INFORMATION TO USERS

This manuscript has been reproduced from the microfilm master. UMI films the text directly from the original or copy submitted. Thus, some thesis and dissertation copies are in typewriter face, while others may be from any type of computer printer.

The quality of this reproduction is dependent upon the quality of the copy submitted. Broken or indistinct print, colored or poor quality illustrations and photographs, print bleedthrough, substandard margins, and improper alignment can adversely affect reproduction.

In the unlikely event that the author did not send UMI a complete manuscript and there are missing pages, these will be noted. Also, if unauthorized copyright material had to be removed, a note will indicate the deletion.

Oversize materials (e.g., maps, drawings, charts) are reproduced by sectioning the original, beginning at the upper left-hand corner and continuing from left to right in equal sections with small overlaps.

ProQuest Information and Learning
300 North Zeeb Road, Ann Arbor, MI 48106-1346 USA
800-521-0600

EXTRACTING TEMPORAL AND SPATIAL INFORMATION FROM REMOTELY SENSED DATA FOR MAPPING WILDLIFE HABITAT

by

Cynthia S.A. Wallace

A Dissertation Submitted to the Faculty of the

DEPARTMENT OF GEOGRAPHY AND REGIONAL DEVELOPMENT

In Partial Fulfillment of the Requirements
For the Degree of

DOCTOR OF PHILOSOPHY

In the Graduate College

THE UNIVERSITY OF ARIZONA

2002

THE UNIVERSITY OF ARIZONA ®
GRADUATE COLLEGE

As members of the Final Examination Committee, we certify that we have read the dissertation prepared by Cynthia S.A. Wallace

entitled Extracting Temporal and Spatial Information from Remotely Sensed Data for Mapping Wildlife Habitat

and recommend that it be accepted as fulfilling the dissertation requirement for the Degree of Doctor of Philosophy

Dr. Stuart E. Marsh	Stuart E. Marsh	11/8/02 Date
Dr. Stephen R. Yool		11/8/02 Date
Dr. Brigitte S. Waldorf		11-8-2002 Date
Dr. Alfredo R. Huete		11/8/02 Date
Dr. Donald E. Myers		11/08/02 Date

Final approval and acceptance of this dissertation is contingent upon the candidate's submission of the final copy of the dissertation to the Graduate College.

I hereby certify that I have read this dissertation prepared under my direction and recommend that it be accepted as fulfilling the dissertation requirement.

Stuart E. Marsh — Dissertation Director

12/02/02 — Date

STATEMENT BY AUTHOR

This dissertation has been submitted in partial fulfillment of requirements for an advanced degree at The University of Arizona and is deposited in the University Library to be made available to borrowers under rules of the Library.

SIGNED: Cynthia A. Wallace

ACKNOWLEDGEMENTS

I have appreciated the guidance and support of many people throughout the process of completing this dissertation.

At the University of Arizona, I am indebted to my advisor, Dr. Stuart Marsh, whose wisdom, understanding, and encouragement helped me realize this goal even though my life outside the university often threatened to derail the effort. I am especially grateful for Stuart's ability to sift through my abundance of ideas and extract those that were viable, keeping my research efforts focused and productive. I thank my committee for their expertise and feedback, which was given freely and promptly: Dr. Stephen Yool, my Master's advisor, who fostered my interest in remote sensing and spatial analysis; Dr. Brigitte Waldorf, who made statistics exciting, meaningful and interesting; Dr. Don Myers, who was always willing to draw on his encyclopedic knowledge of geostatistics to brainstorm new ideas; and Dr. Alfredo Huete, who can make the most complex ideas understandable and who taught this hard-rock geologist to appreciate soil.

Thanks are extended also to the students and staff at the Arizona Remote Sensing Center for providing an exciting and vital research environment in which it is easy and fun to learn all this technical stuff. Special thanks go to Jessica J. Walker, for her scientific and technical insights, her exceptional talent as an editor and wordsmith, her encouragement and faith that I could do this, and for her friendship.

In my community, I want to thank first my mother-in-law, Ouida Wallace, for her love, help, energy and prayers. You know I couldn't have done this without you. I am also grateful to Khalsa Montessori School for providing a vibrant, nurturing, and loving environment for my daughters. Khalsa attracts the most amazing people and many of the families have been there for the girls and for me countless times. I also am grateful to my faith community for their care, concern, support, and prayers. Nobody does this alone.

Finally, I want to thank my children, Margo Marie and Mathea Sierra. I am truly the luckiest Mom in the world and I love you more than you can ever imagine. You have both been remarkably patient and have worked so hard to help me achieve this great goal: taken on such responsibility, done without me more evenings than you would choose, respected my time to work. Thank you for loving me and supporting me through this, and for keeping my priorities straight. We are a perfectly wonderful family.

DEDICATION

Dedicated to my Mother

Doris Jeanne Keller Anderson

January 30, 1926 ~ November 9, 2002

MAGIC CARPET

By: Shel Silverstein

You have a magic carpet
That will whiz you through the air,
To Spain or Maine or Africa
If you just tell it where
So will you let it take you
Where you've never been before,
Or will you buy some drapes to match
And use it
On your
Floor?

TABLE OF CONTENTS

LIST OF FIGURES

ABSTRACT

The research accomplished in this dissertation used both mathematical and statistical techniques to extract and evaluate measures of landscape temporal dynamics and spatial structure from remotely sensed data for the purpose of mapping wildlife habitat. By coupling the landscape measures gleaned from the remotely sensed data with various sets of animal sightings and population data, effective models of habitat preference were created.

Measures of temporal dynamics of vegetation greenness as measured by National Oceanographic and Atmospheric Administration's Advanced Very High Resolution Radiometer (AVHRR) satellite were used to effectively characterize and map season-specific habitat of the Sonoran pronghorn antelope, as well as produce preliminary models of potential yellow-billed cuckoo habitat in Arizona. Various measures that capture different aspects of the temporal dynamics of the landscape were derived from AVHRR Normalized Difference Vegetation Index composite data using three main classes of calculations: basic statistics, standardized principal components analysis, and Fourier analysis. Pronghorn habitat models based on the AVHRR measures correspond visually and statistically to GIS-based models produced using data that represent detailed knowledge of ground-condition.

Measures of temporal dynamics also revealed statistically significant correlations with annual estimates of elk population in selected Arizona Game Management Units, suggesting elk respond to regional environmental changes that can be measured using

satellite data. Such relationships, once verified and established, can be used to help indirectly monitor the population.

Measures of landscape spatial structure derived from IKONOS high spatial resolution (1-m) satellite data using geostatistics effectively map details of Sonoran pronghorn antelope habitat. Local estimates of the nugget, sill, and range variogram parameters calculated within 25x25-meter image windows describe the spatial autocorrelation of the image, permitting classification of all pixels into coherent units whose signature graphs exhibit a classic variogram shape. The variogram parameters captured in these signatures have been shown in previous studies to discriminate between different species-specific vegetation associations.

The synoptic view of the landscape provided by satellite data can inform resource management efforts. The ability to characterize the spatial structure and temporal dynamics of habitat using repeatable remote sensing data allows closer monitoring of the relationship between a species and its landscape.

CHAPTER 1

INTRODUCTION

CONTEXT OF THE PROBLEM

Understanding the relationship between a wildlife species and its environment is a prerequisite to making informed resource management decisions. With increased knowledge, natural resource managers can better anticipate effects of environmental change and take steps to mitigate their consequences. Without this knowledge, incorrect decisions could be made under misguided assumptions about the links between a wildlife species and its environment. Wildlife in danger of extinction are particularly vulnerable, since the margin of error may be low and they are particularly vulnerable to ill-advised management actions, since the margin of error is low and the consequences of missteps are high.

Wildlife habitat maps created through traditional, field-based mapping techniques generally present a static picture of the environment. A vegetation map, for instance, shows the type and location of vegetation as it existed at the time of data collection, but typically conveys no information about the vegetation condition, variations in size and densities, response to climatic events, or seasonal phenology. This lack of information handicaps research attempts to analyze the habitat preferences of mobile species that are highly sensitive to environmental changes.

The ideal basis for a habitat map is one in which researchers frequently and quantitatively assess the environmental changes in a landscape and study the subsequent behavioral responses in a given wildlife species. Remote sensing data, with its high

repeat times and synoptic views, show promise as a key element in this approach: if the data can be used to develop indices of a landscape's temporal dynamics and spatial structure, the relationship between a species and its environment can be more closely monitored and evaluated. Once areas of detailed ground-condition knowledge are reliably linked to an interpretation of remote sensing data, the information can be extrapolated across space and through time, permitting the characterization of landscapes at regional scales.

The specific goal of this research was to assist the optimal management of the endangered Sonoran pronghorn antelope population in Organ Pipe Cactus National Monument (OPCNM), Cabeza Prieta Wildlife Refuge, and the Barry M. Goldwater Range, Arizona, by enhancing current habitat maps through the incorporation of remote sensing data (Figure 1). The broader objective was to develop techniques that could be applied to the creation and refinement of habitat maps of other wildlife species.

BACKGROUND

The research conducted in the context of the attached papers evaluates indices derived from both multitemporal and fine spatial scale remote sensing data. These indices are used to characterize the spatial structure and temporal dynamics of landscapes for mapping and monitoring of wildlife habitat. The regional-scale remote sensing data is synthesized with a database of local-level wildlife sightings to produce temporal and spatial indices that characterize preferred habitat.

(5) The study provides information on the suitability of the proposed Sonoran Desert National Park for the pronghorn. The proposed park, which would be larger than Yellowstone and the Grand Canyon combined, encompasses the current OPCNM, the Cabeza Prieta National Wildlife Refuge, and the Barry M. Goldwater Airforce Range. Results of this study may help scientists and politicians understand the anticipated benefits of protecting such a large and continuous tract of land.

Sonoran pronghorn antelope sighting data were extracted from a database of over 2715 records collected between November 1994 and October 1999 by the Arizona Game and Fish Department (Hervert et al., 1996) and provided to us by OPCNM personnel. Although the locations of the antelope may not necessarily reflect ideal habitat, the operating assumption is that the majority of the sightings occur in landscapes that are in some sense preferred. Field recordings present in the database are:

- Month, day and year of the sighting.
- Name of the observer.
- Identifier number of the collared Pronghorn antelope.
- Location of the Pronghorn sighting in latitude/longitude and Universal Transverse Mercator (UTM) coordinates. [The datum for the UTM coordinates was not specified and is assumed to be NAD27.]
- Number of bucks, does, and fawns observed.
- Size of group observed.
- An indicator that a group contains more than one collared animal.

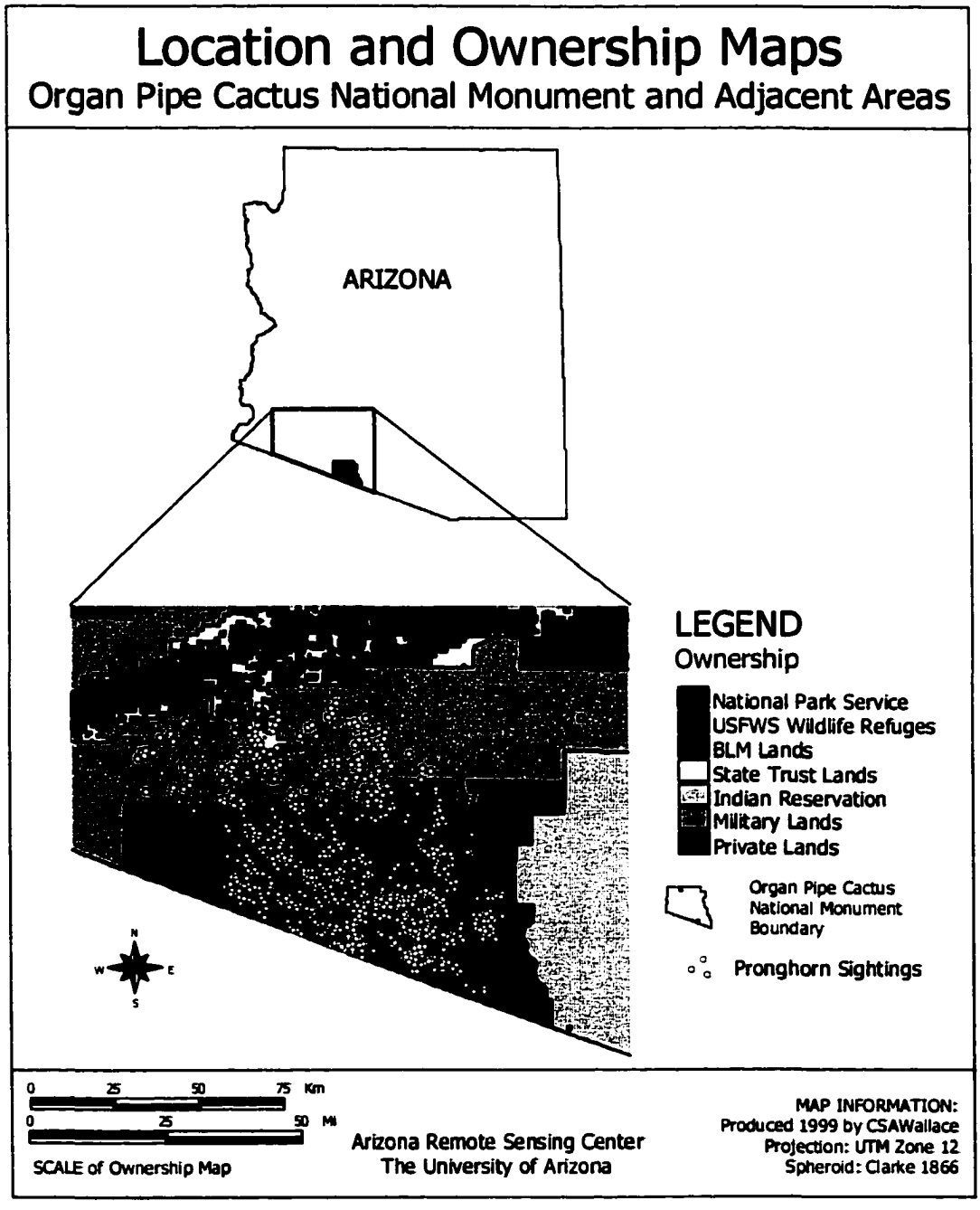


Figure 1: Primary study area: the habitat region of the endangered Sonoran pronghorn antelope population in Organ Pipe Cactus National Monument, Cabeza Prieta Wildlife Refuge, and the Barry M. Goldwater Range, Arizona.

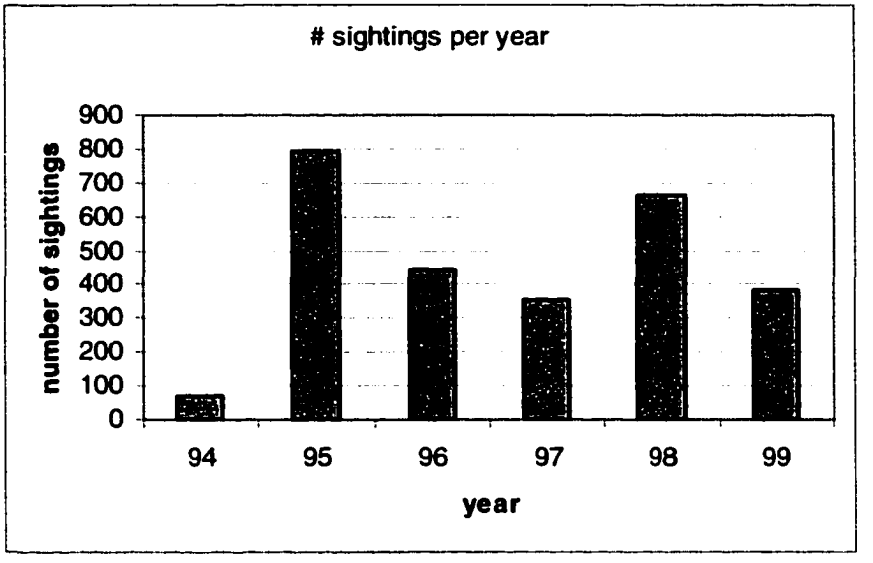


Figure 2: Histogram showing number of Sonoran pronghorn antelope sightings collected for each year.

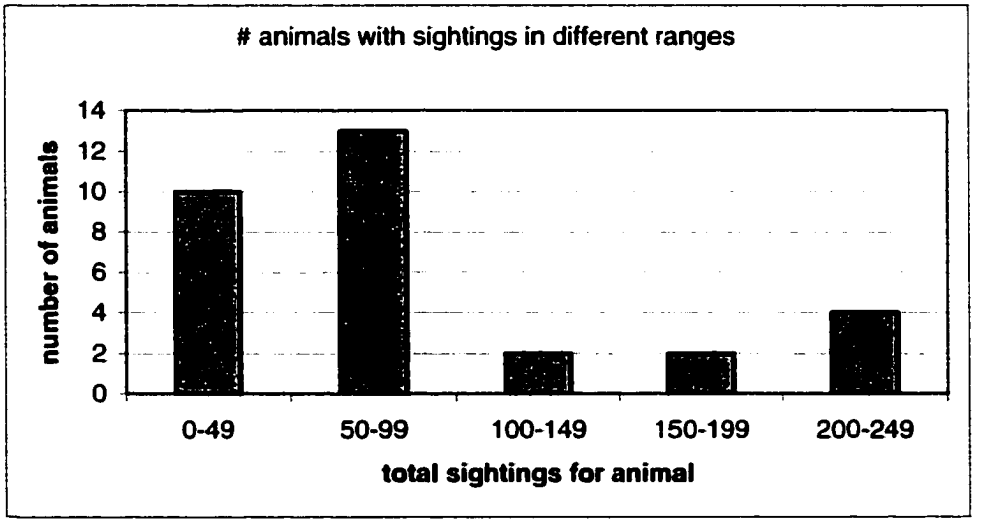


Figure 3: Histogram showing number of individual Sonoran pronghorn antelope per sightings category. Sightings categories stratify the total number of sightings into five equal classes, from 0 to 249, as shown.

Note that the radio-collared animals show an area of overlap in their home ranges as defined by MCP in an area centered about 20 kilometers west of the OPCNM boundary.

The Elk Population in Arizona

The second animal studied in this dissertation is the Arizona elk. This animal was chosen because the debate over the number of elk in Arizona over the past two decades has affected management. Consumptive users (i.e. hunters) have argued that the elk population should have priority over domestic livestock while ranchers argue that the population is too large and is detrimental to their operations. Unfortunately, empirical data from which to make informed decisions are limited.

Elk are known to migrate seasonally, and show a strong tendency to occupy the same geographical areas year after year (Brown, 1994). Although elk habitually return to the same geographical use areas, they do not show social fidelity to a particular group of animals beyond the family unit, which is limited to the cow, her calf, and possibly her yearling. A high degree of exchange of individuals between elk groups is observed on all seasonal ranges (Franklin et al.,1975). Weather patterns appear to trigger the migrations, especially during the late fall and early winter, where snow depth and forage availability triggers migration from summer to winter habitats (Brown, 1994). Because landscape condition is a reflection of weather patterns, remotely sensed information capturing landscape condition was investigated as a tool to provide information on elk migration patterns and abundance.

The Yellow-billed Cuckoo

The third animal studied is the yellow-billed cuckoo. This bird was selected for study because it is currently a candidate for listing as an endangered species. In 1993, the yellow-billed cuckoo was split into two subspecies, eastern: *Coccyzus americanus americanus* and western: *C. a.occidentalis* (Franzreb and Laymon, 1993). The western subspecies of the yellow-billed cuckoo has been in decline throughout the western United States for the past 60 years. The decline is variously ascribed to habitat destruction (Franzreb, 1987), decreased water tables (Phillips et al., 1964), and the use of pesticides (Gaines and Laymon, 1984; Rosenberg, et al., 1991). In February of 1998, 23 groups filed a petition with the United States Fish and Wildlife Service (USFWS) seeking endangered species status for the western subspecies.

The western yellow-billed cuckoo has been associated with cottonwood-willow dominated broadleaf deciduous riparian habitat (Hamilton and Hamilton, 1965; Gaines, 1974; Gaines and Laymon, 1984; Halterman, 1991). However, its exact distribution and habitat requirements in Arizona are not known. Because the contrast between the riparian vegetation species (in terms of vegetation type, density and condition) and the surrounding semi-arid desert vegetation species over much of the cuckoo range in Arizona is so great, we explore the application of remotely sensed information to capture landscape phenology and condition that may be related to seasonal cuckoo migration and nesting habitat.

- Hansen, et al. (2001) assessed the change and fragmentation of caribou habitat in the North Columbia Mountains of British Columbia. Two Landsat scenes were independently classified: an MSS scene from 1975 and a TM scene from 1997, with accuracies of 89.5% and 91.8%, respectively. Comparative analyses of selected spatial metrics for the two time periods revealed the 1997 image displays an overall decrease in abundance of winter habitat; reduced habitat patch geometric complexity, mean proximity, and core area; and increased habitat patch density and edge density.
- Franklin, et al. (2002) incorporated derived images from Landsat TM with topographic data from digital elevation models (DEMs) and GIS inventory data within an Evidential Reasoning (ER) classifier to produce maps of detailed vegetation and land cover for support of grizzly bear habitat mapping in Alberta Canada. They found that combining the multi-source data produces much higher classification accuracies than obtained using the various input data layers alone (85% vs. 45-50% accuracy, respectively), and that the ER classifier performs better than the commonly applied Maximum Likelihood classifier, which produced an accuracy of 71%.

A few wildlife studies have looked at the multi-temporal information from the NOAA AVHRR NDVI data sets. For example, Verlinden and Masago (1997) associated the measures of grass greenness and wildlife species densities obtained from ground and aerial surveys to a single coincident NOAA AVHRR NDVI composite image in the

Kalahari of Botswana. The authors concluded that the density and distribution of some wildlife species correspond to the greenness of preferred vegetation. A recent study by Oindo and Skidmore (2002) examines the relationship between species richness in Kenya (both plants and mammals) and interannual NDVI parameters. The authors show that higher average NDVI corresponds to lower species richness, whereas higher standard deviation and coefficient of variation correspond to higher species richness. They conclude that the NDVI parameters capture environmental factors influencing species richness.

Measuring Spatial Structure using Remote Sensing

A common technique to improve land cover maps produced from remotely sensed imagery is the inclusion of image texture using standard classification techniques (Jensen, 1996). However, recent studies have incorporated more sophisticated measures of spatial autocorrelation using the variogram function of geostatistics. Geostatistics is a class of spatial statistical techniques commonly used to characterize the spatial autocorrelation found in natural systems (Isaaks and Srivastava, 1989; Cressie, 1993). The spatial patterns of a region can be characterized quantitatively by the experimental variogram, which is derived by calculating one-half the average squared data value difference for every pair of data locations and plotting this value against the distance between the data pairs, expressed formally as:

$$\text{gamma(h)} = \frac{1}{2}\frac{1}{N}\sum_{i=1}^{N}[Z(x_i) - Z(x_i + h)]^2 \qquad (1)$$

Where: gamma(h) is the value of the variogram at lag h,
x_i is a data location, h is a lag vector,
$Z(x_i)$ is the data value at location x_i, and
N is the number of data pairs spaced a distance and direction h apart.

The resulting experimental variogram is then used to fit an appropriate theoretical model variogram having specific mathematical properties (Figure 5). The model variogram is fitted visually by the analyst, and each proposed model is evaluated by calculating a "goodness of fit" measure until a suitable fit to the experimental variogram is achieved. The nature of the variability characteristic of the data set is reflected in the overall shape of the model variogram (Isaaks and Srivastava, 1989; Webster and Oliver, 1990). Parameters of the model variogram include information on the distance over which data are correlated (the range); the level of random variation within the data (the nugget); and the total variation present (the sill).

Studies that have applied geostatistics to synthetic and actual images have found direct ties between several scene characteristics and the behavior of the variogram (e.g. Woodcock et al., 1988a and b). For example, the density of the image objects is related to the height of the variogram sill, the size of the image objects is related to the range of the variogram, and the variance in the size distribution of objects is related to the roundness of the variogram near the sill. In addition, the studies show that increasing the image pixel size results in lowered sill value; increased range; and increased nugget.

Other studies use the variogram for land cover mapping applications. For example, geostatistics can discriminate between different vegetation communities in

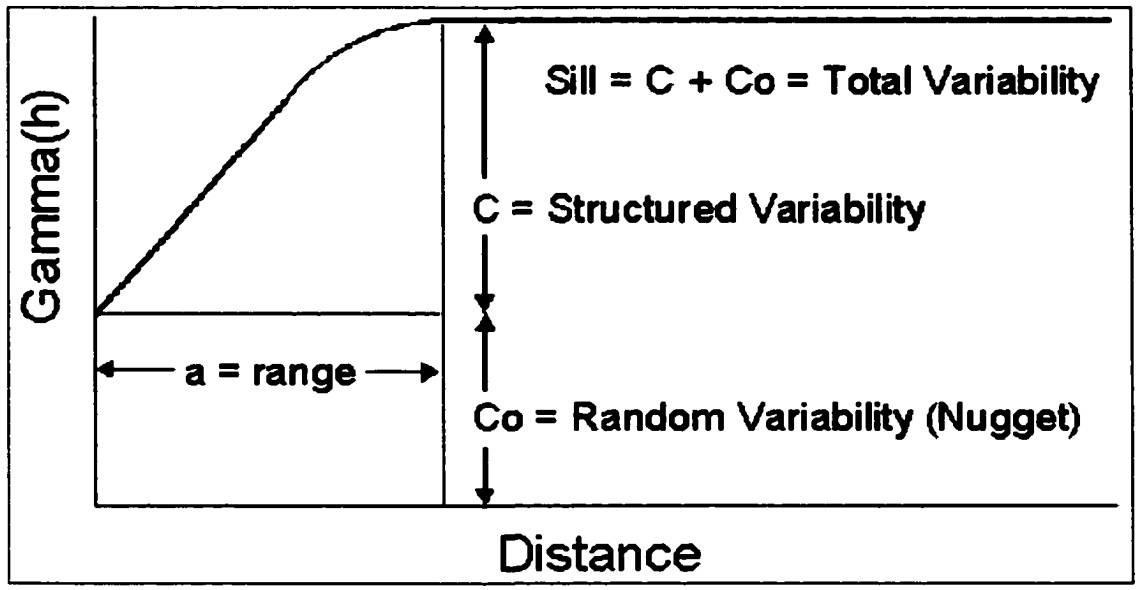


Figure 5: Example of a "Typical Variogram", after Cressie (1993). The horizontal axis plots lag distance, which is the directional distance between pairs of data points. The vertical axis plots the variogram value, termed "gamma", which is 1/2 the average squared difference in data values for all data pairs separated by the corresponding lag distance. Parameters of the model variogram include information on the distance over which data are correlated (the range); the level of random variation within the data (the nugget); and the total variation present (the sill).

remotely-sensed images of arid and semi-arid regions (Wallace et al., 1999; Phinn, 1997). These studies found that the spatial patterns of plant distributions within vegetation communities can be quantified by the variogram parameters if the image pixel sizes are less than three to five meters. In another study, Herzfeld (1996) uses geostatistics to enhance subtle topological features on a submarine fan. She produced new images by calculating the variogram at various lagged distances for all pixels in a sonar image.

atmospheric effects. After 1992, geometric registration of the individual AVHRR bands was accomplished prior to calculating NDVI images. The NDVI is calculated as follows:

$$NDVI = \frac{NIR - red}{NIR + red} \qquad (2)$$

where:
- *NDVI* is the Normalized Difference Vegetation Index
- *NIR* is the reflectance captured by the AVHRR sensor in the near-infrared region of the electromagnetic spectrum
- *red* is the reflectance captured by the AVHRR sensor in the visible red region of the electromagnetic spectrum

NDVI values range from −1 to 1, with non-land surfaces (such as water and snow) typically having negative values and land surfaces typically having positive values. As landscapes become more densely vegetated, the NDVI trends to 1. The calculated NDVI values, which range from −1 to 1 are linearly rescaled into the range 0 to 200, resulting in an 8-bit data value for image processing applications (Eidenshink, 1992).

Temporal profiles of NDVI values derived from AVHRR NDVI data clearly reveal the phenology of vegetated landscapes (Tucker, 1985; Hirosawa et al., 1996; Reed and Sayler 1997). Figure 6 shows examples of the temporal profiles of these data for selected landscapes in a 200x200-km area surrounding Tucson, Arizona. These classes are 2 of 20 that were derived using unsupervised classification of the 62-layer image containing all available 2-week NDVI composites for the years 1993, 1994, and 1995. Even in this relatively small region we find that landscape phenologies range from those with maximum green-up in the summer (Class 17) to those with maximums in the winter

(Class 9). There is also a broad range in the amount of overall vegetation cover, from the barren Wilcox Playa to the densely vegetated "sky-island" coniferous forests.

As described in the literature, measures of landscape temporal dynamics are obtained from time-series of NDVI images in several ways, including:

- Calculation of basic statistical measures, such as the mean, standard deviation, and coefficient of variation (COV). Such measures are typically calculated for the period of one year or one season and compared across years (e.g. Fuller, 1998; Weiss et al., 2001). A variation on this is the calculation of z-score statistics, which normalize the distribution of data values in each image (spatially) and across time to identify anomalous regions representing, for example, extreme change (Yool, pers. Comm.). Application of Standardized Principal Components Analysis (SPCA) to the multi-temporal data set (Eastman and Fulk, 1993; Hirosawa et al., 1996). SPCA is a data compression technique that projects the information content (i.e. the variability) of a data set onto fewer axes called components. For long time series of AVHRR bi-weekly NDVI composites, the first principal component (PC1) represents the characteristic NDVI, and can be shown to be equivalent to the average NDVI for the time period analyzed. The second principal component (PC2) represents "seasonality," and the extremes of PC2 correspond to landscapes with summer green-up and those with winter green-up.

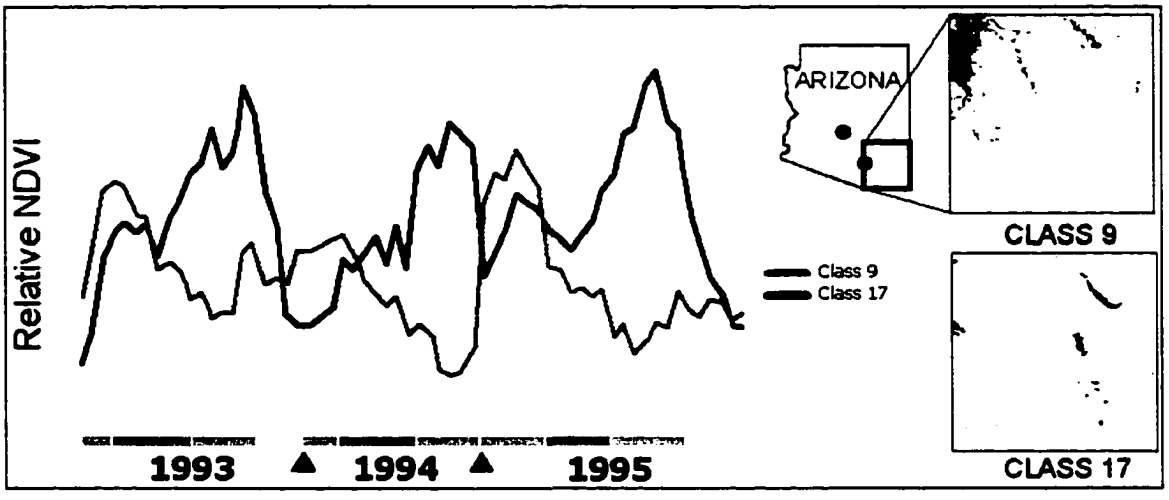


Figure 6: Graph shows the average temporal profile of the pixels composing the two classes shown at right. Class 9 corresponds to the blue profile and has maximum NDVI values in the winter. Class 17 corresponds to the red profile and has maximum NDVI values in the summer. The profile plots relative NDVI on the vertical axis against time on the horizontal axis for the years 1993, 1994 and 1995. The seasons of the year are colored as follows: Winter (JFM) = blue; Spring (AMJ) = green; Summer (JAS) = red; and Fall (OND) = yellow. Note that 1994 contains no data for the fall season.

- Application of curve-fitting techniques to the temporal profile of the NDVI at each pixel. Olsson and Eklundh (1994) applied Fourier series analysis to these data for a single year and extracted the first and second harmonics for the entire African continent to identify landscapes that are mono- or bi-modal. Fourier transforms have also been applied to mapping fire burn areas by extracting the high-frequency phase values from multi-temporal images representing several years of data (D. Johnson, 1999). Other curve-fitting techniques were used by Reed and Sayler (1997), who extracted the onset of greenness, time of peak greenness, duration of the growing season, and the time-integrated NDVI for the state of Colorado. Similar metrics were

derived by Myneni, et al. (1997) for all terrestrial vegetation during the years 1981 through 1991 and used to interpret vegetation responses to global warming.

These studies demonstrate that valuable indices of the temporal dynamics of vegetated landscapes can be derived from long time-series AVHRR NDVI data. In many studies, indices are compared across time to identify trends within landscapes (Hirosawa et al., 1996; Lambin and Strahler, 1994). Such comparisons permit the recognition of landscape stability or evolution and help quantify change and predict its potential consequences. These same indices have also been used to map the landscape (Tucker, 1985; Loveland et al., 1995), to monitor regional to global vegetation patterns (Marsh et al., 1992; Sellers et al. 1994), and to characterize the variability within and between recognized units (Hirosawa et al., 1996).

Statistical Analyses

To create species habitat models the technique utilized contrasted known animal locations (animal sightings or points sampling of known habitat patches) with points randomly selected from some specified subset of the study area. The various temporal or spatial measures derived from remotely sensed data are evaluated for inclusion into the models by determining if a difference can be demonstrated between the measures at known animal locations compared to the measures at random locations.

For most of the analyses, this difference is demonstrated statistically by using a Student's t-test for the difference of means under the assumption of equal variance, using

a significance level of $\alpha = .05$ (Davis, 1986; Salkind, 2000). This test evaluates the null hypothesis that the difference in the means for the two distinct populations is zero, utilizing the t-distributed test-statistic:

$$\frac{\frac{1}{n_1}\sum_{i=1}^{n_1} X_{1i} - \frac{1}{n_2}\sum_{i=1}^{n_2} X_{2i} - (\mu_1 - \mu_2)}{\sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}\left[\frac{n_1+n_2}{n_1 n_2}\right]}} \sim t_{n_1+n_2-2} \qquad (3)$$

where:
X_{1i} are data values from Sample 1 (for example, the actual sightings locations)
X_{2i} are the data values from Sample 2 (for example, the random locations)
n_1 is the number of data values in Sample 1
n_2 is the number of data values in Sample 2
s_1^2 is the estimated variance of X in Population 1
s_2^2 is the estimated variance of X in Population 2
μ_1 is the mean of X in Population 1
μ_2 is the mean of X in Population 2

If the null hypothesis is true, i.e., $\mu_1 = \mu_2$, the observed value of the test statistic takes on the form:

$$\frac{\bar{x}_1 - \bar{x}_2}{\sqrt{\frac{(n_1-1)s_1^2 + (n_2-1)s_2^2}{n_1+n_2-2}\left[\frac{n_1+n_2}{n_1 n_2}\right]}} \qquad (4)$$

If data are highly skewed, a non-parametric test for difference of sample means can be employed that is analogous to the t-test. We apply the Mann-Whitney test for difference in means, which takes the two data sets of n_1 and n_2 samples, combines them, orders the values from smallest to largest, and then assigns a rank to each of the values

from 1 to $n_1 + n_2$. The test for significance compares the mean rank of the samples for each of the two original groups (see http://www.tufts.edu/~gdallal/npar.htm).

Finally, we apply chi-squared tests to evaluate whether the nominal data sets preferentially group the actual sightings data into particular categories as compared to a random arrangement. This test is applied in this research to both identify data layers for input to habitat modeling and to validate produced preference maps. The chi-squared statistic is calculated as follows:

$$\chi^2 = \sum \frac{(O - E)^2}{E} \tag{5}$$

where:
- χ^2 is the chi-square value
- O is the observed frequency
- E is the expected frequency (from Davis, 1986)

The null hypothesis for the chi-squared test is that there is no difference in the proportion or frequency of occurrences in each category. For this study, we are dealing with mapped land cover units, such as vegetation or habitat preference. In such analyses, we would expect that the data, if randomly assigned, would occupy each category according to the size (area) of the category.

The calculated chi-square statistic is compared to a set of critical values to determine whether the null hypothesis should be accepted or rejected. Again, critical values are constructed for various significance levels and depend on the degrees of freedom for your test. The degrees of freedom in a chi-squared test is one less than the number of nominal categories partitioning the data. For example, a chi-square test

evaluating a habitat preference map consisting of three categories: Highly preferred, moderately preferred, and not preferred, would possess two degrees of freedom.

OBJECTIVES

The research goal of this study was to derive information from remotely sensed data that might help managers understand the spatial and temporal distribution of wildlife in Arizona, as a consequence of both seasonal variation and extreme climatic events (i.e., El Niño and La Niña). Current habitat studies commonly use remotely sensed images for land cover mapping (Patterson and Yool, 1998; Jensen, 1996; Price, 1992) and many studies enhance digital image classification using textural or multi-temporal information (Jensen, 1996; Weiss et al., 2001; Herzfeld, 1996). Studies of wildlife habitat also commonly include remotely sensed imagery, but often from only a single source (Verlinden and Masago, 1997; Austin et al., 1996).

My study expands on existing habitat studies because I extract information from images representing several time periods as well as data from two sensors with very different spatial resolutions: the AVHRR (1.1 km) and (1 and 4 meter). These two sensors reveal very different perspectives of the landscape: the AVHRR providing frequent sampling across time to enable analysis of landscape phenology, and the providing detailed spatial information at the scale of individual trees or bushes to enable analysis of vegetation structure. Such synergy of information enhances the ability to delineate wildlife habitat extents and define season-specific habitat characteristics.

In addition, the majority of previous studies have sought to use the remotely sensed data to approximate and extrapolate mapped units defined by researchers on the ground (e.g. Hansen et al., 2001; Franklin et al., 2002; Austin et al., 1996). Such studies define habitat preference by incorporating the mapped units into a model. In contrast, this study uses the distribution of the wildlife sightings within the data space of the various derived temporal and spatial measures to define habitat preferences. By doing this, I make no assumptions about the nature or presence of any natural boundaries between different landscapes that may be sensed as distinctive by the study animal. I am also able to convert the various derived measures to a common scale of standard deviation distances using the statistics of the values observed at the wildlife sightings, as follows:

$$X_{new} = \left| \frac{X_{old} - \mu}{\sigma} \right| \qquad (6)$$

where:
- X_{new} is the new value in standard deviation distance units
- X_{old} is the original derived measure
- μ is the mean of the set of derived measures at the wildlife locations
- σ is the standard deviation of the set of derived measures at the wildlife locations

The common scale allows for comparisons of and calculations between observed preferences for either different derived measures or the same measure for different years.

DISSERTATION FORMAT

This dissertation takes the form of four pre-publication papers, which are found in the Appendices. The overall goal of these papers was to extract information from

remotely sensed data that will help define habitat preferences of wildlife and inform natural resource management efforts.

The first paper evaluates and compares the result of using different geospatial analysis techniques to characterize Sonoran pronghorn antelope (*Antilocapra americana sonoriensis*) habitat in southwestern Arizona. Jessica Walker, a University of Arizona Geography student, developed an initial model of pronghorn habitat in a GIS-environment for her Master's thesis by applying logistic regression to field-based measurements of landscape characteristics available for OPCNM (Walker, 2000). To compare with this model, I developed other habitat models independently by extracting and analyzing measures of landscape temporal dynamics from NOAA AVHRR NDVI data at known pronghorn locations. These data possess fine temporal resolution (daily acquisition) but coarse spatial resolution (1 kilometer). Landscape dynamics were derived using standardized principal component (PCA) and Fourier analysis of the AVHRR data. The U.S. Department of the Interior, Organ Pipe Cactus National Monument and the U.S. Department of Defense, U.S. Air Force sponsored this research.

The second paper presents methods of extracting measures of landscape vegetation structure from imagery using geostatistics. This research was designed originally to provide spatially detailed information within the region identified in the preceding study using the ubiquitous and inexpensive digital orthophoto quarter quadrangle (DOQQ) data produced by the U.S. Geological Survey (USGS), and available for this study through the Arizona Regional Image Archive (ARIA) website (http://aria.arizona.edu/). The intended research is depicted as a flowchart in Figure 7.

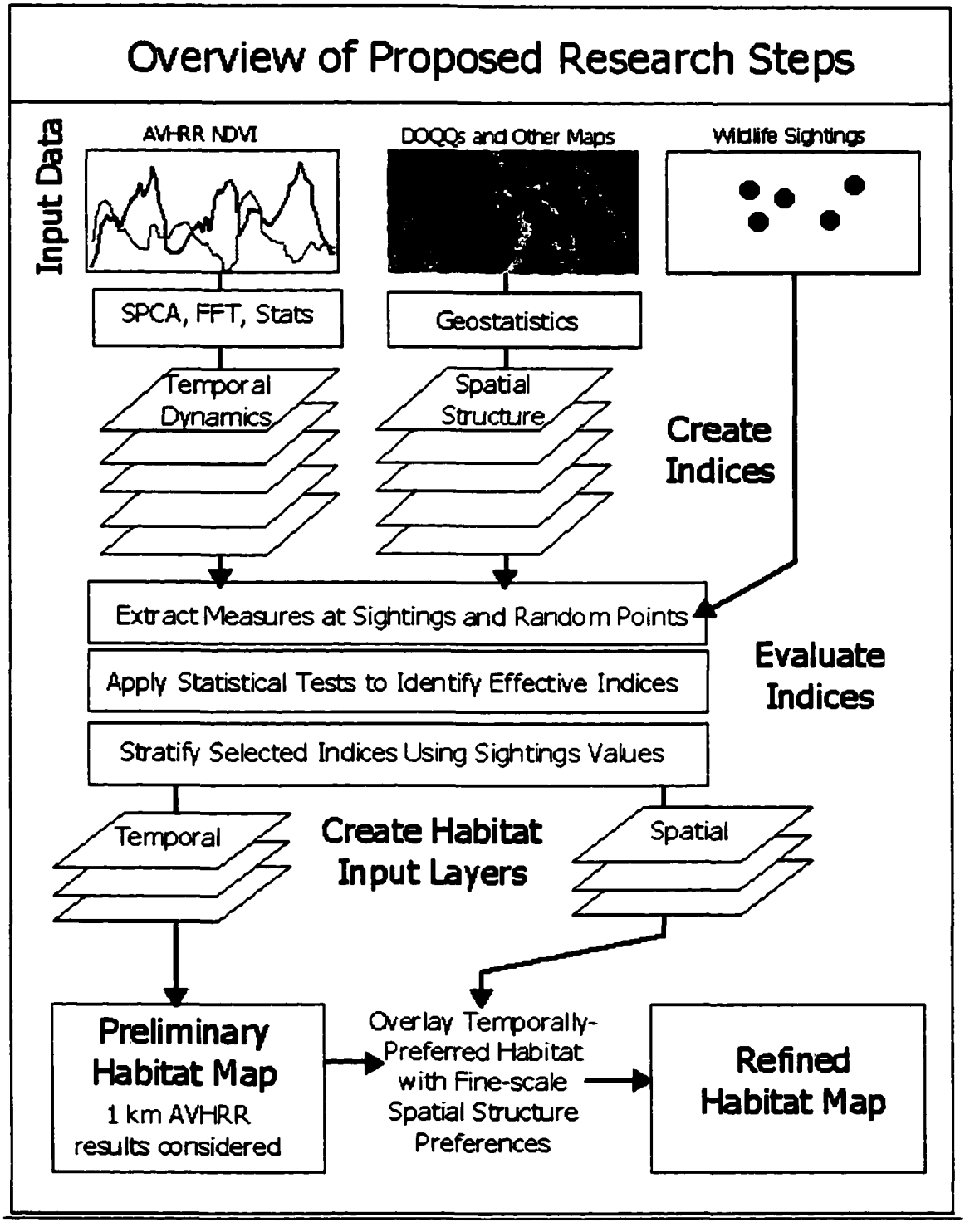


Figure 7: Overview of proposed research steps

When the geostatistical analysis was applied to the OPCNM DOQQs, however, it revealed a noise pattern in the form of broad, sub-horizontal striping in the derived images. Although the origin of this noise remains unclear, consultations with various individuals at the USGS and on campus led us to conclude that removal of this noise would be difficult if not impossible. We therefore elected to purchase data available for a small area northwest of OPCNM on the Barry M. Goldwater Range.

Paper 2 presents the results of the geostatistical analysis of landscape structure for this small subset of the original study area. We extracted and evaluated measures that characterize the spatial structure of vegetated landscapes from satellite imagery for mapping endangered Sonoran pronghorn antelope habitat. Information derived from the satellite imagery was used to stratify the landscape into favorable regions defined by coupling the structural measures with Sonoran pronghorn antelope sighting data.

The third paper explored relationships between satellite greenness measures derived from NOAA AVHRR data and elk population dynamics in selected Arizona game management units. The research was designed to provide information that might help managers understand the distribution of elk in Arizona as a consequence of both seasonal variation and extreme climatic events. The paper summarizes the results of an interdisciplinary study accomplished during the first year under the auspices of the Synergy project. The Synergy project at the University of Arizona is one of several research efforts around the country sponsored by the National Aeronautics and Space Administration (NASA) through contract with Raytheon to demonstrate the applicability of EOS data to the non-research community. Since its inception, the Synergy project has

grown dramatically; it is now focused on serving useful and useable data and products to natural resource managers and decision makers through its RangeView web site: http://rangeview.arizona.edu/.

The fourth paper "AVHRR Imagery Analysis and Habitat Modeling for a Yellow-billed Cuckoo (*Coccyzus americanus occidentalis*) Survey in Arizona" presents the results of another interdisciplinary effort accomplished at the U of A. The branch of the Cooperative Ecosystem Studies Unit (a partnership between the USGS, BLM, US Bureau of Reclamation, NFS, and NPS) housed at the University of Arizona was conducting a census of the Yellow-Billed Cuckoo in Arizona to evaluate its status as a potential threatened or endangered species. Although sightings data were sparse, the riparian habitat of the bird is very distinctive in AVHRR satellite imagery, prompting us to explore these data as a tool for reconnaissance mapping to guide field census surveys. In my position as second author, I accomplished the AVHRR image processing to extract and evaluate measures of temporal dynamics related to predicting potential cuckoo habitat.

The second chapter of the dissertation summarizes significant findings of the dissertation papers, suggests directions for future research, and presents a brief concluding statement.

CHAPTER 2

PRESENT STUDY

SUMMARY

The methods, results and conclusions of this study are described in the papers appended to this dissertation. The following is a summary of the most important findings presented in these papers.

Paper 1 first demonstrated the viability of using high temporal resolution AVHRR data to characterize and map season-specific habitat of the Sonoran pronghorn antelope. Measures of temporal dynamics extracted from these data using standardized principal components analysis and Fourier transforms were effective at discriminating between pronghorn sightings and random points. In the OPCNM, the AVHRR-based models (produced by coupling the derived temporal measures with a set of pronghorn sightings) show a visual and statistical correspondence to GIS-based models (produced by coupling maps of detailed field-based data with the sightings). These results for the local area of OPCNM provide assurance that the AVHRR models capture landscape characteristics that have biophysical meaning. These findings are significant because the AVHRR models derived in this study can be constructed easily, provided that a sightings data set is available and can be matched with coincident AVHRR NDVI imagery (available for the US from 1989 to present). Our results also show that it is vital to consider seasonal differences when characterizing the habitat of this highly mobile species. Significant differences in the landscapes accessed by the Sonoran pronghorn antelope are identified

estimates against the simplest measure of average greenness for each quarter (quarter 1 = JFM, quarter 2 = AMJ, quarter 3 = JAS, and quarter 4 = OND). The regression results suggested a correlation between the seasonal greenness of certain vegetation classes and the annual elk population numbers. We constructed maps showing areas within each vegetation class of each GMU that tend to be greener than average through time, assuming there is a verifiable physical correspondence between the greenness values and elk population numbers, (i.e., that elk seek out areas of higher greenness within their habitat range). Comparison with a set of elk sightings data showed a visual fit between these greener than average areas and the sightings, suggesting these maps could be used to predict elk locations. A main purpose of this study was to resolve debate over the number, distribution and impact of elk herds across the state. Not only did this study provided valuable information toward this goal, it also fostered significant new interaction and productive discussion among the wide variety of stakeholders in the Arizona elk debate.

Paper 4 further expanded the applicability of AVHRR-based measures of landscape temporal dynamics by applying them to a bird species. The analysis combined both area-based data (known habitat patches) and point-based data (bird sightings). Preliminary field studies to validate these models were mixed; the model effectively captured cuckoo habitat in some locations but missed important structural features in others. Data collected in 1999 at the Bill Williams National Wildlife Refuge show actual nesting locations to be in agreement with the habitat prediction model. In other locations the model effectively captured the correct mix of preferred vegetation species (i.e.,

cottonwood, willow) but did not discriminate the correct structural attributes (i.e., layered canopy with mature trees). Regardless, the model is proving to be a valuable tool for predicting potential yellow-billed cuckoo habitat in Arizona and for guiding field studies. Current efforts seek to incorporate spatial information into this model to help define important structural attributes.

FUTURE WORK

Results presented in this dissertation suggest many promising avenues for future research. Although statistically significant relationships between wildlife parameters (sightings, population, habitat patches) and the temporal or spatial measures derived from remotely sensed data are defined, these relationships must be verified and field checked in order to establish definitively the link between an interesting result and a useful tool for natural resource management. In particular, there is a need to explore the biophysical meaning of the temporal and spatial measures in relation to the vegetation communities and landscapes they represent. This understanding would provide insight to the nature of the link between a species and its environment as identified in these results.

With respect to the temporal analyses, it is well established that multi-temporal AVHRR data are effective for mapping vegetation communities (Moody and Johnson, 2001; Jakubauskas and Legates, 2001; Reed and Sayler, 1997; Tucker, et al., 1985). The natural range of variability within recognized vegetation communities, however, should be established (both intra- and inter-annually). The pronghorn analysis (Appendix A) created AVHRR-based models to compare with GIS-based models by compositing

models created for individual years, but the differences noted between the yearly models could be explored more thoroughly. Knowledge about the interannual variability of vegetation communities in the region would inform this effort. Extrapolating the same analyses to other geographic areas, other species, or other temporal data types (e.g. MODIS data) are other areas for future research.

With respect to the spatial analysis, it would be useful to explore the effect of changing the size of the analysis window, which was set at 25x25 meters for this study, and changing the size of the study area, which was limited by data availability. Rapid increases in computer performance capabilities may allow the analyses to be accomplished using a "moving window", which would preserve the spatial resolution of the input image, as opposed to "tiling" the image. Applying a supervised classification to the set of spatial measures could also be investigated. For example, the classifier could be trained to map important landscape types, such as areas dominated by invasive species or fawning habitat. Again, extrapolating the same analyses to other geographic areas, other species, or other data types (e.g. airborne data) are other areas for future research.

Finally, future research should seek to develop truly integrated models that synthesize the temporal and spatial information. The original vision for this research was to accomplish such integration through a simple overlay, where landscapes identified as temporally preferred by the Sonoran pronghorn antelope would be extracted and further refined into preference classes based on their spatial characteristics (see Chapter 1, Figure 7). This goal was not realized because we could not find suitable regional fine spatial resolution data. Development of accessible (i.e., understandable) multi-dimensional

characterizations of the landscape by effectively combining varied types of information (temporal, spatial, spectral, wildlife) in a GIS environment should clearly assist natural resource management.

CONCLUSIONS

The research accomplished for this dissertation used various mathematical and statistical techniques to extract and evaluate measures of landscape temporal dynamics and spatial structure from remotely sensed data for the purpose of mapping wildlife habitat. By coupling the landscape measures gleaned from the remotely sensed data with various sets of animal sightings and population data, effective models of habitat preference were created. The synoptic view of the landscape and vegetation availability provided by these satellite data can inform resource management efforts. The ability to characterize the details of spatial structure and temporal dynamics of habitat using repeatable remote sensing data allows closer monitoring of the relationship between a species and its landscape. For threatened and endangered species, this enhanced source of environmental data may prove invaluable.

Gaines D.A. and S.A. Laymon. (1984) Decline status and preservation of the yellow-billed cuckoo in California. *Western Birds* 15:49-80.

Halterman M. (1991) *Distribution and habitat use of the yellow-billed cuckoo (Coccyzus americanus occidentalis) on the Sacramento River, California, 1987-1990.* MS Thesis, California State University, Chico, California.

Hamilton W.J. and M.E. Hamilton. (1965) Breeding characteristics of yellow-billed cuckoos in Arizona. Pages 405-432 *in* Proceedings of the California Academy of Sciences. Fourth Series.

Hanson, M.J., S.E. Franklin, C.G. Woudsma, and M. Peterson (2001) Caribou habitat mapping and fragmentation analysis using Landsat MSS, TM, and GIS data in the North Columbia Mountains, British Columbia. Canada, *Remote Sensing of Environment* 77:50-65.

Herr, A.M. and L.P. Queen (1997) Crane Habitat Evaluation Using GIS and Remote Sensing. *Photogrammetric Engineering and Remote Sensing.* 59(10):1531-1538.

Hervert, J. J., R. S. Henry, M. T. Brown, and L. A. Piest, (1996) *Sonoran pronghorn population monitoring: progress report*, Nongame and Endangered Wildlife Program Technical Report 110. Arizona Game and Fish Department, Phoenix, Arizona.

Herzfeld, U.C. and C.A. Higginson (1996) Automated Geostatistical Seafloor Classification - Principles, Parameters, Feature Vectors, and Discrimination Criteria: *Computers & Geosciences*, 22:35-52.

Hirosawa, Y., S.E. Marsh and D.H. Kliman (1996) Application of Standardized Principal Component Analysis to Land-Cover Characterization Using Multitemporal AVHRR Data. *Remote Sensing of Environment.* 58:267-281.

Hooge, P.N. and B. Eichenlaub (1997) *Animal movement extension to arcview. ver. 1.1.* Alaska Science Center - Biological Science Office, U.S. Geological Survey, Anchorage, AK, USA.

Huete, A.R., and Jackson, R.D. (1987) Suitability of Spectral Indices for Evaluating Vegetation Characteristics on Arid Rangelands. *Remote Sensing of Environment.* 23:213-232

Hughes, K.S. (1991) *Sonoran Pronghorn Use of Habitat in Southwest Arizona.* The University of Arizona, M.S. Thesis. 86 p.

Isaaks, E.H., and R.M. Srivastava (1989) *An Introduction to Applied Geostatistics;* Oxford University Press, New York.

Jakubauskas, M.E. and Legates, D.R. (2001) Harmonic Analysis of Time-Series AVHRR NDVI Data for Characterizing US Great Plains Land Use/Land Cover. *Proceedings of the XIXth International Society for Photogrammetry and Remote Sensing Congress*, Amsterdam, The Netherlands, July 18-23, 2000.

Jensen, J.R. (1996) *Introductory Digital Image Processing: A Remote Sensing Perspective, 2nd edition.* New Jersey: Prentice-Hall, Inc. 316 p.

Johnson, D. (1999) Automated Identification of Fire Scars and Other Surface Anomalies Using Frequency-Based Analysis of AVHRR Data. *Talk presented at the annual meeting of the Association of American Geographer's,* Honolulu, Hawaii 23-27 March 1999.

Kilgore, M.J. and W.S. Fairbanks (1997) Winter Habitat Selection by Reintroduced Pronghorn on Antelope Island, Great Salt Lake, Utah. *Great Basin Naturalist* 57(2):149-154

Lambin, E.F., A.H. Strahler (1994) Change-Vector Analysis in Multitemporal Space: A Tool To Detect and Categorize Land-Cover Change Processes Using High Temporal Resolution Satellite Data. *Remote Sensing of Environment.* 48:231-244

Loveland, T.R., J.W. Merchant, J.F. Brown, D.O. Ohlen, B.C. Reed, P. Olson and J. Hutchinson (1995) Seasonal Land-Cover Regions of the United States. *Annals of the Association of American Geographers.* 85(2):339-355

Marsh, S.E., J.L. Walsh, C.T. Lee, L. Beck, and C.F. Hutchinson (1992) Comparison of Multitemporal NOAA-AVHRR and SPOT-XS Data for Mapping Land Cover Dynamics in the West African Sahel. *International Journal of Remote Sensing,* 13(16):2997-3016.

Moody, Aaron and David M. Johnson (2001) Land-Surface Phenologies from AVHRR Using the Discrete Fourier Transform, *Remote Sensing of Environment,* 75(3):305-323.

Myneni, R.B., C.D. Keeling, C.J. Tucker, G. Asrar, and R.R. Nemani (1997) Increased Plant Growth in the Northern High Latitudes from 1981-1991; *Internet:* http://www/forestry.umt.edu/ntsg/nature/

Oindo and Skidmore (2002) Interannual Variability of NDVI and Species Richness I Kenya. *International Journal of Remote Sensing.* v23(2):285-298.

Olsson, L. and L. Eklundh (1994) Fourier Series for Analysis of Temporal Sequences of Satellite Sensor Imagery. *International Journal of Remote Sensing,* 15(18):3735-3741.

Patterson, M.W. and Yool, S.R. (1998) Mapping fire-induced vegetation mortality using Landsat Thematic Mapper data: A comparison of linear transformation techniques; *Remote Sensing of Environment* 65: 132-142

Phillips, A., J. Marshall and G. Marshall (1964) *The birds of Arizona*. University of Arizona Press, Tucson, Arizona. 220 pp.

Phinn, S, J. Franklin, A. Hope, D. Stowe and L. Huenneke (1996) Biomass Distribution Mapping Using Airborne Digital Video Imagery and Spatial Statistics in a Semi-Arid Environment. *Journal of Environmental Management;* 47:139-164

Reed, B.C. and K. Sayler (1997) A Method for Deriving Phenological Metrics from Satellite Data, Colorado 1991-1995; *Internet:* http:// geochange.er.usgs.gov /sw/impacts/biology / Phenological-CO

Rosenberg K.V., R.D. Ohmart, W.C. Hunter, B.W. Anderson (1991) Birds of the lower Colorado Valley. University of Arizona Press, Tucson, Arizona. 416 pp.

Salkind, Neil J. (2000) *Statistics for People Who (Think They) Hate Statistics*. Sage Publications, Inc., Thousand Oaks, California, 385p.

Sellers, P.J., C.J. Tucker, G.J. Collatz, S.O. Los, C.O. Justice, D.A. Dazlich, and D.A. Randall (1994) A Global 1 by 1 NDVI Data Set for Climate Studies. Part 2: The Generation of Global Fields of Terrestrial Biophysical Parameters from the NDVI. *International Journal of Remote Sensing*, 15(17):3519-3545.

Tucker, C.J., J.R.G. Townshend and T.E. Goff (1985) African Land-Cover Classification Using Satellite Data. *Science*. 227(4685):369-375.

Verlinden, A. and R. Masago (1997) Satellite Remote Sensing of Habitat Suitability for Ungulates and Ostrich in the Kahlahari of Botswana. *Journal of Arid Environments* 35:563-574.

Walker, J.J. (2000) *A GIS Model of Sonoran Pronghorn Antelope Habitat in Organ Pipe Cactus National Monument, Arizona*. Master's Thesis, Department of Geography and Regional Development, The University of Arizona, Tucson, 121p.

Wallace, C.S.A., J. M. Watts, and S. R. Yool (2000) Characterizing the Landscape Structure of Vegetation Communities in the Mojave Desert Using Geostatistical Techniques. *Computers and Geosciences*, v26:397-410.

Webster, R. and M.A. Oliver (1990) *Statistical Methods in Soil and Land Resource Survey;* Oxford University Press, New York

Weiss, E., S.E. Marsh, and E.S. Pfirman (2001) Application of NOAA-AVHRR NDVI time-series data to assess changes in Saudi Arabia's rangelands, *International Journal of Remote Sensing* 63:1005-1027.

Woodcock, C.E., A.H. Strahler and D.B. Jupp (1988a) The Use of Variograms in Remote Sensing: I. Scene Models and Simulated Images; *Remote Sensing of Environment;* 25:323-348.

Woodcock, C.E., A.H. Strahler and D.B. Jupp (1988b) The Use of Variograms in Remote Sensing: II. Real Digital Images; *Remote Sensing of Environment;* 25:349-379

Modeling Sonoran Pronghorn Antelope Habitat: Comparing Logistic Regression of Geospatial Data with Principal Component and Fourier Analysis of Multi-temporal Remote Sensing Data

Cynthia S.A. Wallace
Jessica J. Walker
Stuart E. Marsh

The University of Arizona
Office of Arid Lands Studies
Arizona Remote Sensing Center
1955 East Sixth Street
Tucson, Arizona 85719
520-621-8574

October 23, 2002

To Be Submitted To: *International Journal of Remote Sensing*

ABSTRACT

This study evaluates and compares the result of using different geospatial analysis techniques to characterize Sonoran pronghorn antelope (*Antilocapra americana sonoriensis*) habitat in southwestern Arizona. The first habitat model was developed in a GIS-environment by applying logistic regression to field-based measurements of landscape characteristics available for Organ Pipe Cactus National Monument. Other habitat models were developed by extracting and analyzing measures of landscape temporal dynamics from NOAA AVHRR NDVI data at known pronghorn locations. Temporal measures were derived using standardized principal component (PCA) and Fourier analysis of the AVHRR data. These measures capture dynamics that include vegetation phenology, and were selected to provide regional information on vegetation cover and communities. Both GIS-based and remote sensing-based modeling approaches relied on an extensive database of pronghorn sightings, and examined the influence of seasonal differences on migration patterns and model results.

Validation results confirm the effectiveness of all models. The GIS models produce a final predictive ability between 66 and 86 percent. The AVHRR models discriminate between sightings and random points, and show a reasonable fit to the GIS models, both visually and statistically. These findings are significant because the AVHRR models can be easily constructed, provided that a sightings data set is available and can be matched with coincident AVHRR NDVI imagery.

1.0 INTRODUCTION

Understanding the relationship between a wildlife species and its environment is a prerequisite to making informed resource management decisions. With increased knowledge, natural resource managers can better anticipate effects of environmental change and take steps to mitigate their consequences. Without this knowledge, incorrect decisions could be made under misguided assumptions about the links between a wildlife species and its environment. Wildlife in danger of extinction are particularly vulnerable, since the margin of error may be low.

If landscapes can be characterized by indices of topography and land cover or temporal dynamics derived from repeatable remote sensing data, the relationship between a species and its landscape can be more closely monitored and evaluated. Effective use of geospatial and remote sensing data tied to areas where detailed knowledge of wildlife locations have been established enables extrapolation of information across space and through time, permitting the characterization of landscapes at regional scales.

Our research was designed to provide additional insight into the optimal management of the Sonoran pronghorn antelope subspecies (*Antilocapra americana sonoriensis)* population in southern Arizona. The pronghorn population faces an uncertain future and is listed as endangered by the United States Fish and Wildlife Service (USFWS), which means that it is vulnerable to extinction throughout all or a significant portion of its habitat range. Environmental resource managers are particularly concerned about the pronghorn population because there are conflicting theories about why its numbers have remained stagnant. The need for more specific and detailed information about the pronghorn's relationship to its habitat is critical, since the population crisis cannot be addressed appropriately until its underlying cause is identified.

We hope to provide additional insight into this problem by producing new habitat maps from a multi-year database of pronghorn sightings using two geospatial analysis techniques. The first approach developed a logistic regression model of habitat based upon detailed field maps of surface topography and cover. The second approach developed models of habitat based upon standardized principal component and Fourier analyses of the Normalized Difference Vegetation Index (NDVI) generated from multitemporal National Oceanic and Atmospheric Administration (NOAA) Advanced Very High Resolution Radiometer (AVHRR) data.

Current habitat maps created through traditional, field-based mapping techniques generally present a static picture of the environment. A vegetation map, for example, shows the type and location of vegetation at the time of data collection. Typically, such maps relay no information about spatial variability in size and densities of the vegetation, its condition, its response to climatic events, or its seasonal phenology. However, wildlife

respond directly to these changes. The use of multitemporal remote sensing data can enhance habitat maps by incorporating dynamic information that the static maps lack. We therefore use a combination of spatially detailed geospatial map layers and digital images in this study to produce a multidimensional characterization of the landscape.

2.0 BACKGROUND

GIS and Habitat Mapping

GIS technology has facilitated the development of habitat suitability models through the ability to integrate and analyze multiple, complex data layers. In studies that combine both nominal and continuous data sets, such as vegetation classes and elevation, logistic regression is commonly applied (e.g. Bian 1997; van Manen 1997). Logistic regression is analogous to classic linear regression, except that it uses a binary dependent variable, such as presence or absence (Garson, 2000). The logistic regression model uses measurements on the independent variables to estimate the likelihood of attaining one of two dependent variable states:

$$p_i = \frac{e^{\alpha+\beta|X|i+\beta 2X\,2i+\cdots+\beta nXni}}{1+e^{\alpha+\beta|X|i+\beta 2X\,2i+\cdots+\beta nXni}} \qquad (1)$$

where $0 \leq P_i \leq 1$ is the probability that an individual i belongs to group 1 (for example 'presence' of n animal), the X_{ji} are the j-th independent variables for case i, the β_j are the weights, and α is a constant (Kvamme, 1985). The model is linearized by taking the natural log of the odds ratio:

$$\ln\frac{p}{1-p} \qquad (2)$$

resulting in the linear logistic model:

$$Y = \alpha + \beta_1 X_{1i} + \beta_2 X_{2i} + \ldots\ldots + \beta_n X_{ni} \qquad (3)$$

If a variable has a non-zero weight β_i, it is an indication that the variable has an effect on the outcome of the presence/absence probability, given the simultaneous interactions with the other included variables. A positive coefficient indicates an increase in the variable whereas a negative coefficient indicates the reverse (Kvamme, 1985).

Studies that applied logistic regression to wildlife habitat modeling include:

- Periar and Itama (1991), who used environmental and location variables to predict the presence/absence of the Mt. Graham red squirrel, and then employed the resulting probability map as a proxy for habitat suitability.
- Bian and West (1997), who constructed a model of elk calving/not-calving conditions in prairie environments using telemetry data for animal locations and various environmental data, including land cover types and distance to anthropogenic features; and
- Mladenoff et. al (1995), who assessed the importance of various landscape-scale features in timber wolf habitat in Wisconsin. They found that the single parameter of road density predicts wolf habitat although they considered a suite of variables that included human population density and proximity, deer population, roads, land cover type, ownership, and patch statistics (size, edge, fractal dimension, contagion).

There is considerable precedence for the application of logistic regression to wildlife studies in the literature. Because several detailed field-based data layers were available for the Organ Pipe Cactus National Monument, and they included both nominal and continuous data (vegetation, soils, topography), logistic regression was selected as an appropriate and logical modeling choice for this local area.

Remote Sensing and Habitat Mapping

Studies of wildlife habitat commonly include remotely sensed imagery, but often from only a single source. Several studies have classified a single Landsat TM image for inclusion in a Geographic Information System (GIS) to define habitat, (Herr and Queen, 1997; Austin et al., 1996; Hansen, et al., 2001; and Franklin, et al., 2002. In another study, Verlinden and Masago (1997) associate the measures of grass greenness and wildlife species densities obtained from ground and aerial surveys to a single coincident NOAA AVHRR Normalized Difference Vegetation Index (NDVI) composite image in the Kalahari of Botswana. The authors concluded that the density and distribution of some wildlife species correspond to the greenness of preferred vegetation.

This study expands on these existing habitat studies because we extract information from images representing several time periods using NOAA AVHRR NDVI data. These data provide frequent sampling across time to enable analysis of landscape temporal dynamics, which includes vegetation phenology. Since an important component of a species habitat is vegetation, which provides both food and shelter, it is reasonable to consider a remotely sensed variable tied to vegetation. Furthermore, since wildlife respond to seasonal and interannual differences in vegetation condition (for example, the green-up of annual forage species), multi-temporal data provided by the NOAA AVHRR sensor are a logical choice.

The ability to extract meaningful information on the type and condition of vegetation from NOAA AVHRR NDVI data is well documented in recent studies. Several studies

extract measures describing the landscape temporal dynamics from these data through application of various image-processing techniques. As described in the literature, measures of landscape temporal dynamics are obtained from time-series of NDVI images in several ways, including:

- Calculation of basic statistical measures, such as the mean, standard deviation, and coefficient of variation (COV). Such measures are typically calculated for the period of one year or one season and compared across years (for example, Fuller, 1998; Weiss et al., 2001). A variation on this is the calculation of z-score statistics, which normalize the distribution of data values in each image (spatially) and across time to identify anomalous regions representing, for example, extreme change (Yool, pers. comm.).

- Application of Standardized Principal Components Analysis (SPCA) to the multi-temporal data set (Eastman and Fulk, 1993; Hirosawa et al., 1996). SPCA is a data compression technique that projects the information content (i.e. the variability) of a data set onto fewer axes called components. For long time series of AVHRR bi-weekly NDVI composites, the first principal component (PC1) represents the characteristic NDVI, and can be shown to be equivalent to the average NDVI for the time period analyzed. The second principal component (PC2) represents "seasonality," and the extremes of PC2 correspond to landscapes with summer green-up and those with winter green-up.

- Application of curve-fitting techniques to the temporal profile of the NDVI at each pixel. Olsson and Eklundh (1994) applied Fourier series analysis to these data for a single year and extracted the first and second harmonics for the entire African continent to identify landscapes that are mono- or bi-modal. Fourier transforms have also been applied to mapping fire burn areas by extracting the high-frequency phase values from a multi-temporal image representing several years of data (D. Johnson, 1999). Other curve-fitting techniques are used by Reed and Sayler (1997), who extracted the onset of greenness, time of peak greenness, duration of the growing season, and the time-integrated NDVI for the state of Colorado. Similar metrics were derived by Myneni et al. (1997) for all terrestrial vegetation during the years 1981 through 1991 and used to interpret vegetation responses to global warming.

The above studies demonstrate that valuable indices of the temporal dynamics of vegetated landscapes can be derived from long time-series AVHRR NDVI data. In many studies, indices are compared across time to identify trends within landscapes (Hirosawa et al., 1996; Lambin and Strahler, 1994). Such comparisons permit the recognition of landscape stability or evolution and help quantify change and predict its potential consequences. These same indices have also been used to map the landscape (Moody and Johnson, 2001; Tucker, 1985; Loveland et al., 1995), to monitor regional to global vegetation patterns (Jakubauskas and Legates, 2001; Reed and Sayler, 1997; Marsh et al.,

1992), and to characterize the variability within and between recognized units (Hirosawa et al., 1996).

3.0 RESEARCH OBJECTIVES

Our objectives were to:

(1). *Develop a GIS-based model of preferred habitat for the local Organ Pipe Cactus National Monument (OPCNM).* Detailed, field-based environmental variables (e.g., vegetation, soils, and topography), available only for the local area of the OPCNM, are integrated with the pronghorn sightings information to create seasonal models of habitat preference using statistical analyses, including logistic regression. Logistic regression compares a dichotomous, dependent variable (such as presence or absence) with a set of independent variables that can include both categorical and numerical data (Bian, 1997). Since some of the independent variables (e.g., vegetation) are qualitative, this particular statistical technique was appropriate for the project. A number of similar habitat studies have successfully implemented this method (Bian, 1997; van Manen, 1997).

(2). *Develop a remote sensing-based model of preferred habitat for the region including OPCNM.* This model is developed by integrating NOAA AVHRR NDVI remotely sensed (RS) data with the pronghorn sightings database to create seasonal models of habitat preference. Measures describing the landscape temporal dynamics are derived from the AVHRR NDVI data through application of basic statistics, standardized principal components analysis (PCA) of NDVI (Hirosawa et al., 1996; Lambin and Strahler, 1994) and Fourier transforms (Olsson and Ecklundh 1994, Jakubauskas and Legates, 2001, Johnson and Moody, 2001). In this study, we evaluate the temporal characteristics of the landscapes occupied by pronghorn sightings, as measured by the indices extracted from the remote sensing data, to determine whether the pronghorn preferentially occupy areas of specific character compared to random locations. The key indices of landscape dynamics, which provide information on vegetation type and phenology, extracted from the remote sensing data are then integrated in a geographic information system (GIS) environment to produce derived habitat preference maps. These maps are evaluated by testing how well they predict the locations of pronghorn sightings withheld from the model development.

(3). *Compare and evaluate GIS-based and remote sensing-based Sonoran pronghorn antelope habitat models.* The results of the remote sensing based modeling are further evaluated by comparing them to seasonal habitat maps derived from detailed field-based GIS data (objective 1). Since both the GIS and RS-based models produce continuous measures of habitat preference, we can compare the models both visually as well as statistically where they overlap in the

local area of OPCNM. Finally, we examine the characteristics of the generated models, to contribute toward an integrated model of the seasonal habitat of the Sonoran pronghorn antelope.

3.0 STUDY AREA AND DATA SOURCES

3.1 Study Area

The project study area encompasses the Organ Pipe Cactus National Monument, Cabeza Prieta Wildlife Refuge, and the Barry M. Goldwater Range in southwestern Arizona (Figure 1). The study, therefore, provides information on the suitability of the proposed Sonoran Desert National Park for the Sonoran pronghorn antelope (*Antilocapra americana sonoriensis).* The proposed park, which would be larger than Yellowstone and the Grand Canyon combined, encompasses the current OPCNM, the Cabeza Prieta National Wildlife Refuge, and the Barry M. Goldwater Airforce Range. Results of this study may help scientists and politicians understand the anticipated benefits of protecting such a large and continuous tract of land

3.2 Data Sources

Four main data sources have been compiled to accomplish this study:

1. *Sonoran Pronghorn Antelope Sightings*: Sonoran pronghorn antelope sighting data were extracted from a database of over 2715 records collected between November 1994 and October 1999 by the Arizona Game and Fish Department (Hervert et al. 1996) and provided to us by OPCNM personnel. Although the locations of the antelope may not necessarily reflect ideal habitat, the operating assumption is that the majority of the sightings occur in landscapes that are in some sense preferred. Field recordings present in the database are:

- Month, day and year of the sighting.
- Name of the observer.
- Identifier number of the collared Pronghorn antelope.
- Location of the Pronghorn sighting in latitude/longitude and Universal Transverse Mercator (UTM) coordinates.
- Number of bucks, does, and fawns observed.
- Size of group observed.
- An indicator that a group contains more than one collared animal.
- Observed activity.
- Additional comments.

Figure 2 shows the distribution of pronghorn sightings by year. During the fall of 1994, 22 animals were radio-collared and located on a weekly basis using fixed-wing aircraft. The majority of the sightings were collected during 1995, and the attrition of sightings observed in the following two years is due to animal mortality. By the fall of 1997, only seven of the original 22 animals were still alive. Nine additional animals were collared between December 1997 and January 1998, accounting for the spike of additional sightings observed in 1998. Additional deaths in 1998 are reflected in the decline in sightings observed for 1999.

Figure 3 shows the number of animals with different quantities of sightings. Four individual pronghorn were responsible for 200 or more of the sightings in the database, and a total of eight pronghorn were each observed at least 100 times. The remaining 23 pronghorn were observed less than 100 times each.

The territory accessed by individual radio-collared pronghorn is shown in Figure 4, which maps the overlapping extents of pronghorn home ranges. For this map, home ranges were calculated as the minimum convex polygon (MCP) containing all the sightings for an individual animal (Hooge and Eichenlaub, 1997). Examples of these MCPs for the three pronghorn responsible for the greatest number of individual sightings (17, 14, and 9) are outlined in Figure 4.

2. *Detailed GIS data for the OPCNM*: OPCNM personnel made available an extensive collection of data sets of the biophysical parameters of the monument. The data sets include coverage of: Geology, Riparian Vegetation, Roads, Soils, Topography, Vegetation, and Washes. We also generated a Digital Elevation Model (DEM) at a resolution of 30 meters by combining the topographic contours, washes, and streams data in a GIS environment. Several variables derived from these coverages, such as slope, aspect and topographic ruggedness, are included in the analysis. Ultimately the 30-m DEM set the spatial resolution of the GIS-based model development.

3. *Multi-temporal NOAA Advanced Very High Resolution Radiometer (AVHRR) Data*: The temporal dynamics of the landscape were characterized using NOAA AVHRR data. The AVHRR data have a relatively coarse spatial resolution of 1-kilometer, but their temporal resolution is extremely fine, since the data are collected daily (EDC 1994). Vegetation communities that are mapped on a regional scale can vary due to differing vegetation type, density, or vigor. Data that are sampled frequently will capture ephemeral events such as the green-up of annual vegetation. For this study, we use the Normalized Difference Vegetation Index (NDVI), which is calculated from the reflectance values of the visible (red) and near-infrared (NIR) regions of the electromagnetic spectrum and is sensitive to various attributes of vegetation (e.g., Huete and Jackson 1987).

Biweekly maximum value composites of the NOAA AVHRR NDVI for the conterminous United States were obtained from the U.S. Geological Survey's (USGS) Earth Resources Observation Systems (EROS) Data Center (EDC), in Sioux Falls, South Dakota. These NDVI data products are compiled for the conterminous U.S. for the years 1989 through 2002 (EDC, 1994) and are available for this study through the Office of Arid Lands Studies at the University of Arizona.

To construct this data set, the daily AVHRR channel data are radiometrically calibrated and corrected for solar illumination angles. The calibration coefficients used after 1990 vary temporally and are based on stable desert targets, thereby incorporating the effects of sensor degradation. After 1992, geometric registration of the individual AVHRR bands was accomplished prior to calculating NDVI images. The NDVI is calculated as follows:

$$NDVI = \frac{NIR - red}{NIR + red} \tag{4}$$

where: *NDVI* is the Normalized Difference Vegetation Index
NIR is the reflectance captured by the AVHRR sensor in the near-infrared region of the electromagnetic spectrum
red is the reflectance captured by the AVHRR sensor in the visible red region of the electromagnetic spectrum

NDVI values range from –1 to 1, with non-land surfaces (such as water and snow) typically assuming negative values and land surfaces typically assuming positive values. As landscapes become more densely vegetated, the NDVI trends to 1. The calculated NDVI values, which range from –1 to 1 are linearly rescaled into the range 0 to 200, resulting in an 8-bit data value for image processing applications (Eidenshink, 1992).

Recording the highest NDVI value detected for each pixel during the two-week period creates the biweekly maximum value composite product. This process produces a relatively cloud-free image of the earth's surface. The NDVI data for an individual year, therefore, consists of 26 images, each recording the highest NDVI value observed during that particular two-week period.

4.0 RESEARCH METHODS

4.1 Habitat Modeling Using Logistic Regression of GIS-based Data

Specific methods utilized in the logistic regression analysis involved: (1) investigation of the relationship of pronghorn sightings by season to specific environmental attributes (soil, geology, and vegetation); physical characteristics such as (elevation, slope, and aspect); and anthropogenic disturbances (roads or unnatural barriers); (2) development of statistical techniques to identify and quantify the relationship between the pronghorn sightings and the variables most strongly characteristic of the pronghorn locations; (3) generation of indices that delimited the most likely areas of future pronghorn occupation, using the most influential landscape variables; (4) application of logistic regression techniques to create a map of occupation likelihood (presence/absence); (5) examination of the resultant likelihood map of pronghorn occupation and explanation of any departures from expected occupation rates in particular areas; and (6) evaluation of the particular procedures and methods employed to create the GIS model for their efficacy at predicting the pronghorn locations.

The following variables were chosen for consideration: the 531 locations of radio-collared pronghorn recorded between November 6, 1994 and October 24, 1999 within Organ Pipe Cactus National Monument (OPCNM): geology, soil, and vegetation classes within OPCNM, roads, washes, and USGS elevation contours. Using those datasets, additional layers were derived: a 30m digital elevation model (DEM), distance to roads, distance to washes, slope, aspect, curvature, topographic relief, and visibility. These variables were chosen because they were considered to be potentially relevant to the pronghorn's choice of habitat, given knowledge of their habits and preferences (Hughes, 1991).

Creation of Random Location Set: As in a prior pronghorn study (Marsh et al., 1999), a database of random locations was also generated for the purposes of statistical comparison. The random sightings, drawn from sites unoccupied by the pronghorn, functioned as a 'neutral' or control group against which to contrast the traits of the subject animal. Theoretically, if the pronghorn are selective about their locations, the mean values of the environmental variables specific to their chosen locations should differ from the mean values of the same variables at random locations (Pereira and Itami, 1991).

For the pronghorn analysis, the region from which to draw random points was constructed by taking the cumulative area formed by equidistant "buffers" around each of the pronghorn sightings. We derived the buffer distance by halving the average daily distance (10km) a pronghorn might reasonably cover in a single day (Hervert et al. 1996). All elevations above 600m were also excluded from consideration, since pronghorn were never spotted in higher elevations. 590 points, rather than 531, were then chosen at

random from the resulting area, since a greater number of random points are required to capture the variation of the environment adequately (Kvamme, 1985).

The random and actual point locations were analyzed statistically to identify the most important landscape parameters that influenced pronghorn location, in order to use these variables as inputs in the final logistic regression model. A strong seasonal split was observed in the pronghorn location data. The average location of each month was calculated (i.e., the centroid of the sightings by month) and plotted, revealing two distinct clusters that we coded into the two seasons 'summer' (June, July, August, and September) and 'non-summer'. Because of this observed seasonality, the analysis was conducted on both an annual and seasonal basis. For validation purposes, approximately 20% of the sightings (100 annual; 50 seasonal) were selected randomly and withheld from each of the models.

Soil and Vegetation: These variables representing nominal data were evaluated statistically through the use of non-parametric chi-square tests. In all cases the sightings data produced significant chi-square statistics, showing that the pronghorn sightings are preferentially clustered within particular classes rather than being distributed uniformly across the landscape. Therefore both the soil and vegetation data sets were selected for inclusion in the logistic regression modeling.

Prior to their input into the model, the original 26 soil and 39 vegetation classes were regrouped into three preference categories to avoid over-fitting the data during the modeling process: any class was coded as "Preferred" (code = 1) if a disproportionately large number of pronghorn were present; "Avoided" (code = 2) if significantly fewer pronghorn were present than expected; and "No Preference" (code = 3) if the number observed was approximately equal to the number expected.

Aspect: The landscape parameter of aspect was treated as a categorical variable given the difficulty of accounting for the circularity of the data, in which 0 and 360 are equivalent. Accordingly, the 360-degree circle was apportioned into 8 quadrants of 45° each for analysis. The Mann-Whitney (Wilcoxon) W test, a non-parametric test analogous to the t-test, showed that the difference of means for the random and sightings points was significant. The aspect variable was, therefore, included in the logistic regression model.

Continuous Data: The continuous data sets were evaluated by using the Student's t-test for the difference of means under the assumption of equal variance, using a significance level of $\alpha = .05$ (Davis, 1986; Salkind, 2000). The t-test is applied to determine whether the means of the continuous data values at random locations and at sightings locations are the same. The test evaluates the null hypothesis that the difference in the means for the two distinct populations is zero, utilizing the t-distributed test-statistic. The t-test results indicate that the means of the random and sightings values are statistically different for all variables except mean distance to washes on either an annual or seasonal basis.

1. Basic statistics: This analysis considers the 26 values at each 1 km^2 pixel (one for each two-week composite) as a collection of data values. The measures calculated are the mean, the standard deviation, and the coefficient of variation. These represent measures of vegetation vigor or density, variability, and variability normalized by the mean, respectively.
2. Standardized principal component analysis (PCA): This analysis considers the data for each year as a stack of 26 images. The measures calculated are the first and second principal components (PC1 and PC2), which have been shown to represent vegetation amount and seasonality, respectively (Hirosawa et al., 1996; Lambin and Strahler, 1994).
3. Fourier analysis: In this analysis, the ordered values at each pixel, from layer 1 through layer 26, are input to Fourier analysis as a vector or waveform. The measures extracted are the additive term (i.e., the best fit of a flat line to the data) and the first frequency magnitude and phase (describing the best fit of a single sine wave to the data vector). These measures represent vegetation amount, variability and seasonality (Jakubauscas and Legate, 2001; Moody and Johnson, 2001).

The ability of the various temporal measures to map landscapes accessed by pronghorn was evaluated by first comparing values observed at actual pronghorn sightings with values at a set of random points. As in the GIS analysis, the random points were selected within a buffered zone of 5 km (1/2 the average daily travel distance of a pronghorn) around the sightings and masked to exclude landscapes not accessed by pronghorn (Walker, 2000). Without exception, t-tests showed that the sightings occupy statistically different landscapes than random points, with sightings located in pixels that are greener and more variable.

The ability of the various temporal measures to map the differences between the seasonal distributions of pronghorn was also evaluated using t-tests. We evaluated the difference between the pixels occupied in the "summer" and "non-summer" seasons. Based on these results, we determined that the measures derived via PCA and Fourier analysis are superior to the basic statistics method for discriminating between seasonal landscapes preferred by the pronghorn antelope.

To facilitate interpretation of the final results, we chose to create two independent models of seasonal pronghorn habitat: one using the PCA-derived temporal measures and the other using the Fourier-derived measures. These two sets of temporal measures are derived using techniques that produce uncorrelated terms, simplifying subsequent multivariate regressions and statistics. The two different sets of models were also evaluated independently, to determine which is preferable based upon validation results, comparisons to each other, and comparison to the GIS-based models. Derived temporal measures for the habitat modeling included:

Technique	*Derived Measure*	*Vegetation Interpretation*
PCA-derived:	PC1	Vegetation amount
	PC2	Vegetation seasonality
Fourier-derived:	Additive term	Vegetation amount
	First frequency phase greenness	Timing of maximum
	First frequency magnitude	Vegetation vigor

Mapping dynamics of landscapes preferred by pronghorn: The selected PCA-derived and Fourier-derived temporal measures were extracted for all pronghorn sightings points, and these data were sorted by year and by season (summer vs. non-summer). From each of these fifteen sets (five years of one annual set and two seasonal sets each), 20 percent of the sightings were randomly selected and removed for subsequent model validation. The models were then generated using the remaining 80 percent of the data as follows:

1. The mean and standard deviation of each data set were calculated. These statistics describe the range of values preferred by the pronghorn antelope for each of the derived temporal measures. Circular statistics were required to calculate the desired statistics for the Fourier phase values. Circular statistics convert each phase angle value into its component sine and cosine values, which are then manipulated to produce the desired measures.

2. The image of a derived temporal measure (e.g., PC1) was normalized using the descriptive statistics calculated above statistics by subtracting the mean from each pixel value and dividing by the standard deviation. This results in data values that describe the directional distance from the mean image value preferred by the antelope in standard deviation units.

3. Calculating the absolute value of the standard deviation units produces an image of standard deviation distance (SDD). In the SDD images, values close to zero define pixels that best match the temporal response of landscapes highly preferred by the pronghorn.

Converting the images of temporal measures to standard deviation distances (SDD) based on the statistics of the pixels corresponding to the pronghorn sightings accomplishes two things; First, it allows for direct calculations involving the various measures because the images are converted to the same units; Second, it allows otherwise abstract measures (in our case, the PCA-derived measures) to be compared from year to year. PCA analysis is scene-specific, so that values obtained for one image (with its particular geographic and

data-value extents) cannot be directly compared to any other image. By converting to SDD units, such comparisons are possible.

Seasonal and Annual Model Development: We elected to combine the images showing preferred pronghorn temporal values into models for each year by simply averaging the pertinent measures. Although statistical methods to optimize the weighting of the various measures could be applied, we felt that the simpler averaging facilitated year-to-year comparisons, provided a reasonable preliminary model to evaluate and refine if subsequently deemed appropriate, and kept the results more easily understood by potential end-users.

The PCA-derived models were created by averaging the corresponding SDD images of PC1 and PC2 for the year and season desired. This produces models in which values between 0 and 0.5 result if both PC1 and PC2 are within one standard deviation of the mean value preferred by the pronghorn, i.e., the pixel has both a favorable vegetation amount (PC1) and a favorable seasonality (PC2).

The Fourier-derived models were created by averaging the corresponding SDD images of the additive term, normalized first frequency phase, and normalized first frequency magnitude for the year and season desired. In this case, values close to zero represent pixels that possess favorable vegetation amount (additive term), timing of greenness (first frequency phase), and variability (first frequency magnitude).

Composite Model Development: Two sets of composite models were created. First, composite models were created to produce single models of each temporal measure (e.g., PC1) for the entire period. They were produced by averaging the corresponding SDD images for each of the temporal measures over all of the years included in the analysis, so that each year is equally weighted. The resulting composites reveal areas that are preferred across all the years with respect to the various attributes related to: vegetation amount (PC1, Fourier additive term), seasonality (PC2), timing of maximum vegetation greenness (Fourier first frequency phase), or landscape variability (Fourier first frequency magnitude).

Secondly, composite models of pronghorn habitat preference were also created to produce a single annual model and single models of each season for the entire period. They were produced by averaging the corresponding models for each of the years included in the analysis. The resulting composites reveal areas that are preferred across all the years, with each year equally weighted. Again, we felt that by weighting each year equally, the results can be evaluated more easily by potential end-users. In addition, we felt that equal weighting provides a reasonable preliminary model to evaluate and refine if subsequently deemed appropriate. For example, if it is determined that a particular year is more representative of "normal" conditions, it can be weighted accordingly. In these

composites, the units are in standard deviation distances, such that values near 0 represent areas that are close to the mean pronghorn preferred values in all of the four years.

Model Validation: One method to validate these models is to evaluate how well the models predict the locations of the pronghorn sightings not used in the model development. Validation was accomplished using the 20 percent of the original pronghorn sightings data points that were withheld for this purpose. The values of the three models for each year (annual, summer and non-summer) at the validation points were collected and tabulated. Ideally, if the model correctly predicts the validation points, their tabulated values will represent a sample from a normal distribution, such that approximately 68 percent of the points will be within one standard deviation of the mean and 95 percent within 2 standard deviations from the mean.

5.0 MODELING RESULTS

5.1 Results of Habitat Modeling Using Logistic Regression of GIS-based Data

Three final models resulted from the analysis: annual, summer, and non-summer. These models are shown in Figure 5.

The resultant annual model was of the form:

$$\text{Logit}(p) = 1.37092 - 0.00867996*\text{Elevation} + 0.000345963*\text{DistanceToRoad} + 0.283676*(\text{SoilCode}=1) + 0.990946*(\text{SoilCode}=2) + 0.992209*(\text{VegCode}=1) \quad (5)$$

where p is the probability of a pronghorn sighting (Walker, 2000).

As described in Kvamme (1985), positive coefficients indicate that increases in the corresponding variables are associated with a greater likelihood of presence (i.e., pronghorn sighting). By extension, the interpretation of a negative coefficient is that an increase in its corresponding variable contributes to the decreased likelihood of a pronghorn sighting. In the above equation, it is evident that an increase in elevation lessens the likelihood of a pronghorn presence, whereas an increase in the distance to roads heightens it. Likewise, the presence of "Preferred" soil and/or vegetation classes (Code=1) contributes favorably to a pronghorn prediction. The positive coefficients on the codes for the "Avoided" soil class (SoilCode=2) is surprising, given the distinction between the "Preferred" and "Avoided" training inputs of the model. One explanation for this weighting is that the determination of whether a class should be "Preferred" or "Avoided" was made on the basis of the expected numbers of pronghorn in those classes, compared to the number that was actually observed. Although there were fewer pronghorn than expected in the "Avoided" classes, those categories tended to host a

considerable number of pronghorn, if only because those land cover classes occupied large areas within OPCNM. Thus, they could contribute to the positive prediction.

We developed the following seasonal models (Walker, 2000):

The non-summer seasonal model took the form:

$$\text{Logit}(p) = 2.2383 - 0.00884412*\text{Elevation} + 0.000243428*\text{DistanceToRoad} + 0.846564*(\text{SoilCode}=1) + 1.06742*(\text{SoilCode}=2) + 0.781246*(\text{VegCode}=1) + 0.890129*(\text{VegCode}=2) \quad (6)$$

and the summer seasonal model took the form:

$$\text{Logit}(p) = 0.929478 - 0.00636021*\text{Elevation} + 0.00012443*\text{DistanceToRoad} + 0.896108*(\text{SoilCode}=1) + 2.35173*(\text{SoilCode}=2) + 0.695839*(\text{VegCode}=1) + 0.744043*(\text{VegCode}=2) \quad (7)$$

The success of the models is evaluated using the chi-square goodness-of-fit test. This test looks at the improvement offered by the model over the constant-only model. For the annual, non-summer and summer models, the goodness of fit test was significant at the 0.05 level, showing the models predict the observed data better than the constant-only model.

Model success is also evaluated using the prediction table, which shows the proportion of correct and incorrect classifications of the binary dependant variable. Each model (Figure 5) is tested by using it to predict the presence/absence of the data points that had been withheld from the model training. To do so, the data layers were first multiplied by their respective regression coefficients, summed together to obtain the linear predictor (Logit(p)), and then converted to probabilities using the formula:

$$P = \exp(\text{Logit}(p))/(1+\exp(\text{Logit}(p))) \quad (8)$$

The resultant grid of values potentially ranged from 0 to 1, representing the likelihood of occurrence (0 absence to 1 presence). A cut-off value is chosen to partition the continuous probability range into the binary 'presence/absence' classification. For this study, we used a cut-off of 0.5 as well as 0.45. The grid was then overlaid with the withheld validation points. Since the points were extracted from the original 531 observed pronghorn sightings, in a perfect model each point would land in a region identified as "presence."

The number of points that were accurately classified determined the predictive ability of the derived model (Table 1). At the more stringent 0.5 cut-off, the annual, non-summer, and summer models correctly predict 70, 66, and 76 percent of the test points, respectively. Reducing the cutoff to 0.45 generally improves the predictive capability of the models. At this cut-off, the annual, non-summer, and summer models predict 77, 74, and 86 percent of the test points (Walker, 2000).

Despite the high degree of mobility of the species and the extent of rangeland in the OPCNM, the pattern of observations indicates that the pronghorn prefer well-defined, specific areas within the Monument and revisit them on a consistent basis. Examination of the seasonal and annual predictive maps of pronghorn habitat reveals that the model fits the general patterns of pronghorn observations: it returns a reasonable overview of the preferred sites and highlights similar areas within OPCNM. Discrepancies fall into two general categories: errors of omission, in which pronghorn were not observed in areas that were considered to be preferred habitat, and errors of commission, in which pronghorn were observed in areas that were not identified as preferred habitat.

Errors of commission cannot be assessed properly, since it is impossible to determine if a pronghorn had visited or would ever occupy a given location (Hodgson et al. 1988). It is possible, however, to attempt to account for some of the more marked patterns of pronghorn locations in the Monument area. In doing so, there are a number of caveats to bear in mind. As Austin et al. (1996) discuss, such models are trained on environmental variables chosen for their likely biological significance, but the final measurements are derived primarily through statistical means. Therefore, the subsequent relationships between the pronghorn distribution and the underlying habitat variables of the predictive maps should not be assumed to represent causality. For instance, the changing, seasonally-dependent relationship between the increase in the distance from roads and the greater likelihood of pronghorn presence should not lead automatically to the assumption that pronghorn are less averse to roads in certain times than others. Although such a conclusion could be valid, depending on the seasonal patterns of traffic flow, the more likely explanation is that the pronghorn are reacting to another variable that is highly correlated with the roads parameter. Further analysis based on the inclusion of major roads adjacent to OPCNM could help clarify this issue.

The next concern arises from the fact that the model was trained using pronghorn observations within a subset of OPCNM territory, but the results were extrapolated and applied to the Monument as a whole. The predictive map thus reflects only the biases and preferences that were evident in the subset area, since no assumptions could be made about the pronghorn preference for soil and vegetation classes that were not present. Any such classes were assigned to the "No Preference" category, implying neither preference nor avoidance, but it is possible and likely that the pronghorn would discriminate among them in some way.

The fundamental premise of the model's application to a wider extent is that pronghorn behavior remains consistent over similar but separated landscapes. This generalization was necessary for the mapping objectives of this project, but it is by no means an unassailable assumption. A more rigorous test of the model's effectiveness would be to apply its results to areas outside of OPCNM where the pronghorn have also been sighted.

With these caveats in mind, it is possible to examine the models for trends and peculiarities. One of the most striking observations is that pronghorn are absent from a region in the southeastern corner of OPCNM that is identified as preferred habitat in the annual and both seasonal models. A possible explanation for the absence is that Arizona State Highway 85, which bisects the eastern edge of the Monument, effectively bars the population from moving eastward. During the entire span of pronghorn observations, not a single sighting occurred east of Highway 85. Compounding the possibly inhibiting factor of the highway is the presence of the border town of Lukeville and the activity around the busy border crossing. An additional explanation for the lack of pronghorn in the southeast corner is the absence of a habitat 'corridor' that would connect the area to the more inhabited territory in the west: according to the model, there is a sizable (~4 km) gap between the two areas that is not identified as preferred habitat.

The seasonal maps clearly show the difference in the areas accessed by the pronghorn. Most striking is the apparent shift of sightings from the low western floodplains in the non-summer months to the bajada slopes in the southwestern and north-central parts of OPCNM during the summer. The seasonal vegetation preferences of the pronghorn sightings seem to fit with observations of pronghorn diet (Hughes 1991). During the relatively cooler non-summer months, the pronghorn are found preferentially in the floodplains to the west, where they can access abundant forbs. During the harsher summer season, the pronghorn occupy the bajada slopes that host perennial chain-fruit cholla (*Opuntia fulgida*), a common and reliable food source during the dry season.

5.2 Results of Habitat Modeling Using Remote Sensing-based Data

Seasonal and Annual Models: To illustrate the modeling steps applied to each of the five years, examples are shown for one year. Figure 6 shows the development of the PCA-derived annual model for 1995. The upper images are the normalized 1995 PC1 and normalized 1995 PC2 images. The bottom image is the PCA-derived 1995 annual model. Figure 7 shows the development of the Fourier-derived annual model for 1995. The upper images are the normalized 1995 additive term, normalized 1995 first frequency phase, and normalized 1995 first frequency magnitude images. The bottom image is the Fourier-derived 1995 annual model.

The resulting models derived for each of the five years are shown in Figures 8 through 12. Inspection of these models reveals a distinct seasonal pattern. During the summer months,

the modeled landscapes are more restricted and are shifted toward the east. This pattern is especially apparent by looking at the position of the OPCNM boundary relative to the summer and non-summer landscapes. The more restricted pattern of the summer months could reflect a preference of the pronghorn for more areally-restricted vegetation types, or it could reflect the fact that the summer models are based on only four months of sightings data rather than the eight months of data included in the non-summer models.

Comparing the PCA-based models with the analogous Fourier-based models reveal that the PCA models tend to generalize the landscape more than the Fourier models (for example, see Figure 8). This is not surprising, since the Fourier models are based on three temporal measures (additive term, magnitude, and phase) to describe the amount and seasonality of vegetation whereas the PCA models are based on two measures (PC1 and PC2). Whether the discrimination offered by the Fourier model is more accurate or more useful than the generalization provided by the PCA model should be evaluated in the field by those using the models.

Seasonal and Annual Models Validation Results: The seasonal and annual models for the individual years were validated using the pronghorn sightings withheld from model development. The validation results (Table 2) show that the models do a reasonable job of predicting the validation points except for the year 1997, in which the Fourier-based models and the non-summer PCA-based model predict a high percentage of validation point values to be greater than two standard deviations from the model mean. These data were revisited and the reason for this discrepancy is unclear. The fact that 1997 is the year represented by the fewest number of sightings may partially explain this phenomenon. Inspection of the images for the Fourier models of 1997 (Figure 10) confirm that this year is highly anomalous compared to the other four years. Furthermore, the normalized images for the 1997 (not shown) reveal the measure that deviates most from the other years is the Fourier phase, and this unusual phase pattern appears to dominate all the 1997 Fourier-based models. Based on these results, the year 1997 was excluded from the composite model creation.

Comparison of the validation results for the Fourier-based and the PCA-based models (Table 2) shows that the PCA-based models tend to include more validation points in the smaller standard deviation distance classes. For example, the PCA annual model for 1995 shows 68% of the validation points within 1 standard deviation of the model mean compared to the Fourier annual model with only 55%. Although these validation results might suggest the PCA-derived models are preferred over the Fourier models, the fact (as noted above) that the PCA-based models tend to generalize the landscape more than the Fourier-based models can also explain this difference.

Composite Models of Temporal Measures: PCA-based composite PC1 and PC2 models are shown in Figures 13 and 14. Fourier-based composite additive term, first

frequency phase, and first frequency magnitude models are shown in Figures 15, 16, and 17.

Inspection of these composites reveals a variety of patterns. For each temporal measure, the summer landscape is more spatially restricted and tends to be shifted somewhat east of the non-summer landscape. This is the same pattern noted for each of the yearly models. The measures of vegetation amount (PC1 and Fourier additive term) map out nearly identical landscapes, and show the largest contiguous areas that overlap from year to year (i.e., with low average SDD). The two seasonality measures, PC2 and Fourier magnitude, also map landscapes that are strikingly similar, but PC2 maps a slightly more extensive pattern. The Fourier phase measure maps the most unique and disseminated landscape, with the least amount of year-to-year overlap (i.e., fewest areas of very low SDD). The phase landscapes do, however, occupy the same general regions as the other seasonality measures. These composites show that both methods, PCA and Fourier, effectively capture information on landscape temporal dynamics. This information is similar for the vegetation amount and seasonality measures, but the Fourier method also provides the unique measure of phase, which provides information on the timing of the maximum greenness.

Composite Seasonal and Annual Models: PCA-based composite summer, non-summer, and annual models are shown in Figure 18. Fourier-based composite summer, non-summer, and annual models are shown in Figure 19.

A visual comparison of the PCA and Fourier composite models reveals many similarities. Both PCA and Fourier methods delineate large patches of unfavorable habitat to the far southwest (south of the prominent bend in the Salt River) and to the east of OPCNM. Both PCA and Fourier delineate some large patches of highly favorable landscapes adjacent to and west of OPCNM as well as to the north of OPCNM. We also see the same general patterns observed with the previous models: the summer models tend to be more spatially restricted and shifted east of the non-summer models, and the PCA models tend to generalize the landscape more than the Fourier models. It is worthwhile to reiterate here that the discrimination offered by the Fourier model may or may not be more accurate or more useful than the PCA-based models, and would preferably be evaluated in the field by those using the models.

Composite Seasonal and Annual Models Validation Results: The composite seasonal and annual models were validated in two ways: first by using the 20 percent of actual pronghorn sightings excluded from model development and, second, by comparing the actual pronghorn sightings to a set of random points. The validation results using the set of withheld sightings are shown in Table 3. As with the validation of the yearly models, we see the expected pattern of points tabulated for the distance classes, suggesting that the models reasonably predict the locations of the pronghorn sightings. As with the yearly models, the composite models show that the PCA-based models tend to include more

validation points in the smaller standard deviation distance classes. As discussed in results of the yearly models, these validation results might suggest that PCA-derived models are preferable to Fourier models, but can be explained by the fact that the PCA-based models tend to generalize the landscape more than the Fourier-based models.

As an additional visual check of the effectiveness of the composite models, we overlaid the entire set of pronghorn sightings. Figure 20 shows the total pronghorn sightings database overlain on the annual PCA-based and Fourier-based composite models. Histograms in Figure 21 show the summer, non-summer and annual distributions of the sightings within the standard deviation distance classes for the PCA and Fourier models as paired histograms.

The visual fit between the geographic extent of the pronghorn sightings and the modeled regions of low SDD is good, indicating that the models are capturing salient landscape characteristics to which the pronghorn are responding. In fact, the Fourier models outline a fan-shaped area of low SDD (favorable) to the west of OPCNM that approximately mirrors the distribution of the pronghorn sightings, suggesting these models may be doing a better job than the PCA models. With respect to the histogram distributions, we see that the PCA models include more sightings in lower SDD classes, however the Fourier model has a tighter distribution.

6.0 DEFINING PREFERRED HABITAT: TOWARD AN INTEGRATED MODEL

Finally, we examine the characteristics of the models generated, to contribute toward an integrated model of seasonal habitat for the Sonoran pronghorn antelope. The results of the remote sensing-based modeling are further evaluated by comparing them to the GIS-based models developed using logistic regression in the Organ Pipe Cactus National Monument (OPCNM). This comparison provides another means to ascertain the appropriateness of the AVHRR-based models. We compared the GIS and AVHRR models in two ways: visually and statistically.

A visual comparison of the OPCNM GIS-based models with the final composite AVHRR-based models suggests that both models (PCA and Fourier) show a reasonable fit to the GIS models. Figures 22 and 23 show side-by-side comparisons of the PCA-based (summer, non-summer, and annual) and the Fourier-based (summer, non-summer, annual) composite models, respectively, with their GIS-based model analog. The PCA models appear to better define areas along the south edge of the monument that are identified in the GIS model, whereas the Fourier models appear to better define features along the northern and central portion of the Monument. In all cases, the extreme eastern part of the monument is considered not preferred, and the western edge is identified as preferred landscape. Summer models are more restricted and occupy slightly higher elevation landscapes than the winter models.

The most notable area of disagreement between the GIS-based and AVHRR-based models is in the north central part of the monument. GIS-based models indicate a distinct, target-shaped pattern that is not evident in the AVHRR-based models. The sightings that define this feature are dominated by an individual pronghorn. Since the GIS-based models were developed using only the 531 sightings within OPCNM, they can be dominated by the sightings from individual animals. In contrast, the AVHRR-based models are developed using the entire set of 2715 sightings, so the effect of movement of a single animal is not as apparent. Furthermore, the fact that the AVHRR models do not identify this local area to be preferred suggests this individual animal is accessing a landscape that is relatively unique compared to the other animals.

A statistical comparison of the models is possible because both models use a continuous scale to describe landscape favorability for pronghorn: probability in the GIS models and standard deviation distances in the AVHRR models. These results must be used with caution, however, since these measures were created using very different types of data at different scales,.

To perform a statistical comparison requires degrading the 30-meter spatial resolution of the GIS models to the I-kilometer spatial resolution of the AVHRR models. This was accomplished by calculating the average probability of the GIS model within each AVHRR pixel footprint. We then performed a linear regression of the GIS and AVHRR model values associated with each 1-kilometer pixel. Inspection of the initial results revealed a few outliers (up to 5 pixels out of 1338) that were removed for the final linear regression analyses. The graphs in Figure 24 show the regression results between GIS probability (y-axis) and the AVHRR standard deviation distance (x-axis).

Overall, the regression results show a weak to moderate correlation between the GIS probabilities and the AVHRR distances. The PCA-based models overall show a slightly stronger correlation to the GIS models compared to the Fourier-based models. In both AVHRR data sets, the non-summer models show the strongest correlation and the annual models show the weakest correlation. Although these correlations are moderate at best, given the disparate types of data utilized, the results indicate that the AVHRR-based models are reasonably effective at capturing landscape characteristics identified in the GIS-based models.

7.0 CONCLUSIONS

We created viable habitat models on the basis of the available environmental and pronghorn sightings data using logistic regression. The ultimate predictive ability of the resulting logistic regression models varied between 66 and 86 percent, as a function of model type (annual, summer, or non-summer) and cut-off value selected. We were able to demonstrate the improvement in detail that the implementation of more sophisticated

statistical techniques can bring to habitat modeling. The research was particularly distinctive in its incorporation of a seasonal component of the pronghorn locations, rather than a uniform model of territorial occupation, as well as its integration of measures of terrain variability that have been absent from other pronghorn studies.

This study has also demonstrated the viability of using high temporal resolution AVHRR data to characterize and map season-specific habitat of the Sonoran pronghorn antelope. The GIS-based models and the AVHRR-based models for the OPCNM show a reasonable correspondence. Based on visual comparisons and regression results at OPCNM, non-summer models exhibit the best fit, annual models the poorest fit, and, in general, the PCA-derived models exhibit a better fit than the Fourier-derived models. Since the GIS-based models directly incorporate detailed, field-based map information, these results for the local area of OPCNM provide assurance that the AVHRR models capture landscape characteristics that have biophysical meaning.

These findings are significant because the AVHRR models derived in this study can be easily constructed, provided that a sightings data set is available and can be matched with coincident AVHRR NDVI imagery (available for the US from 1989 to present). In contrast, the GIS-based models require detailed field-based data sets that are expensive and labor-intensive to develop and are not generally available. The AVHRR models in this study, constructed by simply averaging the pertinent temporal measures and model years, effectively capture salient landscape characteristics to which the pronghorn seem to respond. The integrity of these models is clearly demonstrated by the validation results, which utilized actual sightings withheld from model development.

With respect to the two methods used to extract temporal indices from the AVHRR data (PCA vs. Fourier), this study shows that PCA models tend to generalize the landscape relative to the Fourier models. This is not surprising, since the Fourier models are based on three temporal measures (additive term, magnitude, and phase) to describe the amount and seasonality of vegetation whereas the PCA models are based on two measures (PC1 and PC2). Whether the discrimination offered by the Fourier model is more accurate or more useful than the generalization offered by the PCA model can only be determined in the field by those using maps generated form these models. Although the regression results for OPCNM show stronger correlation between the GIS models and the PCA-derived models when compared to the Fourier-derived models, these results could reflect local conditions within OPCNM that do not generalize to the surrounding region. The visual fit between the various models and the entire sightings database suggests to us that Fourier models may ultimately be more informative.

With respect to all the models generated for this study, our results show that it is vital to consider seasonal differences when characterizing the habitat of this highly mobile species. In all cases, pronghorn avoid the eastern part of OPCNM and prefer landscapes along the western part of the monument. However, significant differences in the

landscapes accessed by the Sonoran pronghorn antelope are identified when considering each season separately. In addition, validation and regression results are much stronger when individual seasons are evaluated.

The observed geographical preferences found in the models fit with field-based observations of pronghorn habitat. The models define unfavorable terrain in the east of OPCNM and favorable terrain to the west. Field studies show that the eastern part of OPCNM contains steep, high elevation terrain, which is unsuited to pronghorn whereas the western part of OPCNM contains low-elevation floodplains known to host pronghorn (Hughes, 1991). The striking seasonal difference defined in the models also seems to fit with observations of pronghorn diet (Hughes, 1991). The models define an apparent shift of sightings from the low western floodplains in the non-summer months to the bajada slopes in the southwestern and north-central parts of OPCNM during the summer. Field studies show that during the relatively cooler non-summer months, the pronghorn are found preferentially in the floodplains to the west, where they access abundant forbs. During the harsher summer season, the pronghorn occupy the bajada slopes that host perennial chain-fruit cholla (*Opuntia fulgida*), a common and reliable food source during the dry season (Hughes, 1991; Hervert, 1996).

The AVHRR models developed in this study can be used to evaluate year-to-year comparisons of landscapes preferred by pronghorn. For example, January 1998 occurs in the middle of a very strong El Niño event. In Arizona, such winters tend to have more abundant precipitation. This climate condition appears to be reflected in the corresponding models for 1998 (Figure 11), which identify the most widespread area as favorable for pronghorn, and which include the westernmost landscapes of any of the models. The coincidence of the strong El Niño conditions and the widespread nature of the 1998 models could be the result of more abundant and widespread forage availability. Such relationships would need to be verified and field checked during the course of such an event. The ability to not only identify meaningful connections between the models and climate but also to make use of such knowledge would require field based knowledge of the study area as well as familiarity with the pertinent ecological processes. However, once established, such connections can be used to help manage wildlife and habitat. For example, it would be possible to predict the most likely places to find pronghorn during a given season in a year with particular climate conditions.

Year-to-year comparisons of the models were possible because the measures derived from the AVHRR data were converted to common units of standard deviation distance (SSD) based on the pronghorn distributions. In the case of PCA, which produces scene-specific and abstract measures, conversion to common units is necessary for between-year comparisons. In the case of Fourier measures, however, there is the option of extracting values from the AVHRR that have biophysical meaning and that can be compared directly between years. For example, the first frequency phase value, calculated as a number between 0 and $2*(\pi)$, represents the interval from January 1 to December 31 and

can be translated into a Julian date that represents the timing of maximum greenness for the year evaluated.

In addition, the second frequency term of the Fourier analysis can be used to evaluate the strength and timing of landscapes displaying a bimodal greenness pattern (Jakubauskus and Legates, 2001). Such a bimodal pattern is typical of the Arizona desert because there are two distinct seasons of high precipitation each year: the summer monsoon and the winter rains. For the current study, the higher frequency Fourier terms were not extracted, as the PCA analysis has no analogous measure. The information contained in the higher frequency Fourier terms, however, could prove fruitful for characterizing habitat in the desert southwest and should be explored.

The AVHRR models reveal an anomalous pattern for the year 1997, reflected especially in the phase value of the Fourier analysis. Because this year was so different when compared to the remaining four years, it was excluded from the composite model creation. However, the unusual pattern observed for 1997 could be defining important aspects of pronghorn habitat and should be explored. It is possible that unusual precipitation patterns account for the observed phase anomaly. Inspection of available climate data and evaluation of higher frequency Fourier terms extracted from the AVHRR data may shed light on the timing of greenness response to precipitation events.

Future research could also evaluate the application of more complex statistical methods (e.g., discriminant analysis, and neural nets) to create models from the original AVHRR temporal measures. In this study, we combined the normalized images showing preferred pronghorn standard deviation distances into models for each year by simply averaging the pertinent SDD measures. As stated previously, we felt that this simple averaging method facilitated year-to-year comparisons, provided a reasonable preliminary model to evaluate, and kept the results more transparent to the potential end-user audience. Although the resulting models are shown to be reasonable, statistical methods to optimize the weighting of the various measures might serve to refine the models.

Finally, to create a truly integrated model, future research could also seek to fine-tune the relationship between the GIS and AVHRR models. The resulting regression coefficients could then be used to extrapolate the relationship calculated within OPCNM to the greater area (including the Barry Goldwater Range and the Cabeza Prieta). Although the moderate correlation identified in this study is encouraging, the current results do not warrant extrapolation with these data. Nevertheless, the potential does exist to use the GIS model to train the AVHRR data. Such an effort assumes that the GIS-based model is more accurate than the AVHRR-based model and would require additional ancillary information, such as regional vegetation cover, to reasonably confine any extrapolation procedure.

ACKNOWLEDGEMENTS

The U.S. Department of the Interior, Organ Pipe Cactus National Monument and the U.S. Department of Defense, U.S. Airforce directly sponsored this research. In addition, we would like to thank the Arizona Game and Fish Department (AGFD) and OPCNM for facilitating access to their data and many helpful suggestions. We would also like to thank Mr. Thomas Potter (OPCNM) and Mr. James DeVos (AGFD) for their invaluable assistance and Mr. Jay Miller for his early contributions to this work.

REFERENCES

Austin, G.E.; Thomas, C.J.; Houston, D.C.; Thompson, D.B.A., 1996. Predicting the spatial distribution of buzzard Buteo buteo nesting areas using a geographical information system and remote sensing. *Journal of Applied Ecology*, v33/6, 1541-1550.

Bian, Ling; West, E., 1997. GIS modeling of elk calving habitat in a prairie environment with statistics. *Photogrammetric Engineering and Remote Sensing*, v63/2, 161-167.

EDC, 1994. *Conterminous United States AVHRR Data on CD ROM - User's Manual*, U.S. Department of Interior, Eros Data Center. Sioux Falls, SD.

Garson, G.D. , Professor of Public Administraiton, North Carolina State University. "Logistic Regression." Web page, [accessed January 2000]. Available at: http://www2.chass.ncsu.edu/garson/pa765/garson.htm.

Herr, A.M. and L.P. Queen (1997) Crane Habitat Evaluation Using GIS and Remote Sensing. *Photogrammetric Engineering and Remote Sensing*. 59(10):1531-1538.

Hervert, J. J., R. S. Henry, M. T. Brown, and L. A. Piest, 1996. *Sonoran pronghorn population monitoring: progress report*, Nongame and Endangered Wildlife Program Technical Report 110. Arizona Game and Fish Department, Phoenix, Arizona.

Herzfeld, U.C. and C.A. Higginson, 1996. Automated Geostatistical Seafloor Classification - Principles, Parameters, Feature Vectors, and Discrimination Criteria: *Computers & Geosciences*, 22:35-52.

Hirosawa, Y., S.E. Marsh and D.H. Kliman, 1996. Application of Standardized Principal Component Analysis to Land-Cover Characterization Using Multitemporal AVHRR Data. *Remote Sensing of Environment*. 58:267-281.

Hodgson, M. E., J. R. Jensen, H. E. Machey, and M. C. Coulter, 1988. Monitoring Wood Stork Foraging Habitat Using Remote Sensing and Geographic Information Systems. *Photogrammetric Engineering and Remote Sensing* 54, no. 11:1601-1607.

Hooge, P.N. and B. Eichenlaub (1997) *Animal movement extension to arcview. ver. 1.1*. Alaska Science Center - Biological Science Office, U.S. Geological Survey, Anchorage, AK, USA.

Huete, A.R., and Jackson, R.D., 1987. Suitability of Spectral Indices for Evaluating Vegetation Characteristics on Arid Rangelands. *Remote Sensing of Environment*. 23:213-232.

Hughes, K.S., 1991. *Sonoran Pronghorn Use of Habitat in Southwest Arizona.* The University of Arizona, M.S. Thesis. 86 p.

Isaaks, E.H., and R.M. Srivastava, 1989. *An Introduction to Applied Geostatistics;* Oxford University Press, New York.

Jakubauskas, M.E. and Legates, D.R., 2001. Harmonic Analysis of Time-Series AVHRR NDVI Data for Characterizing US Great Plains Land Use/Land Cover. *Proceedings of the XIXth International Society for Photogrammetry and Remote Sensing Congress,* Amsterdam, The Netherlands, July 18-23, 2000.

Jakubauskas, M.E., D.R. Legates, and J. Kastens, 2001. Harmonic Analysis of Time-Series AVHRR NDVI data. *Photogrammetric Engineering and Remote Sensing* 67(4):461-470.

Kvamme, K. L., 1985. Determining Empirical Relationships between the natural environment and prehistoric site locations: A Hunter-Gatherer Example. in *For Concordance in Archaeological Analysis.* Ed. C. Carr, 208-39. University of Arkansas, Fayetteville.

Lambin, E.F. and A.H. Strahler, 1994. Change-Vector Analysis in Multitemporal Space: A Tool To Detect and Categorize Land-Cover Change Processes Using High Temporal Resolution Satellite Data. *Remote Sensing of Environment.* 48:231-244.

Leibhold et al., 1993. Geostatistics and Geographic Information Systems in Applied Insect Ecology. Annual Review of Entomology, 38:303-327.

Lomax, R. G., 1998. *Statistical Concepts: a second course for education and the behavioral sciences.* New Jersey: Lawrence Erlbaum Associates.

Marsh, S.E., C.S.A. Wallace, and J.J. Walker, 1999. *Evaluation of Satellite Remote Sensing Methods for Regional Inventory,* Final Report for Organ Pipe Cactus National Monument. Tucson: University of Arizona.

Moody, Aaron and David M. Johnson, 2001. Land-Surface Phenologies from AVHRR Using the Discrete Fourier Transform, *Remote Sensing of Environment,* 75(3):305-323.

Olsson, L. and L. Eklundh, 1994. Fourier Series for Analysis of Temporal Sequences of Satellite Sensor Imagery. *International Journal of Remote Sensing,* 15(18):3735-3741.

Pereira, J. M. C., and R. M. Itami, 1991. GIS-Based habitat modeling using logistic multiple regression: a study of the Mt. Graham red squirrel. *Photogrammetric*

Engineering and Remote Sensing, 57: 11, p. 1475-1486

Phinn, S, J. Franklin, A. Hope, D. Stowe and L. Huenneke, 1996. Biomass Distribution Mapping Using Airborne Digital Video Imagery and Spatial Statistics in a Semi-Arid Environment. *Journal of Environmental Management*, 47:139-164.

Reed, B.C. and Sayler, K., 1997. A Method for Deriving Phenological Metrics from Satellite Date, Colorado 1991-1995. In: *Impact of Climate Change and LandUse in the Southwestern United States, an electronic workshop.* (http://geochange.er.usgs.gov/sw/impacts/biology/Phenological?CO/).

Singer, F.J. and J.E Norland, 1994. Niche Relationships within a guild of Ungulate Species in Yellowstone National Park, Wyoming, Following Release from Artificial Controls. *Canadian Journal of Zoology-Revue Canadienne Zoologie* 72(8):1383-1394.

Van Manen, F.T. and Pelton, M.R., 1997. A GIS model to predict black bear habitat use. *Journal of Forestry*, vol95/8, pp. 338-345.

Verner, J., M. L. Morrison, and C. J. Ralph, 1986. Introduction. *Wildlife 2000: Modeling Habitat Relationships of Terrestrial Vertebrates* J. Verner, M. L. Morrison, and C. J. Ralph, xi – xv, Wisconsin: The University of Wisconsin Press.

Walker, J.J., 2000. *A GIS Model of Sonoran Pronghorn Antelope Habitat in Organ Pipe Cactus National Monument, Arizona.* Master's Thesis, Department of Geography and Regional Development, The University of Arizona, Tucson, 121p.

Wallace, C.S.A., J. M. Watts, and S. R. Yool, 2000. Characterizing the Landscape Structure of Vegetation Communities in the Mojave Desert Using Geostatistical Techniques. *Computers and Geosciences*. 26(4):397-410.

Wallace, C.S.A., 1997. *Geostatistical Evaluation of Ground-based Vegetation Data in a Portion of the Mojave Desert, California.* Master's Thesis, Department of Geography and Regional Development, The University of Arizona, Tucson, 136p.

	Threshold	Correct Points/100	% Correct
Annual	0.45	77	77%
	0.50	70	70%
Non-Summer	0.45	43	86%
	0.50	38	76%
Summer	0.45	37	74%
	0.50	33	66%

Table 1. Logistic regression model validation results utilizing the withheld 20% of sightings.

		Standard Deviation Distances									
1995		0 - 0.5	0.5 - 1.0	1.0 -1.5	1.5 - 2.0	2.0 - 2.5	2.5 - 3.0	More	%<1sd	%<2sd	%>2.0 sd
annual	fourier	7	48	28	13	4	0	0	55	96	4
summer	fourier	10	30	36	22	2	0	0	40	98	2
winter	fourier	12	43	35	6	4	1	0	55	95	5
annual	pca	21	47	22	9	0	0	0	68	100	0
summer	pca	26	34	32	8	0	0	0	60	100	0
winter	pca	26	43	19	12	0	0	0	69	100	0

1996		0 - 0.5	0.5 - 1.0	1.0 -1.5	1.5 - 2.0	2.0 - 2.5	2.5 - 3.0	More	%<1sd	%<2sd	%>2.0 sd
annual	fourier	13	38	42	7	1	0	0	50	99	1
summer	fourier	21	38	33	8	0	0	0	58	100	0
winter	fourier	13	38	36	14	0	0	0	50	100	0
annual	pca	23	44	25	8	0	0	0	67	100	0
summer	pca	25	33	21	21	0	0	0	58	100	0
winter	pca	19	53	19	8	2	0	0	72	98	2

1997		0 - 0.5	0.5 - 1.0	1.0 -1.5	1.5 - 2.0	2.0 - 2.5	2.5 - 3.0	More	%<1sd	%<2sd	%>2.0 sd
annual	fourier	23	27	4	3	1	1	41	49	56	44
summer	fourier	17	29	8	0	4	4	38	46	54	46
winter	fourier	21	28	4	2	2	0	43	49	55	45
annual	pca	20	32	17	20	7	4	0	52	89	11
summer	pca	12	48	24	16	0	0	0	60	100	0
winter	pca	15	33	15	13	13	4	7	48	76	24

1998		0 - 0.5	0.5 - 1.0	1.0 -1.5	1.5 - 2.0	2.0 - 2.5	2.5 - 3.0	More	%<1sd	%<2sd	%>2.0 sd
annual	fourier	13	52	26	4	2	4	0	65	94	6
summer	fourier	20	34	41	0	2	2	0	55	95	5
winter	fourier	12	53	24	4	2	4	0	65	93	7
annual	pca	42	36	14	2	2	2	2	79	95	5
summer	pca	32	39	23	5	2	0	0	70	98	2
winter	pca	45	39	9	1	1	2	2	84	94	6

1999		0 - 0.5	0.5 - 1.0	1.0 -1.5	1.5 - 2.0	2.0 - 2.5	2.5 - 3.0	More	%<1sd	%<2sd	%>2.0 sd
annual	fourier	17	47	26	9	1	0	0	64	99	1
summer	fourier	10	47	33	10	0	0	0	57	100	0
winter	fourier	32	38	15	13	2	0	0	70	98	2
annual	pca	32	43	18	5	0	1	0	75	99	1
summer	pca	37	37	17	10	0	0	0	73	100	0
winter	pca	37	33	24	4	0	2	0	70	98	2

Table 2. Tabulation of the model values extracted at the validation point locations. Models that correctly predict the validation points should produce an approximate normal distribution, with 68% of points < 1 standard deviation and 95% < 2 standard deviations.

MODEL		Standard Deviation Distances									
		0 - 0.5	0.5 - 1.0	1.0 -1.5	1.5 - 2.0	2.0 - 2.5	2.5 - 3.0	More	%<1sd	%<2sd	%>2sd
annual	fourier	0	61	30	8	1	0	0	61	99	1
summer	fourier	1	41	45	12	0	0	0	43	100	0
winter	fourier	0	63	28	7	1	0	0	63	99	1
annual	pca	16	58	16	5	5	0	0	74	95	5
summer	pca	12	54	23	10	1	0	0	66	99	1
winter	pca	17	58	15	4	5	1	0	75	94	6

Table 3. Validation results for actual pronghorn sightings. These are the 20% of the original sightings database that were withheld for validation purposes.

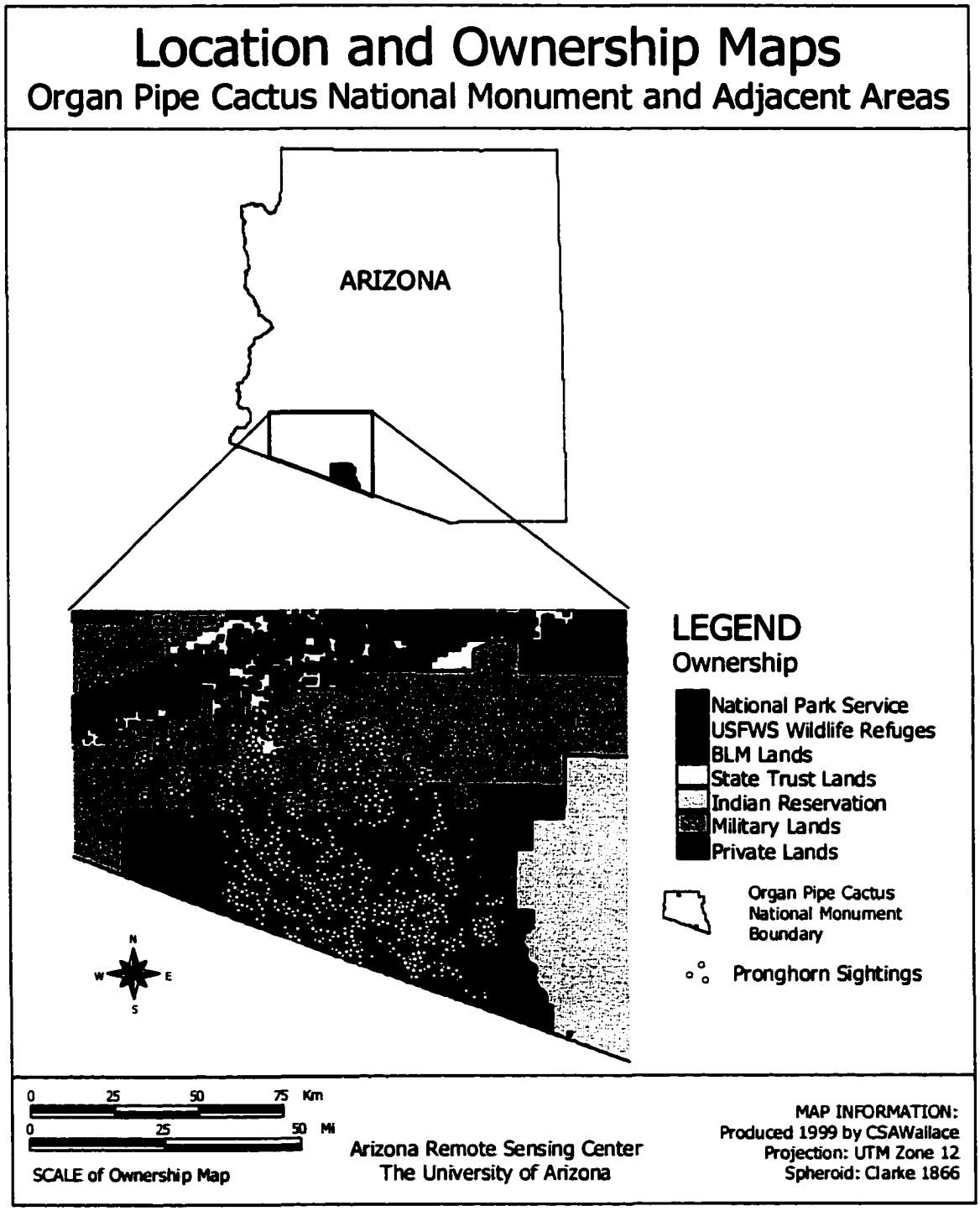


Figure 1: Location and land tenure map of the study area.

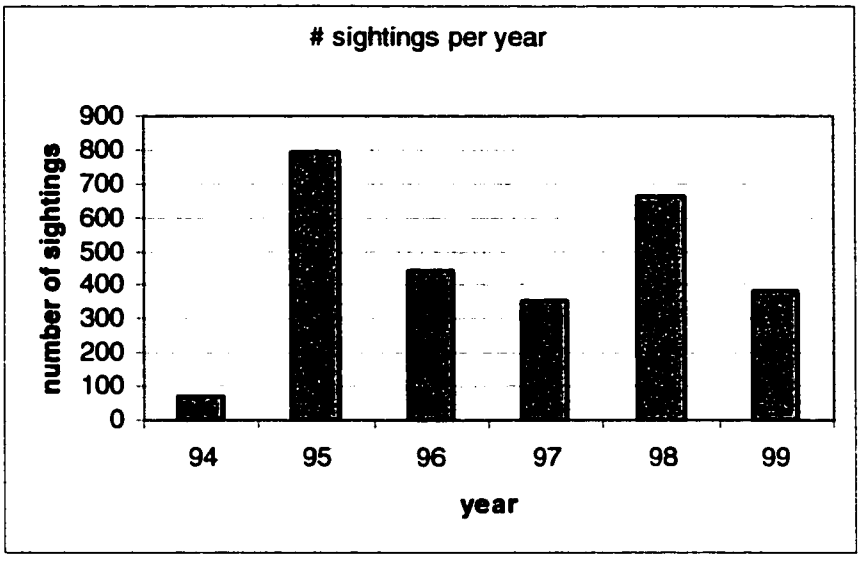


Figure 2: Histogram showing number of Sonoran pronghorn antelope sightings collected for each year.

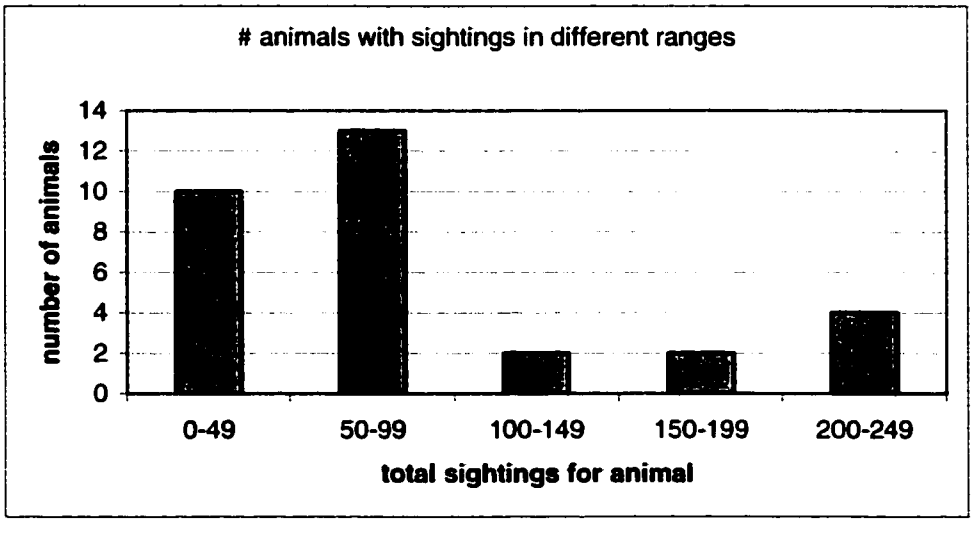


Figure 3: Histogram showing number of individual Sonoran pronghorn antelope per sightings category. Sightings categories stratify the total number of sightings into five equal classes, from 0 to 249, as shown.

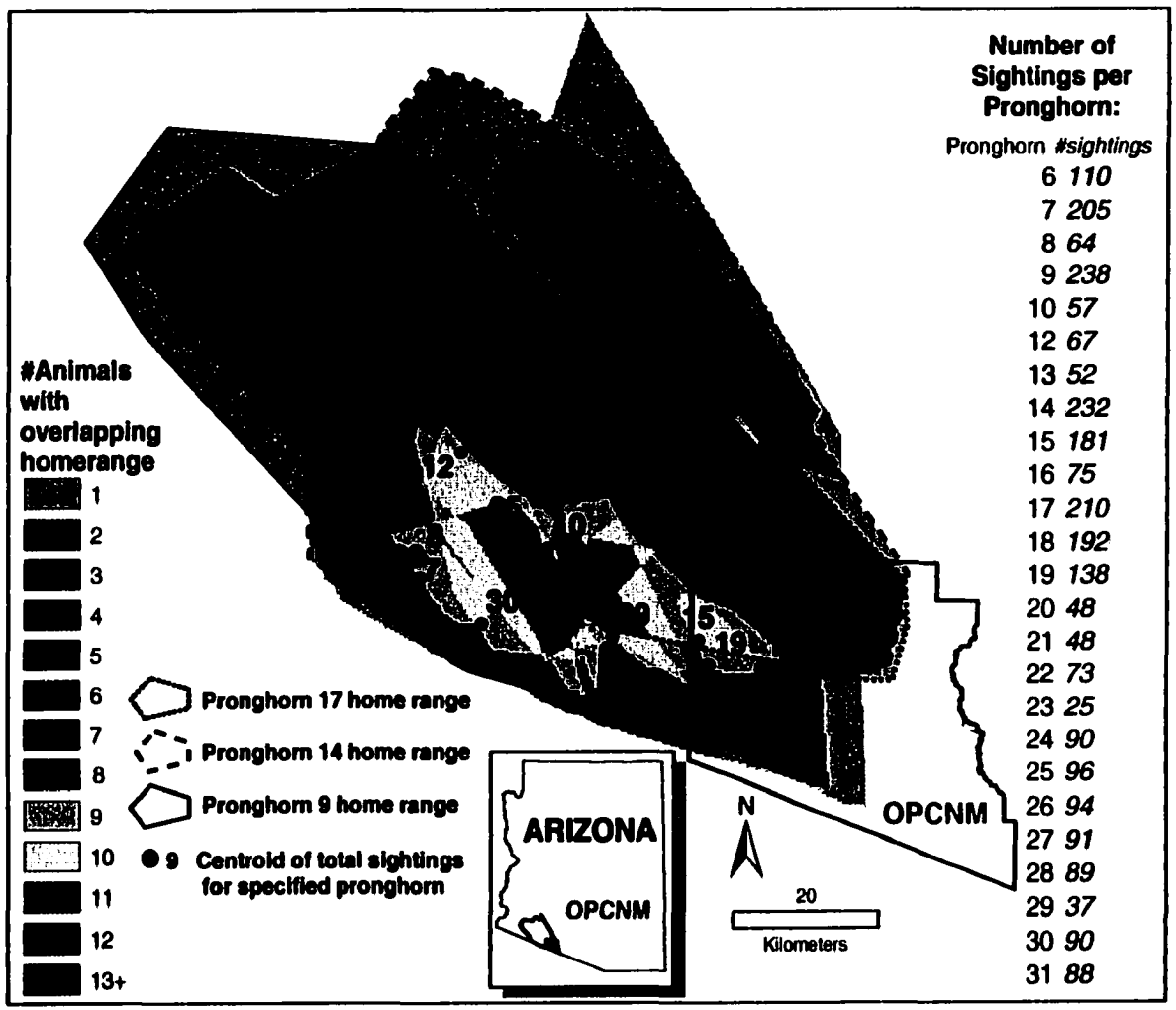


Figure 4: Map showing overlapping pronghorn antelope home range extents. Home ranges are created as the maximum convex polygon containing the sightings, with examples outlined for Pronghorns 12, 17, and 19. The centroids of the sightings for each individual pronghorn are shown as dots labeled with the animal ID number. Deceased pronghorn with less than 10 total sightings were excluded.

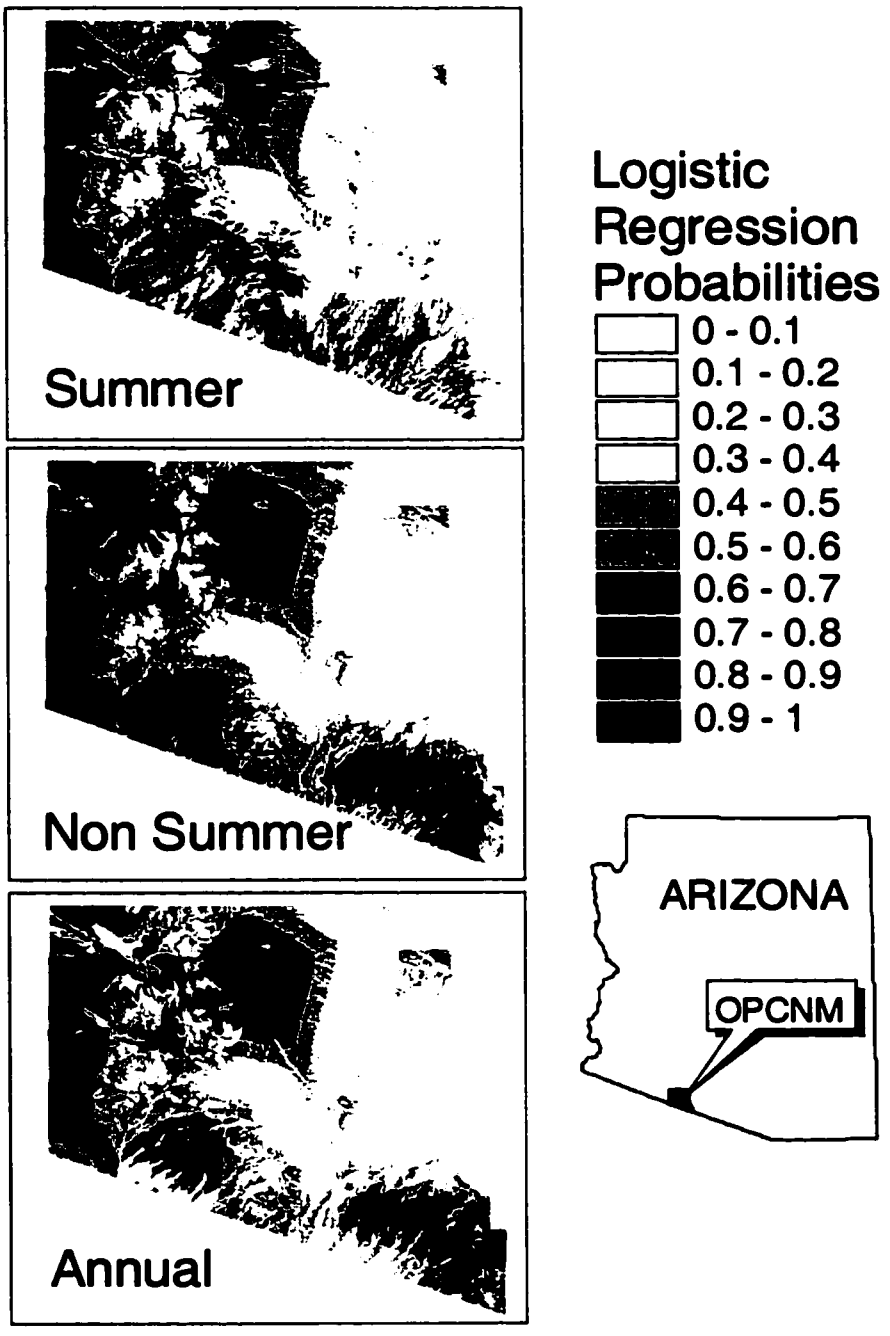


Figure 5: Summer, non-summer and logistic regression model results.

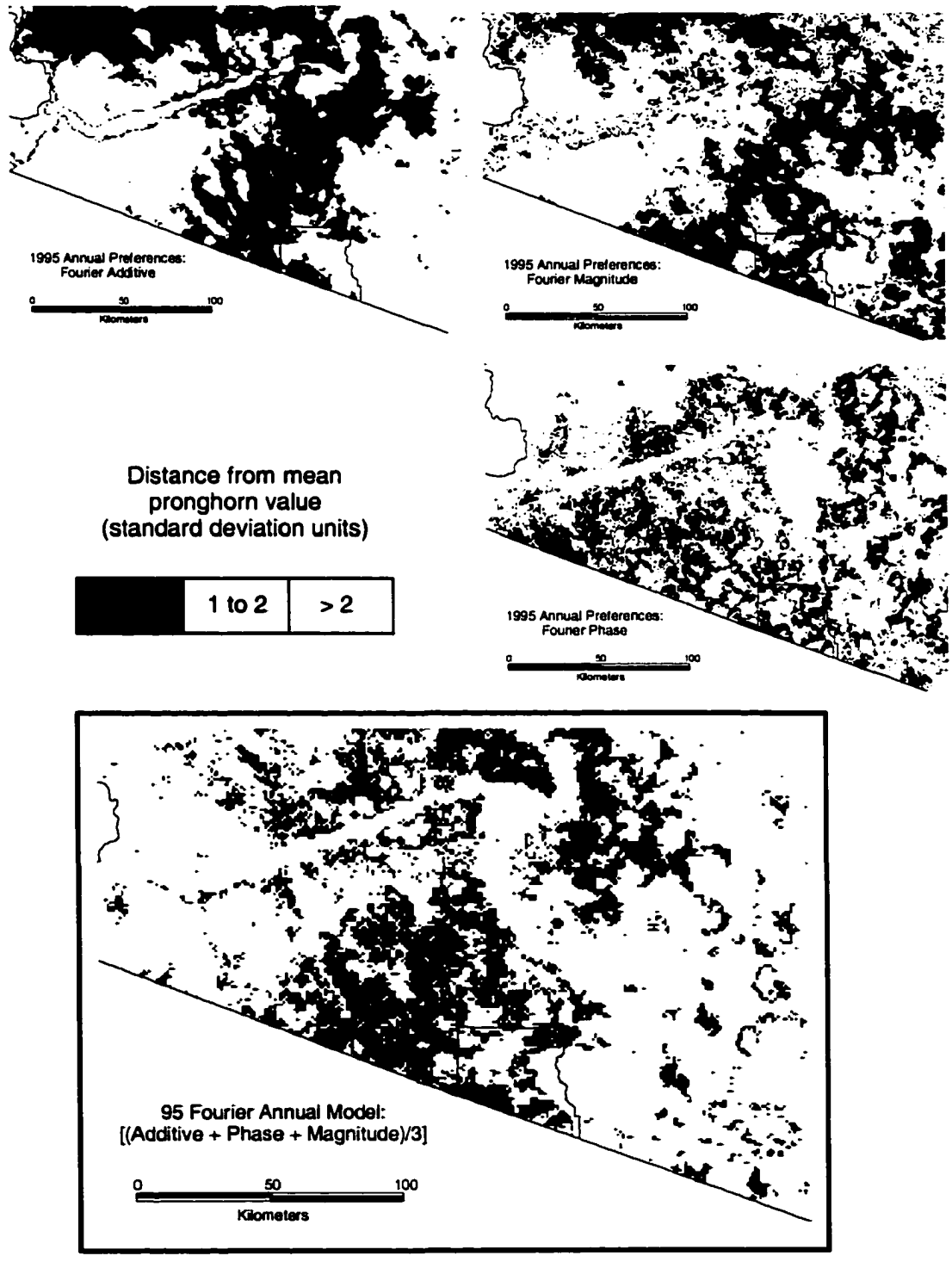


Figure 7: Preferred annual Fourier measures 1995 (top); Fourier Annual Model, 1995 (bottom)

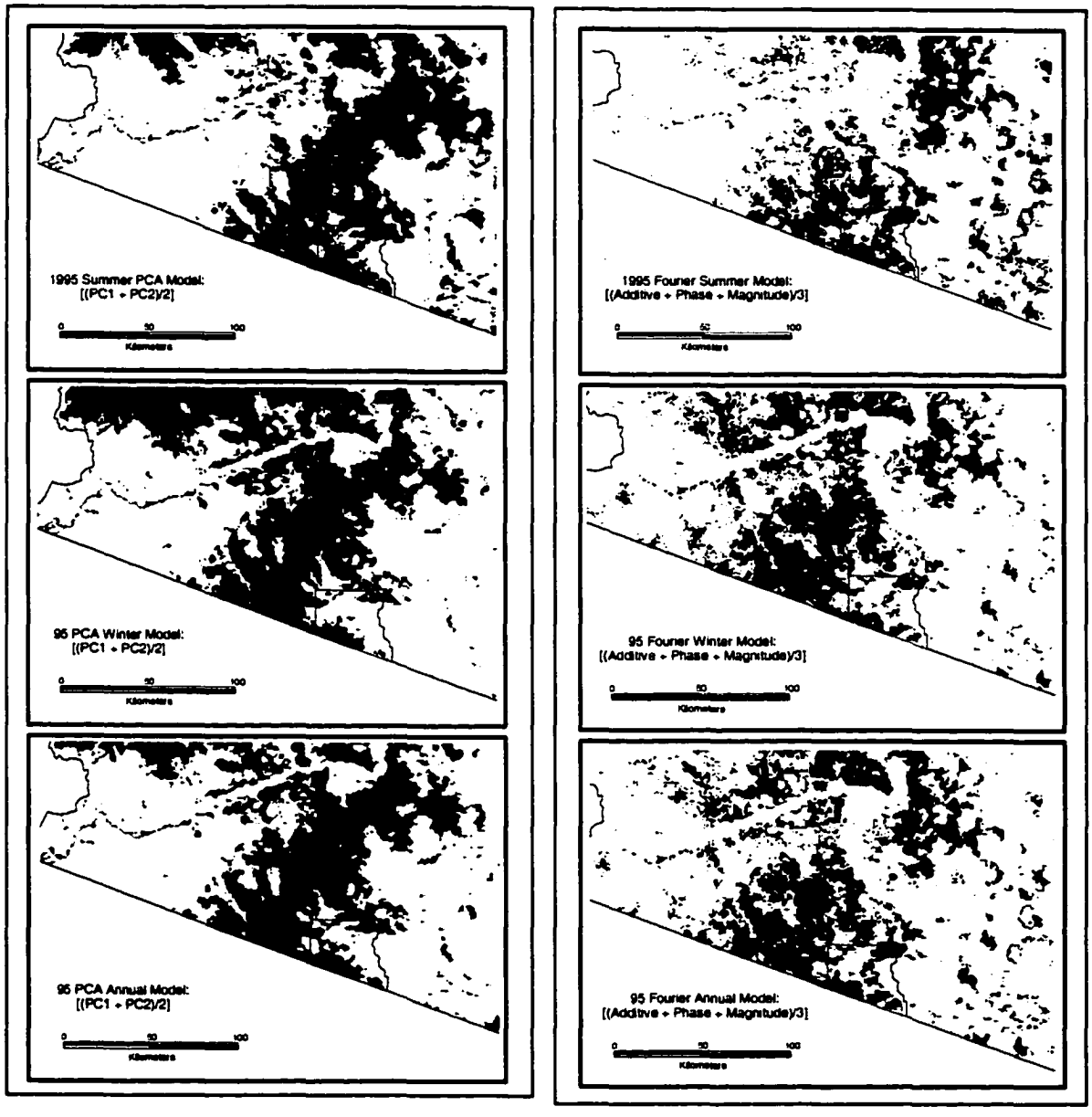


Figure 8: Models for 1995. Left column, from top to bottom, shows the summer, nonsummer, and annual models derived using PCA-based temporal indices. Right column shows the Fourier-based models.

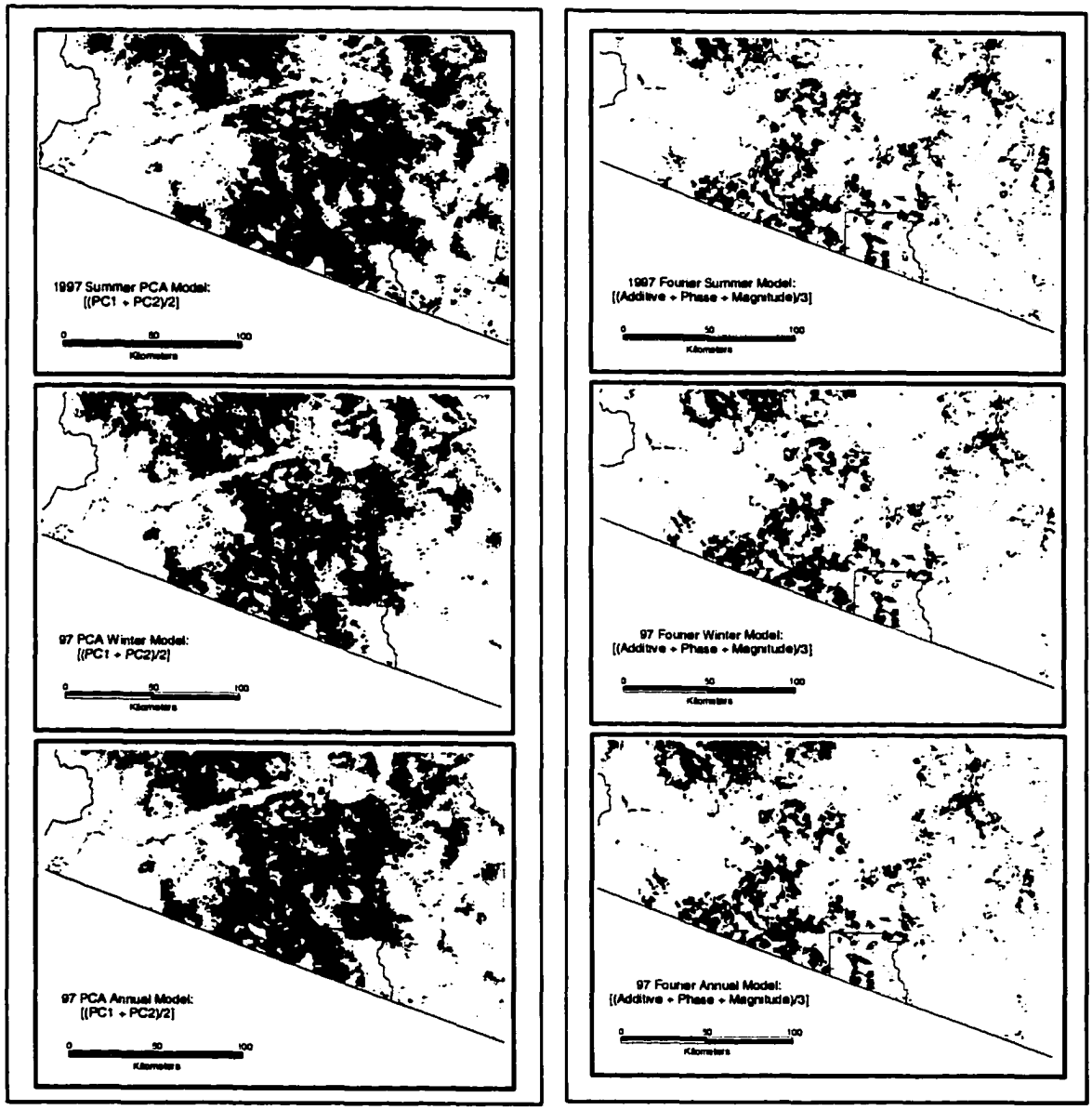


Figure 10: Models for 1997. Left column, from top to bottom, shows the summer, nonsummer, and annual models derived using PCA-based temporal indices. Right column shows the Fourier-based models.

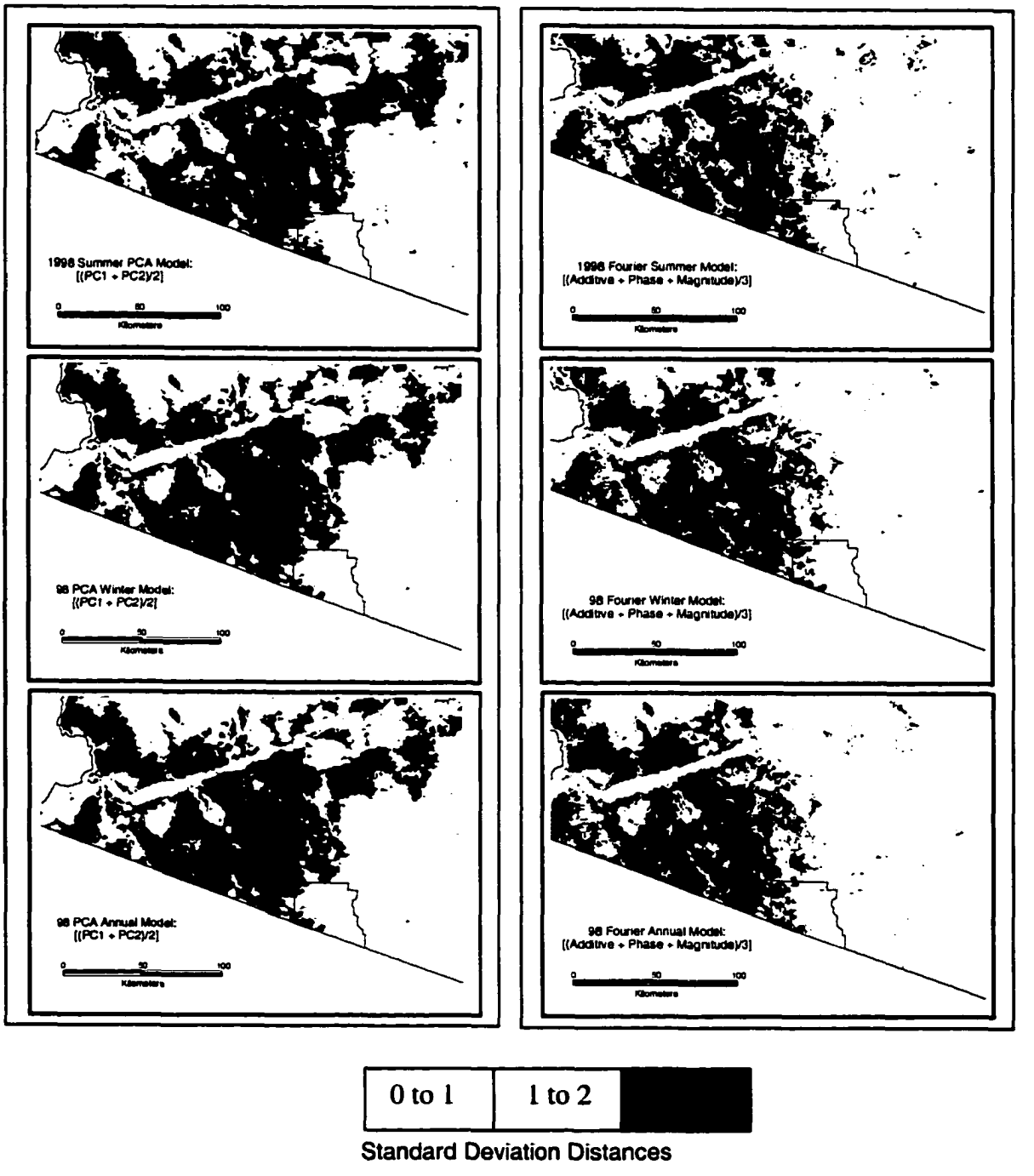


Figure 11: Models for 1999. Left column, from top to bottom, shows the summer, nonsummer, and annual models derived using PCA-based temporal indices. Right column shows the Fourier-based models.

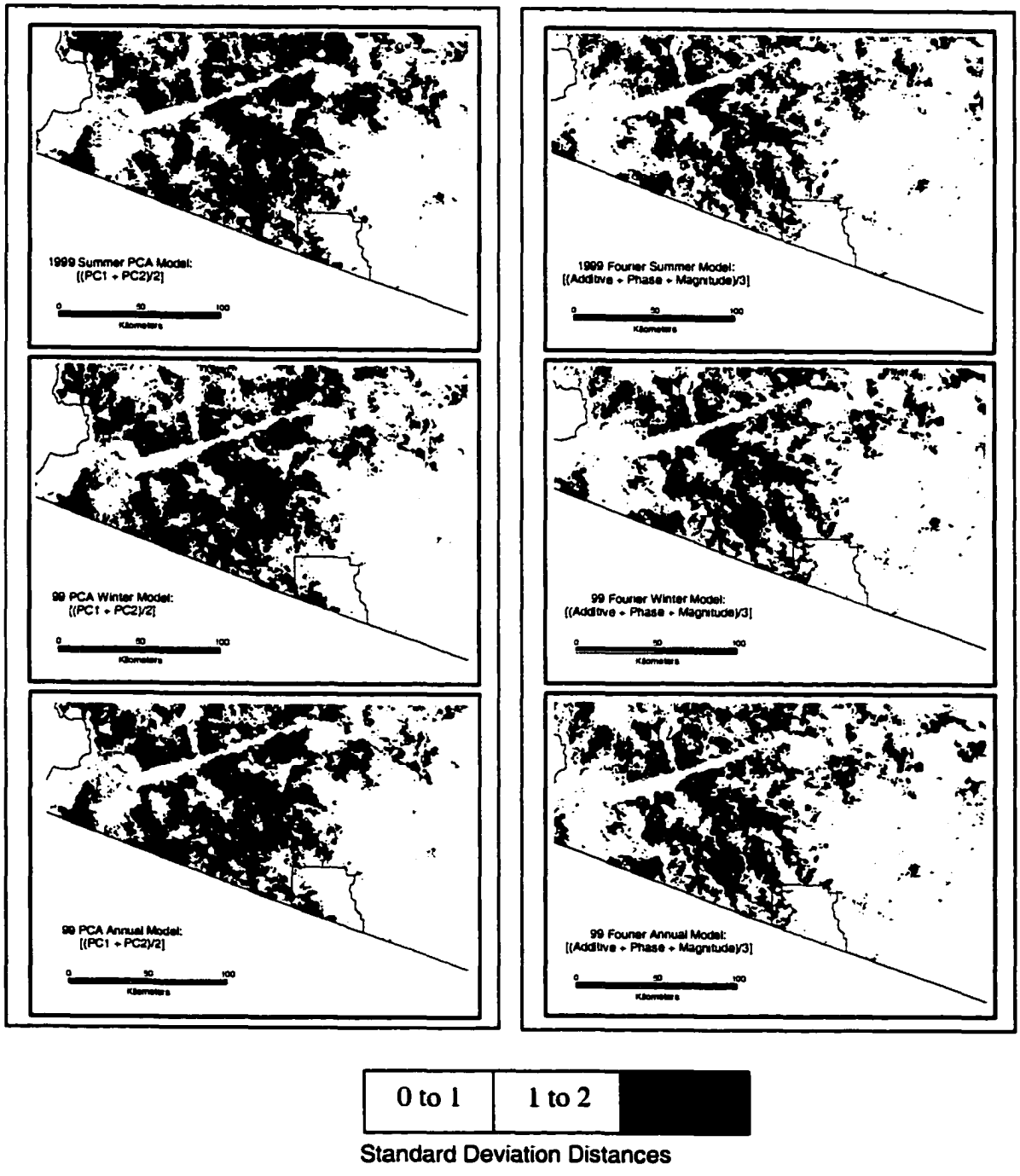


Figure 12: Models for 1999. Left column, from top to bottom, shows the summer, nonsummer, and annual models derived using PCA-based temporal indices. Right column shows the Fourier-based models.

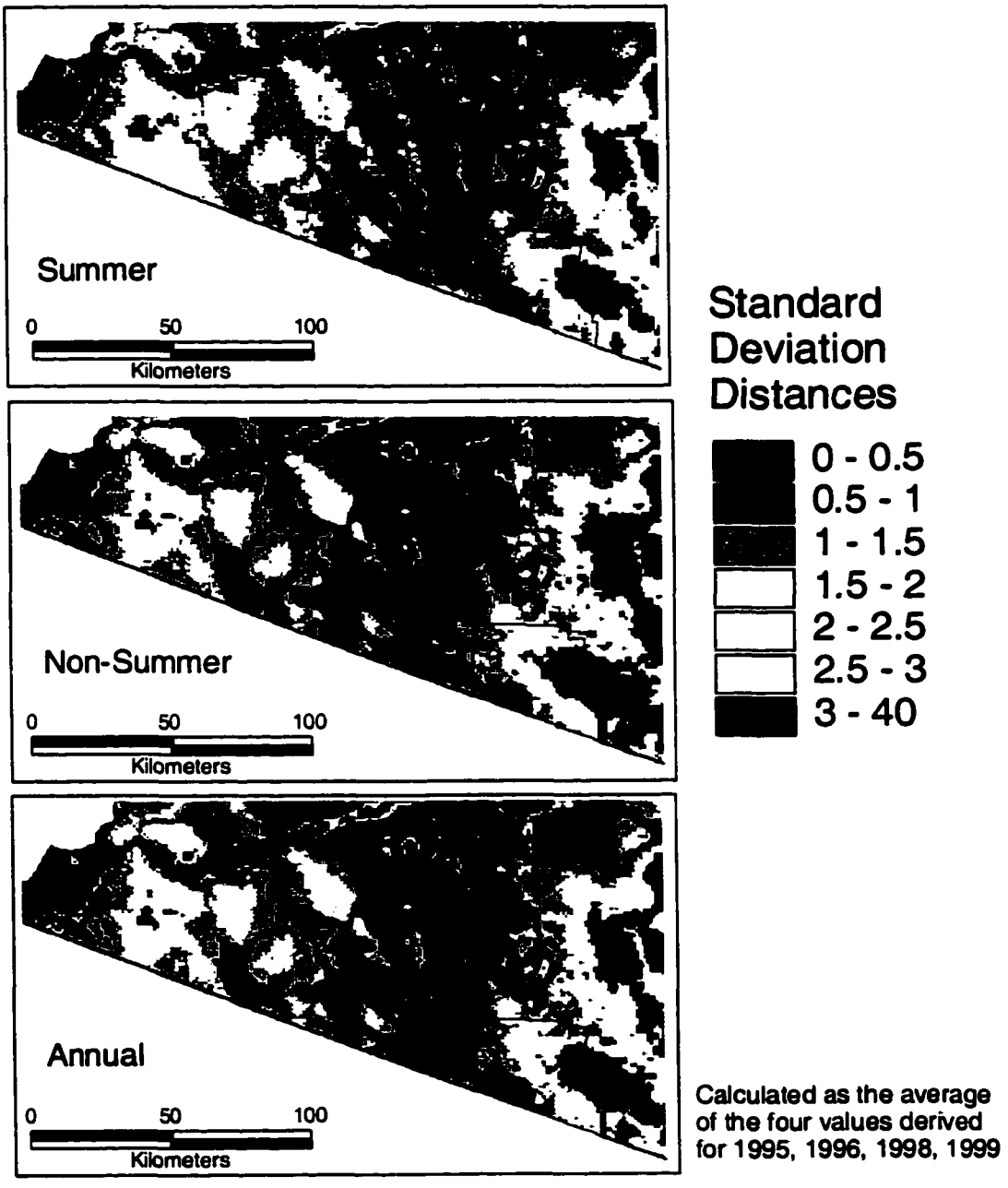


Figure 13: Composite PCA-derived PC1 measures preferred by pronghorn, 1995, 1996, 1998 and 1999. PC1 captures vegetation amount.

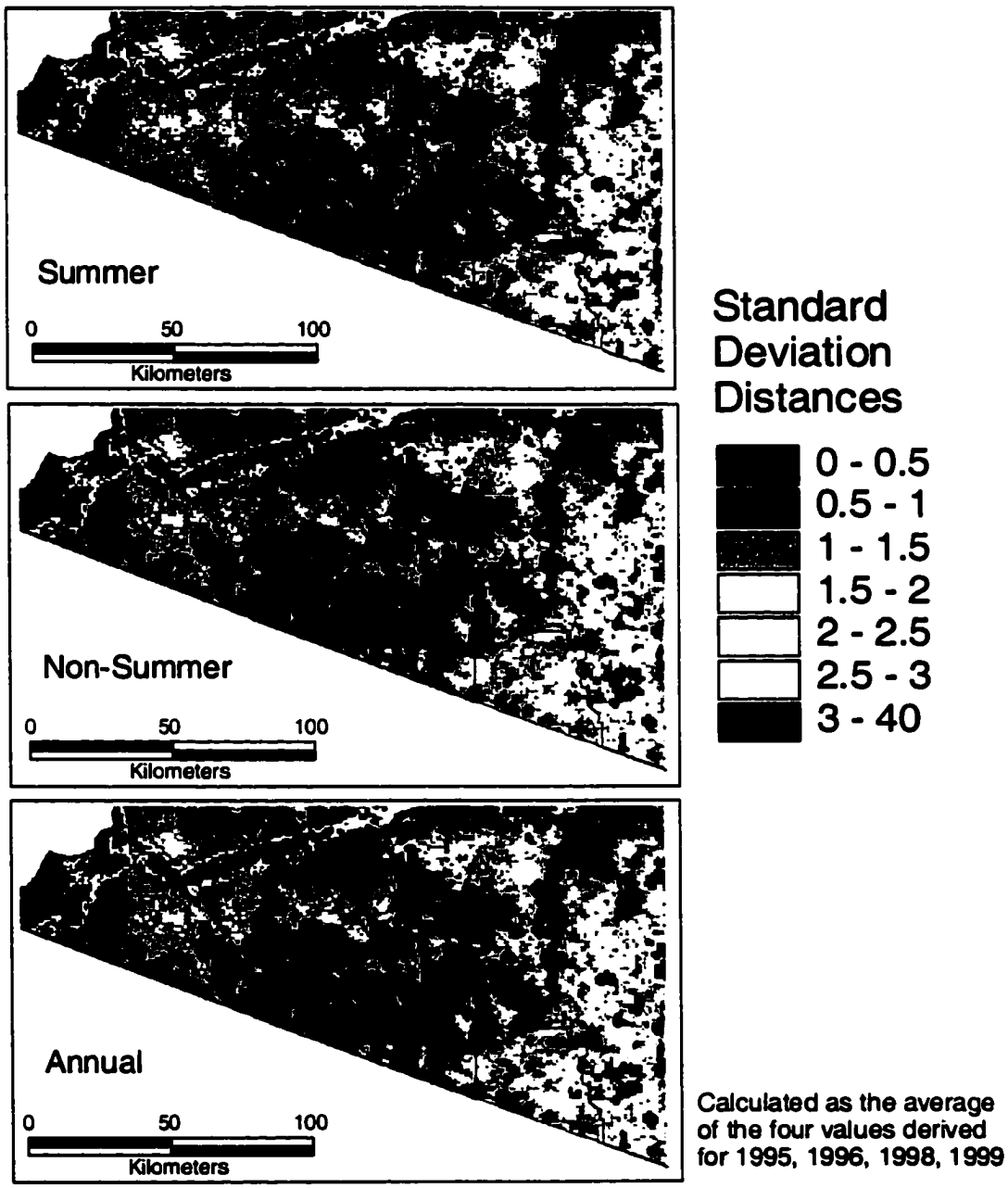


Figure 14: Composite PCA-derived PC2 measures preferred by pronghorn, 1995, 1996, 1998 and 1999. PC2 captures vegetation seasonality.

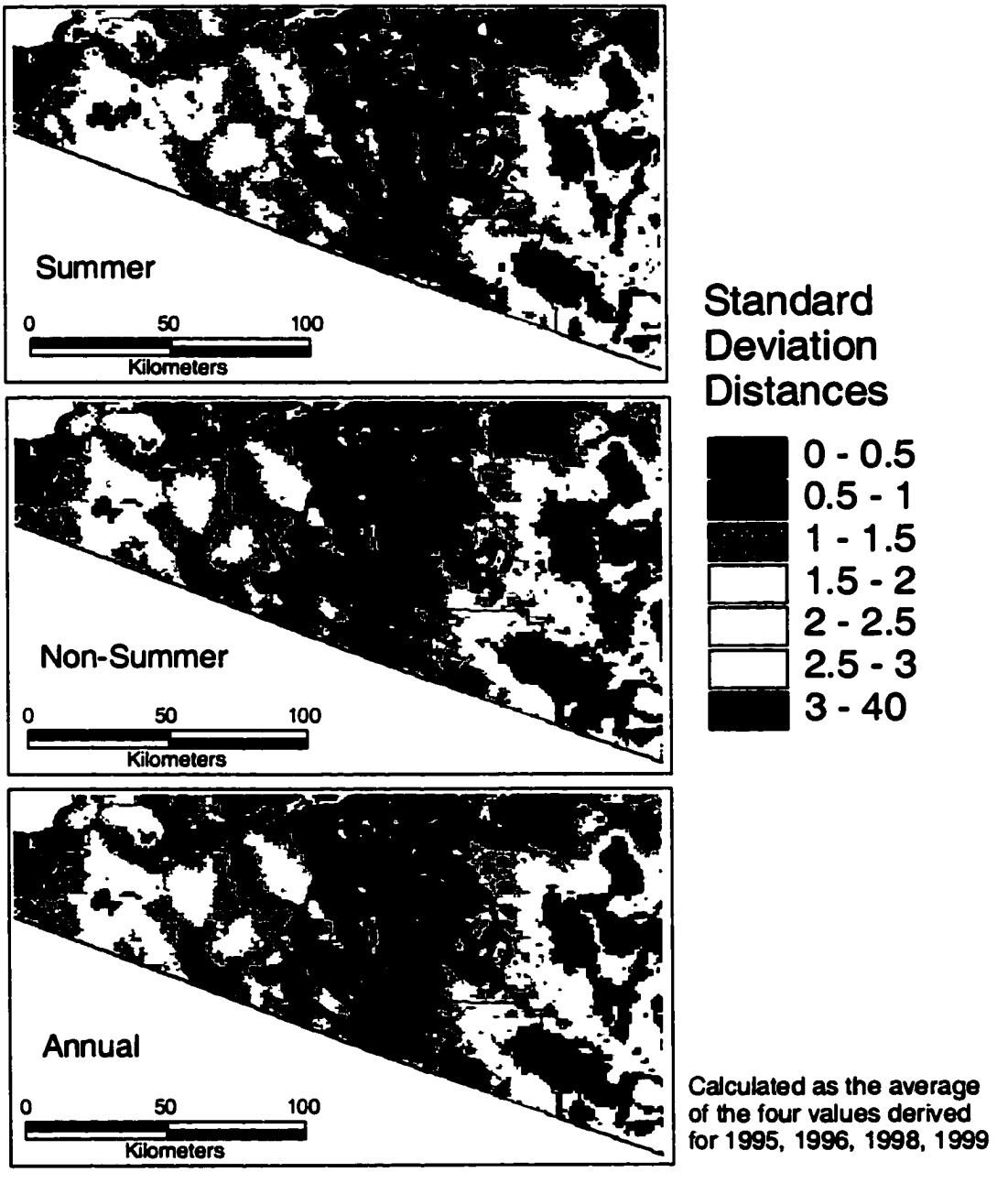


Figure 15: Composite Fourier-derived additive term measures preferred by pronghorn, 1995, 1996, 1998 and 1999. The additive term captures vegetation amount.

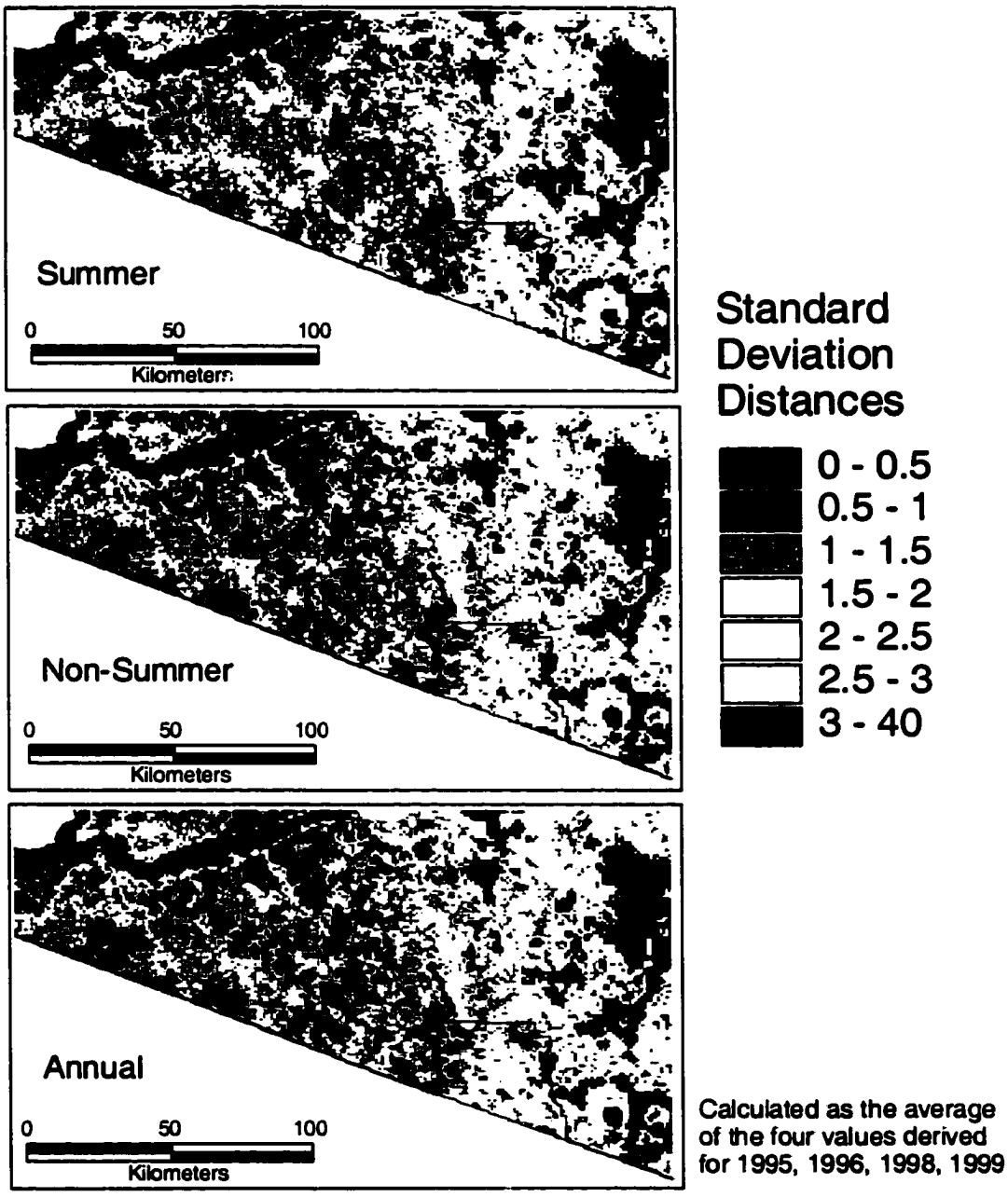


Figure 16: Composite Fourier-derived first frequency phase measures preferred by pronghorn, 1995, 1996, 1998 and 1999. The phase captures vegetation seasonality, in this case the timing of maximum greenness.

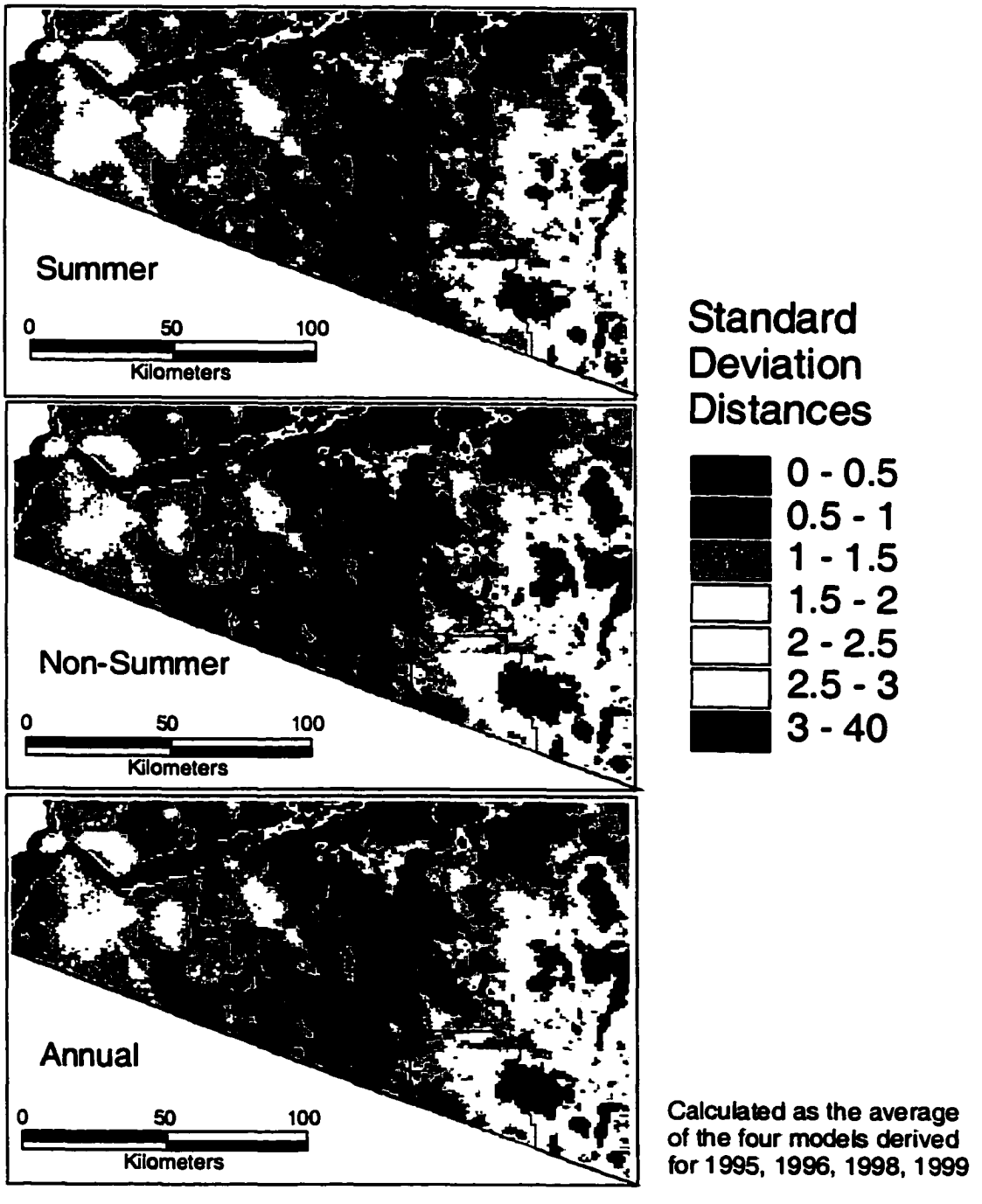


Figure 18: Composite PCA Models.

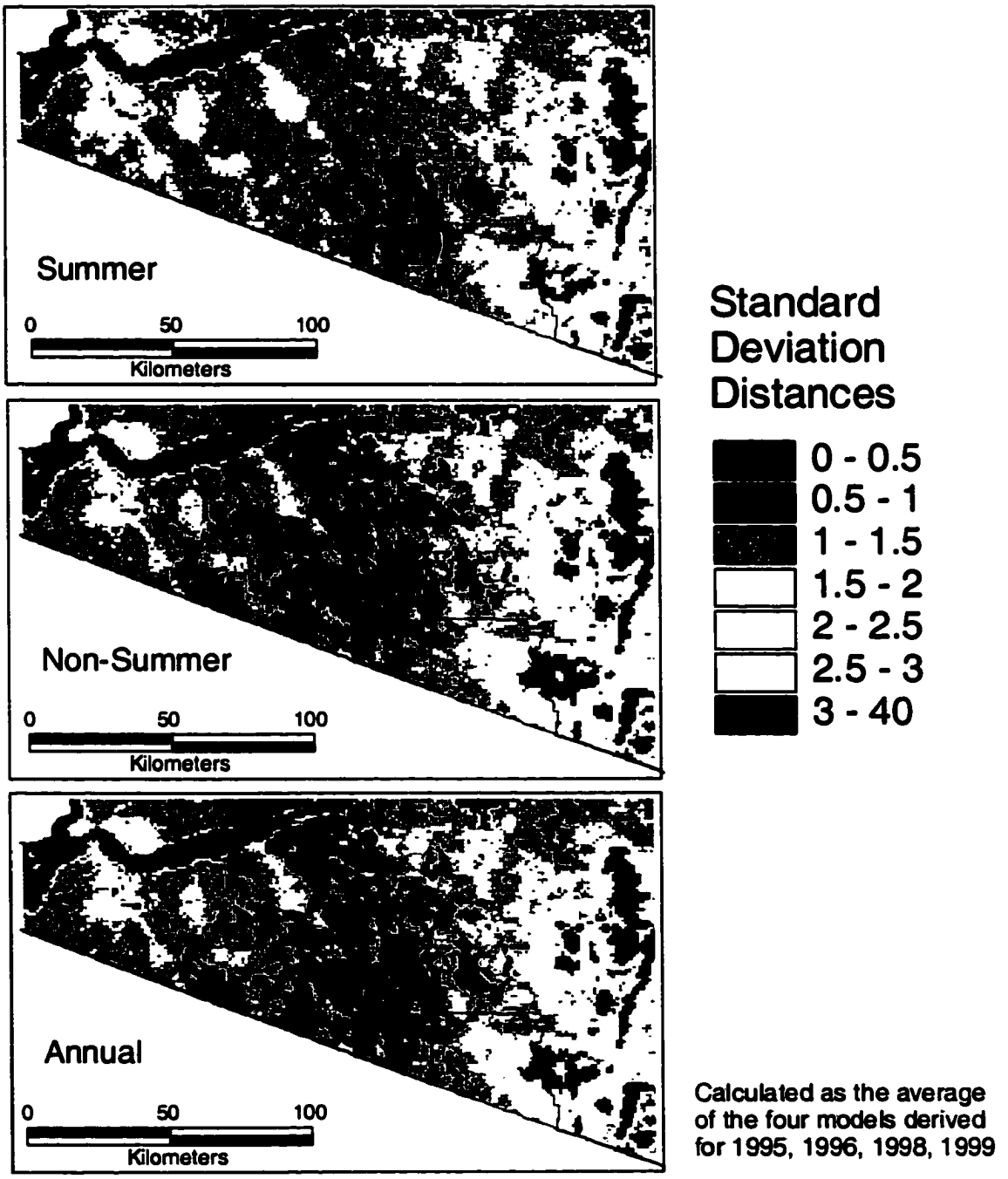


Figure 19: Composite Fourier Models.

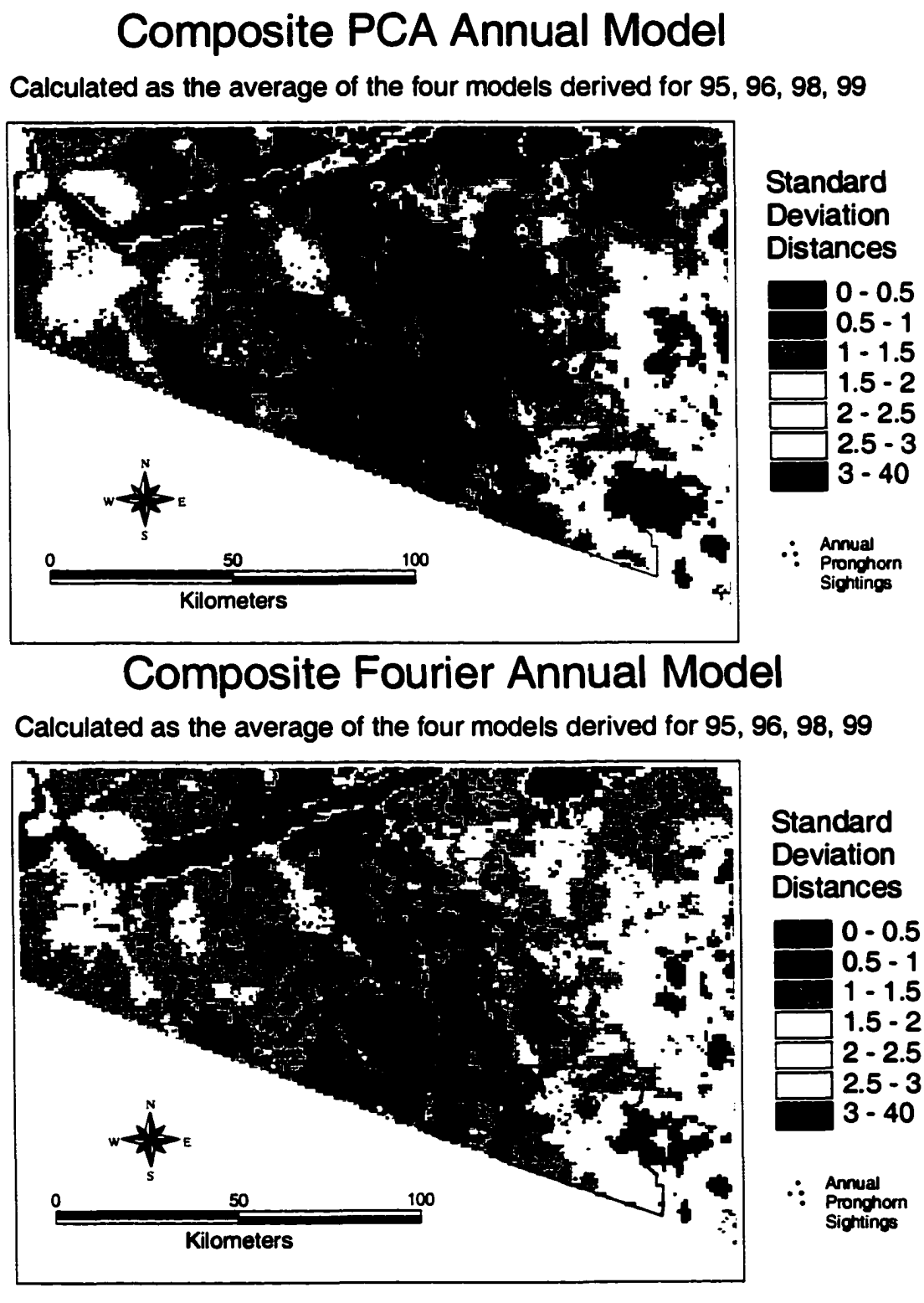


Figure 20: Pronghorn sightings overlain on the annual models.

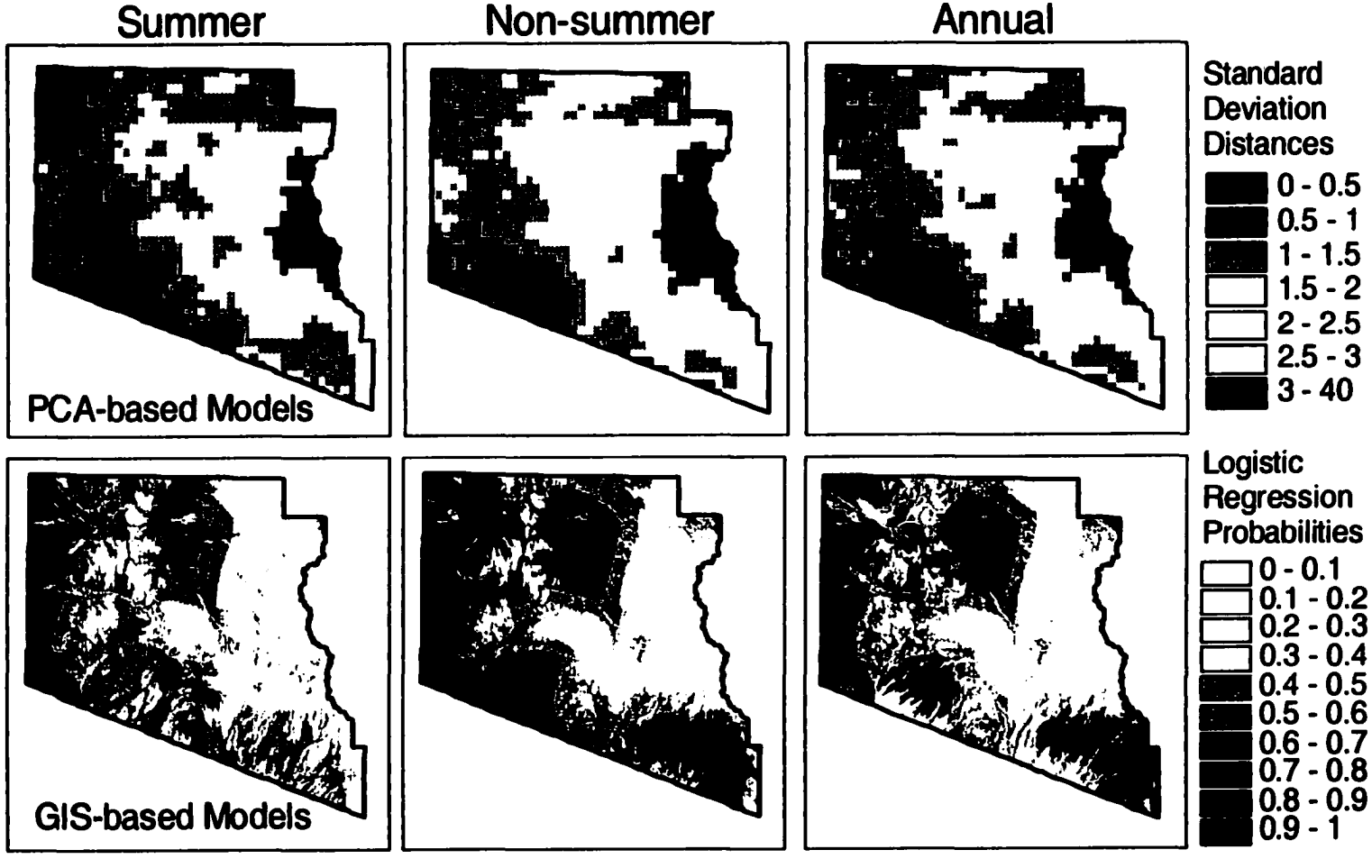


Figure 22: Visual comparison of AVHRR-based PCA models (top) with Logistic regression models (bottom).

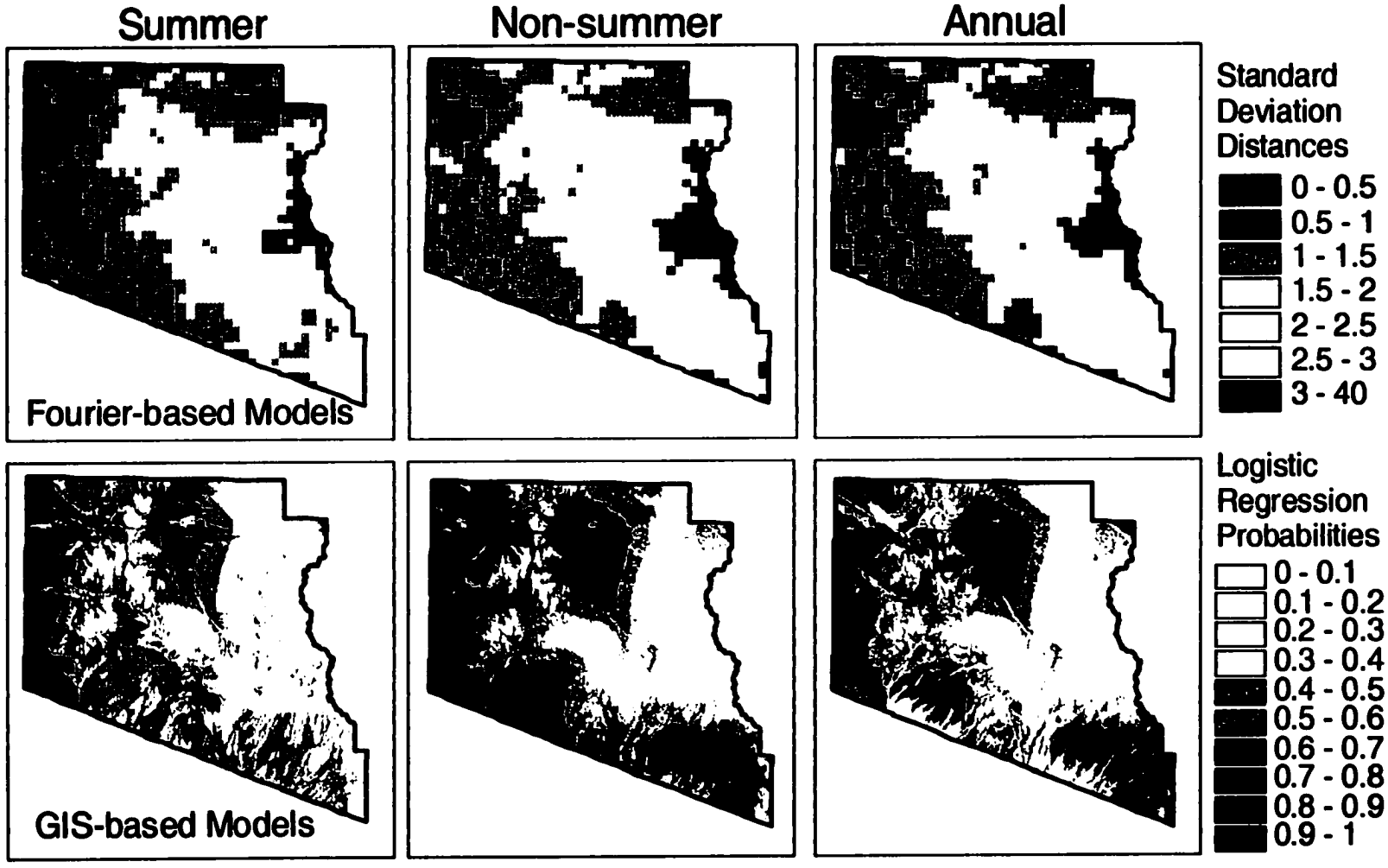


Figure 23: Visual comparison of AVHRR-based Fourier models (top) with Logistic regression non- models (bottom).

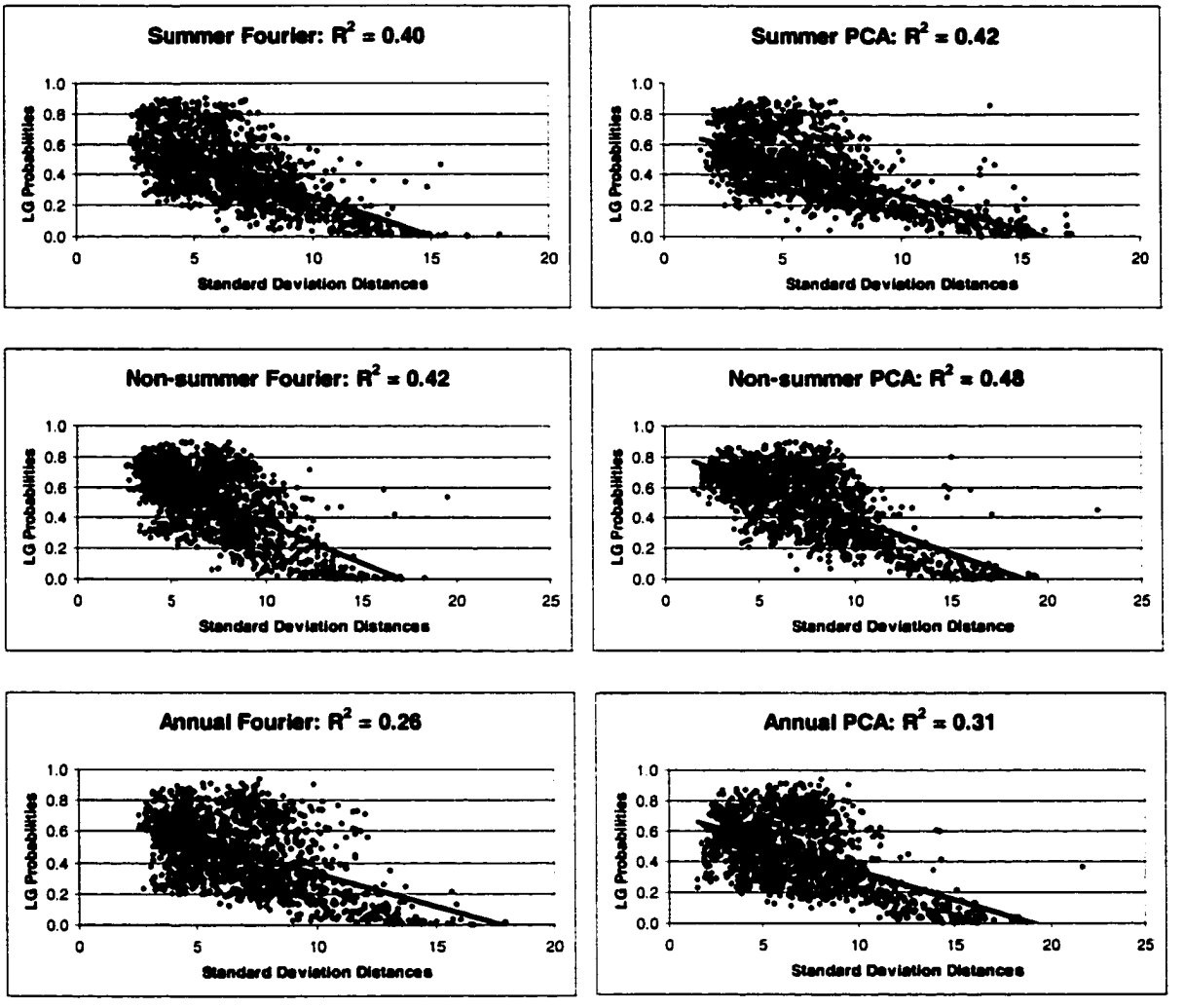


Figure 24: Regression results between GIS-based logistic regression probabilities (y axis) and the AVHRR standard deviation distances (x axis).

APPENDIX B

ABSTRACT

Our study used geostatistics to extract measures that characterize the spatial structure of vegetated landscapes from satellite imagery for mapping endangered Sonoran pronghorn antelope habitat. The fine (1-meter panchromatic and 4-meter multispectral) resolution data provides detailed spatial information at the scale of individual trees or shrubs to enable analysis of vegetation structure and pattern. Information derived from the satellite imagery was evaluated statistically and used to construct two independent models of pronghorn preference by coupling the structural measures with Sonoran pronghorn antelope sighting data: a distribution-based model, which uses the descriptive statistics for the structural measures at pronghorn sightings, and a cluster-based model, which uses the distribution of pronghorn sightings within clusters generated by an unsupervised classification of the structural measure images.

Both models map similar landscapes, and validation results confirm their effectiveness at predicting the locations of an independent set of pronghorn sightings. The geostatistical analysis extracts measures of landscape structure by calculating local estimates of the nugget, sill, and range variogram parameters within 25x25-meter image windows. These variogram parameters, which describe the spatial autocorrelation of the 1-meter image pixels, are shown in previous studies to discriminate between different species-specific vegetation associations. Our study demonstrates the use of the derived parameters in a practical application to define landscapes of preferred habitat for the Sonoran pronghorn antelope. Such information, although not a substitute for field-based knowledge of the landscape and the associated ecological processes, can provide valuable reconnaissance information to guide natural resource management efforts.

abundant forbs in response to winter rains (Hughes, 1991; Kilgore, 1997). Summer landscapes include shrubbier bajada slopes that host perennial chain-fruit cholla (*Opuntia fulgida*), a common and reliable food source during the dry season (Hughes, 1991). Structurally, therefore, non-summer landscapes should be more homogeneous than the summer landscapes. Furthermore, since pronghorn are known to avoid areas with abundant perennial vegetation cover (Irby, 1997; Selting, 1997; Singer, 1994; Singer, 1994; Wood, 1989); pronghorn should occupy more homogeneous landscapes. Our overall goal was therefore to determine if these differences could be mapped using the structural measures derived from fine resolution imagery.

Our research was therefore designed to extract and evaluate measures that characterize the spatial structure of vegetated landscapes from satellite imagery for mapping Sonoran pronghorn antelope habitat. The fine (1-meter) resolution panchromatic data provides detailed spatial information at the scale of individual trees or bushes to enable analysis of vegetation structure and pattern. Information derived from the satellite imagery is used to stratify the landscape into favorable regions defined by coupling the structural measures with Sonoran pronghorn antelope sighting data. By applying statistical and mathematical methods, we translate high spatial resolution remote sensing data into information that should be useful to resource managers at both the regional and local level.

One common technique for improving land cover maps is to include measures of image texture. These measures are generally created by using standard convolution filtering and classification techniques with commonly utilized remote sensing data (e.g. Landsat TM/ETM and Spot) (Jensen, 1996). However, recent studies have incorporated more sophisticated measures of spatial autocorrelation using the variogram function of geostatistics (Stein, et al., 1998; Curran, 2001). Geostatistics is a class of spatial statistical techniques commonly used to characterize the spatial autocorrelation found in natural systems (Isaaks and Srivastava, 1989; Cressie, 1993). The spatial patterns of a region can be characterized quantitatively by the experimental variogram, which is derived by calculating one-half the average squared data value difference for every pair of data locations and plotting this value against the distance between the data pairs, expressed formally as:

$$\text{gamma(h)} = \frac{1}{2}\frac{1}{N}\sum_{i=1}^{N}[Z(x_i) - Z(x_i + h)]^2 \qquad (1)$$

Where: gamma (h) is the value of the variogram at lag h, x_i is a data location, h is a lag vector, $Z(x_i)$ is the data value at location x_i, and N is the number of data pairs spaced a distance and direction h apart. An omni-directional variogram results if data pairs spaced a specified distance for all directions are used in the calculation.

The resulting experimental variogram is then fitted with an appropriate theoretical model variogram having specific mathematical properties (Figure 1). The model variogram is

fitted visually by the analyst, and each proposed model is evaluated by calculating a "goodness of fit" measure until a suitable fit to the experimental variogram is achieved. The nature of the variability characteristic of the data set is reflected in the overall shape of the model variogram (Isaaks and Srivastava 1989; Webster and Oliver 1990; Leibhold 1993). Parameters of the model variogram include information on the distance over which data are correlated (the range); the level of random variation within the data (the nugget); and the total variation present (the sill).

Studies that have applied geostatistics to synthetic and actual images have found direct ties between several scene characteristics and the behavior of the variogram (e.g. Woodcock et al., 1988a and b). This prior research has demonstrated that the density of the image objects is related to the height of the variogram sill, the size of the image objects is related to the range of the variogram, and the variance in the size distribution of objects is related to the roundness of the variogram near the sill. In addition, these studies show that increasing the image pixel size results in lowered sill value; increased range; and increased nugget.

Other studies have used these geostatistical techniques for land cover mapping applications. For example, geostatistics can discriminate between different vegetation communities in remotely sensed images of arid and semi-arid regions (Wallace et al., 1999; Phinn, 1997). These studies found that the spatial patterns of plant distributions within vegetation communities can be quantified by the variogram parameters if the image pixel sizes are less than three to five meters. In another study, Herzfeld and Higginson (1996) use geostatistics to enhance subtle topological features on a submarine fan. This research produced new images by calculating the variogram at various lagged distances for all pixels in a sonar image. These images were then input to standard image classification routines with improved mapping results.

Based on these studies, we have investigated the applicability of these geostatistical techniques for mapping structural information important to wildlife habitat. We developed a new method to calculate and map the local values of the variogram, which are found in previous studies to be distinctive for different vegetation communities (Wallace et al., 1999; Phinn, 1997).

2.0 RESEARCH OBJECTIVES

This research aims to provide insight into the spatial character of habitat. Our objectives were to:

(1). *Derive measures of spatial structure from imagery using geostatistics.* Because the imagery has a spatial resolution of 1-meter, it allows the identification of individual shrub and tree canopies. Spatial measures derived from this imagery therefore capture information on the distribution and pattern of

study area. Although the locations of the antelope may not necessarily reflect ideal habitat, the operating assumption is that the majority of the sightings occur in landscapes that are in some sense preferred by the pronghorn.

2. *imagery:* 1-meter black-and-white (panchromatic) and 4-meter multispectral (blue, green, red, near-infrared) imagery were purchased from Space Imaging for this project. The data were acquired on December 12, 2001, and consisted of the "Geo" product, which has undergone minimal pre-processing and only a standard geometric correction.

3. *Digital Orthophoto Quarter Quadrangles (DOQQs):* Digital color-infrared (CIR) DOQQs produced by the USGS were downloaded from the Arizona Regional Image Archive (ARIA) website (http://aria.arizona.edu/). 12 DOQQs covered the study area. The 1-meter spatial resolution of these images makes them ideal as reference data to accomplish georectification of the imagery. An original attempt to extract spatial structure from the DOQQ data was unsuccessful: the geostatistical techniques identified a noise pattern in these data represented by a sub-horizontal striping across these images. The origin of this noise is unclear, and attempts to remove this noise were unsuccessful, rendering these data unsuited for this analysis.

4.0 RESEARCH METHODS

4.1 Image processing

Since our goal was to evaluate the spatial pattern of individual tree or shrub canopies on the landscape, the fine spatial information of the panchromatic band was merged with the spectral information of the multi-spectral bands. This process, commonly termed "pan-sharpening", was accomplished using the spatial enhancement tool called 'Resolution Merge' in the ERDAS Imagine image processing software suite. To preserve radiometric fidelity, the principal component method of resolution merging was specified and the output designated unsigned 16-bit to match the input data. To preserve the spatial details of the panchromatic data (albeit at some loss of radiometric accuracy) a cubic convolution re-sampling technique was applied. The process takes the 4-meter multi-spectral data and modulates it spatially using the brightness variations present in the 1-meter panchromatic data to produce a multi-spectral image with an effective spatial resolution of 1 meter.

To create a single band image that effectively portrays the distribution of vegetation across the landscape, we calculated the normalized difference vegetation index (NDVI). This index is calculated from the reflectance values in the 4-meter red (band 3) and near-infrared (band 4) bands, where NDVI = NIR-RED/NIR+RED, and produces values between –1 and 1, such that values greater than 0 represent land surfaces and values close

to one represent very dense vegetation. The NDVI represents a measure of vegetation photosynthetic activity, and is sensitive to various biophysical vegetation characteristics, such as biomass and percent cover (e.g. Huete and Jackson, 1987). It is worthwhile to note that it is necessary to first pan-sharpen the multispectral image and then calculate the NDVI.

The combination of the small pixel size and 16-bit radiometric resolution yielded extremely large data files, which complicated storage and data processing considerations. We therefore extracted only the NDVI images and rescaled them to 8-bit data for subsequent processing. The NDVI, calculated from the 16-bit image data, was rescaled to 8-bits using the formula utilized by the U.S. Geological Survey's EROS Data Center (EDC) to preserve their scale and provide a standardized output:

$$NDVI_{8bit} = (NDVI_{calc} + 1) * 200 \qquad (2)$$

This formula rescales the −1 to +1 data range of the calculated NDVI ($NDVI_{calc}$) to a range of 0 to 200 ($NDVI_{8bit}$), such that values greater than 100 represent land surfaces and values close to 200 represent dense vegetation. It was necessary to preserve the absolute NDVI values to combine the two image halves. Because the images were collected on the same day by the same sensor, we did not apply any other radiometric corrections but chose to work with relative reflectances for this analysis.

Inspection of the overlap area between the east and west halves revealed that the images were registered to each other, and the two halves were mosaiced using an overlay function. The image was then georectified in ERDAS Imagine using the CIR DOQQs as reference images. Although ERDAS Imagine software can orthorectify Landsat and Spot data using pre-defined satellite parameters, parameters for the satellite must be extracted on an image-by-image basis. Since this is not a feature available in Imagine at this time, we could not apply a valid terrain correction to the data to produce an orthorectified product. Therefore, it was necessary to avoid areas of high relief that showed significant relief displacement when selecting ground control points.

4.2 Deriving geostatistical measures

The 1-meter NDVI image was clipped to a rectangular area and exported to generic binary data. The local variogram values were calculated using Matlab, a mathematical computation and analysis software package (www.mathworks.com). It was necessary to divide the area into twelve smaller images due to computer performance limitations when running the Matlab routine. The twelve images consisted of two with dimensions 9075x3400 pixels and ten with dimensions 9075x3600 pixels.

The Matlab program written for this analysis segments the image into "tiles" of 25x25 pixels and calculates the value of the variogram for the 1-meter pixels at lags 1 through 8

within each tile. Although we originally considered calculating the variogram measures in a 25x25 meter "moving window", as is typical in many spatial filtering operations (Jensen, 1996), computer performance limitations required us to accomplish a data reduction by "tiling" the image. Eight total lags were chosen to capture the short-range spatial autocorrelation expected as a result of vegetation distributions. Most vegetation communities studied in arid and semi-arid landscapes reveal autocorrelation out to a lag of about 5 meters (Wallace et al., 1999; Phinn, 1997), reflected by the range value. Since the range is associated with the size of image objects (Woodcock et al., 1988a and b), this value likely represents the size of individual tree and shrub canopies. Larger lags can be influenced by topographic characteristics, such as regularly spaced washes.

The size of 25 meters-squared for the tiles was selected since it was large enough to capture a significant number of variogram pairs and it was small enough to represent a reasonable minimum mapping unit for vegetation communities. Calculations were performed on pixel pairs in the ordinal directions, resulting in the following number of total pairs for the specified lags: lag 1, 1200 pairs; lag 2, 1150 pairs; lag 3, 1100 pairs; lag 4, 1050 pairs; lag 5, 1000 pairs; lag 6, 950 pairs; lag 7, 900 pairs; and lag 8, 850 pairs. Although the omni-directional experimental variogram is calculated for all directions, we considered the ordinal directions sufficient for this study, thereby simplifying the programming task.

For each tile, the Matlab routine also selects the first five lags, fits a line to the ordered values, and extracts the intercept at lag 1 and the slope of the line. These values are analogous to the nugget and the [(nugget – sill)/range] (or slope) values of the variogram (see Figure 1). The ten total values calculated (eight lags, intercept, and slope) are then used to populate ten new matrices (i.e., images) that have a pixel size of 25x25 meters. Each new matrix summarizes one of the spatial measures calculated for the original 1-meter NDVI data within the 25x25 meter windows. The twelve small images for each of the ten derived geostatistical measures are recombined in Matlab to produce ten 856x726 matrices (images) covering the initial study area. The matrices were then imported from generic binary data into ERDAS Imagine format images of floating point values for subsequent analysis.

As previously mentioned, preliminary analyses of the derived images revealed significant differences between the east and west halves as a function of their view angle. We therefore subset the derived images and considered only the west half for the remainder of the analysis, which will be referred to as the study area.

Additional image products were also generated. An image of the variogram sill was calculated as the average of lags 6, 7, and 8. These are the lags not included in the derivation of the nugget and slope measures, and inspection of the lag profiles for various pixels suggested their average is a reasonable approximation to a local sill value. This average also fits with the findings of studies that examine the spatial structure of arid and

A total of 150 random points were selected, since a greater number of random points are required to adequately capture landscape variation (Kvamme 1985). As with the pronghorn locations, the image pixels within every 100-meter buffer around each random point were extracted for each of the derived geostatistical measures. The distribution of the pronghorn sightings and the random points in the study area is shown in Figure 3.

4.4 Statistical Analysis

The pixels extracted from the buffers around each pronghorn sighting or random point were treated as separate sample values for analysis, resulting in a total of 3141 samples from pronghorn location buffers and 7707 samples from random location buffers. The data were input to the SPSS statistical software for analysis. Exploratory results revealed the highly skewed nature of the nugget, sill and slope images. This is not surprising, since the variogram is calculated as the average difference squared of the data values separated by particular lag distances. These images were therefore transformed by calculating their square root and the transformed nugget, sill and slope values were extracted for each sample. Although the distributions of the transformed variables are still somewhat skewed, they are not as extreme as the original.

Statistical tests were performed to compare the distributions of geostatistical measures sampled by the pronghorn sightings and the random points. The difference of means for independent samples was tested using both the parametric Student's t-test as well as the non-parametric Mann-Whitney test. Since the sample distributions are skewed, it is possible that the t-test will not identify a significant difference that might exist. In such instances, the Mann-Whitney test, which makes no assumptions about the normality of the data, should be more useful.

The difference of geostatistical measures sampled by the pronghorn sightings and the random point is demonstrated statistically using a Student's t-test for the difference of means under the assumption of equal variance, using a significance level of $\alpha = .05$ (Davis, 1986; Salkind, 2000). This test evaluates the null hypothesis that the difference in the means for the set of spatial measures at pronghorn locations and the set at random points is zero, utilizing the t-distributed test-statistic.

For data that are highly skewed, a non-parametric test for difference of sample means is employed that is analogous to the t-test. We apply the Mann-Whitney test for difference in means, which takes the two data sets of n_1 and n_2 samples, combines them, orders the values from smallest to largest, and then assigns a rank to each of the values from 1 to $n_1 + n_2$. The test for significance compares the mean rank of the samples for each of the two original groups (see http://www.tufts.edu/~gdallal/npar.htm).

The derived geostatistical measures that effectively separate pronghorn sightings from random points, as evidenced by the above tests, were then selected as input to the habitat model.

4.4 Habitat Modeling

Mapping spatial structure of landscapes preferred by pronghorn: The selected images of geostatistical measures were stratified into pronghorn preferences using the distribution of the variables observed for pronghorn sightings samples. Due to the skewed nature of these data, we chose to stratify the image using the observed data values at the quartile breaks. 'Preferred' landscapes were identified as those with values between the 25th and 75th percentile breaks, which include the middle range of values observed for half of the image pixels sampled in the sightings buffers. Landscapes with values less than the 25^{th} percentile or greater than the 75^{th} were identified as 'not preferred'. This produced binary maps of preferred pronghorn spatial structure for each of the derived geostatistical measures.

The selection of the quartile breaks to stratify the image was subjective: in this small study area, these breaks appeared to define a map with an interesting and potentially informative pattern of contrast. If the study area was larger, it may be necessary to shift the thresholds defining this boundary.

Distribution-based Model Development: The binary images of preferred spatial structure, based on the observed data distributions, were combined to create models of Sonoran pronghorn antelope habitat. We elected to combine the binary images of pronghorn preference in two ways: as a union and as an intersection. Creating a union of the binary images produces a map with the maximum preferred area: landscapes are considered preferred if any of the input measures were preferred. Creating an intersection of the binary images produces a map with the minimum preferred area: landscapes are considered preferred only if all of the input measures were preferred. This procedure produced two binary maps of preferred pronghorn spatial structure.

Cluster-based Model Development: For comparison and data exploration purposes, an alternative model of Sonoran pronghorn antelope habitat was created using the original eight lag images. In this process, the eight lag images were first input as an 8-band image to an unsupervised classification routine. This produced coherent clusters of landscape spatial character, based on the combinations of the values for the different lags. Secondly, the distribution of the pronghorn sightings within these clusters was evaluated. A cluster was considered 'preferred' if it contained more sightings than would be expected based on its areal size, 'moderately preferred' if sightings were what would be expected, and 'not preferred' if the number of sightings was less than expected. This produced a map of preferred pronghorn spatial structure with three ordinal categories.

A visual check of the registration was performed by overlaying the DOQQ and NDVI images and systematically sampling the difference between and DOQQ features. An overwhelming majority of these measures were very small (less than 1-meter), although differences up to 81 meters were observed in the higher elevations. A check of the displacement in areas near OPCNM sightings locations revealed a few differences up to 26 meters. Since we analyzed the 100-meter neighborhood around each sighting, the registration errors noted are acceptable for this study. Ideally, however, the imagery should be terrain corrected (orthorectified) since relief displacement may introduce more significant errors and in areas of significantly varying relief.

5.2 Deriving Geostatistical Measures

Images of the nugget, range, sill and slope are shown in Figures 6 and 7. Figure 8 shows images of lag 1 and lag 8. Based upon the frequency distribution of the data, all images (except for the range image) are displayed using a histogram equalization stretch to enhance their textural details.

Inspection of the derived images reveals striking differences between the east and west half prompting us to restrict our study area to the west half. We revisited and re-ran the original image processing procedures to confirm they were appropriate. Since the only significant difference between the two image halves is the view angle, and a slightly different texture is visible in the NDVI image (Figure 4), we conclude that the expression of spatial autocorrelation in the two halves is strongly dependant on this factor. As noted, the same sensor collected the two images one minute apart on December 12, 2001 but with very different view angles.

Metadata reports that the west half used a reverse scan direction with collection azimuth and elevation of 4 and 67 degrees, respectively, whereas the east half used a forward scan direction with collection azimuth and elevation angles of 253 and 85 degrees, respectively. The satellite descends from the north; therefore, the east image was collected with the sensor facing south. As a result, the east image preferentially sees the shadows of the image objects, since the sun is near its winter solstice and is farthest south. This geometry likely explains the increased variability identified in the east image, as evidenced by the overall higher lag values; higher nugget, sill and slope values; and the shorter range. The ability of these geostatistical measures to extract this information could be useful in studies that evaluate the Bi-directional Reflectance Distribution Function (BRDF) properties of landscapes.

5.3 Statistical Analysis

Exploratory statistical analysis of the data revealed the skewed nature of the nugget, sill and slope values, which indicated that we should transform the values by calculating their square root (Lark, 1996). The resulting distributions of the range and the transformed nugget, sill and slope for the images are approximated by the distribution of the random

and the graphs of the cluster signatures are included. Although a total of 10 clusters were requested as input to the unsupervised classification routine, the procedure was able to identify only nine distinct clusters. The classification has stratified the image effectively based on the shape of the variogram; landscapes characterized by high sills (i.e., they possess a high degree of variation) are shown as yellow and red, whereas landscapes characterized by low sills (i.e., they are fairly uniform) are shown as violet. The high sill landscapes include the hilltops and the washes, which host the larger tree and shrub canopies.

The number of pronghorn locations within each cluster was tallied and compared to the number of sightings that would be expected based on the areal proportion of the cluster in the study area (Figure 13). Based on inspection of these results, we created a three-class habitat model by combining the clusters as follows:

Class	*Lag Clusters*	*Rationale*
Highly Preferred	1 and 7	Observed > Expected
Moderately Preferred	3 and 4	Observed ≈ Expected
Not Preferred	5,6,8,9 and 10	Observed < Expected

The observed:expected tallies clearly show that Cluster 1 is highly preferred and that Clusters 3 and 4 are significant components of pronghorn habitat. The very small size of the remaining clusters suggests the results could vary if the study area was larger or more sightings were available; for example, Cluster 7 is considered highly preferred on the basis of a single sighting. The resulting three-class model of Sonoran pronghorn habitat based analysis of pronghorn occupation of image clusters is shown in Figures 14. This model was also filtered twice with a 3x3 majority filter to eliminate isolated pixels. Again we see the hilltops and washes defined as 'not preferred' and the 'preferred' landscapes in the low-relief areas. The classification signatures and the geographic distribution of the individual preferred clusters are shown in Figure 15. We see that the pronghorn are preferentially selecting the more homogeneous landscapes with the smaller sill values.

Model Validation: The validation results show that the distribution-based union model and the cluster-based model are effective at predicting the locations of the AGFD data. Chi-square results, comparing the number of observed pronghorn in the model classes with the number expected based on class area, are significant at the $\alpha = 0.2$ significance level for the binary model (Table 5) and at the $\alpha = 0.1$ significance level for the cluster-based model (Table 6). Significant results were not achieved for the more restrictive distribution-based intersection model (Table 5).

Model Comparison: The distribution-based union model and the cluster-based model are shown as a side-by-side comparison in Figure 16. To facilitate this comparison, the "high" and "moderate" classes of the cluster-based model were combined to form a single "preferred" class. Both models map very similar landscapes, providing reassurance that the models are capturing valid information on pronghorn preferences since they were

derived using independent methods. Although the distribution-based model allows more direct control over the selection of thresholds used to stratify the input images, the cluster-based model allows for a more straightforward interpretation of the landscapes based on the coherent clusters with distinctive variogram-shaped signatures. In addition, the distribution-based model requires a set of wildlife sightings for development. Theoretically, it would be possible to approximate a wildlife habitat model using the cluster-based approach if you had detailed knowledge, perhaps from a different area, of the vegetation community and spatial structure preferences of the species being studied.

6.0 CONCLUSIONS

The method of deriving geostatistical measures from high spatial resolution satellite data described in this paper is shown to be effective at capturing key measures of landscape spatial structure. By applying this method to an NDVI image, we have been able to demonstrate the application of geostatistical measures to wildlife habitat modeling. This is one of the first studies to apply these measures to satellite data. In addition, most studies that apply geostatistics to extract spatial structure from images in support of vegetation mapping demonstrate its viability by comparing results for discrete images from known vegetation types (e.g., Phinn, 1997; Lark, 1996). In contrast, this study presents a method of mapping and evaluating the derived geostatistical measures regionally.

The results from the clustering of the 8-band lag image show that unsupervised classification is effective at extracting coherent landscape units on the basis of their spatial autocorrelation. It is significant that all pixels were assigned to clusters whose signature graphs exhibit a classic variogram shape (Figure 12). This result suggests not only that the 25x25-pixel window used in the analysis is large enough to effectively capture an expression of the vegetation structure of the 1-meter pixel NDVI data, but also that these geostatistical measures can provide key information for potential detailed vegetation mapping. The variogram parameters, captured in the cluster signature graphs, are shown in previous studies to discriminate between different species-specific vegetation associations (Wallace et al., 1999; Phinn, 1997). Furthermore, applying the procedures described herein extracts valuable information about the landscape structure while accomplishing a significant reduction in the data volume, since the input 1-square meter pixel images were output as a series of 25 square meter pixels.

Our study also demonstrates the use of the derived geostatistical measures in a practical application to define landscapes of preferred habitat for the Sonoran pronghorn antelope. We find the pronghorn preferentially accessing more homogeneous landscapes with lower nugget, sill and slope measures relative to the random points. This preference for more homogeneous landscapes fits field-based observations: our sightings data are predominantly from the non-summer months, during which pronghorn are known to access low-elevation flood plains within stabilized dunes (Hughes, 1991; Kilgore, 1997).

These landscapes are characterized by sparse perennial vegetation that host abundant forbs in response to winter rains. Structurally, these landscapes are relatively homogeneous at the scale of our analysis when compared to the shrubbier vegetation communities found in the higher elevation bajadas and hilltops, and this difference is captured in our study. Geostatistical measures of landscape structure from imagery, although not a substitute for field-based knowledge of the landscape and the interaction of pertinent ecological processes, can provide valuable reconnaissance information to guide natural resource management efforts.

Potential application of these procedures include detailed vegetation mapping, monitoring habitat condition, recognizing and mapping habitat fragmentation, and predicting potential wildlife locations within diverse landscapes. The technique can also be applied to monitoring land cover change. For example, the invasion of shrubs into grassland or the invasion of grass into shrubland should be very evident using this method.

ACKNOWLEDGEMENTS

The U.S. Department of the Interior, Organ Pipe Cactus National Monument and the U.S. Department of Defense, U.S. Airforce directly sponsored this research. In addition, we would like to thank the Arizona Game and Fish Department (AGFD) and OPCNM for facilitating access to their data and many helpful suggestions. We would also like to thank Mr. Thomas Potter (OPCNM) and Mr. James DeVos (AGFD) for their invaluable assistance and Mr. Jay Miller for his early contributions to this work. The manuscript benefited from the scientific and technical insights of Jessica Walker as well as from her awesome talent as an editor and wordsmith. The first author is also indebted to the thoughtful critiques of Candice Nelson.

REFERENCES

Cressie, N. (1993) *Statistics for Spatial Data;* New York: Wiley

Curran, P.J. (2001) Remote sensing: using the spatial domain. *Environmental and Ecological Statistics* 8:331-344.

EDC (1994) *Conterminous United States AVHRR Data on CD ROM - User's Manual,* U.S. Department of Interior, EROS Data Center, Sioux Falls, SD.

Hervert, J. J., R. S. Henry, M. T. Brown, and L. A. Piest, 1996. *Sonoran pronghorn population monitoring: progress report*, Nongame and Endangered Wildlife Program Technical Report 110. Arizona Game and Fish Department, Phoenix, Arizona.

Herzfeld, U.C. and C.A. Higginson, 1996. Automated Geostatistical Seafloor Classification - Principles, Parameters, Feature Vectors, and Discrimination Criteria: *Computers & Geosciences,* 22:35-52.

Huete, A.R., and Jackson, R.D., 1987. Suitability of Spectral Indices for Evaluating Vegetation Characteristics on Arid Rangelands. *Remote Sensing of Environment.* 23:213-232.

Hughes, K.S. (1991) *Sonoran Pronghorn Use of Habitat in Southwest Arizona.* The University of Arizona, M.S. Thesis. 86 p.

Irby, L.R., J. Saltiel, W.E. Zidack, J.B. Johnson (1997) Wild Ungulate Damage: Perceptions of Farmers and Ranchers in Montana. Wildlife Society Bulletin 25(2):320-329.

Isaaks, E.H., and R.M. Srivastava (1989) *An Introduction to Applied Geostatistics;* Oxford University Press, New York.

Jensen, J.R. (1996) *Introductory Digital Image Processing: A Remote Sensing Perspective, 2nd edition.* New Jersey: Prentice-Hall, Inc. 316 p.

Kilgore, M.J. and W.S. Fairbanks (1997) Winter Habitat Selection by Reintroduced Pronghorn on Antelope Island, Great Salt Lake, Utah. *Great Basin Naturalist* 57(2):149-154

Kvamme, K. L., 1985. Determining Empirical Relationships between the natural environment and prehistoric site locations: A Hunter-Gatherer Example. in *For Concordance in Archaeological Analysis.* Ed. C. Carr, 208-39. University of Arkansas, Fayetteville.

Walker, J.J., 2000. *A GIS Model of Sonoran Pronghorn Antelope Habitat in Organ Pipe Cactus National Monument, Arizona*. Master's Thesis, Department of Geography and Regional Development, The University of Arizona, Tucson, 121p.

Wallace, C.S.A., J. M. Watts, and S. R. Yool, 2000. Characterizing the Landscape Structure of Vegetation Communities in the Mojave Desert Using Geostatistical Techniques. *Computers and Geosciences*. 26(4):397-410.

Wallace, C.S.A., 1997. *Geostatistical Evaluation of Ground-based Vegetation Data in a Portion of the Mojave Desert, California*. Master's Thesis, Department of Geography and Regional Development, The University of Arizona, Tucson, 136p.

Webster, R. and Oliver, M.A. (1990); *Statistical Methods in Soil and Land Resource Survey;* Oxford University Press, New York

Wisdom, M.J. and Thomas, J.W., 1996. *Elk*, in Krausman, P.R., ed., *Rangeland Wildlife*. Society of Range Management. Denver, CO.

Wood A.K. (1989) Comparative Distribution and Habitat Use by Antelope and Mule Deer. *Journal of Mammalogy* 70(2):335-340.

Woodcock, C.E., A.H. Strahler and D.B. Jupp (1988a); The Use of Variograms in Remote Sensing: I. Scene Models and Simulated Images; *Remote Sensing of Environment;* 25:323-348.

Woodcock, C.E., A.H. Strahler and D.B. Jupp (1988b); The Use of Variograms in Remote Sensing: II. Real Digital Images; *Remote Sensing of Environment;* 25:349-379

Statistics

		NUGGET	RANGE	SILL	SLOPE
N	Valid	7707	7707	7707	7707
	Missing	0	0	0	0
Mean		8.9867	4.6101	31.6948	13.9218
Std. Error of Mean		4.431E-02	6.476E-03	.2327	9.762E-02
Median		7.6158	4.5818	24.4813	10.9545
Mode		6.16	4.67	14.80[a]	8.60
Std. Deviation		3.8903	.5685	20.4254	8.5704
Variance		15.1346	.3232	417.1981	73.4515
Skewness		2.164	.253	2.539	2.429
Std. Error of Skewness		.028	.028	.028	.028
Kurtosis		6.132	-.025	8.514	7.759
Std. Error of Kurtosis		.056	.056	.056	.056
Range		36.33	3.80	189.65	82.44
Minimum		4.24	2.89	8.19	3.32
Maximum		40.57	6.69	197.84	85.76
Percentiles	10	5.8309	3.9035	16.6933	7.4162
	20	6.2460	4.1276	18.7883	8.4262
	25	6.4807	4.2184	19.7400	8.8318
	30	6.7082	4.2969	20.5702	9.2195
	40	7.1414	4.4408	22.3398	10.0000
	50	7.6158	4.5818	24.4813	10.9545
	60	8.3066	4.7208	27.3313	12.1655
	70	9.3274	4.8823	31.4643	13.8564
	75	10.0499	4.9730	34.6266	15.3623
	80	11.0000	5.0813	39.4800	17.3494
	90	14.2478	5.3641	58.5002	25.1834

a. Multiple modes exist. The smallest value is shown

Table 1: Descriptive statistics of the pixels sampled in the 100-meter neighborhoods of random points.

Statistics

		NUGGET	RANGE	SILL	SLOPE
N	Valid	3133	3133	3133	3133
	Missing	0	0	0	0
Mean		8.2528	4.6142	28.8415	12.6572
Std. Error of Mean		6.161E-02	9.998E-03	.3360	.1411
Median		7.2111	4.5904	22.9637	10.2470
Mode		6.16	4.67	19.10	9.11
Std. Deviation		3.4488	.5596	18.8082	7.8954
Variance		11.8939	.3132	353.7495	62.3377
Skewness		2.851	.257	3.303	3.126
Std. Error of Skewness		.044	.044	.044	.044
Kurtosis		13.797	.051	19.457	17.095
Std. Error of Kurtosis		.087	.087	.087	.087
Range		39.80	4.12	274.65	111.55
Minimum		4.12	2.78	8.06	3.16
Maximum		43.92	6.89	282.71	114.71
Percentiles	10	5.5678	3.9286	15.2862	6.7823
	25	6.1644	4.2200	18.4481	8.1854
	40	6.7823	4.4431	20.9300	9.3594
	50	7.2111	4.5904	22.9637	10.2470
	60	7.7460	4.7238	25.3469	11.2250
	75	9.0554	4.9793	31.1582	13.7113
	90	12.3935	5.3510	50.7509	22.0227

Table 2: Descriptive statistics of the pixels sampled in the 100-meter neighborhoods of pronghorn sightings.

Ranks

	TYPE	N	Mean Rank	Sum of Ranks
NUGGET	1.00	3141	4910.65	15424344
	2.00	7707	5633.92	43420632
	Total	10848		
RANGE	1.00	3141	5443.20	17097078
	2.00	7707	5416.88	41747896
	Total	10848		
SILL	1.00	3141	4995.88	15692043
	2.00	7707	5599.19	43152932
	Total	10848		
SLOPE	1.00	3141	4979.36	15640160
	2.00	7707	5605.92	43204816
	Total	10848		

Test Statistics[a]

	NUGGET	RANGE	SILL	SLOPE
Mann-Whitney U	1.0E+07	1.2E+07	1.1E+07	1.1E+07
Wilcoxon W	1.5E+07	4.2E+07	1.6E+07	1.6E+07
Z	-10.911	-.397	-9.100	-9.451
Asymp. Sig. (2-tailed)	.000	.691	.000	.000

a. Grouping Variable: TYPE

Table 4: Results of Mann-Whitney test comparing differences of means for pixel values surrounding pronghorn sightings vs. values surrounding random points.

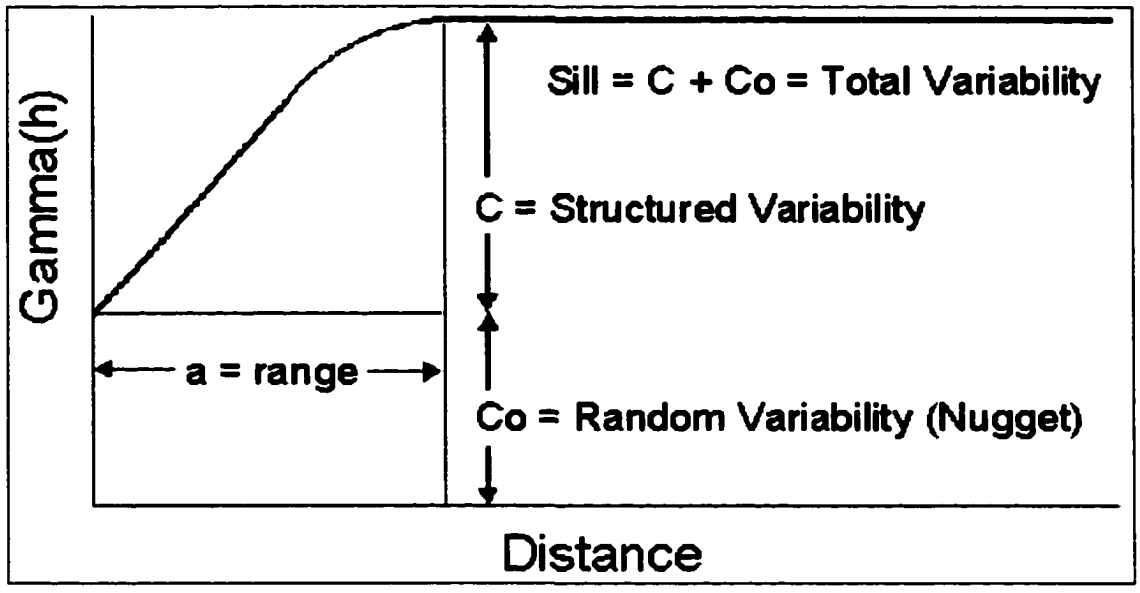


Figure 1: Example of a "Typical Variogram", after Cressie (1993). The horizontal axis plots lag distance, which is the directional distance between pairs of data points. The vertical axis plots the variogram value, termed "gamma", which is 1/2 the average squared difference in data values for all data pairs separated by the corresponding lag distance. Parameters of the model variogram include information on the distance over which data are correlated (the range); the level of random variation within the data (the nugget); and the total variation present (the sill).

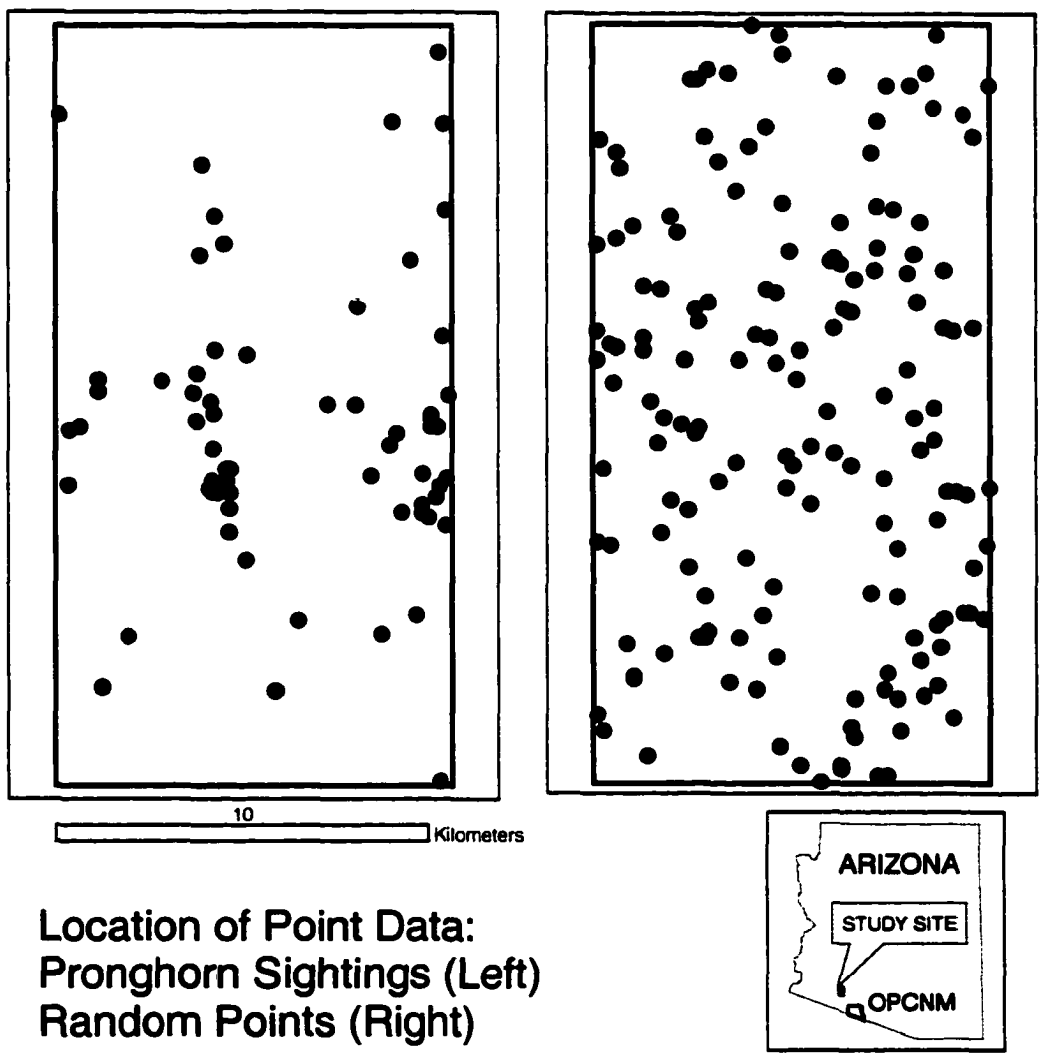


Figure 3: Location of point data.

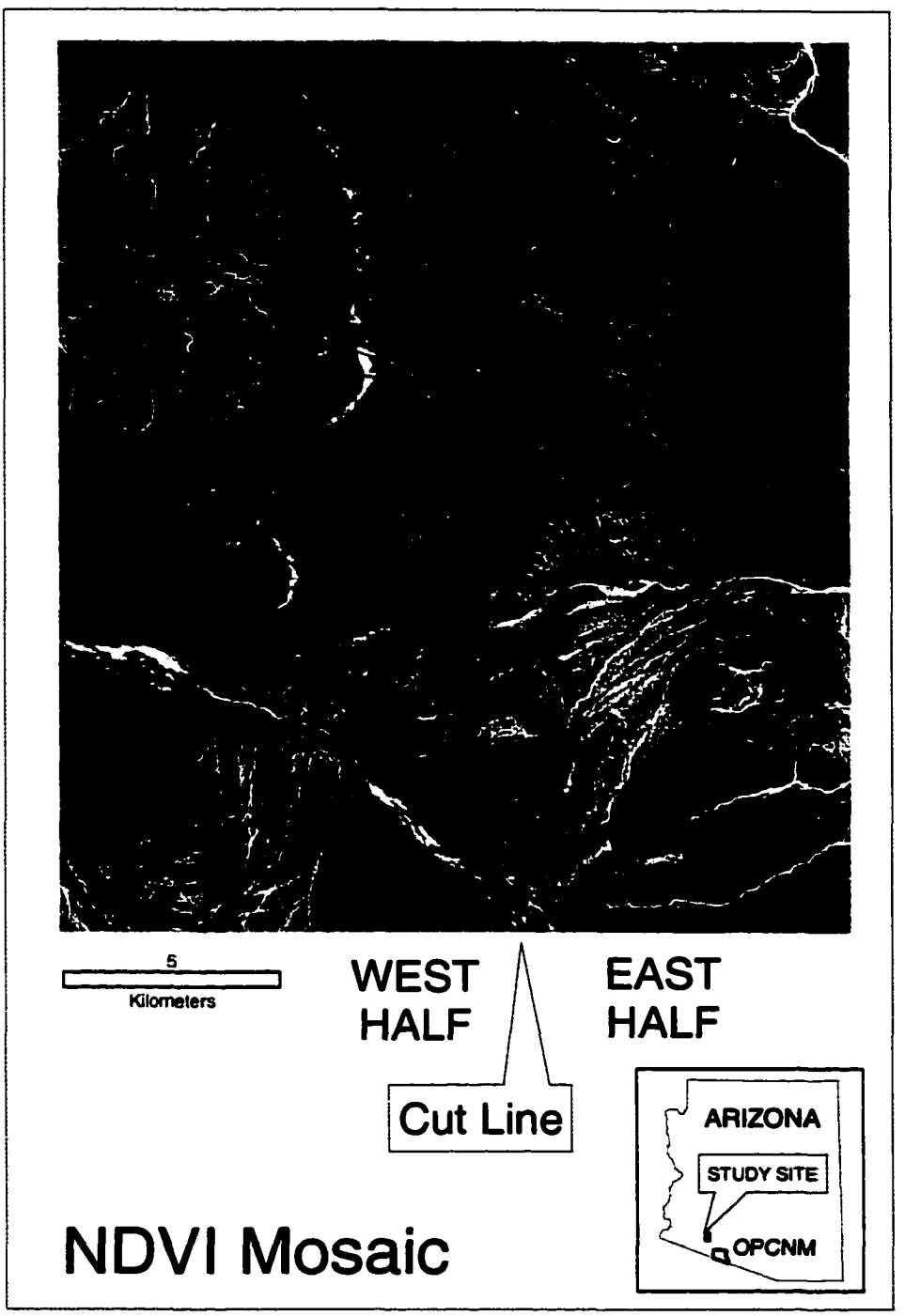


Figure 4: NDVI mosaic.

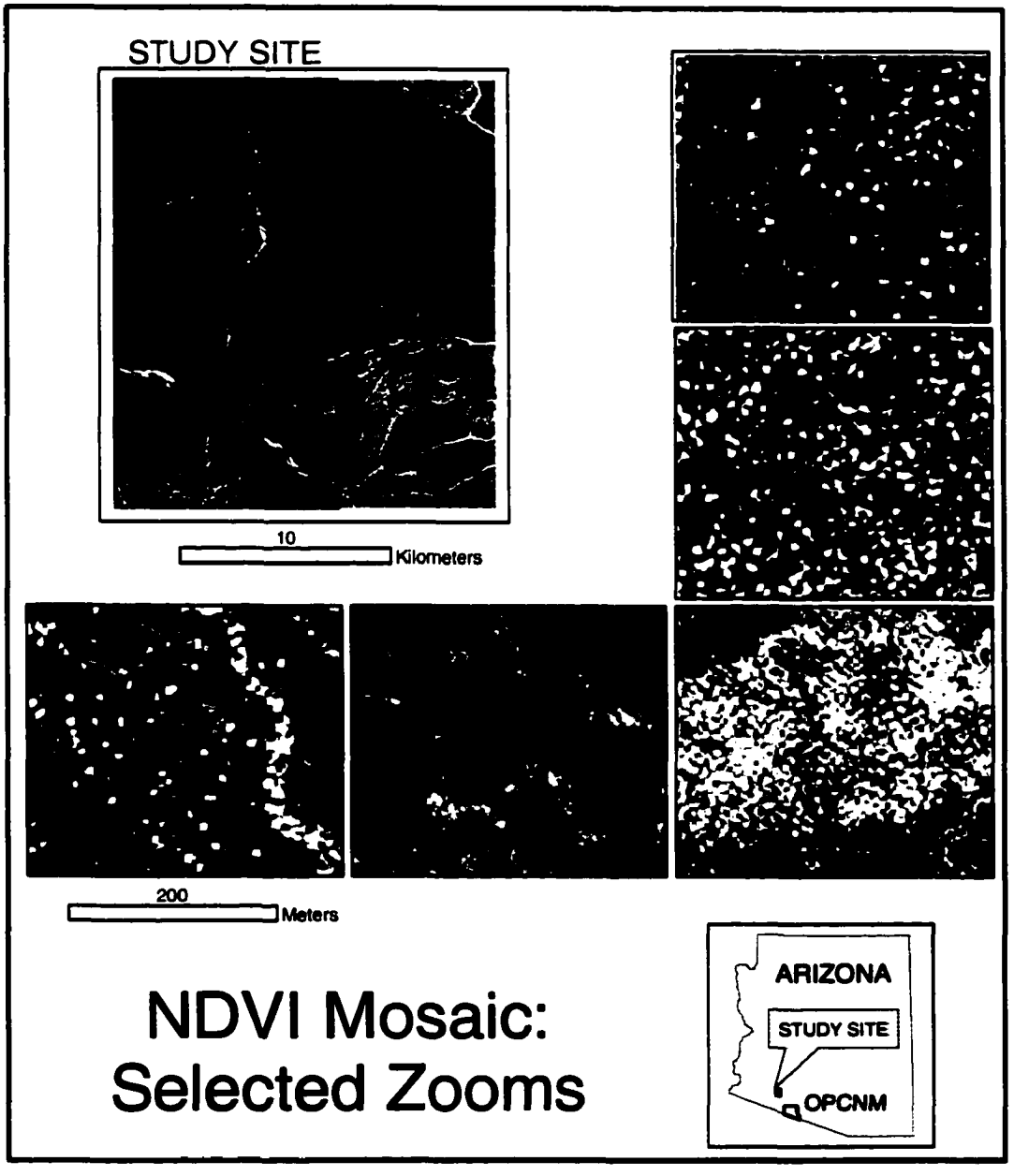


Figure 5: NDVI zooms of selected landscapes.

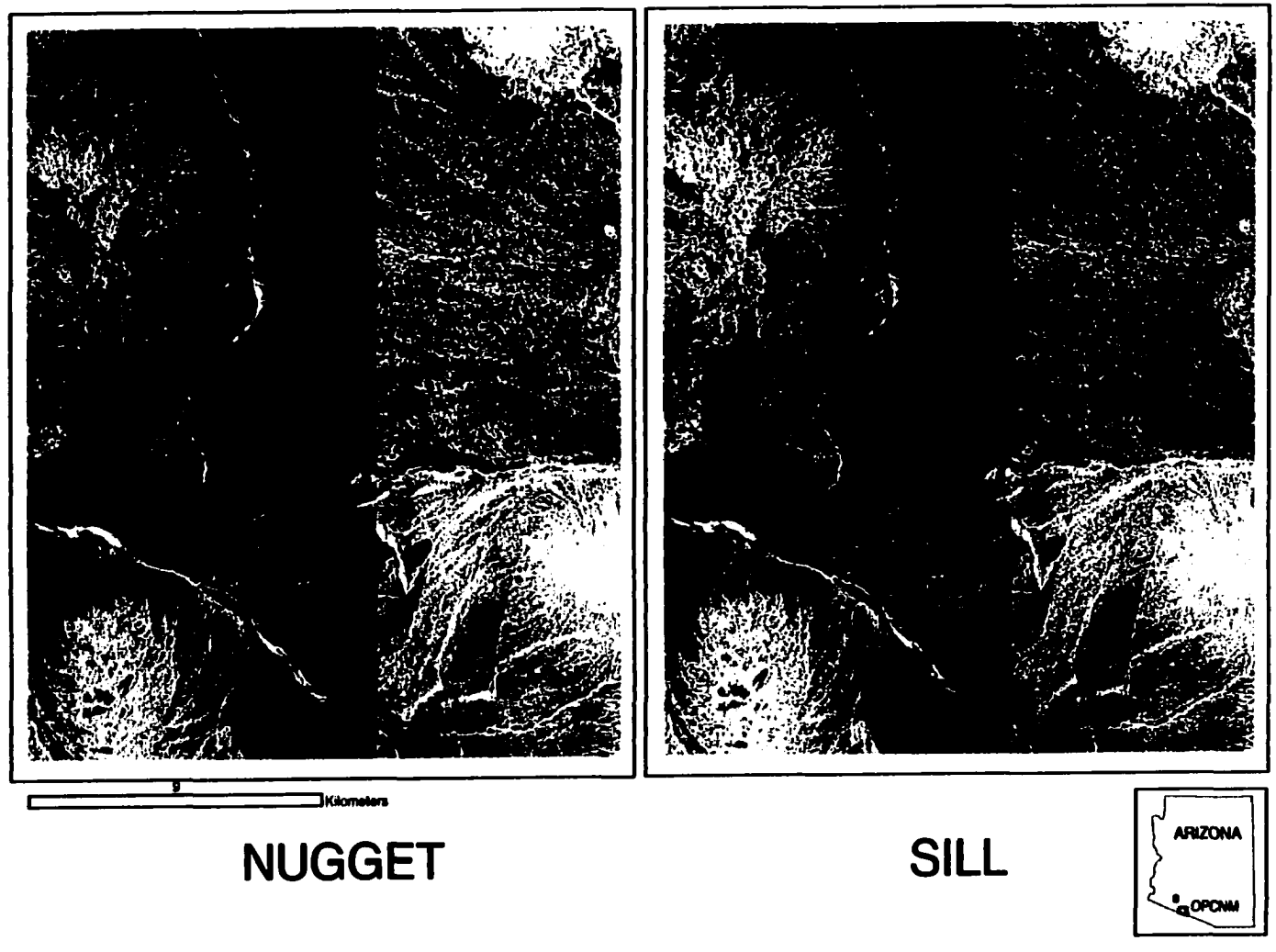


Figure 6: Nugget and Sill images. Note that these are the transformed values, obtained by taking the square-root of the calculated values. Data are displayed using a histogram equalization stretch.

Figure 7: Range and Slope images. Note that the Slope displays the transformed values, obtained by taking the square-root of the calculated values. Data for the Slope image are displayed using a histogram equalization stretch.

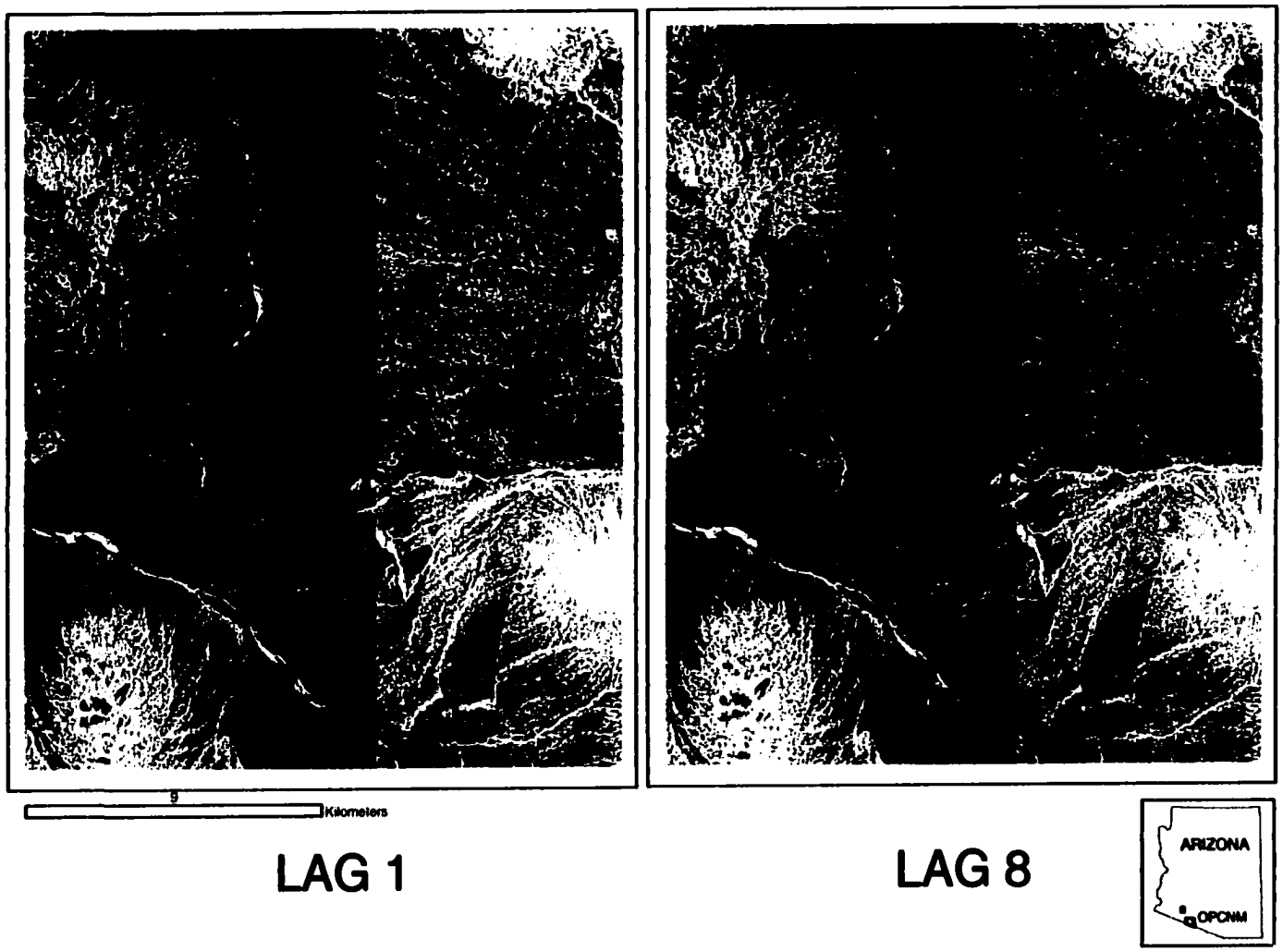


Figure 8: Lag 1 and Lag 8 images. Data are displayed using a histogram equalization stretch.

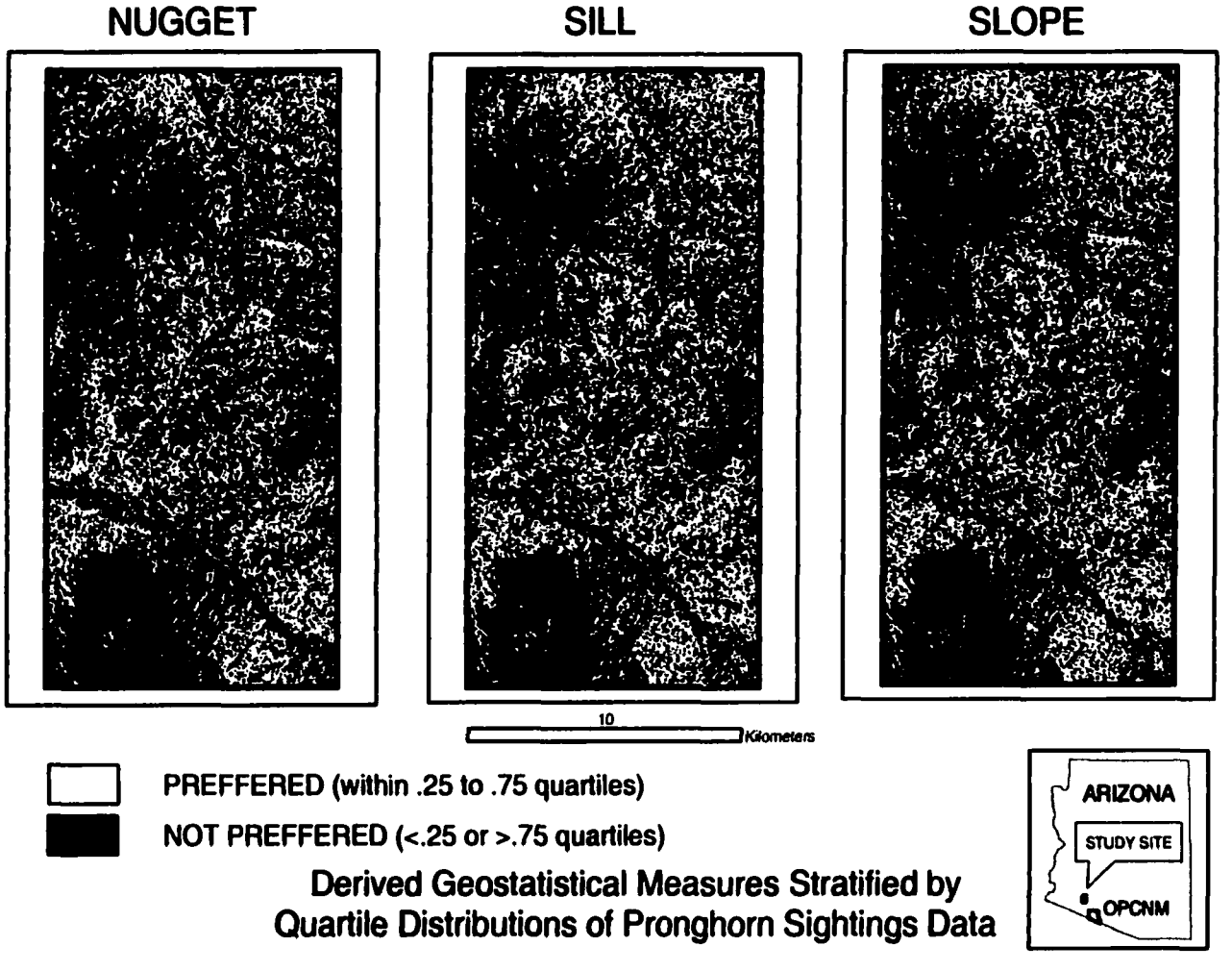


Figure 9: Images of geostatistical measures stratified by quartile distributions of pronghorn sightings data. White areas define landscapes with spatial characteristics preferred by pronghorn.

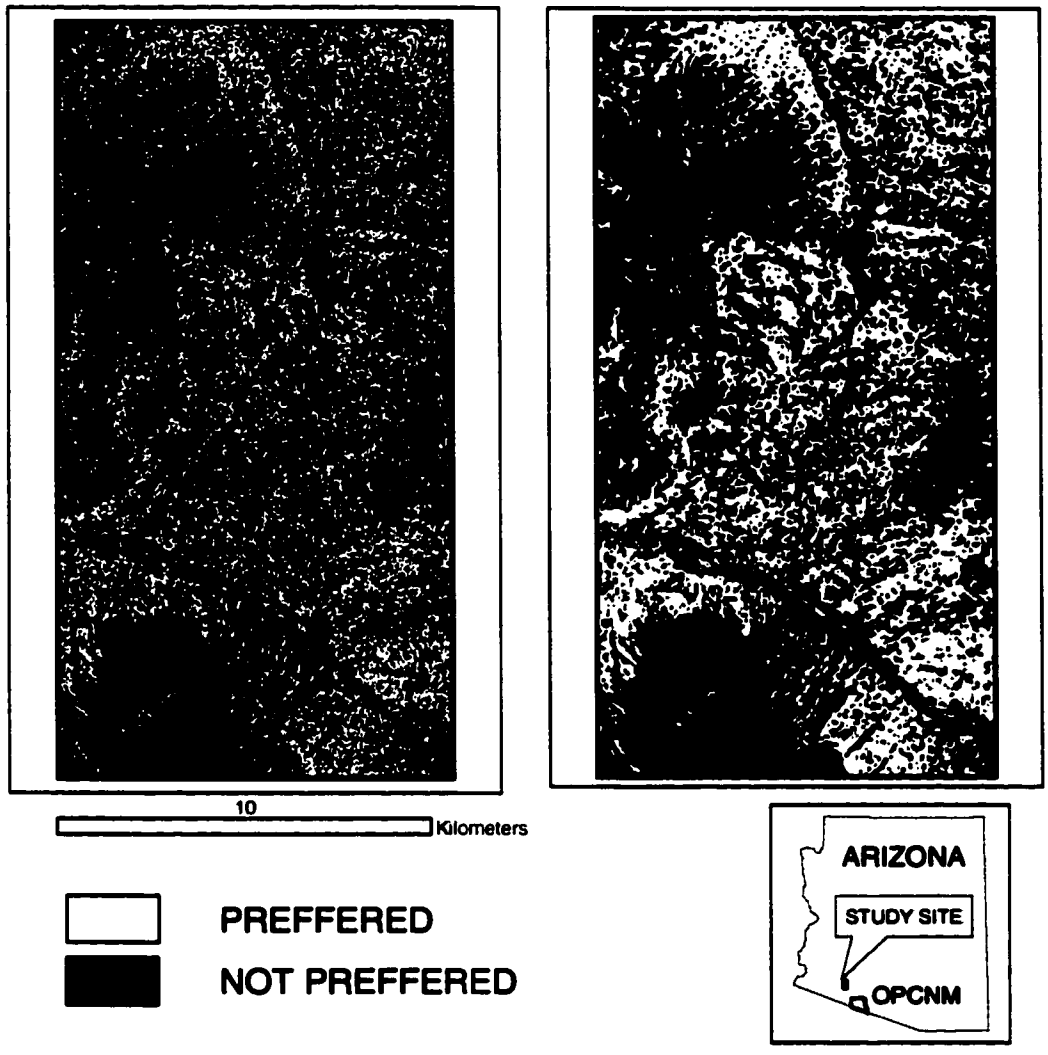


Figure 11: Distribution-based model of Sonoran pronghorn habitat preference minimum. This model is constructed as an intersection of the inputs from Figure 11.

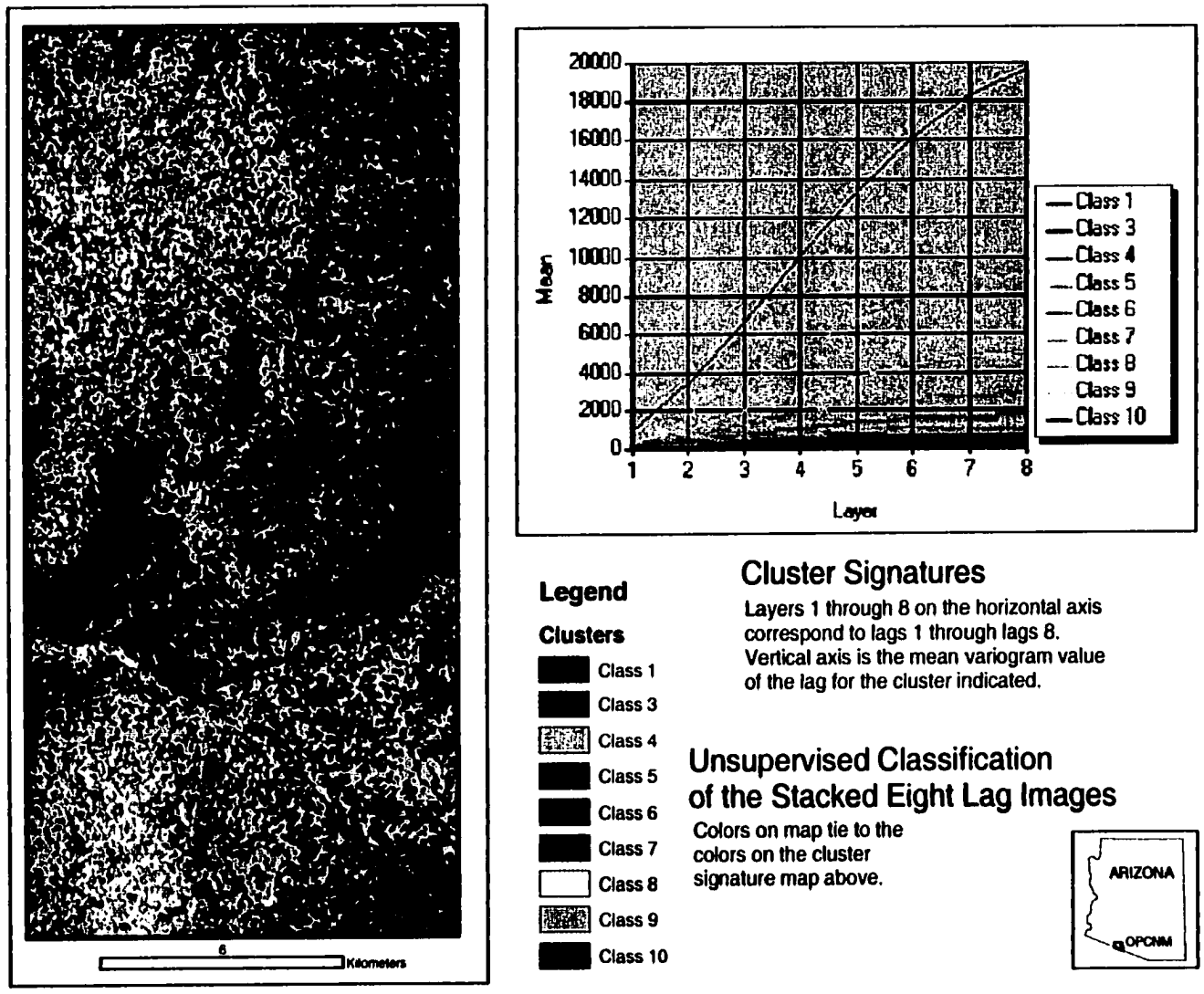


Figure 12: Results of unsupervised classification of the 8-band image composed of lag 1 through lag 8.

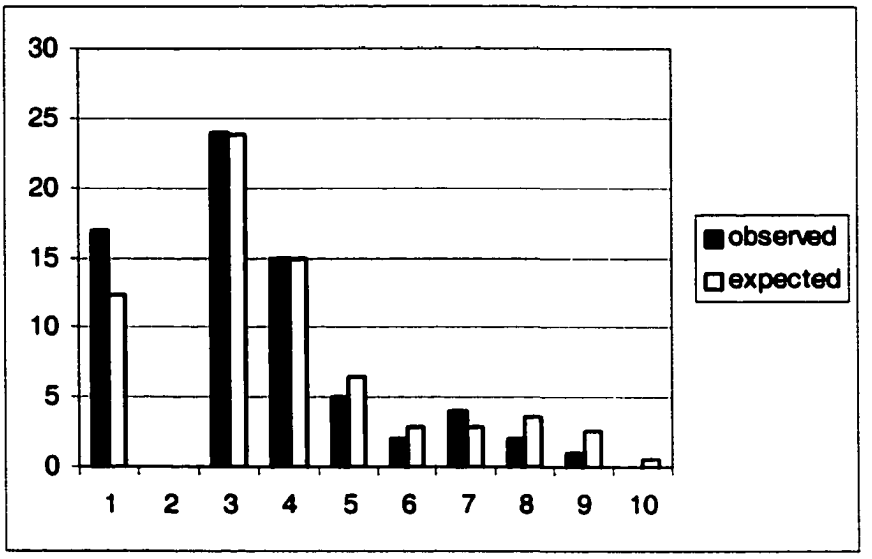


Figure 13: Observed vs. expected pronghorn sightings in the cluster classification.

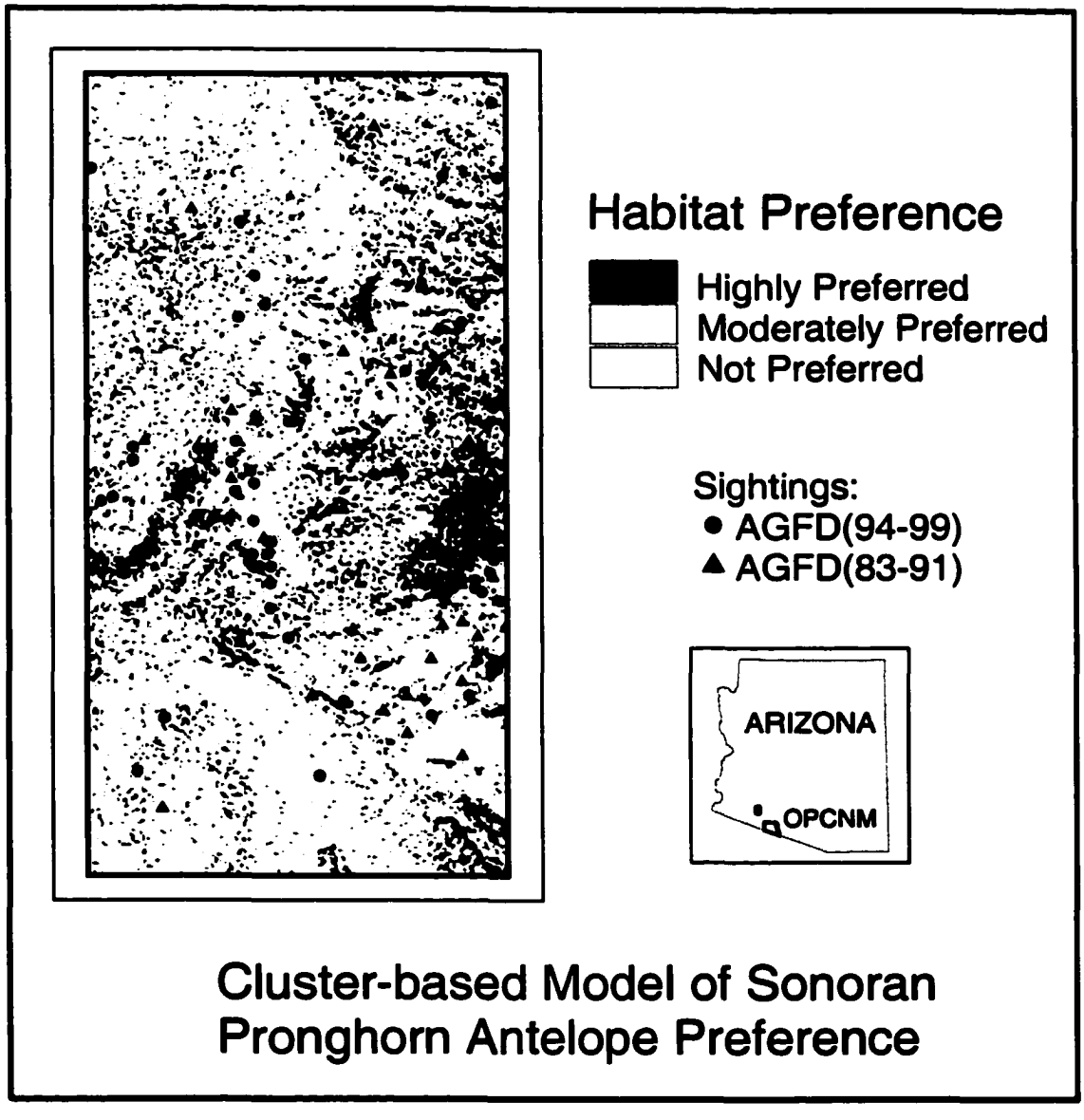
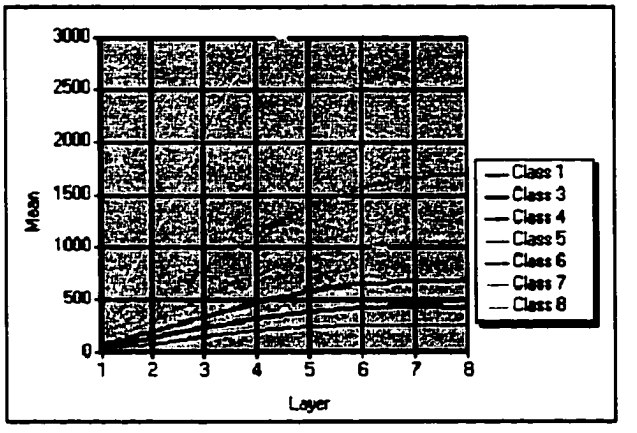


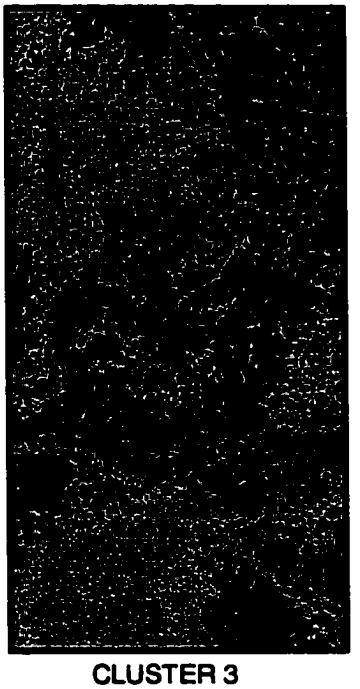
Figure 14: Cluster-based Sonoran pronghorn habitat preference model.



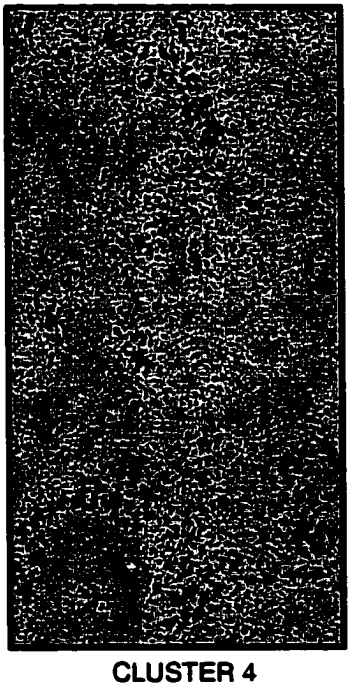
Cluster Signatures

Geographic Distribution of Clusters of High and Moderate Pronghorn Preference

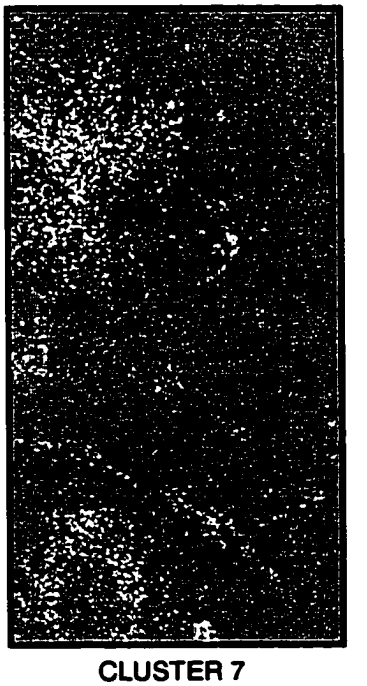
CLUSTER 1



CLUSTER 3



CLUSTER 4



CLUSTER 7

Figure 15: Geographic distributions of the pronghorn highly preferred and moderately preferred clusters.

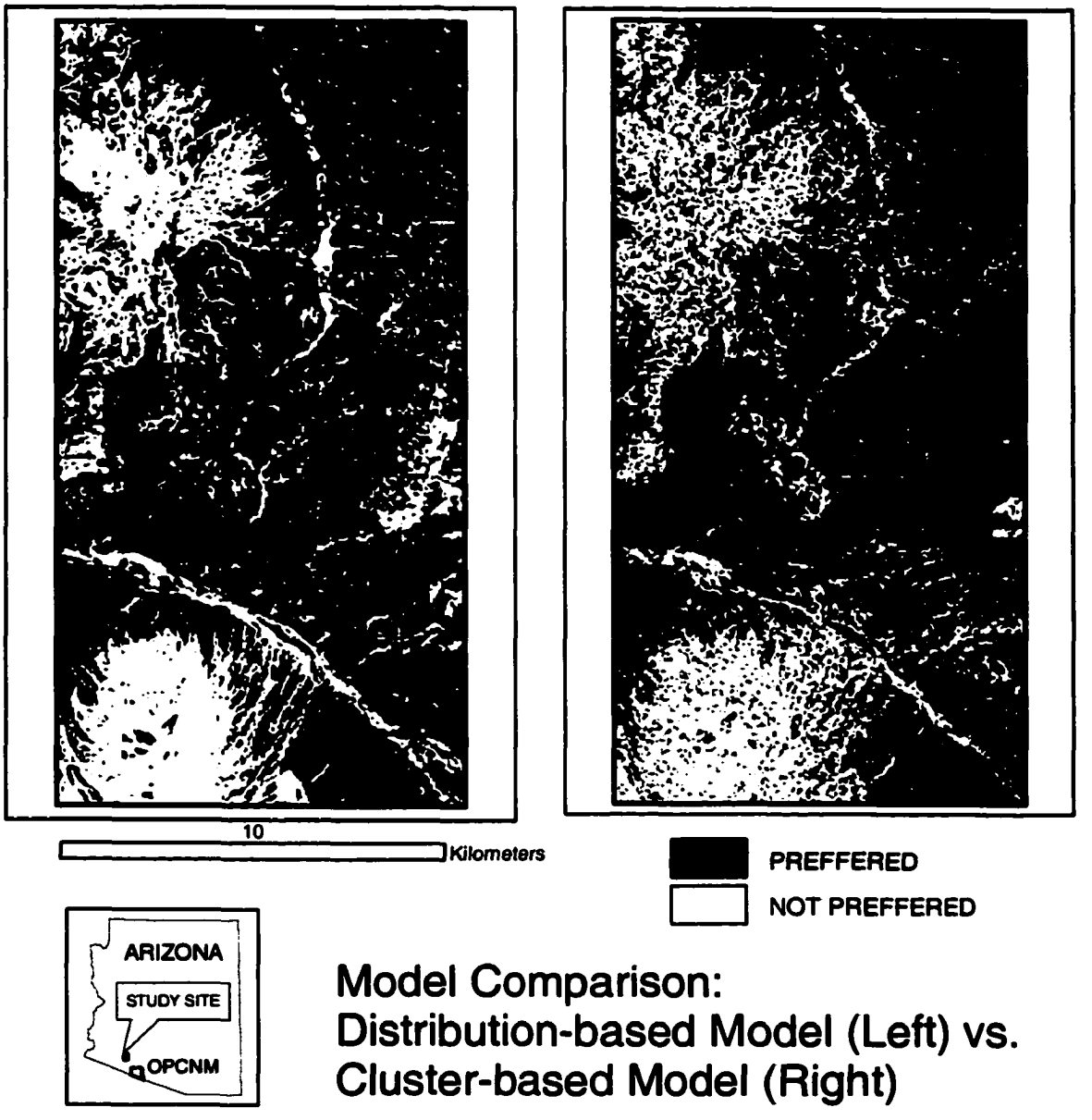


Figure 16: Side-by-side comparison of distribution-based model and cluster-based model.

APPENDIX C

Informing the Elk Debate: Exploring Relationships between Satellite Greenness Measures and Elk Population Dynamics in Arizona

Cynthia S.A. Wallace
Paul R. Krausman
Stuart E. Marsh
Charles F. Hutchinson

The University of Arizona
Office of Arid Lands Studies
Arizona Remote Sensing Center
1955 East Sixth Street
Tucson, Arizona 85719
520-621-8574

December 4, 2002

To Be Submitted To: *The Wildlife Society Bulletin*

INTRODUCTION

Natural resource management in the western United States is an inherently contentious undertaking. Typically, the interests of private landowners are pitted against those of a variety of other public and private interest groups, leaving public land management agencies caught in the middle. Public lands in the West are so vast, and the interest groups concerned are so large in number and diverse in their interests, that debates over natural resource management issues tend to be fierce and loudly public.

The dearth of generally accepted information on resource condition adds to the rancor (Burkhart, 1995, 1997; The Economist, 1998; Harris, 1995; Smith, 1998), since the various parties base their claims and expectations on disparate interpretations of resource health. The mostly semiarid climate of the region further confounds the issue, since conditions tend to vary considerably as a function of interannual variability in precipitation, often driven by El Niño and La Niña climatic events (Figure 1). In the end, debates over the management of natural resources occur against a constantly changing background of inadequate and often outdated information.

Since their reintroduction in 1914, elk (*Cervus elaphus*) have grown into a major management issue in most of the western United States (Wisdom and Thomas, 1996). The majority of the land area occupied by elk is overseen by Federal or state agencies, but individual ranchers have cattle grazing permits over much of this area. Because elk often compete with cattle for forage, and damage infrastructure (i.e., fences, water developments, and crops), ranchers are keenly interested in the management of the population. Environmentalists and hunters share this interest—albeit with different concerns and objectives. All of this attention ensures that the public agencies responsible for managing the elk populations are closely scrutinized when they attempt to judge the condition and health of the resource.

OBJECTIVE

The purpose of this project is to provide information that might help managers understand the distribution of elk in Arizona as a consequence of both seasonal variation and extreme climatic events (i.e., El Niño and La Niña). This understanding could facilitate better estimates of elk distribution and the identification of when and where elk-related conflicts might occur. In addition to improving management of the elk resource, this same information could be used by ranchers to better manage their livestock and vegetation resources. In conjunction with the Arizona State Land Department (ASLD), which is responsible for providing information to manage State Trust lands, research tasks were pursued to establish relationships between the remote sensing data provided by NASA and those factors that help understand and explain elk distribution through time.

The first task involved modeling elk populations over time. Although the Arizona Game and Fish Department (AGFD) collects census data for elk in Arizona, current estimates of elk populations are disputed (Harris, 1995; Noyes, 1995; Foust, 2001; Rawson, 1995). Because of this, we chose to implement a technique developed by Bender and Spencer (1999) for reconstructing elk population based on harvest data, gender ratios, and estimates of male mortality. The data required for the model are collected by the AGFD and tabulated according to individual game management units (GMUs) the areas within which elk are managed. The Bender and Spencer reconstruction technique allows us to derive estimates of elk populations for individual GMUs by year.

The second task involved the use of satellite data to characterize vegetation responses to seasonal and interannual climate variation among vegetation associations within GMUs. This involved the use of NOAA Advanced Very High Resolution Radiometer (AVHRR) time series data to describe temporal vegetation behavior, and Landsat data to describe spatial vegetation distribution in conjunction with U.S. Forest Service vegetation maps. By their very nature, maps created through traditional, field-based mapping techniques present a static picture of the environment. A vegetation map, for instance, shows the type and location of vegetation at the time of data collection. A single map cannot describe vegetation condition, its response to climatic events, or its seasonal phenology. Our intent was to enhance the information content of the static vegetation map by incorporating information on landscape dynamics derived from satellite remote sensing data to produce a multi-dimensional characterization of the landscape.

The final task involved using statistical methods to evaluate relationships between elk population estimates and satellite-derived greenness measures through time. For this task, satellite greenness measures were extracted for each GMU and associated with vegetation associations.

STUDY AREA

The study area is the group of Arizona GMUs shown in Figure 2. These comprise a total of 22,613 km^2, or 7.6% of Arizona's total area. They are a subset of the 79 total GMUs for which the AGFD manages wildlife. The data required to apply Bender and Spencer's reconstruction model can be gleaned from the associated AGFD records. We queried the AGFD survey and harvest databases from 1988 to 1999 and determined that elk were surveyed or hunted in 32 GMUs. Data for elk on Tribal lands were not available. We merged data for units in cases where a population crossed unit boundaries; where survey and harvest data could be combined; and where the AGFD pooled data for their model (e.g., units 5A, 5B, and 6A; 6B and 8). We identified 5 GMUs or combinations of units that had adequate data (i.e., congruous hunt and survey data for ≥ 5 consecutive years between 1988 and 1999) to compare their derived population estimates with landscape temporal dynamics. Elk population estimates were ultimately reconstructed for the following sets of Arizona GMUs covering the years indicated (see Figure 2):

NDVI values range from –1 to 1, with non-land surfaces (such as water and snow) typically having negative values and land surfaces typically having positive values. As landscapes become more densely vegetated, the NDVI trends to 1. The NDVI represents a measure of vegetation photosynthetic activity, and is sensitive to various biophysical vegetation characteristics, such as biomass and percent cover (Huete and Jackson, 1987). To produce a relatively cloud-free image representing actual ground conditions, the maximum value of the daily NDVI images is composited over each 2-week period.

These NDVI composite data sets are produced by the U.S. Geological Survey (USGS) (EDC, 1994) and are available for the state of Arizona through the ARIA website (http://aria.arizona.edu/). To construct this data set, the daily AVHRR channel data are radiometrically calibrated, corrected for variation in solar illumination angles, and geometrically registered, as described in Eidenshink (1992).

(2) *Existing detailed vegetation mapping, fine-scale topography, and terrestrial ecosystem units* for the Coconino National Forest, which lies within the study area. The U.S. Forest Service (USFS) Terrestrial Ecosystem Survey (TES) of the Coconino National Forest identifies 134 mapping units. The initial classifications of the units were based upon delineation of topography, geology, and existing vegetation using stereo pairs of 1:24000-scale aerial photographs. Fieldwork was then conducted to assess the accuracy of these initial units and to collect additional terrain data. The vegetation classification employed in developing the TES units is hierarchical, with broadest categories applicable over larger spatial extents. The classification hierarchy, in descending order, considers vegetation structure, regional phenology, and physiology of dominant species (Miller et al., 1995). For this study, we merged the TES units, using the defined hierarchy, into broad cover classes. The resulting vegetation map covered a portion of the study area (Figure 4).

(3) *Landsat-7 Enhanced Thematic Mapper (ETM+) data.* The high spatial resolution (30-m) Landsat data were used to extrapolate the available TES vegetation mapping into the entire study area. Four Landsat-7 ETM+ scenes were needed to provide complete coverage of the study area: path/row 37/36, 37/35, 36/37, and 36/36. All the scenes were collected during October 2000.

METHODS

Modeling Elk Populations by Reconstruction:

Elk population estimates in central Arizona were calculated using the method described by Bender and Spencer (1999). This model has been used to obtain estimates of elk populations in Washington and Michigan and are comparable to other more labor-intensive and popular estimation techniques, such as mark-resight, total count, and aerial sightability. The reconstruction estimates are derived from harvest data and herd ratios

(bull/cow, calf/cow), and incorporate a measure of annual non-harvest related bull mortality. This approach to estimation was first described by Allen (1942) who reconstructed population size from sex ratios, harvest number, and harvest percentage. Harvest numbers and herd ratios were obtained directly from the AGFD records. An estimate of annual bull mortality was calculated from the percentage of yearling bulls in AGFD surveys, corrected for population trends using the formula (Bender and Spencer, 1999):

$$M_B = 1-[(1-Y_B)*\lambda] \quad (2)$$

Where:
- M_B is the total adult bull mortality rate (fraction per year)
- Y_B is fractional percentage of yearling bulls in the total bull population
- λ is the finite rate of population increase, as estimated below

If the population is age stable and stationary, then $\lambda = 1$ and the yearling bulls effectively replace the bulls that die. To calculate an estimate of the actual λ, we considered the percent population change of males estimated by AGFD models (provided in their survey data) to be equivalent to the instantaneous rate of increase (r). We then calculate λ as follows:

$$\lambda = e^r \quad (3)$$

The population reconstruction is then calculated as follows (Bender and Spencer 1999):

$$N_T = N_B + N_C + N_J \quad (4)$$

Where:
- N_T is the total number of elk
- N_B is the number of bulls
- N_C is the number of cows
- N_J is the number of juveniles

Substituting the following:

$$N_B = K/M_B \quad (5)$$
$$N_C = N_B * R_{C/B} \quad (6)$$
$$N_J = N_C * R_{C/C} \quad (7)$$

Where:
- K is the number of males harvested
- M_B is the mean annual total male mortality rate (calculated above),
- $R_{C/B}$ is the ratio of cows to bulls in pre-hunt surveys, and
- $R_{C/C}$ is the ratio of calves to cows in pre-hunt surveys

Resulting in the following reconstruction equation:

$$N_T = (K/M_B)*(1+R_{C/B}+R_{C/B}*R_{C/C}) \quad (8)$$

We could not calculate confidence intervals because measures of variance were unavailable for some values. Male harvest rates for a staggered-entry Kaplan-Meier estimator were not available, and harvest-per-unit-effort did not have a linear relationship with elk numbers, precluding the use of alternative methods for calculating the mortality rates used in the Bender and Spencer model (1999).

Modeling Landscape Condition and Dynamics:
The following procedures were followed to develop the vegetation map for the study area and extract multitemporal NDVI indices:

(1) *Landsat-7 ETM+ Pre-processing*

Each Landsat-7 ETM+ image was first orthorectified using the corresponding Digital Elevation Model (DEM) and Landsat scene from the 1992 USGS Multi-Resolution Land Characteristics (MRLC) data set as reference images. The MRLC data set consists of images that have been precision corrected for terrain and relief displacement; the images meet quality restrictions of total root mean square errors (RMSEs) of less than 1 pixel (30m) (http://edc.usgs.gov/glis/hyper/guide/mrlc). The four Landsat-7 ETM+ images were orthorectified in ERDAS Imagine, using a minimum of 20 ground control points and a second order polynomial fit. In all cases, final RMSE values were less than 0.5 pixels. Band 8 of each image was rectified separately because of its finer spatial resolution (15-m).

Each image was atmospherically and radiometrically corrected following the methods of Chavez (1996) using a dark-object atmospheric correction algorithm. The model converts the raw digital numbers (DNs) recorded at the satellite to reflectance values by adjusting for sensor calibration factors, sun-earth distance, and solar elevation angle at the time of image acquisition. The conversion is necessary to remove the atmospheric effects in the reflectance data collected at the satellite. Once the data are normalized, the actual ground reflectance values can be compared between Landsat scenes. The orthorectified and atmospherically corrected scenes were assembled into one mosaic comprising all four images.

(2) C*reation of the Vegetation Cover Map*

The Landsat-7 data in combination with various derived data layers was used to create a vegetation map of the study area. Four Landsat-7 ETM+ scenes were needed to generate complete coverage of the study area. The production of a vegetation map involved multiple processing and analytical steps, as described below.

A number of derived and existing images were first evaluated for their utility as input layers in the landscape classification. All images were re-scaled to an 8-bit (0-255) data

range to facilitate data processing. These data sets included existing data, comprised of atmospherically and radiometrically corrected Landsat-7 data, and associated DEM data. Products derived from Landsat-7 data included NDVI, Principal Components, and a texture image created using a 3x3 pixel variance convolution-filter on the 15-m panchromatic band.

We identified the most effective combination of input layers by running an unsupervised classification on various combinations of existing and derived images. Each resulting classification was evaluated by computing a cross-tabulation with the U.S. Forest Service TES data. The ideal classification was one with two characteristics: (1) distinct separability between the classes, i.e., the amount of confusion among class assignments was minimal; and (2) high agreement with the merged TES data in the number and relative distribution of the classified pixels. The combination of layers that yielded the greatest accuracy was composed of the 6 principal component images, the NDVI image, the texture image, and the DEM. This combination was used as the basis for an unsupervised classification with 20 classes. The classification was then compared to the merged TES vegetation map.

We followed standard classification techniques (Jensen, 1996) to produce a final classification map. Any pixel class whose distribution matched that of one of the TES categories was considered to be correctly classified. Classes were merged together if they appeared to reflect acceptable, within-category variations. Any class whose pixels were distributed among two or more TES categories was considered to be incorrectly classified. These classes were extracted from the image, and an unsupervised classification was re-run dividing them into 10 new classes. The distributions of the 10 new classes were examined for their agreement with the TES categories, and were assigned accordingly.

(3) *Characterization of Landscape Temporal Dynamics*

Available bi-weekly composites for each year were combined into representative monthly images by averaging one, two or three composites according to the days of the year spanned by the image. The 2-week composite was assigned to the month that contained the majority of the 14 days. This was necessary because the USGS product initially emphasized the growing season and commonly calculated only one image in January, February, November and December for the earlier years. Creation of monthly images allows the data for each month to have equal weight in the analysis. The 12 monthly images for each year were layer-stacked to create a single multi-layer image for each year.

Because of problems with the AVHRR sensor, the last quarter of 1994 is missing from the USGS database. Composites for this year were estimated by averaging the corresponding images for the years 1990 and 1997. These two years were selected by comparing images from the first three-quarters of all years to 1994. The two years chosen

are visually very similar to those from 1994. In addition, the ENSO conditions leading up to the last quarter of 1994 are also most similar to the ENSO conditions during 1990 and 1997 (see NOAA climate prediction web site: http://www.cpc.ncep.noaa.gov/products/analysis_monitoring/ensostuff/ensoyears.html). Because vegetation conditions in Arizona are heavily influenced by ENSO conditions, we considered it judicious to estimate the missing quarter of 1994 using the two most similar years rather than a simple average of all available years.

Various measures that capture different aspects of the temporal dynamics of the landscape were derived from the AVHRR NDVI data. These measures were produced using three main classes of calculations: basic statistics, standardized principal components analysis (PCA), and Fourier analysis.

Basic statistics were calculated using ERDAS Imagine image processing software. The monthly values at each pixel are treated as a set of 12 data values and the following summary measures were derived:

1. Average of the 12 monthly images
2. Standard deviation of the 12 monthly images
3. Coefficient of variation of the 12 monthly images
4. Average NDVI for Quarter 1 (January, February, March)
5. Average NDVI for Quarter 2 (April, May, June)
6. Average NDVI for Quarter 3 (July, August, September), and
7. Average NDVI for Quarter 4 (October, November, December).

Standardized Principal Components Analysis (PCA), also calculated using ERDAS Imagine, treat the NDVI data as a set of 12 stacked images. PCA is a data compression technique that projects the information content (i.e., the variability) of a multi-dimensional dataset onto orthogonal axes, called components (Jensen, 1996). Several studies apply PCA to long time series of AVHRR bi-weekly NDVI composites (e.g., Tucker, 1985; Eastman and Fulk, 1993; Hirosawa et al., 1996). For these data, it is known that the first principal component (PC1) represents the characteristic NDVI, and can be shown to be analogous to the average greenness for the time period analyzed. The second principal component (PC2) represents "seasonality," and the extremes of PC2 correspond to landscapes with summer green-up and those with winter green-up (Hirosawa et al., 1996).

Using Fourier analysis, the ordered NDVI values at each pixel are input to the analysis as a vector or waveform. Additional measures of the landscape temporal dynamics are derived that describe the shape of the waveform (Jakubauskas and Legates, 2001; Moody and Johnson, 2001). The analysis was accomplished using Matlab, a software package for mathematical computation and analysis (www.mathworks.com). The image data were imported as matrices into Matlab and the following indices were produced:

1. Additive term, which is the best fit of a flat line to the NDVI profile (i.e., the 0th frequency component);

2. Magnitude of the first frequency component, which is the amplitude of the single sine wave that best fits the annual profile; and
3. Phase of the first frequency component, which is the position of the peak (between -pi and +pi) of the single sine wave that best fits the annual profile. The phase data were re-scaled to represent the day of year (0 to 365) of the sine wave peak by taking into consideration the exact dates represented by the first and last biweekly composites.

Fourier indices have proven useful for vegetation studies, including mapping vegetation communities (Moody and Johnson, 2001) and monitoring regional to global vegetation patterns (Jakubauskas and Legates, 2001).

(4) *Linking the Temporal and Spatial Data Sets*

It was first necessary to link the vegetation cover map derived from the Landsat-7 data to the multi-temporal information on vegetation dynamics derived from the AVHRR data in order to analyze the phenology of each vegetation association and its relative contribution to the greenness of the AVHRR pixel. To do this, we determined the constituent vegetation associations in each AVHRR pixel and categorized it according to the dominant vegetation type. Therefore, prior to connecting the data sets, two preliminary steps were performed.

First, as explained above, various measures that capture the temporal dynamics of each AVHRR pixel were calculated for the 11-year period (1989-99): greenness averages for the four quarters (January- March, April-June, July-September, October-December); annual basic statistics (average, standard deviation, and coefficient of variation); the first three components from the PCA; and four Fourier indices (additive component, phase and magnitude of the first frequency component, and the phase, re-scaled to the day of year). Secondly, the areal proportion of each classified vegetation type was calculated within each AVHRR pixel. This processing was accomplished using ArcView and ArcInfo GIS software by overlaying the grid of vegetation classes with the grid of AVHRR pixel locations.

After the preliminary steps were completed, the maps of AVHRR pixels and vegetation maps were joined by evaluating the AVHRR pixels on the basis of their vegetation association. If a pixel contained more than 20% riparian pixels, the entire pixel was considered "riparian" and extracted from the coverage for further analysis. The limit of 20% was chosen empirically: it was the smallest percentage that yielded an acceptable number of AVHRR pixels for analysis. The linear characteristic of riparian habitat meant that it could rarely dominate the 1-km AVHRR pixel spatially, however, the abundance of vegetation in these habitats will dominate the spectral reflectance of the AVHRR pixels. We therefore considered a smaller areal proportion of riparian habitat sufficient to classify the entire AHVRR pixel as riparian. The other AVHRR pixels were classified using a limit of 70%. If a pixel did not have a single vegetation class of that proportion,

shrub and grassland based on the regional Lowe and Brown vegetation map for Arizona (Lowe and Brown, 1973).

Relationships between Annual Elk Population and Satellite Greenness Measures: Using linear regression analyses, we evaluated the relationships between elk population estimates and the satellite-derived landscape greenness measures extracted for each vegetation type within each GMU. Although we applied the analysis to both the Bender and Spencer reconstruction estimates and the AGFD estimates, we are reporting here only the results using the Bender and Spencer population values. The Bender and Spencer estimates produced stronger correlations between the satellite greenness measures than the AGFD estimates. In addition, some stakeholders questioned the reliability of the AGFD data and our intent was to use information considered least controversial by the widest audience.

Initial results revealed Bender and Spencer reconstruction estimates for 1999 diverged dramatically from the patterns of the preceding decade. In many of the GMUs, the 1999 elk estimates were substantially lower than the trend established during the previous years would suggest. In addition, for many of the satellite greenness measures calculated, the elk population visually tracks the greenness measures for all the years except 1999, in which low population estimates diverge from high satellite greenness. These observations suggested the data input to the reconstruction model for 1999 might be suspect. Although the reason for the anomalous behavior of the 1999 estimates is unknown, these population estimates were excluded from further analyses.

The statistical tests for the remaining Bender and Spencer reconstruction estimates and the greenness measures revealed several statistically significant correlations at significance of $\alpha \geq 0.10$ (Table 2). There were moderate to high correlations in unit 9 ($r^2 = 0.53 - 0.91$) and unit 22 ($r^2 = 0.40 - 0.85$), and weak correlations ($r^2 = 0.42 - 0.57$) for the unit grouping 5A, 5B, and 6A in quarters 1 and 2. We did not find significant correlations between elk numbers and vegetation associations in game management units 7A and 7B combined; and the group comprised of 6B, 8, and the Navajo Site. Note that no significant correlations were identified between elk population and landscape condition when the entire set of GMUs were merged (Table 2, bottom). This result highlights the significance of local differences in the landscape dynamics and the importance of analyzing smaller regions that are coherent on a landscape scale.

The strongest correlations resulted from regression of the elk estimates against the simplest measure of greenness for each quarter. Statistically significant correlations were also obtained from regressions with higher, more complex derived measures, such as principal components. Although these correlations are intriguing and warrant further investigation, however, because the connection of the elk estimates with the simplest greenness measures returned the highest correlation results and are also the most readily interpreted, they are the only ones discussed here.

Landscape Temporal Dynamics and Elk Population:
The regression results suggest a correlation between the seasonal greenness of certain vegetation classes and the annual elk population numbers. Although this correlation is not verified by ground-based field data, the fact that it makes sense biologically suggests the relationship may be valid and not simply a specious result. However, it is important to note that numerous factors could account for variations revealed in our results. For example, elk distribution is not likely uniform within GMUs, the boundaries of the GMUs varied over the study period, and the AVHRR data used in this analysis were not water vapor corrected. In fact, inspection of the AVHRR data reveals the effect of persistent cloud cover during the summer rainy (monsoon) season of several years, producing noisy profiles during the monsoon months (July, August and September). It is likely that this noise accounts for the fact that no significant correlations were found for quarter 3 of any year. After the AVHRR data were processed for this study, the USGS recently published water vapor corrected versions of these data sets. The results obtained in this study, therefore, represent a "worst case" scenario with respect to these data.

The observed correlation between seasonal greenness of certain vegetation classes and the annual elk population numbers suggests the elk population is responding to environmental change. The ability to be responsive is further indicative of a population that is reasonably healthy and viable. If correlations between a population and the satellite-derived measures of habitat condition are established, these relationships can be used to help monitor the population. Clearly, if a population is compromised to the point of near extinction, or if factors other than the habitat condition - such as a devastating disease - come into play, then population numbers would be expected to deviate from habitat condition. Such deviations from established relationships could signal areas of concern. In our analysis, we excluded data for 1999 because they obviously deviated from the patterns established by the preceding years. The cause of this deviation should be explored to determine whether it reflects an issue with the input data or reflects an actual drop in elk population due to an external effect.

The stronger correlations noted for GMUs 9 and 22 could reflect environmental or management conditions that are favorable for this analysis. Environmentally, these two GMUs are dominated by desert scrub and pinyon-juniper vegetation classes, both of which have widely-spaced canopy. The satellite measures are, therefore, preferentially capturing information on the ground cover (important as elk forage) when compared to the higher-elevation GMUs with abundant closed canopy forest vegetation. Managerially, these two GMUs have been handled consistently as separate entities. This is in contrast to the other GMUs, which have been handled separately at times and combined with other units at times. It is possible that the more complicated management history makes these data less useable for this analysis. Preliminary results for this study, which considered only GMU 6A, suggest this may be a factor (Hutchinson, et al., 2000). Prior to 1996, GMU 6A was managed separately and after 1996 it was combined with units 5A and 5B. A preliminary study found stronger correlations with vegetation greenness and elk

deviation below the global mean. In the Figures 5 and 6, these pixels are shown as dark gray, regardless of vegetation type.

The above average greenness pixels within each vegetation type could be greener than average for many reasons, including:

1. Vegetation is denser than average for that vegetation class.
2. Vegetation is healthier than average for that vegetation class.
3. There is locally more abundant accessory vegetation, such as annuals and ephemeral species, within the vegetation class.
4. The analysis includes mixed pixels that represent 2 or more vegetation associations, such as pixels dominated by desert scrub but with a high percentage of Pinyon-Juniper. Such conditions explain linear zones of above average vegetation at the contacts between distinct vegetation classes, and do not represent conditions expected to correlate with higher elk population estimates.

We hypothesize that these maps show areas where elk are more likely to be found. The regression results suggest that elk populations are larger during years that landscape conditions are greener. Although these results refer to temporal changes in greenness, we have created these maps to translate that landscape preference spatially. It is reasonable that elk would prefer areas with healthier vegetation and more abundant annual forage. Furthermore, if the denser vegetation areas include denser forage species and not simply denser tree canopy, elk could prefer these as well.

The hypothesis that elk prefer areas that, over time, tend to be of above average greenness was tested using a database of elk locations supplied by the AGFD. These locations are from radio-collared animals that were monitored between 1981 and 1993 (Brown, 1994). Figures 5 and 6 overlay the elk locations on the above average greenness maps and reveal a strong visual correspondence of the elk with areas of above average greenness. These results clearly demonstrate the potential and practical viability of the satellite-based model.

CONCLUSIONS

This research demonstrates the potential use of NASA science results and satellite data for natural resource management. Wildlife management, particularly of elk, is a high priority for the State of Arizona. Since the reintroduction of elk to their historic habitats of the western US in the early 1900's, the elk population has grown dramatically. This growth has led to escalating conflicts between elk advocates and those ranchers and farmers who feel their operations are threatened by the competition of elk for range resources, and the damage they cause to infrastructure (e.g., fences, water developments) and agricultural crops.

The issue is especially contentious because of the public aspect of the management process. Elk populations are managed by public agencies, primarily through the issuance of hunting permits. Moreover, a large part of land on which conflicts occur is publicly owned (i.e., U.S. Forest Service or State Lands) and typically leased to private ranchers. As a consequence, public involvement in management is vital and, for some agencies, legally mandated. Much of the current debate revolves around the credibility and interpretation of information that is brought to bear on the problem. Most specifically, the public wants to know: How many elk are there? Where and when are they a problem for ranchers? How can their management be more effectively pursued?

The project described in this paper produced a number of applied research results that will help to answer some of these important natural resource management questions. The results include:

- Implementation of a new model (Bender and Spencer, 1999) to estimate elk population applied to five game management unit groups in Arizona. The new model suggested greater annual fluctuations in elk population than the AGFD models did. In addition, these data provide a new set of estimates to help wildlife managers better understand the variability of both modeling techniques and elk population. This activity was instrumental in identifying weaknesses that could be corrected in both methods.

- Statistical tests of the correlation between simple greenness measures derived from the NOAA-AVHRR time-series data and elk population estimates derived from the new model revealed moderate to high correlations within three of the five game management unit groups. The highest correlations were between elk population numbers and greenness during specific seasons of the year. The regression results suggest a correlation between the seasonal greenness of certain vegetation classes and the annual elk population numbers. However, these relationships are indeed complex both spatially and temporally and require further analysis before definitive conclusions can be drawn.

- Based on the apparent seasonal association, maps of areas of higher than average greenness were constructed for the three game management unit groups where they were strongest. These maps have the potential of being utilized by wildlife managers for predicting elk locations.

Although it is not surprising that animals appear to respond to differences in vegetation availability – both seasonally and inter-annually – it is significant that these differences can be captured by satellite data. The synoptic view of vegetation availability provided by these satellite data can inform resource management efforts. The ability to characterize the temporal dynamics of vegetation using repeatable remote sensing data allows closer monitoring of the relationship between a species and its landscape.

The methods we develop are directly applicable to the Moderate Resolution Imaging Spectroradiometer (MODIS) data. MODIS data are also collected daily, but the imagery has a much finer spatial resolution (250 m^2). The MODIS deliverables will include an Enhanced Vegetation Index (EVI) that is analogous to the NDVI but incorporates corrections for view-angle, soil background, and atmosphere.

Preliminary results of this research have been distributed within the state by the ASLD and response has been positive. The elk issue touches on a number of different constituencies and we hope that results of this study can shed light on the larger management issue, as well as provide routine information that can help resolve it. An unanticipated and valuable result of this research is that it has fostered significant new interaction with a wide range of end-users from the natural resource management community.

ACKNOWLEDGEMENTS

The authors wish to acknowledge the contributions of our project participants at the University of Arizona included faculty and staff from the Arizona Remote Sensing Center (Eric Pfirman, Jessica Walker, and Kathryn Mauz), the School of Renewable Natural Resources (Larry Howery and Hilary Boyd), the Department of Animal Sciences (Mark Enns and Heather Salazar), and the College of Agriculture's NASA Space Grant Extension Specialist (Barron Orr). The manuscript benefited greatly by the exemplary editing of Jessica Walker.

We also wish to thank Jim DeVoss and Susan Boe at the Arizona Game and Fish Department for providing the elk location data. In addition, we wish to acknowledge our co-Principal Investigator from the State of Arizona, Gene Trobia (Arizona State Cartographer). Partnerships at the State and Federal level were also established the Arizona Geographic Information Council and the U.S Forest Service.

REFERENCES

Allen, D. L. 1942. A pheasant inventory method based on kill records and sex ratios. Transactions of the North American Wildlife Conference 7:329-332.

Bender, L. C., and R. D. Spencer. 1999. Estimating elk population size by reconstruction from harvest data and herd ratios. Wildlife Society Bulletin 27:636-645.

Brown, R.L. 1994. Elk seasonal ranges and migrations in Arizona. Arizona Game and Fish Department Technical Report 15.

Burkhart, B., 1997. Feedback sought on what to do about elk numbers. *The Arizona Republic*, Phoenix, AZ, February 2, 1997.

Burkhart, B., 1995. Hunters under gun to pay for forage eaten by elk. *The Arizona Republic*, Phoenix, AZ, December 17, 1995.

Chavez, P. S., jr. 1996. Image-Based Atmospheric Corrections – Revisited and Improved. *Photogrammetric Engineering and Remote Sensing* 62 (9): 1025-1036.

Davis, J.C. 1986. Statistics and Data Analysis in Geology, Second Edition. John Wiley and Sons, New York, 646p.

Eastman, J.R. and M. Fulk 1993. Long Sequence Time Series Evaluation Using Standardized Principal Components. *Photogrammetric Engineering and Remote Sensing*. 59(6): 991-996.

EDC, 1994. *Conterminous United States AVHRR Data on CD ROM - User's Manual*, U.S. Department of Interior, Eros Data Center, Sioux Falls, SD.

The Economist, 1998. Perspective on 'Natural Regulation' – Elk eat away at Yellowstone Range, Policy. Reprinted in: *The Arizona Republic*, Phoenix, AZ, January 7, 1998.

Eidenshink, J. C. 1992. The 1990 Conterminous U. S. AVHRR Data Set. *Photogrammetric Engineering and Remote Sensing* 58 (6): 809-813.

Foust, T. 2001. High elk numbers to be addressed. *Arizona Daily Star*, Tucson, AZ, April 12, 2001

Harris, R., 1995. The biopolitics of Arizona's Elk – are there too many big-game animals? *The Arizona Republic*, Phoenix, AZ, October 1, 1995.

Holland, L., and R. Fink. 1997. The elk population simulation model. Arizona Game and Fish Department, Phoenix, Arizona, USA.

Huete, A.R., and Jackson, R.D., 1987. Suitability of Spectral Indices for Evaluating Vegetation Characteristics on Arid Rangelands. *Remote Sensing of Environment.* 23:213-232

Hutchinson, C.F., S.E. Marsh, C.S.A. Wallace, J.J. Walker, K. Mauz, P.R. Krausman, H. Boyd, R.M. Enns, H. Salazar, L.D. Howery, E. Trobia 2000. *Informing the elk debate: applying NASA earth observing system (EOS) data to natural resource management conflicts in western states,* Arizona Remote Sensing Center Final Report Submitted to Raytheon Systems Company, University of Arizona, Tucson, Arizona.

Jensen, J.R. 1996. *Introductory Digital Image Processing: A Remote Sensing Perspective, 2nd edition.* New Jersey: Prentice-Hall, Inc. 316 p.

Krausman, P.R., ed., 1996. *Rangeland Wildlife.* Society of Range Management. Denver, CO.

Lowe, C.H. and D.E. Brown, 1973. *The natural vegetation of Arizona.* (Arizona Resources Information Systems Cooperative Publication 2). , Phoenix, Arizona. 53 p.

Marsh, S.E., C.S.A. Wallace, and J.J. Walker, 1999. *Evaluation of Satellite Remote Sensing Methods for Regional Inventory and Mapping of Desert Resources,* Arizona Remote Sensing Center Final Report Submitted to the National Park Service, University of Arizona, Tucson, Arizona.

Miller, G., N. Ambos, P. Boness, D. Reyher, G. Robertson, K. Scalzone, R. Steinke, and T. Subirge, 1995. *Terrestrial Ecosystem Survey of the Coconino National Forest,* USDA Forest Service Southwestern Region.

Noyes, F. 1995. Legislator seeks cuts in elk herd. *Arizona Daily Star,* Tucson, AZ, September 3, 1995

Rawson, W.F. 1995. Ranchers seeking state aid in reducing elk population. *Arizona Daily Star,* Tucson, AZ, August 6, 1995

Risenhoover, K.L. and Bailey, J.A., 1980. Visibility: An important habitat factor for an indigenous, low-elevation Big Horn herd in Colorado. Proceedings of the Biennial Symposium of the Northern Wild Sheep and Goat Council, pp.18-28.

Smith, D., 1998. Elk hunting interest hits all-time high. *The Arizona Republic*, Phoenix, AZ, August 22, 1998.

Tucker, C.J., J.R.G. Townshend and T.E. Goff 1985. African Land-Cover Classification Using Satellite Data. *Science*. 227(4685): 369-375.

Wisdom, M.J. and Thomas, J.W., 1996. *Elk*, in Krausman, P.R., ed., *Rangeland Wildlife*. Society of Range Management. Denver, CO.

Vegetation Class	TES Unit	Producer's Accuracy	User's Accuracy
Riparian and Associated	<100	31.3	49.0
Sonoran Desert Scrub	200+300	89.2	61.4
Pinyon-Juniper Woodland	400	87.6	80.8
Ponderosa Pine - Oak	500	80.3	90.4
Fir, Douglas Fir, Ponderosa	600	58.6	51.5
Spruce, Fir, Douglas Fir	700	78.9	88.4
Alpine Tundra	800	57.5	60.8
Overall Accuracy: 0.81		Kappa: 0.70	

Table 1. Results of the classification accuracy assessment for the merged TES Units.

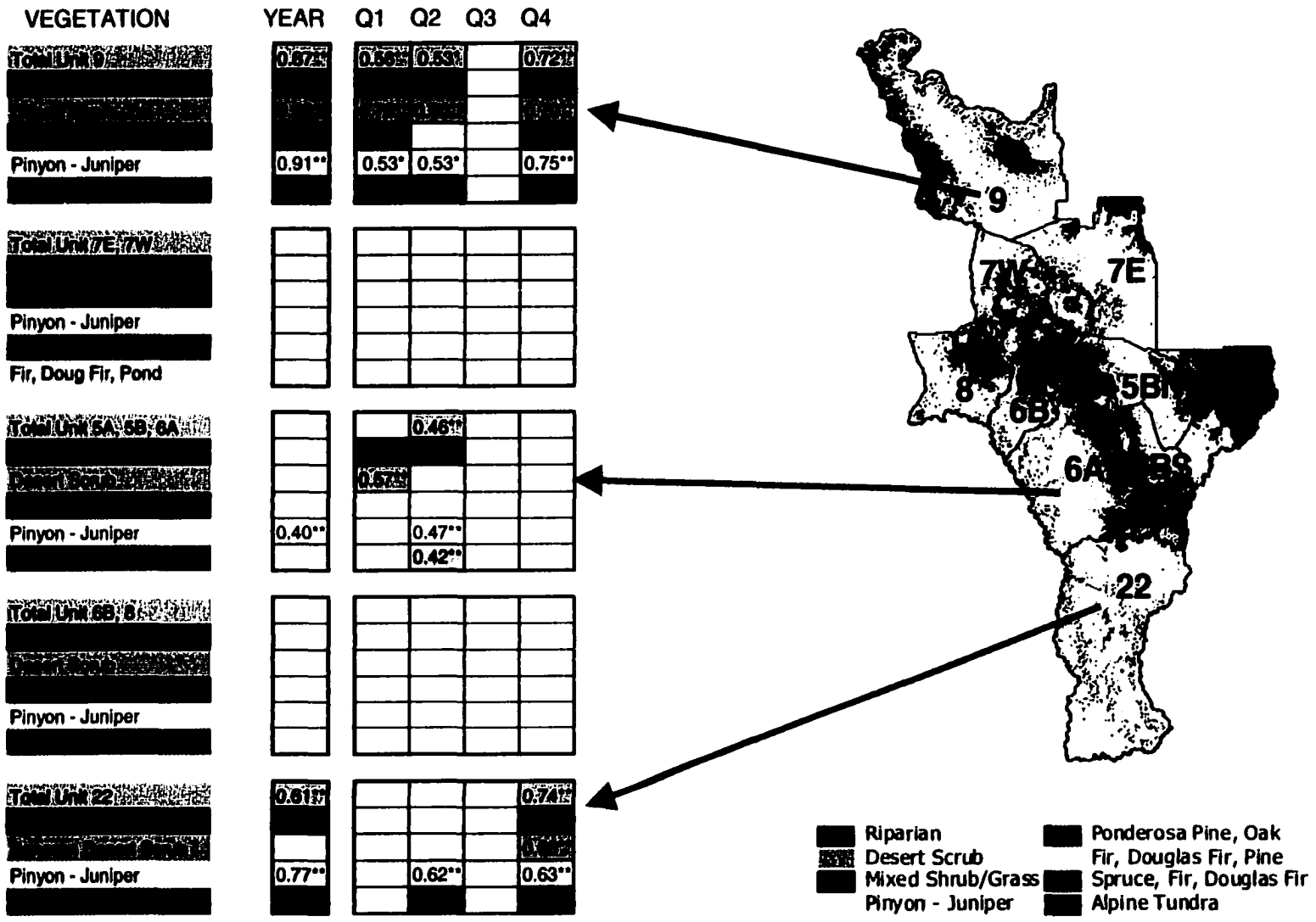


Table 2: Correlations between average greenness measures and the elk population model estimates following Bender and Spencer (1999). The Year column represents the entire year and the Q1-Q4 columns each quarter of the year. A blank cell means there is no significant correlation. (*$\alpha = 0.10\%$ significance level, **$\alpha = 0.05\%$ significance level).

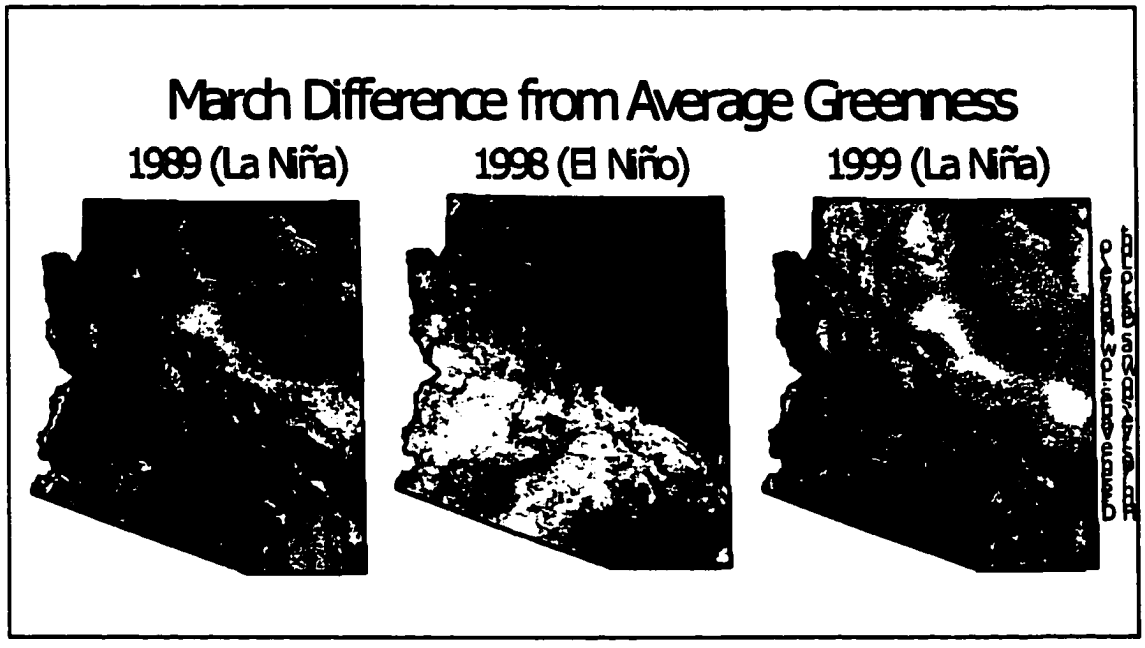


Figure 1. Impact of La Niña and El Niño climatic events on vegetation in Arizona. In these figures, relative "greenness" (or vegetation vigor) is shown in shades of gray, with low to high relative "greenness" depicted as black to white. La Niña winters are dry in Arizona, producing the pattern of relative "greenness" across the state shown in 1989 and 1999. Dark tones in southwest Arizona in the 1989 image (far left) indicate lower amounts of actively growing green vegetation in the desert communities. Light tones along the Mogollon Rim in this same image represent coniferous vegetation, visible due to lack of snow cover. In contrast, El Niño winters in Arizona are unusually wet, producing the pattern of relative "greenness" shown in 1998. The desert regions of southwest Arizona are bright, indicating greater amounts of actively growing green vegetation. Dark tones along the Mogollon Rim represent snow cover, which obscures the satellite signal of the coniferous vegetation.

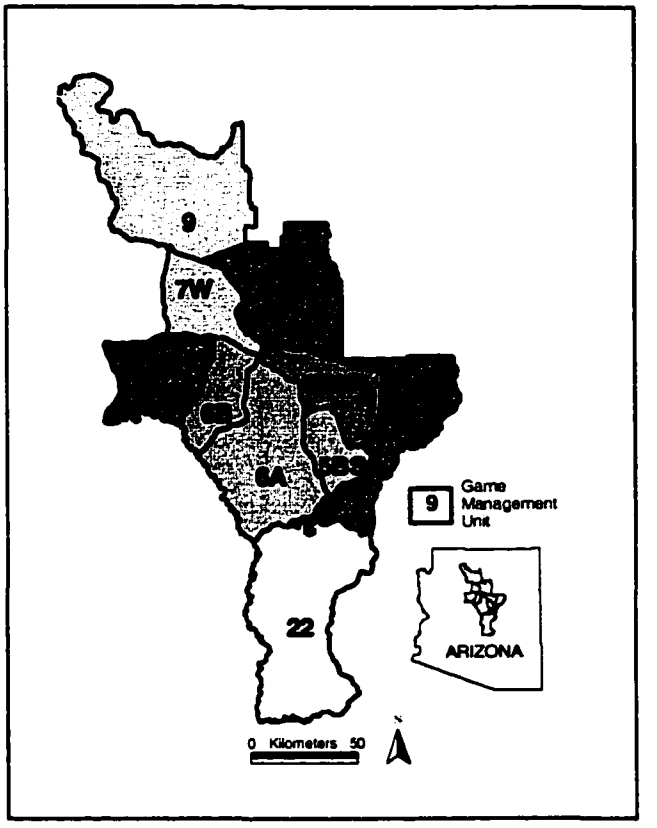


Figure 2: The Arizona Game Management Units in central Arizona selected for this study.

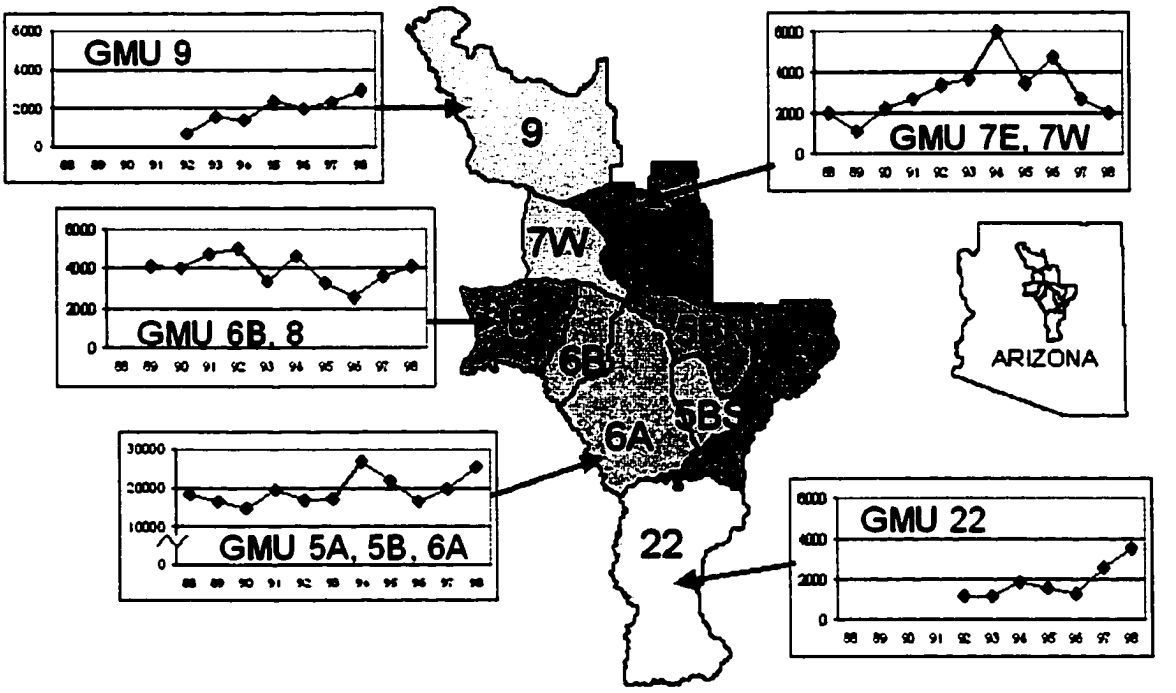


Figure 3: Elk population estimates derived using the Bender and Spencer reconstruction method for selected Arizona Game Management Units in central Arizona, 1988 – 1998.

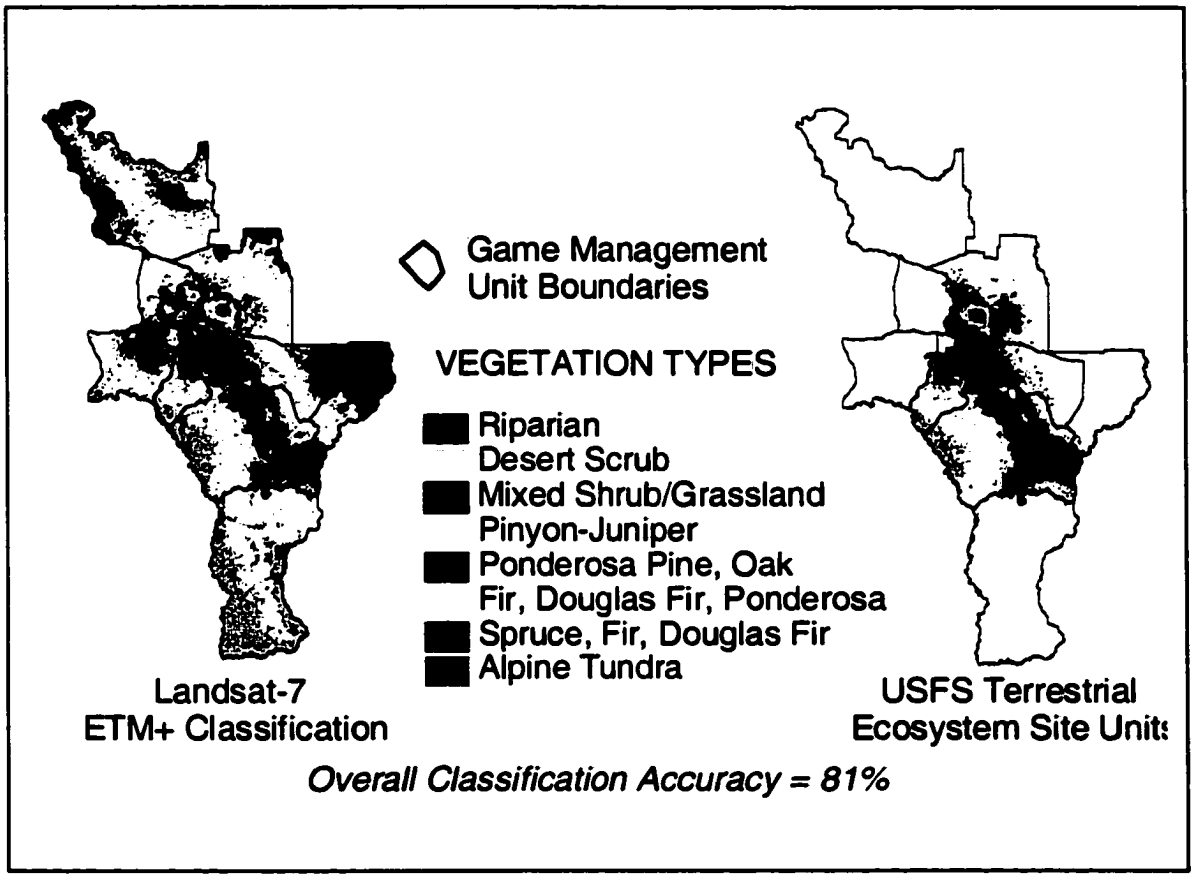


Figure 4. Vegetation classification results (left) compared to the U.S. Forest Service TES map (right).

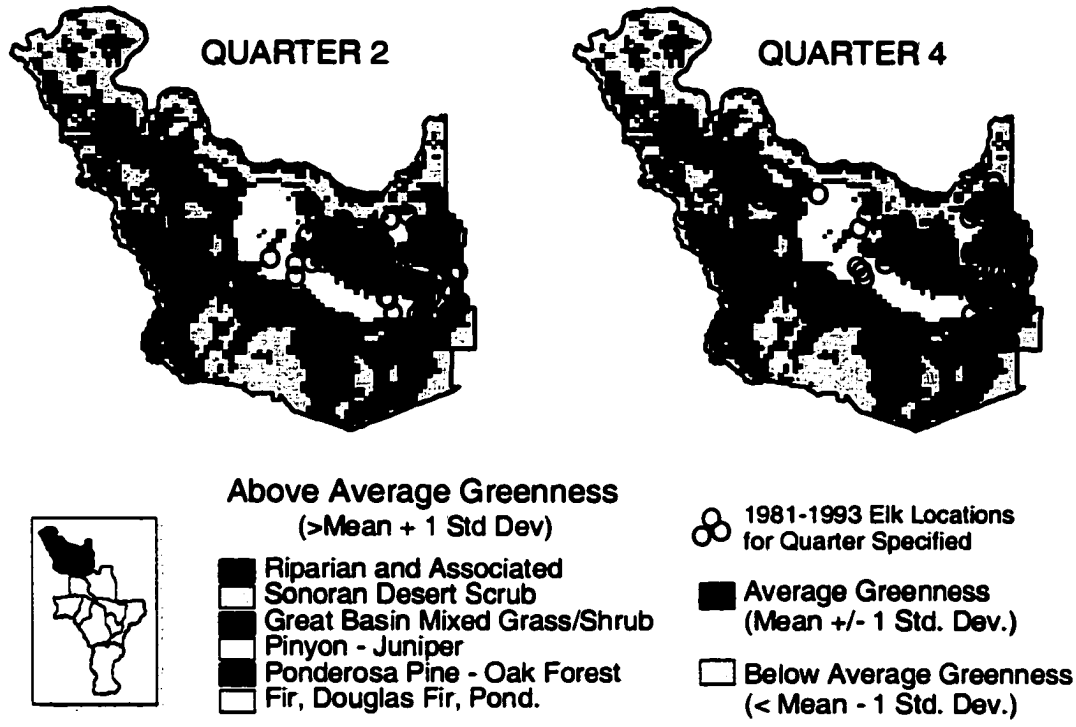


Figure 5. Greenness model results for game management unit 9 with AGFD elk location data.

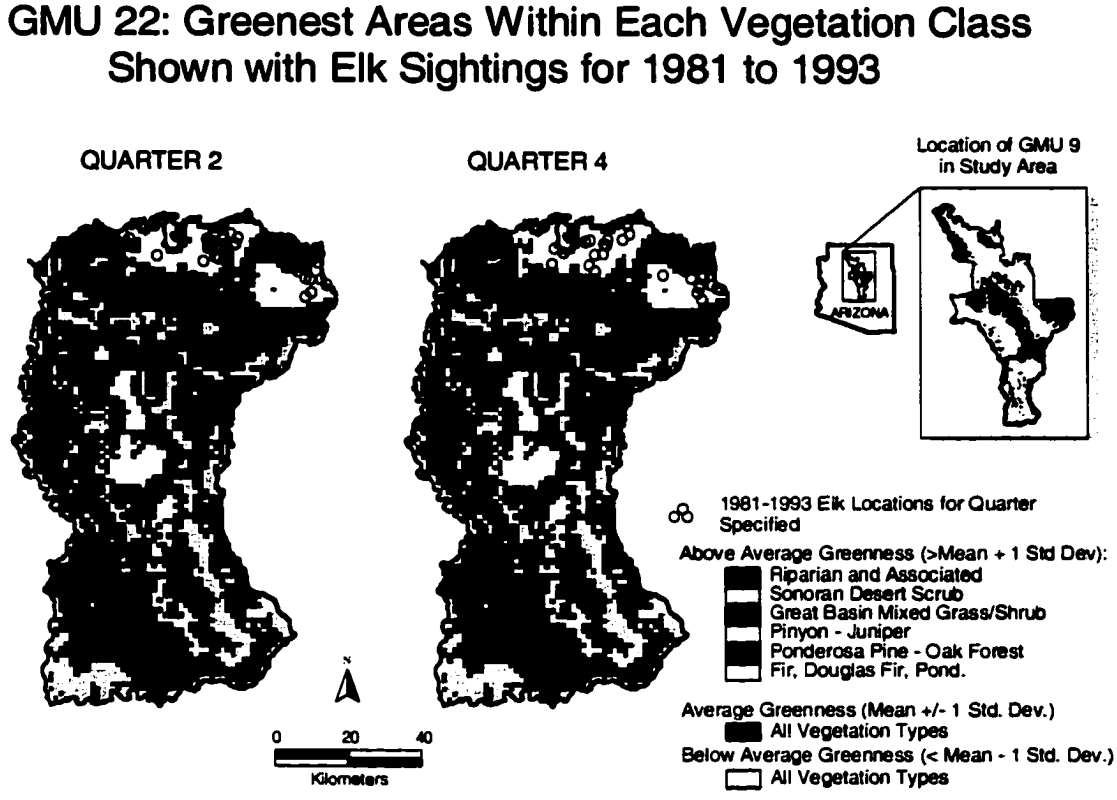


Figure 6. Greenness model results for game management unit 22 with AGFD elk location data.

APPENDIX D

SATELLITE IMAGE ANALYSIS AND HABITAT MODELING FOR A YELLOW-BILLED CUCKOO (*COCCYZUS AMERICANUS OCCIDENTALIS*) SURVEY IN ARIZONA

*Michael R. Kunzmann, Cynthia S.A. Wallace, Alexander R. Rybak, Royden J. Hobbs, Stuart E. Marsh, and William L. Halvorson

USGS Sonoran Desert Field Station, School of Renewable Natural Resources, 125 Biological Sciences East, University of Arizona, Tucson, Arizona 85721 (MRK, ARR, RJH, WLH)

Office of Arid Lands Studies, The University of Arizona, 1955 E. Sixth Street, Tucson, Arizona 85719 (SEM,CSAW)
**Correspondent: mrsk@sherpa.srnr.arizona.edu*

Abstract - The western yellow-billed cuckoo (*Coccyzus americanus occidentalis)* typically nests in mature riparian forests and woodlands along central and southern Arizona drainages and locally along the Virgin River in northern Arizona. Riparian habitat alteration has reduced the breeding range of the yellow-billed cuckoo over the last 60 years. As a result, western yellow-billed cuckoos have been petitioned for listing under the Endangered Species Act. To facilitate current species censusing and monitoring activities, a GIS and concomitant correlation models are relied upon to delineate potential yellow-billed cuckoo habitat and to help determine cost effective field data collection strategies. Initially, GIS and modeling efforts were constrained to legacy data sets such as existing layers of environmental data and historic breeding locations in Arizona. These coarse scaled methods for identifying potential locations of yellow-billed cuckoo breeding habitats have now been refined to include higher resolution spatial data and, using National Oceanic and Atmospheric Administration's (NOAA) Advanced Very High Resolution Radiometer (AVHRR) satellite imagery, a wider array of secondary database attributes required to succinctly define habitat requirements.

To Be Submitted To: *Southwestern Naturalist*

INTRODUCTION

In 1993, the yellow-billed cuckoo was split into two subspecies, eastern: *Coccyzus americanus americanus* and western: *C. a.occidentalis* (Franzreb and Laymon 1993). The yellow-billed cuckoo has been in decline throughout the western United States for the past 60 years. The decline is variously ascribed to habitat destruction (Franzreb, 1987), decreased water tables (Phillips, et al. 1964), and the use of pesticides (Gaines and Laymon, 1984; Rosenberg, et al. 1991). In February of 1998, 23 groups filed a petition with the United States Fish and Wildlife Service (USFWS) seeking endangered species status for the western subspecies. Currently the western yellow-billed cuckoo is still a candidate for listing as endangered.

Historically, the western yellow-billed cuckoo occupied and bred in riparian zones from southern British Columbia to northern Mexico, with an eastern edge in southwestern Idaho, Utah, western Colorado, New Mexico, western Texas, and Chihuahua, Mexico. Current evidence suggests that western yellow-billed cuckoo breeding is now restricted to California, Utah, Arizona, New Mexico, extreme western Texas, Sonora, Chihuahua, and south irregularly to Zacatecas, Mexico (Howell and Webb, 1995, Russell and Monson, 1998). The largest breeding populations of the Western yellow-billed cuckoo in the United States may now be confined to New Mexico and southeastern Arizona. Population estimates of the yellow-billed cuckoo include; less than 50 pair in California (Halterman, 1991); 902 pair along the middle Rio Grande and Pecos rivers in New Mexico (Howe, 1986); and 846 pair along the major rivers of southern Arizona (Groschupf pers. comm.).

The western yellow-billed cuckoo has been associated with cottonwood-willow dominated broadleaf deciduous riparian habitat (Hamilton and Hamilton, 1965; Gaines, 1974; Gaines and Laymon, 1984; Halterman, 1991). However, its exact distribution and habitat requirements in Arizona are not known.

OBJECTIVES

The first objective of this project was to revisit historic locations to confirm the presence and absence of yellow-billed cuckoos and compare sites to known records and to collect and better delineate vegetation characteristics in known yellow-billed cuckoo habitat (Figure 1).

The second objective was to use the NOAA AVHRR satellite imagery coupled with this fieldwork to create a map of potential yellow-billed cuckoo habitat. In this task, we evaluated the temporal dynamics of yellow-billed cuckoo areas as a function of the temporal change in vegetation spectral signatures obtained from satellite imagery. Such changes in vegetation communities are related to features of the landscape including vegetation phenology and condition. We hypothesized that:

1. Indices effectively characterizing the phenology of vegetation and temporal dynamics of the yellow-billed cuckoo landscape can be captured from the multitemporal satellite image data available bi-weekly;

2. Cuckoo habitat patches and sightings can be correlated to the temporal character of the corresponding satellite data; and

3. The landscape preferences of the cuckoo for locations with distinctive temporal signatures can be understood in terms of known cuckoo behavior and the physical characteristics of the landscape.

FIELD SURVEYS

Field surveys were conducted in 1998 to gather information on known habitat locations and begin a census of the Yellow-billed cuckoo.

Yellow-billed Cuckoo Survey Techniques:
California census protocols for this species (Laymon pers. comm.) were modified by a team of experts under Arizona Game and Fish Department (AGFD) leadership. Numerous project cooperators, including the AGFD, the USGS Forest and Rangeland Ecosystems Science Center - Colorado Plateau Field Station, the USGS Western Ecological Research Center – Sonoran Desert Field Station, and numerous private-sector organizations and individuals, have been instrumental in collecting presence and absence data.

Site surveys were conducted in 1998 using a "look-see" approach as described by Bibby et al. (1992). The method relies on prior knowledge of historic habitats, possible habitat preferences, expert opinion, and knowledge of the basic ecology of the species. This approach allowed project cooperators to delineate surveying sites on USGS topographical maps or aerial photographs. The location of all 1998 sites visited were digitized and incorporated into an ArcInfo GIS database for subsequent analysis. Global Positioning System (GPS) receivers and background maps from un-generated Environmental Systems Research Institute (ESRI) ArcInfo coverages, when used in conjunction with Geolink GPS data collection software, were extremely valuable in helping us find difficult survey sites on back-county roads. In addition, Geolink was critical in delineating the shape and size of yellow-billed cuckoo habitat patches (polygons) in the field.

To locate and census the birds, surveyors used a playback recording of the paired cuckoo's contact call ("kowlp" call) to elicit a response from nearby yellow-billed cuckoos (Laymon, 1998). Johnson et al. (1981) recommended this technique when surveying for secretive bird species or species that occupy dense vegetation. Recordings were played every 100 m (328 ft) to detect (attract) yellow-billed cuckoos. If no cuckoos

were detected during the initial listening period, the tape was played for five complete calls with a one-minute delay between call playbacks. In some cases, we did not need to solicit a yellow-billed cuckoo call because the birds were calling when we arrived at the site or during the setup of our vegetation transects. Most bird surveys were conducted between June 15 and August 20, 1998 to coincide with the peak of the yellow-billed cuckoo's breeding season in Arizona (Hamilton and Hamilton, 1965; Nolan and Thompson, 1975).

Vegetation Field Techniques:
We used three different field techniques to characterize the vegetation at each yellow-billed cuckoo site. The techniques selected were: (1) a modified Braun-Blanquet approach using dimensionless releves and prominence rankings (Warren et al. 1982; Bennett et al. 1998), (2) a quantitative sampling method using systematic circular plots to record detailed species composition, canopy-based measurements, percent cover, and other vegetation characteristics, and (3) a simple method of counting all trees that were at least 5 cm diameter at breast height (dbh) and at least 2.0 m in height. Field data were used to calculate similarity coefficients and matrices to help examine the relationships between sites with and without birds. This analysis will be published elsewhere. In addition, site vegetation characteristics were used as "training" sets for calibrating satellite imagery classification techniques.

Preliminary GIS Analysis:
Prior to summer fieldwork a GIS database was developed to facilitate project planning, to digitally delineate historic habitat patches, and to characterize habitat patches by examining polygon size, shape, elevation, and other GIS related variables. Arizona has a predicted 797,876 hectares of yellow-billed cuckoo habitat, approximately 2.7 % of the total land area in Arizona (Halvorson, et al. 2002).

DATA

The following data sets were accessed to accomplish this research:

- Cuckoo sightings, from the 1998 field surveys
- Land cover:
  - Cuckoo habitat, delineated by project cooperators as described in the Field Surveys section above.
  - Riparian vegetation, extracted from the GAP vegetation mapping project.
- NOAA AVHRR Satellite data

The satellite data used in this study were collected by the NOAA AVHRR satellite (Eidenshink, 1992; EDC, 1994) and processed by the U. S. Geological Survey (USGS). Although these data have a relatively coarse spatial resolution (each image pixel is equivalent to a 1-km^2 area on the ground), they possess a high (daily) temporal

resolution. AVHRR data are often used in research on land-cover change since the daily frequency of data availability permits the detection of short-lived vegetation and landscape changes that may be missed by bi-weekly sensors.

A derived image, the Normalized Difference Vegetation Index (NDVI), is generated by the USGS from the reflectance values (NDVI = NIR-RED/NIR+RED) of the AVHRR sensor in the visible red and near-infrared (NIR) regions of the electromagnetic spectrum and is sensitive to various biophysical vegetation characteristics, such as biomass and percent cover (Huete and Jackson, 1987; Price, 1992). The USGS has further calculated a maximum value composite of the daily NDVI images over every 2-week period to produce a relatively cloud free image of the landscape. These composited data for 1998 (26 images) were used in this study to capture the temporal variability of the landscape, including vegetation phenology. A more detailed description of these methods can be found in Kunzmann et al. (1999). The composite multitemporal NDVI images for the entire U.S. from 1989 to the present can be viewed and accessed at the website http://rangeview.arizona.edu/.

METHODS AND RESULTS

Compiling point data:
From the sightings and land cover data, we compiled four sets of point data:

1. The 1998 data with positive bird responses (data set A).
2. The 1998 data with negative bird responses (data set B).
3. Points sampled from polygons representing expert-defined cuckoo habitat patches (data set C).
4. Points sampled from polygons of perennial riparian areas (data set D).

Image processing techniques:
We derived measures describing the temporal dynamics of the landscape from the AVHRR NDVI data. These measures, catalogued in Table 1, were derived using three main classes of calculations: basic statistics, standardized principal components analysis (PCA), and Fourier analysis (Kunzmann, et al. 1999). Applying basic statistics, we treat the 26 values at each 1 km^2 pixel (one for each two-week composite) as a collection of data values and calculate the mean, standard deviation and the coefficient of variation (COV). These measures capture information on vegetation amount, variability, and normalized variability, respectively. Applying standardized PCA, we treat the study area as a stack of images, and derive the individual principal components PC1, PC2 and PC3. When applied to these data, these measures are known to capture information on vegetation vigor or density and two different measures of seasonality, respectively (Hirosawa et al., 1996). Applying Fourier analysis, we treat the ordered 26 values at each pixel as a vector (or waveform) and extract the additive term, first frequency phase and first frequency magnitude. These measures capture information on vegetation vigor or density, timing of maximum greenness, and variability of greenness, respectively. The

phase value, representing the timing of maximum greenness, is rescaled to the Julian date to facilitate interpretation.

Evaluating the landscape dynamics:
Image pixels were extracted for each of the four sets of point data (positive bird responses, negative bird responses, cuckoo habitat, and riparian areas). These data were examined for distinctive temporal attributes by applying statistical t-tests. The objective of applying the t-test was to identify the image-based variables that were significantly different in the cuckoo sightings or cuckoo habitat patches than in the general riparian areas or negative bird responses. These variables were then evaluated to determine whether they should be included in the habitat model.

Results of the t-tests comparing the sets of points described above are given in Table 2. These results can be summarized as follows:

- Locations with positive bird responses are significantly less green and less variable than general riparian areas. General riparian areas include many land cover types ranging from cuckoo habitat to agricultural areas. A significant difference in PC2 suggests a difference in seasonality, but this is not reflected in the Julian date.

- Expert-defined cuckoo habitat patches are also significantly less green but tend to be more variable than general riparian areas. A significant difference in PC3 suggests a difference in seasonality, and the Julian date shows the patches green-up 8 days later on average than the general riparian areas.

- Locations with positive bird responses are less green but equally variable when compared to predicted sites without birds (i.e., negative bird responses). A significant difference is found in the Julian date, where locations with birds green-up 34 days earlier than predicted sites without birds. A significant difference in PC2 also suggests a difference in seasonality between the sightings with and without birds.

An issue of concern with these data sets is that there are only 39 sightings with birds and 25 predicted sites without birds. For this reason, and because t-test results revealed similar relationships between cuckoo habitat and bird sightings, we selected the habitat variables based on the results of the t-tests between the cuckoo habitat points and the riparian points. Based on these results, all variables except PC2 can be used to discriminate between general riparian areas and cuckoo habitat. These variables were selected for evaluation as potential inputs to the final habitat model.

Habitat Preference Class Creation:
The derived images of the selected variables were stratified into cuckoo preference classes based on the observed distribution of the points within cuckoo habitat patches.

The mean and the standard deviation of the image values at these points were calculated and used to partition the data into a three-tiered preference class. Any data up to 1 standard deviation away from the mean were ranked highly preferred; data from 1 to 2 standard deviations away were considered moderately preferred, and anything outside of 2 standard deviations was considered not preferred.

The data ranges specified were then used to create maps of preferred habitat for each image variable. The resulting maps were further evaluated to see how well they predicted the actual sightings data locations. The number of actual cuckoo sightings within each of the three preference categories (Highly Preferred, Moderately Preferred and Not Preferred) for each of the partitioned temporal variables was tallied and the results are shown in Table 3.

The variables of temporal dynamics that best predict the locations of cuckoo sightings were used in the final model. These are shown in bold in Table 3 and include two measures each describing the Vigor/Density, Timing, and Variability of greenness, as follows:

Greenness Vigor/Density:
- Time Series Average NDVI (AZ98_TSAVE)
- Fourier "DC" Productivity (AZ98_PROD)

Timing of the green-up:
- Fourier Phase, rescaled (AZ98_JULIAN)
- Standardized Principal Component 3 (AZ98_PC3)

Variability of the greenness:
- Coefficient of Variation (AZ98_COV)
- Fourier Magnitude (AZ98_MAG)

To create the final habitat model, the six preference maps were reclassified to numerical labels representing the relative quantity of sightings within the categories, such that *"Rare or Not Preferred" = 1, "Uncommon or Somewhat Preferred" = 2, and "Common or Preferred" = 3.* The habitat model was then created by adding together the six stratified data layers, so that each preference variable was equally weighted. This produced a single model for preferred habitat with 13 classes, ranging in value from 6 ('Not Preferred' on all input layers) to 18 ('Preferred' on all input layers). Figure 2 shows the results of the habitat model for the whole state, with the 13 classes collapsed into the three preference classes: high, moderate, and low. However, because yellow-billed cuckoo are primarily found in riparian habitats the final surface generated was clipped (limited) to perennial streams with a 500 m buffer (Figure 3). In this figure, the high and moderate preference classes are combined into a single class of potential habitat.

Model validation:

A commonly used, rigorous test of a model's effectiveness involves first withholding a subset of the total data points from inclusion in the model. The model is 'trained' on the remaining set of data, and is then used to predict the excluded points. The amount of error

ACKNOWLEDGEMENTS

We acknowledge and thank the USGS Biological Resources Division Western Regional Office for project funding to do the field work necessary to conduct this preliminary investigation; Dr. Charles Van Riper, Unit Leader, USGS Colorado Plateau Field Station, Flagstaff, and his staff for census work, data sharing, and digitizing efforts; Mr. Robert Magill, Arizona Game and Fish Department for project coordination; Ms. Murrelet Halterman for her practical advice and yellow-billed cuckoo expertise; The School of Renewable Natural Resources and the following Staff members and graduate students who assisted with the project and related field work: Mr. Peter Bennett, Ms. Patty Guertin, and Mr. Volodymyr Ivakhnyk. We also thank Mr. Douglas Richardson of Baker-GeoResearch Inc., Billings, MT, for his support and the educational use of Geolink mapping software to train a new generation of students on GPS technologies; Ms. Jessica Walker for her assistance with the image processing; and Dr. Lee Graham for her compilation of the Arizona Gap Analysis yellow-billed cuckoo habitat model.

LITERATURE CITED

Bennett, P., M. Kunzmann, W. Halvorson. 1998. Manual for the GAP program: collecting vegetation classification data. USGS Sonoran Desert Field Station. The University of Arizona, Tucson, Arizona. Technical Report.

Bibby, C.J., N.D. Burgess and D.A. Hill. 1992. Bird census techniques. Academic Press. New York, New York. 257 pp.

Eastman, J.R. and M. Fulk. 1993. Long Sequence Time Series Evaluation Using Standardized Principal Components. Photogrammetric Engineering and Remote Sensing. 59(6):991-996.

EDC. 1994. Conterminous United States AVHRR Data on CD ROM - User's Manual, U.S. Department of Interior, Eros Data Center, Sioux Falls, South Dakota.

Eidenshink, J. C. 1992. The 1990 Conterminous U.S. AVHRR Data Set. Photogrammetric Engineering and Remote Sensing. 58:809-813.

Franzreb K. and Laymon, S.A. 1993. A reassessment of the taxonomic status of the yellow-billed cuckoo. Western Birds 24:17-28

Franzreb, K. 1987. Perspectives on managing riparian ecosystems for endangered bird species. Western Birds 18:10-13.

Gaines, D. 1974. Review of the status of the yellow-billed cuckoo in California: Sacramento Valley Populations. Condor 76:204-209.

Gaines D.A. and S.A. Laymon. 1984. Decline status and preservation of the yellow-billed cuckoo in California. Western Birds 15:49-80.

Halterman M. 1991. Distribution and habitat use of the yellow-billed cuckoo (*Coccyzus americanus occidentalis*) on the Sacramento River, California, 1987-1990. MS Thesis, California State University, Chico, California.

Halvorson, W. L. et al. 2002. Arizona GAP Final Report. USGS Sonoran Desert Field Station, University of Arizona, Tucson, Arizona.

Hamilton W.J. and M.E. Hamilton. 1965. Breeding characteristics of yellow-billed cuckoos in Arizona. Pages 405-432 *in* Proceedings of the California Academy of Sciences. Fourth Series.

Hirosawa, Y., S.E. Marsh and D.H. Kliman (1996) Application of Standardized Principal Component Analysis to Land-Cover Characterization Using Multitemporal AVHRR Data. Remote Sensing of Environment. 58:267-281.

Howe, W.H. 1986. Status of the yellow-billed cuckoo (*Coccyzus americanus*) in New Mexico. Final Report to New Mexico Department of Game and Fish. Share with wildlife.

Howell, S.N.G. and S. Webb. 1995. A guide to the birds of Mexico and northern Central America. Oxford University Press, New York, New York.

Huete, A.R., and Jackson, R.D. 1987. Suitability of Spectral Indices for Evaluating Vegetation Characteristics on Arid Rangelands. Remote Sensing of Environment. 23:213-232.

Johnson, R.R., B.T. Brown, L.T. Haight and J.M. Simpson. 1981. Playback recordings as a special avian censusing technique. Pages 68-75 *in* C.J. Ralph and J.M. Scott [eds.], Estimating numbers of terrestrial birds. Studies in avian biology. Allen Press, Inc. Lawrence Kansas.

Kunzmann, M. R., R. Hobbs, C. S. A. Wallace and S. E. Marsh. 1999. AVHRR imagery analysis and habitat modeling as planning tools for conducting yellow-billed cuckoo (*Coccyzus americanus occidentalis*) surveys in Arizona. Electronic Proceedings of 1999 ESRI Conference.

Nolan, V. Jr. and C.F. Thompson. 1975. The occurrence and significance of anomalous reproductive activities in two North American non-parasitic cuckoos *Coccyzus* spp. Ibis 117:496-503.

Phillips, A., J. Marshall and G. Marshall. 1964. The birds of Arizona. University of Arizona Press, Tucson, Arizona. 220 pp.

Price, J.C. 1992. Estimating Vegetation Amount from Visible and Near Infrared Reflectances. Remote Sensing of Environment. 41:29-34.

Rosenberg K.V., R.D. Ohmart, W.C. Hunter, B.W. Anderson. 1991. Birds of the lower Colorado Valley. University of Arizona Press, Tucson, Arizona. 416 pp.

Russell, S.M. and G. Monson. 1998. The birds of Sonora. University of Arizona Press. Tucson, Arizona. pp 360.

Warren, P. L., K. L. Reichardt, D. A. Mouat, B. T. Brown, and R. R. Johnson. 1982. Vegetation of Grand Canyon National Park. USGS Sonoran Desert Field Station, University of Arizona, Tucson, Arizona, Technical Report 9. 140 p.

Table 1. Images Derived for Yellow-billed Cuckoo Surveys.

Image Name	Captures	Definition	Tool Used
AZ98_TSAVE	greenness vigor/density	mean of 26 bi-weekly AVHRR NDVI points	Erdas Imagine
AZ98_TSSDE	greenness variability	standard deviation of pixel profiles	Erdas Imagine
AZ98_COV	greenness variability	coefficient of variation of pixel profiles	Erdas Imagine
AZ98_DC	greenness vigor/density	best-fit flat line to temporal NDVI profile	Matlab
AZ98_MAG	greenness variability	best-fit sine wave to annual NDVI profile	Matlab
AZ98_JULIAN	timing of green-up	peak of annual cycle best-fit sine wave as Julian date (0 to 365)	Matlab
AZ98_PC1	greenness vigor/density	1st principal component (PC1) from standardized principal components analysis (SPCA)	Erdas Imagine
AZ98_PC2	seasonality of green-up	2nd principal component (PC2) from SPCA	Erdas Imagine
AZ98_PC3	seasonality of green-up	3rd principal component (PC3) from SPCA	Erdas Imagine

Using SPSS software, we performed two-tailed student's t-tests for independent means to compare the following sets of data:

1. Bird sightings (set A) vs. a random sample of riparian points (set D).
2. Habitat patch points (set C) vs. a random sample of riparian points (set D).
3. Bird sightings (set A) vs. predicted sites without birds (set B).

Table 3: Number of birds sighted in each category for each partitioned temporal variable.

Preference	AZ98_ TSAVE	AZ98_ TSSD	AZ98_ COV	AZ98_ MAG	AZ98_ PROD	AZ98_ JULIAN	AZ98_ PC3	AZ98_ PC1
Low	1	0	0	0	1	0	1	2
Medium	13	20	13	9	10	15	6	34
High	25	19	26	30	28	24	32	3
Total Birds	39	39	39	39	39	39	39	39

Figure 1: The San Pedro River is characteristic of southern Arizona habitat for the yellow-billed cuckoo. The vegetation community changes from a *Typha latifolia/Baccharis salicifolia* strand to a *Populus fremontii/Celtis recticulata* association as a function of distance from the river.:

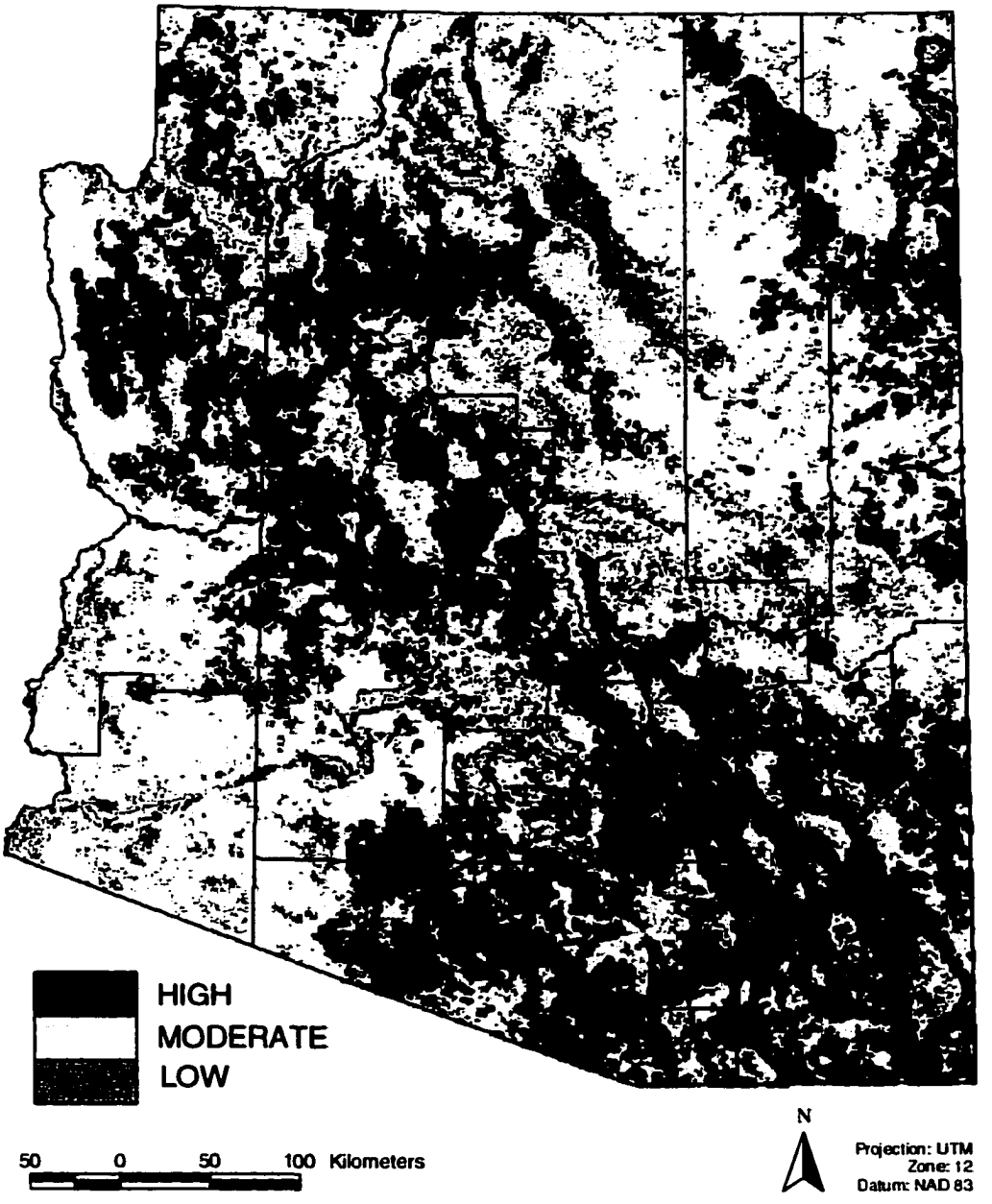


Figure 2. Statewide prediction of yellow-billed cuckoo habitat using AVHRR data. Areas shown as preferred represent landscapes that possess distinctive temporal characteristics similar to those found in known yellow-billed cuckoo habitat.

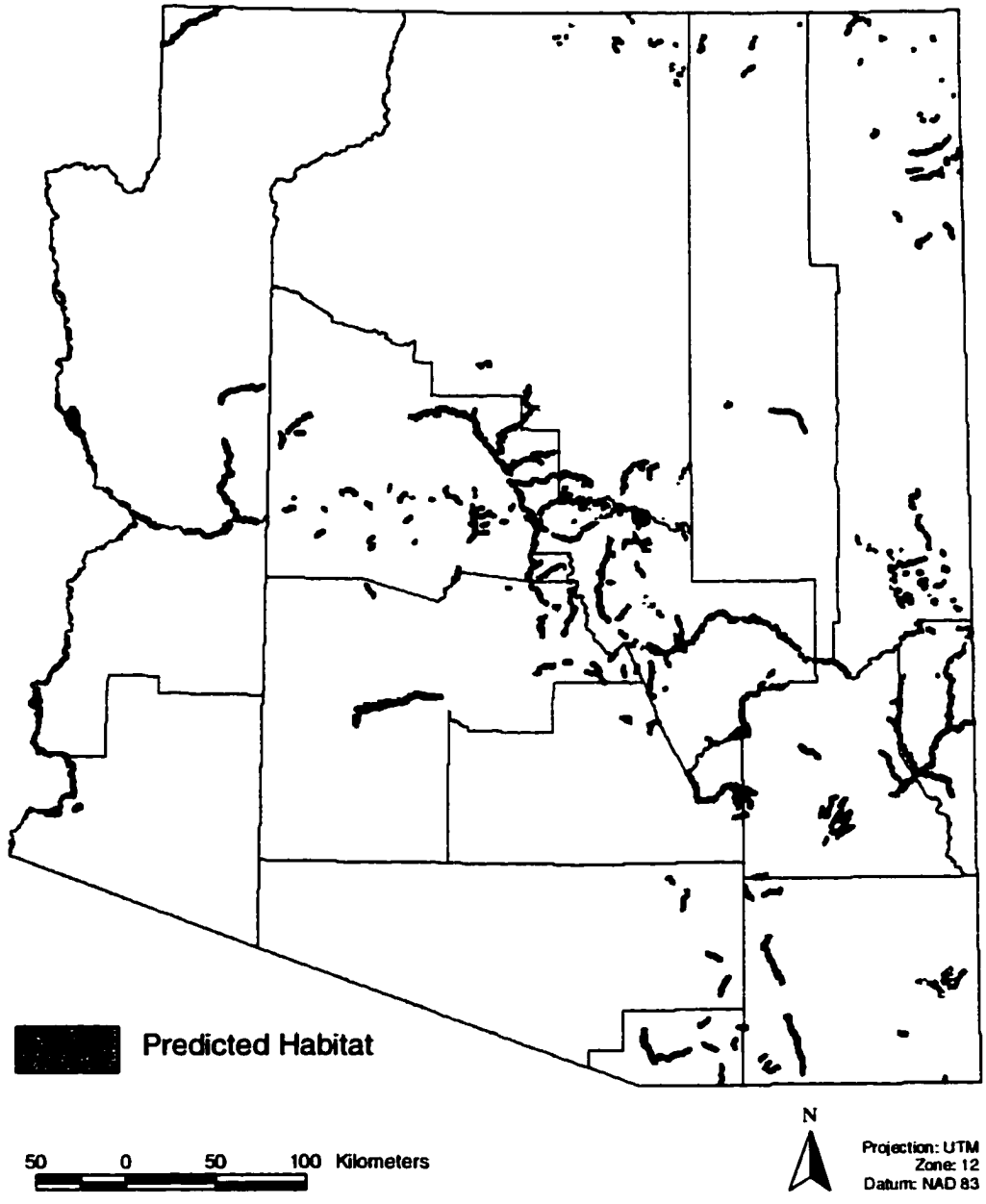


Figure 3: AVHRR predicted yellow-billed cuckoo habitat clipped to perennial water courses buffered to 500m.

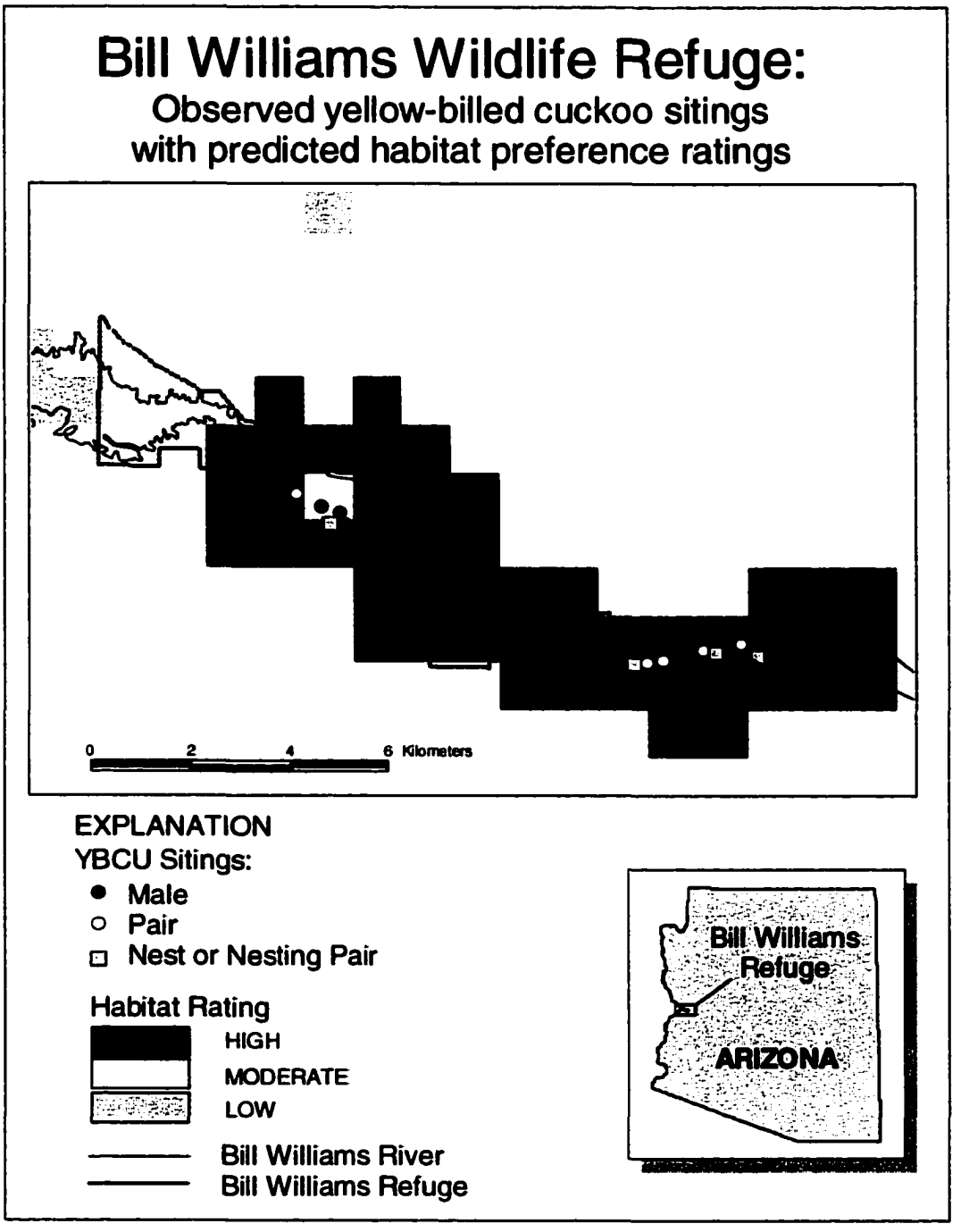


Figure 4: AVHRR predicted yellow-billed cuckoo habitat overlaid with known yellow-billed cuckoo locations on the Bill Williams National Wildlife Refuge.