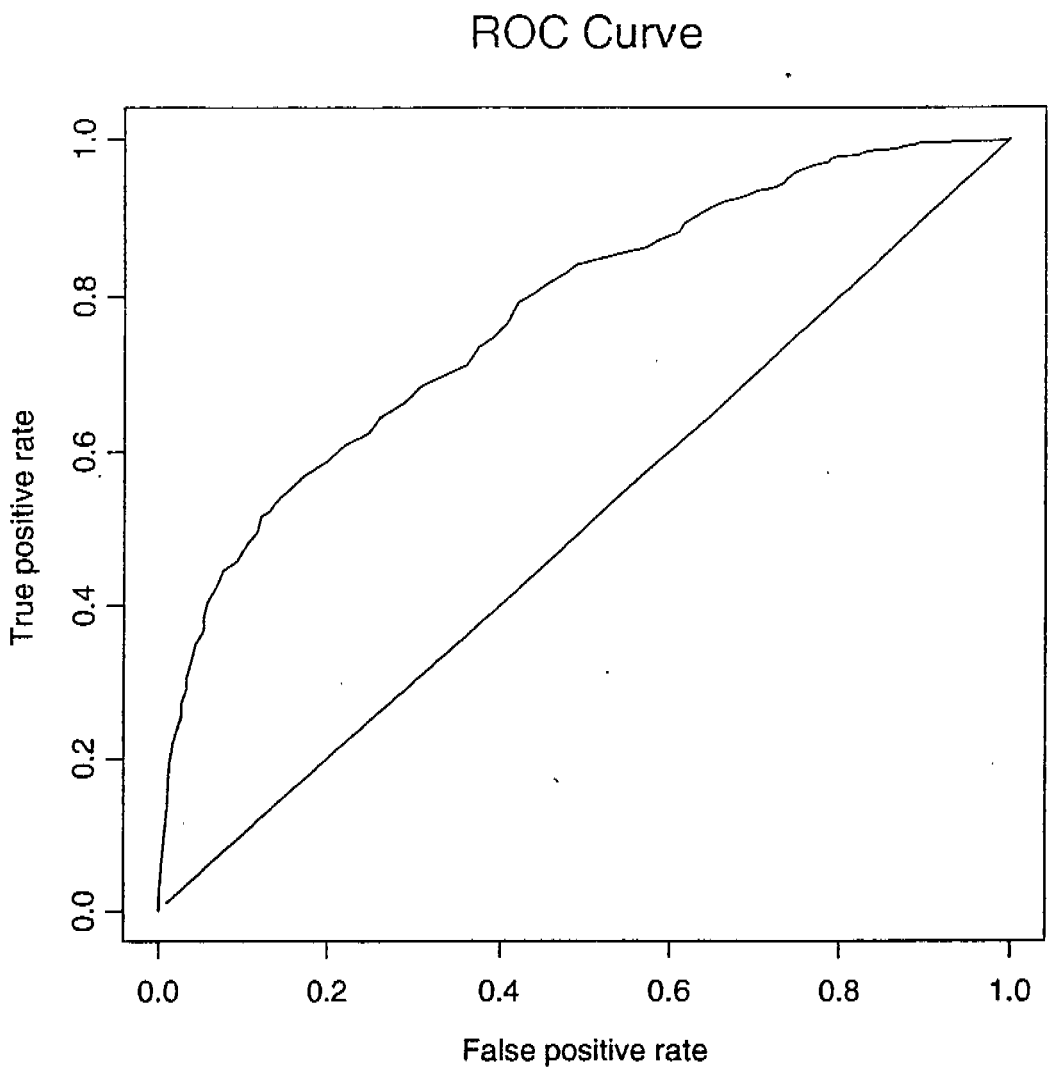


Figure 2.8. Receiver operated characteristic curve used to discriminate between the presence and absence of seedlings and/or sapling trees on the Kaibab National Forest (North Kaibab Ranger District), Arizona. The threshold value that maximized the difference between the true positive and false positive discrimination rates was equal to 0.485.

Table 2.7. Accuracy assessment for the logistic regression model and cross-validation predicting the presence of seedling and/or sapling trees on the Kaibab National Forest (North Kaibab Ranger District), Arizona. Columns represent the field data and the rows represent the model classification. Greater accuracy was achieved predicting seedling/sapling presence than absence.

	Fitted Model		Cross-validation	
	Presence	Absence	Presence	Absence
Presence	985	317	1003	366
Absence	86	168	68	119
Producer's accuracy	92%	35%	94%	24%
User's accuracy	76%	66%	73%	64%
Overall accuracy	74%		72%	
Kappa statistic	0.30		0.22	

When compared with the field data used to generate the original model (Table 2.8), estimates based on the cross-validation of canopy closure yielded similar estimates of average canopy closure on the individual cluster plots. The canopy closure model slightly underestimated the MSW and overestimated MSB, compared to the estimates obtained using the field data. Ninety-five percent confidences interval for canopy closure (Table 8) estimated from the mean variances of the field data (0.627 ± 0.033) and the error estimation associated with the model (0.624 ± 0.034) were similar. However, because we are using a model to estimate canopy closure, more conservative confidence intervals (e.g., Bonferroni simultaneous prediction limits) may be more appropriate (Neter et al. 1985). Conservative intervals for our estimates of canopy closure would therefore become 0.624 ± 0.064.

DISCUSSION AND CONCLUSIONS

Our models of forest structure provide detailed estimates of structural components at a fine-scale (10-m spatial resolution) with relatively high accuracy. These models are superior to traditional forest mapping techniques based on stand data, because the latter are ineffective at predicting structural changes within a given stand (i.e., predictions are limited to the spatial resolution of the stand).

The models presented here described 45% (height of the understory vegetation) to 83% (proportion of pine basal area) of the variability in forest structural components observed on sample plots. The poor ability of the understory vegetation model to describe understory height may be due to inaccessibility of the understory vegetation to the satellite's scanners resulting from the presence of forest canopy. All models provided

Table 2.8. Comparison of sample statistics for estimating average canopy closure on the Kaibab National Forest (North Kaibab Ranger District), Arizona, using field data and the canopy model.

Statistic	Field Data	Canopy Model
Number of plots	174	174
Minimum	0.000	0.000
First quartile	0.523	0.525
Median	0.671	0.680
Mean	0.627	0.624
Third quartile	0.780	0.763
Maximum	0.993	0.966
MSW [1]	0.0071	0.0066
MSB [2]	0.444	0.481
$\hat{V}\left(\hat{\bar{Z}}_e\right)$ [3]	0.00028	0.00031
0.95 bound on the error of estimation	0.033	0.034 (0.064) [4]

[1] MSW – within-cluster mean squared error.

[2] MSB – between-cluster mean squared error.

[3] $\hat{V}\left(\hat{\bar{Z}}_e\right)$ - variance of average canopy closure per sub-plot.

[4] Bonferroni simultaneous prediction limit.

unbiased (p-value > 0.05) or marginally biased estimates of the forest structural components. Notably, the latter bias was not large enough to hinder use of the models for predictive purposes.

Results of 11-fold cross-validation indicated that estimated error variances for the models provided statistically consistent estimates of the true prediction errors associated with the models. These findings suggest that the estimated error variances could be used to assess estimates of uncertainty when the models are applied to non-sample plot locations.

Our ability to calculate estimation uncertainties allows us to develop GIS layers showing the computed estimation errors, as well as place confidence intervals around our estimates, which may be more useful than the estimates themselves because each prediction is associated with a stated level of certainty. We assume that the estimation uncertainties are at, or near, their lower limit because the field and most of the auxiliary data were assumed to be error free. Other sources of error that could have influenced the performance of our models included the sparseness of field plots, errors in vegetation classes (Joy et al. 2001, Chapter 1), structural variations in the forest at spatial scales finer than 10-m, registration errors between remotely sensed and field data, and mapping errors. Because there may be a lower limit (e.g., 15-m) to the spatial accuracy of the Landsat image, some registration errors are likely, where pixels on the resampled image may not have correspond exactly to the 10-m based field data.

The inclusion of abiotic factors such as soils, moisture gradients, climatic and existing forest inventory data could add additional explanatory power to the models for predicting forest structure. Fujisawa (2002) successfully used Landsat TM data,

elevation. and Soil Survey Geographic Data (NRCS 1995) to predict the spatial distribution of white locoweed (*Oxytropis sericea*) in Larimer County, Colorado. Similarly, Kallas (1997:172-175) evaluated the use of climatic, Landsat TM, elevation, and forest inventory data to predict the probability of observing Armilaria root disease on sample plots randomly located on the Black Hills National Forest in South Dakota. In this study, moisture gradients were correlated with elevation and one could have used either variable in the model. However, difficulties associated with using soil, moisture, climate or forest inventory data are that they generally have a coarse spatial resolution and that coverage may not exist (or be current) for the area of interest. Because of these problems, we decided not to use theses types of data in developing our models.

Our models of forest structure can be used to obtain estimates and associated standard error of predictions for any specified geographical region (i.e., nest site, forest stand, management unit, forest) within the KNF using the appropriate formula. This approach to estimating forest structural components is more cost efficient than estimates based on field sampling alone, where plots may or may not be available to provide an estimate, especially if the area of interest is remotely located, small, or irregularly shaped.

All modeling approaches to wildlife habitat are essentially anthropocentric in nature. We model habitat elements at spatial scales *thought* to be important in describing observed patterns of wildlife-habitat relationship (Van Horne 2002). Our initial interest in developing and evaluating models of forest structure to a fine (10-m) spatial resolution on the KNF was to generate a predictive model of the location of nesting goshawks within the landscape (Chapter 3). We therefore modeled forest structure to a scale (small clusters of trees) thought to provide better insight into nest habitat choices made by

goshawks. The spatial detail provided by these models over large, contiguous areas also allows us to identify habitat determinants of goshawk territory quality at various spatial scales, where territory quality is based on a ranking of territories derived from the hawk's long-term demographic performance (Chapter 4). Although the spatial scales at which goshawks select habitat features for nesting and foraging are unknown, we can hypothesize about the scales (fine to large grain) that influence the hawk's behavior and test for correlations between forest structural components and territory quality.

ACKNOWLEDGEMENTS

We thank the many field assistants who contributed to this study: C.M. Erickson, M.A. Gavin, L.J.H. Hunt, A.M. Iñiguez, J.M. Iñiguez, M.A. Kalkhan, D.C. Laing, J.S. Lambert, J.L. Nelson, S.R. Salafsky, J.C. Seyfried, G.S. Stamatellos, R.A. Steffensen, V.L. Thomas, J.D. Wiens, L.E. Williams. R.T. Reynolds provided logistical support and insights into the use of habitat by northern goshawks. P.L. Chapman and R.M. King provided valuable advice on cross-validation techniques. The manuscript was improved with comments from W.J. Gangloff, R.M. King, R.T. Reynolds, M.S. Williams, and 3 anonymous reviewers. This work was financed by Regions 3 (Albuquerque, NM) and the Rocky Mountain Research Station (Fort Collins, CO) of the USDA Forest Service. The North Kaibab Ranger District (Fredonia, AZ) kindly provided housing during the field season.

LITERATURE CITED

Anonymous. 2000. Resource Information System, North Kaibab Ranger District, Fredonia, AZ

Agterberg, F.P. 1984. Trend surface analysis. P. 147-171 in Spatial statistics and models, Gaile, G.L., C.J. Willmott (eds.). Reidel, Dordrecht, The Netherlands.

Avery, T.E., and G.L. Berlin. 1992. Fundamentals of remote sensing and airphoto interpretation. Macmillan Publ. Co., New York. 472 p.

Breiman, L., J. H. Friedman, R.A. Olshen, and C.J. Stone. 1984. Classification and Regression trees. Wadsworth Ind. Group, Belmont, CA. 358 p.

Bunnell, F.L. and D.J. Vales. 1990. Comparison of methods for estimating forest overstory cover: differences among techniques. Can. J. For. Res. 20:101-107.

Burnett, H. 1991. Green island in the sky. Am. For. 97:44-47.

Crocker-Bedford, C. 1990. Goshawk reproduction and forest management. Wildl. Soc. Bull. 18:262-269.

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.

Efron, B. and R.J. Tibshirani. 1993. An introduction to the bootstrap. Chapman & Hall, New York. 436 p.

Egan, J.P. 1975. Signal detection theory and ROC-analysis. Academic Press, New York. 277 p.

ESRI. 1995. ARC/INFO® Software and on-line help manual. Environmental Research Institute, Inc., Redlands, CA.

ESRI. 1998. ArcView® 3.1. Environmental Research Institute, Inc., Redlands, CA.

Friedl, M.A., and C.E. Brodley. 1997. Decision tree classification of land cover from remotely sensed data. Rem. Sens.Environ. 61:399-409.

Fujisawa, H. 2002. Predicting spatial distribution of white locoweed (*Oxytropis sericea*) in Larimer County, Colorado. M.Sc. thesis, Colo. State. Univ., Fort Collins. 61 p.

Guisan, A., and N.E. Zimmermann. 2000. Predictive habitat distribution models in ecology. Ecol. Model. 135:147-186.

Hansen, M., R. Dubayah, and R. Defries. 1996. Classification trees: an alternative to traditional land cover classifiers. J. Remote Sens. 17:1075-1081.

Hayward, G.D., and R.E. Escano. 1989. Goshawk nest-site characteristics in western Montana and northern Idaho. Condor. 91: 476-479.

Hevesi, J.A., Istok, J.D. and Flint, A.L. 1992. Precipitation estimation in mountainous terrain using multivariate geostatistics. Part I: structural analysis. J. Appl. Meteorol. 31:661-676.

Hilden, O. 1965. Habitat selection in birds. Ann. Zool. Fennici 2:53-75.

Isaaks, E.H., and R.M. Srivastava. 1989. An introduction to applied geostatistics. Oxford University Press, New York. 561 p.

Jennings, S.B., N.D. Brown, and D. Sheil. 1999. Assessing forest canopies and understory illumination: canopy closure, canopy cover and other measures. Forestry 72:59-73.

Joy, S.M., R.M. Reich, and R.T. Reynolds. 2001. Modeling small-scale variability in the composition of goshawk habitat on the Kaibab National Forest. Proc. (CD-ROM) Eighth Biennial USDA For. Serv. Remote Sensing Applications Conference, J.D. Greer (ed.), 10-14 April, 2000, Albuquerque, NM. 11 p.

Kallas, M. 1997. Armillaria root disease on the Black Hills National Forest. M.Sc. thesis, Colo. State. Univ., Fort Collins. 172 p.

Kish, L. and M.R. Frankel. 1974. Inference from complex samples (with discussion). J.R. Stat. Soc. B. 36:1-37.

Kravchenko, A. and Bullock, D.G. 1999. A comparative study of interpolation methods for mapping soil properties. Agron. J. 91:393-400.

Lemmon, P.E. 1956. A spherical densiometer for estimating forest overstory density. For. Sci. 2:314-320.

Lemmon, P. E. 1957. A new instrument for measuring forest overstory density. J. For. 55:667-668.

McNab, W.H. 1989. Terrain shape index: quantifying effect of minor landforms on tree height. For. Sci. 35:91-104.

Metzger, K. 1997. Modeling forest stand structure to a ten meter resolution using Landsat TM data. M.Sc. thesis, Colo. State. Univ., Fort Collins. 123 p.

Michaelson, J., D.S. Schmiel, M.A. Friedl, F.W. Davis, and R.O. Dubayah. 1994. Regression tree analysis of satellite and terrain data to guide vegetation sampling and surveys. J. Veget. Sci. 5:673-686.

Moore, K.R., and C.J. Henny. 1983. Nest site characteristics of three co-existing accipiter hawks in northeastern Oregon. Raptor Research. 17: 65-76.

Morrison, M.L., B.G. Marcot, and R.W. Mannan. 1992. Wildlife-habitat relationships: concepts and applications. University of Wisconsin Press, Madison. 364 p.

Neter, J., W. Wasserman, and M.H. Kutner. 1985. Applied linear statistical models. Irwin, Inc., Homewood, IL. 1127 p.

NRCS. 1995. Soils Survey Geographic (SSURGO) Data Base. National Soils Survey Center, Natural Resource Conservation Service.

Reich, R.M., S.M. Joy, and R.T. Reynolds. In review. Predicting the location of northern goshawk nests: modeling the spatial dependency between nest locations and forest structure. Ecol. Model.

Reich, R.M., and R.A. Davis. 1998. On-line spatial library for the S-PLUS© statistical software package. Colo. State Univ., Fort Collins.

Reynolds, R.T., R.T. Graham, M.H. Reiser, R.L. Bassett, P.L. Kennedy, D.A. Boyce, Jr., G. Goodwin, R. Smith, and E.L. Fisher. 1992. Management recommendations for the northern goshawk in the southwestern United States. USDA For. Serv. Gen. Tech. Rep. RM-217. 90 p.

Reynolds, R.T. and S.M. Joy. 1998. Distribution, territory occupancy, dispersal, and demography of northern goshawks on the Kaibab Plateau, Arizona. Final Report, Ariz. Game and Fish Heritage Project No I94045. 76 p.

Reynolds, R.T., S.M. Joy, and D.G. Leslie. 1994. Nest productivity, fidelity, and spacing of northern goshawks in northern Arizona. Studies Avian Biol. 16: 106-113.

Reynolds, R.T., S.M. Joy, J.D. Wiens, S.R. Salafsky, and J.C. Seyfried. In review. Territory and mate fidelity and breeding dispersal in the northern goshawk. Condor.

Reynolds, R.T., E.C. Meslow, and H.M. Wight. 1982. Nesting habitat of coexisting *Accipiter* in Oregon. J. Wildl. Manage. 46: 124-138.

Schloeder, C.A., N.E. Zimmermann, and M.J. Jacobs. 2001. Comparison of methods for interpolating soil properties using limited data. Soil Sci. Soc. Am. J. 65:470-479.

Scheaffer, R.L., Mendenhall, W. and Ott, R.L. 1996. Elementary survey sampling. Duxbury Press, Belmont, CA. 501 p.

Soares A. 1992. Geostatistical estimation of multi-phase structures. Math. Geol. 24:149-160.

Speiser, R., and R. Bosakowski. 1987. Nest site selection by northern goshawks in northern New Jersey and southeast New York. The Condor. 89: 387-394.

Statistical Sciences. 2000. S-PLUS© 4.0 Statistical software package for personal computers. StatSci Division, MathSoft, Inc., Seattle, WA.

Steele, B.M. 2000. Combining multiple classifiers: an application using spatial and remotely sensed information for land cover type mapping. Rem. Sens. Environ. 74:545-556.

Svardson, G. 1949. Competition and habitat selection in birds. Oikos 1:157-174.

Upton, G.J.G. and Fingleton, B. 1985. Spatial data analysis by example. Vol. 1, Point pattern and quantitative data. John Wiley and Sons, New York. 409 p.

Van Horne, B. 2002. Approaches to habitat modeling: the tensions between pattern and process and between specificity and generality. P. 63-72 *in* Predicting species occurrences: issues of accuracy and scale, Scott, J.M. et al. (eds.). Island press, Washington, DC.

White, M.A., and J.T. Vankat. 1993. Middle and high elevation coniferous forest communities of the North Rim region of the Grand Canyon National Park, Arizona, USA. Veget. 109:161-174.

Williams, M.S. 1997. A regression technique accounting for heteroscedastic and asymmetric error. J. Agricult., Biol., and Environ. Stat. 2:108-129.

CHAPTER THREE

PREDICTING THE LOCATION OF NORTHERN GOSHAWK NESTS: MODELING THE SPATIAL DEPENDENCY BETWEEN NEST LOCATIONS AND FOREST STRUCTURE

ABSTRACT

Northern goshawks interact with each other and their environment in a spatially dependent manner. However, predicting the location of active nests (e.g., where eggs are laid) in a given year is difficult due to the secretive nature of the hawks in their forest environment, their annually variable attempts at nesting, and the extent of the area within a home range where they will nest. We used a Gibbsian pairwise potential model to describe the spatial dependency (1) among nest locations and (2) between nest locations and the environment for a large population of goshawks on the Kaibab National Forest's (KNF) North Kaibab Ranger District (NKRD). Nest locations in a given year were regularly distributed at a minimum distance of 1.6 km between active nests; however, as the spatial scale increased (i.e., as distance between the nests increased), the degree of regularity decreased. Important forest predictors for nest locations included canopy closure, total basal area, proportion of basal area in ponderosa pine, spruce, fir, and aspen, maximum height of the understory vegetation, and presence/absence of seedlings and saplings. The probability of an occurrence of an active nest within a 10-m x 10-m

area was modeled using logistic regression. Spatial analysis, using nest spacing and habitat variables, indicated that potential active nest locations were abundant and randomly distributed throughout the NKRD. This supports the supposition that the availability of locations with high potential for nests is not limiting the goshawk population on the study area. Instead, territoriality, and what appear to be non-compressible territories, sets the upper limit to the population. If the forest is managed improperly, however, suitable nest habitat may become rare. Choice of nest location was probably constrained by the availability of high potential locations within spaces defined by neighboring territories. Overall territory density, on the other hand, may reflect the abundance, quality, and accessibility of prey on the study area. This model can be used to evaluate the influence of forest management activities on the goshawk population on the NKRD.

INTRODUCTION

The northern goshawk (*Accipiter gentilis atricapillus*; hereafter goshawk) has been the focus of intensive research for the past decade (Block et al. 1994, Kennedy 1997, Peck 2000) because of suspected population declines due to loss of habitat (Reynolds 1983, 1989, Kenward and Widén 1989, Speiser and Basakowski 1984, Crocker-Bedford 1990, Widén 1997) and changes in forest structure (Reynolds et al. 1992), both resulting from forest management. Many goshawk studies in North America and Europe have focused on the hawks' habitat use, food habits, movements, distribution, demographics, and diets (Block et al. 1994); however, no studies have attempted to use spatially explicit models to describe the spatial dynamics between goshawks and their

environment. Spatially explicit models take into consideration the spatial dependency among objects in the model.

Goshawks interact with conspecifics and their habitat in a spatially dependent manner (Widén 1985, Selås 1997, Reynolds and Joy 1998). By first describing the spatial distribution among active goshawk nests (i.e., nests in which eggs are laid) within a goshawk population and then modeling the interaction between nest locations and forest structure, it may be possible to predict the location of active nests in a given year. Locating active nests is extremely difficult due to the secretive nature of the birds and their annually variable attempts at nesting (Reynolds and Joy 1998), nest concealment, and the extent of area within their home ranges where they will nest.

Spatial statistics have not been used to their fullest potential in animal ecology due to a lack of understanding of these techniques. Although some researchers have used spatial techniques to explore wildlife-habitat relationships (Baker et al.1995, Augustin et al. 1996, Ripple et al. 1997, van Manen and Pelton 1997, Mladenoff et al. 1999, Swindle et al. 1999, Thome et al. 1999, Pearce and Ferrier 2000, Finn et al. 2002), few have recognized their value in exploring manifold spatial dependencies, such as those between species and their environment as well as among species, or among individuals (Reich et al. 2000). The Poisson process has been used widely to describe random spatial patterns, but the model does not take into account the effects of spatial interactions. The Gibbs point process (Diggle et al. 1987, Ripley 1990, Cressie 1991:674-678) is a flexible alternative in which "competitive" intra- or inter-specific interactions are taken into account and described by a pairwise potential function. The Gibbsian pairwise potential model may also be expanded by including environmental variables to identify potential

habitat for a species in a landscape (Reich et al. 1997). In this paper, we fit a Gibbsian pairwise potential model to describe the spatial variability among goshawk nests and their association with forest structure on the NKRD in northern Arizona. We also identify habitat that is more likely to have nests by correlating the location of known nests with environmental variables that account for the coarse-scale variability across the landscape.

STUDY AREA

The study area includes NKRD lands above 2,182 m elevation above sea level (asl). This land comprises the northern two-thirds of the Kaibab Plateau in northern Arizona. The NKRD is bound by the Grand Canyon National Park to the south, steep slopes to the east, and gentle slopes to the north and west that descend a shrub-steppe plain. This 1,285 km^2 area comprises nearly 100% of the tall coniferous forests [excludes most of the low elevation pinyon (*Pinus edulis*)-juniper (*Juniperus* spp.) woodlands] on the NKRD. Six vegetation classes (Fig. 3.1) dominate the study area (Joy et al. 2001, Chapter 1): 1) pinyon-juniper woodlands (106 km^2, 8%) occur at lower elevations (2,182-2,250 m asl) and mix with ponderosa pine (*P. ponderosa*) at transitional zones; 2) ponderosa pine comprises 55% (704 km^2) of the forested area and occurs between 2,250 and 2,550 m asl; 3) mixed-conifer, comprised of ponderosa pine, white fir (*Abies concolor*), Douglas-fir (*Pseudotsuga mensiesii*), and quaking aspen (*Populus tremuloides*) (145 km^2, 11%), occurs between 2,550 and 2,650 m elevation asl; 4) spruce (*Picea pungens*, *P. englemannii*)-dominated mixes (130 km^2, 10%), primarily with subalpine fir (*A. lasiocarpa*), occurs above 2,650 m elevation asl; 5) deciduous (quaking aspen, Gamble's oak [*Quercus* gambeli])-dominated mixes (112 km^2, 9%) occur

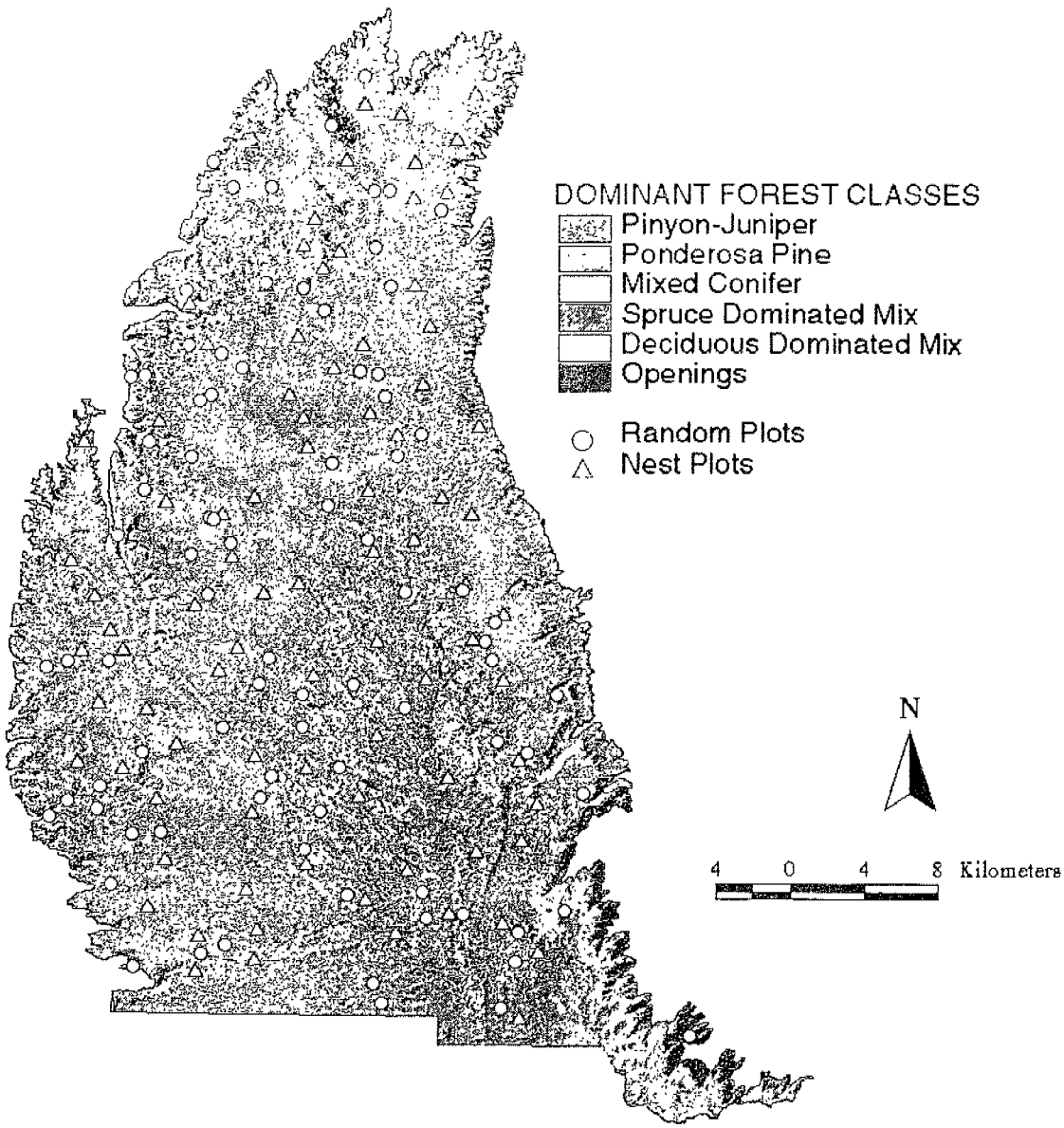


Figure 3.1. Distribution and arrangement of nest plots (triangles) and random plots (circles) used to model forest structure displayed among dominant vegetation classes on the North Kaibab Ranger District, Kaibab National Forest, Arizona.

throughout the forest and are common where extensive disturbance has occurred (Fig. 3.1); and 6) openings (90 km^2, 7%) that contain grasses and herbaceous vegetation include a series of long, narrow meadows and various smaller gaps in the canopy which are scattered throughout the forest.

Nearly all of the KNF has been altered by some form of management during the past 100 years (Pearson 1950, Burnett 1991). By the early-1900s livestock grazing was common and fire suppression had been established. A long-term policy of fire exclusion has resulted in large numbers of shade-tolerant seedlings and saplings throughout the forest creating fuels and closing-in of the historically more-open understory (Weaver 1951). Organized tree harvests in the form of sanitation cuts and single-tree selection began in the 1920s. These harvest regimes continued, along with occasional, small (0.1 km^2) clearcuts in the mixed-conifer, until the late-1970s. Intensive forest management at the stand level (shelterwood, seed, salvage, removal, and thinning cuts) began in the 1980s and continued until 1991, when the NKRD implemented forest management prescriptions designed to enhance the habitat of goshawks and their prey (Reynolds et al. 1992).

The NKRD receives about 67.5 cm of precipitation annually, with winter snowpacks of 2.5-3.0 m (White and Vankat 1993). A drought period typically occurs in May and June, followed by mid- to late-summer thunderstorms and heavy showers.

METHODS

The Data

The data layers used to model spatial dependencies among goshawks and their environment included the location of active nests (i.e., nests in which eggs were laid), Landsat Thematic Mapper (TM) imagery, field measurement, and GIS-derived topographic variables. Nest locations were used to describe the spatial distribution of nests; whereas, the field measurements, Landsat imagery, and topographic variables were used to model forest composition and structure to a 10-m spatial resolution.

Goshawk Nest Locations

Searches for active goshawk nests began in 1991 and continued through 1998. Nests were found by (1) searching on foot (Reynolds 1982), (2) systematically broadcasting goshawk vocalizations from predetermined stations and transects (Kennedy and Stahlecker 1993, Joy et al. 1994), and (3) visiting active nests found in prior years of the study. When the status of a previously-active nest remained unknown, searches of 16 and 24 km^2 areas around that nest were carried out on foot or by broadcasting, respectively, to locate an alternate active nest within the territory. Goshawks may use more than one nest within their territories among breeding years (Reynolds and Wight 1978, Reynolds et al. 1994, Reynolds and Joy 1998). Nest searches began in April and ended after the post-fledging period. Each year, the overall search area on the NKRD was expanded to include more territories. A "territory" (approx. 1.5 km radius) is the area used and defended by a single pair of goshawks during the nesting season and may contain one or more alternate nest trees (Reynolds et al. 1994). When eggs were laid in a nest, the nest (and territory) was considered "active." At each active nest, adults and

juveniles were captured and banded with a USDI Fish and Wildlife Service aluminum leg band and an anodized aluminum colored leg band, the latter marked with unique 2-character alpha-numeric codes readable at up to 50-80 m with 20-40 power spotting scopes. Identifying the individual goshawks allowed us to correctly associate new nests with individual territories. On the study area, territoriality is maintained even in non-breeding years by marked individuals who continue to defend their territory (Reynolds et al. 1994).

Field Data

Models of forest structure were based on the spatial interpolation (see below) of habitat attributes at both active nests and randomly selected plots (Fig. 3.1):

Goshawk nest plots: We measured the forest vegetation immediately surrounding the nest tree at one nest within each of 92 goshawk territories being studied through 1998. In territories containing multiple alternate nests that had been active since 1991, we randomly selected one alternate at which to measure the forest characteristics. At single-nest territories, we measured the vegetation at that nest tree.

Randomly located plots: To describe all the spatial/structural variability on the NKRD, we located 85 random plots throughout the study area. We placed no constraints on the location of random plots (i.e., they were placed irrespective of territories and nests), because we considered all habitat to be potentially available to goshawks for nest site use.

GIS and Landsat TM Data

The GIS database consisted of four topographic variables (elevation, slope, aspect, and landform), 6 bands (1-5, and 7) of Landsat TM data (1997; 22 June; Path 37,

Row 35), and seven variables representing stand structure (percent canopy closure; total basal area; proportions of (a) ponderosa pine, (b) spruce/fir, and (c) aspen in the total basal area; maximum height of the understory vegetation; and the presence of seedlings or saplings). All habitat-related variables were believed to be important to goshawk nest tree selection. Elevation was obtained from USGS digital elevation models (DEM) and used to derive aspect and slope. The DEM was also used to calculate a landform index (McNab 1989), which expresses surface shape as a measure of surface concavity or convexity (computed as the mean slope gradient from the original cell to adjacent cells in 4 directions), a continuous variable. Grid coverages for elevation, slope, aspect, and landform were resampled to 10 m, corresponding to the spatial resolution of the field data (below). Grid coverages representing forest structure were developed by spatially interpolating the random and nest-based field data to a 10-m spatial resolution using trend surface models and regression trees (Chapter 2). Landsat Bands 1-5 and 7, and topographic data were used as predictor variables. All grid manipulations were performed in ArcView® (ESRI 1998).

Field Measurements

Because the spatial variability in forest structure can vary at scales smaller than those determined by the spatial resolution of Landsat TM imagery (i.e., < 30 m), we designed our field sampling to classify forest structure to a 10-m spatial resolution. Sample plots consisted of a cluster of nine 10-m x 10-m subplots that corresponded to a 30-m x 30-m pixel on our Landsat TM imagery, the location of which was verified using a Trimble Navigation Pathfinder™ Asset Surveyor Global Positioning System (estimated accuracy = 1 - 3 m). Field measurements were collected during August and September of

1997. Each plot was established in a north-south, east-west fashion with the coordinate systematically assigned to either the center (nest tree plot) or lower left corner (random plots) of the plot. Vegetative characteristics were recorded on each of the nine 10-m x 10-m subplots and included canopy closure [measured with a concave, spherical densiometer (Lemmon 1956, 1957)], overstory species, total basal area by species (measured with a 20 factor prism), height of the understory vegetation, and the presence of seedlings and saplings.

Spatial Distribution of Active Goshawk Nests

To model the distribution of active goshawk nests, we selected a large (528 km^2) rectangular region within the NKRD. A rectangular region was selected to simplify the algorithm required to adjust for edge effects, while the shape of the rectangular was selected to include as many nests as possible. The spatial location of all active nests in 1998 within the rectangular region B (Fig. 3.2) was assumed to represent the spatial relationship between active goshawk nests and forest structure when the population is at or near full occupancy. In 1998 active nests attained the most continuous spacing (i.e., fewest gaps due to non-nesting territorial pairs or individuals) among all the breeding years studied (Fig. 3.3).

Using the spatial location of each nest in the rectangular region B, a Monte Carlo test (Besag and Diggle 1977) based on the Cramér-von Mises type statistic (Cressie 1991:642)

$$k = \int_0^H \left\{ \hat{K}(h)^{1/2} - \pi^{1/2} h \right\}^2 dh \qquad [3.1]$$

was used to test the null hypothesis of complete spatial randomness (csr); i.e., whether the arrangement of nests within a circular region of radius R, does not differ significantly

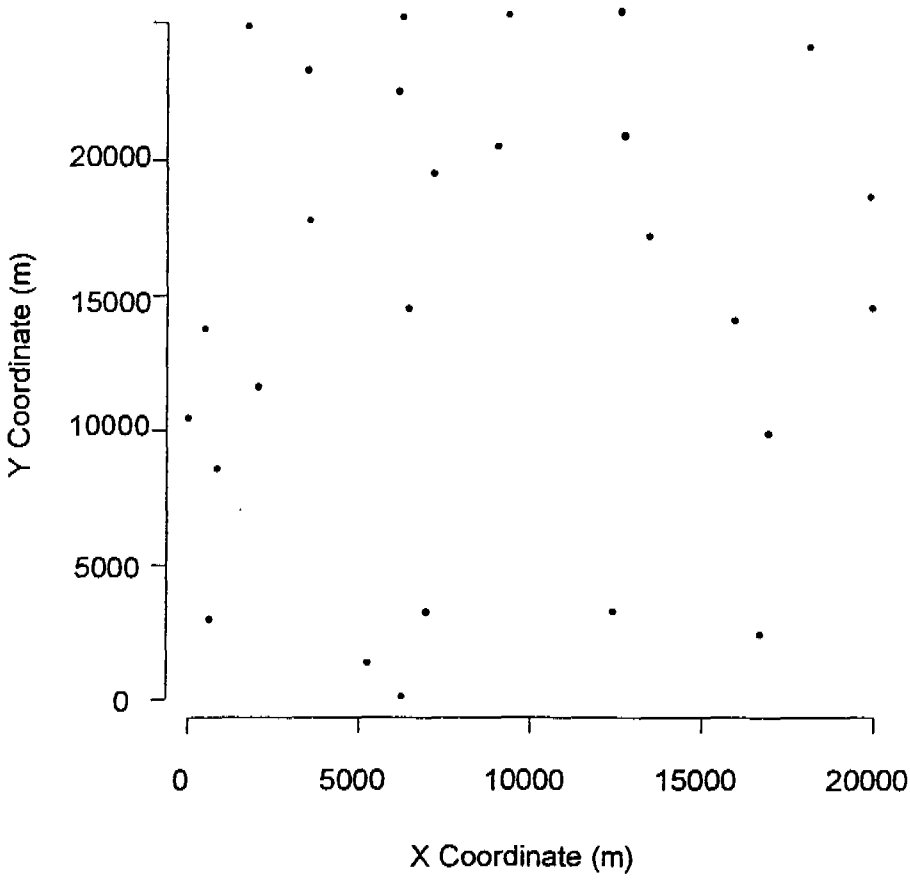


Figure 3.2. Bounded region (B) showing the relative location of 27 active northern goshawk nests from 1998 used to model the spatial relationship between active nests and forest structure.

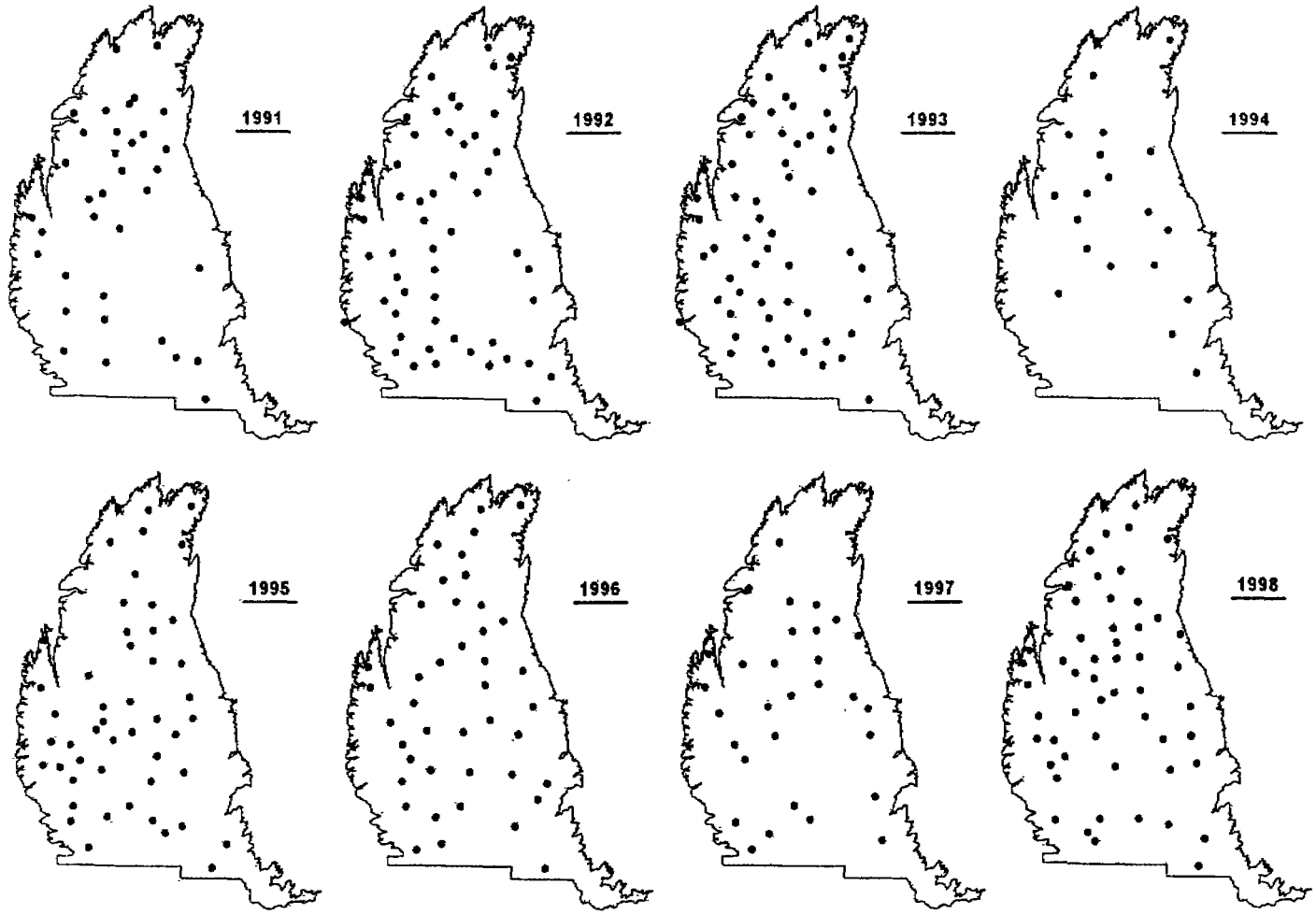


Figure 3.3. The location of active northern goshawk nests between 1991 and 1998 on the North Kaibab Ranger District, Kaibab National Forest, Arizona.

from that expected under the assumption of csr. This was done at 14 spatial scales ranging from 2 km to 16 km in increments of 1 km by simulating values of the test statistic under csr and comparing them to the corresponding statistic calculated from the observed pattern of active goshawk nests. For each simulation, we calculated the empirical K-function $\hat{K}_i(h)$ (Ripley 1977), corrected for edge effect (Cressie 1991:616), and the Cramér-von Mises statistic k. The significance (p-value) of the test was calculated as $\hat{p} = (R + 1 - r)/R$, where R is the number of simulations, and r is the rank of the test statistic associated with the observed point pattern. A small p-value supports the alternative hypothesis of a non-random spatial pattern. All tests were based on 200 realizations of a spatial Poisson process to allow for the calculation of a p-value to the nearest one percent.

Traditional nearest neighbor statistics, which are often used to test nest spacing (e.g., Newton et al. 1977), assume that the nearest neighbors are independent (Cressie 1991:603-606). If applied to mapped data sets such as nests, however, the nearest neighbor measurements are not independent, and one would tend to reject the null hypothesis of csr too often (Cressie 1991:610). In contrast, the K-function and the Cramer von-Mises goodness-of-fit test do not assume that distance measurements are independent. Furthermore, they use information on many spatial scales because they are based on squared distances to the first, second, third,..., kth nearest neighbors.

Gibbsian Pairwise Potential Model

The Gibbsian pairwise potential model is a Markov point process, a flexible class of models in that they simulate both regular (inhibition) and aggregated (contagious) patterns. The primary use of such models has been in the study of regular point patterns

such as those exhibited by the northern goshawk (Reynolds and Joy 1998, Widén 1985), other accipiters (Newton et al. 1977), as well as other raptors (Cade 1960, Ratcliffe 1962, Newton 1979).

The most extreme form of regularity is the direct exclusion from a given area, whether by complete occupancy, allelopathy, or territoriality. Models that describe such phenomenon are termed hard-core models. Every individual in the population has a circular neighborhood or radius 2R within which no other individual can exist. For biological populations that display plasticity of size and shape, the hard-core model may be too extreme. As an alternative, a soft—core model with fixed-range interactions may be used that are less extreme, in that within a given neighborhood of radius R inhibition is not complete, but a competitive effect (i.e., territoriality) is experienced. The degree of territoriality may or may not be a function of the distance between individual pairs (h).

Potential Energy of Goshawk Nests

The location of all N goshawk nests within the bounded region B were assigned coordinates $X=\{X_i=(x_i, y_i) \in B, i = 1, 2, \ldots, N\}$. To model the spatial distribution and association of individual territorial goshawk pairs (i.e., nests), we assumed that the territorial influence between pairs depended on the relative, and not the absolute position of nests. This assumption implies a homogeneous environment. The territorial interaction, or potential energy, Ψ, can be modeled as a function of the Euclidean distance $h_{ij} = \|X_i - X_j\|$ between pairs of nests in which the territorial influence between individual pairs decreases with increasing distance. Thus, the total potential energy for the point process is defined as (Cressie 1991:677):

$$U_N(X) = \sum_{i<j}^{N} \Psi(h_{ij}) . \qquad [3.2]$$

where $U_N(X)$ can be thought of as the total energy required to add a nest to the point pattern. The observed point pattern of goshawk nests can, therefore, be regarded as being distributed according to a Gibbs canonical distribution:

$$f(x) = \exp[-U_N(X)] / Z(\Psi; N), \qquad [3.3]$$

where $Z(\Psi;N)$ is a normalizing constant where the joint probability density integrates to 1. If the normalizing constant exists, the point pattern is said to be stable. A strictly positive pairwise potential (i.e., inhibition process) always yields a stable process, while those with negative potential energy at some specified distances (i.e., contagious process) are generally unstable (Cressie 1991:678). The sign and shape of the potential functions are determined by whether there is inhibition or attraction between nests. Positive values indicate inhibition, while negative values represents attraction. If there are no interactions between nests, the value of the potential function is zero.

Model Parameter Estimation

Consider a family of parameterized pairwise potential functions $\{\Psi_\theta(h);\ \theta \in \Theta\}$. Given a finite set of points in the bounded region B, the likelihood of the potential function $\Psi_\theta(h)$ is given by the Gibbs canonical distribution (Eq. 3.3). The maximum likelihood estimate of θ is obtained by finding a $\hat{\theta}$ that maximizes Eq. 3.3. Maximization requires computing the normalizing constant $Z(\psi; N)$, which is not usually available in closed form (i.e., where an explicit solution exists). Ogata and Tanemura (1981) use the cluster-expansion method of statistical mechanics (Ogata and Tanemura 1981, Cressie 1991:682) to obtain an approximation of the normalizing constant, conditioned on the number of points in B:

$$Z(\psi; N) = |B|^N \left(1 - a(\theta)/|B|\right)^{N(N-1)/2} \qquad [3.4]$$

where

$$a(\theta) = 2\pi \int_0^\infty h[1 - \exp\{-\psi_\theta(h)\}]\,dh\,, \qquad [3.5]$$

is the second cluster integral, and |B| is the area of the bounded region B. In their approximation, only pairwise interactions were considered; higher order interactions were assumed to be negligible. Cressie (1991:683) points out that this approximation holds only for stable pair-potentials, and may not be valid for unstable pair-potentials that require higher-order interactions such as a Markov cluster process. Combining Eq. 3.3 and 3.4 leads to the approximate log likelihood function:

$$\log L(\theta \mid X) = \sum_{i<j}^{N} \Psi_\theta(\| X_i - X_j \|) - \frac{1}{2} N(N-1) \log\left(1 - \frac{a(\theta)}{|B|}\right), \qquad [3.6]$$

which can be solved using nonlinear optimization procedures.

To use this relationship in describing the spatial distribution and association of individual nests, one must be able to mathematically describe the interaction potentials of a spatial point pattern. Three parameterized potential functions proposed by Ogata and Tanemura (1981 and 1985) are available to describe the interactions observed in the distribution of the goshawk nests:

$$\text{PF1}: \ \psi_\theta(h) = -\log[1 + (\alpha h - 1) e^{-\beta h^2}] \quad \theta = (\alpha, \beta), \alpha \geq 0, \beta > 0 \qquad [3.7]$$

$$\text{PF2}: \ \psi_\theta(h) = -\log[1 + (\alpha - 1) e^{-\beta h^2}] \quad \theta = (\alpha, \beta), \alpha \geq 0, \beta > 0 \qquad [3.8]$$

$$\text{PF3}: \ \psi_\theta(h) = \beta(\sigma/h)^{12} - \alpha(\sigma/h)^6 \quad \theta = (\alpha, \beta, \sigma), \beta > 0\,. \qquad [3.9]$$

All three potential functions can model both repulsive and attractive forces. The parameter, α, controls the type of force between a pair of points, while β and σ are scaling parameters. The potential function PF1 represents a purely repulsive potential when $\alpha = 0$, and has both repulsive and attractive potentials when $\alpha > 0$. The potential

function PF2 is repulsive when $0 \leq \alpha < 1$, independent when $\alpha = 1$, and attractive when $\alpha > 1$. The potential for PF3 is purely repulsive when $\alpha < 0$, and attractive when $\alpha > 0$. The second cluster integral, $a(\theta)$, for the three potential functions are given by:

$$\text{PF1}: a(\alpha, \beta) = (\pi/\beta)(1 - \alpha\sqrt{\pi/\beta}/2) \qquad [3.10]$$

$$\text{PF2}: a(\alpha, \beta) = \pi(1 - \alpha)/\beta \qquad [3.11]$$

$$\text{PF3}: a(\alpha, \beta, \sigma) = -\frac{\pi}{6}\beta^{1/6}\sigma^2 \sum_{k=0}^{\infty} \frac{1}{k!}\Gamma\left(\frac{6k-2}{12}\right)\alpha^k \beta^{-k/2}. \qquad [3.12]$$

The pairwise potential models PF1-PF3 were fit to the point data of the individual nests using a nonlinear least squares procedure to obtain an estimate of the parameter vector $\theta = (\alpha, \beta)$ or $\theta = (\alpha, \beta, \sigma)$ that maximized the approximate log likelihood (Eq. 3.6). Akaike's (1977) AIC, was used to select the best model among the three possible models (PF1-PF3).

Potential Energy Between Nests and Forest Structure

To include environmental heterogeneity in the model, the total potential energy was redefined as follows:

$$U_N(X) = \sum_{i<j}^{N} \Psi(h_{ij}) + \sum_{i=1}^{N} \phi(z_i), \qquad [3.13]$$

where $\phi(z_i)$ is a measure of the interaction of individual nests with the environment (i.e., forest structure). If we assume that the presence, or absence, of an individual nest is correlated to a set of known environmental variables we can, for example, define the probability of observing a goshawk nest at a given location as π. The potential energy associated with this location can be expressed as (Reich et al. 1997):

$$\phi(z) = \frac{1}{\pi} - 1 = f(\text{environmental variables}). \qquad [3.14]$$

Large positive values indicate "poor" nest locations while small values indicate "good" nest locations. We define a "good" area as one that has a higher probability of containing an active nest (as in, the probability of finding an active nest is better) than other areas. Good areas, however, do not necessarily confer greater fitness on the birds using those sites (Van Horne 1983, Vickery et al. 1992). Rather, reproductive success and other measures of fitness are influenced by habitat characteristics, food, and life history strategies used throughout the home range (Newton et al. 1977, Reynolds et al. 1992, Kostrzewa 1996). Furthermore, the presence of good habitat alone does not guarantee that a nest will be present because the value of an area as a nest location is dependent upon the arrangement of both the fine- and coarse-scale (i.e., landscape scale) variability in the landscape (Ricklefs 1987), territoriality, and population density.

Modeling Nest Site Suitability

To model the potential energy associated with forest structure we used a multiple logistic regression model (Hosmer and Lemeshow 1989, Manly et al. 1993):

$$\pi = \frac{e^{\beta_0+\beta_1 z_1+\cdots+\beta_k z_k}}{1+e^{\beta_0+\beta_1 z_1+\cdots+\beta_k z_k}}, \qquad [3.15]$$

where π is the probability of observing a goshawk nest, $z_1, \ldots, z_k$ are independent predictor variables, and $\beta_1, \ldots, \beta_k$ are logistic coefficients. Independent variables considered in the model included topographic data (elevation, slope, aspect, landform) and forest structure (total basal area, proportion of pine, aspen, spruce-fir basal area, height of understory vegetation, and presence of seedlings). The final form of the model was based on a forward selection process that eliminated independent variables with high *p*-values. Coefficients from the logistic regression model indicate the direction of change (positive - increase, negative - decrease) required by an independent variable to maximize

the probability of an occurrence of an active nest, given the topographic and environmental constraints imposed by other independent variables.

Preliminary analysis indicated that the functional form of the logistic regression model differed among vegetation classes in that not all of the independent variables were important in all vegetation classes. To account for these differences, we added dummy variables to the model. After fitting the logistic regression, a final model, composed of significant variables and coefficients, was used to create a map of the probability distribution of nest locations. We standardized (Neter et al. 1985:262) the regression coefficients for the logistic model to compare the relative strength of individual variables within each model, as well as across vegetation classes.

We used classification error rates to evaluate the fit of the model. To calculate classification rates, we compared the probability from the logistic regression models, a continuous variable, to a cutoff value. Each 10-m x 10-m pixel of the NKRD was categorized into a dichotomous variable with a value of 1 or 0, representing good and poor nest locations, respectively. To determine the optimal cutoff value, we compared model results to those that would be obtained from a random process. The optimal cutoff value was selected by maximizing the improvement of model predictions over a null model of random habitat selection [i.e., maximizing the difference between the proportion of nest pixels correctly classified and the proportion of the NKRD classified as good nest habitat (Pierera and Itami 1991, Ozesmi and Mitsch 1997)]. This process considered the trade off between maximizing the correct classification of good nest habitat by selecting a lower cutoff value, and minimizing the area classified as good habitat by selecting a higher cutoff value.

Leave-one-out cross-validation (Efron and Tibshirani 1993; 240) was used to generate the mean cutoff value and its associated standard deviation. This mean optimal cutoff value was used to generate a grid surface showing the location of good and poor nest locations. All grid cell values over the optimal cutoff value were assigned a value of 1, while cell values less than the optimal cutoff were assigned a value of 0. The logistic regression model was also used to generate a grid surface of potential energy associated with forest structure (Eq. 3.14).

Simulating the Spatial Distribution of Goshawk Nests

To simulate a point pattern of goshawk nests in a given year, the point process was conditioned on N, the total number of nests observed in the bounded region, B. Using an algorithm proposed by Ogata and Tanemura (1989) the following steps were used to simulate the two components (spatial interactions among nests and forest structure) of the spatial distribution of goshawk nests:

<u>*Step 1*</u>. Randomly locate the first nest ($t = 1$) within the bounded region B. If forest structure is taken into consideration, the location ($X_t = \{x_t, y_t \in B; t = 1\}$) of the first nest is selected proportional to $\exp[-U_1(X)]$, where $U_1(X)$ is the potential energy associated with forest structure (Eq. 3.14). The nest site is selected with probability proportional to the suitability of the site, which is based on the logistic regression model. A low potential energy would indicate a good site, while a high potential energy would indicate a poor site for a nest. If forest structure in not considered in the location of nest sites, the location of the nest is chosen from a uniform distribution on the bounded region B.

Step 2. For the second and successive steps (t, $t = 2, 3, \ldots N$), two additional locations are chosen: $X_t' = \{x_t' \, y_t' \in B; t = 2, \ldots, N\}$ and $X_t^* = \{x_t^*, \, y_t^* \in B; t = 2, \ldots, N\}$ using the procedures outlined in *Step 1*.

Step 3A. If the spatial interaction between nests is not being considered, the total potential energies, $U_t'(X)$ and $U_t^*(X)$, associated with the two locations obtained in *Step 2* are computed (Eq. 3.14) and compared. The location, X_t' or X_t^*, that minimizes the total potential energy is selected as the new location to add to the point pattern.

Step 3B. If the spatial interaction between nests is taken into consideration, the total potential energies, $U_t'(X)$ and $U_t^*(X)$, associated with the two locations obtained in *Step 2* are computed using Eq. 3.13. If $\min\{U_t'(X), U_t^*(X)\} < U_{t-1}(X)$, the new location, X_{t+1} is taken as $\min\{U_t'(X), U_t^*(X)\}$. If $\min\{U_t'(X), U_t^*(X)\} \geq U_{n-1}(t)$, a uniform random number, ξ, on the interval (0,1) is computed. If ξ is less than $\exp[U_{t-1}(X) - \min\{U_t'(X), U_t^*(X)\}]$, location X_{t+1} is taken to be $\min\{U_t'(X), U_t^*(X)\}$. Otherwise, no new nest is added to the point pattern in this step.

Steps 2 and 3 are repeated until all N nests have been located within the bounds of the population.

Step 4. The last step in the simulation was to apply the Metropolis algorithm (Cressie 1991:679, Ogata and Tanemura 1989) to adjust the initial point pattern to a state of equilibrium. This is accomplished by randomly selecting one of the N simulated nest

locations $X_t' = \{x_t', y_t' \in B; t = 1, ..., N\}$. Next, a new location is randomly selected in such a way that the coordinates $\{x_t^*, yt^*\}$ lie in a square with vertices at the point $x_t' \pm \delta$ and $y_t' \pm \delta$, while all other N-1 nests have the same position. The total potential energies associated with the two point patterns are computed and compared using the procedures described in *Step* 3B. If the total potential energy for the point pattern with one of the nest moved slightly is less than the potential energy for the original point pattern, the nest is moved to this new location. This process is repeated until the point pattern converges to a state of equilibrium. To ensure this convergence, δ, the maximum single step displacement allowed in passing from one state to the next, was selected so as to reject one-half of the trial states (Cressie 1991:680). Other than this recommendation, no information is available in the literature on how many steps are required for convergence (Cressie 1991:680). In simulating the spatial distribution of the nests we used 78 x 200 Monte Carlo steps and a δ = 30 m.

Ogata and Tanemura (1985) suggest one way to evaluate the equilibrium assumption is to examine the stationarity of the time series (t) of the total potential energy of the simulated point pattern. If we graph the change in total potential energy as a function of time, one would expect the sample mean of the time series to equal zero (Ogata and Tanemura 1985). If a significant bias exists, this would indicate the point process is non-stationary and alternative models should be considered.

The goodness-of-fit of the point process model was assessed by comparing the transformed empirical K-function ($\hat{L}(h) = \left\{\hat{K}(h)/\pi\right\}^{1/2}$) (Ripley 1977), corrected for edge effect (Cressie 1991:615-618), to the transformed K-functions from 200 simulated realizations of the model. The simulations were used in constructing confidence

envelopes based on the minimum and maximum transformed K-function to test the null hypothesis of no significant differences at the $\alpha = 0.005$ level. If, for any distance, the observed transformed K-function falls above or below the confidence envelopes the null hypothesis is rejected at the appropriate level of significance.

The first set of evaluations was based on the point process model describing the spatial interaction between individual northern goshawk nests. Next, we evaluated the component describing the spatial relationship between individual nests and forest structure. Finally, we combined the two components together to simulate the spatial distribution of goshawk nests based on the spatial interaction between individual nests and forest structure. To assess the degree of agreement between the distribution of predicted nest points and that of active nests, I used a chi-square goodness-of-fit test to assess differences in the probabilities of locating a nest between the predicted points and active nests in 1998.

RESULTS

Modeling Nest Site Suitability

The mean optimum probability cut off from the logistic regression used to distinguish good from poor nest locations was 48% $\pm$1.5% (95% confidence intervals; SD=0.008). Based on this threshold, approximately one-third (410 km^2, 33%) of the NKRD was classified as good nest habitat (Fig. 3.4). None of the pinyon-juniper sites were classified as good (Table 3.1), while 38% (267 km^2) and 35% (45 km^2) of pure ponderosa pine and spruce-dominated sites were classified as good, respectively. Only 24% (35 km^2) of mixed-conifer sites were classified as good nest locations; whereas,

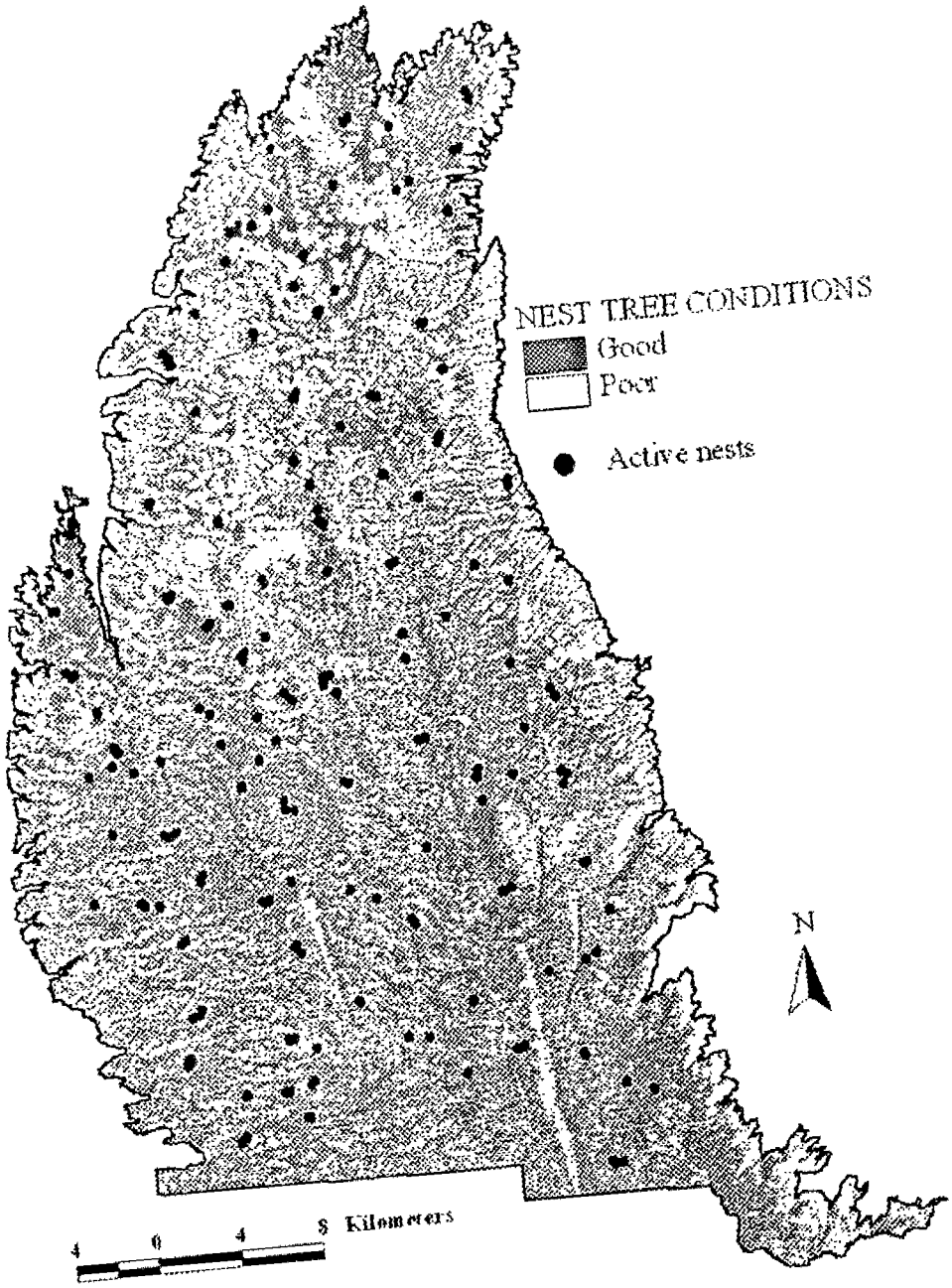


Figure 3.4. Spatial distribution of estimated "good" and "poor" locations for northern goshawk nests on the North Kaibab Ranger District, Kaibab National Forest, Arizona, and all nests active between 1991 and 1998.

Table 3.1. Distribution of estimated good and poor northern goshawk nest habitat by vegetative class on the North Kaibab Ranger District, Kaibab National Forest, Arizona.

Vegetation Class	Good (%)	Poor (%)
Pinyon-Juniper	0	100
Ponderosa Pine	38	62
Mixed-conifer	24	76
Spruce-Dominated Mix	35	65
Deciduous-Dominated Mix	48	52
Opening	14	86
All Vegetation Classes	33	66

48% (54 km^2) of deciduous sites provided good nest locations. Open areas obviously do not contain trees for nests; however, in our model 14% (13 km^2) of openings (Table 3.1) were classified as good nest locations.

Trends in Nest Habitat Use

Between 1991 and 1998, the number of active nests on the study area ranged from a low of 19 (1994) to a high of 55 (1993), representing 204 unique nest locations (out of 344 nest attempts) on 94 unique territories (Table 3.2). The majority (147; 72%) of nest locations, representing 51 territories, were in good sites, while 57 nests (28%), representing 43 territories, were in poor sites (Table 3.3; Fig. 3.4). The largest proportion (79%) of nests in good habitat was in the ponderosa pine class. The fewest ($\leq$ 5%) nests in good habitat were found in deciduous-dominated and mixed-conifer forests. Of the 57 nests in poor habitat, over half (54%) were also in ponderosa pine, while almost a third (30%) were in the mixed-conifer class. Regardless of vegetation class, however, nearly 80% (45 of 57) of nests in poor sites were found within 10 m of a good site.

Nest Habitat

Important variables (Tables 3.4, 3.5) from the logistic regression model and their standardized coefficients, which discriminated between good and poor nest site locations, varied with vegetation class. In ponderosa pine, the likelihood that a stand contained a nest improved with increasing total basal area (above 29 m^2/ha), but smaller proportions of spruce-fir basal area (less than 5.5%) and, especially, aspen basal area (less than 7.9%). Denser canopy closures, flatter slopes, and understory vegetation taller than 0.5 m also improved the probability of a nest location. In the mixed-conifer zone, the likelihood of nest habitat use was greater on steeper (> 8%) slopes with easterly exposure

Table 3.2. Total number of territories and active northern goshawk nests between 1991 and 1998 above 2,182 m in elevation on the North Kaibab Ranger District, Kaibab National Forest, Arizona.

	Year								
	1991	1992	1993	1994	1995	1996	1997	1998	Total
Total territories monitored	36	58	72	87	95	102	105	105	660
New territories with active nests	36	21	13	3	10	8	0	4	94
Active nests	35	52	55	19	48	42	30	53	344
New active alternate nests	35	37	33	9	29	21	15	25	204

Table 3.3. Number of active nests between 1991 and 1998 by estimated suitability (good, poor) of nest locations and vegetative class on the North Kaibab Ranger District, Kaibab National Forest, Arizona.

Vegetation class	Good		Poor		Total Number of Nests
	Number of Nests	Percent	Number of Nests	Percent	
Pinyon-Juniper	0	0	0	0	0
Ponderosa Pine	116	79	31	54	147
Mixed-conifer	8	5	17	30	25
Spruce-Dominated Mix	17	12	5	9	22
Deciduous-Dominated Mix	6	4	4	7	10
Openings	0	0	0	0	0
Total	147	100	57	100	204

Table 3.4. Standardized regression coefficients for variables that maximize the likelihood of a northern goshawk nest occurring in a vegetative class on the North Kaibab Ranger District, Kaibab National Forest, Arizona. The magnitude and direction of the coefficients (positive - increase, negative - decrease) are comparable within and between models.

Variable	Vegetation class					
	Pinyon-Juniper	Ponderosa Pine	Mixed-conifer	Spruce-Dominated Mix	Deciduous-Dominated Mix	Openings
Aspect	----	----	-0.082	-0.172	0.570	----
Slope (%)	-0.373	-0.044	0.041	0.007	0.653	-0.001
Elevation (m)	----	----	-0.016	0.077	----	-0.052
Landform	----	----	0.067	-0.083	-0.324	----
Total BA[1] (m^2/ha)	0.050	0.102	0.116	-0.040	0.112	0.032
Ponderosa Pine BA[2]	----	----	----	0.689	----	----
Spruce-fir BA[2]	-0.051	-0.042	-0.096	----	0.639	-0.067
Aspen BA[2]	0.000	-0.109	0.076	----	-0.236	0.020
Canopy[3]	-0.004	0.003	0.002	0.002	-0.002	-0.001

Table 3.4. (continued)

Understory Height (m)	0.053	0.046	0.039	-0.192	-0.486	0.062
Seedlings[4]	0.062	0.061	0.128	0.091	0.039	0.053

[1] Basal area

[2] Proportion of total BA (m^2/ha)

[3] Proportion of canopy closure

[4] Presence or absence

Table 3.5. Means for variables that maximize the likelihood of a northern goshawk nest occurring in a vegetative class on the North Kaibab Ranger District, Kaibab National Forest, Arizona.

Variable	Vegetation class					
	Pinyon-Juniper	Ponderosa Pine	Mixed-conifer	Spruce-Dominated Mix	Deciduous-Dominated Mix	Openings
Aspect	----	----	181	155	129	----
Slope (%)	18	6	8	7	9	7
Elevation (m)	----	----	2605	2682	----	2490
Landform	----	----	0.002	-0.046	0.327	----
Total BA[1] (m^2 /ha)	17	29	39	36	30	2
Ponderosa Pine BA[2]	----	----	----	0.228	----	----
Spruce-fir BA[2]	0.068	0.055	0.706	----	0.442	0.052
Aspen BA[2]	0.000	0.079	0.138	----	0.825	0.087
Canopy[3]	0.800	0.931	1.035	1.038	1.057	0.267
Understory Height (m)	0.947	0.506	0.499	0.682	0.592	0.772

Table 3.5.(continued)

Seedlings[4]	0.371	0.639	0.887	0.944	0.897	0.222

[1] Basal area

[2] Proportion of total BA (m^2/ha)

[3] Proportion of canopy closure

[4] Presence or absence

and in drainages, particularly where smaller proportions of spruce and fir, but greater proportions of aspen basal area, occur. Elevations lower than approximately 2,600 m als, understory vegetation taller than 0.5 m, dense canopy closures and, in particular, seedlings and saplings also improved the likelihood for nest habitat in the mixed-conifer forest type. In spruce-dominated areas, higher nest-use was associated with less total basal area -- although proportions of ponderosa pine grater than 23%, particularly concurrent with shorter (< 7 m) understory heights -- and somewhat greater canopy closure. Flatter, east-facing slopes, higher elevations than approximately 2,680 m asl, and gradual ridges on the landscape also increase the likelihood for locating a nest in spruce-dominated landscapes. In deciduous-dominated forests, nest site use was enhanced by the presence of ridges and, especially, steeper (> 9%) slopes with south or south-west facing aspects, shorter (< 6 m) understory vegetation, and greater amounts of total basal area, including larger proportions (> 44%) of spruce-fir basal area, but lower proportions (< 82%) of aspen basal area. Smaller canopy closures, more typical of spruce-fir than of aspen, also improve the potential for nesting. It follows that openings, which are devoid of trees, require greater amounts of total basal area than 2 m^2/ha to improve their potential for nest site use. Greater amounts of aspen, which are generally a seral species in open, following a disturbance, increase nest use potential in particular. Seedlings, saplings, and taller understory vegetation are also favored. According to our logistic model, none of the pinyon-juniper forest class was considered "good" nest habitat. Nonetheless, we derived coefficients for the variables that would maximize the likelihood of a nest occurrence in this forest type. These conditions included flatter slopes (< 18%) and the presence of seedlings and saplings, greater total basal area (> 17

m^2/ha), but smaller proportions (< 7%) of spruce-fir basal area, and a slightly more open (< 80%) canopy. Overall, our model suggests that the presence of seedlings and/or saplings improves nest habitat in all vegetation classes.

Simulating the Spatial Distribution of Nests

The transformed K-function (Fig. 3.5) of the spatial distribution of individual goshawk nests (N = 27) in the rectangular region B shows some territoriality as the empirical K-function extends below the lower simulation envelope for distances less than 2 km. The minimum distance observed between active nests in 1998 was 1.6 km. This indicates that there are fewer pairs of nests within a 2 km distances than what we would expect if the nests were randomly distributed, and that those nests were regularly distributed. At distances greater than 2 km, the empirical K-function is contained within the simulation envelopes, indicating that the spatial distribution of goshawk nests does not differ significantly from a random spatial pattern. The Cramér-von Mises goodness-of-fit statistic also indicated some non-randomness in the spatial distribution of goshawk nests (Table 3.6). The p-value associated with this test was ≤ 0.14 for all distances ≤ 16 km. The strongest degree of non-randomness ($p < 0.05$) was observed for distances less than 6 km.

When the Gibbsian pairwise potential model was fit to the nest point data, model PF2 ($\hat{\alpha} = 0.005204$, $\hat{\beta} = 0.005923$) (Fig. 3.6) was selected as the best fitting model based on the AIC. The shape of the potential function suggests that individual nesting pairs of goshawks have a repulsive tendency toward one another and that the territorial effects between individual pairs decrease with increasing distance between nests (i.e., soft-core model). The point at which the potential energy approaches zero ($\approx$ 20 km) provides an

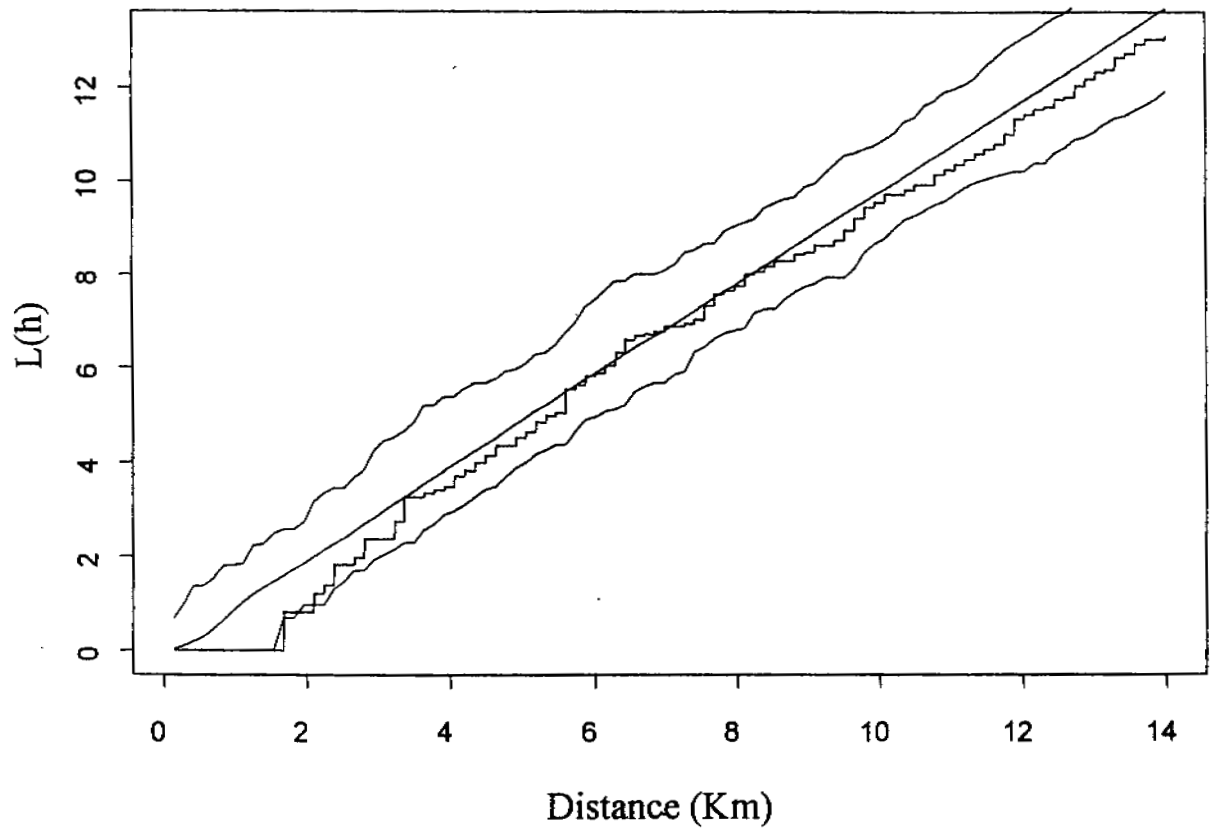


Figure 3.5. Plot of the transformed K-function, $L(h) = \{K(h)/\pi\}^{1/2}$, against distance h, used to model the spatial arrangement of individual northern goshawk nests on the bounded region (B) on the North Kaibab Ranger District, Kaibab National Forest, Arizona. The stair-step line represents the empirical K-function calculated from the data; continuous lines represent the upper, average, and lower 99% simulation envelopes for 200 realizations of a spatial Poisson process.

Table 3.6. Results of the Cramér von-Mises goodness-of-fit test used to test the null hypothesis that northern goshawk nests in 1998 were randomly distributed on the North Kaibab Ranger District, Kaibab National Forest, Arizona.

Distance (km)	Test Statistic	P-value
2	282.25	0.00
3	238.38	0.03
4	195.17	0.00
5	163.32	0.01
6	136.43	0.02
7	117.68	0.06
8	109.72	0.13
9	95.31	0.13
10	98.55	0.10
11	96.42	0.09
12	101.12	0.03
13	104.97	0.13
14	107.94	0.11
15	110.31	0.14
16	119.35	0.08

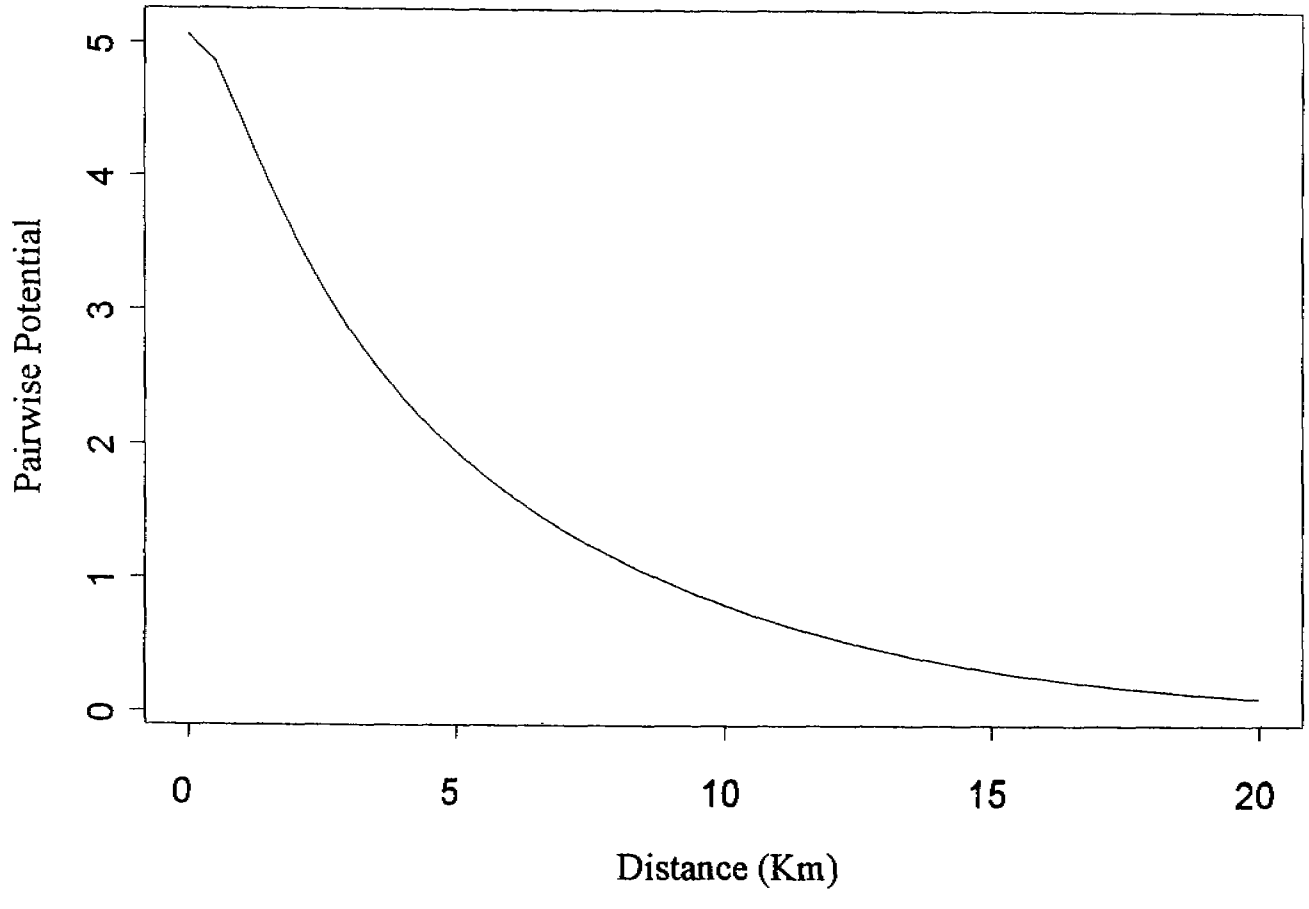


Figure 3.6. Plot of the fitted pairwise potential model (PF2) for individual northern goshawk nests on the bounded region (B) on the North Kaibab Ranger District, Kaibab National Forest, Arizona.

estimate of the maximum zone (circular area) of territoriality around individual nests. This result corroborates the above-mentioned results.

The transformed empirical K-function for the component of the point process model that describes the spatial interaction between individual nests (Fig. 3.7a) is contained within the bounds of the simulation intervals indicating the model provides a good fit to the data. In the range of 5.5 km to 9.5 km, the point process model shows a more regular pattern than that observed in the data. Territories defended by goshawks may be irregular in shape, especially in years when neighboring pairs are not breeding, and their nests may be located near the edge of their territories. Thus, at coarser scales there may be a tendency for some type of clustering of nests. In contrast, the model assumes the nests are at the center of their territories and exhibit an equal territorial force in all directions, resulting in a more regular pattern at all scales. The fact that the empirical K-function is contained within the simulation envelopes suggest the following two hypotheses: 1) the distribution of goshawk nests are spatially independent of forest structure; and 2) there is enough available habitat for nests on the study area as to not limit the spatial distribution of individual goshawk nests.

Except for distances less than 2 km, the transformed empirical K-function for the forest structure component of the point process model (Fig. 3.7b) is contained within the bounds of the simulation intervals. This graph looks similar to the one obtained when we tested for csr (Fig. 3.5), suggesting that if we allocate nests using the potential energy associated with forest structure we generate a pattern similar to that of a random one. This result supports the second hypothesis that the current availability of good nest

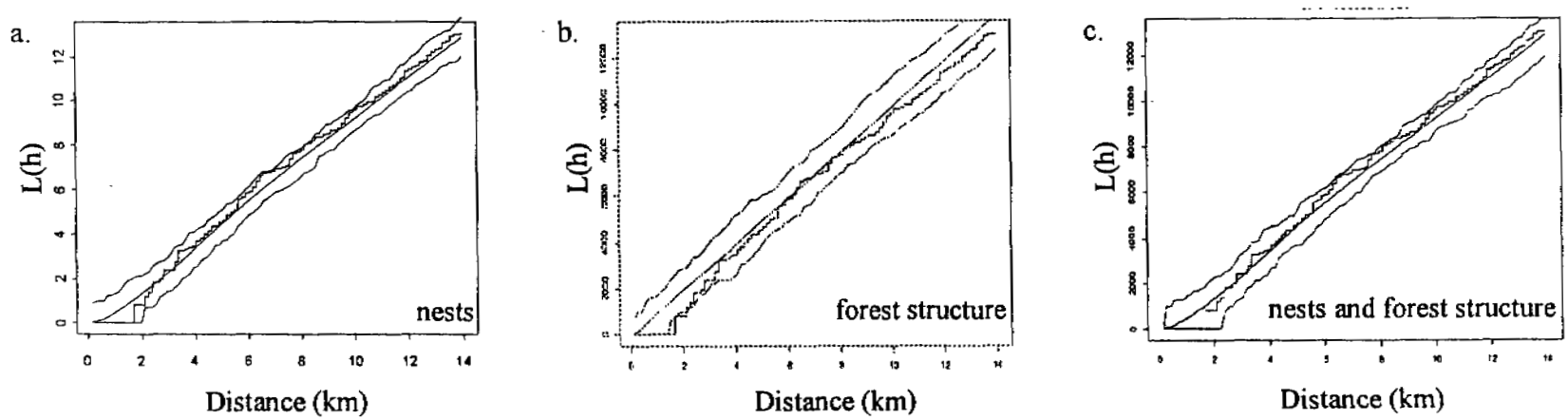


Figure 3.7. Plot of the transformed K-function, $L(h) = \{K(h)/\pi\}^{1/2}$, against distance h, used to model the spatial arrangement of individual northern goshawk nests on the study area on the North Kaibab Ranger District, Kaibab National Forest, Arizona. The stairstep line represents the empirical K-function calculated from the data and the continuous lines represent the upper, average, and lower 99% simulation envelopes for 200 realizations of the (a) nest component of the point process model, (b) forest component of the point process model, (c) point process model that takes into consideration the territoriality between individual active nests and forest structure.

locations on the study area is not a factor limiting the spatial distribution of active goshawk nests.

The transformed empirical K-function for the complete model (Fig. 3.7c) is contained within the bounds of the simulation intervals indicating that the spatial model is capable of describing the distribution of nests on the study area, and in turn, provides a measure of the spatial dependency among individual nests and forest structure. Realization of the final model allows us to predict the location of 27 nest points within the bounded region B (Fig. 3.8) and 96 nest points on the entire KNF (Fig. 3.9). The distribution of nest site probabilities associated with the predicted points depicted in Figure 3.9 did not differ ($\chi^2 = 11.14$, df = 9, P-value = 0.266) from the nest site probabilities associated with active nest in 1998 on the study area (Table 3.7).

DISCUSSION AND CONCLUSIONS

We present a flexible point process model that describes the spatial dependency between the location of active goshawk nests and forest structure. The model assumes that individual nests are distributed according to the potential energy associated with the structure of the forest and a conspecific competitive effect (territoriality).

On the NKRD, it is apparent that suitable nest habitat is not limiting the distribution and abundance of goshawks. Instead, territoriality and what appear to be non-compressible territories limit the distribution and abundance of the nesting population. Within territories, choices of nest locations appear to be limited by the availability of sites with "good" nest habitat structure and topography (see Reynolds et al. 1992). While territory size and density probably reflect the abundance, distribution,

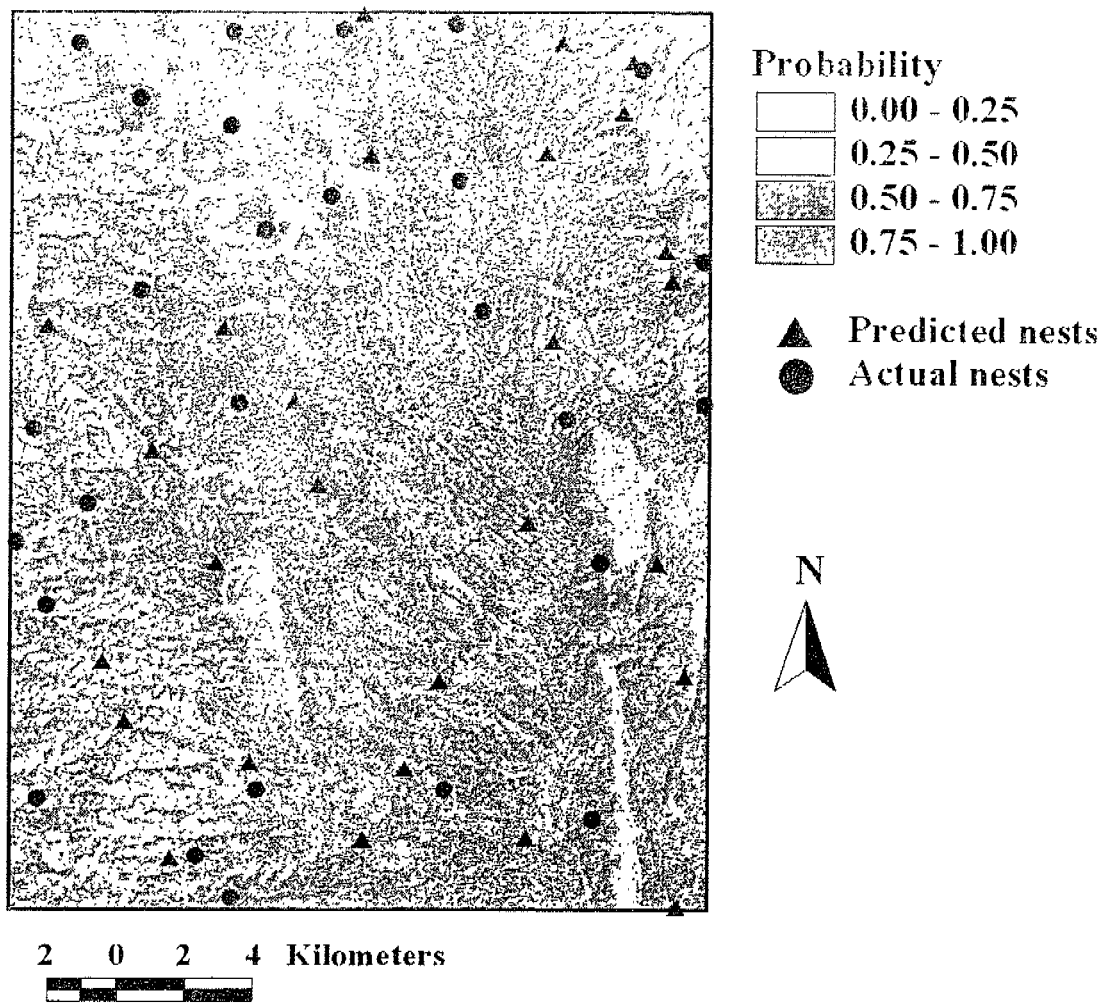


Figure 3.8. Realization of the point process model (triangles) that takes into consideration the territoriality between individual northern goshawk nests and forest structure on the North Kaibab Ranger District, Kaibab National Forest, Arizona. The locations of 27 active northern goshawk nests (circles) used in fitting the model are plotted for comparison. The point patterns are overlaid on a surface showing the probability of finding a northern goshawk nest within the bounded region (B) on the study area associated with forest structure. Areas with a low probability (poor net areas) are lighter in color and areas with a high probability (good nest areas) are darker in color.

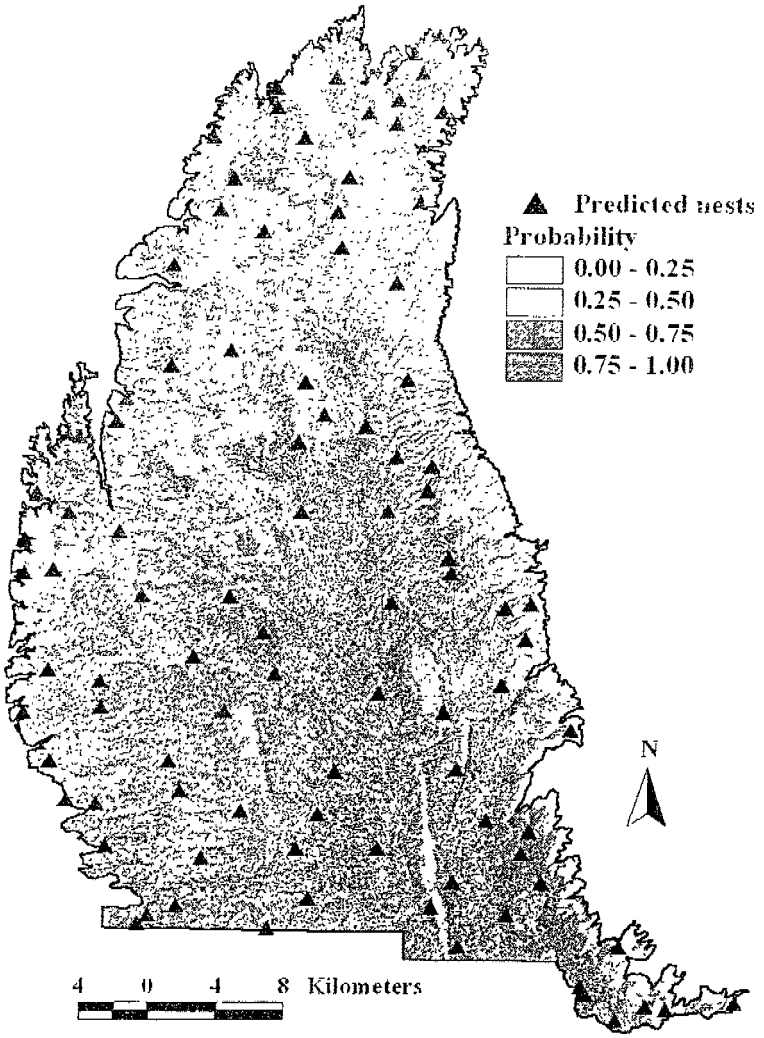


Figure 3.9. Realization of the point process model (triangles) that takes into consideration the territoriality between individual northern goshawk nests and forest structure on the North Kaibab Ranger District, Kaibab National Forest, Arizona. The predicted point pattern of nests is overlaid on a surface showing the probability associated with forest structure of finding a northern goshawk nest within the study area. Areas with a low probability (poor nest areas) are lighter in color and areas with a high probability (good nest areas) are darker in color. The probabilities associated with each simulated point do not differ ($\chi^2 = 11.14$, df = 9, P-value = 0.266) from those of actual nests.

Table 3.7. Distribution of probabilities of finding a northern goshawk nest associated with predicted and observed (1998) nest points on the North Kaibab Ranger District, Kaibab National Forest, Arizona.

Observed Nests		Predicted nests	
Probability	Frequency	Probability	Frequency
0.00-0.10	9	0.00-0.10	11
0.10-0.20	5	0.10-0.20	3
0.20-0.30	5	0.20-0.30	6
0.30-0.40	11	0.30-0.40	8
0.40-0.50	12	0.40-0.50	9
0.50-0.60	12	0.50-0.60	11
0.60-0.70	7	0.60-0.70	13
0.70-0.80	7	0.70-0.80	11
0.80-0.90	11	0.80-0.90	12
0.90-1.00	17	0.90-1.00	14

quality. and accessibility of prey on the study area (Newton et al. 1977, Nilsson et al. 1982, Kenward and Widén 1989, Kenward and Widén 1997, Kenward et al. 2001). Nest locations in a given year were regularly distributed with a minimum of 1.6 km between active nests. Although goshawks need only a small patch (about 0.01-0.10 km^2) of suitable habitat for nests, the reproductive "quality" of those sites should, at least, be partially determined by the capabilities of the surrounding habitat to support populations of diverse prey and provide appropriate foraging opportunities for goshawks (Reynolds et al. 1992, Widén 1997). Although the quality of territories was not determined as part of this study, it was found to vary across the study area (Chapter 4). In reality, only a portion of the good nest sites may actually have successful breeders. Regardless, alterations of forest structure in large areas due to improper management or natural disturbances may reduce the quality of or eliminate nest habitat to the point that the distribution of goshawk territories in our model is affected.

In our final model, the total effect on nest location is dominated by territoriality; however, within territories, the distribution of high and low potential habitat, based on forest structure, plays an important role in nest location. These relationships may reflect the management history on the NKRD. While many forests in the southwest received heavy railroad logging in the 1800 and early 1990s, the Kaibab Plateau experienced limited harvesting (Pearson 1950) because it is isolated from the railroad by the Grand Canyon. Management on the NKRD since the 1960s has been variable; some areas have been intensively managed (i.e., seed tree, shelterwood cuts, clearcuts), while others received less tree cutting (i.e., thinning, individual tree selection). Intensively managed areas generally contain lower quality nest habitat; while the less heavily managed areas

have higher quality nest habitat (S. M. Joy, pers. obs.). With the implementation of management to enhance goshawk nest and foraging habitat (Reynolds et al. 1992), the forest structure should be restored to historical conditions and more of the study area should become suitable habitat.

The varying importance and direction (positive–increase, negative–decrease) of forest structural coefficients in the potential model reflect the probability of sites within each vegetation class to contain a goshawk nest. Increased total basal area in all vegetation classes, except the spruce-dominated type, was favored. Less spruce-fir and aspen in ponderosa pine; greater proportions of ponderosa pine in spruce-dominated forests; less spruce-fir, but more aspen in the mixed-conifer forest; and less aspen, but more spruce-fir in the deciduous-dominated forest are particularly important. In the ponderosa pine zone, more spruce and fir would increase the density of smaller trees and may restrict access by the goshawk to its nest; whereas, more aspen might decrease the protection conferred at nests by adjacent crowns, especially early in the nesting period prior to leaf-out. In spruce-dominated habitat, ponderosa pine trees provide large horizontal branches typically used for nest substrates, easier access to the nest by adults, and would occupy a greater proportion of the canopy, conveying greater above-canopy protection. Because the mixed-conifer forest is typically dense in both the overstory and understory (S. M. Joy, pers. obs.), increasing amounts of aspen basal area in a site suggest that mixed-conifer forests containing mature aspen trees improve nest habitat quality. Mature aspen trees within mixed-conifer forests provide a more open understory, large open crowns for nest placement, and easier access to nests. Mixed-conifer and spruce-fir forests containing some mature aspen trees provide more opportunities for

optimal nest tree use than pure aspen stands, probably due to the latter's openness prior to leaf-out. Our model suggests that regenerating large openings created by management or natural disturbance, particularly with aspen, is needed to restore nest habitat in openings and increase the availability of high quality sites. Aspen is generally seral to disturbance and, subsequently, provides the canopy closure needed by more shade-tolerant pines to regenerate.

Within the ponderosa pine forest type, nest habitat is also enhanced by greater canopy closures and flatter slopes typically associated with pine regeneration. Less canopy closures and steeper, southeast-facing slopes associated with ridges improve nest habitat in the deciduous-dominated forest types by potentially encouraging the regeneration of ponderosa pine, rather than spruce-fir species. Steeper slopes associated with drainages at elevations below 2,600 m asl, easterly-facing exposures, and dense canopy closures, improve nest habitat within the mixed-conifer forest. However, nests in mixed conifer are typically found in trees (usually ponderosa pine) associated with ridges. Perhaps at lower elevations, east facing slopes in drainages encourage the regeneration ponderosa pine trees or aspen, which would eventually provide greater canopy coverage as well as a greater number of useable nest trees and limit the amount of fir regeneration. The habitat characteristics that create high quality nest sites in spruce-dominated forests -- east-facing exposures with slightly elevated slopes leading to ridges -- would most likely enhance the growth of more spruce and fir. Pinyon-juniper, which tends to grow on steep, hot, west-facing slopes above 2,182 m asl on the NKRD, would be improved for nests by flatter slopes and less canopy closure. Although goshawk nests are not found in pinyon-juniper forests on the study area, they do occur within cooler

stands of ponderosa pine in drainages that extend into the pinyon-juniper zone. It seems the habitat characteristics identified by our model provide cooler sites that may encourage the growth of ponderosa pine.

Our model suggests that the presence of seedlings and/or saplings "improves" nest habitat in all vegetation classes. Nevertheless, the nature of tree regeneration in a nest area varied widely. In some areas, seedlings/saplings may be small and few, such that they do not impose a physical or visual barrier for nesting hawks. However, as saplings increase in size, they probably hinder goshawk flight to and from nest trees. They also increase the risk of forest-killing fire and loss of eggs or nestlings. Regardless, the presence or absence of seedlings and saplings alone is insufficient to provide a biologically meaningful index of nest site potential. Shrubs and herbaceous understory height may also be a poor predictor for similar reasons.

Although the majority (86%) of openings on the study area were classified as poor nest habitat, some openings (14%) were classified as good habitat. Within ponderosa pine and mixed-conifer forests on the NKRD, small (10-m x 10-m) openings are common. These small openings may represent some of the 14% that fell in good nest habitat; whereas, some openings classified as good nest habitat may be classification errors attributed to the "open" vegetation class. Openings contained the highest (23%) classification error rate of all vegetation classes (Joy et al. 2001, Chapter 1).

Between 1991 and 1998, 57 out of 204 active nests were in "poor" nest habitat. Of the 57, 80% (45) were found within 10 m of a "good" nest site, regardless of vegetation class. The classification of these nest locations as poor could be due to errors in the mapping of nest trees or registration of the Landsat information. Mapping errors would

tend to lower the significance of the logistic regression model, indicating that good nest locations are more randomly distributed (Stoms et al. 1992). That is, we would be unable to discriminate nest from random sites. However, we believe the majority of our nests were recorded to within 3 m of their actual locations. Alternatively, if the spatial resolution of our models did not capture the geographic scale at which goshawks choose nest trees (e.g., if nest trees were selected based on local prey availability), we might also expect more nests to be in poor sites. Non-nesting territorial goshawks may have introduced spaces into the distribution of nests, thus confusing our estimation of good and poor nest habitat. Finally, some goshawks may chose to nest in poor habitat. Despite all the possible sources of error in our analysis, we believe that territorial interactions (Ozesmi and Mitsch 1997) among breeding goshawks, as well as potential interactions with other raptors (e.g., R. T. Reynolds, unpubl. data, Janes 1984), explain why not all of the active nests were located in good sites.

Treating forest structural components as one continuous variable allows the introduction of environmental heterogeneity into the point process model, which in turn allows a modeling of the spatial interaction between goshawk pairs at nests, both on a local and regional level. Such a model is useful in simulating the effects that changes in the forest, either through manmade or natural disturbances (e.g., fires), have on the spatial dynamics of a goshawk population. This can be accomplished by systematically changing the potential energy associated with forest structure, and then observing how this change influences the spatial distribution of goshawk nests. As some nest sites become unsuitable because of harvesting or other types of disturbances, goshawk pairs may move to an alternate nest within their territory. The location of alternate nests will

depend on the availability of suitable sites. The model also provides information on the total potential occupancy of the forest. Moreover, when we incorporate the demographics of the goshawk population into the point process model, it should be possible to study the spatio-temporal behavior of the goshawk population as influenced by forest management activities.

The information derived from such a model should benefit researchers and managers interested in ecosystem processes by providing a better understanding of the influence that coarse- and fine-scale spatial variability have on the abundance and productivity of goshawk populations. Our approach to modeling the spatial dynamics of an individual species with their habitat can be used in a variety of applications and study areas. However, inferences from the model generated here should not be made beyond the scope of the study area. In alternate study areas, where goshawks may occur in lower densities than on the NKRD, where nest spacing might be irregular, or where habitat data exist at a coarser resolution, a new point process model should be developed.

ACKNOWLEDGEMENTS

We are grateful to the many field technicians that contributed to the goshawk monitoring effort between 1991 and 1998. Special thanks to V. L. Thomas for assistance with ArcView® and Avenue programming. We thank B. S. Cade, P. Kaval, M. L. Farnsworth, R. M. King, M. S. Williams, and 2 anonymous reviewers for thorough reviews of an earlier draft of the manuscript. The North Kaibab Ranger District (Fredonia, AZ) provided housing and logistical during field seasons. Regions 3

(Albuquerque, NM) and the Rocky Mountain Research Station (Fort Collins, CO) of the USDA Forest Service funded this research effort.

LITERATURE CITED

Akaike, H. 1977. On entropy maximization principle. Pages 27-41 *in* P. R. Krishnaiah, ed. Applications of Statistics. North-Holland, Amsterdam.

Augustin, N.H., M.A. Mugglestone, and S.T. Buckland. 1996. An autologistic model for the spatial distribution of wildlife. J. Appl. Ecol. 33:339-347.

Baker, B.W., B.S. Cade, W.L. Mangus, and J.L. McMillen. 1995. Spatial analysis of sandhill crane nesting habitat. J. Wildl. Manage. 59:752-758.

Besag, J.E., and P.J. Diggle. 1977. Simple Monte Carlo tests for spatial patterns. Appl. Stat. 26:327-333.

Block, W.M., M.L. Morrison, and M.H. Reiser. 1994. The northern goshawk: ecology and management. Studies Avian Bio. No 16. 136 p.

Burnett, H. 1991. Green island in the sky. Amer. For. 97:44-47.

Cade, T.J. 1960. Ecology of the peregrine and gyrfalcon populations in Alaska. Univ. Calif. Publ. Zool. 63:151-290.

Cressie, N. 1991. Statistics for spatial data. John Wiley & Sons, NY. 900 p.

Crocker-Bedford, C. 1990. Goshawk reproduction and forest management. Wildl. Soc. Bull. 18:262-269.

Diggle, P.J., D.J. Gates, and A. Stibbard. 1987. A nonparametric estimator for pair-wise interaction point processes. Biometrics 74:763-770.

Efron, B., and Tibshirani, R.J. 1993. An introduction to the bootstrap. Chapman & Hall, Inc., NY. 436 p.

ESRI. 1998. ArcView® 3.1 on-line help manual. Environ. Res. Inst., Inc., Redlands, CA.

Finn, S.P., J.M. Marzluff, and D.E. Varland. 2002. Effects of landscape and local habitat attributes on northern goshawk site occupancy in western Washington. For. Sci. 48:427-436.

Hosmer, D.W., and S. Lemeshow. 1989. Applied logistic regression. John Wiley & Sons, NY. 307 p.

Janes, S.W. 1984. Influences of territory composition and interspecific competition on red-tailed hawk reproductive success. Ecology 65:862-870.

Joy, S.M., R.T. Reynolds, and D.G. Leslie. 1994. Northern goshawk broadcast surveys: hawk response variables and survey costs. Studies Avian Biol. 16:24-31.

Kennedy, P.L. 1997. The northern goshawk (*Accipiter gentilis atricapillus*): Is there evidence of a population decline? J. Raptor Res. 31:95-106.

Kennedy, P.L. and D.W. Stahlecker. 1993. Responsiveness of nesting northern goshawks to taped broadcasts of three conspecific calls. J. Wildl. Manage. 57:249-257.

Kenward, R.E. 1982. Goshawk hunting behaviour, and range size as a function of food and habitat availability. J. Anim. Ecol. 51:69-80.

Kenward, R.E., and P. Widén. 1989. Do goshawks Accipiter gentilis need forests? Some conservation lessons from radio tracking. Pages 561-567 in B.U. Meyburg and R.D. Chancellor, eds. Raptors in the modern world. World Working Group on Birds of Prey and Owls, Berlin, Germany.

Kenward, R.E., R. T. Clarke, K. H. Hodder, and S.S. Walls. 2001. Density and linkage estimators of home range: nearest-neighbor clustering defines multinuclear cores. Ecology 82:1905-1920.

Kostrzewa, A. 1996. A comparative study of nest-site occupancy and breeding performance as indicators for nesting-habitat quality in three European raptor species. Ethol. Ecol. Evol. 8:1-18.

Lemmon, P.E. 1956. A spherical densiometer for estimating forest overstory density. For. Sci. 2:314-320.

Lemmon, P.E. 1957. A new instrument for measuring forest overstory density. J. For. 55:667-668.

Manly, B.F.J., L.L. McDonald, and D.L. Thomas. 1993. Resource selection by animals. Statistical design and analysis for field studies. Chapman and Hall, London, England. 175 p.

McNab, W.H. 1989. Terrain shape index: quantifying effect of minor landforms on tree height. For. Sci. 35:91-104.

Mladenoff, D.J., T.A. Sickley, and A.P. Wydeven. 1999. Predicting gray wolf landscape recolonization: logistic regression models vs. new field data. Ecol. Applic. 9:37-44.

Neter, J., W. Wasserman, and M.H. Kutner. 1985. Applied linear statistical models: regression, analysis of variance, and experimental designs. Richard D. Irwin, Inc., Homewood, IL. 1127 p.

Newton, I., M. Marquiss, D.N. Weir, and D. Moss. 1977. Spacing of sparrowhawk nesting territories. J. Anim. Ecol. 46:425-441.

Newton, I. 1979. Population Ecology of Raptors. Buteo Books, Vermillion, South Dakota. 399 p.

Nilsson, I., S.G. Nilsson, and M. Sylvén. 1982. Diet choice, resource depression, and the regular nest spacing of birds of prey. Biol J. Linnean Soc. 18:1-9.

Ogata, Y., and M. Tanemura. 1981. Estimation of interactive potentials of spatial point patterns through the maximum likelihood procedure. Ann. Instit. Stat. Math., Part B. 33:315-338.

Ogata, Y., and M. Tanemura. 1985. Estimation of interactive potentials of marked spatial point patterns through the maximum likelihood methods. Biometrics 41:421-433.

Ogata, Y., and M. Tanemura. 1989. Likelihood estimation of soft-core interaction potentials for Gibbsian point patterns. Ann. Inst. Stat. Math. 41:583-600.

Ozesmi, U. and W.J. Mitsch. 1997. A spatial habitat model for the marsh-breeding red-winged blackbird (*Agelaius phoeniceus* L.) in coastal Lake Erie wetlands. Ecol. Model. 101:138-152.

Pearce, J., and S. Ferrier. 2000. Evaluating the predictive performance of habitat models developed using logistic regression. Ecol. Model. 133:225-245.

Pearson, G.A. 1950. Management of ponderosa pine in the southwest. US Dept. of Agric., Forest Service Monograph. No. 6. 213 p.

Peck, J. 2000. Seeing the forest through the eyes of a hawk: an evaluation of recent efforts to protect northern goshawk populations in southwestern forests. Nat. Res. J. 40:125-156.

Pierera, J.M.C., and R.M. Itami. 1991. GIS-based habitat modeling using logistic multiple regression: a study of the Mt. Graham red squirrel. Photog. Eng. Rem. Sens. 57:1475-1486.

Ratcliffe, P.A. 1962. Breeding density of the peregrine *Falco peregrinus* and raven *Corvus corax*. Ibis 104:3-39.

Reich, R.M., C.D. Bonham, and K.L. Metzger. 1997. Modeling small-scale spatial interactions of shortgrass prairie species. Ecol. Model. 101:163-174.

Reich, R.M., J. Lundquist, and V.A. Bravo. 2000. Spatial relationship of resident and migratory birds and canopy openings in diseased ponderosa pine forests. J. Environ. Model. Softw. 15:189-197.

Reynolds, R.T. 1982. North American accipiter hawks. Pages 288-289 *in* D. E. Davis, ed. Handbook of Census Methods for Terrestrial Vertebrates, CRC Press, Inc., Boca Raton, FL.

Reynolds, R.T. 1983. Management of western coniferous forest habitat for nesting Accipiter hawks. U.S.D.A. For. Serv., Gen. Tech. Rep. RM-102, Rocky Mount. For. Range Exp. Stat., Fort Collins, CO. 7 p.

Reynolds, R.T. 1989. Accipiters. Pages 92-102 *in* B.G. Pendleton, G.E. Ruibal, D. L. Krahe, K. Steenhof, M. N. Kochert, M. N. LeFranc, Jr., eds. Proc. of the Western Raptor Management Symposium and Workshop. Nat. Wildl. Fed. Sci. Tech. Series No. 12. Nat. Wildl. Fed., Washington, DC.

Reynolds, R. T., R. T. Graham, M. H. Reiser, R. L. Bassett, P. L. Kennedy, D. A. Boyce, Jr., G. Goodwin, R. Smith, and E. L. Fisher. 1992. Management recommendations for the northern goshawk in the southwestern United States. U.S.D.A. For. Serv., Gen. Tech. Rep. RM-217, Rocky Mount. For. Range Exp. Stat., Fort Collins, CO. 90 p.

Reynolds, R.T., and S.M. Joy. 1998. Distribution, territory occupancy, dispersal, and demography of northern goshawks on the Kaibab Plateau, Arizona. Final Rep., Ariz. Game and Fish Herit. Proj. No I94045. 76 p.

Reynolds, R.T., S.M. Joy, and D.G. Leslie. 1994. Nest productivity, fidelity, and spacing of northern goshawks in northern Arizona. Studies Avian Biol. 16: 106-113.

Reynolds, R.T., and H.M. Wight. 1978. Distribution, density, and productivity of *accipiter* hawks breeding in Oregon. Wilson Bull. 90:182-196.

Ricklefs, R.E. 1987. Community diversity: relative role of local and regional processes. Science 235:167-171.

Ripley, B.D. 1977. Modeling spatial patterns. J. Royal Stat. Soc., Series B 39:172-192.

Ripley, B.D. 1990. Gibbsian interaction models. Pages. 3-25 *in* D.A. Griffith, ed. Spatial Statistics: Past, present, and future. Inst. Math. Geogr., Syracuse Univ., NY.

Ripple, W. J., P. D. Lattin, K. T. Hershey, F. F. Wagner, and E. C. Meslow. 1997. Landscape composition and pattern around northern spotted owl nest sites in southwest Oregon. J. Wildl. Manage. 61:151-158.

Selås, V. 1997. Influence of prey availability on re-establishment of goshawk *Accipiter gentilis* nesting territories. Ornis Fenn. 74:113-120.

Speiser, R., and T. Bosakowski. 1984. History, status, and future management of goshawk nesting in New Jersey. Rec. New Jersey Birds. 10:29-33.

Stoms, D.M., F.W. Davis, and C.B. Cogan. 1992. Sensitivity of wildlife habitat models to uncertainty in GIS data. Photogr. Eng. Rem. Sens. 58:843-850

Swindle, K.A., W.J. Ripple, E.C. Meslow, and D. Schafer. 1999. Old-forest distribution around spotted owl nests in the central Cascade Mountains, Oregon. J. Wildl. Manage. 63:1212-1221.

Thome, D.M., C.J. Zael, and L.V. Diller. 1999. Forest stand characteristics and reproduction of northern spotted owl in managed north-coastal California forests. J. Wildl. Manage. 63:44-59.

Van Horne, B. 1983. Density as a misleading indicator of habitat quality. J. Wildl. Manage. 47:893-901.

van Manen, F.T., and M.R. Pelton. 1997. A GIS model to predict black bear habitat use. J. For. 95:6-12.

Vickery, P.D., M.L. Hunter, Jr., and J.V. Wells. 1992. Use of a new reproductive index to evaluate relationships between habitat quality and breeding success. Auk 109:697-705.

Weaver, H. 1951. Fire as an ecological factor in the southwestern ponderosa pine forests. J. For. 49:93-98.

White, M.A., and J.T. Vankat. 1993. Middle and high elevation coniferous forest communities of the North Rim region of the Grand Canyon National Park, Arizona, USA. Vegetat. 109:161-174.

Widén, P. 1985. Breeding and movements of goshawk in boreal forests in Sweden. Holarct. Ecol. 8:273-279.

Widén, P. 1997. How, and why, is the goshawk (*Accipiter gentilis*) affected by modern forest management in Fennoscandia? J. Raptor Res. 31:107-113.

CHAPTER FOUR

CORRELATES OF NORTHERN GOSHAWK TERRITORY QUALITY

ABSTRACT

I identified qualitative differences between potential "source" and "sink" habitats for northern goshawks (*Accipiter gentilis*; hereafter, goshawk) and random locations on the North Kaibab Ranger District (NKRD), Kaibab National Forest (KNF), Arizona by comparing vegetative characteristics among random plots and plots at nests on goshawk territories with higher (source) and lower (sink) reproductive performance. Differences in territory quality were based on the annual rate of egg laying and total fledglings produced during the study period (1991-2000). Neither the recruitment of young, first-time breeders nor the tenure of females on territories influenced territory quality. However, more females were replaced on lower than higher quality territories, suggesting that a female's experience on a territory (vis-à-vis her ability to fledge more young) may influence the territory's assessment of quality. Habitat characteristics (the proportion of pinyon-juniper, ponderosa pine, mixed conifer, spruce- and deciduous-dominated mixes, openings, and vegetative diversity) were characterized at five spatial scales within circle and ring plots to evaluate the amount and spatial arrangement of vegetation, respectively, around goshawk and random sample plots. The majority of differences found in the vegetative components, which were similar for circle and ring plots, occurred between

higher quality territories and random plots. Although differences in vegetative composition and diversity between source and sink territories were not found, the use additional variables, such as forest structural components, may discriminate better between these territory classes. Logistic models for predicting goshawk territory locations and territories of higher quality based on select vegetative characteristics were > 80% accurate. Although the predictive models presented here can be used to assess the probable effects of forest management on territory quality, they should not be used as targets for implementing forest management practices.

INTRODUCTION

Two goals of ecological studies are to (1) determine how landscape heterogeneity, including anthropogenic-caused changes in species composition, structure and pattern, affect the demographic processes of animal species, and (2) develop predictive models for resource management (Risser el al. 1984, Morrison et al. 1992). Animal populations typically occupy a range of habitat conditions and are believed to fill landscapes according to their fitness in a potential habitat (Brown 1969, Fretwell and Lucas 1970, Pulliam 1988). Individuals occupy higher-quality habitats first, and then progressively settle into habitats of lesser quality. In higher quality habitats, reproduction is generally sufficient to sustain population numbers; whereas, in poor habitats, mortality typically exceeds reproduction (Wiens 1986, Pulliam 1988). Some populations in lower quality habitats, or "sinks," persist due to immigration by individuals from higher quality or "source" habitats (Pulliam 1988). This is due, in part, to dominant individuals forcing subordinates to emigrate or disperse from higher quality habitats to sinks. An adequate

proportion of breeding habitats in landscape must therefore be of higher quality to maintain stable breeding populations (Wiens 1986, Pulliam 1988, Howe et al. 1991). Because the relative abundance and spatial arrangement of source-sink habitats can affect population persistence (Pulliam and Danielson 1991, Donovan et al. 1995), conservation and management should focus on assessments of habitat quality and their relationships to source-sink population dynamics.

Bird and mammal populations in forests are increasingly affected by habitat loss and changes in structure associated with landscape-altering human activities, such as tree harvests, livestock grazing, and fire control. Similarly, management-related habitat changes are thought to be responsible for declining populations of northern goshawks (*Accipiter gentilis;* hereafter, goshawk) (Reynolds 1983, 1989; Speiser and Basakowski 1984, Crocker-Bedford 1990, Reynolds et al. 1992; Kennedy 1997, Peck 2000) either by eliminating critical habitat or creating population sinks. As a result, goshawks are considered a "sensitive" species by most National Forests, "special status species" by most National Parks, and "Category II" species by the Fish and Wildlife Service (FWS). Since 1992, the FWS has been petitioned three times to list the goshawk as "threatened."

Although goshawks occur in a wide range of forest types and structures, they typically place their nest in forests with large trees, relatively high canopy closure, and open understories (Reynolds et al.1982, Moore and Henny 1983, Speiser and Basakowski 1987, Hayward and Escano 1989, Siders and Kennedy 1996, Squires and Ruggerio 1996). Goshawks typically forage in mature to old forests, but younger forests, edges, and openings are used (Widén 1989, Bright-Smith and Mannan 1994, Hargis et al. 1994, Younk and Berchard 1994). The use of diverse habitat conditions by hunting goshawks

is reflected in their diets; birds and mammals occupying old forests, edges, and openings are typically eaten (Reynolds et al. 1992, Kenward and Widén 1989). However, despite our knowledge of the habitats used by goshawks for nesting and foraging, we know little about the role that habitat plays in influencing goshawk reproductive success or population persistence.

Since the theoretical development of source-sink models in the 1980s, few raptor studies (Korpimäki 1988, Newton 1991, Kostrzewa 1996, Franklin 1997, Ripple et al. 1997, Thome et al. 1999, Linkhart 2001, Rodewald and Yahner 2001, Finn et al. 2002) have linked direct measures of individual fitness (reproduction, survival) to spatially explicit models of landscape composition, structure, or arrangement, in part due to the time and labor involved in quantifying demographic performance (Wiens 1973). Reynolds and Joy (1998) believe than more than 8-10 years of reproduction data and >50 goshawk territories are required to accurately estimate goshawk demographic performance. Nevertheless, previous attempts to correlate goshawk demographic rates with habitat features (Allison 1996, Desimone 1997, McGrath 1997, Finn et al. 2002) have included few years (≤ 4) and territories (≤ 50) to discriminate habitat-related demographic performance from that due to the influence of individual experience or environmental (weather, prey) stochasticity. Therefore, especially in long-lived species such as goshawks, long periods of population monitoring are recommended (Pulliam 1988, Van Horne 1983, Vickery et al. 1992, Franklin 1997, Van Horne et al. 1997). To date, only one goshawk study (McClaren et al. 2002) has been conducted at sufficient temporal or spatial scales to infer differences in goshawk territory quality from

demographic performance, and comparisons with habitat were not made because minimal differences in nest productivity were found due to the spatial arrangement of nests.

In this study, I assess differences in goshawk territory quality (higher, lower) based on reproductive success, evaluate the influence of age and prior experience on territory quality, identify the amount and spatial arrangement of vegetative characteristics that distinguish between the levels of territory quality and random locations, and develop a predictive model of territory quality. The study takes place on the NKRD in northern Arizona, where we (Joy et al. 1994, Reynolds et al. 1994, Reynolds and Joy 1998) have studied a population of nesting goshawks over a ten-year period (1991-2000).

STUDY AREA

The study area includes 1,281 km^2 of the NKRD on the KNF above 2,182 m above mean sea level (asl) in northern Arizona (Fig. 4.1). Located in the northern 2/3rds of the oval-shaped (95 km x 55 km) Kaibab Plateau, the study area descends steeply to the east and gently to the north and west sides into shrub-steppe plains. To the south, the NKRD shares a border with Grand Canyon National Park-North Rim. The study area is comprised of many drainages, moderately sloping valleys, a few of which have long, narrow (< 1 km wide) meadows containing grasses and herbaceous vegetation, and east-west and north-south oriented canyons. The forest receives approx. 67 cm of annual precipitation (White and Vankat 1993) from late-summer, heavy monsoon showers and snowpacks of 3-6 m.

Six forest classes were identified on the study area (Joy et al. 2001, Chapter 1). (1) Pinyon (*Pinus edulis*)-juniper (*Juniperus* spp.) woodlands (106 km^2) occur approx.

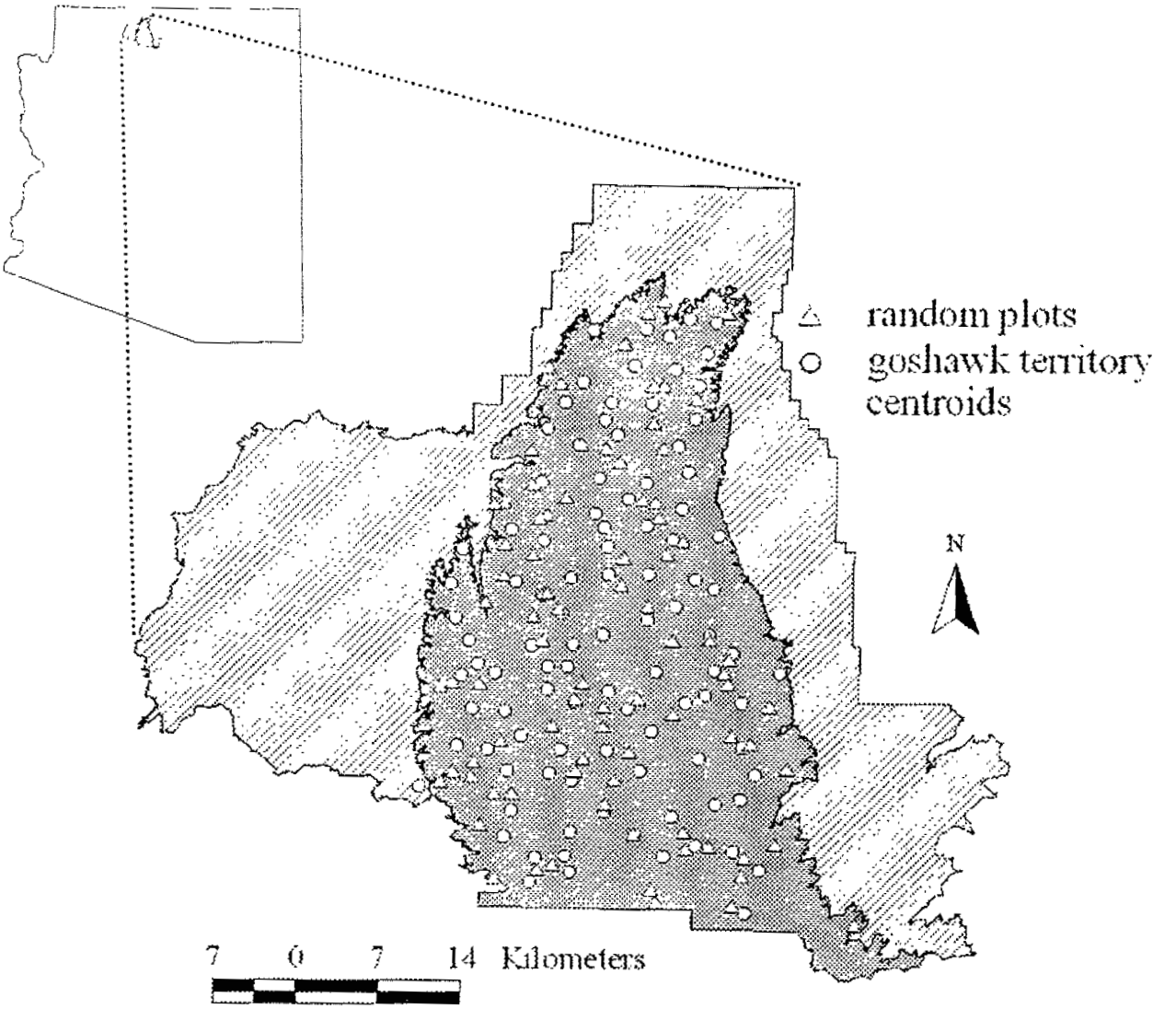


Figure 4.1. Location of the study area (shaded region) ≥ 2,182 m above mean sea level on the North Kaibab Ranger District (hatched region) of the Kaibab National Forest, Arizona, including random plots (triangles) and goshawk territory centroids (circles). Centroids were calculated as the mean coordinate of all active nests between 1991 and 2000 within a territory, weighted by the number of years each nests was active. Active nests are ones in which ≥ 1 egg was laid.

between 2,182-2,250 m and mix with ponderosa pine (*P. ponderosa*) at higher, transitional zones; (2) ponderosa pine comprises over 50% (714 km^2) of the forested area and occurs approx. between 2,250 and 2,550 m asl; (3) mixed conifer (145 km^2), which occurs approx. between 2,550 and 2,650 m asl, is comprised of ponderosa pine mixed with white (*Abies concolor*), Douglas-fir (*Pseudotsuga mensiesii*), quaking aspen (*Populus tremuloides*) and with spruce (*Picea pungens, P. englemannii*) at higher elevations; (4) spruce-dominated mixes occur with sub-alpine fir (*A. lasiocarpa*) (130 km^2) above approx. 2,650 m asl; (5) deciduous (*quaking aspen, Quercus gambeli*)-dominated mixes (112 km^2) occur sparsely throughout the forest and are common where extensive disturbance (fire, logging) has occurred; and (6) Openings (74 km^2) occurred throughout the study area. Openings ranged in size from 100 m^2 to 2.15 km^2.

Forest management began on the NKRD in the 1920s in the form of single-tree (sanitation) selection. Sanitation continued, along with a cluster of small (0.1 km^2) clearcuts (totaling 9.9 km^2) in the south-central part of the study area, until the late-1970s when intensive stand-level management (shelterwood, seed, salvage, removal, and thinning cuts) began. Intensive stand management continued until 1991 when the NKRD implemented forest management designed to enhance the habitat of goshawks and their prey (Reynolds et al. 1992). Livestock grazing was common on the NKRD between the late 1800s and the mid-1920s. Fire suppression since the early 1900s has affected the composition and structure of all forested areas on the NKRD by encouraging regeneration of shade-tolerant tree species.

Some management-caused changes appear to affect the reproductive performance of goshawks by altering the composition and abundance of prey resources. For example,

the Kaibab squirrel (*Sciurus aberti kaibabensis*), a primary goshawk prey species, depends on ponderosa pine for food (cambium, inner bark, cones) and nest sites (Hall 1981). A reduction in historically frequent ground fire through fire suppression has resulted in denser forests and prevented regeneration of ponderosa pine, a shade intolerant species. Loss of ponderosa pine will result inn loss of the Kaibab squirrel, eventually reduced the carrying capacity of these forests for goshawks.

METHODS

Terminology

"Habitat" refers to the local biotic, climatic, and edaphic conditions that make up the goshawk's environment. "Territory" (approx. 1.5 km radius) is the area used and defended by a single pair of goshawks during the nesting season. Each territory may contain one or more alternate "nest areas" (Reynolds et al. 1992), i.e., a 0.15-0.20 km^2 area surrounding a nest tree in which adults build their nests, roost, and may prepare prey for their young. A territory is considered "occupied" in a given year if a pair or single goshawk has been observed in one or more alternate nest area within that territory on two or more occasions, or once if accompanied by molted feathers, feces, prey remains, or recent nest construction. When eggs were laid in a nest, the nest (and territory) is "active." Active nests that fail to fledge young (i.e., in which eggs or nestlings were lost) are "failed." The number of young fledged from nests is the "productivity" or "reproductive output" for that territory. For nests in which eggs are laid, but no young are fledged, productivity is zero. Movements by adults from one breeding territory to another between nesting seasons is termed "breeding dispersal." "Landscape,"

"landscape-scale," and "landscape-level" refer to land surfaces and associated habitats occurring across several kilometers (Turner and Gardner 1991). In this study, landscapes range from the size of a goshawk territory to the entire study area.

Population Sampling

Between 1991 and 2000, field crews (6-20 personnel) searched annually for active goshawk nests. Each year, all previously known nests and territories were monitored and systematic searches were conducted for new nests and territories. By the year 2000, 90% of the study area had been searched. Where the status of pairs on territories remained unknown after initial early spring nest visits, searches on foot (Reynolds 1982) within a 0.50 km radius were conducted around previously active nests and by systematically broadcasting goshawk vocalizations from predetermined stations and transects (Kennedy and Stahlecker 1993, Joy et al. 1994) within a 1.50-km radius around the last known active nests. Both search methods required detecting goshawks by sight, vocalization, or sign of their presence (molted feathers, feces, plucked prey, nest construction). Each year, visits to nests on territories and follow-up searches began in April and ended at the close of post-fledgling dependency (mid-August). All known nest structures in all territories were visited by the first week of incubation to avoid missing early nest failures. The location of active nests was verified using a Trimble Navigation Pathfinder™ Asset Surveyor Global Positioning System (GPS) with an estimated accuracy of 1-3 m. Because new territories were discovered every year, there were a variable number of years of data on resident pairs.

During each breeding season, visits to active nests were made weekly to determine the status of nests and causes of nest failure, and to capture, band, or resight

breeding adults. Adults were trapped during the nestling and early-fledgling periods in nest areas using dho-gaza nets arranged in a "V" around a live great horned owl (*Bubo virginianus*) (Bloom 1987). All birds were fitted with USFWS aluminum leg bands and anodized aluminum colored leg bands marked with unique 2-character alpha-numeric codes readable at up to 50-80 m with 20-40 power spotting scopes (Reynolds et al. 1994). Marked individuals provided confirmation of territory ownership when alternate nests were used. Nest trees were climbed late in the nestling period to band nestlings. Counts of the number of young observed in nests within 10 days of fledging (age of banding) were used to assess annual productivity.

The size of goshawk territories (i.e., defended area) on the study area was unknown; however, the mean distance between active nests each year (n = 10 yr) was rarely less than 3 km ($\bar{x}$ = 3.5 km, SD = 0.2 km; R. T. Reynolds and S. M. Joy, unpubl. data). I therefore assumed that goshawk pairs defended an area within a minimum radius of 1.5 km from their active nest. Because the location of alternate nests can be as much as 1.6 km apart among breeding years (n = 191, $\bar{x}$ = 0.5 km, SD = 46 m, R. T. Reynolds and S. M. Joy, unpubl. data), the habitat components concentrated within each defended area may shift annually. To account for this annual shift in core-area use, I examined habitat characteristics on a nest-by-nest basis and averaged the results for each territory with >1 nest.

Qualitative Ranking of Territories

Direct estimates of some vital demographic variables, such as survival, are hard to obtain (Raphael et al. 1996) due to the difficulty in recapturing or resighting individuals and the complicating influence of factors such as immigration and emigration. As a

result, Vickery et al. (1992) suggested that indirect measures of fitness, such as nesting behavior, could be used as surrogates when direct measures of fitness were not available or reliable. The annual rate of egg laying, which is generally a function of prey availability (Newton 1989, Korpimäki 1990), and number of young that survive to reproduce (Franklin 1997) are examples of such measures. However, too few goshawk young (n = 18) have been found breeding to use the number of young that survive to reproduce as a measure of reproductive performance (Reynolds and Joy 1998). I, therefore, used the frequency of annual egg laying and production of fledglings (0-4) to infer territory quality. The annual egg-laying rate indicates the investment made by adults in raising young. Many pairs of goshawks initiate nest building during courtship, but fail to lay eggs. Because the failure to lay eggs results primarily from lack of food (Newton 1979, 1992), birds were assumed to lay eggs more often on higher quality territories where prey were assumed to be more readily available or abundant. Furthermore, preliminary work (Reynolds et al., in prep.) has shown that the probability of egg laying on the study area is inversely related to the amount of selective forest harvesting (e.g., shelterwood, seed-tree, and overstory removal) that has occurred within a 1.2 km radius of a territory center.

Because the number of young that survive to fledging is strongly correlated with the availability of supplemental food after hatching (Ward and Kennedy 1996), I expected the annual production of fledglings on higher quality territories to exceed that of lower quality territories due to differences in prey resources. Nest failures were included in measures of young fledged (i.e., number fledged = 0).

To derive comparative, summary measures of territory performance for the demographic variables, I used general linear models (GLM, SAS Inst. Inc. 2000). This approach accounted for annual variation in environmental factors affecting the population as a whole, such as weather or variable prey populations. The GLM procedure was also insensitive to missing values resulting from differences the total years a territory was under study and inter-annual variation in the annual frequency of egg laying on territories. The least-squares means option (LSMEANS, SAS Inst. Inc. 2000) provided relative estimates of each variable by territory.

To categorize territories by levels of goshawk reproductive performance (i.e., potential population sources and sinks), I used cluster analyses. Individual values j (j = 1, 2, ..., n) for each demographic variable mean i (i = 1, 2, ..., n) were standardized prior to clustering using the function (Romesburg 1984:78-79):

$$z_{ij} = \frac{x_{ij} - \bar{x}_i}{s_i}, \qquad [4.1]$$

where x_{ij} denotes each data value, $\bar{x}_i$ is the mean value for the variable i, and s_i is the standard deviation for variable i. To determine the appropriate number of clusters (seeds) to use in the final procedure, I clustered the data using Ward's minimum variance clustering method (PROC CLUSTER, SAS Inst., Inc. 2000), a seed-based hierarchical algorithm that allows exploration of the appropriate number of clusters to use in the final clustering procedure. Clusters were generated from sequential nearest neighbor comparisons of the Euclidean distance between all pairs of standardized scores. To produce the best, final cluster fit for each territory, I then used the K-means clustering algorithm (PROC FASTCLUS, SAS Inst., Inc. 2000). K-means is a nonhierarchical

clustering method that uses nearest centroid sorting to iteratively minimize the Euclidean distances from the cluster means. Cluster analyses were run on each demographic variable (frequency of egg laying, productivity) alone and together to find the model that distinguished best among the levels of goshawk territory quality. Multiresponse permutation procedures (MRPP; Mielke and Berry 2001) were used to test for distributional differences in the clusters at the $\alpha = 0.05$ level.

Goshawk Age and Experience

Some measures of reproductive performance in raptors may be strongly influenced by age at breeding or the fidelity of adults to territories, confounding the effects of habitat with age or experience (Newton 1991, 1989, Newton and Wyllie 1992, Koenig et al. 1992). To assess the potential influence of age and/or experience on the outcome of the cluster analyses of territory quality, I examined male and female recruitment rates, and the turnover and tenure rates of females on territories. Newton (1985) found that inexperienced Sparrowhawks (*Accipiter nisus*) tended to produce fewer young than experienced hawks, and that the proportion of these first-time breeders was greater on lower quality territories than on higher quality territories. I therefore tested for differences in the number of male and female recruits (known birds banded as fledglings that joined the breeding population) between the levels of territory quality using Fisher's Exact Tests (TABLES; SPSS, Inc. 2002), evaluated at the $\alpha = 0.05$ significance level. In the analysis, female and male recruitments were treated separately. A number of young (3rd or 4th year plumage) previously unbanded hawks also bred on the study area. Because of their young age (all known recruits were between 3-6 yrs of age when they first bred; R. T. Reynolds and S. M. Joy, unpubl. data) and lack of prior capture history

on known territories, these young hawks were assumed to be first-time breeders. I ran additional comparisons of territory quality using (1) these previously unbanded, young hawks and (2) the young hawks + recruits. Multiple incidences of recruitment on a territory (which only occurred on 3 territories) were treated as one recruitment incident in the analyses.

I evaluated two related measures of individual experience on territories that reflected the territory's history: female turnover and tenure rates. Newton and Wyllie (1992) found sparrowhawks that stayed on the same territory year after year had higher breeding success than birds that moved. Here, I used only females because their recapture/resight histories were more complete than those of males (males were often trap-shy or otherwise elusive). Female turnover was assessed as the number of times a female was replaced on a territory / the number of female recapture or resight opportunities on that territory. Females that returned to their original territories after a different female had bred there were treated as turnovers. Recapture or resight "opportunities" included only years in which a female was captured or resighted on the territory; whereas, failure to capture or resight a female, either due to non-breeding, nest failure, or trap/resight shyness, was not counted. Female tenure was assessed as the mean number of years each female remained on a territory / the number of female breeders on the territory. I tested each experience variable for no difference between the levels of territory quality using Mann-Whitney U Tests (NPAR; SPSS, In. 2002), evaluated at the $\alpha = 0.05$ significance level.

Newton (1991) reported that 70% of sparrowhawks that changed territories moved to territories classified as high or higher in occupancy rates and nesting success

than their previous territories. To determine whether movements between territories by nesting adults were related to territory quality, I assessed breeding dispersal annually for all banded hawks.

Habitat Modeling

One limitation to developing models of habitat associations with demographic performance includes the spatial resolution of the habitat maps. Because spatial resolution is typically sacrificed for generality in the mapping of large areas (Jensen 1986:186-187), fine-grain characterization of habitat composition and spatial structure across the landscape is difficult to obtain. For wide-ranging species such as raptors, detailed habitat characterizations are especially problematic. Therefore, forest composition was classified to a 10-m x 10-m spatial resolution (Joy et al. 2001, Chapter 1), finer than the spatial resolution of the Landsat TM imagery used to develop the classification (approx. 30 m x 30 m) to provide detailed landscape-level information for the study area. Classifications were based on field sampling that occurred in 1997 and the interpretation of Landsat TM imagery from the same time period. Vegetation composition comprised six dominant types (pinyon-juniper, ponderosa pine, mixed conifer, spruce-dominated mix and deciduous-dominated mix forests, and openings; see Study Area) modeled to a cross-validation (10-fold) accuracy of 74.5%. Because most forests are a composite of several, interspersed species, the Shannon-Weaver Index (H') (Lincoln et al. 1998), which measures the average degree of habitat diversity, was also used as a habitat variable. H' was calculated as:

$$H' = -\sum_{i=1}^{F} p_i \ln(P_i), \qquad [4.2]$$

where p_i is the proportion of an area in vegetation type i (i = 1, 2, ..., F). Small values of H' indicated lower diversity, while high values indicated a diverse forest composition.

Correlating Habitat Quality with Environmental Variables

To investigate whether goshawk territories with higher and lower reproductive performance are characterized by the composition or spatial arrangement of vegetative characteristics surrounding them, circle and ring plots, respectively, centered on the sample points were used. Because differences in territory quality may reflect vegetative characteristics resembling non-nest areas, random plots were also included in this assessment. Random plots ($\underline{n}$ = 85; Fig. 4.1) were generated using simple random sampling irrespective of territories. Habitat at random locations was treated as a third level of "habitat quality" in the analyses. To assess differences in the amount (proportion) and diversity of vegetative types across the levels of habitat qualities, I used five nested, concentric circles (0.07, 0.28, 1.13, 4.52, and 18.1 km^2 or 0.15, 0.30, 0.60, 1.20, and 2.4 km radii) from nest and random sampling plots (Fig. 4.2a). The spatial arrangement of habitat was assessed using five concentric, non-overlapping rings of increasing radius (0.07, 0.16, 0.35, 0.72, 1.48 km^2 or 0.15, 0.15-0.30, 0.30-0.60, 0.60-1.20, and 1.20-2.40 km radii) around the nest and random sampling plots (Fig. 4.2b), where the cumulative radii for rings were equal to the radii used for the circular plots. Similar sampling structures have been used in studies of northern spotted owl (*Strix occidentalis caurina*) (Ripple et al. 1997, Swindle et al. 1999, Thome et al. 1999) and goshawk (Finn et al. 2002) habitat. The proportions of habitat for random plots and nests occurring on the edge of the study area were estimated for the portion of the circle or ring that fell within the study area only. All data queries were performed in ArcView® (ESRI

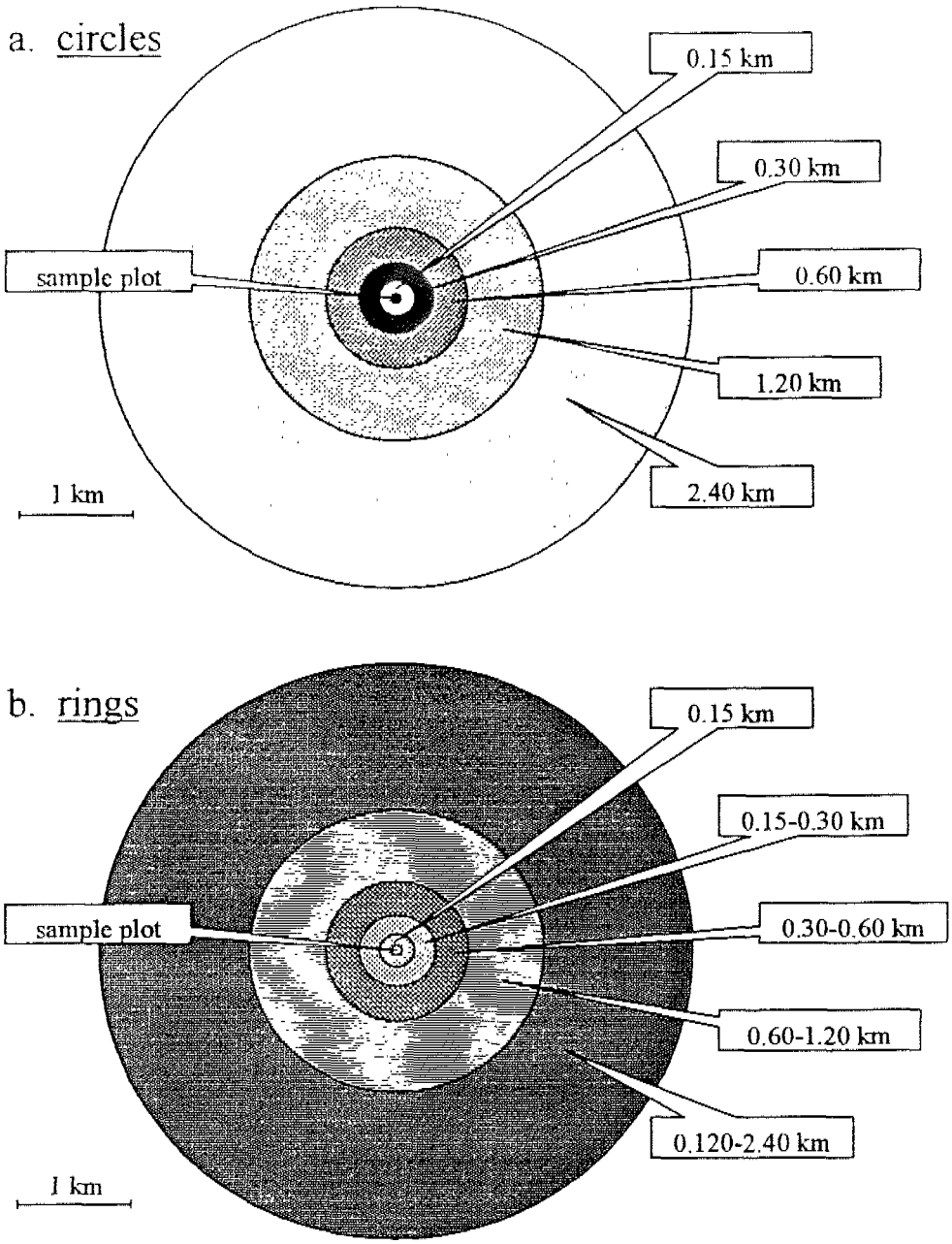


Figure 4.2. Configuration of circle (a) and ring (b) plots used to assess the amount and spatial arrangement of habitat components, respectively, around random plots (n = 85) and northern goshawk territories demonstrating higher (n = 56) and lower (n = 44) reproductive performance on the Kaibab National Forest (North Kaibab Ranger District) in Arizona.

1998) using a customized application. Values at multiple alternate nests were averaged to the territory. One territory was dropped from analyses because more than half of the habitat circles and rings fell outside the study area boundary.

Concentric distances incorporated three home range components thought to be biologically and behaviorally important to the goshawk's survival: the nest area (NA), post fledging-family area (PFA), and foraging area (FA) (Reynolds et al. 1992). The NA, an area of approximately 0.19 km radius (0.12 km^2) in size, is where nesting hawks focus their breeding season movements and behavior. The PFA, which extends approximately 0.57 km in radius (1.70 km^2) beyond the NA, is where fledglings spend the majority of time while still dependent on parents for food. The FA, which extends an additional 2.0 km in radius (21.9 km^2) beyond the NA and PFA, is where adults search for food during the nesting season. The PFA and FA are assumed to represent general locations for their respective activities and not an individual's use of a precise point within the home range; i.e., actual use of these areas may not follow a circular area.

Mixed-model analysis of variance (ANOVA) (PROC MIXED, LSMEANS; SAS Inst., Inc., 2000) was used to test for differences in the proportion of each vegetative type among levels of habitat quality (random plot, lower or higher quality nest plots) and circle or ring sizes, and their interactions. The sample plots within each level of habitat quality were the basic sampling unit used, and were assumed to be random samples of habitat quality. Levels of habitat quality and circle or ring sizes, as well as their interactions, were treated as fixed effects in the model. Prior to running the final models, each ANOVA was tested for spatially correlated residuals (PROC VARIOGRAM AND PROC INSIGHT; SAS Inst., Inc. 2000). Because territory plots represented the average

of several alternate nest sampling plots, territory locations used to examine spatial correlations were based on the mean coordinate of all alternate nests used within the territory weighted by the number of years each alternate nests was active. Measurements of habitat variables at the different radii were treated as repeated measures because circles and rings were nested and dependent (Baker et al. 1995). Where appropriate and to improve model fit [lowest AIC (Akaike 1977)], a covariate term (power or quadratic) that fit the functional form between the observed data and circle size or ring location was introduced. Standard ANOVA tests were evaluated at the $\alpha = 0.10$ significance level. Pairwise comparisons among levels of territory quality for significant variables were based on adjusted p-values [ADJUST=TUKEY (SAS Inst., Inc. 2000) for models without a covariate term and Bonferroni simultaneous probabilities (0.10/6 = 0.017, Miller 1981) for models with a covariate].

Because ANOVA tests measure central tendency and are sensitive to skewed data distributions, and hence violations of normality assumptions, I also evaluated differences in habitat quality for circle and ring plots using MRPP (Mielke and Berry 2001). Vegetative types were considered simultaneously and individually, the latter along with vegetative diversity. Circle and ring sizes were treated as repeated measures in these analyses. Because MRPP does not yield interactions easily, I tested individually for differences in the proportion of each vegetative type and diversity within each circle or ring size. For pairwise comparisons, I used the Peritz Closure Method (Petrondas and Ruben 1983) for assessing significance, which maintains Type I error rates at or below the specified α. All test (univariate, pairwise) were evaluated at the $\alpha = 0.05$ significance level. MRPP and the Peritz Closure Method were also used to compare the frequency

and size distribution within 0.15, 0.30, and 0.60-km radii around goshawk and random plots.

Habitat management recommendations (Reynolds et al. 1992) for the goshawk and its prey throughout the southwestern United States encourage no openings within the goshawk NA and small (0.008 km^2), interspersed openings within the PFA. I, therefore, examined the size and frequency of openings within the random plots, and higher and lower quality territories to determine whether their size distribution, cumulative area, and frequency differed among plots. Fifty percent of the random plots, and higher and lower quality territories were selected for this analysis using simple random sampling. Within each territory, one randomly selected nest plot was used. The GIS grid containing openings was clipped (CLIP, ESRI 1998) to 0.15-km, 0.30-km, and 0.60-km radii circle sizes and polygonized using ArcView®(ESRI 1998). Ring plots were not used in this assessment because they create false polygon boundaries. The number of polygons in the opening class, total area occupied by openings, and maximum, minimum and average size of the openings for the three circle radii were tested simultaneously for differences among higher and lower quality territories and random plots using MRPP (Mielke and Berry 2001). All response variables were standardized to a common unit (Mielke and Berry 2001:20). Univariate MRPP tests were also performed on each variable to assess its contribution to the differences.

Predicting the Location and Quality of Territories

To predict whether a point is likely to be a goshawk territory or random location, or a territory of higher or lower quality, I used both classification trees (CT) (Breiman et al. 1984, De'Ath and Fabricius 2000) and logistic regression models (Hosmer and

Lemeshow 2000). For classification trees, all proportional habitat variables were considered simultaneously. To avoid the unit sum constraint within the logistic regression models, I dropped the variable (proportion of deciduous-dominated mix) with the greatest spatial distribution on the study area and that contributed least to each model. This model also produced the lowest AIC. Whereas, advantages of using CT models include their straightforward use in predicting > 2 classes of data, ease of interpretation, robustness of results, and definitive nature of the classification criteria, CT models provide no estimates of uncertainty with each prediction. Alternatively, logistic regression models predict the probability of belonging to a class, as well as providing confidence limits associated with each prediction. However, logistic models are difficult to apply to > 2 classes of data, their application and interpretation may be more difficult than CT models due to the functional form of the logistic equation, and decisions regarding class membership may be more difficult due to borderline probabilities. Models with the highest cross-validated classification accuracies were chosen as the best predictors.

Three forms of the CT model (circles alone, rings alone, and circles + rings) (CART; Steinberg and Colla 1997) were used to predict differences among territories with higher and lower reproductive potential and random locations, and differences between territories of higher and lower quality alone. For comparisons with the logistic regression models, higher and lower quality territories were pooled in CT predictions of territory locations versus random plots for circles and rings. To avoid over-fitting the CT models, 10-fold cross-validation (Efron and Tibshirani 1993) was used to identify the tree

size that minimized the total deviance associated with the tree. Ten-fold cross-validation was also used to assess the accuracy of the CT models.

Logistic regression (PROC LOGISTIC; SAS Inst., Inc. 2000) was used to predict (1) territory locations (versus random locations) and (2) higher quality territories (versus lower) for circles and rings. Both the backward selection and best subset options were specified for the selection of significant variables in the regression models, where the number of variables producing the lowest AIC (Akaike 1977) was selected. The probability equation takes the form (Hosmer and Lemeshow 2000):

$$\pi(x) = \frac{e^{\beta_0+\beta_1 x_1+\beta_2 x_2+\ldots+\beta_k x_k}}{1+e^{\beta_0+\beta_1 x_1+\beta_2 x_2+\ldots+\beta_k x_k}}, \qquad [4.3]$$

where $\pi(x)$ is the probability of belonging to the reference class. Leave-one-out cross-validation (Efron and Tibshirani 1993) was used to estimate classification accuracies for the logistic models. Cross-validated prediction probabilities of ≥ 0.5 were assigned to the event predicted.

RESULTS

Population Sampling

Between 1991 and 2000, 454 active nests representing 101 goshawk territories were studied and a total of 771 young were produced on the study area. We captured and color banded 257 adult (136 females, 121 males) and 553 nestling (275 females, 277 males, 1 unknown sex) goshawks. Seventeen percent (77) of nests failed to produce young in one or more years. Both the annual frequency of egg-laying ($p < 0.001$, df = 100, Type III SS = 52.82, F = 3.28) and the number of young fledged ($p = 0.0260$, df = 100, Type III SS = 127.21, F = 1.35) differed among territories. The number of years a

territory was active ranged from 1–10, depending on annual frequency of egg laying and year the nest was found during the study. Relative, adjusted egg laying rates for territories (based on least squares means) ranged from 0.1 for territories that contained active nests in few of the years territories were monitored to 1.13 for territories containing active nests in all of the years monitored. Between zero and four young were produced annually (relative, adjusted productivity ranged from 0.40–3.32 young).

Qualitative Ranking of Territories

The standardized annual frequency of egg laying and productivity produced a simpler (i.e., with fewer intermediate points) two-axis cluster structure than either variable used alone for classifying territories of variable reproductive quality. The optimal number of clusters produced was two (Fig. 4.3), and they differed (P = 0.00) in distribution from one another. Fifty-six territories were classified as higher quality and 46 as lower quality (Fig. 4.4). Adjusted (least squares) means for the annual rate of egg laying was 0.79 (SD = 0.24) for higher and 0.49 (SD = 0.23) for lower quality territories. Mean productivity for higher and lower quality territories was 2.03 (SD = 0.43) and 1.22 (SD = 0.48), respectively.

Goshawk Age and Experience

A total of 33 banded recruits (20 females, 13 males) and 21 young, unbanded hawks (18 females, 13 males) joined the breeding population between 1991 and 2000. No differences were found (Table 4.1) in the number of recruits, young hawks, or recruits + young hawks between levels of territory quality. Likewise, females tended to stay on their territories for similar amounts of time ($F = 2.27$, $P = 0.14$) on higher ($n = 51$, $\bar{x} = 2.61$, SE = 0.19) and lower ($n = 41$, $\bar{x} = 1.73$, SE = 0.16) quality territories. A larger

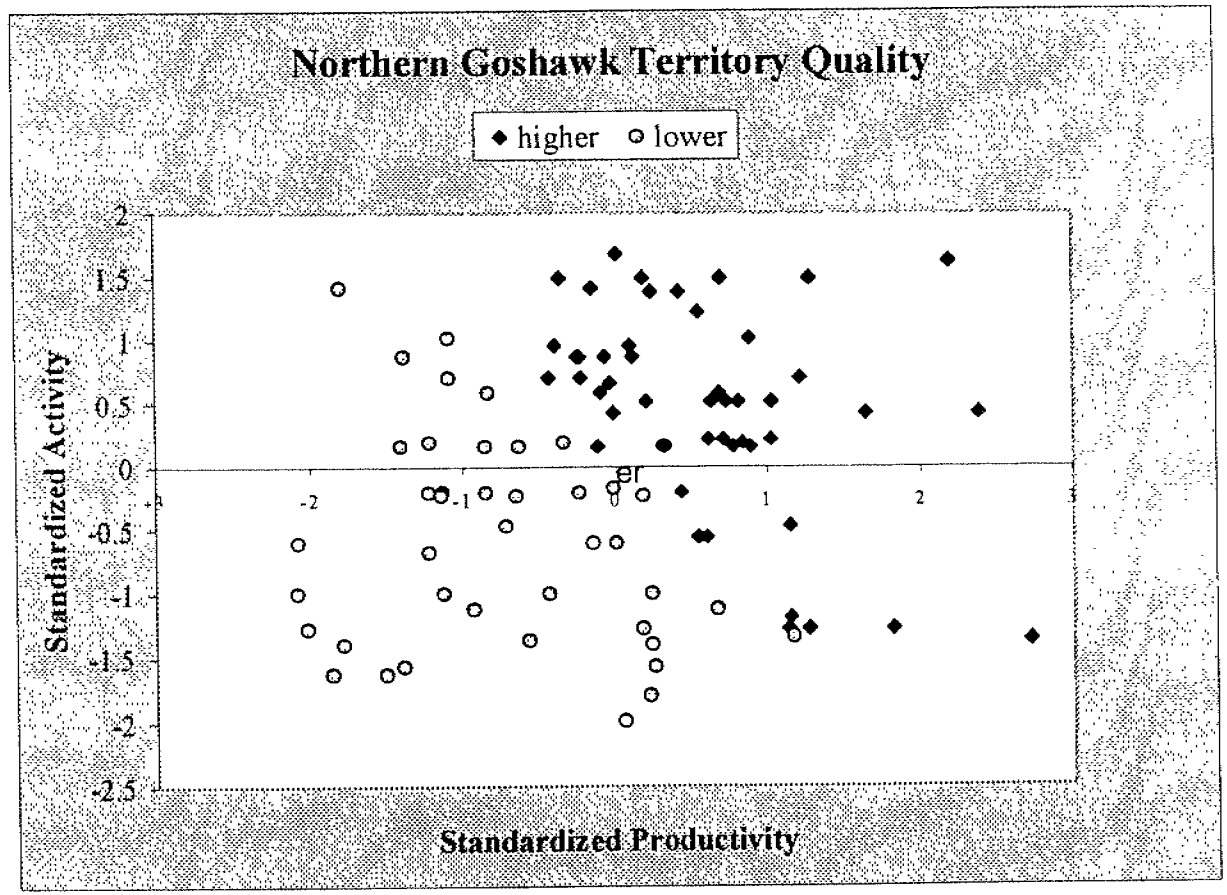


Figure 4.3. Clusters distinguishing ($P = 0.00$) higher from lower quality northern goshawk territories on Kaibab National Forest (North Kaibab Ranger District) in Arizona. Clusters were based on Euclidean distances between standardized least squares means for the annual rate of egg laying and number of young produced.

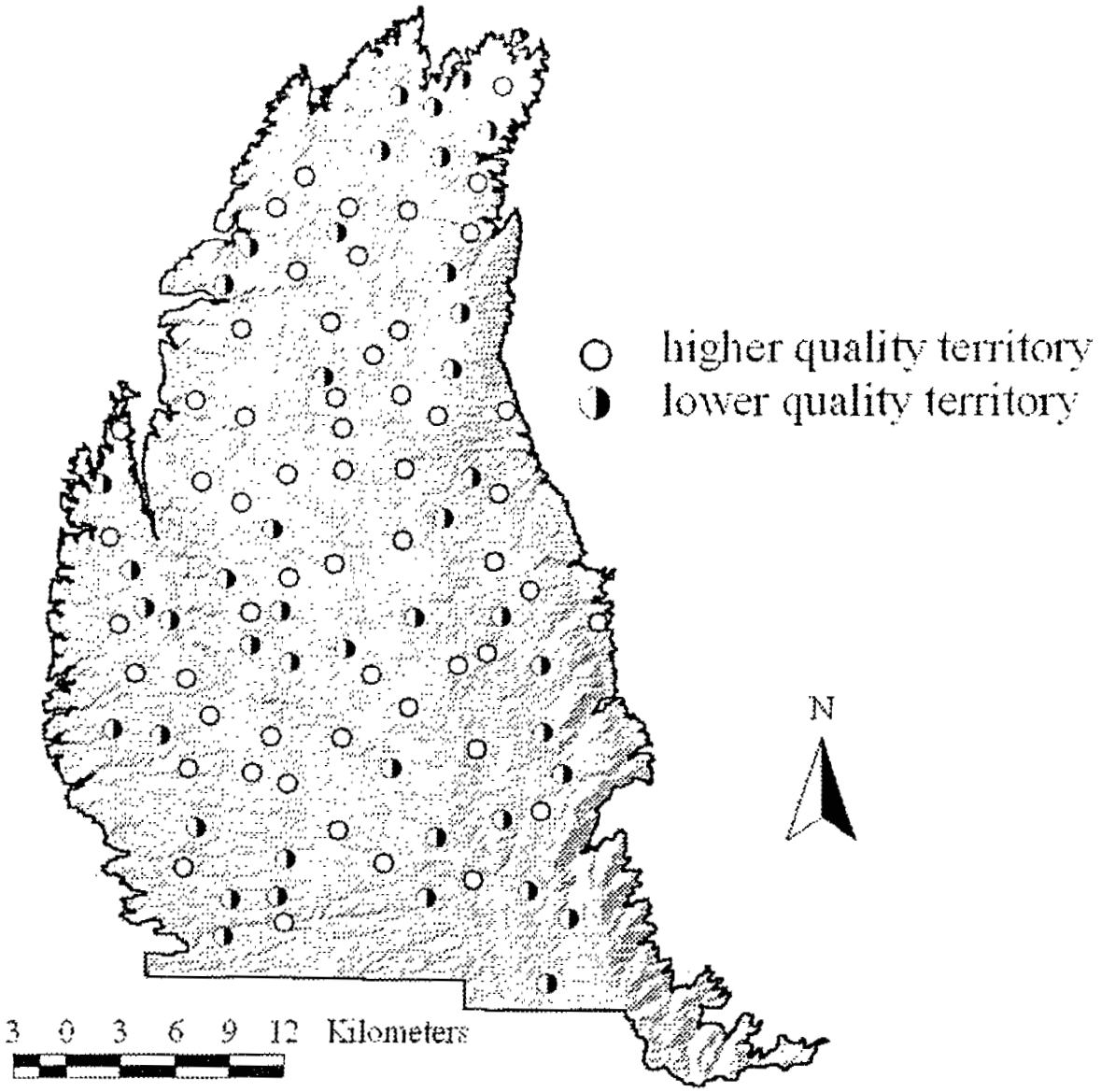


Figure 4.4. Distribution of higher (n = 56) and lower (n = 44) quality northern goshawk territories on the Kaibab National Forest (North Kaibab Ranger District), Arizona, based on cluster analysis of annual rates of egg laying and number of young fledged from nests. Territory locations are represented by the mean coordinate of all active nests between 1991 and 2000 used within a territory, weighted by the number of years each nests was active.

Table 4.1. Fisher's Exact Tests of the number of young northern goshawks (banded recruits and unbanded hawks) that joined the breeding population on 101 territories on the Kaibab National Forest (North Kaibab Ranger District), Arizona, evaluated between levels of territory quality (higher, lower).

Sex	Status	Territory Quality	Count	Fisher's Exact Test
Female	recruit	higher	12	P = 0.803
		lower	8	
	young	higher	12	P = 0.433
		lower	6	
	recruit + young	higher	22	P = 0.300
		lower	13	
Male	recruit	higher	8	P = 0.769
		lower	5	
	young	higher	9	P = 0.375
		lower	4	
	recruit + young	higher	17	P = 0.261
		lower	9	

number (F = 9.28, P = 0.00) of female turnovers, however, occurred on lower (n = 29, $\bar{x}$ = 1.11, SE = 0.10) than higher (n = 46, $\bar{x}$ = 0.60, SE = 0.05) quality territories.

Between the years 1991 and 2000, 19 breeding adult goshawks changed territories. Eight adults (42%; 5 females, 3 males) moved from higher quality territories to territories of similar quality; one adult (5%; female) moved from a higher to lower quality territory and then moved back to its original higher quality territory in the next year. Ten adults (53%; 6 females, 4 males) left lower quality territories; seven (70%; 5 females, 2 males) moved to higher quality territories, and three (30%; males) remained on territories of the same, lower quality.

Correlating Habitat Quality with Environmental Variables

Although use of covariates (power or quadratic terms) improved the performance of three (proportion of ponderosa pine and openings, and H' for circles) of the 14 (7 models for circles and 7 for rings) ANOVA models of habitat variables, they did not change the outcome of the models. For simplicity, therefore, all ANOVA model results are presented in their linear form (Tables 4.2, 4.3). Significant differences between random locations and territories of higher or lower quality (designated as the main effect "HQ") and their associated interactions [HQ × circle size (designated as "CS") or ring size (designated as "RS")] were of greatest interest in this study. None of the models had spatially correlated residuals.

The amount (Table 4.2, Fig. 4.5a-d) and spatial arrangement (Table 4.3, Fig. 4.6a-d) of four of the seven habitat variables differed between higher quality goshawk territories and random locations. The overall significance of tests for both circle and ring plots were similar. More ponderosa pine ($P \leq 0.07$), less deciduous-dominated vegetation

Table 4.2. Results of ANOVA tests for differences in the amount of vegetative types and diversity of vegetation around plots at northern goshawk nets and random locations on the Kaibab National Forest (North Kaibab Ranger District) in Arizona. Measurements of habitat variables within nested, concentric circles were treated as repeated measures. Asterisks indicate significance at the $\alpha = 0.10$ level.

Habitat component	Variable	DF	Denominator DF	F value	Pr > F
Pinyon-Juniper[1]	HQ[2]	2	182	0.45	0.5468
	CS[3]	4	728	7.99	<0.0001*
	HQ × CS	8	728	1.51	0.0529*
Ponderosa Pine[1]	HQ	2	182	2.51	0.0840*
	CS	4	728	7.20	<0.0001*
	HQ × CS	8	728	1.51	0.1482
Mixed Conifer[1]	HQ	2	182	0.44	0.6435
	CS	4	728	0.03	0.9987
	HQ × CS	8	728	4.64	<0.0001*
Spruce-dominated Mix[1]	HQ	2	182	1.14	0.3230
	CS	4	728	1.11	0.3526
	HQ × CS	8	728	2.54	0.0100*
Deciduous-dominated Mix[1]	HQ	2	182	6.15	0.0026*
	CS	4	728	1.40	0.2309
	HQ × CS	8	728	2.38	0.0155*

Table 4.2. (continued)

Openings[1]	HQ	2	182	9.57	0.0001*
	CS	4	728	3.48	0.0079*
	HQ × CS	8	728	4.02	0.0001*
Diversity (H')[4]	HQ	2	182	3.24	0.0415*
	CS	4	728	66.84	<0.0001*
	HQ × CS	8	728	1.11	0.3549

[1] Proportion

[2] Habitat Quality – Random location, higher and lower northern goshawk reproductive performance

[3] Circle Sizes 1-5 = 0.07, 0.28, 1.13, 4.52, and 18.1 km^2; radii = 0.15, 0.30, 0.60, 1.20, and 2.4 km

[4] Shannon-Weaver Index

Table 4.3. Results of ANOVA tests for differences in the spatial arrangement of vegetative types and diversity of vegetation around plots at northern goshawk nets and random locations on the Kaibab National Forest (North Kaibab Ranger District) in Arizona. Measurements of habitat variables within non-overlapping, concentric rings were treated as repeated measures. Asterisks indicate significance at the $\alpha = 0.10$ level.

Habitat component	Effect	DF	Denominator DF	F value	Pr > F
Pinyon-Juniper[1]	HQ[2]	2	182	0.45	0.6361
	RS[3]	4	728	7.99	<0.0001*
	HQ × RS	8	728	1.51	0.1484
Ponderosa Pine[1]	HQ	2	182	2.43	0.0909*
	RS	4	728	5.25	0.0004*
	HQ × RS	8	728	1.41	0.1905
Mixed Conifer[1]	HQ	2	182	0.29	0.7482
	RS	4	728	0.04	0.9973
	HQ × RS	8	728	4.05	0.0001*
Spruce-dominated Mix[1]	HQ	2	182	1.07	0.3468
	RS	4	728	1.34	0.2527
	HQ × RS	8	728	2.03	0.0410*
Deciduous-dominated Mix[1]	HQ	2	182	6.14	0.0026*
	RS	4	728	1.25	0.2885
	HQ × RS	8	728	2.06	0.0370*

Table 4.3. (continued)

Openings[1]	HQ	2	182	9.40	0.0001*
	RS	4	728	2.87	0.0225*
	HQ × RS	8	728	3.33	0.0009*
Diversity (*H'*)[4]	HQ	2	182	3.03	0.0508*
	RS	4	728	49.32	<0.0001*
	HQ × RS	8	728	0.98	0.4500*

[1] Proportion

[2] Habitat Quality – Random location, higher and lower northern goshawk reproductive performance

[3] Ring sizes 1-5 = 0.07, 0.16, 0.35, 0.72, 1.48 km^2; radii = 0.15, 0.15-0.30, 0.30-0.60, 0.60-1.20, and 1.2-2.40 km

[4] Shannon-Weaver Index

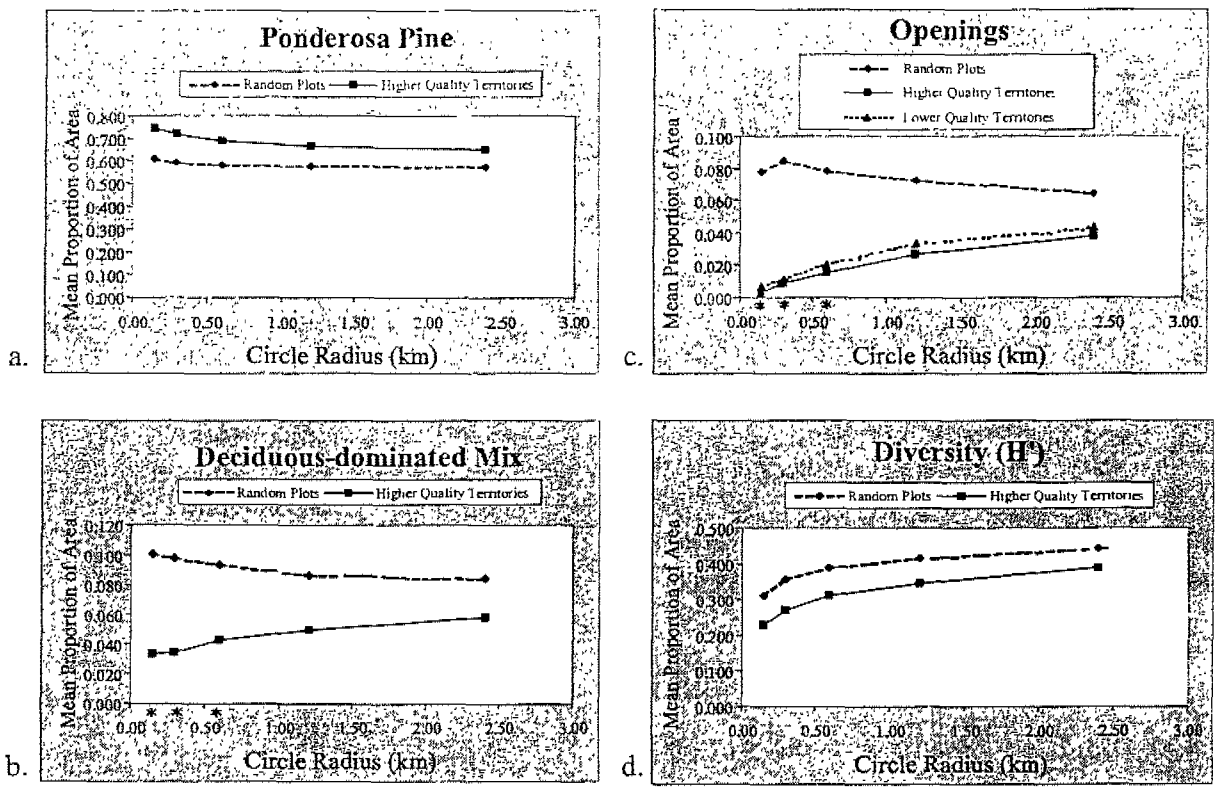


Figure 4.5. Significant relationships from analysis of variance tests of the amount (proportion) of habitat components around nests on northern goshawks territories of higher and lower quality and random plots on the Kaibab National Forest (North Kaibab Ranger District) in Arizona. Habitat proportions were measured at 5 spatial scales (0.15, 0.30, 0.60, 1.20, and 2.4 km radii) in concentric, overlapping circles around sample plots. Territory quality was based on the annual rate of egg laying and number of young produced at nests between 1991 and 2000. Asterisks indicate significant differences in landscape scales (distances) between main effects.

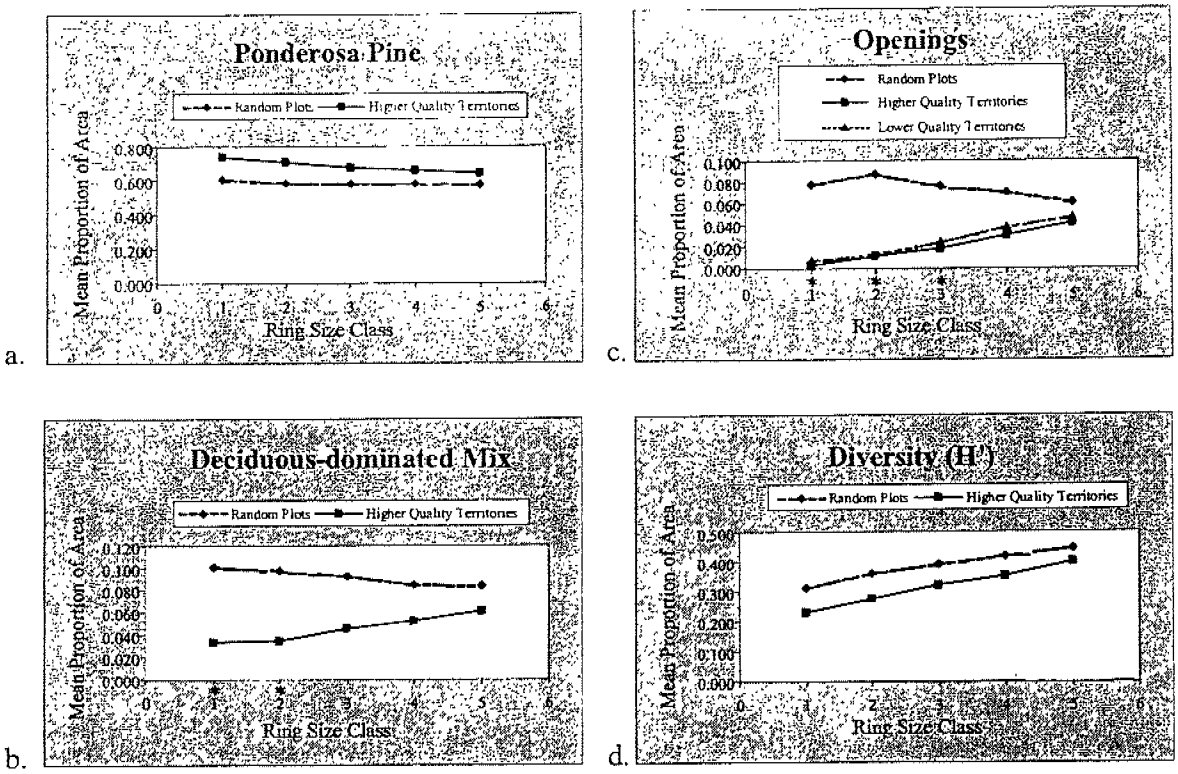


Figure 4.6. Significant relationships in analysis of variance tests of the spatial arrangement of habitat components around nests on northern goshawks territories of higher and lower quality and random plots on the Kaibab National Forest (North Kaibab Ranger District) in Arizona. Spatial arrangement was measured as the proportion of habitat at 5 spatial scales (1=0.15, 2=0.15-0.30, 3=0.30-0.60, 4=0.60-1.20, and 15=.20-2.40 km radii) in concentric, non-overlapping rings around sample plots. Territory quality was based on annual rates of egg laying and the number of young produced at nests between 1991 and 2000. Asterisks indicate significant differences in landscape scales (distances) between main effects.

(P = 0.00), less area in openings (P = 0.00), and lower diversity of vegetative types (P ≤ 0.04) distinguished higher quality territories from random plots (Table 4.4). A smaller proportion of area in openings (P = 0.00) also occurred on lower quality territories than on random plots (Table 4.4).

On higher quality territories, the amount (Table 4.5) of deciduous vegetation and forest openings increased with distance for both circle (Fig. 4.5b, c) and ring (Fig. 4.6b, c) plots; whereas, proportions of these habitats decreased with distance from random locations. Differences in these relationships were significant up to 0.60 km from the sample plots for deciduous vegetation (P < 0.06) and openings (P < 0.02). The proportion of openings at distances of up to 0.60 km was also important (P < 0.05) in distinguishing lower quality territories from random locations. As circle and ring size increased, the difference between main effects in the models became smaller due to overlap in the measurement of habitat variables (Table 4.5). ANOVA detected no difference (P > 0.10) between territories with lower and higher reproductive success or between lower quality territories and random plots (Table 4.4), except that noted for openings.

MRPP tests for distributional differences in the mean proportions of vegetative types (considered simultaneously) differed (P = 0.01) between higher quality territories and random locations for circle and ring plots (Table 4.6). Results of univariate tests of vegetative type, which were consistent between circle and ring plots (Tables 4.4, 4.6), showed that less deciduous-dominated vegetation and smaller proportions of openings (P = 0.00) were present on higher quality territories than random locations. Less area (P = 0.00) in openings was also present on low quality goshawk territories than on random

Table 4.4. Proportions of vegetative types and diversity for higher and lower quality northern goshawk territories and random locations on the Kaibab National Forest (North Kaibab Ranger District) in Arizona. Values represent least square means and standard error estimates for concentric, overlapping circles and non-overlapping rings, representing the amount and spatial arrangement of vegetative types around sample plots, respectively. Asterisks (*, **) indicate significance at the $\alpha = 0.10$ level.

	Habitat Quality											
	Circle Plots[1]						Ring Plots[2]					
	Goshawk Territories[3]				Random		Goshawk Territories				Random	
	Higher		Lower		Plots		Higher		Lower		Plots	
	(n = 56)		(n = 44)		(n = 85)		(n = 56)		(n = 44)		(n = 85)	
Habitat Component	$\bar{x}$	SE	$\bar{x}$	SE	$\bar{x}$	SE	$\bar{x}$	SE	$\bar{x}$	SE	$\bar{x}$	SE
Pinyon-Juniper[4]	0.053	0.013	0.043	0.015	0.063	0.010	0.056	0.012	0.047	0.014	0.063	0.010
Ponderosa Pine[4]	0.694*	0.037	0.630	0.041	0.588*	0.030	0.688*	0.036	0.627	0.040	0.586*	0.029
Mixed Conifer[4]	0.112	0.017	0.121	0.020	0.099	0.014	0.111	0.017	0.119	0.019	0.102	0.014
Spruce-Dominated Mix[4]	0.080	0.019	0.119	0.022	0.082	0.016	0.079	0.019	0.117	0.022	0.083	0.015

Table 4.4 (continued)

Deciduous-Dominated Mix[4]	0.043*	0.011	0.063	0.013	0.093*	0.009	0.045*	0.010	0.065	0.012	0.092*	0.009
Openings[4]	0.019*	0.011	0.023**	0.013	0.076*,**	0.009	0.021*	0.011	0.026**	0.012	0.075*,**	0.009
Diversity (H')[5]	0.309*	0.023	0.356	0.026	0.384*	0.018	0.316*	0.022	0.360	0.025	0.386*	0.018

[1] Circle radii = 0.15, 0.30, 0.60, 1.20, and 2.4 km; areas = 0.07, 0.28, 1.13, 4.52, and 18.1 km^2

[2] Ring radii = 0.15, 0.15-0.30, 0.30-0.60, 0.60-1.20, and 1.20-2.40 km; areas = 07, 0.16, 0.35, 0.72, 1.48 km^2

[3] Means of > 1 nest per territory

[4] Proportion

[5] Shannon-Weaver Index

Table 4.5. Proportion of vegetative types and diversity for higher and lower quality northern goshawk territories and random locations by plot (circle, ring) size on the Kaibab National Forest (North Kaibab Ranger District) in Arizona. Values represent least square means and standard error estimates for each concentric, overlapping circle and non-overlapping ring size, representing the amount and spatial arrangement of vegetative types around sample plots, respectively. Asterisks (*, **) indicate significance at the $\alpha = 0.017$ level.

	Habitat Quality – Size 1[1, 2]											
	Circle Plots						Ring Plots					
	Goshawk Territories[3]				Random		Goshawk Territories				Random	
	Higher		Lower		Plots		Higher		Lower		Plots	
	(n = 56)		(n = 44)		(n = 85)		(n = 56)		(n = 44)		(n = 85)	
Habitat Component	$\bar{x}$	SE	$\bar{x}$	SE	$\bar{x}$	SE	$\bar{x}$	SE	$\bar{x}$	SE	$\bar{x}$	SE
Pinyon-Juniper[4]	0.023	0.014	0.016	0.016	0.057	0.011	0.023	0.014	0.016	0.016	0.057	0.011
Ponderosa Pine[4]	0.743	0.038	0.646	0.043	0.609	0.031	0.743	0.037	0.646	0.042	0.609	0.030
Mixed Conifer[4]	0.110	0.018	0.144	0.020	0.079	0.015	0.110	0.018	0.144	0.020	0.079	0.015
Spruce-Dominated Mix[4]	0.087	0.020	0.130	0.022	0.077	0.016	0.087	0.020	0.130	0.022	0.077	0.016

Table 4.5 (continued)

Deciduous-Dominated Mix[4]	0.033*	0.012	0.057	0.013	0.101*	0.010	0.033*	0.012	0.057	0.013	0.101*	0.010
Openings[4]	0.003*	0.012	0.007**	0.014	0.078*,**	0.010	0.003*	0.012	0.007**	0.014	0.078*,**	0.010
Diversity (H')[5]	0.231	0.024	0.295	0.027	0.312	0.019	0.231	0.024	0.295	0.027	0.312	0.019
	Habitat Quality – Size 2[1, 2]											
Habitat Component	$\bar{x}$	SE	$\bar{x}$	SE	$\bar{x}$	SE	$\bar{x}$	SE	$\bar{x}$	SE	$\bar{x}$	SE
Pinyon-Juniper[4]	0.040	0.014	0.032	0.016	0.061	0.011	0.045	0.014	0.038	0.016	0.063	0.011
Ponderosa Pine[4]	0.718	0.038	0.654	0.043	0.592	0.031	0.710	0.037	0.656	0.042	0.586	0.030
Mixed Conifer[4]	0.118	0.018	0.122	0.020	0.092	0.015	0.120	0.018	0.115	0.020	0.096	0.015
Spruce-Dominated Mix[4]	0.082	0.020	0.126	0.022	0.072	0.016	0.080	0.020	0.125	0.022	0.071	0.016
Deciduous-Dominated Mix[4]	0.034*	0.012	0.055	0.013	0.098*	0.010	0.034*	0.012	0.054	0.013	0.098*	0.010
Openings[4]	0.009*	0.012	0.011**	0.014	0.085*,**	0.010	0.011*	0.012	0.013**	0.014	0.087*,**	0.010
Diversity (H')[5]	0.270	0.024	0.321	0.027	0.358	0.019	0.277	0.024	0.322	0.027	0.362	0.019
	Habitat Quality – Size 3[1, 2]											
Habitat Component	$\bar{x}$	SE	$\bar{x}$	SE	$\bar{x}$	SE	$\bar{x}$	SE	$\bar{x}$	SE	$\bar{x}$	SE

Table 4.5 (continued)

Pinyon-Juniper[4]	0.055	0.014	0.048	0.016	0.064	0.011	0.061	0.014	0.053	0.016	0.064	0.011
Ponderosa Pine[4]	0.689	0.038	0.635	0.043	0.584	0.031	0.679	0.037	0.629	0.042	0.582	0.030
Mixed Conifer[4]	0.117	0.018	0.114	0.020	0.102	0.015	0.117	0.018	0.111	0.020	0.105	0.015
Spruce-Dominated Mix[4]	0.081	0.020	0.124	0.022	0.078	0.016	0.081	0.020	0.123	0.022	0.080	0.016
Deciduous-Dominated Mix[4]	0.042*	0.012	0.059	0.013	0.094*	0.010	0.045*	0.012	0.061	0.013	0.092*	0.010
Openings[4]	0.016*	0.012	0.021**	0.014	0.079*,**	0.010	0.018	0.012	0.024	0.014	0.076	0.010
Diversity (H')[5]	0.311	0.024	0.352	0.027	0.389	0.019	0.320	0.024	0.357	0.027	0.392	0.019

	Habitat Quality – Size 4[1,2]											
Habitat Component	$\bar{x}$	SE	$\bar{x}$	SE	$\bar{x}$	SE	$\bar{x}$	SE	$\bar{x}$	SE	$\bar{x}$	SE
Pinyon-Juniper[4]	0.070	0.014	0.056	0.016	0.066	0.011	0.075	0.014	0.058	0.016	0.067	0.011
Ponderosa Pine[4]	0.668	0.038	0.616	0.043	0.580	0.031	0.661	0.037	0.610	0.042	0.578	0.030
Mixed Conifer[4]	0.108	0.018	0.114	0.020	0.109	0.015	0.105	0.018	0.115	0.020	0.111	0.015
Spruce-Dominated Mix[4]	0.077	0.020	0.110	0.022	0.086	0.016	0.076	0.020	0.105	0.022	0.089	0.016

Table 4.5 (continued)

Deciduous-Dominated Mix[4]	0.050	0.012	0.070	0.013	0.087	0.010	0.052	0.012	0.074	0.013	0.084	0.010
Openings[4]	0.027	0.012	0.034	0.014	0.072	0.010	0.030	0.012	0.039	0.014	0.070	0.010
Diversity (H')[5]	0.345	0.024	0.388	0.027	0.416	0.019	0.353	0.024	0.395	0.027	0.419	0.019
	Habitat Quality – Size 5[1, 2]											
Habitat Component	$\bar{x}$	SE	$\bar{x}$	SE	$\bar{x}$	SE	$\bar{x}$	SE	$\bar{x}$	SE	$\bar{x}$	SE
Pinyon-Juniper[4]	0.075	0.014	0.065	0.016	0.065	0.011	0.077	0.014	0.069	0.016	0.065	0.011
Ponderosa Pine[4]	0.651	0.038	0.599	0.043	0.576	0.031	0.645	0.037	0.593	0.042	0.575	0.030
Mixed Conifer[4]	0.105	0.018	0.111	0.020	0.115	0.015	0.104	0.018	0.110	0.020	0.117	0.015
Spruce-Dominated Mix[4]	0.073	0.020	0.105	0.022	0.095	0.016	0.071	0.020	0.103	0.022	0.098	0.016
Deciduous-Dominated Mix[4]	0.058	0.012	0.076	0.013	0.085	0.010	0.061	0.012	0.078	0.013	0.084	0.010
Openings[4]	0.038	0.012	0.044	0.014	0.065	0.010	0.042	0.012	0.047	0.014	0.061	0.010
Diversity (H')[5]	0.389	0.024	0.424	0.027	0.442	0.019	0.400	0.024	0.431	0.027	0.444	0.019

[1] Circle sizes 1-5 correspond to radii = 0.15, 0.30, 0.60, 1.20, and 2.4 km, and areas = 0.07, 0.28, 1.13, 4.52, and 18.1 km^2

Table 4.5 (continued)

[2] Ring sizes 1-5 correspond to radii = 0.15, 0.15-0.30, 0.30-0.60, 0.60-1.20, and 1.20-2.40 km, and areas = 07, 0.16, 0.35, 0.72, 1.48 km^2

[3] Means of > 1 nest per territory

[4] Proportion

[5] Shannon-Weaver Index

Table 4.6. Results of MRPP tests of the distribution in the proportion of vegetative types and vegetative diversity for higher and lower quality goshawk territories and random locations for all circle or ring plot sizes considered simultaneously, on the Kaibab National Forest (North Kaibab Ranger District), Arizona. Where significant effects in habitat quality occur, results of pairwise tests of the habitat quality are provided. Asterisks indicate significance at the $\alpha = 0.05$ level. The Peritz Closure Method (Petrondas and Ruben 1983) was used for pairwise comparisons.

Habitat Component	Habitat Quality[1]	Circle Plots[2] P-value	Ring Plots[3] P-value
All (except Diversity)	All	0.0.29*	0.035*
	R vs. H	0.012*	0.014*
	R vs. L	0.137	0.149
	H vs. L	0.258	0.278
Pinyon-Juniper[4]	All	0.564	0.630
Ponderosa Pine[4]	All	0.074	0.084
Mixed Conifer[4]	All	0.229	0.258
Spruce-dominated Mix[4]	All	0.392	0.400
Deciduous-dominated Mix[4]	All	0.000*	0.001*
	R vs. H	0.000*	0.000*
	R vs. L	0.083	0.086
	H vs. L	0.078	0.087
Openings[4]	All	0.000*	0.000*
	R vs. H	0.000*	0.000*

Table 4.6. (continued)

Openings[4]	R vs. L	0.000*	0.000*
	H vs. L	1.000	1.000
Diversity (H')[5]	All	0.051	0.066

[1] Levels of habitat quality = random (R) location, higher (H) and lower (L) goshawk reproductive performance

[2] Circle radii = 0.15, 0.30, 0.60, 1.20, and 2.4 km; areas = 0.07, 0.28, 1.13, 4.52, and 18.1 km^2

[3] Ring radii = 0.15, 0.15-0.30, 0.30-0.60, 0.60-1.20, and 1.20-2.40 km; areas = 07, 0.16, 0.35, 0.72, 1.48 km^2

[4] Proportion

[5] Shannon-Weaver Index

plots. These results differed from the ANOVA results in which four vegetative types were found to produce significant relationships. However, when MRPP tests were run on the individual circle or ring sizes for each vegetative type and diversity (Table 4.7), results similar to those of the ANOVA emerged: more ponderosa pine (up to a 0.30-km radius), less deciduous-dominates mixes (up to a 1.20-km radius), fewer openings (up to a 1.20-km radius), and lower vegetative diversity (0.30-km radius for circles, 0.15-km - 0.30-km radius for rings) occurred on high quality goshawk territories than on random locations for both circle and ring plots. More mixed conifer (up to a 0.15-km radius) and less deciduous-dominated habitat (0.30-km radius for circles, 0.15 km - 0.30-km radius for rings), and fewer openings (up to a 0.60-km radius) also distinguished ($P > 0.05$) low quality territories from random locations for both circle and ring plots.

Openings modeled on the study area (Chapter 1) ranged in size from 100 m^2 to 2.15 km^2. Within a 0.15-km radius from nests, openings on both higher and lower quality territories ranged from an average of 0.0% to 6.3% of the total area; whereas, these proportions ranged from 0.0% to 32% for random plots. Within 0.30-km radius, the ranges were 0.0-8.8% and 0.0-48.8% for high/low and random plots, respectively. Within a 0.60-km radius, the sizes of openings ranged up to 10.8% of the total area for nest plots and up to 41.5% for random plots. Openings in random plots were more numerous and larger, occupied more total area, and were greater in maximum, minimum, and average size than on either higher or lower quality territories ($P < 0.05$ for all tests). Results of the tests for the individual response variables did not differ.

In contrast to the ANOVA results, differences found using MRPP (Table 4.7) between goshawk territories and random locations included a greater number of

Table 4.7. Results of univariate MRPP tests of the distribution in the proportion of vegetative types and vegetative diversity for higher and lower quality goshawk territories and random locations for each circle and ring plot size on the Kaibab National Forest (North Kaibab Ranger District), Arizona. Where significant effects in habitat quality (HQ) occur, results of significant pairwise tests of habitat quality are provided. Asterisks indicate significance at the $\alpha = 0.05$ level. The Peritz Closure Method (Petrondas and Ruben 1983) was used for pairwise comparisons.

Habitat Component	Plot Sizes[1, 2]	Habitat Quality[3]	Circle Plot P-value	Ring Plot P-value
Pinyon-Juniper[4]	1-5	—	>0.050	>0.050
Ponderosa Pine[4]	1		0.039*	0.039*
		R vs. H	0.011*	0.011*
	2		0.038*	0.042*
		R vs. H	0.012*	0.014*
	3-5	—	>0.050	>0.050
Mixed Conifer[4]	1		0.040*	0.040*
		R vs. L	0.010*	0.010*
	2-5	—	>0.050	>0.050
Spruce-dominated Mix[4]	1-5	—	>0.050	>0.050
Deciduous-dominated Mix[4]	1		0.001*	0.001*
		R vs. H	0.000*	0.000*
	2		0.000*	0.000*

Table 4.7. (continued)

Deciduous-dominated Mix[4]	2	R vs. H	0.000*	0.000*
		R vs. L	0.030*	0.026*
	3		0.001*	0.004*
		R vs. H	0.000*	0.001*
	4		0.013*	0.037*
		R vs. H	0.003*	0.010*
	5	—	>0.050	>0.050
Openings[4]	1-2		0.000*	0.000*
		R vs. H	0.000*	0.000*
		R vs. L	0.000*	0.000*
	3		0.000*	0.000*
		R vs. H	0.000*	0.000*
		R vs. L	0.001*	0.002*
	4		0.000*	0.009*
		R vs. H	0.011*	0.004*
	5	—	0.065	0.263
Diversity (H')[5]	1	—	>0.050	>0.050
	2		0.025*	0.032*
		R vs. H	0.006*	0.008*
	3-5	—	>0.050	>0.050

[1] Circle sizes 1-5 correspond to radii = 0.15, 0.30, 0.60, 1.20, and 2.4 km, and areas = 0.07, 0.28, 1.13, 4.52, and 18.1 km^2

Table 4.7 (continued)

[2] Ring sizes 1-5 correspond to radii = 0.15, 0.15-0.30, 0.30-0.60, 0.60-1.20, and 1.20-2.40 km, and areas = 07, 0.16, 0.35, 0.72, 1.48 km^2

[3] Levels of habitat quality = random (R) location, higher (H) and lower (L) goshawk reproductive performance

[4] Proportion

[5] Shannon-Weaver Index

significant relationships at specific distances, as well as farther distances from the sample plot. Nonetheless, relationships were most common within 0.30-km and 0.60-km radii of the sample plots.

Predicting the Location and Quality of Territories

The three-level (random/higher/lower) CT did not perform as well as the two-level (random/higher or random/lower) models at predicting territory location or quality. Overall, however, logistic regression (Tables 4.8, 4.9) produced higher cross-validation classification accuracies than either the two- or three-level CTs in determining whether a sample plot falls within a territory versus a random location, or in a higher versus lower quality territory. Model outcomes did not differ between the backward elimination or best subset selection procedures for significant variables. The models for circles and rings performed comparably; however, misclassification (off-diagonal) rates were slightly higher for the circle data; therefore, only models based on the ring data are discussed.

The highest prediction accuracy (85%; Table 4.8) for the probability that a point on the study area belongs to a goshawk territory (versus a random location) is associated with the proportions of mixed conifer habitat and openings within 0.15 km, ponderosa pine and spruce-dominated mixes between 0.15-0.30 km, pinyon-juniper between 0.30-0.60 km, and spruce-dominated mixes between 1.20-2.40 km of the sample plot. The proportion of openings within 0.15 km of the plot had the strongest negative influence on the prediction; whereas, the proportion of spruce-dominated mixes between 0.15-km - 0.30-km radii had the strongest positive influence. The model resulted in 15% of territories being misclassified as random plots. The probability of a sample plot

Table 4.8. Significant vegetative coefficients for predicting the location of northern goshawk territories vs. random sites on the Kaibab National Forest (North Kaibab Ranger District) in Arizona fitted by logistic regression (cross-validation accuracy = 85%). The model was run at 5 spatial scales (concentric, non-overlapping rings of radii 0.15, 0.15-0.30, 0.30-0.60, 0.60-1.20, and 1.20-2.40 km) around the sample plots.

Variable	Scale (radius, km)	Parameter Estimate	SE	Wald χ^2	*P*-value	Lower 95% Confidence Limit	Upper 95% Confidence Limit
Intercept		-6.15	1.91	10.34	0.0013	-10.16	-2.64
Mixed Conifer[1]	0.00-0.15	9.28	2.64	12.29	0.0005	4.50	14.92
Opening[1]	0.00-0.15	-20.04	8.71	5.29	0.0214	-40.08	-5.06
Ponderosa Pine[1]	0.15-0.30	6.81	1.97	11.93	0.0006	3.19	10.94
Spruce-dominated Mix[1]	0.15-0.30	19.71	4.84	16.56	<0.0001	11.10	30.08
Pinyon-Juniper[1]	0.30-0.60	9.42	2.91	10.50	0.0012	4.05	15.51
Spruce-dominated Mix[1]	1.20-2.40	-11.87	4.62	6.59	0.0102	-21.62	-3.46

[1] Proportion

Table 4.9. Significant habitat coefficients for predicting higher (vs. lower) quality nests for northern goshawks on the Kaibab National Forest (North Kaibab Ranger District) in Arizona fitted by logistic regression (overall cross-validation accuracy = 82%). The model was run at 5 spatial scales (concentric, non-overlapping rings of radius 0.15, 0.15-0.30, 0.30-0.60, 0.60-1.20, and 1.20-2.40 km around the sample plots).

Variable	Scale (radius, km)	Parameter Estimate	SE	Wald χ^2	*P*-value	Lower 95% Confidence Limit	Upper 95% Confidence Limit
Intercept		-2.05	1.20	2.94	0.0863	-4.60	0.16
Ponderosa Pine[1]	0.00-0.15	2.60	1.29	4.09	0.0432	0.22	5.33
Mixed-Conifer[1]	0.00-0.15	-9.58	4.14	5.35	0.0207	-18.58	-2.076
Mixed-Conifer[1]	0.15-0.30	14.50	5.16	7.89	0.0050	5.46	25.96

[1] Proportion

belonging to a random location can be predicted with a cross-validation classification accuracy of 66% (34% misclassification rate).

Three variables (Table 4.9) were useful in predicting whether a new nest on a territory was likely to contain eggs in a given year and to produce the greatest number of young: the proportions of ponderosa pine and mixed-conifer habitats within 0.15 km of the plot and mixed conifer between 0.15-0.30 km of the plot. The composition of habitat close to (within a 0.30-km radius of) the nest thus appeared to be of great importance in predicting site quality based on reproductive effort. Of these three variables, only the proportion of mixed conifer vegetation within 0.15 km of the nest plot was a negative predictor in the model. Alternatively, the proportion of mixed-conifer habitat between 0.15 and 0.30 km had the greatest positive influence on predicting higher territory quality. Although the cross-validation accuracy for predicting a higher quality goshawk location was high (82%) and the misclassification rate was low (18%), the model did not perform well at predicting lower quality goshawk territories (accuracy = 34%, misclassification = 66%).

DISCUSSION AND CONCLUSIONS

Differences in the reproductive quality of goshawk territories (based on the frequency of egg laying and number of young produced) were described best using two levels of quality (higher and lower). Neither the age of the breeder (male or female) nor the length of time a female stayed on a territory explained the differences in territory quality. More females, however, were replaced on lower quality territories than on higher quality territories, suggesting that a female's experience on a territory, especially

in the first year of breeding, may influence reproductive success. Newton and Wyllie (1992) found European sparrowhawks that stayed on the same territory fledged young more often than those that moved. Resident females are more likely to have a greater knowledge of nest habitat, prey availability, and sources of predation than first-time breeders on a territory. The degree to which a female's experience affects the territory's qualitative ranking in this study is unclear, however, because the number of young produced was only one of two variables on which territory quality was based. The second reproductive variable, frequency of egg laying, is influenced more by prey availability than the female's experience on a territory, and I would expect both resident and first-time female breeders to lay eggs at similar frequencies on their territories.

When presented with the opportunity to move (generally due to the non-return of a mate; Reynolds et al., in review), goshawks on lower quality territories moved to higher quality territories, while goshawks on higher quality territories generally maintained territory quality after moving. Although these data are anecdotal, they imply that differences in territory quality exist and that goshawks will improve their likelihood for long-term reproductive success when possible. Newton (1991) found that in south Scotland the majority (70%) of breeding Sparrowhawks also moved to a nest site of the same or higher quality than their previous nest site.

The most common habitat variables that distinguished higher and lower quality goshawk territories from random locations for both the amount and spatial arrangement of vegetative types included the proportions ponderosa pine, mixed conifer, and deciduous-dominated habitat, openings, and the overall diversity of vegetative types. Significant relationships did not vary greatly between circle and ring analyses, indicating

that either the amount (circles) or arrangement (rings) of vegetative types and diversity could be used on the study area to distinguish between random locations and nests with higher or lower reproductive success. Tests of both central tendency (ANOVA) and data dispersion (MRPP) were useful in identifying important effects.

The majority of significant relationships within vegetative types occurred between higher quality territories and random locations. In particular, the amount of deciduous-dominated vegetation (less) and openings (fewer) within a 0.60-km (ANOVA) and 1.2-km (MRPP) radius of the sample plots were important. These distances incorporated the NA, PFA, and 11% of the FA (Reynolds et al. 1992) within a goshawk's home range.

The NA provides critical habitat for nests, and may include more than one nest. The PFA, in addition to providing a staging ground for young goshawks to learn to hunt while still receiving food from their parents, appears to correspond to the defended area around nests (i.e., the territory). The PFA also provides food and cover for a number of the goshawk's prey species (Reynolds et al. 1992).

Greater proportions of ponderosa pine and lower diversity of vegetative types distinguish habitat in higher quality territories from random locations. These relationships were particularly important near (within 0.30 km) the sample plots. Although ponderosa pine comprised 55% of the study area, 70% of nests on territories were found in this forest class (Chapter 1). Ponderosa pine trees provide large, protective crowns with easy access to nests, as well as habitat for and access to the goshawk's prey. Historically, forests of ponderosa pine were maintained by frequent (2-4 yr), low-intensity surface fires that resulted in same age clumps (groups) of trees separated by small openings and maintained relatively open understories. Fire suppression and

selective tree cutting during the past century have allowed pine and fir regeneration to fill the below-canopy space and inter-cluster gaps, thereby changing the structure and species diversity of the pine forest. The importance of lower species diversity near nest in higher quality territories may be a function of selection for greater proportions of "pure" ponderosa pine, containing less fir regeneration.

Less deciduous-dominated habitat was found at higher and lower quality territories than at random plots. Although goshawks will nest in aspen trees, aspen crowns provide less protection from predators and greater exposure to inclement weather. Aspen stands also provide less concealment while hunting, but are good trees for cavity excavation by hole-nesting prey such as woodpeckers.

Although goshawks occasionally forage along forest openings using longer flight times, their hunting behavior typically employs a short-sit-and-wait and short flight pattern, where the hawks use multiple perches under the forest canopy to search for prey and travel times vary with forest type (Kenward 1982, Widén 1984). In addition to providing less suitable habitat for the goshawk's dominant prey [Kaibab and red squirrels (*Sciurus* spp.)], forest openings may diminish the concealment of nests. The number and size of forest openings within a goshawk's territory and foraging range are therefore important to the goshawk's reproductive success. Fewer and smaller openings, as well as less total area in openings, on higher and lower quality territories than on random plots illustrate this point.

Significant trends between random locations and higher or lower quality goshawk territories may reflect historical forest management practices on the study area. Goshawks there exhibited strong territory fidelity (Reynolds et al., in review), where only

5% of males and 6% of females dispersed to new breeding sites. Moreover, prior to 1991, the NKRD maintained 0.16-km to 0.20-km radius "buffers" of no-harvest around goshawk nests (M. Siders, pers. comm.), minimizing the number of forest openings near nests, and allowing pine and fir regeneration to grow. Because the majority of nests used in this study were found at or near nests discovered prior to 1991, the vegetative characteristics at territories may reflect this buffer history. However, because the area within historical buffers comprises only 1% of the 0.60-km radius area of important habitat relationships surrounding nests, this explanation for the differences is unconvincing. The only exception to this caveat may include the higher amounts of mixed-conifer habitat within 0.15 km of nests in lower quality territories than at random locations, which may reflect regeneration that occurred within the buffer zone.

In this study, I was unable to detect differences between levels (lower, higher) of reproductive fitness, and thus between source and sink habitat, for either the amount or spatial arrangement of vegetative types. This may be the result of poor resolution between clusters for territories falling along the axis determined by the standardized least squares means for the number of young fledged (Fig. 4.3). Better resolution between clusters may be achieved using upper and lower variable percentiles, as demonstrated by Thome et al. (1999), or by using additional, relevant demographic variables (Manly 1986:133) such as fecundity or survival (Newton 1989, Franklin 1997) in the cluster analysis. Alternatively, use of additional habitat variables, such as forest structural components (Chapter 2), might better discriminate between higher and lower quality territories, as forest structure may be as, if not more, important to goshawk reproductive success than composition of forest vegetation. This analysis is currently underway.

MANAGEMENT IMPLICATIONS

If higher quality territories represent goshawk population "sources," thus contributing to persistence of the species on the study area, then the habitat conditions therein should not be altered greatly beyond the vegetative characteristics identified in this study. Altering the habitat greatly or, similarly, refraining from any form of habitat maintenance over time, may change the proportion of vegetative types, thereby creating more random-like habitat conditions. Conservatively, management recommendations based on this research should be carried out with caution until the assessment of additional habitat variables relevant to the demographic performance of goshawks on the study area (such as the horizontal and vertical structure of the forest) are made.

The spatial arrangement of the proportions of vegetative types was most important in predicting whether a site belongs to a goshawk territory or random plot on the study area, and whether a territory would be of higher or lower reproductive quality. The proportions of pinyon-juniper (0.30-0.60-km radius), mixed conifer (0.00-0.15-km radius), ponderosa pine (0.15-0.30-km radius), and spruce-dominated (0.15-0.30-km radius) habitats near the sample plots positively influenced territory location; whereas, the proportion of openings (0.00-0.15-km radius) strongly lowered the probability of a territory location. That the proportion of spruce-dominated mixes between 1.20 and 2.40 km (corresponding to the FA) of a sample plot also negatively influenced the probability of finding a territory suggests that adequate foraging habitat is also important to the goshawk. The typically dense, understory filled, spruce-mixed forest limits access to prey.

The amount of ponderosa pine within 0.15 km and mixed conifer within 0.30 km of sample locations, corresponding to the NA and a portion of the PFA within territories, were associated with higher quality territories. These habitat types are typically dominated by large ponderosa pine crowns (mixed conifer being comprised mostly of mature ponderosa pine with fir regeneration), suggesting that once this criterion is met, the NA and especially the PFA may support a variety of forest conditions. Once a nest is established, the spatial arrangement of vegetative types within the foraging area does little to differentiate further higher from lower quality habitat. Stronger relationships are expected to emerge, however, using variables of forest structure (including horizontal and vertical measures) due to their influence on the access to and resource availability for prey.

Because significant habitat variables in these analyses reflected past management on the study area, the regression coefficients herein (Tables 4.8, 4.9) may be used to predict the effects of future forest management on goshawk territory location and site quality. However, model parameters should not be used to estimate management targets to maintain goshawk habitat. In addition, the predictive use of these equations is limited to the study area on the NKRD. For forests elsewhere, site-specific functions should be developed.

ACKNOWLEDGEMENTS

This research was supported by Region 3 (Albuquerque, NM) and the Rocky Mountain Research Station (Fort Collins, CO) of the USDA Forest Service. The NKRD (Fredonia, AZ) provided housing and logistical support throughout the study. I thank the

countless field technicians and volunteers that assisted with goshawk monitoring between the years 1991 and 2000. P. L. Chapman and R. M. King kindly provided statistical advice. J. Lambert compiled a portion of the data used in the analyses. I am grateful to V. L. Thomas for providing customized ArcView® applications. P. Kaval provided useful comments on an earlier draft of the manuscript.

LITERATURE CITED

Akaike, H. 1977. On entropy maximization principle. Pages 27-41 *in* P. R. Krishnaiah, ed. Applications of Statistics. North-Holland, Amsterdam.

Allison, B. 1996. A landscape characterization of nesting northern goshawk habitat in northern California using remote sensing and GIS. M.Sc. Thesis, S. Illinois Univ., Edwardsville. 77 p.

Baker, B. W., B. S. Cade, W. L. Mangus, and J. L. McMillen. 1995. Spatial analysis of sandhill crane nesting habitat. J. Wildl. Manage. 59:752-758.

Bloom, P. H. 1987. Capturing and handling raptors. Pages 99-123 *in* B. A. G. Pendleton, B. A. Milsap, K. W. Kline, and D. M. Bird, eds. Raptor Management Techniques Manual. Natl. Wildl. Fed. Sci. Tech. Ser. No. 10, Washington, DC.

Breiman, L., Friedman J. H., Olshen R. A., and. Stone C. J. 1984. Classification and Regression trees. Wadsworth Ind. Group, Belmont, CA. 358 p.

Bright-Smith and D. J., and R. W. Mannan. 1994. Habitat use by breeding male northern goshawks. Stud. Avian Biol. 16: 58-65.

Brown, J. L. 1969. Territorial behavior and population regulation in birds. A review and re-evaluation. Wilson Bull. 81:293-329.

Crocker-Bedford, C. 1990. Goshawk reproduction and forest management. Wildl. Soc. Bull. 18:262-269.

De'Ath, G., and Fabricius K. E. 2000. Classification and regression trees: a powerful yet simple technique for ecological data analysis. Ecology 81:3178-3192.

Desimone, S. M. 1997. Occupancy rates and habitat relationships of northern goshawks in historic nesting areas in Oregon. M.Sc. Thesis, Oregon State Univ, Corvallis. 78 p.

Donovan, T. M., Thompson III, F. R., Faaborg, J., and Probst, J. R. 1995. Reproductive success of migratory birds in habitat sources and sinks. Conserv. Biol. 9:1380-1395.

Efron, B. and R. J. Tibshirani. 1993. An introduction to the bootstrap. Chapman & Hall, NY. 436 p.

ESRI. 1998. ArcView® 3.1. Environmental Research Institute, Inc., Redlands, CA.

Finn, S. P., J. M. Marzluff, and D. E. Varland. 2002. Effects of landscape and local habitat attributes on northern goshawk site occupancy in western Washington. For. Sci. 48:427-436.

Franklin, A. B. 1997. Factors affecting temporal and spatial variation in northern spotted owl populations in northwest California. Ph.D. Diss., Colo. State Univ., Fort Collins. 185 p.

Fretwell, S. D., and H. L. Lucas. 1970. On territorial behavior and other factors influencing habitat distribution in birds. I. Theoretical development. Acta Biotheor. 19:16-36.

Hall, J. G. 1981. A field study of the Kaibab squirrel in Grand Canyon National Park. Wildl. Monogr. 75:6-54.

Hargis, C. D., C. McCarthy, and R. D. Perloff. 1994. Home ranges and habitats of northern goshawks in eastern California. Stud. Avian Biol. 16:66-74.

Hayward, G. D., and R. E. Escano. 1989. Goshawk nest-site characteristics in western Montana and northern Idaho. Condor. 91:476-479.

Hosmer, D. W., Jr., and S. Lemenshow. 2000. Applied logistic regression. John Wiley & Sons, Inc., NY. 373 p.

Howe, R. W., G. J. Davis, and V. Mosca. 1991. The demographic significance of 'sink' populations. Biol. Conserv. 57:239-255.

Jensen, J. R. 1986. Introduction to digital image processing: A remote sensing perspective. Prentice-Hall, Englewood Cliffs, NJ. 379 p.

Joy, S. M., R. M. Reich, and R. T. Reynolds. 2001. Modeling small-scale variability in the composition of goshawk habitat on the Kaibab National Forest. *In* J.D. Greer, ed. Proc. (CD-ROM) Eighth Biennial USDA For. Serv. Remote Sensing Applications Conference, 10-14 April, 2000, Albuquerque, NM. 11 p.

Joy, S. M., R. T. Reynolds, and D. G. Leslie. 1994. Northern goshawk broadcast surveys: hawk response variables and survey costs. Stud. Avian Biol. 16:24-31.

Kennedy, P. L. 1997. The northern goshawk (Accipiter gentilis atricapillus): is there evidence of a population decline? J. Raptor Res. 31:95-106.

Kennedy, P. L., and D. W. Stahlecker. 1993. Responsiveness of nesting northern goshawks to taped broadcasts of three conspecific calls. J. Wildl. Manage. 57:249-257.

Kenward, R. 1982. Goshawk hunting behaviour, and range size as a function of food and habitat availability. J. Anim. Ecol. 51:69-80.

Kenward, R. E., and P. Widén. 1989. Do goshawks Accipiter gentilis need forests? Some conservation lessons from radio tracking. Pages 561-567 in B. U. Meyburg and R. D. Chancellor, eds. Raptors in the Modern World. World Working Group on Birds of Prey and Owls, Berlin, Germany.

Koenig, W. D., F. A. Pitelka, W. J. Carmen, R. L. Mumme, and M. T. Stanback. 1992. The evolution of delayed dispersal in cooperative breeders. Q. Rev. Biol. 67:111-150.

Korpimäki, E. 1988. Effects of territory quality on occupancy, breeding performance and breeding dispersal in Tengmalm's owl. J. Anim. Ecol. 57:97-108.

Korpimäki, E. 1990. Body mass of breeding Tengmalm's owls Aegolis funerus: seasonal, between-year, site and age-related variation. Ornis Scand. 21:169-178.

Kostrzewa, A. 1996. A comparative study of nest-site occupancy and breeding performance as indicators for nesting-habitat quality in three European raptor species. Ethol., Ecol., Evol. 8:1-8.

Lincoln, R., G. Boxshall, and P. Clark. 1998. A dictionary of ecology, evolution and systematics. Cambridge University Press, UK. 361 p.

Linkhart, B. D. 2001. Life history characteristics and habitat quality of flammulated owls (Otus flammeolus) in Colorado. Ph.D. Diss., Univ. of Colo., Boulder. 206 p.

Manly, B. F. J. 1986. Multivariate statistical methods: A primer. Chapman & Hall, London, UK. 215 p.

McClaren, E. L., P. L. Kennedy, and S. R. Dewey. 2002. Do some northern goshawk nest areas consistently fledge more young than others? Condor 104:343-352.

McGrath, M. T. 1997. Northern goshawk habitat analysis in managed forest landscapes. M.Sc. Thesis, Oregon State Univ., Corvallis. 127 p.

Mielke, P. W., and K. J. Berry. 2001. Permutation methods: a distance function approach. Springer-Verlag, NY. 352 p.

Miller, R. G., Jr. 1981. Simultaneous Statistical Inference. Springer-Verlag, NY. 299 p.

Moore, K. R., and C. J. Henny. 1983. Nest site characteristics of three co-existing accipiter hawks in northeastern Oregon. J. Raptor Res. 17: 65-76.

Morison, M. L., B. G. Marcot, and R. W. Mannan. 1992. Wildlife-habitat relationships. Univ. of Wisconsin Press, Madison. 343 p.

Newton, I. 1979. Population ecology of raptors. Buteo Books, Vermillion, SD. 399 p.

Newton, I. 1985. Lifetime reproductive output of female Sparrowhawks. J. Anim. Ecol.54:241-253.

Newton, I. 1989. Sparrowhawk. Pages 279-296 *in* I. Newton (ed.), Lifetime Reproduction in Birds. Academic Press, Ltd., London. 479 p.

Newton, I. 1991. Habitat variation and population regulation in Sparrowhawks. Ibis 133:76-88.

Newton, I. 1992. Lifetime reproduction in birds. Academic Press, London. 479 p.

Newton, I., and I. Wyllie. 1992. Fidelity to nesting territory among European Sparrowhawks in three areas. J. Raptor Res. 26:108-114.

Peck, J. 2000. Seeing the forest through the eyes of a hawk: an evaluation of recent efforts to protect northern goshawk populations in southwestern forests. Nat. Res. J. 40:125-156.

Petrondas, D. A., and G. K. Ruben. 1983. Multiple comparisons by rerandomization tests. J. Am. Statist. Assoc. 78:949-957.

Pulliam, H. R. 1988. Sources, sinks and population regulation. Amer. Natural. 132:652-661.

Pulliam, H. R., and B. J. Danielson. 1991. Sources, sinks, and habitat selection: A landscape perspective on population dynamics. Amer. Natural. 137:850-866.

Raphael, M. G., R. G. Anthony, S. DeStesfano, E. D. Forsman, A. B. Franklin, R. Holthausen, E. C. Meslow, and B. R. Noon. 1996. Use, interpretation, and implications of demographic analyses of northern spotted owl populations. Stud. Avian Biol. 17:102-112.

Reynolds, R. T. 1982. North American accipiter hawks. Pages 288-289 *in* D. E. Davis, ed. Handbook of Census Methods for Terrestrial Vertebrates. CRC Press, Inc., Boca Raton, FL.

Reynolds, R.T. 1983. Management of western coniferous forest habitat for nesting Accipiter hawks. U.S.D.A. For. Serv., Gen. Tech. Rep. RM-102, Rocky Mount. For. Range Range Exp. Stat., Fort Collins, CO. 7 p.

Reynolds, R.T. 1989. Accipiters. Pages 92-102 *in* B. G. Pendleton, G. E. Ruibal, D. L. Krahe, K. Steenhof, M. N. Kochert, M. N. LeFranc, Jr., eds. Proc. of the Western Raptor Management Symposium and Workshop. Natl. Wildl. Fed. Sci. Tech. Ser. No. 12, Washington, DC. 317 p.

Reynolds, R.T., and S. M. Joy. 1998. Distribution, territory occupancy, dispersal, and demography of northern goshawks on the Kaibab Plateau, Arizona. Final Rep., Ariz. Game and Fish Heritage Proj. No I94045. 76 p.

Reynolds, R. T., E. C. Meslow, and H. M. Wight. 1982. Nesting habitat of coexisting Accipiter in Oregon. J. Wildl. Manage. 46: 124-138.

Reynolds, R. T., R. T. Graham, M. H. Reiser, R. L. Bassett, P. L. Kennedy, D. A. Boyce, Jr., G. Goodwin, R. Smith, and E. L. Fisher. 1992. Management recommendations for the northern goshawk in the southwestern United States. U.S.D.A. For. Serv., Gen. Tech. Rep. RM-217, Rocky Mount. For. Range Exp. Stat., Fort Collins, CO. 90 p.

Reynolds, R.T., S. M. Joy, and D. G. Leslie. 1994. Nest productivity, fidelity, and spacing of northern goshawks in northern Arizona. Stud. Avian Biol. 16: 106-113.

Reynolds, R. T., S. M. Joy, and Robin M. Reich. In prep. Effects of Anthropogenic and natural disturbances on frequency of breeding by goshawks on the Kaibab Plateau.

Reynolds, R. T., S. M. Joy, J. D. Wiens, S. R. Salafsky, and J. C. Seyfried. In review. Territory and mate fidelity and breeding dispersal in the northern goshawk. Condor.

Ripple, JW. J., P. D. Lattin, K. T. Hershey, F. F. Wagner, and E. C. Meslow. 1997. Landscape composition and pattern around northern spotted owl nest sites in southwest Oregon. J. Wildl. Manage. 61:151-158.

Risser, P. G., J. R. Karr, and R. T. T. Forman. 1984. Landscape ecology: directions and approaches. Illinois Nat. Hist. Survey Special Publ. No. 2. 17 p.

Rodewald, A. D., and R. H. Yahner. 2001. Avian nesting success in forested landscapes: influence of landscape composition, stand and nest-patch microhabitat, and biotic interactions. Auk 118:1018-1028.

Romesburg, H. C. 1984. Cluster analysis for researchers. Lifetime Learning Publications, Belmont, CA. 334 p.

SAS Inst. Inc. 2000. The SAS System for Windows, release 8.1. SAS Institute, Cary, NC.

Siders, M. S., and P. L. Kennedy. 1996. Forest structural characteristics of accipiter nesting habitat: is there an allometric relationship? Condor 98:123-132.

Speiser, R., and T. Bosakowski. 1984. History, status, and future management of goshawk nesting in New Jersey. Rec. New Jersey Birds. 10:29-33.

Speiser, R., and R. Bosakowski. 1987. Nest site selection by northern goshawks in northern New Jersey and southeast New York. Condor. 89:387-394.

SPSS, Inc. 2002. SPSS Statistical Software, vers. 11.0. Chicago, IL.

Squires, J. R., and L. Ruggiero. 1996. Nest-site preference of northern goshawks in southcentral Wyoming. J. Wildl. Manage. 60:170-177.

Steinberg, D., and P. Colla. 1997. CART--Classification and Regression Trees software. Salford Systems, San Diego, CA.

Swindle, K. A., W. J. Ripple, E. C. Meslow, and D. Schafer. 1999. Old-forest distribution around spotted owl nests in the central Cascade Mountains, Oregon. J. Wildl. Manage. 63:1212-1221.

Thome, D. M., C. J. Zael, and L. V. Diller. 1999. Forest stand characteristics and reproduction of northern spotted owl in managed north-coastal California forests. J. Wildl. Manage. 63:44-59.

Turner, M. G., and R. H. Gardner. 1991. Quantitative methods in landscape ecology. Springer-Verlag, NY, Inc. 536 p.

Van Horne, B. 1983. Density as a misleading indicator of habitat quality. J. Wildl. Manage. 47:893-901.

Van Horne, B., G. S. Olson, R.L. Schooley, J. G. Corn, and K. P. Burnham. 1997. Effects of drought and prolonged winter on Townsend's ground squirrel demography in shrubsteppe habitats. Ecol. Monogr. 67:295-315.

Vickery, P. D., M. L. Hunter, Jr., and J. V. Wells. 1992. Use of a new reproductive index to evaluate relationship between habitat quality and breeding success. Auk 109:697-705.

Ward, J. M. and P. L. Kennedy. 1996. Effects of supplemental food on survival and movements of juvenile northern goshawks. Auk 113(1):200-208.

White, M. A., and J. T. Vankat. 1993. Middle and high elevation coniferous forest communities of the North Rim region of the Grand Canyon National Park, Arizona, USA. Vegetat. 109:161-174.

Widén, P. 1984. Activity patterns and time-budget in goshawk *Accipiter gentilis* in a boreal forest area in Sweden. Ornis Fenn. 61:109-112.

Widén, P. 1989. The hunting habits of goshawks Accipiter gentilis in boreal forests of central Sweden. Ibis 131:205-231.

Wiens, J. A. 1973. Interterritorial habitat variation in grasshopper and savannah sparrows. Ecol. 54:877-884.

Wiens, J. A. 1986. Spatial scale and temporal variation in studies of shrubsteppe birds. Pages 154-172 *in* J. Diamond and T. Case, eds. Community Ecology. Harper Row Publ., NY.

Younk, J. V., and M. J. Bechard. 1994. Breeding ecology of the northern goshawk in high-elevation aspen forests of northern Nevada. Stud. Avian Biol. 16:119-121.

DISSERTATION SUMMARY

OVERVIEW

The management and conservation of wildlife species, such as the goshawk, require not only knowledge of the species' life history characteristics, but also knowledge of the optimal composition and spatial arrangement of the habitat resources that meet the species' biological needs. Fine resolution habitat models are integral to assessing wildlife-habitat relationships at multiple scales. In Chapters 1 and 2, I showed that it is possible to model the composition and structure of forest vegetation on a large (1,285 km^2) and diverse study area to a fine resolution (10 m x 10 m) from broader-scale (30 m x 30 m) Landsat TM imagery. With field data collected at a desired resolution as a main variable and Landsat data as an auxiliary variable, I generated fine-scale inferences about characteristics of the goshawk's environment previously unavailable from conventional Landsat TM classifications. In Chapter 3, I predicted the location of active nests throughout the study area using a dynamic spatial simulation model that incorporated the interaction of these fine-scale habitat models with a sample of active goshawk nests whose spatial distribution reflected the bird's territorial behavior. Although territoriality plays a much greater role in nest placement on the study area than the availability of nest site habitat, this relationship could change with excessive levels of natural or management-caused disturbance to forests. Notably, this model makes no inferences about the reproductive performance of individual pairs of hawks at predicted nest

locations. Moreover, factors that were not addressed here, such as the distribution and abundance of prey or the composition and structure of foraging habitat, may also influence nest location.

The dynamic spatial simulation model developed here brings wildlife habitat modeling to a new level of sophistication where multi-way relationships between a species and components of its environment are explored simultaneously. This dynamic spatial simulation model provides a more realistic understanding of the influences that habitat components have on a species' life history characteristics. Furthermore, this approach to modeling more closely simulates ecosystem-based processes, an important concept in the conservation of species, especially top-level predators such as the goshawk whose populations are sensitive to changes in complex food webs caused by natural and management related habitat changes.

In Chapter 4, I explored the relationship between tree species composition at multiple scales within goshawk territories and goshawk reproductive performance (annual rate of egg laying and productivity) to understand the influence that the amount and arrangement of habitat have on goshawk population persistence. Territories with high reproductive success (i.e., high quality territories) were distinguished from random locations by a greater proportion of ponderosa pine, fewer deciduous trees and openings, and lower species diversity. Low quality territories also had fewer openings than random locations. Smaller proportions of deciduous-dominated vegetation and fewer openings, particularly near (0.0-0.6 km from) nests, suggested possible selection by goshawks for these habitat characteristics. Although I was unable to detect differences in the composition or diversity of vegetation between territories with higher and lower

reproductive performance, this apparent lack of distinction may reflect the habitat components tested. Use of alternative habitat variables that influence demographic performance, such as forest structural components, may provide a more appropriate test of these differences.

FUTURE WORK

In the coming months, I will explore the relationship between forest structure and goshawk reproductive performance at multiple scales, as with forest species composition. Habitat features (composition/structure) that are correlated with territory quality will be incorporated into the dynamic spatial simulation model to assess the effects of territorial behavior and forest species composition and structure on population performance (Fig. S.1). In addition, by introducing alternate management scenarios into the habitat models, I should be able to assess the probable effects of management-related forest changes on nest location and reproductive success, and track changes in these variables over time as the forest regenerates.

Eventually, information on the fecundity and survival of goshawks, as well as the distribution of prey species, will be incorporated into the modeling process to improve the predictive power and scope of the dynamic spatial simulation model. This level of integration will incorporate the effects of foraging habitat quality, an important ecological component missing from the current models.

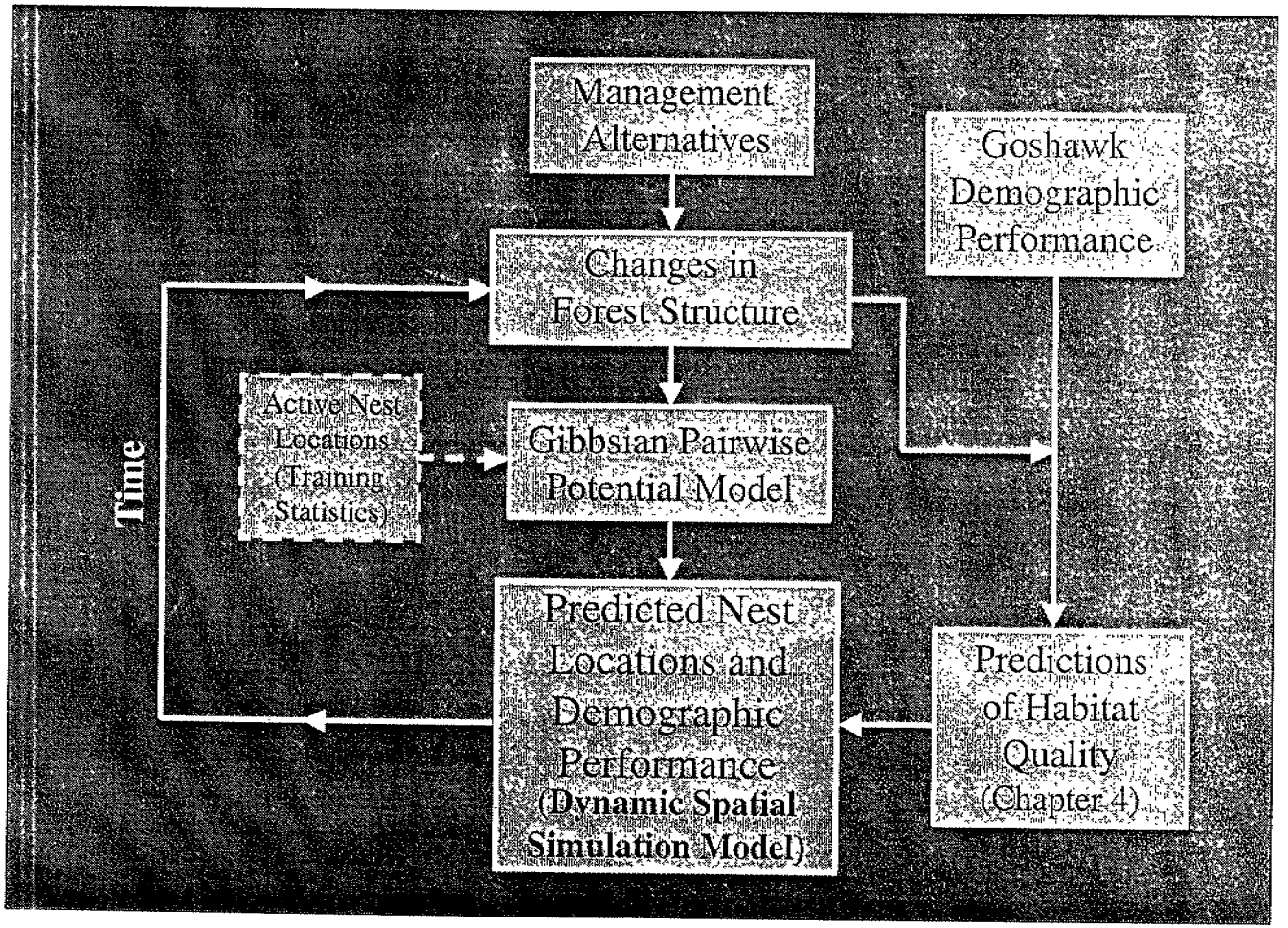


Figure S.1. Overview of the dynamic spatial simulation model that includes predictions of northern goshawk habitat quality and changes in forest structure due to the implementation of management alternatives on the Kaibab National Forest (North Kaibab Ranger District) in Arizona.