# LANDSCAPE-SCALE MODELING OF VEGETATION LAND COVER AND SONGBIRD HABITAT, PINALEÑOS MOUNTAINS, ARIZONA

By J. Judson Wynne

A Thesis

Submitted in Partial Fulfillment

Of the Requirements for the Degree of

Master of Science

in Environmental Sciences and Policy

Northern Arizona University

May 2003

Approved:

______________________________

William M. Block,
Ph.D., Co-Chair

______________________________

Thomas D. Sisk,
Ph.D., Co-Chair

______________________________

David Ostergren, Ph.D.

______________________________

Kathryn A. Thomas, Ph.D.

# ABSTRACT

## LANDSCAPE-SCALE MODELING OF VEGETATION LAND COVER AND SONGBIRD HABITAT, PINALEÑOS MOUNTAINS, ARIZONA

J. JUDSON WYNNE

The availability of remotely sensed imagery and geographic information systems (GIS) data has resulted in these data being increasingly used in guiding making land management decisions. For making the best management decisions, land managers and researchers must understand and identify potential sources of error. Using a competing models framework, I modeled vegetation land cover and songbird habitat on the Pinaleños Mountains, southeastern Arizona. I compared six land cover classification maps derived from Landsat ETM+ imagery. I used various combinations of two spectral correction (atmospheric and sun-angle correction, ASAC) and three enhancement (principal components analysis, normalized difference vegetation index and tasseled-cap transformation) techniques. Elevation and aspect data was also used to reclassify vegetation types to their known elevational and aspect boundaries. An independent verification dataset (collected 2001) was used for the accuracy assessment. The highest accuracy classification map was the "ASAC with principal components analysis (PCA) and normalized difference vegetation index (NDVI) reclassifying mixed coniferous and oak-juniper without reclassifying pine-oak" (overall accuracy 71.2%, $p$-value $< 0.001$). This study demonstrated an increase in overall accuracy of 15.4% (from 55.8% to 71.2%) when comparing the uncorrected to the ASAC with PCA and NDVI classification. I recommend examination of spectral enhancement techniques most applicable to the classification objectives, and the use of a haze and sun-angle correction algorithms. Songbird habitat was modeled using a three-year dataset (1993-1995) of bird survey points, habitat information based on information derived from literature, and coarse landscape-scale variables. Models were validated using a 2002 dataset. I modeled habitat using classification tree and logistic regression models for eight songbird species within a competing models framework. I tested and/ or evaluated all datasets used in all phases of the modeling process. GIS data were considered of the highest quality. Sample sizes were considered low in statistical power (< 30 samples for presence and absence) and the sample design was inappropriate for landscape-scale habitat modeling. The verification dataset was collected during the 2002 drought. Six of eight yielded accuracy values performed well (> 60%) and were comparable to other studies using similar habitat variables. Low predictive success of these models was potentially due to a combination of inappropriate study design, small sample size, environmental stochasticity in the verification dataset, and lack of finer-scale GIS data. Although use of these models in guiding management decisions is limited, the criteria developed provide a systematic framework for evaluating data quality for modeling wildlife-habitat relationships. I recommend using a method similar to the one presented here for evaluating model datasets and to potentially reduce the extent of error propagation. This will ultimately provide land managers with higher quality habitat models and thus an ability to make better management decisions.

## ACKNOWLEDGMENTS

I am indepted to many people for without whom this project could not have been completed. I would like to acknowledge and thank Dr. William M. Block, USDA-Forest Service, Rocky Mountain Research Station for funding both my graduate studies and field research. I owe a tremendous amount of gratitude to Drs. Thomas D. Sisk, Joy Mast, Kathryn A. Thomas and David Ostergren for serving on my committee, improving upon this thesis, and fast-tracking my defense. I also thank Lisa Angle, George Asmus, Walter Friauf, Genice Froehlich, Lorean Hennings and Larry Jones, USDA-Forest Service, Coronado National Forest, Safford Ranger District for their friendship and endless logistical field support. A special thank you to the Larson Family for the extended use of their cabin at Turkey Flat. I am forever grateful to John DeDecker, Chuck LaRue, Brian McGuiness, Jennifer McNichol, Amanda Moors, Maria Santos, and Mike Stoddard for braving rugged terrain and every climatic condition the Pinaleños has to offer; your comradere, positive attitudes, suggestions and support made our time on “my” mountain a most memorable experience. A tremendous thank you to Lisa Taylor and Doug Friedman for “Woody-sitting.” I would also like to thank the Sky Island Alliance for providing volunteers for fieldwork. Also, I greatly appreciate the advice, support and direction from my *compadres* in the Sisk Applied Ecology Lab and my friends and colleagues at the USDA FS-Rocky Mountain Research Station and USGS-Colorado Plateau Field Station. David Vanderzanden, USDA FS, Remote Sensing Applications Center, Salt Lake City, Utah and Dr. G. Lennis Berlin, Department of Geography, NAU provided direction, instruction and advice in remote sensing; I truly appreciated your assistance and support. Special thanks to James Battin, Charles Drost and Maria Santos for suggestions on improving this thesis, Dr. John Prather for GIS and statistical support, Steve Andariese and Jay Thompson for access to the NAU School of Forestry, Remote Sensing Lab, Kate Woodman-Strait, Information Technology, NAU for MS Access database support, and Jeff Jenness for endless GIS support and philosophical debates.

I extend my deepest gratitude and esteem to Ethan Aumack, John DeDecker, Chas Jones, Nesta McKnight, Maria Santos, and Lisa Taylor for their friendship, support and guidance down the winding road of life. Also, special thanks to Nesta, Justin and Brian and numerous other musicians for a multitude of pickin' and grinnin' sessions over the years. Music, art, literature, philosophy, and science are interconnected. The Ancients knew it, why did so many Moderns forget?

I am most grateful to the most important people in my life, my family. Their unconditional support and love has enabled me to pursue my goals and live my dreams. Thank you Grandmom, Pop, Mom, Janie, Lynn, my sister Kimberly and my four-legged best friend of 15 years, Woodstock. Special remembrance to my Grandmother, Mrs. Ida Mae Wynne, who passed away in 1999, and the lady that carried me to the ice machine when I was but a *litt'l fella*, my other Mother, Ms. Janie Jackson, who passed away in 2000. I dedicate the completion of this thesis and my degree to you both.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

PREFACE

This thesis was written as two independent chapters. Both of which are formatted to fulfil the format requirements of the journal where I will ultimately submit them for publication. Because these chapters were drafted to "stand alone," there was some redundancy of content. Additionally, each chapter has its own methods, literature cited, tables and figures sections.

Chapter one analyses various combinations of spectral correction and enhancement techniques of Landsat ETM+ imagery for developing vegetation land cover maps. The second chapter focuses on modeling songbirds-habitat and the issues of error inherent in the retrospective, verification and GIS datasets, which are used for modeling. I will submit both chapters for publication in the Wildlife Society Bulletin.

# *1.0 Introduction*

The use of satellite imagery-derived vegetation maps and spatially explicit predictive habitat models are of particular importance in the fields of conservation biology, wildlife management (Stoms et al 1992), and ecology (Graetz 1990). Graetz (1990) suggests a global understanding of ecological processes is nearly impossible without the intensive and extensive use of remotely sensed imagery. Stoms et al (1992), Pearce and Ferrier (2000), Penhollow and Stauffer (2000), Wright et al. (2000), and Brugnach et al. (2003) emphasize the importance of spatially explicit habitat modeling in understanding wildlife-habitat relationships.

Development of these products generally requires field data (for model building and testing), remotely sensed imagery and GIS-based information. For the produced data products to be used within a management context, they should be defensible (Starfield 1997). However, errors in these data products and their subsequent uncertainty may challenge their defensibility. Lane and Chandler (2003) suggest although information used in these applications has become increasingly easy to generate, and user control has improved due to user-friendly software programs, these data are not always of sufficient quality. However, despite data errors and error propagation, users of these data products rarely understand or address the limitations and quality of these datasets (Chrisman 1987).

Prior to accepting satellite-derived data products and spatially explicit habitat models *prima facie*, the data used in developing and verifying these products must be evaluated for accuracy and appropriateness. I have developed two evaluative criteria for illuminating potential sources of error that may occur and subsequently propagate in the development of satellite imagery derived vegetation maps and spatially explicit habitat models. Although a complete understanding of data quality and determining how errors propagate is often beyond our capabilities (Chrisman 1987, Stoms et al. 1992), identifying where the potential sources of errors may occur can lead to higher quality data products.

My thesis is presented as two separate studies. In the first study, I used satellite imagery and other available datasets to: (1) develop a vegetation map to be used for developing songbird-habitat models, and (2) compare the accuracy of vegetation maps derived from Landsat 7 ETM+ using combinations of two spectral correction (atmospheric haze and sun-angle correction) and three spectral enhancement techniques (principal components analysis, tasseled-cap transformation and normalized vegetation index), and an independent dataset to assess vegetation map accuracy. My objectives for this study were to (1) evaluate the usefulness of the selected vegetation map for modeling songbird habitat and, (2) identify the best combination of these techniques for creating a supervised classification of vegetation map with the highest accuracy. This effort is highly relevant, because managers increasingly rely on vegetation maps derived from remotely sensed imagery, yet data for image classification are often limited. In the second study, I used the highest quality vegetation map in concert with other available datasets to develop and test predictive habitat models for songbirds. The purpose of this study was to develop an evaluative criteria to assist resource managers in processing, evaluating and developing data used in building and testing predictive wildlife-habitat relationship (WHR) models. I used eight passerine bird species, available GIS information, a three-year (1993-95) retrospective dataset (for model building), information derived from peer-reviewed literature, and a one-year (2002) verification dataset (for model testing) of bird presence data for the Pinaleño Mountains, southeastern Arizona to develop and test these criteria. Specifically, I: (1) evaluated the quality of retrospective and verification data, as well as GIS information; (2) compared results of parametric and nonparametric mechanistic WHR models to literature-derived information models; and, (3) evaluated the usefulness of model predictions to land managers. Both of these studies are representative of issues faced by most resource managers, where management challenges require the development of improved spatial data products, but data are often limited and resources are insufficient to fund major surveys.

## *1.1 Literature Cited*

Brugnach, M., J. Bolte and G. Bradshaw. 2003. Determining the significance of threshold uncertainty in rule-based classification models. Ecological Modeling 160: 63-76.

Chrisman, N. 1987. The accuracy of map overlays: a reassessment. Landscape and Urban Planning 14: 427-439.

Graetz, R.D. 1990. Remote sensing of terrestrial ecosystem structure: an ecologists pragmatic view. In Remote sensing of biosphere functioning (R.J. Hobbs, H.A. Mooney, eds.) Springer, New York, pps. 5-30.

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

Penhollow, M. & D. Stauffer, 2000. Large-scale habitat relationships of neotropical migratory birds in Virginia. Journal of Wildlife Management 64 (2): 363-372.

Starfield, A.M. 1997. A pragmatic approach to modeling for wildlife management. Journal of Wildlife Management 61 (2): 261-270.

Stoms, D.M., F.W. Davis and C.B. Cogan 1992. Sensitivity of wildlife habitat models to uncertainties in GIS information. Photogrammetric Engineering and Remote Sensing 58 (6): 843-850.

Wright, A., A. Fielding, C. Wheater 2000. Predicting the distribution of Eurasian badger (Meles meles) over an urbanized landscape: a GIS approach. Photogrammetric Engineering and Remote Sensing 66 (4): 423-428.

# 2. An Analysis of Remote Sensing Techniques for Improving Vegetation Map Accuracy: Implications for Land Management

### 2.1.1 Abstract

Remote sensing technology has various applications in resource management, including the mapping of vegetation, land use and wildlife habitat. The value of remote sensing is widely appreciated in wildlife biology. However, its appropriate application is impossible without informed approaches to image processing and accuracy assessment of resultant data products. In the absence of incisive image processing and error analysis, remotely sensed data may be applied in ways that obscure, rather than illuminate, our understanding of vegetation cover and wildlife habitat. Many state and federal agencies, eager to save time and money, have pursued Landsat imagery as a surrogate for extensive field studies of habitat quality and vegetation distribution. Remote sensing technology has provided extensive coverage of areas much larger than those previously mapped, and the availability of Landsat imagery has facilitated an acceptable shortcut to time- and labor -intensive habitat mapping. However, many of these data products are untested, and little guidance is available to the resource practitioner regarding practical and efficient image processing and accuracy assessment techniques. To evaluate the utility of Landsat data for creating a vegetation map to assist in the songbird habitat management, I developed and compared six vegetation maps, derived from various combinations of two spectral correction (atmospheric haze and sun-angle correction) and three enhancement (principal components analysis, normalized difference vegetation index and tasseled-cap transformation) techniques.

Elevation and aspect data were also used to constrain vegetation types within their known elevational and aspect boundaries. A “competing models” approach was then used to select the highest quality vegetation map. The vegetation map with the highest accuracy was corrected for atmospheric haze and sun-angle, corrected with PCA and NDVI, and involved refinement mixed coniferous and oak juniper by known elevational and aspect thresholds-overall accuracy of the resultant vegetation map was 71.2% ($p < 0.001$), with user accuracies ranging from a low of 30.8% in pine-oak, to 80.9% in oak juniper. This study strongly suggests pre- and post-processing techniques are often of great importance in developing useful vegetation maps, and adoption of any data product without first assessing accuracy is ill advised. Overall accuracy in this study was lower than expected, but likely reflects the upper end of accuracies obtainable from vegetation maps based on Landsat data and employing techniques currently available to most resource managers. Given the reliance of managers on baseline vegetation maps and the critical value of ensuing decisions, accuracy assessment and evaluation within the land management context should be mandatory for all uses of remotely sensed data. It is not safe to assume remotely sensed data will provide improved vegetation maps, simply because they may offer the advantages of higher resolution intensive base-line field investigations of the management area.

### *2.1.2 Introduction*

Remote sensing technology has various earth-science applications including mapping geology, land use and forest types (Chavez 1996). Graetz (1990) emphasized the importance of remotely sensed imagery to the field of ecology, asserting a global understanding of ecological processes is impossible without its intensive and extensive use. This technology is of particular importance in assisting in the decision making process concerning the use of natural resources. Sustainable management of any ecosystem requires, among other things, a thorough understanding of the distribution of vegetation across the landscape (Schmidt and Skidmore 2003). However, agencies no longer have the time, money and personnel to conduct extensive field studies (Lins and Kleckner 1996), which may assist in quantifying vegetation distributions. Therefore, a system for producing maps quickly and economically using easily accessible data is highly desirable (Cardillo et al. 1999). The use and interpretation of multispectral satellite imagery provides an alternative to large-scale field efforts, but often this approach is applied without an attempt to quantify errors in vegetation mapping. Thus, the adequacy of these applications may remain unverified. Nevertheless, many state and federal programs have adopted the use of multispectral satellite imagery for environmental monitoring and assessment projects, and these data are often the primary source of information for interpreting systems at the landscape scale (Lins and Kleckner 1996).

Although satellite imagery processing and interpretation is becoming increasingly important to land managers, there are serious practical limitations concerning remote sensing techniques and technologies, based on some fundamental properties of instrumentation, ubiquitous heterogeneity and difficulties of interpreting natural processes using these technologies (Okin et al. 2001). Classifying vegetation accurately within densely vegetated mountainous terrain is often difficult, due to the dynamic range and low amplitude of the radiance within the visible portion of the electromagnetic spectrum (Chavez 1992). Additional problems include the definition of mutually exclusive vegetation classes and the delineation of transition zones (areas where one vegetation type gradates into another type, Avery and Berlin 1992). For these and other reasons, maps derived from remotely sensed data are often of dubious accuracy, which may make these data unsuitable for guiding management decisions (Townshend 1992, Foody 1999).

To overcome some of these limitations, processing techniques, such as spectral correction and enhancement techniques, can be applied to improve information content of the data (Crist et al. 1986, Chavez 1996, Jensen 1996). Appropriate application of these techniques may improve the ability to resolve differences between vegetation types, increasing the accuracy of vegetation mapping efforts.

The purpose of this study was to (1) develop a vegetation land cover map useful for developing songbird-habitat models, and (2) compare the accuracy of vegetation maps derived from Landsat 7 ETM+, using two spectral correction (atmospheric haze and sun-angle correction) and three spectral enhancement techniques (principal components analysis, tasseled-cap transformation and normalized vegetation index). My objective was to evaluate the usefulness of this product for mapping songbird habitat and identify the best combination of these techniques for creating a supervised classification of vegetation map with the highest accuracy. This effort is highly relevant, because managers increasingly rely on vegetation maps derived from remotely sensed imagery, yet data for image classification are often limited. Additionally, formal accuracy assessments of spatial data products are not routinely executed. Thus, this study addresses the general question of the utility of Landsat-based vegetation maps developed from limited ground data. I chose, as a case study, an approximately 470 $km^2$ on the Coronado National Forest, southeastern Arizona. Various vegetation data sets were available. However, none were of sufficient quality to guide bird habitat management decisions. Therefore, I devised the criteria elucidated in this paper for creating a vegetation land cover map to assist in modeling bird-habitat relationships. This case is representative of many public lands, where management challenges require the development of improved spatial data products, but data are limited and resources are insufficient to fund major vegetation surveys.

### *2.1.3 Study Area*

The Pinaleños Mountains are located in Graham County, Arizona, approximately 200km north by northeast of Tucson, Arizona (Figure 1). Delineating the northern extent of the Madrean Archipelago, this sky island is managed by the USDA -Forest Service Coronado National Forest, Safford Ranger District. The planametric area of the study area is approximately 410 $km^2$ and its surface area, adjusted to account for the varied topography (*sensu* Jenness 2001), is approximately 470 $km^2$. The area was defined by a lower elevational limit of 1455 m, which corresponded to the lower limits of the avian research efforts (W. Block, *unpublished data*).

## *2.2 Methods*

Land cover maps were generated via supervised classification of Landsat 7 ETM+ imagery (capture date: 12 November 1999). This supervised classification was undertaken in collaboration with, and guidance from, the USDA FS Remote Sensing Application Center, Salt Lake City, Utah and the Department of Geography, Northern Arizona University, Flagstaff, Arizona. Similar technical support is available to most National Forest offices. A USDA Forest Service Stage 2 Stand Exam and photo interpretations of 2-meter resolution digital orthophoto quads were used for creating the training dataset. I used a common algorithm, the maximum-likelihood (Foody 1999, Franco-Lopez et al. 2001, ERDAS 2001) for conducting all supervised classifications.

### *2.2.1 Model Assumptions*

I made the following simplifying assumptions in selecting the highest quality vegetation map:

1) A “hard” image classification (*sensu* Foody 1999) was used for classifying land cover, which assumes no spectral mixing of vegetation types within each 30-meter pixel.
2) The data in the satellite imagery are normally distributed; this is a requisite for using the maximum-likelihood algorithm for classifying land cover.
3) The vegetation classes, selected *a priori*, are mutually exclusive and represented in crisp sets (Foody 1999).

### *2.2.3 Vegetation Mapping Process Description*

This process was comprised of seven steps (Figure 1), as follows:

1) Image preprocessing – All remotely sensed imagery and GIS information were reprojected into a common projection system. Three sets of imagery were created in this phase: 1) one set uncorrected for atmospheric haze and sun-angle, 2) another set corrected for atmospheric haze and 3) a third set corrected for atmospheric haze and sun-angle.

2) Image processing – Principal components analysis (PCA), normalized difference vegetation index (NDVI) and tasseled-cap transformation (TCT) spectral enhancement techniques were applied to the three aforementioned sets of imagery.

3) Layer stack development – Various combinations of the aforementioned spectral correction and enhancement techniques were produced to create six different layer stacks. A *layer stack* is a combination of individual images to create a "stacked" dataset for analysis. For example, bands 1, 2 and 3 (blue, green and red) of Landsat ETM+ imagery may be stacked to create a color composite.

4) Supervised Classification – Spectral signatures were derived from each layer stack. Next, supervised classifications were performed, on each of the six layer stacks based, on a USDA FS Stage 2 Stand Exam and photointerpretations of aerial photography.

5) Accuracy Assessment I – Using an independent reference dataset of 344 points collected in 2001, an accuracy assessments were conducted. Only vegetation maps with overall accuracy ≥50% were retained for further analysis.

6) Image Post-processing – For all vegetation maps with overall accuracy of ≥50%, I used elevation and aspect information derived from habitat associations guide and digital topographic information to refine vegetation classes based on elevational and aspect constraints.

7) Accuracy assessment II – These refined vegetation maps were then accuracy assessed against an independent reference dataset. Overall accuracy, statistical significance, and producer and user accuracy of each vegetation class per supervised classification were generated.

8) "Best" vegetation map selection – The vegetation map with the highest overall accuracy, statistical significance, and producer and user accuracy of each vegetation class was identified as the "best" vegetation map.

Although multiple feedback loops can occur throughout this process whereby analysts and land managers may need to reprocess their data products (Figure 1), only the feedback loop from "image post-processing" to "accuracy assessment" was required in this analysis. Details regarding methods employed in each section of the vegetation mapping process are provided below.

### *2.2.4 Image Preprocessing*

All GIS and remote sensing data were reprojected into UTM, Zone 12, and NAD 1927. The minimum mapping unit was 30 meters, which was also the pixel size.

Once all GIS and remote sensing data were reprojected, three separate images were derived from Landsat 7 ETM+ scene. The first image was an *uncorrected* image without any spectral correction techniques applied. The second image was an *atmospheric haze-corrected* image, and the third was an *atmospheric haze and sun-angle corrected* image.

#### *2.2.4.1 Atmospheric Haze Correction*

Atmospheric scattering occurs when certain wavelengths of the ultraviolet, visible and infrared bands of the electromagnetic spectrum are impeded by obstructions as they enter the earths' atmosphere. Gaseous molecules, suspended particulates and clouds all influence the amount of scatter. The result is atmospheric haze, which can affect the spectral information used in developing remotely sensed imagery. As the earth's atmosphere scatters, absorbs and refracts light, the amount of electromagnetic energy reaching the sensors on the remotely sensed detectors is affected to various degrees (Chavez 1996). The greater the atmospheric haze, the greater the influence on the system's detectors.

Optimal atmospheric haze correction should be image-specific and requires *in situ* field measurements of atmosphere during the satellite overflight (Chavez 1996). Because this information was not available for the Landsat image used, I used the Chavez (1996) Cosine of Solar Zenith Angle (COST) model for this correction. The purpose of the atmospheric haze correction algorithm is to convert the spectral reflectance values generated by satellite sensors to the actual ground reflectance values (i.e., absolute surface reflectance values, Chavez 1996). This approach uses the average of the transmittance values computed by using *in situ* field measurements of atmosphere. Chavez (1966) model, when compared to published and unpublished calculations of atmospheric haze correction were as accurate as those calculated using *in situ* field measurements. Chavez's model, when compared to published and unpublished calculations of atmospheric haze correction, was as accurate as those calculated using *in situ* field measurements.

#### *2.2.4.2 Sun-Angle Correction*

Solar illumination on *terra firma* varies with topography (Chavez 1992). Furthermore, shadows created by topographic relief will introduce added complexities in correcting for low sun-angles (Nunez 1980). Satellite image capture is dependent on the satellite's location at a given time, rather than the location of the sun. Subsequently, when imagery is captured in areas of high topographic relief and when the sun is not at its zenith, topographic shadowing can result. The purpose of the sun-angle correction algorithm is to elevate the sun-angle to the solar zenith (or nadir). This algorithm normalizes brightness values or (digital numbers, DNs) approximating the "value that would be obtained if the image was captured when the sun was at nadir. Sun-angle was also corrected using the COST model (Chavez 1996).

### *2.2.5 Image Processing*

I used three spectral enhancement techniques when attempting to resolve differences between vegetation types. Enhancement techniques are used to: 1) reduce the number of layers considered, 2) provide a more direct association between the spectral reflectance and processes the ground, and 3) highlight information most important to the analyst or user, in this case wildlife biologists interested in bird habitat (Crist et al. 1986). Using bands 1-5 and 7 of the uncorrected, atmospheric haze corrected, and atmospheric haze and sun-angle corrected images, I conducted principal components analyses, tasseled-cap transformation and normalized difference vegetation index. Band 6 was omitted from analysis because this image is a 60-meter resolution thermal band. I compared the combined TCT with the two developed PCA layers and Band 4 to the combined NDVI with the two developed PCA layers and Band 4. This comparison was selected because both TCT and NDVI are important in illuminating differences between vegetation classes. The tasseled-cap algorithm utilizes all six bands, whereas the NDVI uses only the red and near infrared bands in extracting information related to vegetation cover.

### *2.2.5.1 Principal Components Analysis*

Principal component analysis is a common multivariate statistical approach for identifying the most important sources of variance in a multiband images (Basteon and Curtiss 1996, Ricotta et al. 1999). This technique can remove or reduce redundancy in multiband images (Ricotta et al. 1999) condensing these data into a few transformed PCA layers (Basteon and Curtiss 1996, Jensen 1996, ERDAS 2001). I used PCA to transform correlated bands into a single PCA axis image. Landsat TM imagery, bands 1, 2 and 3, and bands 5 and 7 are highly correlated (G. Lennis Berlin, personal communication, Jensen 1996, ERDAS 2001). Principal components are ranked in terms of the amount of variance explained (Ricotta et al. 1999). I applied PCA to bands 1, 2 and 3, retaining principal component axis one ($PC_1$) for use in subsequent vegetation maps, which corresponds to the highest amount of variance within the data, retained. This process was repeated for bands 5 and 7, and $PC_1$ was again retained. Band 4, typically uncorrelated to any other bands was retained without PCA.

### *2.2.5.2 Tasseled-cap transformation*

The tasseled-cap transformation algorithm was developed to: 1) correct for atmospheric haze when *in situ* field measurements for atmospheric haze correction are lacking, and 2) reveal forest attributes such as species, age and structure (Huang et al. 2002). Using a multiple logistic regression equation, this transformation uses six bands of Landsat 7 ETM+ imagery (excluding the 60m resolution thermal band) and compresses these data into three layers – brightness, greenness and wetness (Crist et al. 1986). The dependent variables are either brightness, greenness or wetness and the independent variables are the six TM band reflectance values (Crist et al. 1986). This approach generates linear combinations of the original bands and often captures 95% or more of the data variability (Crist et al. 1986).

#### *2.2.5.3 Normalized difference vegetation index*

Vegetation has low red reflectance (i.e., red portion of the electromagnetic spectrum) due to the absorption of chlorophyll and high near-infrared (NIR) reflectance (i.e., NIR portion of electromagnetic spectrum) due to scattering of leaf mesophyll (Tucker 1978, Hurcom and Harrison 1998). The Normalized Difference Vegetation Index (NDVI) is correlated with net primary production (Sellers 1987) and reflective leaf density (Birky 2001). Sellers (1985) discovered NDVI provides a near-linear relationship to canopy photosynthetically active radiation (PAR) and bulk stomatal resistance. NDVI is calculated from the reflected solar radiation in the near-infrared (NIR) and red (RED) wavelength bands via the algorithm (Carlson and Ripley 1997, ERDAS 1999). For Landsat ETM+ imagery, band 3 is the red wavelength and band 4 is near infrared. NDVI uses the following equation:

$$NDVI = (NIR - RED)/(NIR + RED).$$

### *2.2.6 Layer Stack Development*

A layer stack is a process whereby multiple imagery is overlay to form a stack of images for analysis. I developed six layer stacks for supervised classification. The first layer stacks consisted of combined bands 1, 2, and 3 $PC_1$ layer, combined bands 5 and 7 $PC_1$ layer, band 4 and the TCT brightness, greenness, and wetness layers. The second layer stack consisted of combined bands 1, 2, and 3 $PC_1$ layer, combined bands 5 and 7 $PC_1$ layer, band 4 and the NDVI of the uncorrected image, respectively. The third layer stack contained the combined bands 1, 2, and 3 $PC_1$ layer, combined bands 5 and 7 $PC_1$ layer, band 4 and the three TCT layers. The forth layer stack consisted of the combined bands 1, 2, and 3 $PC_1$ layer, combined bands 5 and 7 $PC_1$ layer, band 4 and the NDVI of the atmospheric haze corrected image. The fifth layer stack contained the combined bands 1, 2, and 3 $PC_1$ layer, combined bands 5 and 7 $PC_1$ layer, band 4 and the three TCT layers. The final layer stack consisted of the combined bands 1, 2, and 3 $PC_1$ layer, combined bands 5 and 7 $PC_1$ layer, band 4 and the NDVI of the atmospheric haze and sun-angle corrected image.

#### *2.2.6.1 Training Data*

Training datasets, which were used for deriving spectral reflectance values of each vegetation class, consisted of a digitized USDA Forest Service Stage 2 Stand exam and photointerpretations of a 2-meter resolution Digital orthophoto quarter quad (DOQQ). The Stage 2 Stand exam was conducted to identify the extent of available habitat for the USDI FWS endangered Mount Graham red squirrel (Lisa Angle, USDA Forest Service, Coronado National Forest, Safford Ranger District, *personal communication*). Because this species occurs primarily in the upper elevation mixed-coniferous forest, this assessment included mainly mixed coniferous forest polygons interspersed with several pine-oak woodland polygons.

Digital orthophoto quads are black and white aerial photographs. Spectral reflectance values for oak-juniper woodlands and grassland were derived by photointerpretations of two-meter mosaiced digital ortho-photo quarter quads (DOQQ).

I used five vegetation classes (grassland/ upland meadow, oak-juniper woodland, pine-oak woodland, mixed-coniferous forest and bug kill). Four of the types were based on the dominant species and life form of vegetation and the fifth represented a mixture of vegetation types, all characterized by "bugkill." Low elevation grassland and high elevation (upland) meadows were found to have similar spectral signatures and were combined. Oak-juniper woodland occurs between 1270 and 1640 meters elevation, without accounting for aspect influence, and 1270 to 1970 between aspects of 135 to 225 degrees (USDA 1997b); species include *Quercus arizonica, Q. emoryi, Q. turbenella, Q.* hybrid, *Juniperus deppeana, J. monosperma,* and *J. osteosperma*. Pine-oak woodland occurs between 1640 and 2420 meters elevation without aspect influence and 1970 to 2850 meters elevation between aspects of 135 to 225 degrees (USDA 1997a); dominant tree species include *Pinus leiophylla, P. ponderosa, Quercus hypoleucoides, Q. rugosa* and *Q.* hybrid. Mixed -coniferous forest occur at elevations ≥ 2430 meters without aspect influence and ≥ 2850 meters for aspects between 135 to 225 degrees (USDA 1997b); dominant woody vegetation includes *Abies* spp., *Picea pungens, Pinus strobiformis* and *Pseudotsuga menziesii*. The vegetation class defined as "bug kill" was located at the upper elevations at and around the summit of Mount Graham. This area was devastated by a moth (*Nepytia janetae*) outbreak; this infestation killed many trees leaving a distinct signature easily seen in uncorrected and unenhanced satellite imagery. During 1999, the year the Landsat ETM+ imagery was captured, the lower elevational limit of this outbreak was 3015 meters (Steve Dudley, USDA FS, Rocky Mountain Research Station, Flagstaff Lab, personal communication).

#### *2.2.6.2 Reference Data*

I collected an independent dataset, collected in June-August 2001, for model verification. The reference data, representing 344 sample points, were collected at random across the five vegetation classes. Efforts were made to obtain an adequate sample size between 45 (Fitzpatrick-Lins 1981) and 50 (Hay 1979, Congalton and Green 1999) for verification the supervised classifications. I collected data at one-half kilometer intervals along traversable roads and hiking trails spanning the elevational gradient of the mountain range, identifying vegetation type by visual assessment. At each sample site, I estimated a circle of sixty-meter radius centered on the point, and divided the plot into four quadrants based on cardinal directions. For each quadrant I estimated the percent cover, for each of five vegetation types. Cover estimates for each of the quadrants were summed for each sample location. A plot containing ≥ 75% of one of the five vegetation types was assigned to that type (e.g., if the summed quadrants of a sample site was summed to 85% oak juniper, the sample site was classified as oak juniper).

Approximately 800 person-hours were dedicated to this data collection effort, for the sole purpose of assessing the accuracy of the various vegetation maps. While the sampling design produced relatively coarse assessment data, and some spatial autocorrelation certainly exists within the dataset, it represents a reasonable compromise between the ideal data for accuracy assessments (G. Lennis Berlin, personal communication) and the practical reality of limited time and budgets faced by managers.

#### *2.2.6.3 Supervised Classification*

I conducted supervised classifications on all six layer stacks. Because spectral signatures usually correspond to a mixture of several land surface types for satellite imagery between 10-30 meters (Price 1994), I used a coarse-scale assessment for classifying land cover.

### *2.2.7 Image post-processing*

As it is important to consider the influences of ecological processes on landscape development, it is equally important to consider the influence of topography, which govern ecological processes (Hadley 1994). Inclusion of topographic variables in modeling vegetation can prove critical to characterizing landscape (Brown 1994). Because there was considerable confusion can result in transition zones (i.e., pine-oak woodlands, Avery and Berlin 1992), additional information such as elevation, may serve to further resolve and ultimately identify differences between vegetation types (Skidmore and Turner 1989). I used elevation and aspect in concert with the supervised classification data to resolve the potential overlap between the different vegetation types.

#### *2.2.7.1 Selection of Elevation Data*

Prior to conducting this analysis, I needed to identify the highest quality elevation dataset. I evaluated three elevation grids (two elevation models and a radar image of elevation). These models were Digital Elevation Model (DEM), Shuttle Radar Topographic Mapping Mission data (SRTM) and National Elevation Dataset (NED). I reviewed the metadata associated with each dataset and analyzed each for "no data values" (i.e., empty cells with "0" value) within each layer to determine which was most accurate. Once the best dataset was selected, I used ERDAS Imagine 8.5 to derive slope and aspect grids. I did not consider the algorithms used for generating these grids as a source of potential error. Jones (1998) concluded creation of the best slope grid, regardless of the algorithm used, depends on the quality of the elevation model. I extrapolated this statement to include aspect because ERDAS Imagine 8.5 uses a similar convolutional method for deriving aspect. This algorithm is a third-order finite difference technique, which uses eight neighboring elevation cell values to derive the center grid value. Thus, by selecting the best elevation model, the amount of error in the slope and aspect grids was minimized.

#### *2.2.7.2 Identification of Vegetation Maps for Refinement Procedure*

Image post-processing techniques were applied to all vegetation maps with ≥ 50% overall accuracy and identified as statistically significant ($p \leq 0.05$). Vegetation maps, which did not meet these criteria, were omitted from further analysis.

*2.2.7.3 Refinement of Vegetation Maps by Elevation and Aspect*

Once the best elevation dataset was selected, I used the Modeler module in ERDAS Imagine 8.5 to evaluate several different combinations of constraining vegetation type by elevation and aspect. Elevation and aspect parameters were identified for each of the four vegetation types using the USDA FS Arizona-New Mexico habitat associations guide (USDA 1997a, 1997b).

Vegetation classes for each map were refined based on elevation and aspect constraints as identified by USDA FS (1997a, 1997b). I used an exploratory process using numerous combinations of vegetation types with associated elevation and aspect constraints to obtain the highest accuracy vegetation map. Elevational thresholds were based on where, elevationally, the vegetation class has been identified to occur. Aspect thresholds were based on where, elevationally and based on aspect where the vegetation class has been identified to occur. For example, mixed conifer forests will occur at a higher elevation on north facing slopes and lower elevation on south facing slopes. I refined each vegetation map with elevation and aspect using three different approaches: 1) all vegetation maps were modeled by constraining all four vegetation types by elevation and aspect thresholds; because grasslands and upland meadows occur throughout the elevational gradient, this vegetation type was excluded from this analysis; 2) oak-juniper woodland and mixed coniferous forest were refined using the defined thresholds per vegetation class; and, 3) the three vegetation classes were refined, using the aforementioned criteria, except when oak-juniper occurred within the pine-oak transition belt. For example, one of the conditional statements used in this process was “if pine-oak or mixed coniferous and less than 1818 meters in elevation, then refined as oak-juniper.”

### *2.2.8 Accuracy Assessment*

Although data quality and accuracy in landscape analyses are of the utmost importance, these areas have received little attention in the management literature (Hess 1994, Hess and Bay 1997, Luck and Wu 2002). Using an ArcView 3x extension, the Cohen’s Kappa statistic (see Congalton and Green 1999) was used to conduct an accuracy assessment by running the reference dataset against each vegetation map. For each vegetation map, classification tables were created and overall accuracy, and producer and user accuracy were calculated (Congalton 1991). Most accuracy assessments generally provide only overall accuracy (Congalton and Green 1999). I have provided a hypothetical example of “vegetated” and “not vegetated” for the sake of illustration. Overall accuracy is calculated by summing the major diagonal cells within the classification table containing correctly classified cells and dividing by the total sample size ($(A+E)/I$, Figure 3). Simply providing overall accuracy can be misleading with respect to model predictability (Congalton 1991, Congalton and Green 1999, Pearce and Ferrier 2000). Therefore, I used the classification table to calculate two additional indices describing model performance- producer and user accuracy. *Producer accuracy* uses the reference data to calculate the accuracy per class (e.g, Vegetated Class = $A/G$, Not Vegetated = $E/H$) because the producer of the vegetation map wants to know how well a certain category can be classified (Congalton 1991). I divided the total number of correct pixels of a given class by the total number of correctly classified pixels from that class (Congalton and Green 1999). *User accuracy* is employed the classified data to calculate the number of correctly classified pixels (e.g., Vegetated Class = *A/C*, Not Vegetated = *E/F*) because the user of the vegetation map is typically most interested in the probability a pixel classified within a category is representative of that category in geographic space (Congalton and Green 1999). For example, consider the producer accuracy of the "vegetated" class is 89% and user accuracy is 60% (based on Congalton and Green 1999). This means the producer of the map claims 89% of the time an area defined on the ground as "vegetated" was classified as "vegetated." However, the user of the map will discover when the map indicates an area is "vegetated," there is a 60% chance it will actually be "vegetated" on the ground.

#### *2.2.8.1 Accuracy Assessment Interpretation*

The accuracy assessments of all six vegetation maps were analyzed. All vegetation maps with either ≤50% overall accuracy or $p > 0.05$ were removed from further analysis.

### *2.2.9 Selection of Best Vegetation Map*

A *competing models approach* was used to select the "best" vegetation map. I conducted this approach in a stepped process. First, I refined all vegetation maps using Approach 1. These maps were accuracy assessed and the refined vegetation map with the highest overall accuracy, user accuracy, specificity and sensitivity, lowest omission and commission errors, and significant $p$-value (≤0.05) were identified. This procedure was repeated for refinement approaches 2 and 3. All significant refined vegetation maps were identified and another accuracy assessment was conducted using these maps. This was done to identify whether these maps differed statistically. Once this assessment was conducted, overall accuracy, producer and user accuracy, and *p*-values were compared. I considered the "best" vegetation map to have the highest overall accuracy, producer and user accuracy, and a significant *p*-value ($\leq 0.05$).

### *2.3 Results*

Results are presented for accuracies of initial image classification phase and refinement phase, identification of the best elevation map, as well as the selection of the best vegetation map.

#### *2.3.1 Reference Data*

The field data collected resulted in the following sample sizes per vegetation type: mixed coniferous (n= 133), pine-oak (n= 39), oak-juniper (n= 123), grassland/ upland meadow (n= 45), and bug kill (n= 4).

#### *2.3.2 Identification of Vegetation Maps for Refinement Procedure*

Of the six land cover maps whose accuracies were assessed, only three had overall accuracies $\geq 50\%$ and the vegetation maps were statistically significant (*p*-value $< 0.05$, Table 1). These were 1) the spectrally uncorrected image classification with PCA and TCT (overall accuracy 52%, *p*-value $< 0.001$), 2) the atmospheric haze corrected vegetation map with PCA and NDVI (overall accuracy 52%, *p*-value <

0.001), and 3) the atmospheric haze and sun-angle corrected with PCA and NDVI (overall accuracy 55.8%, *p*-value < 0.001).

### *2.3.3 Improvements using Elevation and Aspect*

#### *2.3.3.1 Selection of Elevation Data*

National Elevation Data (NED) were identified as the best elevation map. This dataset meets the "best available" data standards of the National Spatial Data Infrastructure (Gesch et al. 2002). The NED dataset is a "seamless" elevation data for the entire United States. Created from traditional DEMs, this dataset was produced using: 1) a feathering approach to eliminate edge matching errors when one has to manually mosaic two or more 7.5 minute elevation maps; 2) an interpolation algorithm to fill slivers of missing data along edges of mosaiced images; and, 3) a filtering technique to eliminate linear striations on the image, which, unless removed, act as noise and increase the amount of error in the dataset (Gesch et al. 2002). Because USGS-produced DEMs were used as base data for creating NED data, the NED data was deemed to be more accurate. Shuttle Radar Topographic Mapping mission data is also a seamless dataset of the United States. However, within the study area this dataset contained numerous "no data" values and the elevation map was deemed incomplete for the study area.

### *2.3.4 Selection of Best Vegetation Map*

The three approaches for refining vegetation classes by elevation and aspect resulted in overall accuracies >70.0% (Table 2, Figure 4). These maps were: 1) atmospheric haze corrected, with PCA and NDVI used to refine mixed conifer and oak-juniper only (overall accuracy 70.9%, $p$-value < 0.001); 2) atmospheric haze and sun-angled corrected, with PCA and NDVI used to refine mixed conifer and oak-juniper (overall accuracy 71.2%, $p$-value < 0.001); and, 3) atmospheric haze and sun-angled corrected, with PCA and NDVI used to refine mixed coniferous, pine-oak and oak-juniper, except where oak-juniper occurs within the pine-oak elevation and aspect zone (overall accuracy 70.4%, $p$-value < 0.001). Hereafter, these vegetation maps will be referred to as first, second and third vegetation map, respectively. The second vegetation map (listed as number 2 above) was considered the "best" because, with the exception of mixed coniferous, it had the highest user accuracies (mixed coniferous = 67.9 %, pine-oak = 30.8, oak-juniper = 80.9%, grassland/ upland meadow = 73.3% and bugkill = 75.0%). Producer accuracy was highly variable among vegetation types (mixed coniferous = 95.4 %, pine-oak = 10.3, oak-juniper = 72.4%, grassland/ upland meadow = 48.9% and bugkill = 75.0%).

## *2.4 Discussion*

### *2.4.1 Comparison of Vegetation Maps*

The three data products resulting from the refinement procedures were of similar overall accuracies, and the selection of a "best" vegetation map was made based on producer and user accuracies for particular vegetation classes. The main difference between the first and second vegetation maps was the former had a lower user accuracy for mixed conifer (67.9%) and the latter vegetation map had a much higher and acceptable user accuracy (80%). The third vegetation map, where oak-juniper was refined, except where oak-juniper occurred within the pine-oak elevation and aspect zone, appeared to grossly overestimate the occurrence of pine-oak woodlands. This overestimation is not quantifiable. I believe this vegetation map overestimated pine-oak woodland, based on my knowledge of the mountain range, as well as the dynamics of transition zones. Avery and Berlin (1992) and Goodchild (1994) suggest confusion may arise with continuous variation and/ or slow transitions across ecotones. These transition zones may extend hundreds of meters and typically are gently undulating and the boundaries are fuzzy rather than crisp (Tichý 1999).

### *2.4.2 Training Data*

Because of the difficulties in mapping transition zones, large samples of training data for the pine oak woodland would permit better resolution of this vegetation class than I was able to achieve. For the USDA FS Stage 2 Stand Exam data, there were 31 individual study plots for pine-oak woodland. Because many plots of pine-oak were adjacent to other pine-oak plots, I aggregated these plots into nine sample areas. A larger sample sizes for the pine-oak woodlands might have captured a greater amount of this variability, and thus permitted me to better resolve this vegetation class. However, the objective of the USDA FS was not to map pine-oak woodland. Thus, small sample sizes of pine-oak woodland are not the result of an inadequate sample design, but rather the lack of training data available to resolve the pine-oak transition zone.

For the land cover class, "bug kill," this infestation is restricted to the highest elevations surrounding Mount Graham. There are two areas where the *Nepytia* moth infestation occurred. The main outbreak encompassed the summit and surrounding areas; a lesser outbreak occurred to the northeast of Mount Graham. Information provided by USDA Forest Service, Rocky Mountain Research Station researchers combined with photointerpretations of high-resolution DOQQs provided the spectral signatures for the vegetation maps. Because the outbreak was highly confined to a small geographic area (~162 hectares) training data for this class were not separated by great distances. However, delineation of this feature was not a critical component of developing vegetation maps. A simple masking procedure of the *Nepytia* infestation from the mapping of vegetation classes process would have rendered a similar result.

### *2.4.3 Reference Data*

The use of alternative sampling designs for the collection of reference data may improve the accuracy assessment, as well as the power of statistical tests of image classification results. Because some vegetation classes are more abundant than others, a stratified sampling strategy may provide a data set better suited to the estimation of accuracy in each class (Goodchild 1994). Fitzpatrick-Lins (1981) recommends using a GIS to select sample sites using a dual stratification of geography and land cover categories. However, stratified random sampling is often too costly to consider in heterogeneous forests (Skidmore and Turner 1989). Time expenditure in reaching sites delineated in areas of high topographic variability would have resulted in many fewer samples in this study. Due to both budgetary and time constraints, I was forced to use a more conventional, but less ideal, random sampling approach.

An additional limitation to the reference dataset is the sample size. Although there were > 100 samples of mixed coniferous forest and oak-juniper woodland and exactly 45 for grassland/ upland meadow, there were fewer than the recommended 45-50 samples (Hay 1979, Fitzpatrick-Lins 1981, Congalton and Green 1999) for pine-oak woodlands (n=39), grassland/ open meadows and bug kill (n=4). The small sample size for pine-oak woodlands probably contributes to the low user accuracy rating for this vegetation class.

### *2.4.5 Sources of Error*

Image classification error can be caused by mixed pixels (Skidmore et al. 1988, Steele et al. 1998, Goodchild 1994, Price 1994), poor spectral separation (Price 1992, Okin et al. 2001), lack of cross-calibration of sensors on the satellite (Hall et al. 1991), registration error of the imagery (Goodchild 1994), technician bias (Foody 1999), or a combination of these factors. Of the sources of error identified above, the issue of sub-pixel analysis has received considerable attention in the literature. Researchers and users of remotely sensed data are often unaware of the potential problems introduced by the use of the pixel as a unit of analysis (Fisher 1997).

Mixed pixels can occur when multiple vegetation types occur within a pixel or when the appropriate vegetation type occurs, but is obscured by the reflectance value of bare soil (Price 1992). Because distinctions between vegetation types are rarely clearcut (Goodchild 1994), spectra from one type may be similar to a mixture of spectra from other vegetation types (Price 1994), which can further confound the vegetation map development process.

Poor spectral separation between classes may be pronounced when canopy cover is low. In these conditions, the pixel may become dominated by the reflective radiation of the soil, which may swamp the reflective radiation of the dominant vegetation type (Price 1992). Poor separation is particularly common in arid and semiarid environments and, because it is generally characterized by low spectral contrasts, can be difficult to model (Okin et al. 2001). In arid environments, some vegetation types are spectrally indeterminate (Okin et al. 2001). “Hard” image classifications, such as the one presented here, assume each pixel fit conveniently into each of the subjectively established classes (Price 1997). Because landscapes are highly variable, this rarely occurs in nature.

Because the Landsat imagery used was at 30-meter resolution and the vegetation on the Pinaleños Mountains is highly heterogeneous due to topographic variability, spectral mixing of pixels and subsequent errors from pixel mixing likely occurred. Poor spectral separation between lowland oak-juniper (user accuracy = 80.9%) and grassland/ upland meadow (user accuracy = 73.3%) did not appear to significantly influence the overall image classification accuracy. However, classification of the pine oak transition zone was poor (user accuracy = 30.8%). Although spectral confusion of pixels within this transition zone is probable, a small sample size of training sites (n = 9) and a small sample size of pine-oak sample sites for verification (n = 39) also contributed to the poor classification of this vegetation class.

Cross-calibration of sensors requires: 1) standardization of sensors; 2) effects from time-varying gains in sensor electronics and optics; and, 3) changes over time in the processing of ground data (Hall et al. 1991). One can remove the first two effects using the internal calibration sources; however, the sensor’s optical train, which undergoes an unknown amount of degradation with time (Hall et al. 1991), may further confound cross-calibration efforts. Once satellite imagery are captured, these data are then georegistered to the ground; this approach is not perfect and error is inherent

(Goodchild 1994). In practice, resource managers must rely on the operators of the remote sensing platform to ensure proper calibration.

Analyst bias occurs throughout the vegetation map development process. Selection of the vegetation classes and spectral signatures for the supervised classification are subjective. Foody (1999) asserts classification of remotely sensed imagery is subjective and the resultant quality of the vegetation map is highly correlated to analyst decisions and idiosyncrasies. To a certain extent, this is inevitable. However, such biases could be minimized by the adoption of a set of standard practices that would direct work conducted within a particular management area or agency. Vegetation classes were reasonable due to the resolution of Landsat imagery and the availability of training data. Additionally, the availability of high resolution (2-meter) DOQQs and my knowledge of the study area enabled me to effectively resolve oak-juniper (user accuracy = 80.9%) and grassland/ upland meadows (user accuracy = 73.3%), which were supported by the user accuracies of these two vegetation types.

### *2.4.6 Alternative Methods to Classifying Land Cover*

There are several applications for extracting sub-pixel information, as well as variants of more traditional methods, which may serve to resolve some of these issues. These include fuzzy set analysis (Fisher and Pathirana 1990), neural networks (Tatem et al. 2002) and "hardening" the data set by recoding the pixels by dominant class (Foody 1999). In cases where such approaches are not practical, high image classification accuracies may be unobtainable, and managers must judge whether the available vegetation maps are suitable for the intended applications. Furthermore, the use of quantitative accuracy assessments are required to properly judge these vegetation map development procedures.

A multitude of spectral enhancement techniques currently exist, and numerous new techniques are under development. I analyzed the effectiveness of only three in this assessment, although these were carefully selected for their promise and practicality. Another technique, which was not explored here, is *band ratios.* Brown (1994) suggests combining bands 3, 4, and 5 with a *band ratio* of bands 4 and 5, which he found helpful in reducing topographic effects (Brown 1994). Other approaches may be useful in different regions and other vegetation types. At this point, I cannot generalize beyond this case study, except to state similar techniques should be considered for other topographically diverse forested sites.

Finer scale vegetation training data combined with higher resolution data, such as low altitude aerial photography (Price 1994), digital videography or SPOT imagery, may provide a higher accuracy land cover map. A detailed inventory of stand and canopy characteristics is required to improve our understanding of forest ecosystems, as well as for developing methods for classifying and mapping forests at a landscape scale (Treitz and Howarth 2000). However, the cost of imagery increases rapidly with increasing resolution, and costs quickly become prohibitive for most management applications.

The use of multitemporal imagery in classifying land cover can capture interannual variation (i.e., responses of vegetation to annual precipitation or comparing seasonal variations) and, thus, may increase the image classification accuracy (Defries and Townshend 1994). Wallin and others (1992) found significant year by data interaction when modeling vegetation using Advanced Very High Resolution Radiometer data, which they suspect was due to interannual variation in seasonal rainfall. Inclusion of multiple year imagery when interpreting remotely sensed imagery could capture this variability and lead to higher image classification accuracy, particularly in arid regions with high interannual variation in precipitation.

Finally, the usefulness of any single vegetation map, regardless of accuracy, is limited by image resolution (Tatem et al. 2002). Because ecological systems often operate at multiple scales, it is often difficult to identify a single resolution which will is most suitable for resolving differences in vegetation land cover (Treitz and Howarth 2000). In practice, remote sensing applications in resource management tend to utilize inexpensive or free remote sensing data (such as Landsat imagery), which are usually collected at one spatial resolution. Managers must be cognizant of this to use the resultant data products appropriately. Theoretically, interpretation at multiple scales, and analysis of the differences images analyzed at multiple scales, can improve resolution and expand the utility of remotely sensed data, but this remains a challenging endeavor (Treitz and Howarth 2000). Moreover, from a practical perspective, the expense of collecting training and reference data at multiple scales is generally prohibitive. However, future advances in multi-scale spatial analyses will likely resolve some of these difficulties and, ultimately, provide managers with more flexible data products to guide management decisions.

### 2.5 *Conclusion*

Because most landscapes are highly heterogeneous, especially in areas with high topographic variability, the spectral enhancement techniques employed here should not be seen as a prescription to be followed uncritically for supervised classifications of vegetation in other regions. However, the improvements in the vegetation maps resulting from these techniques suggest both pre- and post-processing techniques should be investigated in other vegetation mapping efforts. The accuracy improvement in this study were due, in large measure, to spectral correction techniques and the refinement of vegetation class by elevation and aspect, but accuracy was also the result of the relatively simple land cover classes selected *a priori*, the availability of training data, and the ability to resolve low elevation vegetation using photointerpretations. The combination of atmospheric haze and sun-angle correction with a using combined bands 1, 2 and 3 $PC_1$ and bands 5 and 7 $PC_1$, band 4 and an NDVI worked well for this study area. Results from this study indicate the use of atmospheric haze and sun angle correction algorithms greatly improved the vegetation map accuracy of land cover in the Pinaleños Mountains. A similar approach may serve to map land cover of other sky islands in southeastern Arizona, southwestern New Mexico and northern Mexico, and these techniques should be examined, in concert with other spectral enhancement techniques, for classifying land cover in other areas with high topographic variability. Furthermore, for any vegetation mapping project, appropriate application of spectral correction and enhancement techniques offer promise for improving management utility in a cost-effective manner. Of course, none of this could be achieved or properly evaluated without quantitative accuracy assessments, which must become a standard part of any application of remotely sensed data to real-world management issues.

### *2.6 Management Implications*

These management implications are derived from "lessons learned" during this study. This information is presented in a checklist format for clarity and easy reference.

(1) Clearly define the objectives of the vegetation mapping effort.

(2) Identify *a priori* the vegetation classes to be classified. These should be selected based on the resolution of the remotely sensed data use in the vegetation mapping effort and the availability of training data.

(3) Ensure (1) and (2) are in agreement.

(4) If available, use remotely sensed data from multiple years when classifying land cover. This will help avoid biases due to interannual variability of the system of study.

(5) Evaluate spectral enhancement techniques and identify those most applicable to the study site and objectives of the vegetation mapping effort.

(6) Correct remotely sensed imagery for haze and sun-angle.

(7) If training data is available, ensure these data meet the needs of the vegetation mapping effort. If not, the training data should be collected within a sampling effort designed to meet vegetation map needs.

(8) Employ a competing models approach for vegetation maps developed under different assumptions and with different analytical techniques. This will assist in identifying the best-available techniques are evaluated and the "best" resulting vegetation map selected.

(9) Use a multi-criteria vegetation map selection process. Select the "best" vegetation map only after assessing overall accuracy, statistical significance of vegetation map, and the producer and user accuracies of each vegetation class.

### *2.7 Literature Cited*


Anhold, J., G. Asmus, R. Smith and B. Lefevre. 2003. Status Paper: Mount Graham Forest Insect Outbreaks, Coronado National Forest, May 16, 2003. Coronado National Forest, Safford Ranger District, Safford, Arizona. Pps3.

Avery, T.E. and G.L. Berlin 1992. Fundamentals of remote sensing and aerial interpretation, fifth edition. Macmillan Publishing Company, New York. Pps 472.

Birky, A.K. 2001. NDVI and a simple model of deciduous forest seasonal dynamics. Ecological Modeling 143: 43-58.

Brown, D.G. 1994. Predicting vegetation types at treeline using topography and biophysical disturbance variables. Journal of Vegetation Science, 5: 641-656.

Cardillo, M., D. Macdonald, S. Rushton 1999. Predicting mammal species richness and distributions: testing the effectiveness of satellite-derived land cover data. Landscape Ecology 14: 423-435.

Carlson, T.N. and Ripley, D.A. 1997. On the relation between NDVI, fractional vegetation cover, and leaf area index. Remote Sensing and Environment 62: 241-252.

Chavez, P.S. 1992. Comparison of spatial variability in visible and near-infrared spectral images. Photogrammetric Engineering and Remote Sensing 58: 957-964.

Chavez, P.S. 1996. Image-based atmospheric haze corrections – revisited and improved. Photogrammetric Engineering and Remote Sensing 62: 1025-1036.

Congalton, R.G. 1991. A review of assessing the accuracy of classifications of remotely sensed data. Remote Sensing of Environment 7: 33-46.

Congalton, R. and K. Green. 1999. Assessing the accuracy of remotely sensed data: principles and practices. Lewis Publishers, Boca Raton. Pps. 137.

Crist, E.P., R. Laurin and R.C. Cicone 1986. Vegetation and soils information contained in transformed thematic mapper data. *In* Proceedings of IGARSS' 1986 Symposium, 1465-70. Ref. ESA SP-254. Paris: European Space Agency. Pps6.

DeFries, R.S., and Townshend, J.R.G. 1994. NDVI-derived land cover classifications at a global scale. International Journal of Remote Sensing 15: 3567-3586.

ERDAS 2001. ERDAS Imagine Field Guide, 5th Edition: Revised and Expanded. Atlanta, Georgia pps. 698.

Fisher, P.F. and S. Pathirana. 1990. The evaluation of fuzzy membership of land cover classes in the surburban zone. Remote Sensing of Environment. 34: 331-347.

Fisher, P.F. 1997. The pixel: a snare and a delusion. International Journal of Remote Sensing 18: 679-685.

Fitzpatrick-Lins, K. 1981. Comparison of sampling procedures and data analysis for a land-use and land-cover map. Photogrammetric Engineering and Remote Sensing 47: 343-351.

Foody, G.M. 1999. The continuum of classification fuzziness in thematic mapping. Photogrammetric Engineering and Remote Sensing 65 (4): 443-451.

Franco-Lopez, H. A.K. Ek and M.E. Bauer 2001. Estimation and mapping of forest stand density, volume and cover type using *k*-nearest neighbors method. Remote Sensing of Environment 77: 251-274.

Gesch, D, M. Oimoen, S. Greenlee, C. Nelson, M. Steuck and D. Tyler 2002. The National Elevation Dataset. Photogrammetric Engineering and Remote Sensing 68 (1): 5-11.

Goodchild, M.F. 1994. Integrating GIS and remote sensing for vegetation analysis and modeling: methodological issues. Journal of Vegetation Science 5: 615-626.

Graetz, R.D. 1990. Remote sensing of terrestrial ecosystem structure: an ecologists pragmatic view. In Remote sensing of biosphere functioning (R.J. Hobbs, H.A. Mooney, eds.) Springer, New York, pps. 5-30.

Hadley, K.S. 1994. The role of disturbance, topography and forest structure in the development of a montane forest landscape. Bulletin of the Torrey Botanical Club 121: 47-61.

Hay, A.M. 1979. Sampling designs to test land-use map accuracy. Photogrammetric Engineering and Remote Sensing 45: 529-533.

Hess, G.R. 1994. Pattern and error in landscape ecology: a commentary. Landscape ecology 9: 3-5.

Hess, G.R. and J.M. Bay 1997. Generating confidence intervals for composition-based landscape indexes. Landscape Ecology 12: 309-320.

Huang, C., Wylie, B., Homer, C., Yang, L., and Zylstra, G., 2002. Derivation of a Tasseled cap transformation based on Landsat 7 at-satellite reflectance: International Journal of Remote Sensing 23: 1741-1748.

Hurcom, S.J. and Harrison, A. R. 1998. The NDVI and spectral decomposition for semi-arid vegetation abundance estimation. International Journal of Remote Sensing 19: 3109-3125.

Jenness, J. 2001. Surface areas and ratios extension. Jenness Enterprises, Inc. http://www.jennessent.com/arcview/surface_areas_ratios.htm Date accessed: 25 March 2003.

Jensen, J. R. 1996. Introductory digital image processing: a remote sensing perspective. Second edition. Prentice Hall International, London. Pps316.

Jones, K. 1998. A comparison of two approaches to ranking algorithms used to compute hill slopes. Geoinformatica 2 (3): 235-256.

Lins, K.F. and R.L. Kleckner, 1996. *In* Gap Analysis: A landscape approach to biodiversity planning (Scott, J., T. Hart and F. Davis, editors). American Society for Photogrammetry and Remote Sensing, Bethesda, Maryland. Pps. 57-65.

Luck, M. and J. Wu 2002. A gradient analysis of urban landscape pattern: a case study from the Phoenix metropolitan region, Arizona, USA. Landscape Ecology 17: 327-339.

Nuñez, M. 1980. The calculation of solar and net radiation in mountainous terrain. Journal of Biogeography 7: 173-186.

Okin, G.S., D.A. Roberts, B. Murray and W.J. Okin 2001. Practical limits on hyperspectal vegetation discrimination in arid and semiarid environments. Remote Sensing of Environment 77: 212-225.

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

Price, J.C. 1992. Estimating vegetation amount from visible and near infrared reflectances. Remote Sensing of Environment 41: 29-34.

Price, J.C. 1994. How unique are spectial signatures? Remote Sensing of Environment 49:181-186.

Ricotta, C., G.C. Avena and F. Volpe 1999. The influence of principal components analysis on the spatial structure of a multispectral dataset. International Journal of Remote Sensing 20: 3367-3376.

Schmidt, K.S. and A.K. Skidmore 2003. Spectral discrimination of vegetation types in a coastal wetland. Remote Sensing of Environment 85: 92-108.

Sellers, P.J. 1985. Canopy reflectance, photosynthesis and transpiration. International Journal of Remote Sensing 6: 1335-1372.

Sellers, P.J. 1987. Canopy reflectance, photosynthesis and transpiration II. Remote Sensing of Environment 21: 143-183.

Skidmore, A.K., G.W. Forbes and D.J. Carpenter 1988. Non-parametric tests of overlap in multispectral classification. International Journal of Remote Sensing 9 (4): 777-785.

Skidmore, A.K. and Turner, B.J. 1989. Assessing the accuracy of resource inventory maps. In H.G. Lund (ed.), *Proceedings of Global Natural Resource Monitoring and Assessments, Vol. 2* (524-535), Baltimore: ASPRS.

Steele, B.M., J.C. Winne and R.L. Redmond. 1998. Estimation and mapping of misclassification probabilities for thematic land cover maps. Remote Sensing of Environment 66: 192-202.

Tatem, A.J., H.G. Lewis, P.M. Atkinson and M.S. Nixon. 2002. Super-resolution land cover pattern prediction using Hopefield neural network. Remote Sensing of Environment 79: 1-14.

Tichý, L. 1999. Predictive modeling of the potential natural park vegetation pattern at Podyjí National Park, Czech Republic. Folia Geobotanica 34: 243-252.

Townshend, J.R.G. 1992. Land cover. International Journal of Remote Sensing 13: 1319-1328.

Treitz, P. and P. Howarth 2000. High spatial resolution remote sensing data for forest ecosystem classification: an examination of spatial scale. Remote Sensing of Environment 72: 268-289.

Tucker, C.J. 1978. A comparison of satellite sensor bands for vegetation monitoring. Photogrammetric Engineering and Remote Sensing. 44: 1369-1380.

United States Department of Agriculture, Forest Service (USDAFS) 1997a. Plant Associations of Arizona and New Mexico, Volume 2: Woodlands. USDAFS, Southwestern Region, Habitat Typing Guide. Washington, D.C. pps196.

United States Department of Agriculture, Forest Service (USDAFS) 1997b. Plant Associations of Arizona and New Mexico, Edition 3. Volume 1: Forests. USDAFS, Southwestern Region, Habitat Typing Guide. Washington, D.C. pps291.

United States Geological Survey (USGS) 1991. Remote Sensing Field Guide-Desert. Desert Processes Working Group. ETL-0588. U.S. Geological Survey, Flagstaff, Arizona.

Wallin, D.O., C.C.H. Elliott, H.H. Shugart, C.J. Tucker and F. Wilhelmi. 1992. Satellite remote sensing of breeding habitat for an African weaver-bird. Landscape ecology 7: 87-99.

### *2.8 Tables*

Table 1. These vegetation classes for supervised classifications were not refined by elevational and aspect. Overall accuracy, classification statistical significance (*p*-value), and user and producer accuracies per vegetation class for all six supervised classifications.

| | Uncorrected with [1]PCA and [2]NDVI layers | Uncorrected with PCA and [3]TCT layers | Atmospheric haze corrected with PCA and NDVI layers | Atmospheric haze corrected with PCA and TCT layers | Atmospheric Haze and Sun-angle corrected with PCA and NDVI layers | Atmospheric Haze and Sun-angle corrected with PCA and TCT layers |
|---|---|---|---|---|---|---|
| *p*- value | <0.001 | <0.001 | <0.001 | <0.001 | <0.001 | <0.001 |
| Overall Accuracy | 49.4 | 49.4 | 52.0 | 47.7 | 55.8 | 47.3 |
| User Accuracy | | | | | | |
| Oak-juniper | 62.6 | 62.4 | 70.7 | 71.5 | 69.1 | 54.5 |
| Pine-oak | 74.5 | 74.4 | 66.7 | 15.4 | 20.5 | 5.1 |
| Mixed-conifer | 34.6 | 34.6 | 34.5 | 36.8 | 57.1 | 45.9 |
| Grassland/ meadow | 31.1 | 31.1 | 37.8 | 37.8 | 46.7 | 64.4 |
| Bug kill | 100 | 100 | 75.0 | 100 | 50.0 | 100 |
| Producer Accuracy | | | | | | |
| Oak-juniper | 70.6 | 70.6 | 71.3 | 68.2 | 65.9 | 61.5 |
| Pine-oak | 21.5 | 21.5 | 20.8 | 8.1 | 14.0 | 4.1 |
| Mixed-conifer | 69.7 | 69.7 | 76.7 | 58.3 | 62.8 | 60.4 |
| Grassland/ meadow | 87.5 | 87.5 | 73.9 | 70.8 | 70.0 | 50.0 |
| Bug kill | 22.2 | 22.2 | 21.4 | 12.1 | 28.6 | 14.8 |

[1]PCA- principal components analysis
[2]NDVI – normalized difference vegetation index
[3]TCT- tasseled cap transformation

Table 2. These are the final classifications, which had the highest accuracies. Vegetation classes for supervised classifications were refined using various elevational and aspect constraints identified by vegetation associations. Overall accuracy, classification statistical significance (*p*-value), and user and producer accuracies per vegetation class for all six supervised classifications.

| | Atmospheric haze corrected with PCA and NDVI layers, refined for elevation and aspect of mixed conifer and oak-juniper only | Atmospheric haze and sun-angle corrected with PCA and NDVI layers, refined for elevation and aspect of mixed conifer and oak-juniper only | Atmospheric haze and sun-angle corrected with PCA and NDVI layers, refined for elevation and aspect of mixed conifer, pine-oak and oak-juniper except where oak-juniper occurs within pine-oak elevation and aspect zone |
|---|---|---|---|
| *p*- value | <0.001 | <0.001 | <0.001 |
| Overall Accuracy | 70.9 | 71.2 | 70.4 |
| User Accuracy | | | |
| Oak-juniper | 79.8 | 80.9 | 80.9 |
| Pine-oak | 33.3 | 30.8 | 30.8 |
| Mixed-conifer | 73.8 | 67.9 | 80.0 |
| Grassland/ meadow | 63.3 | 73.3 | 73.3 |
| Bug kill | 50.0 | 75.0 | 75.0 |
| Producer Accuracy | | | |
| Oak-juniper | 74.0 | 72.4 | 72.4 |
| Pine-oak | 25.6 | 10.3 | 51.3 |
| Mixed-conifer | 91.0 | 95.5 | 81.2 |
| Grassland/ meadow | 42.2 | 48.9 | 48.9 |
| Bug kill | 75.0 | 75.0 | 75.0 |

## *2.9 Figures*

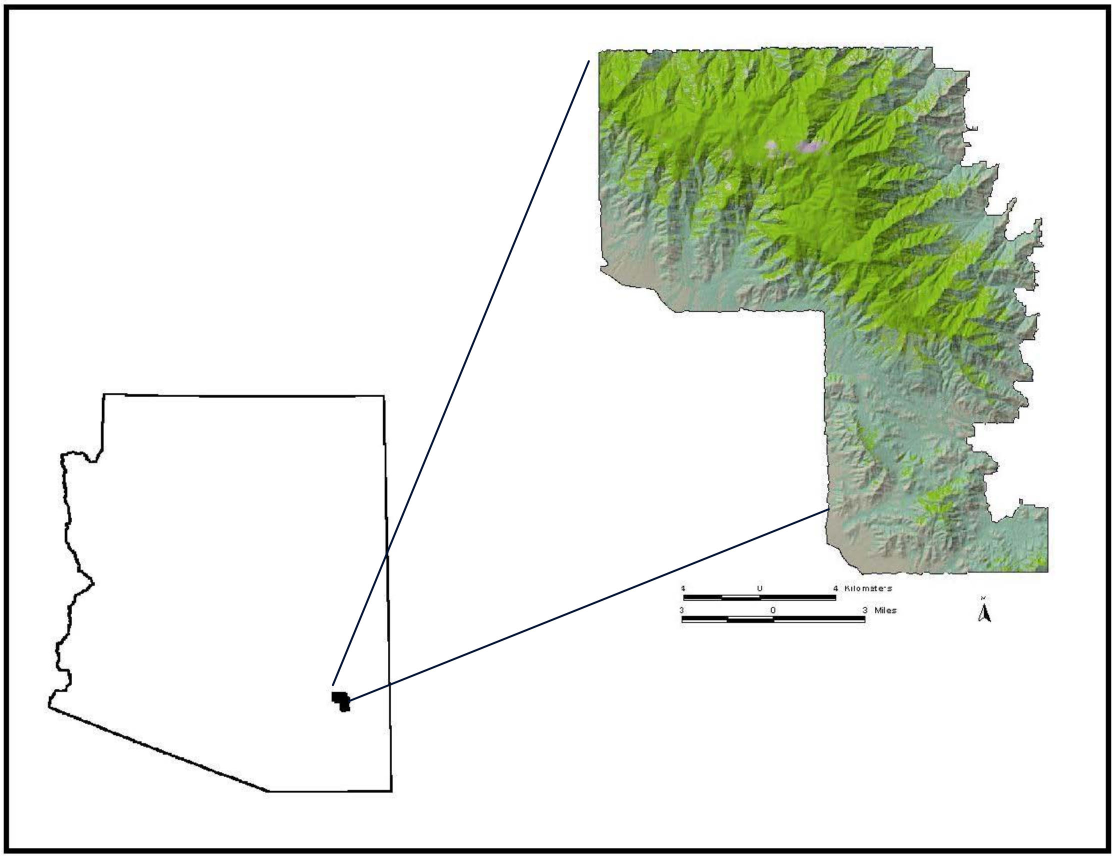


Figure 1. Locator map of the Pinaleños Mountains, southeastern Arizona.

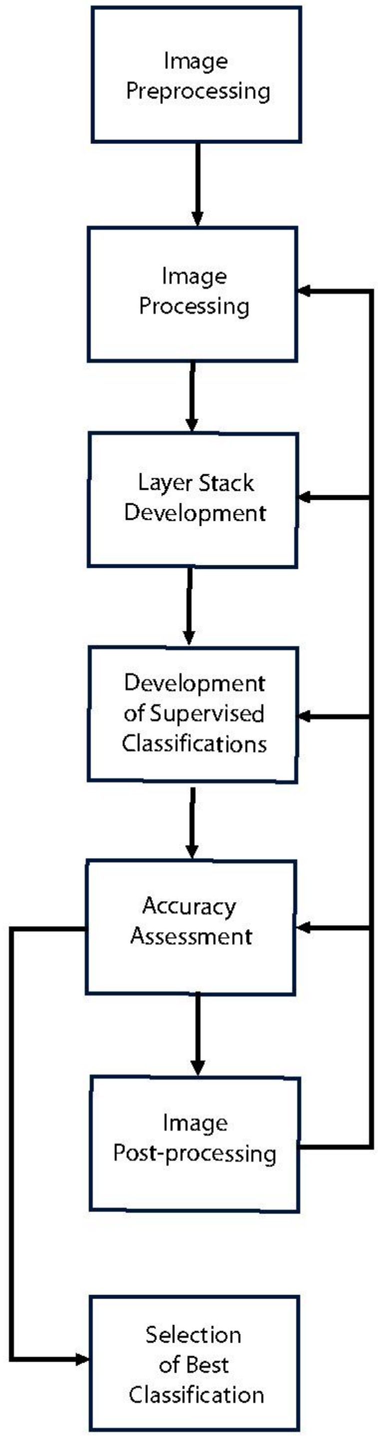


Figure 2. A schematic flow diagram of the modeling process. This process flows from top to bottom and consists of an image preprocessing, image processing, development of layer stacks, development of supervised classifications, accuracy assessment, image post-processing, a feedback loop to a second accuracy assessment, and selection of best model. Additional feedback loops may occur throughout this process; however, these additional feedback loops were not needed in this study.

| | | Reference Data | | |
|---|---|---|---|---|
| | | Observed Vegetated | Observed Non-vegetated | |
| Classification Data | Predicted Vegetated | *A* | *B* | *C* |
| | Predicted Non-vegetated | *D* | *E* | *F* |
| | | *G* | *H* | *I* |

Figure 3. Simplified classification table, using two hypothetical classes, to describe the agreement between observed and predicted values and calculating overall accuracy, and producer and user accuracy.

Overall Classification Accuracy = $(A+E)/I$

Producer Accuracy – Vegetated Class = $A/G$, Not Vegetated = $E/H$

User Accuracy – Vegetated Class = $A/C$, Not Vegetated = $E/F$

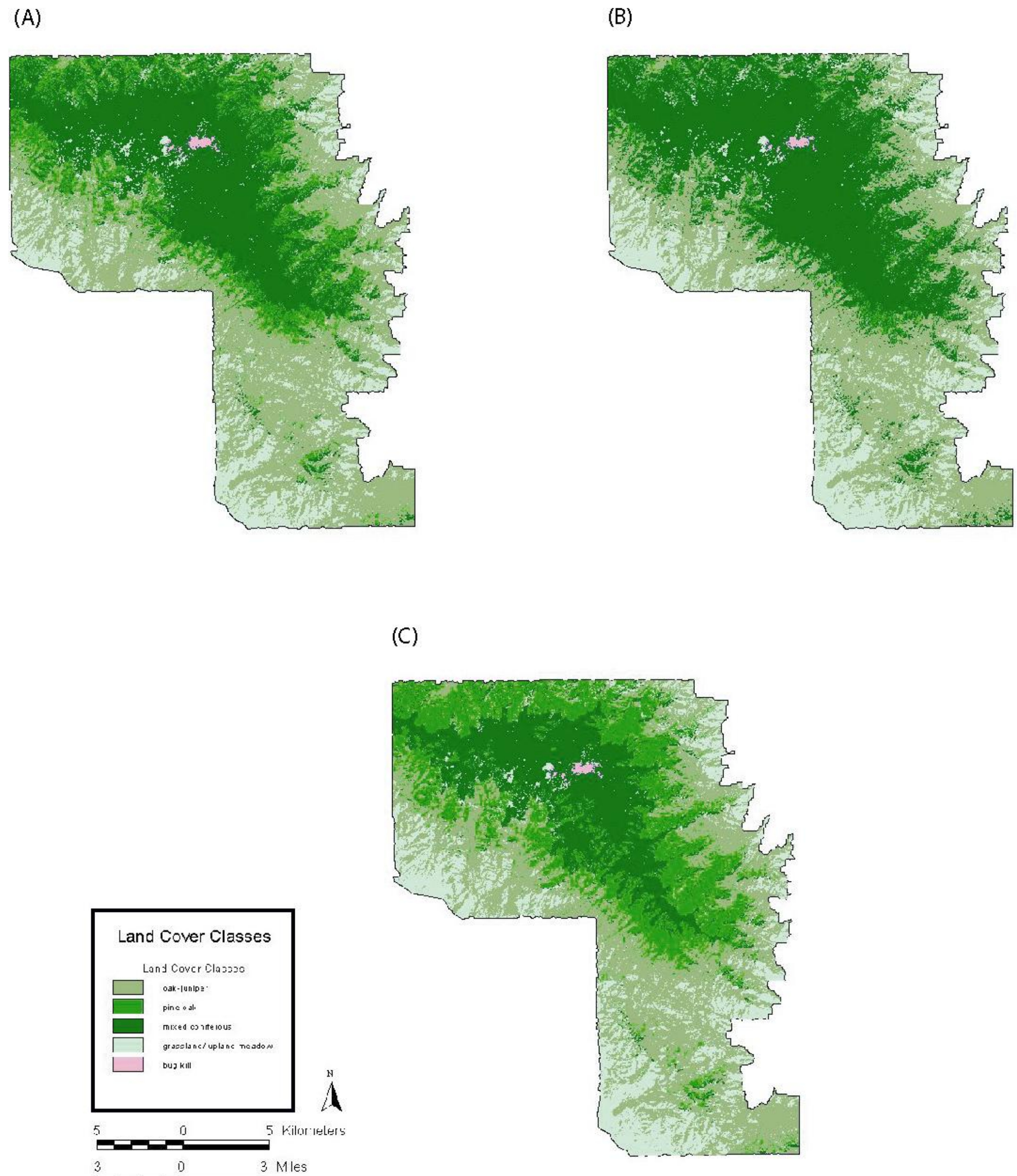


Figure 4. The three “best” land cover classification maps. (A) Haze-corrected with PCA and NDVI refined mixed coniferous and oak-juniper only (overall accuracy 70.9%, $p$-value < 0.001), (B) Haze and sun-angled corrected with PCA and NDVI refined mixed coniferous and oak-juniper only (overall accuracy 71.2%, $p$-value < 0.001), and, (C) Haze and sun-angled corrected with PCA and NDVI refined mixed coniferous, pine-oak and oak-juniper except where oak-juniper occurs within the pine-oak elevation and aspect zone (overall accuracy 70.4%, $p$-value < 0.001).

## *3. A Landscape-scale Approach for Predicting Songbird Occurrence: An Evaluative Criteria for Selecting Models for Wildlife Management*

### *3.1.1 Abstract*

Wildlife-habitat relationship models are employed routinely in making resource-management decisions. Understanding and identifying potential sources of model error is imperative to providing resource managers with the highest quality habitat models. A three-year dataset (1993-1995) of bird survey points, habitat information derived from literature, and coarse landscape-scale variables were used to develop models. Models were validated using a 2002 dataset of bird presence from point-counting stations. Coarse-scale landscape variables included slope, aspect, elevation, vegetation type, and distance to springs and streams. Using a competing models framework, I modeled habitat at the landscape-scale using classification tree and logistic regression models for eight songbird species on the Pinaleños Mountains, southeastern Arizona. Classification tree output and literature-derived information were used for creating predicted distribution maps with a GIS, accuracy assessed using the 2002 dataset and a Cohen's Kappa, and selected using a multi-criteria selection approach. A stepwise logistic regression and literature-derived information forced regression were validated using a regression equation; the best-fitted model was selected using a combined Akaike's Information Criterion (AIC) and multi-criteria selection approach. I tested and/ or evaluated all datasets used in all phases of the modeling process. GIS information were considered of the highest quality and finest scale available. The best available elevation model was used for elevation, slope and

aspect, vegetation land cover (overall accuracy = 71.2%), and distance to springs and streams maps were used. Additionally, the verification dataset was collected during the 2002 drought. Although none species' models attained 80% accuracy, six of eight yielded accuracy values better than chance and comparable to other studies using similar habitat variables. Sample sizes were considered low in statistical power (< 30 samples for presence and absence), the sample design was not appropriate for landscape-scale habitat modeling, and interannual variation occurred for the bridled titmouse across years for the elevation variable. Low predictive success of these models was probably due to a combination of inappropriate study design, small sample size, environmental stochasticity in the verification dataset, and lack of fine-scale GIS information. Although use of these models in guiding management decisions is considered limited, the criteria developed provides a systematic framework for evaluating data quality for modeling wildlife-habitat relationships. I recommend using a method, which includes elements identified here for evaluating model datasets for the potential errors and potentially reduce the extent of error propagation. This will ultimately provide natural resource managers with higher quality wildlife-habitat relationship models.

### *3.1.2 Introduction*

Spatially explicit predictive models are an important tool for understanding wildlife-habitat relationships and guiding natural resource management decisions (Stoms et al 1992, Pearce and Ferrier 2000, Penhollow and Stauffer 2000, Wright et al. 2000, Brugnach et al. 2003). For predictive models to be a useful tool in the decision making process, they must be accurate, general, and easy to apply (Van Horne and Weins 1991). Bolger et al. (1997) suggest modeling wildlife-habitat relationships at the landscape scale may actually be more appropriate because land-management decisions are made at the landscape scale. When GIS information are available, landscape scale models are often simple, inexpensive to generate and may provide information useful within a management context.

However, prior to accepting a habitat model and its output *prima facie*, data used in creating and validating the model must be evaluated for its accuracy and appropriateness. Assessment of these data may reduce some of the model uncertainties and error. Failure to account for potential sources of error can lead to inaccurate results and interpretations. When management decisions are based on models whose data are not thoroughly evaluated, wrong decisions may be made.

Generally, the four types of information can be used in spatially explicit modeling. These are retrospective, verification data (*sensu* Carrol et al. 1999), literature-derived and GIS information. Retrospective data are used to build the models. Verification data are used for testing model predictions. GIS information were also used build models and test model predictions. Each of these data were evaluated for data quality prior to modeling. Retrospective and verification data should be tested and evaluated for interannual variability, environmental stochasticity, appropriate study design, adequate sample size and spatial autocorrelation. GIS information were verified using photointerpretation techniques; additionally, multiple land cover maps and digital elevation data were compared and the best dataset was selected.

Retrospective and verification data are used for generating and testing the model, respectively. When either dataset represents multiple years interannual variability must be considered before pooling data across years. Failure to assess a multi-year dataset for interannual variation can affect statistical analyses and model inference because annual variation in habitat use is expected for most terrestrial vertebrates (Schooley 1994). Climatic factors vary annually resulting in variability in available food resources and microclimates. This results in temporal fluctuations in population densities, which may alter the spatial distribution of occurrence (Van Horne 1983). Environmental stochasticity includes perturbations such as fire, inclement weather and anthropogenic habitat disturbance. An organism may respond to such an event by temporarily shifting how it selects habitat (Morrison et al. 1998). When this occurs, and models are based on these data, inaccurate model predictions may result (Gutzwiller and Barrow 2001). Appropriate study design and adequate sample size are usually coupled. When a study is designed for a specific research question, the appropriate sample size required to answer the question is generally a criteria for the study design. Low statistical power (Rao 1998) and poor model performance (Dettmers and Bart 1999) can result when the study design is not tailored to the research question and inadequate sample sizes are used. Datasets not designed to address the specific research question may result in undersampling and contribute to overall model error (Edwards et al. 1996). Spatial autocorrelation occurs when samples occur in a regular pattern in geographical space, which result in a significant amount of redundancy across samples (Odland 1988). Thus, the closer two observations are to one another, the more likely they are to be similar. By measuring the degree of similarity between samples, a test of spatial autocorrelation provides a measurement of overall independence between sites. The assumption of sample independence is violated when spatial autocorrelation is significant.

Because collecting field data for building models is time, labor, and monetarily intensive, many wildlife-habitat spatially explicit models are often built using information derived from the peer-reviewed literature (Stoms 1992, Clark et al. 1993, Scott et al. 1996, Merrill et al. 1999, Wisdom et al. 2000). Perhaps the most well known expert opinion/ literature search based project is the GAP Analysis Program, which is currently providing distribution maps of all vertebrate species in the United States (Scott et al. 1996). Errors associated with using peer-reviewed literature may occur by combining multiple studies of habitat use or selection, which used divergent data collection methods.

GIS information used in spatially explicit modeling (e.g., elevation, topographic derivatives, point, line and polygon locations of water sources, human habitation areas, etc.) should be evaluated for quality and selected judiciously. A review of the metadata file associated with the GIS file can provide information pertaining to how the GIS was created and may provide an accuracy rating of the file. If the information represents points (e.g., springs, buildings, mines) or lines (e.g., roads or streams), the positional accuracy of these data can be tested by overlaying these data on high-resolution aerial photographs. Although GIS information have become easy to generate and user control has improved due to increasingly user-friendly software programs, these data are not always of sufficient quality (Lane and Chandler 2003). MacKinnon and DeWulf (1994) found vegetation classification maps based upon LANDSAT TM imagery was limited due to the coarse scale of the imagery. Sources of error in land cover maps may be attributed to topographic variability, landscape complexity and land-use patterns (Steele et al. 1998). Jones (1998) found differences in the performance of various algorithms used to derive slope from an elevation model. Although some performed better than others, each algorithm produced varying degrees of error. Even when the "best" method for creating GIS information is selected, calculations, interpolations and combinations of map data are less exact than the original map layer (Guisan and Zimmerman 2000). However, a complete understanding of data quality and determining how errors propagate across a several layer GIS-based habitat map is often beyond our capabilities (Chrisman 1987, Stoms et al. 1992). Subsequently, errors will always occur in GIS information and these errors will likely propagate in a multiple-layer GIS model regardless of data quality. Interestingly, despite data errors and subsequent error propagation in GIS information, users rarely understand or address the limitations and quality of these datasets (Chrisman 1987). By critically selecting GIS information and evaluating those data for potential sources of error, one can maximize data quality and thus reduce the amount of uncertainty in a GIS-based model.

The purpose of this study is to develop an evaluative criteria to assist resource managers in processing, evaluating and developing data used in building and testing predictive wildlife-habitat relationship (WHR) models. I used eight passerine bird species, available GIS information, and a three-year (1993-95) retrospective dataset and a one-year (2002) verification dataset of bird presence data for the Pinaleño Mountains, southeastern Arizona to develop and test these criteria. Specifically, I: (1) evaluated the quality of retrospective and verification data, as well as GIS information; (2) compared results of parametric and nonparametric mechanistic WHR models to literature-derived information models; and, (3) evaluated the usefulness of model predictions to land managers.

### *3.1.3 Study Area*

The Pinaleños Mountains are located in Graham County, Arizona, approximately 200km north by northeast of Tucson, Arizona (Figure 1). Delineating the northern extent of the Madrean Archipelago, this sky island is managed by the USDA -Forest Service Coronado National Forest, Safford Ranger District. The planametric area of the study area is approximately 410 $km^2$ and its surface area, adjusted to account for the varied topography (*sensu* Jenness 2001), is approximately 470 $km^2$. The area was defined by a lower elevational limit of 1455 m, which corresponded to the lower limits of the avian research efforts (W. Block, *unpublished data*).

### *3.2 Methods*

#### *3.2.1 Model Assumptions*

I used the following assumptions in the modeling effort:

1) Because of limited samples, "presence" for each species was defined by one or more observations at a point.
2) To prevent overfitting the models of all species, subsets of the verification dataset were created depending on the primary vegetation type(s) where the species occurs. For example, points occurring in the mixed-coniferous forest were removed from Mexican jay (*Aphelocoma ultamarina*) models because the known natural history characteristics suggest this species rarely occurs in this type.
3) The GIS information used in the modeling effort will adequately capture landscape-scale habitat selection of passerine birds.

#### *3.2.2 Data Used in Model Development*

##### *3.2.2.1 Modeling Process Description*

The modeling process was comprised of five general components: (1) evaluation of data used in model development; (2) development of habitat models; (3) assessment of model accuracy; (4) selection of best models; and, (5) interpretation of model results (Figure 2). Specifically, this process occurred as follows:

1) Evaluation and selection best GIS information layers – All known available GIS information for the study area were obtained. These data were evaluated and the best data layers were selected. Retrospective data were evaluated for appropriateness of use in landscape level habitat use models. Tests for interannual variation were conducted on the retrospective data.

2) Development of habitat models – This occurred as two separate procedures.

   Procedure 1: Classification tree analyses of retrospective data were conducted and variables strongly correlated with habitat use for each species were identified. Habitat variables were identified from the literature. Using these variables, predictive distribution maps were developed using ArcView 3x. Accuracy assessments were conducted and the best models were selected.

   Procedure 2: For each species, I conducted three different logistic regression procedures. Akaike's Information Criterion (AIC) was used to select the best model. An accuracy assessment was conducted.

3) Assessment of model accuracy – Using information derived from CART and the literature GIS-based distribution maps were built and tested against the verification data using a Cohen's Kappa statistic. Logistic regression models were tested using the verification data with the regression equation of the best regression model. Error matrices were developed.

4) Selection of best overall model – By comparing the accuracy of both CART/GIS-based and logistic regression models, the model with the best performance was selected.

5) Interpretation of results – The results of the evaluations of both GIS and retrospective data, as well as overall model selection and performance were interpreted and conclusions were drawn. Inferences were made regarding the potential influences on model performance.

#### *3.2.2.2 GIS Information*

Because time and cost effectiveness are paramount in resource management, devising a system for producing maps quickly and economically using easily accessible data is highly desirable (Cardillo et al. 1999). For the Pinaleños Mountains, only landscape-level data were readily available. I selected this coarse-scale assessment to 1) utilize data commonly available to land managers and 2) test the accuracy of landscape-scale habitat models. These variables are: vegetation land cover type (oak-juniper woodland, pine-oak woodland, mixed coniferous forest, grassland/ upland meadow), elevation (meters), slope (percentage), aspect (degrees), and distance to springs and streams (meters; Table 1; Figure 3). Minimum mapping unit was 30 meters and data were projected in UTM, Zone 12, and NAD 1927.

The vegetation land cover map was generated via supervised classification of a LANDSAT 7 ETM+ imagery (Bands 1-3, 5 and 7; capture date: 12 November 1999). A USDA Forest Service Stage 2 Stand Exam and photointerpretation of 2-meter resolution digital orthophoto quads was used for creating the training dataset. A stage 2 stand exam is a detailed inventory of vegetation using the USDA FS plant association classifications for forests and woodlands (USDA 1997a, 1997b). I classified this land cover map using four coarse-scale classes (grassland/ upland meadow, oak-juniper woodland, pine-oak woodland and mixed-coniferous forest, and bug kill). Low elevation grassland and high elevation (upland) meadows have similar spectral signatures and were combined. Oak-juniper woodland occur between 1270 and 1640 meters elevation without aspect influence and 1270 to 1970 between aspects of 135 to 225 degrees (USDA 1997b); species include *Quercus arizonica, Q. emoryi, Q. turbenella, Q.* hybrid, *Juniperus deppeana, J. monosperma,* and *J. osteosperma*. Pine-oak woodland occur between 1640 and 2420 meters elevation without aspect influence and 1970 to 2850 meters elevation between aspects of 135 to 225 degrees (USDA 1997b); dominant tree species include *Pinus leiophylla, P. ponderosa, Quercus hypoleucoides, Q. rugosa* and *Q.* hybrid. Mixed -coniferous forest occur at elevations ≥ 2430 meters without aspect influence and and ≥ 2850 meters for aspects between 135 to 225 degrees (USDA 1997a); dominant woody vegetation includes *Abies* spp., *Picea pungens, Pinus strobiformis* and *Pseudotsuga menziesii*. The vegetation class defined as "bug kill" was located at the upper elevations at and around the summit of Mount Graham. From 1997 to 1998, an approximate 400-acre area of spruce-fir forest was devastated by a defoliating looper, (geometrid moth, *Nepytia janetae*, Anhold et al. 2003). In 1998, an exotic spruce aphid (*Elatobium abietinum*) was discovered on the mountain range. These infestations combined have resulted in an area, which a distinct signature resolvable from uncorrected and unenhanced satellite imagery. During 1999, the year the Landsat ETM+ imagery was captured, the lower elevational limit of the insect damage was approximately 3015 meters (Steve Dudley, USDA FS, Rocky Mountain Research Station, Flagstaff Lab, personal communication). An independent dataset, collected in June-August 2001, was used for model verification. This dataset, collected in July – August 2001 (representing 344 sample points), were collected at random across the four vegetation classes. Sample points were collected, at one-half kilometer intervals along traversable roads and hiking trails spanning the elevational gradient of the mountain range. Vegetation type was defined by visual assessment. At each sample site, I estimated a circle of sixty-meters centered on the point, and divided the plot into four quadrants based on cardinal directions. For each quadrant I estimated the percent cover, for each of five vegetation types. Cover estimates for each of the quadrants were summed for each sample location. A plot containing ≥ 75% of one of the five vegetation types was assigned to that type (e.g., if the summed quadrants of a sample site was summed to 85% oak juniper, the sample site was classified as oak juniper). Due to the coarseness of this assessment, this method was acceptable (G. Lennis Berlin, personal communication). Map accuracy was calculated by Cohen's Kappa statistic, and errors of commission and omission (Congalton and Green 1999).

I evaluated three elevation grids (two elevation models and a radar image of elevation). These models were Digital Elevation Model (DEM), Shuttle Radar Topographic Mapping Mission data (SRTM) and National Elevation Dataset (NED). I reviewed the metadata associated with each dataset and analyzed each for "no data values" (i.e., empty cells with "0" value) within each layer to determine which was most accurate. Once the best dataset was selected, I used ERDAS Imagine 8.5 to derive slope and aspect grids. I did not consider the algorithms used for generating these grids as a source of potential error. Jones (1998) concluded the creation of the best slope grid, regardless of the algorithm used, and depended on the quality of the elevation model. I extrapolated this statement to include aspect because ERDAS Imagine 8.5 uses a similar convolutional method for deriving aspect. This algorithm is a third-order finite difference technique, which uses eight neighboring elevation cell values to derive the center grid value. Thus, by selecting the best elevation model, the amount of error in the slope and aspect grids was minimized.

Distance to springs and streams was derived from the Arizona Land Resource Information System (2000) coverages of Arizona springs and streams. These coverages were enhanced using digital raster graphics of 1:24,000 USGS topographic maps and 2-meter resolution digital orthophoto quads. I consulted with District hydrologist, Charles Duncan, USDA Forest Service, Coronado National Forest, Safford Ranger District to obtain an expert opinion regarding the areas most likely to contain water during the bird breeding season (April through July).

#### *3.2.2.3 Retrospective Data*

From April through July 1993 – 1995, the USDA Forest Service Forest and Range Experimental Station, under the direction of W. Block, conducted bird surveys along six transects consisting of 72 point counting stations. These transects were established within oak-juniper woodland, pine-oak woodland, and mixed-coniferous forest. Counting stations were spaced at 300-meter intervals using a systematic-random sampling design (Cochran 1977). Point count station locations were plotted on 1:24,000 USGS topographic maps. Birds were counted using the variable-radius point count method (Reynolds et al. 1980) and distance to each observation was estimated. Each point was surveyed three times per year. To reduce observer bias, site visits were conducted by at least three different observers per point per year. Surveys were conducted between 0.5 hours before and completed four hours after sunrise. Habitat information was collected at or derived from each point. Vegetation data were collected using a circular plot technique; plots were 36m in diameter. Elevation was estimated from topographic maps. Slope was measured in degrees using a clinometer. Aspect was measured using a compass. In addition to the information collected in the field, ALRIS GIS coverage files of streams and springs were used to derive distance values (in meters) to these water sources.

From the 1993-1995 bird study, I used a subset of the bird species recorded for this modeling effort. First, I selected all observations for all species ≤ 120 meters from each point. I selected this distance to reflect the distance of three 30-meter grid cells, which maintains the mapping interval of 30 meters. Then, bird species were selected based on their level of detectability, as well as an "equal" weighting of presence to absence within the species' dataset, which resulted in nine study species. These species selected for modeling were Mexican jay (*Aphelocoma ultramarina*), bridled titmouse (*Baelophus wollweberi*), red-faced warbler (*Cardellina rubrifrons*), yellow-rumped warbler (*Dendroica coronata*), spotted towhee (*Pipilo maculatus*), broad-tailed hummingbird (*Selasphorus platycercus)*, Bewick's wren (*Thryomanes bewickii*), and warbling vireo (*Vireo gilvus*; Table 2). The Bewick's wren, bridled titmouse and Mexican jay are year-round residents and broad-tailed hummingbird, spotted towhee, red-faced warbler, yellow-rumped warbler and warbling vireo are migratory. Generally, warblers occur at upper elevations conifer forests, the warbling vireo in pine oak woodlands, the Bewick's wren, bridled titmouse and Mexican jay at lower elevations within oak-juniper woodland and pine-oak woodlands, and the broad-tailed hummingbird and spotted towhee may occur across the entire elevational gradient where habitat exists.

#### *3.2.2.4 Verification Data*

Additional bird counting stations were established and bird presence data were recorded in 2002 during the bird-breeding season (May to June). These data were used for model verification. Fourteen transects, representing 103 points, were established in oak-juniper woodland, pine-oak woodlands and mixed-coniferous forests. All transects were established within one kilometer of a road. Locations of point count stations were recorded using a GARMIN GPS12 global positioning system. This device does not permit for differential correction. Positional accuracy of this device is ~15 meters (GARMIN 1999). Because study plots were 120 meters in diameter, the coordinates would fall well within this boundary.

Bird survey methods used were the same for the 1993-95 bird survey except each counting station was surveyed twice. Petit et al. (1997) concluded that two visits resulted in surveying approximately 90% of the birds at a site. Finding this estimation acceptable, I decided upon more points and fewer visits per point to better cover the study area and capture variability. Except for vegetation data, habitat information was collected and derived using the same techniques as the retrospective dataset. Methods used for vegetation data collection roughly followed the line-intercept method (Canfield 1941). Vegetation data were collected by establishing two 60-meter belts, one meter wide, in a "cross-hair" configuration centered over the point count station, which resulted in four 30-meter spokes situated at 90 degree angles from the next closest spoke. The one meter square at the center point was marked and vegetation within this area was recored only once. Orientation of the first axis was selected at random; the second axis was situated 90 degrees from the first axis. First, I recorded all vegetation falling on the centerline of each one-meter wide belt (at 0.5m; Canfield 1941); secondly, I recorded all vegetation falling within the one-meter wide belt.

#### *3.2.2.5 Literature-derived Information*

For each study species, available literature was reviewed and biologically important variables were identified. Because the Birds of North America Species Accounts (Academy of Natural Sciences, Philadelphia, PA/ American Ornithologists' Union) is currently the most comprehensive information available on North American birds, I used this information as the basis of identifying habitat variables. If these accounts did not provide regionally specific information, additional literature was reviewed. Because inclusion of GIS-based models based on literature-derived information was simply to compare these models to models based on empirical datasets, when multiple studies were used to compile habitat parameters of each species, evaluation of the different methods for collecting habitat use and selection information was beyond the scope of this study.

### *3.2.3 Statistical Analyses*

I used GIS information, the retrospective dataset and literature-derived information to create spatially explicit models to predict coarse scale landscape variables related to habitat use. Models were generated used non-parametric classification and regression tree analysis (CART; Brieman et al. 1984), and parametric multiple logistic regression (Hosmer and Lemeshow 2000) and variables identified via literature-derived information. CART tends to perform better than stepwise logistic procedures (Brieman et al. 1984). Because GAP uses a coarse scale for deriving distribution maps, I evaluated the appropriateness of this approach by comparing literature-derived information models to mechanistic models. Predictions were tested using the verification dataset.

#### *3.2.3.1 Data Pooling Across Years-Retrospective Dataset*

I pooled data across years for the 1993-95 dataset. I used presence-only data from this dataset to analyze significant differences across-variable per year for each species. Interannual variability was tested using a MANOVA with Bonferonni correction (Schooley 1994) and was considered significant at $p \leq 0.05$. If there was no significant interannual variability, then temporal variability was considered acceptable and data between years was pooled (Schooley 1994).

### *3.2.3.2 Classification Tree Analyses*

I used CART (Salford Systems 2000) program for identifying variables driving habitat use. CART offers a flexible and simple approach for modeling complex ecological relationships (De'ath and Fabricus 2000). Classification trees provide a unique way of identifying the variables most likely to describe the system of study. CART is designed to handle binary response variables (in this case, present or absent). When a variable is identified as an important predictor, the variable is split dichotomously at a threshold within the range of numeric values (e.g., ≤2010 meters elevation for presence; Figure 4). CART requires large sample sizes (≥ 200 samples, Brieman et al. 1984). When this sample size requirement is not met, CART models may not reveal patterns in the data (Breiman et al. 1984). None of the species' datasets in this study met this requirement.

I generated CART models using two approaches. First, CART was run using the Gini index method. At each split on the tree, this index splits the largest category into a separate group. This method entailed loading all variables into CART and running the model; the output defines the "organic tree." Second, a "simulated tree" was produced using a variable shaving technique. Variable shaving involves running the model with all variables, identifying the most important variable, removing it and re-running the model; this procedure was repeated up to three times until a tree was built or three variables were identified. The variables identified were considered habitat descriptors for each species. For selecting significant nodes, a majority rule was used (Breiman et al. 1984), which consists of selecting the node containing the majority of samples of presence.

### *3.2.3.3 Multiple Logistic Regression*

Regression models are highly useful when attempting to describe the relationship between the response (dependent) variable and one or more explanatory (independent) variables (Hosmer and Lemeshow 2000). This method was designed for addressing the relationship of a binary or dichotomous response variable with its covariates. Because logistic regression tends to fit data in a nonlinear fashion (Hosmer and Lemeshow 2000) and ecological data often have a nonlinear distribution (Morrison et al. 1998), logistic regression should serve to best explain the relationship between the response and explanatory variables. This method has been used routinely in modeling observational data against a suite of predictor variables (e.g., Brennan et al. 1986, Pereira and Itami 1991, Guisan and Zimmerman 2000, Pearce and Ferrier 2000, Compton et al. 2002). Subsequently, due to its usefulness in predicting habitat this approach was chosen for comparison to CART.

Using the appropriate transformation, I transformed all non-normally distributed variables in an attempt to normalize the variables. Correlations among variables were tested using the Pearson correlation coefficient. If the correlations' $R^2$ value was > 0.5, and the variable was not biological significant, the covariate with the largest absolute value was removed. For all models, classification cutoffs were weighted based on percentage of point count stations where the species was detected (Hosmer and

Lemeshow 2000); for example, if there were 60 presences and 40 absences, the classification cut of would be 0.6. Moran's I statistic (Odland 1988) was used to test for spatial autocorrelation. Spatial autocorrelation was significant with a Moran's I, $p \leq 0.05$.

Three models were produced using logistic regression; these models were a stepwise logistic regression, a forced logistic regression model using literature-derived information and a forced logistic regression model using literature-derived information with interaction terms. The stepwise procedure was a backward stepwise logistic model; this approach involves loading all variables into the model, and then removing one variable at a time until the model with variables that best fit the data have been identified. The literature-derived forced logistic model was developed by loading all variables identified from the literature as being correlates of habitat use for each species. Finally, the literature-derived information with interaction terms forced logistic model was required for comparison to the other two models because (1) significantly correlated variables with biological significance were retained and (2) regardless of statistical significance, many of the variables used were correlated (e.g., elevation with vegetation type and vegetation type with aspect).

#### *3.2.3.4 Accuracy Assessment*

For each model, error matrices were created and overall accuracy, model sensitivity, specificity, commission and omission errors were calculated (Figure 5). In most accuracy assessments, overall accuracy is usually the only statistic provided

(Congalton and Green 1999). Overall accuracy is calculated by summing the major diagonal cells containing correctly classified cells and dividing by the total sample size (i.e., $( A + E ) / I$ ). Simply providing overall accuracy can be misleading regarding model predictability (Congalton 1991, Congalton and Green 1999, Pearce and Ferrier 2000). Therefore, I used the error matrix to calculate four additional indices regarding model performance. *Sensitivity* is an agreement index, which is described as a proportion of the total number of correctly predicted presences and the total number of observed presence ($A / G$). The proportion of the total number of correctly predicted absences and the total number of observed absences ($E / H$) is model *specificity*. An error of exclusion (*omission error*) is the total number of samples predicted to be absent divided by the total number of observed absences ($E / G$). An inclusion error (*commission error*) is described as and is the total number of samples identified as present divided by the total number of predicted absences ($B / H$). To assess model significance, a standardized $Z$ test was used and a $p$-value was derived.

#### *3.2.3.4.1 CART/ GIS-based Model Accuracy*

I used ArcView 3x and the Spatial Analyst extension to create predictive distribution maps for each species' habitat. A map based on the organic tree, simulated tree and literature-derived models were generated. Using the Kappa Statistic extension (refer to Congalton and Green 1999 for procedure), I conducted an accuracy assessment by running the verification dataset against each predictive distribution maps.

#### *3.2.3.4.2 Logistic Regression Model Accuracy*

Overall accuracy, commission and omission errors, and sensitivity and specificity were derived for each model. The error rate of each model was calculated by running the independent datasets through the equations created by the logistic regression models. Then, I assessed how many of the predicted values matched with the observed values using the defined cutoff point for presence. Thereafter, these observed values were tallied and the error matrices were constructed.

### *3.2.3.5 Competing Model Approach*

When modeling wildlife-habitat relationships, comparative evaluations are superior to selecting just one modeling procedure (Austin 2002). I used a competing models approach to compare parametric and non-parametric tests, and models based on literature information. There were three spatial models (CART organic tree, CART induced tree and a model based on literature-derived information) and three regression models (stepwise logistic, stepwise logistic with biological interaction terms, and a forced regression model based on literature-derived information). For both CART and logistic regression models, a competing models framework was employed. When selecting of the "best" model, the approach must be objective and repeatable (Burnham and Anderson 1998). First, the best spatial and regression models were selected. Secondly, the best overall model was selected. For the CART/ GIS-based models, the organic tree, simulated tree and literature-derived information model were compared and a multi-criteria selection process selected the model with the best fit. The "best" model contained the highest overall accuracy, lowest commission and omission errors, highest model specificity and sensitivity and significant *p*-value ($\leq 0.05$).

The best models for logistic regression, stepwise logistic, forced logistic with literature-derived information and forced logistic with literature-derived information with interaction terms were selected using AIC (Burnham and Anderson 1998). When AIC values are similar between models indicating similar model fits, a multi-criteria selection process was used. I considered the "best" model to have the highest $R^2$ value, the largest Hosmer-Lemshow goodness of fit p-value, the highest overall model accuracy, lowest commission and omission errors, highest model specificity and sensitivity, and when possible, no spatial autocorrelation.

## *3.3 Results*

### *3.3.1 Data Quality*

#### *3.3.1.1 GIS Information*

For the land cover classification map, overall classification accuracy was 71.2% significant at $p < 0.001$. Low classification accuracy in pine-oak woodland transition zone contributed to the lower than generally accepted ≥80% overall accuracy. Because most study species selected oak-juniper and pine-oak or mixed-coniferous and pine-oak, I concluded the low accuracy of the transition zone habitat should not dramatically affect model output.

National Elevation Data (NED) were identified as the “best available” elevation map. This dataset meets the “best available” data standards of the National Spatial Data Infrastructure (Gesch et al. 2002). The NED dataset is a “seamless” elevation data for the entire United States. Created from traditional DEMs, this dataset was produced using: (1) a feathering approach to eliminate edge matching errors when one has to manually mosaic two or more 7.5 minute elevation maps; (2) an interpolation algorithm to fill slivers of missing data along edges of mosaic-ed images; and, (3) an filtering technique to eliminate linear striations on the image, which, unless removed, act as noise and increase the amount of error in the dataset (Gesch et al. 2002). Because USGS-produced DEMs were used as base data for creating NED data, the NED data were deemed more accurate. Shuttle Radar Topographic Mapping mission data are also a seamless dataset of the United States. However, within the study area, this dataset contained numerous “no data” values and the elevation map was deemed incomplete for the study area.

Considerable disagreement existed between the Arizona Land Resource Information System-produced data for streams and springs with high-resolution (2m) digital orthophoto quad maps and digital raster graphics. A majority of these features had to be re-digitized to improve the precision of these map layers. Because the springs and streams maps were enhanced using a combination of photo interpretations of fine resolution (2-meter resolution) aerial photographs and expert opinion, this is the “best available” data of these features.

#### *3.3.1.2 Retrospective Data*

Interannual variability for the bridled titmouse by elevation between years 1993-1994 and years 1994-1995 was significant. However, due to small sample sizes for each year, I still pooled the data. The study design used for collecting the retrospective data was not initially intended for use in this landscape study so sample sizes were not optimal for each study species.

#### *3.3.1.3 Verification Data*

Verification data were collected during the 2002 drought. The extent of the effects of this stochastic event was not quantified.

### *3.3.2 Statistical Analysis*

#### *3.3.2.1 CART Analyses*

In a study of five taxa, Cablk et al. (2002) concluded CART analysis tended to capture much of the spatial variability and autocorrelation. This suggests CART may not be as susceptible to spatial autocorrelation as other global models (e.g., logistic regression, John Prather, *personal communications*). Therefore, tests for spatial autocorrelation were not conducted for these analyses. Additionally, none of the species' datasets met the sample size requirement (≥200) as suggested by Breiman et al. (1981). The variable shaving technique (simulated tree) was believed to serve as a surrogate for this sample size requirement; however, none of the models using this technique fit the data better than the CART organic tree or literature-derived information models.

#### *3.3.2.2 Logistic Regression Analyses*

Only the spotted towhee model had significant Moran's I spatial autocorrelation ($p \leq 0.05$).

### *3.3.3 Model Output*

#### *3.3.3.1 Selection of Best Models*

When comparing performance of CART/ GIS-based models to the multiple logistic regression models, three of four CART/GIS-based models (bridled titmouse, broad-tailed hummingbird, Mexican jay) out performed logistic regression models (Table 3). CART organic tree models had the best fit for broad-tailed hummingbird (Overall Accuracy= 57.3, $p < 0.05$, $N_C$ (commission) = 0.049, $N_o$ (omission) = 0.379, Sensitivity = 0.875, Specificity = 0.381, covariate: elevation ≥ 1894 meters) and Mexican jay (Overall Accuracy = 63.6, $p = 0.09$, $N_C = 0.364$, $N_o = 0.364$, Sensitivity = 0.636, Specificity = 0.636, covariate: oak-juniper). The literature-derived information model performed best for the bridled titmouse (Overall Accuracy = 61.4, $p < 0.01$, $N_C = 0.386$, $N_o = 0$, Sensitivity = 1, Specificity = 0.469, covariates: oak-juniper, pine-oak, grassland, elevation ≤ 1818 meters). For the Bewick's wren (Overall Accuracy = 70.5, $R^2 = 0.86$, HL G-O-F (Hosmer-Lemeshow Goodness of Fit) = 0.991, $N_C = 0.023$, $N_O = 0.272$, Sensitivity = 0.944, Specificity = 0.538, covariates (none significant): oak-juniper, grassland, elevation), the forced logistic regression model of literature-derived information fit the data best. By default, the red-faced warbler (Overall Accuracy = 68.9, $R^2 = 0.795$, HL G-O-F = 0.727, $N_C = 0.162$, $N_O = 0.147$, Sensitivity = 0.647, Specificity = 0.725, covariate: aspect, *slope, mixed-coniferous, oak juniper, *distance to springs) and yellow-rumped warbler (Overall Accuracy = 68.9, $R^2$ = 0.339, HL G-O-F = 0.402, $N_C$ = 0.116, $N_O$ = 0.581, Sensitivity = 0.884, Specificity = 0.419, covariates: negative correlation with slope and *mixed-coniferous and interaction of aspect with mixed-coniferous) backward stepwise logistic regression models and the spotted towhee (Overall Accuracy = 66.0, $R^2$ = 0.406, HL G-O-F = 0.378, $N_C$ = 0.117, $N_O$ = 0.223, Sensitivity = 0.613, Specificity = 0.681, covariates: mixed-coniferous, pine oak, a negative correlation with oak juniper and a negative correlation with elevation) literature-derived information forced regression model fit the datasets best because CART/GIS-based models for these species were not significant. The warbling vireo was not reported because performance of all models were below 50% overall accuracy. For logistic regression models, an asterisk is used to connote significance of covariates ($p<0.05$).

These results were similar to those obtained in other predictive WHR models of bird habitat distribution. In the Unita Mountains, Utah, Lawler and Edwards (2002) predicted nest habitat of four cavity-nesting bird species. Of these, the models for mountain chickadee (*Parus gambeli*) and tree swallow (*Tachycineta bicolor*) had overall accuracies of 50% and 75%, respectively. Models for the northern flicker (*Colaptes auratus*) and red-naped sapsuckers (*Sphyrapicus nuchalis*) had overall accuracies of 84% and 80%, respectively. Penhollow and Stauffer (2000), in the Quantico Marine Corps Base, Virginia modeled habitat of 12 neotropical and short-distant migrants. They predicted habitat for five species with overall accuracies ≤ 70%, three species with overall accuracies between 70 and 75%, three species with overall accuracies of 80% and one species with an overall accuracy of 90%.

Regarding the WHR predictive models for the eight species examined in this study, low predictive power is probably due to a combination of coarse-scale GIS data (particularly the coarse scale vegetation classification map), small sample size of presence and absence data for each species, suboptimal sampling design in the retrospective dataset (which was appropriate for the original objective, the elucidations of bird/habitat relationships), and environmental stochasticity in the verification data.

### *3.4 Discussion*

This study follows Levins (1966) theoretical modeling paradigm. With all models, there are three components: simplicity, accuracy and generality; one can have two at most with the third being diminished. Although the landscape-scale variables were simple and easy to understand by land managers, accuracy of all models was below the 80% acceptance level. Model predictability was sacrificed because: (1) fine scale GIS information were not available; (2) small sample sizes of the retrospective data; (3) regarding retrospective data, the study was not designed for question proposed; and, (4) verification data were collected during a severe drought year.

### 3.4.1 Data Quality

#### 3.4.1.1 GIS Information

Although our GIS information were of the highest quality available, finer scale variables may either improve or diminish model accuracy. Bird-habitat relationships can be scale dependent (Weins et al. 1987) and vegetation structure often influences habitat selection by birds (MacArthur and Wilson 1967, Pearman 2002). Modeling habitat at such a fine scale may better identify parameters driving habitat selection. Conversely, "noise" within a system may be obscured at a coarse scale (Tobalske 2002), which by modeling at a finer-scale inquiry may result in lower predictive accuracy. Use of coarse landscape-scale GIS information was due to availability and testing of coarse-scale habitat models. Because the "proper" scale of inquiry is often obscure and ecological responses may occur at fine or broad scales or both (Weins 1989); therefore, when data are available, multiple scales of habitat selection should be investigated.

When a study is designed specifically for modeling wildlife habitat relationships I recommend a multiple scale approach similar to Johnson's (1980) concept of habitat selection. It is defined as: First order – the physical or geographic range; Second-order – home range of individual or social group; Third-order – sites where "patterns of utilization" occur, such as feeding, roosting, perching and nesting, and; Forth order – i.e., actual procurement of plant species or parts of plants. I modeled habitat by combining the second and third order. Ideally, I would have rather separated the two orders, but due to data availability, these orders were combined. The "home range of the individual or social group" and "feeding, roosting and nesting" were inferred using an unlimited radius point count methodology. I assumed if a bird was singing, it was using the habitat where it was detected.

Hall (1997) investigated habitat use at multiple scales for elegant trogon. She found this pine-oak woodland species also requires nest cavities in large diameter sycamores within riparian areas within its home range. When this approach is not possible, a study should either: (1) consider the differences between geographic range, home range and habitat requisites when defining the scale of inquiry (Trani-Griep 2002) or (2) ensure the scale of inquiry is compatible with the goals of the study (Tobalske 2002).

There are several abiotic and biotic factors, which may improve model performance. Dettmers and Bart (1999) suggest abiotic factors may improve overall accuracy; these variables include topographic roughness index, geology, climate and precipitation. Biotic factors such as soil, patch size, distance to nearest disparate patch and interspersion of habitat types may also be of some value. None of these were considered in this modeling effort. Although these metrics were not within the scope of our study, they may explain away additional variability and improve overall accuracy.

#### *3.4.1.2 Retrospective and Verification Data*

Levins' (1966) theoretical framework for model building is realized when uncertainties of the retrospective and verification data are considered. The retrospective dataset contains small sample sizes for each species and it was not collected in an appropriate manner conducive to the landscape-level habitat modeling. Data used to answer a question different from purpose of the initial data collection (i.e., a divergent question from the initial study), may result in undersampling and will contribute to the overall error (Dettmers and Bart 1999, Edwards et al. 1996: 267). Cockran (1977) indicate inadequate sample size result in lack of confidence and precision.

Verification data were collected during the worst drought on record. Uncertainty exists whether all study species were selecting habitat differently during a severe drought year. Mexican jays were observed in the mixed-coniferous forest. Bushtits and Bullock's orioles (though not considered within this study) were observed in mixed conifer and aspen forests. Although I observed individuals of species usually associated with lower elevations at upper elevations, I cannot be certain if this was related to the drought. There is also the possibility most species returned to their same territories as the previous years but had low nest fidelity. If this is the case, I may still observe these species in their respective habitats.

Positional accuracy was a concern regarding both datasets. During 1993-95, counting stations were plotted on a topographic map. These counting stations were evenly spaced and on average at 300-meter intervals. During the 2002, point locations were recorded using GPS. Due to high topographic variability, although counting stations along transects were paced at 300 meters, at times counting stations occurred at ≤ 150-meter intervals. For 2002, each surveyor who lay transects calibrated their pace prior to establishing transects, so this was not a factor. I determined this placement of points in 2002 was due to topographic variability. Although sample points were laid at 300 meter paced intervals, undulating topography resulted in the spacing being considerable less. I believe there may be similar errors in point placement for the 1993-95 data. Furthermore, positional error may frequently occur in areas of high topographic variability. Fielding and Haworth (1995) question the value of modeling organisms occurring in areas with a high degree of heterogeneity. Bolger et al. (1997) suggests predictions of species distributions may be inaccurate when landscape patterns are heterogeneous.

An additional potential source of error is observer variability. Errors occur due to unequal audiological abilities among observers and this error may be propagated, to an unknown degree, through the use of unlimited-radius point counts (Petit et al. 1997). Additionally, individual observer skill in bird identification differs. Therefore, observer bias occurs between surveyors during a given season. This may be partially addressed by having a static set of surveyors counting birds at a given point (i.e., three different biologists survey a given point three times during a season). Additionally, different teams of observers collected the retrospective and verification datasets. The overall propagated error of these factors combined is unknown.

#### *3.4.1.3 Model Performance*

Two CART models, two stepwise logistic regression models, one GIS-based literature-derived information and two forced regression models of literature-derived information fit the datasets best. These models were the broad-tailed hummingbird and Mexican jay CART organic tree models, the bridled titmouse GIS-based literature-derived information model, and the red-faced warbler and yellow-rumped warbler backward stepwise logistic regression models, and the spotted towhee and Bewick's wren forced logistic regression model of literature-derived information.

For the broad-tailed hummingbird, the "CART organic tree" fit the data best identifying elevation ($\geq$ 1894m). Because of the coarse scale of this assessment, the upper elevational extent of pine oak and all of mixed-coniferous forests, as well as riparian areas, which likely contain a higher density of flowering plants, is captured within this variable identified by CART. Although this model predicts habitat use slightly better than chance (overall accuracy = 57.3%), the sensitivity value is high indicating the model predicted presence, when the species was actually "present" quite well. However, because this species selects areas with higher densities of flowering understory plants (William and Calder 1992), a map of understory would be more appropriate. Unfortunately, there is still no method for deriving understory cover from satellite imagery.

The model with the best fit for the Mexican jay was the CART organic tree. This model identified oak-juniper woodland as the only predictor variable. Mexican jays inhabit oak and oak-juniper woodland, pine-oak woodlands and riparian areas containing oaks, and may occur at higher densities in areas with high densities of oak

(Brown 1994). A finer scale vegetation map and canopy cover map may greatly improve the accuracy of this model.

The forced logistic regression of literature-derived information performed best for the spotted towhee. This model identified positive correlations for mixed-coniferous and pine-oak, and negative correlations for oak-juniper and elevations (all nonsignificant at $p > 0.05$). In the intermountain west, the spotted towhee occurs at mid-elevations in a variety of vegetation associations characterized as dense, shrubby, thickets (Greenlaw 1996). Although this model does make sense biologically, samples are significantly spatially autocorrelated ($p < 0.01$); thus, independence across sample sites is questionable.

Then red-faced warbler backward stepwise logistic regression analysis identified positive correlations for aspect, slope ($p < 0.05$), mixed conifer forest, oak-juniper woodland and distance to springs ($p < 0.05$). This species requires mid- to high-elevation montane coniferous forests, which include fir (Abies), pine (Pinus) (Price 1888), open pine-oak forests and stream and snow melt drainages (Martin and Barber 1995). These analyses identified oak-juniper (though nonsignificant) as a predictor variable for this species. Red-faced warblers rarely venture down to this elevation. Additionally, the model identified a significant negative correlation for distance to springs. This does not make biological sense because this species nests in canyons and drainages where conditions are wet (Martin and Barber 1995).

The best-fitted model for the yellow-faced warbler was the backward stepwise logistic regression. This model identified negative correlations with slope (not significant) and mixed-coniferous ($p < 0.05$) and a positive correlation with the interaction term of aspect with mixed-coniferous (not significant). This species occurs primarily in mature coniferous and mixed-coniferous-deciduous forests (Franzreb 1978, Hunt and Flaspohler 1998). Marshall (1957), in his study of pine-oak woodlands of the Madrean Archipelago, observed yellow-rumped warbler in at the upper elevational extent of pine-oak woodlands. Because mixed-coniferous forests and elevation are highly correlated (with the exception of lower elevation north facing slopes and low elevation steep sided canyons), this model potentially captures upper elevation and mixed-coniferous forest as important predictor variables. Although mixed-coniferous was significant and overall accuracy was considerably better than chance, this model served to verify, in a very general sense, our current understanding of this species.

### *3.5 Conclusion*

Overall, five of the models developed in this study performed reasonably well (overall accuracy > 68%). While none of the models attained an overall accuracy of ≥ 80%, often regarded as the acceptance threshold for a 'good' predictive model (Mosher et al. 1986, Wright et al. 2000), the results are comparable to those from other habitat mapping projects executed under similar conditions with similar constraints. For this study, it is reasonable to assume the majority of model uncertainties and subsequent errors occurred due to inadequate retrospective and verification datasets. The retrospective data used to develop the predictive WHR models was collected for the purpose of studying bird-habitat relations within a small (0.1 ha) study plots (Block *unpublished data*). Although this study design served it's original application well, the use of these data for predicting bird habitat, over a much larger (0.1 ha) area, was suboptimal from a spatial modeling perspective. Clearly, larger sample sizes for each species, which would offer better coverage of the mapping region, would likely, produce more accurate and management-relevant habitat maps. However, as a management case study, this study is representative of many WHR modeling challenges, where management questions demand better understanding of the spatial distribution of habitats, but time and available resources limit the amount and extend of additional data collection. In such cases, wildlife ecologists must make the best use of available data. Within this context, this study demonstrates a practical test of the value of similar mapping efforts, many of which have been undertaken since the advent of easy-to-use spatial analysis software and GIS. The accuracies obtained in this effort, while falling short of the high standards set for state-of-the-art modeling efforts, suggest the predictive maps generated by this effort improve our understanding of the distribution of bird habitats for selected species on the Pinaleños Mountains. Given the constraints of small sample sizes, high topographic variability, and coarse-scale validation data, these results are encouraging. Whether these WHR models provide information useful to land managers will require further scrutiny.

### *3.6 Management Implications*

This study illustrates several areas where model error and uncertainty occur, as well as how it may affect the reliability of model predictions. To maximize model performance, all datasets should be subjected to the highest level of scrutiny. Determining how errors propagate within a multi-layered GIS study will perhaps remain a fruitful endeavor (Edwards et al. 1996). Therefore, it is imperative for habitat modelers to critically evaluate GIS information and use only the highest quality datasets. This can be accomplished by the following recommendations, which will result in making wildlife habitat relationship models most useful to land managers.

1) Ideally, the study should be designed to answer the research question proposed, which will ensure the data are correctly parameterized and an adequate sample size is obtained.

2) When possible, multiple year dataset should be used for creating and validating models, which will assist in capturing variability of the system of study. For example, verification data were collected during the 2002 severe drought.

3) If pooling data between years, tests for interannual variations should be conducted.

4) When possible, multiple scales of habitat selection should be investigated.

5) In areas of high topographic variability, alternatives to traditional linear transect methods should be investigated.

6) Test for spatial autocorrelation should always be conducted. When designing a study to model vertebrate distributions on the landscape, a "bird shot" sample design, whereby survey sites are *a priori* selected across the study area may assist to reduce the likelihood of spatial autocorrelation.

7) Test model predictions with an independent empirical dataset.

### *3.7 Literature Cited*


Arizona Land Resource Information System (ALRIS) 2000. Arizona Land Resource Information System. Arizona State Land Department, Phoenix, Arizona. http://www.land.state.az.us/alris/index.html date accessed: 10 October 2003.

Austin, M. 2002. Case studies of the use of environmental gradients in vegetation and fauna modeling: theory and practice in Australia and New Zealand. *In* Predicting Species Occurrences: Issues of Accuracy and Scale (J. Scott, R. Heglund, M. Morrison et al.). Island Press, Washington. Pps 73-82.

Bolger, D., T. Scott, and J. Rotenberry. 1997. Breeding bird abundance in an urbanizing landscape in coastal southern California. Conservation Biology 11 (2): 406-421.

Breiman, L. J. Friedman, F. Olshen, C. Stone. Classification and regression tress. Chapman and Hall Publishing, New York.

Brennan, L. W. Block, R. Gutiérrez 1986. The use of multivariate statistics for developing habitat suitability index models. *In* Wildlife 2000 (J. Verner, M. Morrison, C. Ralph, eds), University of Wisconsin Press. Pps. 177-182.

Brown, J. 1994. Mexican jay (Aphelocoma ultramarina). *In* Birds of North America 118 (A. Poole and F. Gill, eds). Academy of Natural Sciences, Philadelphia, and American Ornithologists' Union, Washington D.C. 15pps.

Brugnach, M., J. Bolte and G. Bradshaw. 2003. Determining the significance of threshold uncertainty in rule-based classification models. Ecological Modeling 160: 63-76.

Burnham, K. and D. Anderson 1998. Model selection and inference: a practical information-theoretic approach. Springer-Verlag, New York, New York. Pps. 353.

Cablk, M., D. White and A. Kiester 2002. Assessment of spatial autocorrelation in empirical models in ecology. *In* Predicting Species Occurrences: Issues of Accuracy and Scale (J. Scott, R. Heglund, M. Morrison et al.). Island Press, Washington. Pps 429-440.

Canfield, R. 1941. Application of the line interception method in sampling range vegetation. Journal of Forestry 39:388-394.

Cardillo, M., D. Macdonald, S. Rushton 1999. Predicting mammal species richness and distributions: testing the effectiveness of satellite-derived land cover data. Landscape Ecology 14: 423-435.

Carrol, C., W. Zielinski, R. Noss. 1999. Using presence-absence data to build and test spatial habitat models for the fisher in the Klamath region, U.S.A. Conservation Biology 13 (6): 1344-1359.

Chrisman, N. 1987. The accuracy of map overlays: a reassessment. Landscape and Urban Planning 14: 427-439.

Clark, J.D., J.E. Dunn, and K.G. Smith 1993. A multivariate approach of female black bear habitat use for a geographic information system. Journal of Wildlife Management 57: 519-526.

Cockran, W. 1977. Sampling techniques. Third edition. Wiley, New York.

Compton, B., J. Rhymer and M. McCollough 2002. Habitat selection by wood turtles (*Clemmys insculpta*): an application of paired logistic tests. Ecology 83 (3): 833-843.

Congalton, R.G. 1991. A review of assessing the accuracy of classifications of remotely sensed data. Remote Sensing of Environment 7: 33-46.

Congalton, R. and K. Green. 1999. Assessing the accuracy of remotely sensed data: principles and practices. Lewis Publishers, Boca Raton. Pps. 137.

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

Dettmers, R. and J. Bart. 1999. A GIS method applied to predicting forest songbird habitat. Ecological Modeling 9 (1): 152-163.

Edwards, T., E. Deshler, D. Foster and G. Moisen. 1996. Adequacy of wildlife habitat relation models for estimating spatial distributions of terrestrial vertebrates. Conservation Biology 10 (1): 263-270.

Fielding A. and P. Haworth. 1995. Testing the generality of bird-habitat models. Conservation Biology 9 (6): 1466-1481.

Franzreb, K. 1978. Tree species used by bird in logged and unlogged mixed-coniferous forests. Wilson Bulletin 90: 221-238.

Garmin, Inc., 1999. Garmin 12 Personal Navigator, Owner's Manual and Reference. Garmin International, Inc. Olathe, Kansas. Pps. 66.

Gesch, D, M. Oimoen, S. Greenlee, C. Nelson, M. Steuck and D. Tyler 2002. The National Elevation Dataset. Photogrammetric Engineering and Remote Sensing 68 (1): 5-11.

Greenlaw, J. 1996. Spotted towhee (*Pipilo maculatus*). *In* Birds of North America 263 (A. Poole and F. Gill, eds). Academy of Natural Sciences, Philadelphia, and American Ornithologists' Union, Washington D.C. 32 pps.

Guisan, A. and N. Zimmerman 2000. Predictive habitat distribution models in ecology. 135: 147-186.

Gutzwiller, K. and W. Barrow. 2001. Bird-landscape relations in the Chihuahuan Desert: coping with uncertainties about predictive models. Ecological Applications 11 (5): 1517-1532.

Hall, L. 1996. Habitat selection by the elegant trogon (*Trogon elegans*) at multiple scales. Ph.D. dissertation, University of Arizona, Tucson, AZ.

Hosmer, D. and S. Lemeshow 2000. Applied logistic regression. Second edition. Wiley-Interscience Publications, John Wiley and Sons, Inc. pps. 373.

Hunt, P and D. Flaspohler 1998. Yellow-rumped warbler (*Dendroica coronata*). *In* Birds of North America 376 (A. Poole and F. Gill, eds). Academy of Natural Sciences, Philadelphia, and American Ornithologists' Union, Washington D.C. 27 pps.

Jenness, J. 2001. Surface areas and ratios extension. Jenness Enterprises, Inc.
http://www.jennessent.com/arcview/surface_areas_ratios.htm Date accessed: 25 March 2003.

Johnson, D. 1980. The comparison of usage and availability measurements for evaluating resource preference. Ecology 6 (1): 65-71.

Jones, K. 1998. A comparison of two approaches to ranking algorithms used to compute hill slopes. Geoinformatica 2 (3): 235-256.

Lane, S. and J. Chandler 2003. The generation of high quality topographic data for hydrology and geomorphology: new data sources, new applications and new problems. Earth Surface Processes and Landforms 28: 229-230.

Lawler, J. and T. Edwards, Jr. 2001. Landscape patterns as habitat predictors: building and testing models for cavity-nesting birds in the Unita Mountains of Utah, USA. Landscape Ecology 17: 233-245.

Levins, R. 1966. The strategy of model building in population biology. American Scientist 54: 421-431.

MacArthur, R. and E. Wilson . 1967. The theory of island biogeography. Monographs in population biology 1: 1-203.

MacKinnon, J. and R. DeWulf 1994. Designing protected areas for giant pandas in China *In* Mapping the diversity of nature (Miller, R., editor). Chapman and Hall, London. Pps. 127-142.

Martin T. and P. Barber 1995. Red-faced warbler (*Cardellina rubrifrons*): *In* The Birds of North America, No. 375 (A. Poole and F. Gill, eds.). The Birds of North America, Inc., Philadelphia, PA. 14pps.

Marshall, J. 1957. Birds of the pine-oak woodland in southern Arizona and adjacent Mexico. Pacific Coast Avifauna 32: 125.

Merrill, T., D.J. Mattson, R.G. Wright, H.B. Quigley 1999. Defining landscapes suitable for restoration of grizzly bears Ursus arctos in Idaho. Biological Conservation 87: 231-248.

Morrison, M., B. Marcot, and R. Mannan 1998. Wildlife Habitat Relationships: Concepts and Applications, 2nd Edition. Pps. 458.

Mosher, J., K. Titus and M. Fuller 1986. Developing a practical model to predict nesting habitat of woodland hawks. *In* Wildlife 2000: Modeling habitat relationships of terrestrial vertebrates (J. Verner, M. Morrison, C. Ralph, *editors*). The University of Wisconsin Press, Madison, Wisconsin. Pps. 31-35.

Nocedal, J. & M. Ficken 1998. Bridled Titmouse (*Baeolophus wollweberii*). *In* The Birds of North America, No. 375 (A. Poole and F. Gill, eds.). The Birds of North America, Inc., Philadelphia, PA.

Odland, J. 1988. Spatial Autocorrelation. Sage Publications, Inc. 87pps.

Pereira, J.M.C. & R.M. Itami, 1991. GIS-based habitat modeling using logistic multiple regression: a study of the Mt. Graham red squirrel. Photogrammetric Engineering & Remote Sensing 57 (11): 1475-1486.

Pearman, P. 2002. The scale of community structure: habitat variation and avian guilds in tropical forest understory. Ecological Monographs 72 (1) 19-39.

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

Penhollow, M. & D. Stauffer, 2000. Large-scale habitat relationships of neotropical migratory birds in Virginia. Journal of Wildlife Management 64 (2): 363-372.

Petit, D., L. Petit, V. Saab and T. Martin 1997. Fixed-radius point counts in forests: factors influencing effectiveness and efficiency. *In* Monitoring Bird Populations by Point Counts (C. Ralph, J. Sauer and S. Droege, eds), USDA Forest Service, Pacific Southwest Research Station, General Technical Report PSW-GTR-149. pps. 49-56.

Price, W. 1888. Nesting of the Red-faced warbler (*Cardellina rubrifrons*) in the Huachuca Mountains, southern Arizona. Auk 5: 385-386.

Rao, P. 1998. Statistical research methods in the life sciences. Brooks/ Cole Publishing Company, Pacific Grove, CA. Pps. 889.

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

Salford Systems 2000. CART Version 4.0. Salford Systems, Inc. San Diego, CA.

Schooley, R. 1994. Annual variation in habitat selection: patterns concealed by pooled data. Journal of Wildlife Management 58 (2): 367-374.

Scott, J., T. Hart, F. Davis 1996. Gap Analysis: A landscape approach to biodiversity planning. American Society for Photogrammetry and Remote Sensing, Bethesda, Maryland. Pps. 320.

Steele, B., J. Winne, and R. Redmond 1998. Estimation and Mapping of Misclassification Probabilities for Thematic Land Cover Maps. Remote Sensing and the Environment 66: 192-202.

Stoms, D.M., F.W. Davis and C.B. Cogan 1992. Sensitivity of wildlife habitat models to uncertainties in GIS information. Photogrammetric Engineering and Remote Sensing 58 (6): 843-850.

Trani-Griep, M. 2002. The influence of spatial scale on landscape pattern description and wildlife habitat assessment. *In* Predicting Species Occurrences: Issues of Accuracy and Scale (J. Scott, R. Heglund, M. Morrison et al.). Island Press, Washington. Pps 197-204.

Tobalske, C. 2002. Effects of spatial scale on the predictive ability of habitat models for the green woodpecker in Switzerland. *In* Predicting Species Occurrences: Issues of Accuracy and Scale (J. Scott, R. Heglund, M. Morrison et al.). Island Press, Washington. Pps 197-204.

United States Department of Agriculture, Forest Service (USDAFS) 1997a. Plant Associations of Arizona and New Mexico, Edition 3. Volume 1: Forests. USDAFS, Southwestern Region, Habitat Typing Guide. Washington, D.C. pps291.

United States Department of Agriculture, Forest Service (USDAFS) 1997b. Plant Associations of Arizona and New Mexico, Volume 2: Woodlands. USDAFS, Southwestern Region, Habitat Typing Guide. Washington, D.C. pps196.

Van Horne, B. 1983. Density as a misleading indicator of habitat quality. Journal of Wildlife Management 47: 893-901.

Van Horne, B. and J. Weins, 1991. Forest bird habitat suitability models and the development of general habitat models. Fish Wildlife Research 8. U.S. Fish and Wildlife Service.

Weins, J. 1989. Spatial scaling in ecology. Functional Ecology 3: 385-397.

Weins, J., J. Rottenberry and B. Van Horne. 1987. Habitat occupancy patterns in North American shrubsteppe birds: the effects of spatial scale. Oikos 48: 132-147.

Williams, A. and L.L. Calder. 1992. Broad-tailed hummingbird (*Selasphorus platycercus*). *In* The Birds of North America, No. 16 (A. Poole and F. Gill, eds.). The Birds of North America, Inc., Philadelphia, PA.

Wisdom, M., R. Holthausen, B. Wales, Cl Hargis, V. Saab, D. Lee, W. Hann, T. Rich, M. Rowland, W. Murphy, M. Eames 2000. Source habitats for terrestrial vertebrates of focus in the Interior Columbia Basin: broad-scale trends and management implications: Volume 2- Group Level Results. U.S. Department of Agriculture Forest Service, General Technical Report PNW-GTR-485.

Wright, A., A. Fielding, C. Wheater 2000. Predicting the distribution of Eurasian badger (Meles meles) over an urbanized landscape: a GIS approach. Photogrammetric Engineering and Remote Sensing 66 (4): 423-428.

## *3.8 Tables*

Table 1. Landscape-scale variables used in all predictive habitat models (all variables have resolution of 30m).

| Type | Date | Source | Units | Type of data | Classes |
|---|---|---|---|---|---|
| [1]Landover | Nov 99 | Landsat 7 ETM+ | Meters | Categorical | MC – Mixed-coniferous forest<br>PO – Pine-oak woodland<br>OJ – Oak-juniper woodland<br>GR – Grasslands/meadows |
| Elevation | | NED | Meters | Continuous | |
| [2]Aspect | | NED | Degrees | Continuous | |
| [2]Slope | | NED | Percent | Continuous | |
| [3]Distance to streams | 2002 | ALRIS* | Meters | Continuous | |
| [3]Distance to springs | 2002 | ALRIS* | Meters | Continuous | |

[1]Land cover data was derived from a supervised classification using LANDSAT 7 ETM+ (November 1999) imagery and four classes were created.
[2]Aspect and slope were derived from NED elevation data.
[3]Locations of springs and streams were enhanced from Arizona Land Resource Inventory System (ALRIS) using photointerpretations of 2-meter resolution digital orthophoto quarter quad.

Table 2. Common name, species' code and scientific name of the species used in the modeling effort. For the 1993-95 and 2002 datasets, the number of counting stations where birds were detected (or presence, *P*), number of counting stations where birds were not detected (or absences, *A*), and total number of counting stations (*N*) are also provided.

| | | | 1993-1995 | | | 2002 | | |
|---|---|---|---|---|---|---|---|---|
| Common name | Code | Scientific name | *P* | *A* | *N* | *P* | *A* | *N* |
| Mexican jay | MEJA | *Aphelocoma ultramarina* | 21 | 26 | 47 | 12 | 32 | 44 |
| Bridled titmouse | BRTI | *Baelophus wollweberi* | 26 | 21 | 47 | 12 | 32 | 44 |
| Red-faced warbler | RFWA | *Cardellina rubrifrons* | 37 | 11 | 48 | 34 | 40 | 74 |
| Yellow-rumped warbler | YRWA | *Dendroica coronata* | 37 | 11 | 48 | 43 | 31 | 74 |
| Spotted Towhee | SPTO | *Pipilo maculates* | 50 | 22 | 72 | 31 | 72 | 103 |
| Broad-tailed hummingbird | BTHU | *Selasphorus platycercus* | 56 | 16 | 72 | 40 | 63 | 103 |
| Bewick's wren | BEWR | *Thryomanes bewickii* | 29 | 18 | 47 | 18 | 26 | 44 |
| Warbling vireo | WAVI | *Vireo gilvus* | 30 | 18 | 48 | 17 | 57 | 74 |

Table 3. Only the models with the "best" performance are listed. The models selected were backward stepwise regression, organic classification tree, GIS-based literature-based and forced logistic regression of literature-based. Moran's I spatial autocorrelation statistic was ran only for parametric mechanistic models (logistic regression).

| | Overall Accuracy (%) | Sensitivity (%) | Specificity (%) | Commission (%) | Omission (%) | $R^2$ | Significance | Moran's I | AIC |
|---|---|---|---|---|---|---|---|---|---|
| *Backward stepwise* | | | | | | | | | |
| [a]Red-faced warbler | 68.9 | 64.7 | 41.9 | 0.116 | 0.581 | 0.399 | [b]0.402 | NS | selected |
| [a]Yellow-rumped warbler | 68.9 | 88.4 | 41.9 | 0.116 | 0.581 | 0.399 | [b]0.402 | NS | selected |
| *Forced logistic regression (literature-based)* | | | | | | | | | |
| Bewick's wren | 70.5 | 94.4 | 53.9 | 0.023 | 0.272 | 0.86 | [b]1 | NS | selected |
| Spotted towhee | 66 | 61.3 | 68.1 | 0.117 | 0.223 | 0.406 | [b]0.378 | NS | selected |
| *CART (organic tree)* | | | | | | | | | |
| [a]Broad-tailed hummingbird | 57.3 | 87.5 | 38.1 | 0.049 | 0.379 | - | 0.03 | - | - |
| Mexican jay | 63.6 | 63.6 | 63.6 | 0.364 | 0.364 | - | 0.09 | - | - |
| *GIS-based (literature based)* | | | | | | | | | |
| Bridled titmouse | 61.4 | 100 | 46.9 | 0.386 | 0 | - | 0.03 | - | - |

[a]GIS-based models nonsignificant, BSR selected by default
[b]Hosmer-Lemshow Goodness of Fit

## *3.9 Figures*

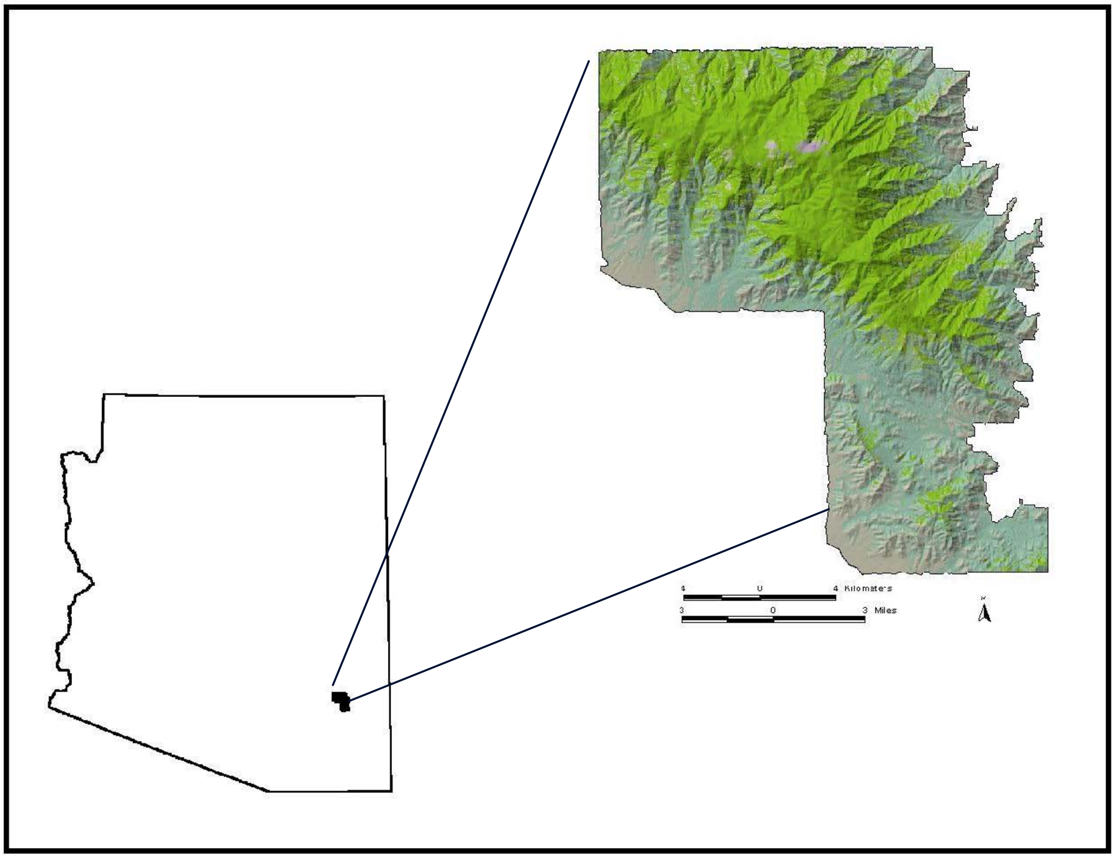


Figure 2.1. Locator map of the Pinaleños Mountains, southeastern Arizona.

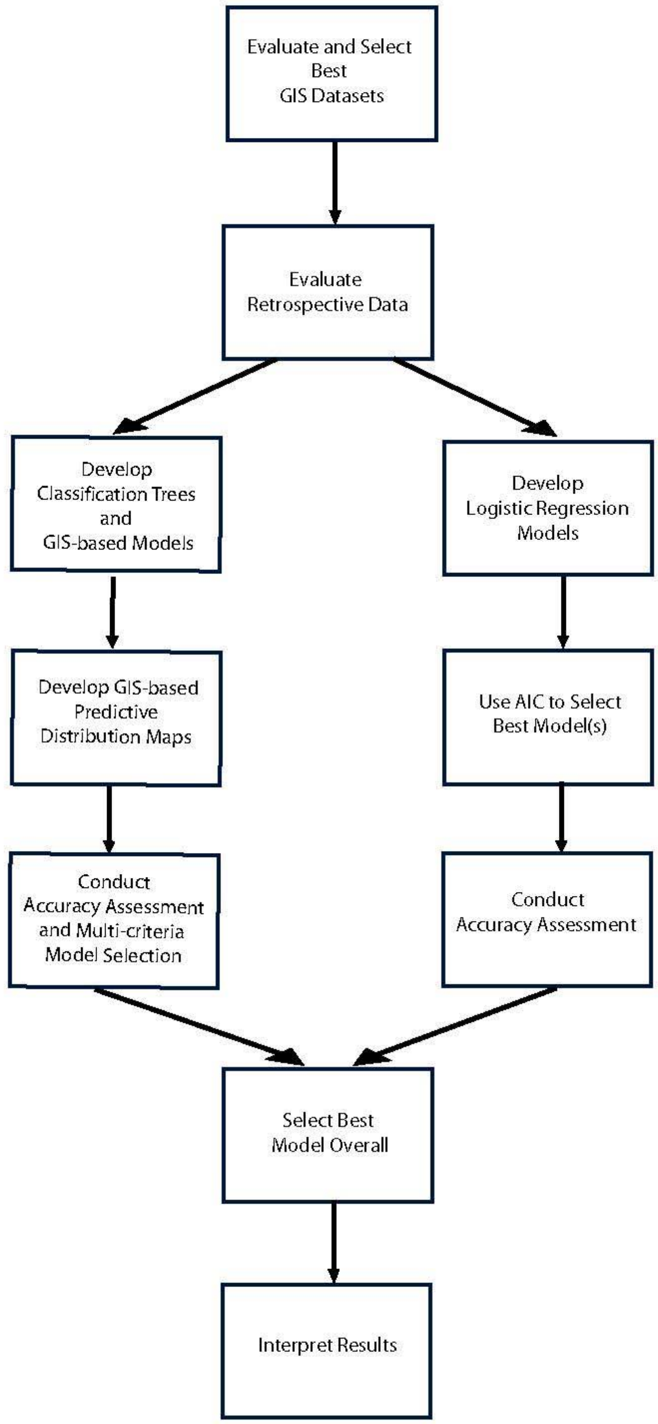


Figure 2.2. A schematic flow diagram used in the habitat modeling process. GIS, retrospective and verification data are evaluated, CART, GIS-based and logistic regression models are developed, accuracy for all models are assessed, the best models are selected and model results are interpreted.

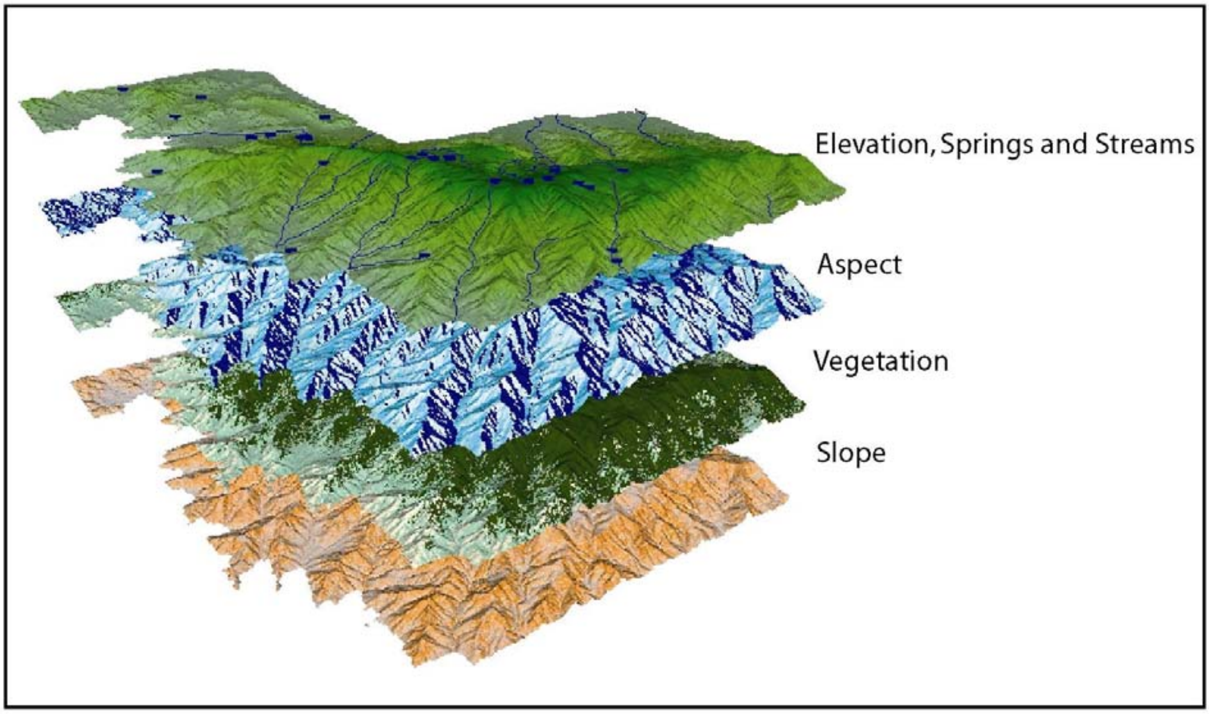


Figure 3. This layer stack, produced using ArcView 3x 3-D Analyst, is a graphical illustration of the six landscape-scale variables used in the modeling effort. Elevation, springs and streams are presented in one layer (at top).

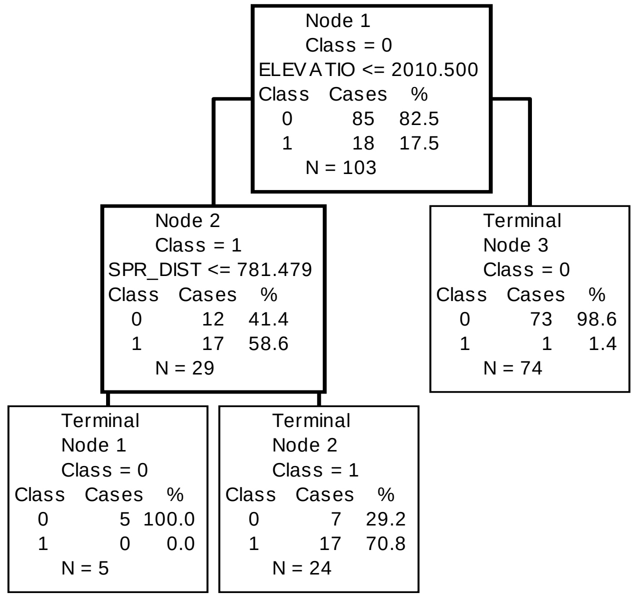


Figure 4. Example of a classification tree, which is for illustrative purposes only. Nodes 1 and 2 are considered parent nodes. Terminal nodes 1, 2 and 3 are considered daughter nodes.

| | Observed Present | Observed Absent | |
|---|---|---|---|
| Predicted Present | *A* | *B* | *C* |
| Predicted Absent | *D* | *E* | *F* |
| | *G* | *H* | *I* |

Figure 5.  Classification matrix used to describe the agreement between observed and predicted values and calculating overall accuracy, model sensitivity, specificity, commission and omission errors, and producer and user accuracy.