INTEGRATED APPROACH TO PREDICTIVE MODELING: A CASE STUDY FROM THE UPPER XINGU (MATTO GROSSO, BRAZIL)

By

JOSEPH CHRISTIAN RUSSELL

A DISSERTATION PRESENTED TO THE GRADUATE SCHOOL OF THE UNIVERSITY OF FLORIDA IN PARTIAL FULFILLMENT OF THE REQUIREMENTS FOR THE DEGREE OF DOCTOR OF PHILOSOPHY

UNIVERSITY OF FLORIDA

2005
Copyright 2005

by

Joseph Christian Russell
ACKNOWLEDGMENTS

I would like to acknowledge all those who played an integral part in my study’s development. My thanks go to my committee members: Dr. Ken Sassaman, Dr. Michael Binford, Dr. Richard Stepp, and most especially to Dr. Michael Heckenberger. The data presented in this dissertation came as a direct result of funding secured by Dr. Heckenberger, and it is thanks to his dedication to his students, and to his spirit of academic collaboration that this study was made possible. I would like to thank my fiancée, Michelle, without whose support none of this would have been possible. Additionally, I would like to thank Cooper, for reminding me that everything can be made better through the judicious use of sleep. Finally, I extend heart-felt thanks to my dog, Jasper, for reminding me daily that there is certainly more to life than writing, and that running around outdoors and catching sticks has its own rewards.
## TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>Section</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>ACKNOWLEDGMENTS</td>
<td>iii</td>
</tr>
<tr>
<td>LIST OF TABLES</td>
<td>vii</td>
</tr>
<tr>
<td>LIST OF FIGURES</td>
<td>viii</td>
</tr>
<tr>
<td>ABSTRACT</td>
<td>x</td>
</tr>
<tr>
<td>CHAPTER</td>
<td></td>
</tr>
<tr>
<td>1 INTRODUCTION</td>
<td>1</td>
</tr>
<tr>
<td>Selection of a Region</td>
<td>1</td>
</tr>
<tr>
<td>Synthetic Approaches</td>
<td>6</td>
</tr>
<tr>
<td>Organization</td>
<td>9</td>
</tr>
<tr>
<td>2 AMAZONIAN ECOLOGY</td>
<td>12</td>
</tr>
<tr>
<td>Xingu Indigenous Park</td>
<td>14</td>
</tr>
<tr>
<td>Description of the Region</td>
<td>18</td>
</tr>
<tr>
<td>Early Attempts to Characterize the Region</td>
<td>20</td>
</tr>
<tr>
<td>Varzea-Terra Firme Dichotomy</td>
<td>24</td>
</tr>
<tr>
<td>Importance of the Upper Xingu</td>
<td>40</td>
</tr>
<tr>
<td>3 PREDICTIVE MODELING AND GIS</td>
<td>42</td>
</tr>
<tr>
<td>Defining the Model</td>
<td>42</td>
</tr>
<tr>
<td>Predictive Modeling and Archaeology</td>
<td>42</td>
</tr>
<tr>
<td>Differing Approaches to the Modeling Problem</td>
<td>49</td>
</tr>
<tr>
<td>Model Testing</td>
<td>58</td>
</tr>
<tr>
<td>Refining the Model</td>
<td>60</td>
</tr>
<tr>
<td>Criticisms of GIS and Predictive Models</td>
<td>60</td>
</tr>
<tr>
<td>What Drives the Model?</td>
<td>65</td>
</tr>
<tr>
<td>Meshing of Archaeological Theory with Predictive Models</td>
<td>69</td>
</tr>
<tr>
<td>Integration of Landscape</td>
<td>73</td>
</tr>
<tr>
<td>Refined Approach</td>
<td>74</td>
</tr>
</tbody>
</table>
4 GLOBAL POSITIONING SYSTEM (GPS) ..............................................................78
   Using GPS in Archaeological Survey .........................................................78
   GPS Use in the Upper Xingu .................................................................88
5 REMOTE SENSING .........................................................................................99
   Brief Overview .........................................................................................99
   Processing Imagery for Analysis .........................................................103
6 METHODS .......................................................................................................121
   Traditional Analysis Techniques .........................................................122
   Realization of an Integrated Approach .................................................124
   Pre-classification Techniques .............................................................129
   Classification .........................................................................................131
   Results ....................................................................................................133
   Data Transformations ..........................................................................134
      Normalized Differential Vegetation Index (NDVI) .........................136
      Transformed Normalized Differential Vegetation Index (TNDVI) ....141
      Transform Vegetation Index (TVI) ....................................................142
      Simple subtraction Vegetation Index (SVI) .......................................142
      Devised Band Ratios .........................................................................142
      Soil Adjusted Vegetation Index (SAVI) .............................................142
      Tassel Cap Transformation .............................................................143
      Decorrelation ....................................................................................145
      Principal Components Analysis .....................................................146
   Revised Methodology ..........................................................................149
7 RESULTS ........................................................................................................153
   Post-Classification Procedures and GIS Manipulation .......................153
   Discussion of the Classification Process .............................................164
8 DISCUSSION .................................................................................................174
   Summary and Conclusions of Methods .................................................175
   Land Use/Land Change in Amazonia ....................................................188
      Perspectives ......................................................................................188
      Future Directions .............................................................................191
APPENDIX
A CONFUSION MATRICES, CLASS SPECTRAL PROFILES, AND SEPARABILITY VALUES .........................................................194
LIST OF REFERENCES .................................................................................................................. 217
BIOGRAPHICAL SKETCH ........................................................................................................... 268
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>6-1.</td>
<td>The PCA statistics for 2002 (August) transformation</td>
<td>147</td>
</tr>
<tr>
<td>6-2.</td>
<td>The PCA statistics for 2003 (May) transformation</td>
<td>148</td>
</tr>
<tr>
<td>7-3.</td>
<td>The 2002 (August) accuracy results for combined classes</td>
<td>158</td>
</tr>
<tr>
<td>7-4.</td>
<td>The 2003 (May) accuracy results for combined classes</td>
<td>158</td>
</tr>
<tr>
<td>A-1.</td>
<td>The 2002 (August) combined classes confusion matrix</td>
<td>194</td>
</tr>
<tr>
<td>A-2.</td>
<td>The 2003 (May) combined classes confusion matrix</td>
<td>197</td>
</tr>
<tr>
<td>A-3.</td>
<td>The spectral profiles for the 2002 (August) combined classification</td>
<td>200</td>
</tr>
<tr>
<td>A-4.</td>
<td>The spectral profiles for the 2003 (May) combined classification</td>
<td>205</td>
</tr>
<tr>
<td>A-5.</td>
<td>The separability values for the 2002 (August) classes</td>
<td>210</td>
</tr>
<tr>
<td>A-6.</td>
<td>The separability values for the 2003 (May) classes</td>
<td>213</td>
</tr>
</tbody>
</table>
LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>1-1. Overview of Brazil and location of the Upper Xingu study area</td>
<td>3</td>
</tr>
<tr>
<td>1-2. Map of the Upper Xingu project study area showing known archaeological site locations (adapted from unpublished project imagery)</td>
<td>4</td>
</tr>
<tr>
<td>2-1. Overview of study area and Xingu Park southern boundaries in the Upper Xingu basin, Matto Grosso state, Brazil (5, 4, 3 color composite, May 2003)</td>
<td>16</td>
</tr>
<tr>
<td>4-1. GPS survey map of MTFX-6</td>
<td>90</td>
</tr>
<tr>
<td>4-2. GPS survey map of MTFX-13</td>
<td>90</td>
</tr>
<tr>
<td>4-3. 5, 4, 3 composite Landsat 7 image (August 2002) with generalized legend for a rough interpretation of land-cover classes</td>
<td>91</td>
</tr>
<tr>
<td>4-4. MTFX-6 and MTFX-13 survey data integrated into a GIS with a 5, 4, 3 color composite from Landsat 7 imagery</td>
<td>92</td>
</tr>
<tr>
<td>4-5. GPS survey map of MTFX-18 integrated into a GIS with a 5, 4, 3 color composite from Landsat 7 imagery</td>
<td>93</td>
</tr>
<tr>
<td>4-6. GPS survey maps of MTFX-19 and MTFX 20 integrated into a GIS with a 5, 4, 3 color composite from Landsat 7 imagery</td>
<td>94</td>
</tr>
<tr>
<td>4-7. Composite GPS survey map of the Nokugu cluster of sites integrated into a GIS with a 5,4,3 color composite from Landsat 7 imagery</td>
<td>95</td>
</tr>
<tr>
<td>4-8. Composite GPS survey map of Kuikugu cluster of sites integrated into a GIS with a 5, 4, 3 color composite from Landsat 7 imagery</td>
<td>96</td>
</tr>
<tr>
<td>6-1. Methodological flowchart</td>
<td>128</td>
</tr>
<tr>
<td>6-2. Detail of the Upper Xingu study region (2002 NDVI transform)</td>
<td>139</td>
</tr>
<tr>
<td>6-3. Detail of the Upper Xingu study region (2003 NDVI transform)</td>
<td>140</td>
</tr>
<tr>
<td>7-1. Detail of the 2002 (August) supervised classification</td>
<td>156</td>
</tr>
</tbody>
</table>
7-2. Detail of the 2003 (May) supervised classification

7-3. Detail of predicted site locations (PC1 of vegetation indices composite, PC1 of soil indices composite, PC2 of Landsat 7, August 2002)
The objective of this research was to assess the degree to which an implementation of GPS-based survey techniques, and the extraction of information through a combination of GIS and remotely sensed imagery, in conjunction with more formal archaeological techniques and ethnoarchaeological research methods, can add to the formation of a predictive modeling effort aimed at developing a heuristic model capable of identifying additional significant site locations based on the vegetative signature unique to anthropogenic soils in the area. The paucity of data for the Upper Xingu region of Brazil (and Amazonia in general) makes it an excellent location for the use of new technologically-informed approaches to studying tropical land use, and providing an initial step toward an empirical understanding of the anthropogenic aspects of landscape formation.

Important questions facing Amazonian anthropology today are the nature of prehistoric and present human-environment interaction, and the impacts of human
settlement in non-western tropical settings. This dissertation addresses some of the problems of quantifying and qualifying the nature of the landscape transformation resulting from long-term human occupation and settlement of the Upper Xingu region of southern Amazonia. I have examined the nature of the Xinguano impact on their local environment, how they transformed their immediate landscape, and how the results of this work undermine current theories of a fundamentally non-dynamic environment coupled with models of low-density population.

There is a growing acknowledgement by archaeologists of the fundamental utility of GIS, GPS survey, and the integration of remotely sensed imagery into current studies. This begs more in-depth investigations into the extent to which each technology may be utilized. The application of the predictive model of archaeological site locations described in this study allowed rapid identification of areas with a high probability of past human occupation. This has the potential to direct more efficient archaeological exploration of the Indigenous Park of the Xingu. In all, some 1800 km² of predicted site locations were identified within the region of study. During the course of the Upper Xingu Project, headed by Michael Heckenberger, a small sample of this region was discovered to have undergone massive alteration by human agents (Heckenberger et al. 2003). The pattern of anthropogenic vegetation uncovered through this investigation adds to a mounting body of evidence that would seem to reaffirm this hypothesis.
CHAPTER 1
INTRODUCTION

Selection of a Region

In anthropology, geography, geology, ecology, and numerous other disciplines, a consensus of opinion has been reached regarding the dramatic impact human society can have on the world around it. Often times, these arguments are couched in terms of the adverse nature of human impact on the environment and the best ways to minimize potentially harmful consequences of human activities. In much the same vein, few would argue that modern ecosystems are static communities. Instead, they are viewed as manifestations of continuous and complex interactions among organisms and between organisms and the abiotic elements of their environment. Such interactions are not simply modern phenomena. This complex interplay of actors and elements has always existed, and the state of modern ecosystems is very much a product of their histories.

The sum of the analyses in this dissertation represent an effort to understand how humans and the environment within which they live act upon one another over time, manifested as transformations writ large on the landscape. My study sought to add to a body of evidence informing on larger questions critical to the debate about the role of humans in the formation of landscapes, and the impacts of long-term human occupation on the environment in non-western tropical settings. Specifically, this research posits how one might be able to quantify what those impacts might look like in the present day by deriving a specific signature for vegetation growing in past settlement areas, and then asking how one might then identify such areas when a predictive model based on that
signature is extrapolated out to the larger region. The methodological approach used in this dissertation is an initial step to an empirical understanding of the human-environment dynamic at play in the Upper Xingu. The rationale behind my study was to understand the Upper Xingu region, not just as an object of study, but also as the home for thousands of people who are coping with a heterogeneous and diverse environment on a daily basis in the present, mimicking the lifeways of ancestral populations who did much the same in the past (Heckenberger 1996, 2005), and providing a further point of discussion on questions of what exactly is the nature of the human impact on the environment in this region, and how do these impacts challenge theories of a fundamentally non-dynamic environment?

Simply stating that human beings have a dramatic impact on their surrounding environments would seem, on the surface, a common-sensical premise. However, in the minds of some, this has yet to be proven within Amazonia. Certainly little archaeology has been conducted in Amazonia (and especially in the Upper Xingu), and yet some researchers would have others believe that the human populations of this region were somehow less capable of transforming their surroundings to meet their needs. Additionally, the Upper Xingu, while researched, remains largely untreated in the anthropological literature. Using data gathered first-hand, and informed by the efforts of other researchers currently working (or who have worked) in this and similar regions, I tried to understand, and to qualify and quantify, the effects of long-term human occupation and transformations of the environment that accompany such activities through an integrated approach of utilizing GPS survey, geographic information systems,
and remotely sensed imagery. My study focused on the area occupied by the Indigenous Park of the Xingu, as well as surrounding territory (Figure 1-1).

Figure 1-1. Overview of Brazil and location of the Upper Xingu study area

The bulk of this dissertation is predicated on the results of more than a decade of anthropological research conducted by Dr. Michael Heckenberger (University of Florida, Department of Anthropology). The ethnographic materials used in this dissertation are a result of Dr. Heckenberger’s ongoing partnership with the region’s Kuikuro. The bulk of the GPS survey was conducted by myself and other project members in cooperation with Kuikuro informants as part of the larger National Science Foundation funded project “Late Prehistoric Social Complexity in Southern Amazonia (Upper Xingu, Brazil).” The primary purpose of this research was to document regional settlement patterns (size, placement, and form of occupation sites) and determine if there was a rank-order settlement hierarchy. Research to date is summarized in The Ecology of Power: Archaeology and
Memory in the Southern Amazon (Heckenberger 2005) and some recent findings have been reported in Science (Heckenberger et al. 2003).

Heckenberger’s efforts in the Kuikuro study area have identified numerous prehistoric settlements (Figure 1-2), primarily through major earthworks. Detailed mapping was conducted over most major earthworks at six prehistoric archaeological sites, demonstrating a remarkably complex and integrated regional plan. The earthworks include: 1) excavated ditches in and around ancient settlements (up to 2.5 km long and 5 meters deep); 2) linear mounds bordering major roads and circular plazas (averaging about 0.5 to 1.0 meter high); and 3) a variety of wetland features (bridges, weirs, artificial

Figure 1-2. Map of the Upper Xingu project study area showing known archaeological site locations (adapted from unpublished project imagery)
ponds, raised causeways, canals, and other structures) that are often less obvious but clearly integrated into the regional plan (Heckenberger et al. 2003; Heckenberger 2005).

The research conducted thus far in as part of Upper Xingu study has revealed a rank-order hierarchy, including a number of larger first-order occupation sites, with multiple plazas, and earthen features that seem to indicate the remains of large population centers (Heckenberger et al. 2003; Heckenberger 2005). There also appear to be some second-order settlements, with single or double plazas, comparatively fewer earthworks, with evidence of ties to larger centers, as well as some third-order plaza villages that have one plaza, are much smaller, and have earthworks only in the immediate area of the plaza (Heckenberger et al. 2003; Heckenberger 2005).

Thus, the Upper Xingu provides a wide array of data sources upon which to draw, including preserved landscapes, indigenous populations, and archaeological sites. What is lacking is any meaningful attempt to produce information using methods that are technologically based. Thus, this region provides an opportunity to form some guidelines for integrating new technologies with traditional archaeology to refute older theories and perhaps posit new questions.

The model of low density prehistoric populations in the Amazon is outdated, and generally agreed on to no longer accurately reflect past settlement patterns. In light of recent modern research into the nature of human-environment interactions, we have begun to understand the Amazon, and the people who have occupied that region, in a new light. At issue is the vast territory to be covered, and a means by which we can quantify and qualify precisely what these interactions are, and how humans transform their environment over time. The amount of land area to be studied seems, on the surface,
prohibitively expansive for traditional archaeological investigative techniques.

Thankfully, new technologies vastly increase our ability to operate over larger scales and with greater efficiency and rapidity than previously possible.

**Synthetic Approaches**

Use of newer technologies that can be integrated into most methodologies to alter how we archive, manage, analyze, and display data is becoming more widespread in archaeological research. The potential of these tools can be realized only if their varied functions are correctly understood, and if they can effectively interact with one another. If recent literature is any indication, the archaeological world has become increasingly curious and interested in the application of Geographic Information Systems (GIS), Global Positioning Systems (GPS), and Remote Sensing (RS) to anthropological problems. The main impediments to full adoption of these new tools (and a move away from entrenched methodological approaches) hinges on the difficulty in using software packages, which are sometimes complex to implement and operate; and the difficulty in integrating varied datasets, in different formats, coming from different sources, and based on fundamentally different models of analysis. The answer to these problems will be the availability of interoperable geoprocessing and image-processing technologies.

By meshing traditional archaeology with an integrated approach to these new technologies, this dissertation attempted to understand the long-term aspects of human/environment interaction in the Upper Xingu region by interpreting archaeological evidence and remotely sensed imagery within the context of a new methodological approach to data collection, processing, and analysis. This integrative approach sought to extract the strengths from each system to contribute to a synthesized methodological approach for predicting site locations within the Upper Xingu of Brazil using spectral
signatures of specific training sites as the basis of defining location parameters. The goal of this research is not to replace ground-truthed or ground-informed archaeological investigations; merely to develop a method to increase the efficiency with which specific regions of interest might be targeted. These findings will have implications, not just for archaeologists, but also, to other disciplines, as well as advancing our general knowledge of the region.

This study resulted in a realized model, but perhaps its value is greater in demonstrating how various analysis and data collection platforms may be integrated. One of the greatest obstacles to successful synthesis of technology into the design of an archeological model is that the archaeological data are often widely varied in content (anthropic or natural facts, sites, features, historical texts), thus the information is not homogeneous (aerial/digital/near infra-red photographs, remote sensed images, drawings, plans, contour surveys, paper/digital maps, CAD layers, GIS databases, field notebooks, and so on). Such is the power of a GIS that varied data sources can be treated as different layers of information. Thus, a predictive model derived from the informed use of multiple data sources can be realized.

An elementary assumption in any predictive model of prehistoric activity loci is that some locations are more suitable for conducting specific activities than others, or that some locations were of greater strategic, religious, or other cultural importance. It is also assumed that any attributes contributing to this differential suitability can be quantified and represented using modern data sources, or can be represented by proxy data sources. To date, the majority of predictive models formulated by archaeologists in the United States demonstrated that the variables chosen by archaeologists to correlate with
archaeological sites often represent noncultural aspects of the landscape. Though these variables form the backbone of the “prediction,” even the archaeologists themselves oftentimes explicitly acknowledged that the variables might have no causal relationship with the placement of sites. Such variables are considered indicators. While these variables often have no causal relationship with the placement of sites, when taken in combination, they invariably result in the discovery of sites. Contemporary efforts at archaeological prediction explicitly or implicitly assume that environmental factors are the primary (even exclusive) determinants of most human behavior. The causal link between site locations and natural, independent variables is usually considered multivariate, meaning that sites were positioned to take advantage of an optimal combination of all the resources critical to the cultural group under investigation (Ebert and Kohler 1988:107).

The fundamental flaw in models of this sort (setting aside the environmentally deterministic overtones) is the subjectivity involved in assessing the potential of prehistoric archaeological site existence while relying primarily on the expertise and experience of archaeologists. While subjective knowledge should never be discounted (and, in fact, may be one of the most valuable resources available to the modeler), such efforts are flawed from the standpoint of replicability outside of very specific regions. My model differs from many of the existing models in its attempt to utilize indigenous knowledge and contemporary cultural patterns, and incorporate that understanding into the modeling process.

Any effort to predict where people chose to settle in the past requires some understanding of the people and their activities. Initially, it might seem burdensome and
problematic to introduce anthropological explanation and culturally informed variables into a discussion of practical archaeological prediction. However, it is only in the context of explanation and explanatory modeling utilizing data from numerous sources to derive variables outside of simplistic geomorphological feature proximics so common in modeling to date that archaeologists can successfully predict the locations and other characteristics of the materials that make up the archaeological record. Variables that determine the locations of occupational sites are not static properties of the environment that can be measured easily from topographic or environmental maps (Ebert and Kohler 1988:158-159). No one would argue that geographical or environmental variables (often characterized by landscape features such as river flats or terraces, joining or narrowing bodies of water, soil drainage among others) are effective for finding sites, but it is questionable whether they are finding the total range of site types available or only the sites they expect to find. Models based on culturally uniform variables are valid to a certain extent, and often times will find more sites than would be found by pure chance. However, adding other types of data, and introducing various platforms of data collection broadens the scope of modeling efforts, expanding the range of identified sites, ultimately proving more useful for managing archaeological resources in areas where insufficient archaeological work has been conducted.

**Organization**

The organization of this dissertation reflects the multi-disciplinary character of the research. Each chapter is devoted to a topic supporting the development of a methodological approach to the integrative use of technology in the Upper Xingu region to expand our understanding of human-environment interaction in that region. Chapter 2 describes the theoretical, environmental, and cultural setting for the study. The Upper
Xingu is placed squarely into the context of the Amazon region and the ongoing debate within anthropological circles governing the capacity of the environment to support large-scale settlement of human populations.

The next three chapters provide an introduction to each of the systems that will be used is provided, including a short synopsis of how each software/data collection platform operates. Chapter 3 addresses the theoretical development of predictive modeling approaching in general, and in anthropology in particular. The provided overview of predictive models includes commentary on their inception, development, and the fine-tuning of their components. In addition, some of the inherent weaknesses of current modeling trends are demonstrated, and a list of critiques of predictive models outlined. A short section is also provided regarding the development of predictive models pertinent to these investigations. This chapter also presents the first of three technologies used in this study, geographic information systems (or GIS), and explains its utility, especially in light of modeling efforts.

Chapter 4 contains an explanation of the second utilized technology, GPS. The overall system of GPS is explained in detail. Additionally, arguments are made for the greater inclusion of GPS in archaeological survey, including an explanation of the many benefits to using GPS over traditional survey methodologies. Finally, I present the GPS survey data collected in the Upper Xingu, demonstrating the utility of GPS survey methods as a critical aspect of integrative approach methodologies. Data collected in the field supported the creation of a database including size of archaeological sites, and was essential to the overall mapping of the project area. Moreover, spectral data referring to
the specific regions of interest were tied directly to georeferenced points and were integrated into the database.

Chapter 5 expounds on the utility of remote sensed imagery, and specifically Landsat imagery, to provide a backdrop for the formulation of an overall approach methodology.

In Chapter 6, the integrative approach is finally realized. As this study involves the use of Landsat imagery, in combination with a variety of image processing techniques, GPS surveyed archaeological data sets, and the integration of these two types of data within a GIS to form a cohesive whole, this chapter is devoted to the method of classification of the overall image, and the results of those classifications, as well as the output of the final analysis product.

Chapter 7 presents the analysis of the classified imagery, as well as a reporting of the findings and their implications to the larger arguments regarding the degree to which populations with the Xingu altered their surrounds.

In the final chapter, concluding remarks are addressed. Questions and hypotheses are discussed based on the main findings of the dissertation. Further studies are also suggested in search of a better understanding of the relationship of humans and their environment in the Amazon.
CHAPTER 2
AMAZONIAN ECOLOGY

The Upper Xingu region of Brazil provides an excellent backdrop for addressing questions relating to the nature of the landscape transformation that has occurred as a result of long-term prehistoric human occupation, specifically in regards to discussing what the dynamic interaction of the Xinguano and local environments, and how they transformed their landscape. The paucity of data for this region makes it an excellent candidate for the use of new approaches to researching land use in tropical settings. This dissertation sought to apply technologically-informed research techniques to provide an initial step towards an empirical understanding of landscape formation through the formation of heuristic model designed to predict archaeological site locations through the spectral signature of anthropogenic vegetation.

The Amazon retains a prominent position in both Western popular and scientific imagination as a region relatively untouched by human hands. This concept of a “pristine” environment is belied by the archaeological evidence suggesting the presence of large, densely settled, and integrated regional populations (Heckenberger 1996; Heckenberger et al. 1999; Heckenberger et al. 2003; Heckenberger 2005; Petersen and Heckenberger 2001; Porro 1996; Roosevelt 1980, 1991; Whitehead 1998), a concept that has long been suggested by several researchers in the area (e.g., Carneiro 1970; Denevan 1976; Lathrap 1968, 1970). Demonstrable evidence for large, complex social formations in the Upper Xingu, and the extensive influence these populations may have had on the natural environment has important implications for how anthropologists, ecologists, as
well as researchers in many other disciplines, view the non-Western peoples indigenous to the region, and the nature of the landscape transformation these peoples may have performed.

The attention to anthropogenic forests, beginning two decades ago, was a paradigm shift in Amazonian ecology and ethnology, and has motivated research on human-environmental interactions in the region (Balee 1987, 1989, 1993; Balee and Campbell 1990; Brondizio 1996; Brondizio et al. 1994, 1996, 2002; Carniero 1983, 1985, 1987, 1995; Heckenberger 1996, 1998, 1999, 2005; Heckenberger et al. 1999, 2001, 2003). However, perceptions of the transformative nature of human occupation on surrounding environments are certainly not a new concept. As early as the mid 1800s, von Humbolt pondered “plant geography,” a collective study of vegetation, and argued, “the vegetation of a region was an expression of the physical environment, and also a direct influence on Mankind, both materially and spiritually” (Nicolson 1987: 177). George Marsh’s Man and Nature or Physical Geography as Modified by Human Action (1864) also grew out of this growing acknowledgement of the fundamental link between human populations and the environments within which they lived. This concept of the dynamic nature of human/environmental interaction remained an undercurrent through much of the 20th century in anthropology (Boas 1887), ecology (Clements 1936; Sayce 1938), and in general writings and thought (Sauer 1956; Bates 1956; Mumford 1956). This trajectory of literature regarding the anthropogenic nature of environments, and the transformative processes human actors perform in the formation of landscapes formed the foundation for much of modern thought on the subject.
In light of the pedigree and abundant nature of literature regarding these issues, it does seem surprising that theories of a fundamentally non-dynamic environment would have persisted for so long, or that some researchers would have others believe that the human populations of this region were somehow less capable of transforming their surroundings to meet their needs. Modern considerations of the Amazon demonstrate a near-universal agreement of the widespread anthropogenic transformations of vegetation. Indeed, Heckenberger et al. (2003) effectively correlated spatial and structural patterns of vegetation and archaeological remains with settlement distribution, adding to the body of evidence of these large-scale transformations. Through the use of an integrated technological approach, I sought to add to this body of evidence, specifically in describing how researchers might assess the extent of landscape transformation through a heuristic modeling effort.

**Xingu Indigenous Park**

As stated in the introductory chapter, this dissertation centered on the Xingu Indigenous Park (or PIX after the Brazilian designation of Parque Indigena do Xingu), specifically encompassing most of the Upper Xingu cultural area, as well as areas immediately to the north of the Xingu River proper (Figure 2-1). The Upper Xingu contains a number of diverse ethnic groups including the Aueti, Kalapalo, Kamayurá, Kuikuro, Matipu, Mehinaku, Nahukuá, Trumai, Wauja, and Yawalapiti. While retaining distinct cultural differences, these groups maintain a highly interconnected network of specialized trade, marriages, and inter-village rituals. The PIX is located in the northeastern part of the State of Mato Grosso, in the southern part of the Brazilian Amazon. Comprising some 26420.03 km², the park encompasses enormous biodiversity and a number of distinct ecological communities, ranging from savannas and dryer, semi
deciduous forests to the south, to the other end of the scale with Amazonian
ombrophylous forest to the north (including extremely dense primary and secondary
forest cover, fields, flood land forests, terra firme forests, and forests on terra preta).
Climatically, the region alternates between a wet season (November to April) and a dry
season during the rest of the year.

The Park was formed July 31, 1961, with adjustments made in 1968, and 1971,
with the final perimeter demarcation completed in 1978. Deemed a “National Park,” this
region was intended to serve both as a protected ecosystem and as a haven for the
indigenous populations that guided its creation. With the formation of Funai in 1967
(replacing the SPI, or Indian Protection Service), the “National Park” designation was
dropped in favor of the “Indigenous Park” moniker.
The Indigenous Park of the Xingu consists of three generalized regions: one to the north (known as the Lower Xingu), one in the central region (the Middle Xingu) and one to the south (the Upper Xingu). The southern regions harbor the primary feeder rivers of the Xingu river system and contain a closely knit federation of distinct cultural groups. With final demarcation of the boundaries of the park in 1978, only a few years elapsed before the first incidents involving encroachment into park boundaries. At first, trespassing was in the form of hunters and fishermen at an individual, low-impact level.
However, by the end of the Twentieth Century, numerous forest fires on cattle ranches to the northeast, as well as the encroachment of lumber companies to the west began to compromise the park borders. Additionally, occupation and agricultural practices near the headwaters of the park river system caused increasing pollution of the internal water supply. This pollution, in addition to other forms of encroachment on park boundaries, are among the most pressing issues to the populations resident in the region. Indigenous groups in the park, however, have proactively defended the sovereignty of park boundaries through litigation. Two territorial lawsuits, in particular, resulted in the Wawi and Batovi Indigenous Lands, of the Suyá and the Wauja respectively. Ratified in 1998, these expanded areas brought the total land area of the Park to approximately 2,797,491 hectares.

The question of monitoring the territorial boundaries is paramount in the minds of the indigenous population, often discussed in meetings of leaders, assemblies of the Xingu Indigenous Land Association (ATIX), and in meetings with the Funai and federal and state environmental agencies (IBAMA and the State Environmental Foundation, or FEMA). Eleven vigilance posts have been established thus far to protect and buffer areas that allow direct access to the Park, as in the case where highway BR-080 borders park boundaries or where major waterways intersect the park.

While vital to protecting the interests of the indigenous populations, this series of posts alone is insufficient deterrent to boundary violations. Other systems have been put into place to assist the park inhabitants in maintaining the integrity of park borders. For example, a coordinated partnership between ATIX and ISA led to an ongoing project (the Borders Project), that systematically maps deforestation for trend analysis and use in
litigation, as well as identifies new vectors of settlement and occupation of areas in close proximity to park borders. This joint partnership has taken on the task of constructing training programs for the vigilance post inhabitants, restoring and maintaining boundary markers delimiting the physical limits of the park, and maintaining a geo-referenced database of all ranch owners bordering the Park. The Borders Project allows park inhabitants to closely follow breaches of park boundaries, mobilizes resident communities, and provides a medium for dialog through inter-village discussions and with the public agencies responsible (FUNAI, IBAMA, and the state government).

**Description of the Region**

The most striking feature of the Amazon Basin are its diversity of ecosystems and the conglomeration of many distinct and different cultures (contrary to traditional portrayals of limited ecological and cultural diversity). Though research efforts into every aspect of the Amazon (from anthropological studies to botanical categorizations), have made numerous leaps from the relative paucity of information catalogued by the mid-20th century, we still lack a robust understanding of the complexity of the region. Efforts to construct prehistory and protohistoric periods of tropical lowland South America are severely hindered by the fragmentary and often unreliable nature of the cultural data in some instances, and by the sheer lack of any real information at all in others. Despite the efforts of numerous researchers to begin to understand the vast areas of the Amazon Basin, a great deal of the area is still unknown archeologically. Thus, the Amazon has often been portrayed as a pristine habitat, with indigenous peoples viewed as little more than Stone Age peoples living in harmony with the ancient, unchanged forest. Such models have been summarily dismantled by evidence of significant changes in
environment and human adaptation before the European conquest of the Americas. Perhaps the most important pieces of evidence to date suggest that ancient Amazonians made considerable impacts on the habitat in areas adjacent to their settlements, and that some distinctive forest patterns once thought purely natural now may be seen as having been influenced by past human activities.

Anthropologists have made numerous attempts to categorize (and organize) the Amazon into a number of distinct cultural areas, and much of the current thought governing potential social complexity in the Amazon can be attributed to an evolving understanding of the cultural history of indigenous populations. Early studies on the region tried to explain the apparent lack of social complexity by the existence of “limiting factors,” thought responsible for relatively low population densities (Heckenberger 1996, 1998, Heckenberger et al. 1999, 2003). Meggers (1954) suggested that levels of social complexity are limited by the agricultural potential of soils in the Amazon. Others have argued that protein sources, not necessarily the production of highly caloric crops, have limited population densities throughout the region (Gross 1975; Ross 1978).

Recent archeological evidence from the region shatters many preconceived notions of the nature of complexity in pre-Columbian Amazonia, proving the existence of complex prehistoric cultures in the Amazon (Lathrap 1970; Myers 1973; Roosevelt 1987, 1989a, 1989b; Heckenberger 1999, Heckenberger et al 2003). Agronomic and biophysical surveys demonstrate a diversity of soils and land resources, including areas suitable for large-scale agriculture (Falesi 1974; Cochrane and Sánchez 1982; Nicholaides et al. 1984). The very process of colonization, evidenced today in ethnographic studies of population migrations, and shown in the past through
archaeological investigations, suggests that previous views of the Amazon were overly
simplistic, and that much more complex systems are in play, requiring a fundamental
shift in the way researchers construct their questions and seek answers if we hope to ever
understand the potential of the Amazon for development in the past and in the present
(Moran 1981, 1984; Schmink and Wood 1984; Bunker 1985; Hecht and Cockburn 1990;
Stewart 1994).

**Early Attempts to Characterize the Region**

Steward provided the first substantial body of information governing cultural
development in South America with the publication of the Handbook of South American
Indians (HSAI). In his general theory, Steward promoted the viewpoint of cultural
ecology, which portrayed cultures as functionally integrated wholes, possessing
technology designed to maximize ecological exploitation. In Amazonia, the environment
was viewed as the primary limiting factor in cultural development (Roosevelt 1980: 3).
He classified South American cultures according to a list of characteristics focused on
exploitation of local environments. The focus on environment and material culture,
especially those items utilized in exploiting resources, over the symbolic-ritual
complexes enabled Steward to create circumscribed geographic boundaries wherein
shared cultural traits delineated “culture areas.”

Steward’s Handbook laid the groundwork for an environmental determinist, or eco-
functionalist (Heckenberger 2005), perspective in Amazonian scholarship where direct
linkages were suggested between population density and natural resources. Steward was
the first to maintain that variance in populations was due to ecological constraints, and
for a number of decades this line of reasoning persisted. Typical exchanges between
Amazonian scholars consisted of debate over the role of the environment in conditioning
human adaptive strategies and guiding the course of cultural development in the region. The environment was viewed in terms of absolute constraints it posed to human societies and cultural evolution, as typified in Betty Meggers’ (1954, 1996) version of Steward’s environmental determinist viewpoint, which she believed would better explain cultural development in general and tropical lowland culture in particular.

Meggers seized on Steward’s “Tropical Forest Tribes” with its core of unifying traits. Using Steward’s definitions, the Tropical Forest culture embodies a number of signature traits, notably a reliance on hunting and/or fishing, gathering, and rudimentary slash-and-burn cultivation for subsistence. She strengthened Steward's argument by collating evidence on the poverty and fragility of tropical soils. Meggers believed that both patterns of settlement and formation of a synthesized culture were a direct result of the amount and quality of cultivable land available to a given population. Meggers in effect argued that the highly leached soil substrate and rapid decomposition of organic matter in humid tropical climates imposed absolute limitations on the possibilities of intensive cultivation and soil improvement (which in turn were determinants of the slash-and-burn cultivation system prevalent among upland horticulturalists by ethnographic accounts), resulting in local population pressure which could be felt even under low density unless people lived in small, nomadic communities.

In “Environmental Limitation on the Development of Culture” (1954), Meggers asserted that the primary locus of interaction between a culture and its environment can be found in subsistence activities, and that the most vital aspect of the environment, from the point of view of the participants, is its suitability for food production. She outlined the importance of agriculture in cultural development and identified the effects of soil
fertility, climate and rainfall as critical to the productivity of agriculture. In turn, productivity regulates population size and other factors related to the survival of the group (1954). Her working hypothesis was that there must exist a casual relationship between the type of environment and the maximum cultural development that a given environment can support.

Meggers’s hypothesis was the central theme in discussions concerning Amazonian cultural development for decades, centering on whether or not various subsistence economies could provide the basis for sedentism and population growth. In general, Meggers’s assumed that the limiting factors on population size and density were tied to the ability of human cultures to adapt to heterogeneous environmental conditions. She dismissed the possibility of the presence of large-scale sedentary populations in Amazonia, due to environmentally limiting factors inconducive, in her mind, to supporting large populations. In essence, Meggers’ dismissed any ability of prehistoric populations to manage environmental conditions to achieve levels of subsistence necessary for population growth and management (Gross et al. 1979; Moran 1982).

Utilizing the trope of the “Tropical Forest Tribe” (denoting a limited population and essentially placing a cap on the developmental sequence of the peoples under investigation), Meggers has tenaciously held onto a belief that the small populations that characterize the impermanent ethnographic communities of modern times can sufficiently characterize the full extent of variability of cultural groups of the past (Roosevelt and Meggers 1996, 2001; Heckenberger et al. 1999, 2001). Deemed the “Standard Model” (Viveiros de Castro 1996), this outdated point of view was subsumed by revisionist thinking expanding on Julian Steward’s initial observations (1946) that perhaps the
different environments represented by floodplain and upland areas might hold the key to some of the cultural variation he observed. In gross terms, regional specialists typically delineate two ecological zones within Amazonian: *terra firme* was defined as land not subject to annual inundation, with elevations varying from immediately above flood level to several hundred meters, and *varzea* as the flood plain of “white water” rivers, laden with fertile silt. Some authors question, however, whether their distinction is particularly relevant to cultural diversity, at least when framed in such simplistic, dichotomous terms (Whitehead 1996; Heckenberger et al. 1999, Heckenberger 2005).

Certainly there is widespread acknowledgement on the part of most researchers in the Amazon that, at least within the *varzea* regions, large, sedentary societies could and did exist. Archaeological evidence points to ceramic-making Archaic Sedentary Shellfishers (evidenced by fluvial and maritime shell middens) that formed settled villages as early as 7500 BP. By 6000 to 4000 BP, there is evidence of initial riverine agriculture based on diversified root crops, fruit trees, and numerous other tropical forest resources in parts of Amazonia (Heckenberger et al. 1999; Lathrap 1970, 1977; Neves 2001; Neves et al. 2003; Oliver 2001; Petersen et al. 2001; Roosevelt et al. 1991). Later (2000-3000 BP) sites on the Orinoco River demonstrate the widespread adoption of ceramic griddles (postulated to have been used in manioc preparation) and pottery forms generally associated with the Amazonian Barrancoid, or Incised-Modeled ceramic tradition, providing evidence of early complex societies with intra-regional interaction. By around 3000 BP, sites in the Central Amazon and the middle Orinoco exhibit evidence of institutionalized social hierarchy, with a growth of regional chiefdom-like

The Marajoara culture at the mouth of the Amazon (developing *in situ* by ca. AD 400) has been revealed as a complex chiefdom boasting numerous villages or towns containing numerous domestic structures, cemeteries, and ceremonial complexes (Roosevelt 1991, 1994, 1999; Scheen 1997, 2001, 2004). The complexity of the mound construction suggests long-term habitation and ceremonial use of the sites, sometimes in groups of up to 40 related mounds; these are interpreted as small regional polities (Scheer 2005). Cultural distinct polychrome pottery began to be produced by AD 400-600, reaching a florescence by 1000 AD. This mound-building culture used ceramic urns for burial, both in cemeteries and accompanying single/few individual burials.

Within the middle Amazon, Amazonian Borrancoid persisted until about 900 AD at sites like Acutuba, Manacupuru, and Paredao (Heckeberger et al. 1998). Incised and punctuated styles became dominant along the Lower Amazon by 1000 AD. The best examples of this ceramic style are Tapajonica from Santarem, at the mouth of the Tapajos River, and the related Konduri style upstream around the Trombetas River. When Europeans encountered the Tapajos in the 1500s, they observed a continuous strip of settled villages along the riverbanks, organized as chiefdoms, with village chiefs subordinate to a paramount chief (Medina 1998, Porro 1996). Agricultural production likely centered on the cultivation of both manioc and maize (Roosevelt 1991).

**Varzea-Terra Firme Dichotomy**

To explain archaeological evidence of complex social organization in numerous floodplain areas and Contact-period accounts of large Amerindian groups with complex social structure, many researchers (e.g. Carneiro 1970, 1986, 1995; Deneven 1984, 1996;
Lathrap 1960, 1970; Meggers 1996; Roosevelt 1980, 1989, 1991, 1994) have hypothesized a *varzea* model emphasizing that the rich floodplain soils could, in fact, support large, fully sedentary villages (see Heckenberger et al. 1999). Termed the “*varzea-terra firme* dichotomy,” this hypothesis was designed to account for the distribution of fine ceramics, large populations, and monumental architecture within the floodplain regions of the Amazon and its Andean derived tributaries. These elements were believed to be representative of complex social formation suggesting the possibility of early chiefdom-level organization of cultures within this region. Proponents of this model suggest that the higher productivity of the soils located in *varzea* regions provided the basis for sedentary populations to settle into these riverine areas, and, conversely, that the reduced fertility of soils in the *terra firme* limit growth (thus, inherent in this theoretical model is a bias against the formation of complex social organization in the *terra firme* regions).

This revisionist view accounted for the increasing archaeological evidence of the formation of complex social organization in the *varzea*, as well as the seeming lack of complexity in the upland areas. In a bizarre leap of logic, many were led many to believe that the ethnographic present of displaced, disenfranchised groups could sufficiently explain the nature of *terra firme* social organization in the past, but, in an amazing display of duplicitous logic, they found that the model was insufficient to explain the level of social organization in floodplain areas clearly documented in Contact periods accounts. Continued excavation of sites within the floodplains, and the lack of any real archaeological investigations into the development of more complex social organization in the upland areas, seemed to bolster the *varzea* hypothesis.
This *varzea-terra firme* dichotomy became the launching point of a number of discourses attempting to unravel the apparent lack of development in interfluvial regions. Lathrap (1968) maintained an adaptivist view, expanding on the division between *terra firme*-interfluvial groups and *varzea* populations engaged in more sedentary root-crop farming combined with fishing and the hunting of aquatic mammals and reptiles. Lathrap placed primacy on the contrast between the two ecological zones in terms of differential productivity, both in agriculture and animal bio-mass. Lathrap’s thesis is that most of the groups inhabiting the tropical inland are the remains of evolved agricultural societies forced into an environment unsuitable to the basic economic pattern (1968). They had to rely on the hunting of forest game to provide the protein and fat essential to the diet. As game was often scarce, they were forced into nomadism, a decline in agricultural productivity, and a still greater dependence on wild food (Lathrap 1968). This led to Lathrap’s formation of the “Tropical Forest Culture” (1970), significantly different from early categorizations of upland settlement groups in that it addressed shared economic and cultural patterns independent of Morgan-esque evolutionary stages of cultural development.

Carneiro also felt that earlier explanations were too simplistic and static. Late nineteenth century travelers' opinions that the soils of the region were rich and productive were quickly followed by suggestions that soils were so poor that they could not support complex cultures or intensive cultivation. While it is true that tropical rain forests typically have the highest biomass and species diversity of any terrestrial ecosystem, the selfsame climatic conditions that create an ideal environment for a wide variety of flora is also ideal for chemical weathering, erosion, and the destruction of humus. Soils in many
parts of the Amazon Basin are generally considered poor, leached, acidic, and capable of producing for only a couple of years before being abandoned due to fertility decline (McNeil 1964). The soils of the upland tropical forests are dominated by highly acidic and nutrient-poor zones. The biological diversity of the upland forest regions is directly related to nutrient recycling and a number of complex, extremely localized adaptations of plants regional conditions, as well as land management approaches practiced by local populations in the past and present rather than any inherent qualities of the soils of the regions (Posey and Balee 1989). Furthermore, indigenous technologies and a wide range of domesticated and semi-domesticated plants provide diverse avenues for increased productivity.

Within the some floodplain regions, soils are rejuvenated each year. As a matter of course, it is more useful to say the humid tropics include some of the poorest and some of the richest soils in the world (Sanchez 1976, 1977). In 1954, Meggers set out to demonstrate that the poor and acid soils of the humid tropics could only support small-scale societies living by swidden cultivation because soil fertility could only be sustained from the nutrients released by the slash-and-burn method of cultivation. Denevan (1976: 225) called attention to the extremely low estimates of population density made by Julian Steward (1946), who suggested densities of barely 0.2 persons per square kilometer for lowland South America. Of course overall population density would have been highly variable, due to any number of localized conditions like productivity of soils, availability of edible flora, and concentrations of suitable fauna.

In general, it has been suggested (and may well be true) that the upland interfluvial forests had lower population densities than savannas and floodplains in lowland South
America. Nonetheless, it remains to be widely demonstrated that this is uniformly or even typically the case. Prehistoric populations over the centuries seem to have brought about significant environmental modification that has enhanced the net returns to human populations and ensured that familiar environments are present from one generation to the next. For a long time after the initial arrival of the Europeans, the upland forests offered refuge to native populations fleeing from the incursions of colonial and national societies. Carneiro challenged several of Meggers' notions about the limits imposed on Amazonian culture by a rainforest environment (Carneiro 1970, 1983, 1985, 1995). In a study among the Kuikuru of the upper Xingu River, he estimated the carrying capacity of a slash-and-burn system, and his calculations suggested a much denser population limit and a more elaborate cultural development than Meggers' estimations (Carneiro 1995). With ethnographic data from the Kuikuru, Carneiro states that "with slash-and-burn as the only limiting factor, a village of some 2,000 persons could live on a permanent basis where the Kuikuru do now" (Carneiro 1961: 232 cf. Carneiro 1995: 48). It turned out that any soil limitations could be corrected through the addition of organic matter, mulches, irrigation, and overall proper soil management. Carneiro's studies flew in the face of Meggers’ ecological limitation of soil fertility, and his studies have mostly focused on the role of other ecological factors related to food crops in the evolution of Amazonian chiefdoms.

Carneiro developed a hypothesis to explain the rise of complex social organization in the varzea that actually began to trend away from environmental determinism, seeking explanation in “such things as the availability of varzea, the huge quantities of aquatic food resources found in the waters of the Amazon, and other ecological factors, including
population pressure and warfare” (1995: 52). Roosevelt repudiated Carneiro’s environmental circumscription theory, offering her maize hypothesis (1980) as an alternative means of overcoming the believed environmental limitations of the Amazon and providing the impetus for the formation of chiefdom-level social organization. Roosevelt’s maize hypothesis has itself been brought under severe scrutiny (Carneiro 1995) and all but dismissed by Roosevelt herself. To date, the general disagreement between all parties involved over the nature of environmental limitations has been heated, and there has been an unfortunate tendency to focus on single limiting factors (Beckerman 1979 and Gross 1975 on protein availability; Meggers 1954 vs. Carneiro 1957 on soil fertility).

In spite of an ever-increasing body of knowledge about numerous aspects of Amazonia demonstrating irrefutably just how variable tropical habitats can be, there still seems to be some impetus to hold onto vestigial linkage with existing hypotheses based on earlier viewpoints. For some inexplicable reason, many researchers in the Amazon continue to suggest that the findings from one site are generalizable to the entire region. Most anthropologists whole-hearted accept the varzea-terra firme dichotomy, categorizing as terra firme adaptation a slew of findings from ecologically disparate regions as the Xingu Basin, the Rio Negro Basin, and the central Brazilian savannas. At its heart, however, the dichotomy is replete with false assumptions and broad generalizations that render the entire basis of the theorem untenable.

Scientific studies across numerous disciples have begun to recognize the complexity of the of the region, and the fallacy in derivation to so simple a picture (Moran 1993, 1995). On the one hand, some basic findings were generalized for the
entire region as 'unifying principles': a general scarcity of nutrients in the soil (Sombroek 1984); a tightly closed, continuous recycling of nutrients within the biomass of the forest (Jordan 1989); extreme diversity of the biota (Prance and Lovejoy 1985); and the regional recycling of a large part of the rainwater, crucial for the maintenance of a climate affected by pluvial processes (Schubart and Salati 1980; Dickinson 1987). On the other hand, the heterogeneity of such a broad territory was recognized in several fields. Among other regional elements, Amazonian river types were differentiated into whitewater (e.g., Solimões-Amazonas, Madeira), clearwater (e.g., Tapajós, Xingu), and blackwater (e.g., Negro) (Sioli 1984). Sombroek (1984) divided soils into three major categories: well-drained soils of the uplands, imperfectly drained soils in the sedimentary parts of the region, and poorly drained soils in varzeas and igapos. Pires (1984) and Pires and Prance (1985) have described several vegetation types. A myriad of papers have addressed the human dimension behind the environmental heterogeneity (Balee 1989; Deneven 1984; Hames and Vickers 1983; Meggers 1974; Moran 1974, 1995; Moran and Herrera 1984; Posey and Balee 1989; Sponsel 1992).

Recent scholarship has begun to refute previous positions, providing evidentiary information regarding the numerous techniques used by indigenous groups in social (regional trade, kinship ties), technological (food storage, resource procurement) and ecological (settlement in proximity to multiple ecological zones) realms to buffer from potential environmental constraints (Denevan 1996; Posey 1983). In addition, substantial evidence is available for the modification of the landscape by native populations, leading to the creation of novel resources and demonstrating a greater degree of human agency with the environment (Balee 1989; Clay 1988; Denevan 1992). Perhaps the most
powerful evidence comes from Heckenberger’s 1999 publication that demonstrated the existence of large, fully sedentary populations in a broad range of ecological settings in pre-contact Amazonia (Heckenberger et al. 1999), further discrediting environmental determinants of cultural evolution.

The tropical rain forest has long had a reputation for being pristine. Strong evidence suggests, however, that much of the vegetative cover in the Amazon Basin is actually the result of careful cultivation, rooted in anthropogenic processes. The high biodiversity of the Amazonian Basin places it unquestioningly at the top of a list of the most diverse ecosystems on the planet. However, the scale of “ecosystem” puts researchers at a disadvantage to talk about the reasons why we see the variability of flora and fauna, where the origins of such biodiversity herald. In keeping with principles of historical ecological theory, the resultant image that we have of the overall distribution of so many different types of organisms is actually the result of natural disturbances in environmental systems (meaning naturally occurring changes to the environment) being met with human response, and vise versa. Brown and Lugo (1990: 4) estimate that today about forty percent of the tropical forest in Latin America is secondary as a result of human clearing and that most of the remainder has had some modification despite current low population densities. Many soils of the lowland humid tropics have long been considered too infertile to support sustainable agriculture. The primary thrust of this argument relies on the assumption that due to the rapid decomposition of organic matter as a result of high temperatures, large amounts of precipitation, and the lack of stabilizing minerals in the humid tropics, sustainable agriculture simply would not be possible. Increasingly strong evidence refutes such research, instead implying that permanent or
semi-permanent agriculture can itself create sustainable fertile soils in a cyclical self-sustaining and a self-renewing.

Modifications of human origin have dramatic impacts on regrowth opportunities for various flora (and, by extension, fauna). For example, species composition in fallow periods of slash-and-burn horticultural techniques (adopted by many ingenious groups) differ significantly from natural gaps in canopy, altering “species composition of the mature forest on a long-term scale” (Walschburger and Von Hildebrand 1991: 262). This is not to suggest that human modification of tropical forests, be it past or present, is limited to a simple “slash and burn” model. Rather, the Amazonian Basin is composed of a sequence of crafted landscapes, operating as a functional system at local levels, and mirrored in its functionality at more regional levels. Large expanses of what are viewed as “natural” forested areas were and are created landscapes conforming to a larger design in which the kinds, numbers, and distributions of useful species are managed by human populations. As previously indicated, a vital mechanism in forest management is manipulation of slash-and-burn fallow periods in an effort to replenish or attract specific types of vegetation or faunal species of interest to local populations. The effort expended in selection, transplantation, and protection of specific forms of “wild” vegetation is a premeditated labor that belies the simplicity with which some regard slash-and-burn cultivation, suggesting instead a preconceived notion of well-manicured environments (Denevan and Padoch 1987). Balee (1987, 1989) has expounded on the “anthropogenic” forests in Amazonia in which numerous species have been carefully manipulated to increase the availability of specific species without adversely affecting local biodiversity. These include specialized forests (babassu, Brazil nuts, lianas, palms, bamboo),
composing approximately 12% of the total upland forest in the Brazilian Amazon (Balee 1989: 14). Balee (1989:14) concludes "large portions of Amazonian forests appear to exhibit the continuing effects of past human interference."

What has become painfully obvious is that older models attempting to characterize the variation in social organization and complexity in the occupation of upland and riverine ecological zones based upon the differential perceived productivity of these regions have vastly underestimated the ability of indigenous populations to fully exploit a given environment. While the agricultural potential of soils in the interfluvial regions of the Amazon may not have the same nutrient level, they are, nonetheless, more than sufficient to allow intensive production of manioc. This coupled with the high availability of protein in the form of aquatic resources could provide more than adequate calories to sustain large populations, allowing sedentary groups to grow and thrive. Manioc cultivation, even in the poorest soils, has been hypothesized to be perfectly adequate for providing a stable economic foundation for large, settled populations (Carneiro 1987, 1995; Heckenberger 1998; Heckenberger et al. 1999; Petersen et al. 1999.). Anthropogenic soils, known as terra preta do índio, anthrosols, and terra preta arqueológica, or Amazonian Dark Earths (ADE)(Smith 1980; Kern and Kampf 1989; Kern et al. 2003; Woods et al. 2003), are highly coveted areas for cultivation (Smith 1980). The presence of terra preta along the edges of the large floodplains, as well as in the uplands, has been hypothesized to demarcate areas of intense cultivation (Balee 1989: 10-12; Smith 1980). Terra preta anthrosols are closely associated with flat-tops of escarpments of the well-drained Tertiary Plateau (Terra firme), where they form patches resulting from ancient middens (waste deposits). These soils contain charcoal and
cultural waste from prehistoric burning and settlement with high carbon, nitrogen, calcium, and phosphorus content leading to a specific vegetative signature (Kern and Kampf 1989; Zech et al. 1990; Kern and Costa 1997; and Glaser et al. 2000; Woods et al. 2003) and are often associated with types of forests that are believed to be the product of long-term management by native Amazonians (Balee 1989). When several species dominate in terra firme forests, human activity is usually implicated (Anderson 1983; Balee 1989; Balee and Campbell 1990). In numerous published inventories of “virgin” tropical rain forest in the Lower Amazon, certain species are more common than expected (Campbell et al. 1986) and inferring either direct human intervention, or, at the very least, significant human-induced impacts on local ecosystems. Most importantly, anthropogenic upland forests that grow on past occupation sites are commonly extensive enough to be detected in remote sensing data due to the distinctive texture and reflectivity of their vegetation.

Debate may have once centered on theories of natural origins for terra preta (Smith 1980), it is now indisputable that such deposits are, in fact, human artifacts (Heckenberger et al. 1999; Woods 1995; Woods and McCann 1999). It is highly improbable upland soils well outside the normal floodplain environments could ever attain such high concentrations of P, Ca, K, and C without some sort of anthropogenic influence, evidenced by the chemical composition (Eidt 1977), spatial distribution (typically on high landforms where depositional processes could only be cultural), and the presence of artifacts throughout the modified horizon. Sites of terra preta usually occur in areas averaging 20ha (Smith 1980; Zech et al. 1990; McCann et al. 2001), but very large terra preta sites up to 350ha have also been reported (Smith 1999). Smith
(1999) has shown evidence that terra preta is not restricted to varzea, but also occurs on
terra firma regions as well. The similarity of the texture and mineralogy with that of
surrounding soils (Zech et al. 1990), in conjunction with the occurrence of pre-
Columbian ceramics and other artifacts in the upper horizons of terra preta soils
(Sombroek 1966; Smith 1980) seems to suggest that the occurrence of this “black earth”
is actually an indicator of human activity, and methodical enhancing of poorer soils with
fertilizing materials through the introduction of soil organic matter (SOM) and nutrients
(1990) and supported by later scholars (McCann et al. 2001; Woods and McCann 1999)
suggests that chemical and biological processes work in conjunction to form deposits of
terra preta once a threshold of nutrient retention capacity and biological activity is
reached as a result of cultural activity (burning, deposition, etc.).

The absence of significant ADE deposits in many contemporary Amerindian and
mestizo settlements (Heckenberger et al. 1999) would suggest that itinerancy, low-density
settlement, and traditional slash-and-burn agricultural practices are not the patterns of
settlement and land-use that led terra preta production. The ability to enrich localized
environments to suit changing cultural demands would indicate that human agency is an
incredibly powerful force (not theory, but and actualizing mechanism. The possibility
that these anthrosols were intentionally modified to expand livelihood opportunities in
previously impoverished landscapes, as suggested by Herrera et al. (1992) and Mora et al.
(1991) is an incredibly powerful statement of the impact that humans, historically or
prehistorically, had in constructing landscapes. An emerging view stresses the ecological
praxis of native Amazonians (intensive agricultural practices and intentional soil
modification), or stated more simply, “soils were a constraint, but people overcame them” (Mann 2000: 788). However, we must consider the possibility that terra preta may well be a simple byproduct of changes in settlement and livelihood. Selection of settlement locals are certainly intentional, as are the choices of agricultural practices undertaken in a particular environment, and waste treatment and other activities carried out within settlements. But does the intentionality of acts leading to the production of enriched soils necessarily mean that the end result is also intentional? Terra preta has thus generated a great deal of interest among the research community in Amazonia owing to a broad recognition of the import of discussions of “black earth” formation and its ties to Amazonian cultural history (Whitehead 1998).

Research on terra preta has suggested that, even in areas with low agricultural potential, soils could be modified in such a way as to dramatically increase productive potential. To assess intentionality and the degree of intentionality in terra preta formation, it is important first to identify benefits derived from these soils once formed. Despite widespread acceptance of anthropogenic origins and soil enrichment (Kern et al. 1989; Woods 1995; Woods and McCann 1999), evidence for the importance of these soils in traditional agriculture has been mostly anecdotal. Furthermore, historical factors that played a role in defining the cultural importance of terra preta through time, leading to its treatment as an important resource, have yet to be addressed. McCann et al. (2001) discuss indigenous practices of soil inoculation with ash, organic material, and microorganisms, and the implications of these practices for tropical soil management. Yet linkages between intentional agricultural intensification and terra preta formation, and between terra preta and an increase in human carrying capacity, have been
insufficiently discussed. Furthermore, while the presence of terra preta continues to provide evidence of relatively permanent settlement in Amazonia prior to European arrival (Denevan 1996; Smith 1980), the suggestion that these richer pockets of soil made possible this transition to a more sedentary lifestyle is as yet unproven, but a highly seductive hypothesis. If Peterson et al. (1999) are correct in their assumption that terra preta was the result of human occupation rather than the genesis of it, then certainly it would have served as an important incentive for continuous occupation. Additionally, if, as has been suggested, (Peterson et al. 1999) terra preta was, in fact, the direct result of intensive occupation, then this would have created something analogous to a living organism (self-perpetuating and self-sustaining as accumulations of human waste acted as a vector of renewal).

Contributions of recent historical ecological literature indicate indigenous groups of the Amazon Basin may have actively modified the environments in which they live (Balee 1989, 1998; Posey 1985, 1998). Most literature in this vein refutes the concept of human adaptation to environment, but the various researchers place different levels of emphasis on the environment in this dialectic. Whitehead (1998) takes perhaps the strongest stance against environmental determinism, proposing a fully “historical” ecology, wherein human agency provides an independent variable in the equation of environmental dynamics maximizing the role of human decision-making (“ecological praxis”) in constructing the landscape. This is not meant to remove outside pressures, however. Localized cultural behavior practices are greatly influenced by cultural and environmental contexts, and broader cultural processes (regional interactions), environmental impacts of past cultural activity, and unstructured or unmodified
environments themselves act to affect change and effect decision-making at the local and individual level. Political-economic influences can have a massive impact in subsuming human agency in localized environments. Additionally, past actions of local populations (landscape degradation, soil enrichment, and other anthropogenic impacts on the environment) affect the resources available to future occupants. Within a historical context, the very nature of a changing environment and the varied properties of any given setting will undeniably interact with cultural processes to influence the range outcomes.

Outdated models emphasizing uniform infertility of soils across the broad landscape of Amazon are patently false. The view that *terra firme* adaptations were ecologically restrictive, dictating the absence (or at least the near absence) of sedentism simply reifies the “Tropical Forest Tribe” model (now some 60 years old). The Amazon, however, has been revealed as an extremely diverse environmental complex with a multiplicity of ecosystems. The original model, with its archaic belief in homogenous ecozones, has been proven insufficient to describe the diversity revealed through ethnographic study (Morgan Schmidt, a PhD candidate, Department of Geography, UF, is currently conducting pedo-archaeological studies related to the Upper Xingu project and a detailed discussion of ADE in the Upper Xingu is not given here).

Those that seek to make a distinction between the fertile soils of the varzea and the comparative infertility of soils in upland regions have also fallen by the wayside, suffering from generalizing principles when evidence suggests that simple divisions into varzea and terra firme cannot encapsulate the diversity of the Amazonian ecosystem. Numerous floodplains soils along “black water” or “clear water” rivers have demonstrated soil characteristics even poorer than many of the “impoverished” upland
soils (Peterson et al. 1999: 11). It is important to note that a number of this model’s proponents have acknowledged the likelihood of both widespread sedentary populations and social complexity in both *varzea* settings and in bluff regions along the margins of upland and riverine areas (Carneiro 1986, 1995; Lathrap 1970; Meggers 1996; Roosevelt 1980, 1994), but the idea that cultures exhibiting these traits were restricted to these areas is simply untenable.

The over-emphasis of the *varzea* as the sole area of agricultural potential has been endemic in much of the though governing the possibility of resource potential significant enough to provide a basis for sedentary populations and social complexity. In regions like the lower Negro and Upper Xingu, sedentary villages were supported by intensive *terra firme* agriculture (Denevan 1992; Heckenberger 1998; Heckenberger et al. 1999; Peterson et al. 1999). Additionally, the vast potential of aquatic resources, both in floodplain and upland areas would suggest that people would have been “pulled” to all riverine settings, not just those located in *varzea* floodplains (Carneiro 1995; Heckenberger 1998; Peterson et al. 1999). Even allowing for strict adherence to ecologically determinist principles governing the “carrying capacity” of the *terra firme*, certainly the combination of manioc agriculture and aquatic resources could have provided a substantial subsistence base allowing for much larger prehistoric populations that what we have observed in the ethnographic present. If we couple that with the evidence that past populations were actively transforming their landscape into a hugely productive agricultural zone through intensive occupation (producing the *terra preta* soils we find today), then we have a powerful model describing how Amerindian populations were not only selective in their use of the landscape (Heckenberger 1999),
but rather than degrade areas through intensive occupation, soils could be improved through both intentional and unintentional modifications.

An inescapable conclusion is that we can no longer view prehistoric settlement of Amerindian populations as small, impermanent, or autonomous groups barely scraping by an existence within a homogenous landscape. Nor can we allow the projection of the ethnographic present (replete with the disenfranchised, displaced, decimated, vastly impacted results of European contact and expansion) onto the past. We must also avoid the trap of strict adherence to old theories of environmental determinism. Certainly local ecology differentially constrains cultural development, but the relationship between culture and environment is a dynamic one. Culture can be molded to fit an environmental setting, and environment can be modified to better suit a particular cultural need. The central questions we must ask ourselves in adopting these deterministic perspectives are can we ever hope to adequately understand the full scope of variability of the Amazonian ecosystem as a whole, or be able to characterize the range of cultural adjustments to it? Can we ever hope to separate constraints imposed by the environment and contrast them to those imposed internally? The view promoted here demands that any statements of the casual nature of environmental factors on cultural change that do not address the reciprocal nature of the relationship should be greeted with skepticism.

**Importance of the Upper Xingu**

The Upper Xingu provides a unique example of an Amazonian cultural lifeway that supported large, densely settled, and integrated regional populations over the past 1000 years (Heckenberger et al. 2003). Archaeological evidence has shown (Heckenberger et al. 2003) evidence of large, well-engineered public works (such as plazas, roads, moats, and bridges) in and between pre-Columbian settlements.
This would seem to suggest a highly modified environment, having much more in common with other regional-scale complex prehistoric societies found elsewhere in the Americas than with the commonly held misbelieve of small-scale, localized populations constrained by their environment, and relegated to the lower echelons of cultural complexity. But can the archaeological evidence alone speak to the questions of to what extent the environments were modified, and if there was a pattern to the modification? In short, yes, but only with great difficulty. The nature of the region is such that archaeological research is time consuming. Site density, even within the relatively small geographic confines of the Upper Xingu archeological work conducted by Mike Heckenberger since 1996, seems substantial enough to place the feasibility of garnering either enough time or money to fully survey the entirety of that area, let alone the extents of the PIX, utilizing traditional survey techniques far beyond our reach. And yet the import of discovering that large-scale regional complex societies did, in fact, exist in this area, and the possibility of dramatically reversing years of researchers relegating tropical forest regions, such as the Upper Xingu, to something akin to a “cultural backwater” begs further investigation. The question is, how can we approach studies of this area in such a way to mitigate both the time and expense of traditional survey? Through the integration of new technologies, we can provide a means to augment current research, including archaeological and ethnographic fieldwork, with remote-sensed data analysis, GPS surveying, unified in a geographic information systems database, in an attempt to further illuminate the interplay of human populations and their environment in non-western tropical forest settings.
CHAPTER 3
PREDICTIVE MODELING AND GIS

Any sort of in-depth study seeking to address research question in Amazonia utilizing tradition field techniques will, at some point, become hampered by issues of accessibility and other logistical problems. The Upper Xingu is no exception, as the area is indeed difficult to access physically and academically, with little data available for the region. This makes the use of new technologies, including GPS, GIS, and remote-sensed data (especially satellite imagery) even more critical to a successful exploration of some of the issues mentioned in the previous chapter. Each of these tools can augment traditional archaeological exploration, with the additive effect producing a robust database of information, leading to more effective models capable of moving us closer to answering questions of long-term human-environment interaction, landscape change, and cultural development in the region. This dissertation sought to formulate an inductive predictive model (not as an explanatory device, but rather as a heuristic model) using vegetative signatures of known archaeological sites to extrapolate out to a wider scale of possible site locations within a specific region of the Xingu.

Defining the Model

Predictive Modeling and Archaeology

An archaeological predictive model is a tool that indicates the relative probability of encountering an archaeological site. Parker (1985) sees predictive modeling as a natural outgrowth of the theories and methodologies of spatial archaeology and predictive modeling has become the focus of a number of archaeological studies (e.g., Allen et. al.
predictive modeling is an avenue of research within archaeology that has gained prominence over the past two decades, specifically with the development of new technologies of remote-sensing, mapping, and GIS. Predictive modeling for archaeology is defined as a “... simplified set of testable hypotheses, based either on behavioral assumptions or on empirical correlations, which at a minimum attempts to predict the loci of past human activities resulting in the deposition of artifacts or alteration of the landscape” (Kohler 1988:33).

To implement a predictive model, a set of georeferenced parameters (factors) that are related to site occurrence is collected, stored, and manipulated. Thus, the underlying theoretical basis for modeling site locations is the hypothesis that environmental attributes can be correlated with archaeological site locations. Known archaeological evidence has to be acquired and classified into chronologically and typologically homogeneous groups. By confronting landscape and archaeological data sets, a heuristic or statistical model is then built up which links the spatial variability of such parameters with the occurrence of the sites (Jacoli and Carrara 1996). The integration of GIS, remote-sensing techniques, and GPS survey may provide new opportunities for identifying, analyzing, and interpreting archaeological sites makes it possible to both define the relations of known sites with the environmental context, and identify new sites. The theoretical and applied aspects of conducting predictive modeling in conjunction with archaeological studies are relatively new. They have their basis in studies conducted during the 1950s and 1960s, but remained the bailiwick of relatively few until they gained prominence during the late 1970s, 1980s, and early 1990s (coinciding with a surge
The literature concerning predictive modeling increased exponentially during that period of time (thus, much of my discussion is framed by the character of the literature of that period). And while, in the past, many archaeologists steadfastly refuse to accept the possible value of predictive models in determining possible site locations (Kohler and Parker 1986:396), with predictive models were viewed by many as an expensive exercise to discover the obvious, and regarded as suspect or unreliable or being limited in value (Kohler and Parker 1986:398), it is important to remember that archaeologists, on the whole, have moved beyond these myopic viewpoints. Most researchers would now acknowledge the utility of predictive models as an important element of contemporary archaeological investigations, with even greater potential in the future, particularly in little know and/or remote regions. In point of fact, most current modeling efforts have begun to develop entirely new methodological approaches into the modeling process, recognizing the importance of the theoretical contributions of landscape archaeology, historical ecology, and advances in thought regarding settlement patterning (Crumley and Marquardt 1990; Crumley 1994; Lookabill 1998; Witcher 1999; Chuchill 2000; Perkins 2000; Wise 2000; Wheatly and Gillings 2002).

The strength of predictive modeling is in providing a unified framework that includes testing and self-correction components, and contains an element that goes far beyond mere number crunching to arrive at an accurate depiction of the past. The mathematical “purity” of predictive modeling efforts (a result of at least two generations of archaeologists’ processual, positivistic approaches) is one of the greatest strengths of current modeling programs, allowing for a replicability of the model, and allowing for
“tweaking” of a model by altering simple (or sometimes complex) variables. However, strict adherence to only those variables that can be enumerated has led to an exclusion of those elements of culture that make it so unique. The archaeological past is composed of tens of thousands or more unique cultural groups: different times, different peoples, and different places. Each culture has its own unique history, its own chronology of events, and it is this uniqueness that gets lost in the midst of current predictive modeling efforts. Incorporating culture history into modeling efforts can breathe “life” into the model, adding new dimensions and increasing the ability of the model to accurately predict, or at least approximate, certain behaviors.

Models are partial representations of a theory, formulated in a manner that enables the archaeologists to test the theory by means of empirical data. When new patterns are found, hypotheses are formed to explain them, and models are built to test the hypotheses (Warren 1990). While there are many approaches to predictive modeling, all must choose between various kinds of units of analysis, dependent and independent variables, types of models and decision rules, and modeling testing procedures. The effective and efficient application of the predictive archaeological models of the past was severely hampered by the labor required to manually measure map variables in large-scale projects. In fact, past applications of high-resolution models were virtually impossible without restricting sample size and the range of variables investigated. Nearly all of these limitations have been successfully overcome in recent years, however, through the application of GIS.

There are two traditional ways to go about developing models of where prehistoric peoples located themselves. One approach lies in a close examination of the
anthropological and historical literature of a region in an effort to deduce the kinds of locations that past peoples may have selected to place their camps and settlements. Researchers must determine relevant variables and then arrive at specific values for each of those variables (a difficult task even in the present, infinitely more so when dealing with materials from the more distant past). For example, we can assume that nearly all peoples would build their residences on level ground, but exactly how level does the ground have to be? A permanent settlement would almost certainly need to be located near a secure water source, but how near, how secure, and of what type? Furthermore, other variables may come into play, such as defense, that make a narrow focus on single or even few variables unrealistic. Like wise, there are always exceptions to the rule (some people do live on slopes, over water, or position their settlements between, not near, water sources). The degree of freedom in selection of these variables, and the error tolerances inherent in the establishment of any fixed value for any one variable, could result in outcomes with widely varying results.

In the face of such issues, the task of accurately describing the past seems daunting. However, utilizing advances in both theory and methodology, it may be possible to gain insight into those environmental and cultural variables deemed of critical importance by the actual groups under study, thus revealing a more robust and more useful form of predictive modeling. First, any modeling effort should incorporate some element of archaeological field survey. By measuring and analyzing environmental variables at known archaeological sites, it is conceivable to ascertain which variables may have been critical in deciding on past site location with some degree of success. By coupling this process with extensive use of available ethnographic materials and analogy, particularly
within specific historic trajectories, one can more accurately determine which features may have been the most important, eliminating those that have little or no bearing for the culture concerned, simplifying the variables, and creating a more elegant model. The reason such studies have not historically been popular is that the amount of labor involved often makes them impractical. However, the advent of GIS has permitted virtually any sort of map variable to be computer-encoded, and combined with other variables to yield complex modeling outcomes over larger areas with more precision, better error correction capabilities, less labor, and the ability to easily model several different iterations of element values. Data updates and corrections, once slow and costly, can be conducted quickly and efficiently. Using this technology, measurements can be completed in a matter of seconds. The only limits to constructing complex simulations and models of behavior are limited only by imagination, as the time required for processing data sets has been relegated to insignificance (Kvamme 1989).

Many archaeologists of the period recognized the value of GIS-based analyses, especially in the realm of archaeological predictive modeling. For the most part, archaeologists would agree that the spatial distribution of sites is largely dependent on a wide spectrum of features (such as landform, soil type, water proximity, vegetation cover, climatic conditions, etc.) that characterize the environmental context where sites are located. Throughout the later 1980s and 1990s, predictive models were increasingly applied in investigations attempting to both explain the spatial distribution of sites already known, and predict where new sites are most likely to occur (Kvamme 1989; Warren 1990). Still, it is critical to recognize the limitations of a GIS, and of predictive modeling within a GIS environment. “Predictive models are probability statements; they
are not “facts,” and cannot substitute for facts in any application requiring the use of hard data about specific individuals as decision making criteria” (Wildesen 1974 in Kohler 1988).

Judicious use of GIS can give practical integration of spatial structures (habitations, soils, river drainage), but to gain a real understanding of past and current relationships among environmental and human systems requires that archaeologists research culturally specific temporal and spatial perspectives. Much current research in GIS and archaeology revolves around ways to incorporate landscape perspectives as well as spatial archaeology. The development of more methodological approaches to the integration of GIS requires a fundamental understanding of how a GIS enables the creation, organization, and management of multidimensional archaeological data sets. Vector GIS is essentially a simple relational database wherein records describe the attributes of a real world geographic entity that is then linked to a geographically referenced digital representation of vector information (point, line, or polygon) (Maschner 1996). A GIS commonly incorporates several different types of geographic entities. Similar mapped geographic entities are generally organized and stored in their own georeferenced layers. When georeferenced layers are placed in the same projected coordinate system, the layers can be overlain, allowing the researcher to investigate the relationships and patterns between phenomena by examining and analyzing alternate combinations of layers, the response to a query, or the results of an analysis.

All archaeological data has a spatial aspect; individual artifacts have provenience, features are located within a site, sites are part of a mosaic of habitation and activity areas, travel routes and borders, and all are located in, and affected by, diverse
environments. How one manages these data can have a dramatic impact upon the efficiency and practicality of attaining research goals. Two benefits of using GIS for archaeological data management are the rapid creation of maps to visualize the archaeological record and the ability to query the database using spatial parameters. The scale of the GIS can range from individual excavation units to regional or national registries.

More than simply a two-dimensional map design program, the real power of GIS lies in its modeling abilities. The general mapping capabilities of GIS are a basic function, and although it is generally more efficient than manual or computer assisted means, it only begins to address the potential of GIS for archaeological data management. These visual representations of archaeological data in a georeferenced map can be used to produce archaeological "sensitivity" maps that indicate which geographical locations are more sensitive than others for cultural resources. The dependability of such predictability models is a function of their performance. This can be examined and tested by comparing the results of a predictive model to archaeological field survey results. By comparing model predictions against known archaeological site locations, it is possible to determine, with specifiable confidence, how accurately a model performs. In fact, this very approach gives us confidence in a model and allows us to use it as a predictive tool.

**Differing Approaches to the Modeling Problem**

Prior research and settlement theory demonstrated that open-air site placements were most often a function of a matrix of environmental factors that have been found to be quite consistent from study to study (e.g., Jochim 1976; Roper 1979; Shermer and Tiffany 1985). As a rule, the variables chosen are restricted to those that reflect relatively stable landform characteristics through time, such as elevation, slope, and aspect, to
insure that there is some correspondence between modern map-measured data and the prehistoric-early historic environment. Some potentially important variables, such as plant community composition and water table elevation, are notoriously sensitive to climatic changes and, as a result, are difficult to use without recourse to proxy measures (Kohler and Parker 1986: 415).

Since the particular environmental variables most suited to a particular model depend in part on the physical nature of the region under investigation and cannot be determined completely a priori (that is, without analysis), most modeling projects initially measure a relatively large number of landform, hydrological, soil, and geologic characteristics, including slope, aspect, elevation, local relief, landform type, horizontal distance to the nearest permanent water and stream confluence, and distance to streams. Characteristics with low predictive power are filtered out throughout the model development process. In large areas, some characteristics may prove to be important in only some sub-areas, necessitating the construction of multiple models. Justification for the adoption of specific characteristics and procedures for operationalizing their measurement can be found in Hasenstab (1990), Kvamme (1986), Kvamme and Kohler (1988), Parker (1985, 1986), and Roper (1979), among other authors. General reviews of the issues involved are provided by Judge and Sebastian (1988), Kohler and Parker (1986), and Kvamme (1990).

Kvamme likens inductive predictive models to the supervised classification of remote sensed imagery: (1) training samples representing the locations of known archaeological classes of interest (e.g. classes representing different types of prehistoric activity, time periods, cultural types, or amounts of activity) are established through on-
the-ground field survey; (2) a statistical or other classifier is developed for these classes, including an archaeologically “empty” class, based on patterns exhibited by the classes on the GIS encoded environmental data; (3) the classification rules are applied on a cell-by-cell bases by the GIS to classify the entire region of study, where each cell is assigned to the archaeological class to which its environmental characteristics are most similar, or the empty class; and (4) the accuracy of the classification is assessed through test samples of known class membership. The classification of the region resulting from the approach constitutes a location model of the region’s archaeological phenomena as defined by the sample data (Kvamme 1989).

Although this "empirical correlation" procedure was, by necessity, used in the formation of earlier predictive locational models, the importance of social and political factors in the spatial location of settlements have been recognized and incorporated into more recent efforts (Crumley and Marquardt 1990, Madry and Crumpley 1990, Madry 1996), and, arguably, have been in place since the advent of central-place theory (Christaller 1966) and earlier gravity models. As a result, consideration of land-use choice derived from “habitual behavior” derived from cultural norms, traditions and spiritual proscriptions became the norm in developing models, rather than an overriding consideration of the economic attractiveness of a specific locality (Kohler and Parker 1986:435 citing Wright and Dirks 1983). Factors related to actions having little archaeological visibility, such as spiritual influences, may have resulted in activities being located in less "typical" locations. Choice of activity location may also be the result of historical events that override environmental considerations. This is where the practice of forming predictive models becomes less science and more of an art.
Researching historical events within a culture, becoming intimate with the ethnographic material, and incorporating it into the variable-forming process is perhaps the most difficult part of model building. This requires extensive data, and even that may not be enough to fully encapsulate even a small range of behaviors.

Despite the conundrum of establishing ties to the cultural history of a site, some archaeologists have made valiant attempts to bring the ethnographic past into their modeling efforts. Flannery (1976) and Reynolds (1976) discuss social factors that condition site placement. Jochim (1976:12) details criteria of economic relevance and assumes that "the determination of resource use tends to precede and condition the site placements and demographic arrangements of a hunter-gatherer group."

There are as many procedures developed for quantitatively determining the correlation between environment and site presence as there are researchers working on such problems (see Warren 1990, Kohler and Parker 1986, Carr 1985 for some of the more highly utilized methodologies). Of particular interest are logistic regression (numerical approach), weighted value approaches, and Dempster-Shafer theory, also known as weight-of-evidence modeling. Logistic regression’s established use, broad literature, and numerous case makes it particularly well suited for predictive modeling endeavors. Predictive models utilizing a numerical approach employ multivariate statistics as a means of identifying associations among variables. Because statistical methodology discovers meaningful associations among variables from known site information, it is important that the known site information be representative of the site population. The most intricate and complex probabilistic designs are of little use if the population sample is not the same as the target population. It is then beholden upon
archaeologists employing the numerical approach to ensure the validity of their assertions by carefully evaluating the nature of the existing database. The data must be representative, and, perhaps more importantly, the researcher should determine whether known site locations reflect the actual distribution of archaeological sites, or simply reflect where archaeologists have conducted their surveys.

It is also critical for the archaeologist to recognize that the physical and cultural environment has changed over time (a careful consideration of a diachronic perspective). Too often, we are lulled into a synchronic view of the environment. To the contrary, the environment is dynamic, and it is only with an eye towards a diachronic study that we can even begin to comprehend the past (changes in environment may well have affected the choice of activity location through time). Kohler and Parker state that:

“... Despite numerous studies in diverse areas indicating change in site location through time in response to changes in adaptation type, and despite evidence that within any adaptation type, functional subsets of sites may have differing environmental determinants, most empirical correlative models aggregate sites of all types and ages together for prediction" (1986:408).

This remains a chief problem with numerically derived models. Often times these models show a lackadaisical treatment of, or completely fail to address temporal considerations. There are some researchers that have chosen to avoid the issue of time altogether, and develop generalized models that, by their very nature, are extremely limited in their applicability (and perhaps not of any use at all). Other researchers have sidestepped temporal considerations by suggesting that discernible patterns of human behavior crosscut considerations of time (Kvamme 1992:23).

The numerical approach certainly retains value as an approach to predictive modeling, allowing the researcher to uncover associations between site locations and variables. The drawback to this approach, however, is that it requires a high degree of
statistical training and competence in order to develop the model, interpret the results, and replicate the outcome. Invariably, a great deal of interpretation is required to relate the results to real-world geography. More often than not, the generalizing nature of the models make them unsuitable for addressing issue of localized cultural practice (thus of little use to researchers interested in specific peoples at specific times). While the numerical approach can and is still used to generate valid predictive models, many archaeologists have turned to different approaches to help eliminate some of these complications.

The weighted value approach, or “graphical approach,” details the development of predictive models utilizing environmental variables, but distinguishes itself by using a graphical methodology derived from map overlay techniques. Advances in GIS applications have allowed researchers gather data from different variables represented on separate computer map layers. These map layers can be combined in ways that can readily identify areas spatially associated with landscape characteristics determined to be pertinent to the questions being asked by the researcher. Utilizing different combinations of variables, the researcher is able to address various stages in the predictive modeling process, fine-tuning the outcome of the model. Database queries of overlain data sets allow researchers to extract those geographical regions that contain desired associations between variables and sites, which can then be evaluated using statistical techniques. Statistical techniques are no longer used as the primary means of discovering sites, but rather as a method of evaluating the strength of the association between variables and sites after the model is applied.
Within these parameters, models either are physically generated by an intersection or weighted value method. The intersection method begins with the basic assumption that all variables used in the generation of a predictive model contribute equally to the determination of site location potential. Calculating high, medium, low potential areas is simply a process of determining where the greatest number of variables that converge in a given location. The weighted value method begins with the basic assumption that each variable contributes differently to the final determination of site location potential. This is accomplished by developing and applying a weighting scale that effectively ranks variables numerically. Site potential is determined by the arithmetic addition of all variables. Areas of high potential will have the largest numeric values and areas of low potential will have smallest numeric values. During the development of numerical predictive models, a number of issues must be considered. These include the “representativeness” of the variables to the behaviors being modeled, the quality of databases consulted, the scale at which modeled should take place, and the manner in which potential is presented.

Weight-of-evidence modeling, perhaps, holds great potential for further investigation, as it is perhaps better suited to the uncertainties associated with archaeological data and environments that cannot be directly observed. Dempster-Shafer theory is a “soft” decision support tool that works extremely well with Fuzzy Set theory or Bayesian statistics. It allows for and deals with incomplete datasets, expert knowledge, anecdotal evidence, experience, and ignorance. Each line of evidence used in the model (like the proximity to fresh water resources) supports one of three hypotheses: (1) that sites are located near water (site), (2) that sites are not located near water
(nonsite), or (3) that we are uncertain about which hypothesis the evidence supports (site/nonsite). Each line of evidence is weighted and a cumulative probability map generated. Conversely, logistic regression analyses can be considered a “hard” decision tool because it does not account for fuzzy datasets or uncertainty; it is assumed datasets are ‘perfect.’ Although this can be seen as a detriment, the statistical robustness of logistic regression makes it a more clean-cut predictive tool. Logistic regression has advantages over other statistical methods such as discriminate function analysis; (1) it has less restrictive assumptions; (2) tends to be a more powerful and consistent; (3) accepts a mix of scales of data (nominal, ordinal, interval and ratio) (Warren 1990).

Both univariate and multivariate statistical models are commonly used to identify variables on which the distributional differences of dependent variables (resources present/absent) are most pronounced. A variety of statistical tests are performed on the univariate descriptive statistics to elicit trends, and a logistic regression technique is most often used to explore multivariate differences. Many researchers have adopted multiple logistic regression models for analysis. The rationales are varied, but primarily center on the fundamental nature of such models in that they make no assumptions about the distribution of the data (representing a nonparametric technique), are robust classifiers regardless of distributional form (an essential attribute in environmentally diverse regions), and can handle nominal, ordinal, and interval level independent variables. Although multiple discriminant analysis, maximum distance classifiers, quadratic classification procedures, and maximum likelihood distance classification techniques all have their adherents (e.g., Bradley et al. 1986; Custer et al. 1986; Kvamme 1983, 1985,

The findings of these two different modeling strategies are used, at least initially, to formulate decision rules (in essence, describing weights of environmental variables that indicate whether archaeological resources are likely present or not). These decision rules can then be applied to any unsurveyed area to determine whether the model specifies that a site will be present. Since a fundamental issue in locational modeling is in determining what weight or rank should be applied to each of the independent variables investigated, a decision point will be selected by calibrating the decision rule at first to sample data (Kvamme 1988). Presumably, the application of decision rules captures a pattern left behind in the archaeological record. In archaeology, a locational model is essentially a decision rule that assigns specific areas to certain classes based on environmental or other non-archaeological characteristics of the area under question. The model predicts archaeological sites when it assigns value to an area of unknown archaeological class membership representing either a presence or absence or archaeological components. This pattern recognition and classification methodology facilitates exploration by abstracting the environmental patterns exhibited by surveyed regions that contain archaeological resources and mapping them across unsurveyed regions through GIS.

GIS is notorious for the amount of initial processing that is required unless the researcher is fortunate enough to acquire existing digital data (although it may be a wolf in sheep’s clothing with uncertain data quality, undesired projections and compatibility issues make using pre-existing digital data troublesome sometimes). At issue is the fundamental truism that raw archaeological data is can be extremely hard to come by in
many areas, much less data sets in a digital format. Analog maps have to be scanned or digitized, contour lines need to be generated to interpolate a digital elevation model (DEM) (Kvamme 1990), and all layers need to be georeferenced so they share the same projection and datum. Even data in digital form almost always needs to be processed in some way, either to make it compatible with a specific GIS format, or to classify or otherwise restructure the data itself. Once all the data is in the GIS, formatted, georeferenced, and processed the actual analysis can be conducted.

**Model Testing**

Simply stated, an archaeological predictive model is simply a set of decision rules that assign areas in a study to one of a number of mutually exclusive categories based on environmental or other non-archaeological characteristics of the locations. Assessment of model performance and accuracy are obviously necessary, for, at the very least, a predictive model must be able to perform better than random chance alone. Model testing involves the determination of the *a priori* or chance probability of the occurrence of certain archaeological events and an independent test of the model’s effectiveness against this probability. Presumably, the identification of the key non-archaeological characteristics of particular locations that are associated with the presence or absence of an archaeological resource is a guarantee that the predictive model will be more effective than the random model, but it must be demonstrated that this is so. In addition, a good test will specify the degree of effectiveness of the predictive archaeological model over the random-chance model. An *a priori* probability is the probability that a random geographical location does or does not contain archaeological resources. As a random-chance locational model, it provides a baseline that helps define what other models must accomplish. In regional studies, random-chance models can be calculated by determining
the relative frequency of the presence or absence of resources in a random sample of
surveyed areas (e.g., Kvamme 1983, 1988; Parker 1985:187). For instance, if 100
surveyed locations contain resources and 900 do not in a surveyed sample of 1000, then
the probability that a given location contains resources by chance is 0.1 or 0.9 that it does
not. Since the probability of correctly attributing a location positive for resources is no
better than chance, these probabilities can be considered random-chance.

The predictive power of a model is determined by calculating its percent of correct
predictions in the test sample and comparing this percent with the likelihood of a correct
prediction by chance alone. These calculations determine the model's specific percent of
predictive accuracy over chance alone. The main method of assessing model
performance in archaeology is usually some form of cross-tabulation that compares the
actual and model assigned presence or absence of resources. One of a number of
statistical tests can then be used to determine the significance of these frequencies (e.g.,
Congalton et al. 1983; Kvamme 1988, 1990). Since the performance or accuracy of the
model is evaluated statistically, field data must be collected within a sampling framework
that utilizes the basic principles of elementary probability theory. This foundation also
gives a model the ability to assign a probability to the occurrence of archaeological
resources in a given area.

The effective and efficient application of early predictive archaeological models
was crippled by the sheer volume of labor required to manually measure map variables
(Judge and Sebastian 1988). For all practical purposes, the application of high-resolution
models was impossible without severely restricting sample size and/or the range of
variables investigated. The advent of GIS, however, has allowed modelers to overcome
these constraints by automating much of the entire process (e.g., Kvamme 1986, 1989, 1990). Unlike traditional database management systems, GIS has a spatial or mappable component that allows the capture, efficient manipulation, analysis, and storage of geographical information. In addition, GIS is easily capable of producing maps of this information in various formats on a video monitor or on paper, and, because the information is coded electronically, it is very easy to update and improve models in relatively short timeframes.

Refining the Model

Criticisms of GIS and Predictive Models

There have been a number of criticisms of archaeologically-based predictive modeling efforts utilizing GIS as the primary analysis tool. As mentioned before, the power of GIS lies in the visual display of volumes of information, and the ability to make such databases infinitely more accessible. However, the graphical power of GIS can create its own compounded conundrum of issues. Models generated in GIS environments can easily become just another “pretty picture,” completely bereft of theory or explanatory power. The main mechanism for many inductive locational models is the statistical analysis that determines which environmental variable(s) are indeed influencing site selection, or at the least, correlate with site presence. Thus, many predictive models generated using GIS reek of an environmental deterministic approach, and, even worse, seem static and do not adequately reflect the dynamic adaptability of humans, or the dynamic nature of the natural environment. This speaks to the accuracy of such derived modeling approaches. Determining the accuracy of a model involves comparing virtual indicators to actual circumstances. Ideally, development of a model should be from a random archaeological site sample so that inherent biases are removed.
Sample-based modeling approaches in archaeology face ubiquitous problems that limit the predictive accuracy of models and should be considered when using modeling results. A key problem is the accessibility of archaeological resources. In any region, many archaeological sites will have been destroyed by erosion or human activities. Other sites will be deeply buried, well hidden in sealed rock shelters or under dense vegetation, covered by towns or lakes, on property to which access is denied or so small as to easily fall between transects in a field survey. Predictive models, because they are necessarily based on sample survey data, are only sensitive to the types of archaeological resources included in the initial samples. This means that they are usually only sensitive to certain types of surface distributions, for the distribution of geologically buried sites is rarely explored in a systematic manner.

A second problem is the difficulty in archaeology of satisfying statistical assumptions, such as the requirement of multivariate normality or homogeneity of variance. For this reason, modelers usually remain somewhat skeptical of statistical indicators of the importance of independent variables in the developmental phase of model building and employ robust mathematical procedures to identify decision rules. It is in the better-controlled testing phase where the requirements of statistical assumptions can be more fully met that statistical inference and probability theory play their primary role.

A third problem is the presence of patterned variation (i.e., spatial autocorrelation) in the distribution of archaeological phenomena. Its existence violates the assumption of independent observations and generally results in overestimates of the significance of
independent variables. This problem can be partially controlled by adopting a sampling procedure that widely separates surveyed parcels of land.

Many other kinds of problems exist. The importance in site location of social and political factors ("sociocultural noise"). The difficulty of considering such factors is one reason why nearly all predictive archaeological models have accuracy rates less than 85 percent. It is also the main reason why field survey must remain an integral component of cultural resource management.

The most common critique of archaeological predictive modeling is that it is often not grounded in anthropological or archaeological theory. Predictive models have been primarily descriptive, for the most part confirming what we already know. Trends in modeling efforts have led to simple aggregations of sites and the flattening of the temporal dimension resulting in models that adequately predict the average known site, but fail to address the maximums of variability. In a standard distribution of sites, a large proportion of sites will fall within a middle-range, which we can address as “type-sites,” or, more to the point, “common.” However, it is the uncommon, or non-type sites from which archaeologists may be able to glean the most information. Because a large number of current correlational models are based on observed archaeological sites, they will inevitably be biased towards predictions of sites about which information already exists rather than for “outliers,” which may have a greater potential for new information. By associating sites representing many different functional, chronological, and cultural types into a single open-air class, a great deal of locational, or cultural, variability is introduced to the modeling problem. Nevertheless, in past literature, some researchers asserted that
there were common locational tendencies that may crosscut functional categories, such as preferences for level ground or proximity to water (Kvamme 1985).

Predictive modeling has become critical as a means of identifying landscape variables that are consistently correlated with known site distributions. By identifying these correlates, researchers are better able to identifying uninvestigated localities that have a high probability of containing sites based upon their geographic similarity to known sites. There is a danger in this, however. Simply identifying new site locations based upon the attributes of known site locations is not really making progress in the investigation of the unknown. Instead, we are simply investigating more of the same sorts of sites, with the added bonus of identifying areas that have not yet been surveyed (we are merely modeling existing assumptions and expectations). This is not to say that such undertakings are any less important than other pursuits, but I submit that we, as archaeologists, can take it one step farther. For us to make predictions of the unknown, we must step outside what is “expected” and employ a modeling rationale that does not build exclusivity into its results. Thus, the underlying flaw in correlation models is exposed. Such models are exceedingly good at illustrating the probable location of any number of like sites based upon an approximate “type,” but without a more substantial theoretical foundation, they cannot be expected to produce information governing “why” or “how” such sites came to be.

The importance of infusing some measure of anthropological or archaeological theory into modeling efforts remains a critical goal. Butzer states, “When the intellectual framework is too narrow, the results, no matter how elaborately programmed, cannot hope to allow high level interpretations” (Butzer 1978). And Ebert states “predictive
modeling will be transformed into a worthwhile adjunct to archaeology and archaeological thinking only by the formulation of a body of explanatory propositions linking contemporary correlations with the past. In other words, it is productive, explanatory thought, and not computers, that can potentially raise predictive modeling above an anecdotal level (Ebert 2000).

Our ability to generate predictive models rests on two fundamental assumptions: First, that prehistoric people made settlement choices based on particular characteristics in the natural environment; second, that those chosen environment factors can be mapped within the modern environment in a given area of interest. With this in mind, and with a sufficient sample, it should be possible to distinguish between places where archaeological sites should or should not appear. Thus, in order to implement a predictive model, first a set of georeferenced parameters related to site occurrence must be collected, stored, and manipulated. Known archaeological evidence has to be acquired and classified into chronologically and typologically homogeneous groups. By combining both landscape and archaeological data sets, a heuristic or statistical model is then constructed, allowing the archaeologist to link the spatial variability of such parameters with the occurrence of the sites (Kvamme 1989).

Our ability to grasp and accurately measure the variability between elements of site and non-site areas is critical for accurately modeling settlement patterns across the landscape. These predictive models help not only to location possible sites, but they make it possible for land mangers to get a sense of the expected distribution of the resources under their care. The goal, then, of predictive modeling is to establish a correlation between certain environmental parameters and known archaeological site
locations, build a statistical model based on that relationship, and apply the model to unsurveyed land. GIS remains the most capable tool for performing these tasks. GIS has spawned a revolution in spatial thinking, making a fundamental change in the way human spatial behavior is studied.

The physical space occupied by a population serves as the primary frame of reference by which they locate themselves and their varied activities relative to all other occurrences. There are many cultural influences upon space and place: communication routes, personal taste, landscape history, proximity to other sites, aesthetics, and ritual, to name but a few. They have been omitted, for the most part, because they are unknown in sufficient detail to allow for their operationalization in a GIS framework. If the archaeological record was that complete, there would be no need for predictive models in the first place. However, this is not to say that there is no possibility for either a stand-alone cognitive model or a coupled cognitive-environmental model. “If people’s actions are systematically patterned by their beliefs, the patterning (if not their beliefs, as such) can be embodied in the archaeological record” (Renfrew et al. 1982:11). Research (like viewshed and cost surface analyses) continues to delve into how human behavior could be incorporated in a GIS framework.

What Drives the Model?

In essence, the necessary information needed to generate accurate modeling projects depends on the purpose of the project. In most academic projects, the goal is to model the locational behavior of different functional, chronological, and cultural types of occupations (components). By contrast, the goal of most cultural resource management projects is to conserve resources and limit cost by identifying areas with and without resources regardless of the nature of the resources themselves. Given this goal, and the
difficulty involved in clearly identifying meaningful functional and cultural types of occupations that are securely anchored in time in most archaeological sites, it is not surprising that, in early modeling attempts (in the 1970s, 1980s, and 1990s) the most frequently used dependent variable in these contexts were relatively simplistic, binary cases of the presence or absence of archaeological materials (e.g., Bradley et al. 1986; Kvamme 1984, 1986, 1990; Parker 1985; Warren 1990). Because these approaches lumped occupations of various kinds together, it incorporated a great deal of locational variability that reduces the potential predictive power of those models (e.g., Judge 1973; Roper 1979). However, such approaches had the advantage of minimizing complexity by focusing on defined events that form a mutually exclusive, exhaustive, and non-ambiguous partitioning of the region being investigated and of producing large sample sizes, because of the use of the single “resource present/absent” class.

Such approaches also depended on common locational tendencies that were perceived to crosscut functional and cultural categories, such as proximity to water and preference for level ground, and that many locations in a region were unsuitable for most kinds of activities for similar environmental reasons, such as the presence of swamps or very steep slopes (e.g., Kvamme 1985; Kvamme and Jochim 1989). Many powerful predictive models have been built using this simple solution to defining all the possible events that can occur in a land parcel (e.g., Kvamme 1989, 1990; Parker 1985).

Other choices of dependent variables in early locational modeling efforts included multiple site types (e.g., Kvamme 1988; Parker 1986), counts of artifact density (Green 1973; Nance et al. 1983; Zubrow and Harbaugh 1978), and various measures of site significance (e.g., James and Knudson 1983). The advantages and disadvantages of these

A variety of independent variables have been used in archaeological models of locational behavior, including sociocultural and radiometric (e.g., Custer et al. 1986) characteristics and positional parameters (Parker 1985). Historically, many modeling projects have focused on the economic component of site location, as environmental factors were generally considered intimately related to locational decisions by groups without advanced transportation. It is the archaeological signatures of these populations that were the main focus of many cultural resource management surveys during the past few decades in most parts of North America (e.g., Jochim 1976; Wood 1978). The crux of the argument historically centered on the belief that these types of societies placed emphasis on economic transactions within a regional environment and that these populations tended to minimize the time and effort they expended in these transactions (Kvamme 1990: 271). The effect was to encourage location close to important environmental resources. The focus on environmental or biophysical characteristics of geographical locations (such as slope, soil type, elevation, plant community type, and distance to water) is also a practical one as these variables are relatively easy to identify today through measurements or observations made on maps, aerial photographs, remotely sensed data sets, and even computer-generated spatial information sources, such as GIS. Environmentally based predictive locational models work by correlating the location of a sample of sites with the environmental characteristics of the land parcels they are in and predicting that other, unknown sites will be present in parcels with similar sets of
characteristics. The goal is to define those characteristics of physical locations that have some bearing on the distribution of archaeological resources in a study area.

In some ways, these approaches to modeling mirrored the processual/post-processual debate that still rages within archaeological circles. While this study does tend to emphasize a recording and analysis of pattern in a processual, positivistic way, it is by no means limited to the theoretical bent of the Processual School proponents. In fact, much of this dissertation is grounded in a historicist, post-processual approach. The middle ground, however, seems to be the best place to situate this research.

Within archaeological regional or landscape studies, current or in the past, with or without the use of GIS, the general approach has been concerned with discerning and interpreting patterns of archaeological land use and settlement. The archaeological record certainly retains pattern, but pattern can also be found in numerous factors affecting our interpretation of the record (namely erosion and deposition, land use, and research bias). This raises the specter of how can we interpret the archaeological record if we cannot separate the patterns we wish to illuminate from the “background noise” wrought by other processes? If human patterning is only one of a number of factors determining the patterns we find in the archaeological record, what hope do we have of modeling these processes, and what, in the final analysis, does this say about the models we develop?

An examination of the literature reveals some of the most basic environmental variables used in predictive models: elevation, slope, aspect, and distance to water (Kvamme 1985; Parker 1985; Altschul 1990; Carmichael 1990; Warren 1990). Culturally relevant environmental variables can actually be derived from any number of
sources (ethnographic analogy, Jochim 1976; simulation models, Gunn 1979; impressionistic models, Gardner 1978; correlation studies, Bettinger 1977; or linear programming approaches, Keene 1979). Dean (1983:11) has raised an issue that, rather than attempting to find areas that match a number of criteria, people may actually look for only a select few critical variables in their surroundings when identifying and selecting activity locations. Many would decry such an approach as environmentally determinist, especially in light of pendulum-like swings of archaeological theory from the scientific, rule-finding approach of the 1960s and 1970s to the humanistic, historicist approach of postmodernism in 1980s and 1990s where many believed that such approaches were a thing to distance oneself from for fear of rebuke. This highlights concerns about the analysts' ability to choose “meaningful” environmental variables for inclusion in the modeling process.

Meshing of Archaeological Theory with Predictive Models

Anthropologists and archaeologists are acutely aware of the spatial aspect of the material culture they aggressively utilize in cultural reconstruction efforts. However, the methodological approaches to concepts of space and place have been inconsistent from one theoretical paradigm to the next. 19th century Diffusionist thinking resulted in the formation of culture area types, such as the regional divisions laid out by Alfred Kroeber. A fundamental shift in theoretical approach resulted in many archaeologists adopting a methodology of charting similar artifact traits to generate spatially and temporally defined archaeological traditions and phases. In Europe, the Austro-German school of anthrogeography (1880-1900) introduced the notion of the “Kulturkriese”; a formalized methodology of mapping cultural behaviors and material cultural initially over large regions, and, with later refinement, smaller spatial scales. Spatially considerations
remained prevalent in European archaeology. “Comparative analysis of archaeological distribution maps had become a standard, if intuitive procedure in European archaeology” (Clarke 1977: 2). Archaeologists formally trained or teaching in geography, such as O. G. S. Crawford and H. J. Fleure, published archaeological findings in geography publications including the Royal Geographical Society’s Geographical Journal and National Geographic. In the 1930’s, C. Fox’s “Archaeology of the Cambridge Region” combined archaeological and environmental distribution maps over time similar to modern GIS-assisted research. By the 1930’s and 1940’s, American anthropology had enthusiastically begun to reject diffusionist theories in favor of re-emerging evolutionary explanations. “Anthropology turned inward and sought to demonstrate the roles of history, place, and locality as the primary means by which an understanding of human cultural diversity could be appreciated. “Space, thus, became passive and sterile as an analytical concept” (Aldenderfer and Maschner 1996:6).

Ecological anthropological theory and the adoption of the ecosystem concept (Julian Steward) was key to the reintegration of spatial thinking in anthropology, at least at smaller scales (Aldenderfer and Maschner 1996:7). Again, paradigmatic shifts in theoretical approaches drove archaeologists to place increasing emphasis on spatial analysis as central aspect of their research, and a fundamental means of answering archaeological questions. Willey, in his classic Prehistoric Settlement Patterns in the Viru Valley (1953), defined settlement patterns as “the way in which man disposed himself over the landscape on which he lived. It refers to dwellings, to their arrangement, and to the nature and disposition of other buildings pertaining to community life” (Willey
1953). Ecological spatial variability was soon incorporated into more generalized settlement pattern research as its popularity grew.

The emergence of the New Archaeology (and with it a drive for quantitative, scientific means of explanation) drove neo-scientific archaeologists to seek answers outside their discipline. They integrated concepts such as Von Thunen’s economic distance, Christaller’s central place, and Chisholm’s catchment area in an effort to standardize the practice of spatial analysis, to produce a more formulaic means of approaching such analysis (Clarke 1977). Clarke’s 1977 seminal treatise Spatial Archaeology (as well as Models in Archaeology (1972) and Analytical Archaeology (1968)), placed spatial studies at the forefront in methodological approaches to answering archaeological questions. Clarke defined “spatial archaeology” as:

“the retrieval of information from archaeological spatial relationships and the study of the spatial consequences of former hominid activity patterns within and between features and structures and their articulation within sites, site systems, and their environments: the study of the flow and integration of activities within and between structures, sites and resource spaces from the micro to the semi-micro and macro scales of aggregation” (Clarke 1977:9)

The growth of landscape archaeology from its nacent form (spatial archaeology, which focused on the spatial analysis, as well as examining demographics, as well as social and economic interaction) drew archaeologists to examine issues revolving around social aspects of the land, at how people perceived the landscape. “Landscape is the spatial manifestation of the relations between humans and their environment” (Crumley and Marquardt 1990). Sites were not viewed as independent units, but rather, as a part of a network of habitation areas, roads, ritual spaces, rivers, landforms, and resource extraction sites “that societies use and imbue with meaning” (Crumley and Marquardt 1990). Landscapes were relative, particular to the individual or group of individuals
Landscape archaeologists were concerned with the social creation of place, not necessarily the analysis of space (Knapp and Ashmore 1999).

New approaches are allowing researchers to examine the dynamic interchange between humans and their environment. Often, anthropologists have assumed that "culture has triumphed over nature." This conclusion is simply “wrong-headed” and overly simplified, and harkens back to an outdated nature-culture dualism. As anthropologists, we are constantly bombarded by introductory anthropological texts that tell the story of human evolution in environmental terms, and further confiscating readers by separating our “evolution” from our “history” and denying the environment a meaningful role in the development of culture. Instead, values, beliefs and issues, history, and culture constitute the key elements of the explanatory framework. Too often, we find a schism within anthropological circles: in one camp, cultural materialists and other environmental determinists, while diametrically opposed are groups of anthropologists who claim the transcendental nature of culture to rise above the environment within which it is placed. The first is an exercise in humility, denying the interaction of man with his environment and the capacity of human cultures to alter their landscapes in dramatic ways. The second is the ultimate form of hubris, disavowing the massive controls and constraints nature can put upon culture. While few are actually at either end of this theoretical spectrum, many anthropologists find themselves sliding towards on end or the other. Historical ecological approaches have actually taken steps to put us in the center of this continuum. Historical ecology is merely a framework of perspective, allowing researchers to investigate the recursive nature of human-environment interaction (Crumley 1994).
Integration of Landscape

Historical ecology gives us a powerful tool with which we can modify static models and make them dynamic. Incorporating a study of dialectical human-environmental relationships within a given study area requires interactive long-term sequences and a study of changing landscapes. This integration of landscape is critical in our development of a working model. The landscape is a human construction of the environment, imbued with cultural and social significance and giving an area a sense of “place” (Crumley 1994; Tilley 1994; Balee 1998; Knapp and Ashmore 1999). In today’s terminology, the concept of landscape has begun to incorporate socio-symbolic dimensions, allow researchers to envision the landscape as “an entity that exists by virtue of its being perceived, experienced, and contextualized by people” (Knapp and Ashmore 1999:1).

This broader definition of landscape has allowed anthropologists in particular to stress the relationship of people to their environment both horizontally through space and “vertically” through time. Thus, evidence for the historical interrelatedness of humans and environments may be read in the landscape. By identifying the mark that human cultural practice has made upon its environment, changing human attitudes may also be identified and their effects studied. By utilizing historical ecological approaches, we surpass the limitations of simple landscape ecology (the study of structure, function, and change of a heterogeneous land area composed of interacting ecosystems) and move into a framework within which we can conduct a study of past ecosystems by charting the change in landscapes over time, and utilizing data recovered for these studies to add an element of history to our selected predictive modeling variables. The symbolic structures that form an environment do not operate independently of the people who conceptualize
them. Attempts to resolve a symbolic or ideological landscape without a discussion of the “practice” of the individuals inhabiting that landscape quickly become impossibly convoluted; ultimately remaining unresolved (simply a “you can’t get there from here” problem).

What implications does this have for the future of predictive models? We need to decide the scale at which we feel most comfortable applying our model. Regional models may posses generalizing principles that allow us to understand human interaction at a large scale, but the details of that interaction become fuzzy due to the resolution of the model. More accurate studies can only be performed utilizing specific perspectives. This dramatically increases the accuracy of the model, but vastly reduces its generalizing potential. It becomes an exercise in selecting the lesser of two evils based upon the questions being asked by the researcher.

Our job may not be a futile as it sounds. Granted the above choice sounds like a difficult one to make. In either instance, however, the introduction of historically informed environmental analyses into such studies offers an important opportunity for anthropologists, and archaeologists to make strides towards improving the way in which they model human-environmental interaction. The true test will be to see if we can get beyond simple numerical representations of cultural phenomena and start understanding the way in which cultural landscapes are produced.

Refined Approach

Kohler and Parker (1986:433) have stated "Perhaps in building predictive models we are too ready to make the assumption that only a complex multivariate model can adequately account for human locational behavior, when in fact, a few (proxy?) variables, observed in the highly correlated data base that is our environment, may be sufficient for
forming locational decisions”. I take issue with this, for as I see it, human behavior
simply cannot be modeled upon a few variables, and correlational data my not be as solid
as it appears on the surface.

Certainly, we can find correlation in slope, distance to water, and other
environmental variables to the occurrence of certain site groups, but such correlations
may well be the result of erosion, secondary deposition, or differential visibility of
archaeological materials on the surface, rather than directly related to
human/environmental interaction. If we recognize anything at all, it should be that the
patterns of human behavior we are attempting to uncover can and do occur at different
spatial and temporal scales. This dissertation is about the detection and description of
such patterns, and the development of a heuristic device as a starting point for discussion
about how Xinguanos may have modified their environment, and how we can assess the
extent of those modifications through the development of a predictive model. Since it is
absolutely critical that we first understand how the ways in which we study the past affect
our understanding of it, much of this research is given over to a methodological approach,
including how we develop new ways of collecting and recording data, how we analyze
the data and determine some structure or pattern, and how we attach meaning to the
results of our analyses.

The modeling approach described in this dissertation is based on the assumption
that human behavior is driven by a number of cultural and environmental factors. The
physical locations of past occupational areas are assumed to have been the result of
informed choice on the part of the indigenous groups being considered in this research
(i.e., the choice of where to settle was arrived at based upon some inter-group interaction,
as well as an interaction of the population with their environment). This means that significant regional patterning should exist in the distribution of archaeological resources, an implication of the assumption supported by numerous studies of settlement data (e.g., Judge 1973; Kvamme 1985; Roper 1979).

It is outside the purview of this research to derive what environmental factors may have been at play, nor can we easily derive what cultural factors may have influenced settlement choices (although ethnographic accounts certainly give us insight into what kinds of locations were more suitable than others). In fact, we would be quite remise if we based our model on simple correlations of proximity to water, or soil drainage values. Archaeologists often take a static, classificatory approach to the environment, even when the human variables happened to be considered part of a dynamic system. Church et al. decry the use of modern environmental data in APMs as well: “…all the predictive models to date have relied on variable expressed in the contemporary environment. … To expect a model based on present-day conditions to be of use to modeling the site locations of say, Paleoindian sites, is a tenuous assumption at best” (Church et al. 2000). The pattern must be found outside of these types of relationships. While geographic information systems can give practical integration of spatial structures (habitations, soils, river drainage), practical understanding of past and current relationships among these environmental and human systems require a culturally specific temporal and spatial perspective applied at a regional scale.

Two broad assumptions must be made before a predictive model may be employed. First, archaeological resource locations must be assumed to be nonrandomly distributed with respect to identifiable environmental variables; and second, site samples can be
obtained that are sufficiently representative of resource locations of the region under study. To satisfy this second assumption, this study utilized targeted sampling, based upon informed knowledge derived from ethnographic investigations about what sorts of resources may have been of particular interest to the progenitors of the archaeological site locations in the Upper Xingu. The process of targeted sampling will be described in more detail in the remote sensing portion of this dissertation.

The goal of this research is to find new ways of studying past human/environmental interaction utilizing technological advances which can aid in the discerning of pattern over large spatial scales. By utilizing an integrated approach to these technologies, it may be possible to formulate a model of what sorts of areas might be likely to contain traces of past human occupation, and thus contribute to a larger understanding of the interplay of humans and the landscape within which they existed in the Upper Xingu region.
CHAPTER 4
GLOBAL POSITIONING SYSTEM (GPS)

The first steps in providing information for the development of any modeling effort are the collection of spatial and attribute data. For the purposes of this study, a synthetic approach was utilized, incorporating GPS spatial data collected over the course of two field seasons, which was then augmented by ethnoarchaeological data, and unified in a GIS database to present a unified platform for analysis of remotely sensed imagery. The GPS was used as both survey tool and as a general storage device for georeferenced attribute data, thus serving two critical functions during the course of this project. In the following text, a brief description of the overall system behind GPS is provided, as well as an in-depth treatment of how such a system may be more effectively incorporated into archaeological investigations.

Using GPS in Archaeological Survey

Historically, archaeologists have been highly skeptical of the degree to which measurements obtained via GPS signals can be relied upon and utilized in mapping and surveying. To be sure, there are a number of sources of possible error and biases that enter into the process of obtaining precise positioning with a Global Positioning System. The combined magnitude of these error factors influences the accuracy of the positioning results. Biases may be defined as being those systematic errors that cause the true measurements to be different from observed measurements by a constant, predictable, or systematic amount. Biases may have physical bases, but may also enter at the data processing stage through imperfect knowledge of constants, for example any "fixed"
parameters such as the satellite orbit, station coordinates, velocity of light, etc. Both the positions of each satellite in the constellation and their on-board atomic clocks are constantly monitored, each of the satellite vehicles drift slightly from their predicted orbits over time, and their on-board atomic clocks are can never remain completely synchronized. Additionally, satellite transmission can often be disrupted as it travels through both the troposphere and ionosphere. Signals reaching the ground antenna are subject to "multipathing." Random measurement errors can dramatically affect the accuracy of precise position observations. Random errors are simply unpredictable events (in magnitude and sign) and are due to any one on many factors. The chief sources of random error are the "resolution" of the measurement scale, random internal instrumental effects, and, very rarely but quite possible, some external, highly localized condition such as micro-meteorological events, local signal interference, and the like. Systematic errors, on the other had, occur according to some pattern. For example, errors of this type may be induced by the instrument, the observer, the physical or environmental conditions, but at a constant magnitude, or such errors might be the result of incorrect application of calibration data. Finally, gross errors are the result of overt blunders. By their very nature, these types of errors are often so conspicuous that they can be easily identified and corrected immediately. If, however, these errors go unnoticed, they can have detrimental impacts on the final product.

GPS relative positioning, also called differential positioning, employs two GPS receivers simultaneously tracking the same satellites to determine their relative coordinates. Of the two receivers, one is selected as a reference, or base, which remains stationary at a site with precisely known coordinates. The other receiver, known as the
rover or remote receiver, has its coordinates unknown. The rover receiver may or may not be stationary, depending on the type of the GPS operation. Thus, higher accuracies are generally possible if the relative position of two GPS receivers, simultaneously tracking the same satellites, can be derived.

The principle behind this observation is quite simple: because a wide variety of sources of error can ultimately affect the absolute position of two or more GPS users to almost the same extent, these errors largely cancel each other out when differential or relative positioning is executed. There are a wide variety of differential positioning procedures, but each method shares a common thread in that the position of the GPS receiver of interest is derived relative to another fixed, or reference, receiver whose absolute coordinates in the satellite datum are assumed to be known. Thus, the ability to derive precise positions via a GPS essentially requires measurement of the baseline components between simultaneously observing receivers.

Ultimately, accuracy is dependent on a number of variables. Precision in position measurement clearly depend on whether the user is moving or stationary as stationary observations permit an improvement in precision due to the effect of averaging of positions over time. Accuracy is also dependent on whether or not the data can must be processed in the field, or if it can be subjected to post-processing at a later date before implementation. Real-time positioning requires a "robust" but less precise technique to be used. The luxury of post-processing the data permits more sophisticated modeling and processing of GPS data to minimize the magnitude of residual biases and errors. Again, the level of measurement noise has a considerable influence on position precision, with carrier phase measurements permitting a higher accuracy than pseudo-range
measurements. Accuracy is influenced by the degree of redundancy in the measurements (for example: the total number of SVs within view of the receiver (dependent upon the elevation cutoff angle), the number of receiver tracking channels, the ability of the system to track other GPS constellations like GLOSNASS, and the number of observations.

GPS relative positioning has the advantage of consistently providing a higher degree of accuracy than autonomous positioning observation. Even higher levels of accuracy can be obtained by utilizing carrier-phase measurements as opposed to pseudo range measurements, primarily due to the principle that measurements of two (or more) receivers simultaneously tracking any given SV will elicit similar errors and biases (Langley 1993). The similarity of error measurement is a function of the distance between the two receivers (the closer they are in proximity to one another, the more similar the levels of error will be). If the difference between the two measurements is calculated, then it should follow that any similarity in error should cancel out and be removed from the measurement equation. Static GPS surveying is a relative positioning technique that depends on the carrier-phase measurements (Hoffmann-Wellenhof et al. 1994). A static methodology is designed to utilize two or more stationary receivers to simultaneously track the same SVs. A designated base receiver is set up over a point with precisely known coordinates, such as a geodetic benchmark. The other receiver is set up over an unknown point. By collecting simultaneous measurements at both the base and remote receivers over a specified period of time, a large amount of averaged positions can then be post-processed to determine the precise location of the unknown point in relation to the coordinates of the geodetic benchmark. More applicable to archaeological survey, however, is the fast, or rapid, static surveying technique. Rapid
static survey also uses two or more receivers simultaneously tracking the same satellites and using carrier-phase measurements. The difference from the aforementioned static technique is that rapid static survey requires that only the base station remain stationary over a known point. The rover need remain stationary over unknown points for only a short period of time, and then can be moved to another point (Hoffmann-Wellenhof et al. 1994). Post-processing of data collected in either of these static modes may elicit either a fixed solution or a float solution (dependent on whether enough common data was collected from the receivers for the software to fix ambiguity parameters at integer).

Fast-track, or “stop-and-go,” provides for the most rapid form of static survey. This method also employs two or more GPS receivers simultaneously tracking the same satellites with a stationary base receiver and any number of rovers that travel between unknown points, taking very brief (one- to two-second recording rate for a period of about 30 seconds per each stop) measurements. As long as the initial integer ambiguity can be determined (through initialization), one can hope for approximately centimeter-level positioning accuracy as long as there is a minimum of four common satellites simultaneously tracked by both the base and the rover receivers at all times. In order to perform properly, the rover and base station must always maintain at least four common satellites, even when the rover is being moved, otherwise the initialization process must be repeated again.

The last carrier-phase relative positioning technique that will be discussed is called RTK surveying. Again, it employs two or more receivers, with one operating as a base station, and one or more rovers. The method is most useful when: 1) the survey involves a large number of unknown points located in the vicinity the base station; 2) the
coordinates of the unknown points are required in real time; and 3) the line of sight, the propagation path, is relatively unobstructed (Langley 1998). This is actually the preferred method by many users because of the relative ease of use and the ability to determine position in real-time as opposed to the other methods that require some degree of post-processing. The base receiver measurements and coordinates are transmitted to the rover receiver a radio beacon. The built-in software in a rover receiver combines and processes the GPS measurements collected at both the base and the rover receivers to obtain the rover coordinates.

More commonly used methods include real-time differential GPS (DGPS), a code-based relative positioning technique that employs two or more receivers simultaneously tracking the same satellites and based on the fact that the GPS errors in the measured pseudo ranges are essentially the same at both the base and the rover, as long as the baseline length is within a few hundred kilometers. Again, base receiver remains stationary over the known point. The built-in software in the base receiver uses the precisely known base coordinates as well as the satellite coordinates, derived from the navigation message, to compute the ranges to each satellite in view. The software measures the difference between the computed ranges and the measured code pseudo ranges to obtain the pseudo range errors (or DGPS corrections). These corrections are transmitted in a standard format called Radio Technical Commission for Maritime Service (RTCM) to the rover. The rover then applies the DGPS corrections to correct the measured pseudo ranges at the rover. Finally, the corrected pseudo ranges are used to compute the rover coordinates.
The accuracy obtained with this method varies between a sub meter and about 5m, depending on the base-rover distance, the transmission rate of the RTCM DGPS corrections, and the performance of the C/A-code receivers (Langley 1998). Higher accuracy is obtained with short base-rover separation, high transmission rate, and carrier-smoothed C/A-code ranges. Real-time measurements are often preferable over post-processed observations because the positioning data as well as the accuracy measures can be obtained while remaining in the field (obviously this is a critical issue for archaeologists as it allows for continuous data collection without the need to return to a home office or otherwise connect to some data-transfer system to access post-processing data files). The ability to receive corrections in the field inevitably leads to higher productivity compared with post processing. Post processing, however, will generally lead to more accurate results, primarily because of the inherent abilities of the post processing software to do editing and cleaning of the collected GPS data.

Real-time DGPS operations require a communication, or radio, link to transmit the information from the base receiver to the rover receiver. DGPS corrections are typically transmitted at 200 Kbps utilizing very high and ultrahigh frequency (VHF/UHF) bands (OMNISTAR, http://www.omnistar.com/; RACAL LandStar, http://www.racal-landstar.com/). Dedicated radio link equipment is available to consumers to transmit base station information using the VHF/UHF band. While expensive, these types of radio link systems are able to provide exceptional line-of-sight coverage and have the ability to penetrate into buildings and other obstructions.

A number of various GPS correction services are readily available at various levels of accuracy and cost. Very high levels of accuracy can be achieved utilizing any of the
highly precise permanent GPS reference station networks established by several organizations around the world (IGS and the Continuously Operating Reference Station (CORS). These services are available free of charge. The Canadian Active Control System (CACS) is another regional GPS service, which is available to users for a small charge. Reference stations within these systems receive signals on a continuous basis, and thus provide some of the most accurate corrections possible. A number of other countries have established their own internal networks of reference stations along coastal areas (primarily designed to enhance the safety of marine navigation), which continuously broadcast real-time DGPS corrections in RTCM format. GPS receivers capable of accepting RTCM corrections can reach levels of accuracy in the range of sub meter to a few meters. At the commercial level, two real-time DGPS correction services are widely used. One broadcasts the DGPS corrections through FM broadcast stations, while the other transmits through communication satellites. These systems are called wide-area differential GPS (WADGPS). Both systems require a special receiver to decode the DGPS correction information, which would be interfaced to the GPS rover receiver to output positional information at the meter-level accuracy. WADGPS systems have several advantages over conventional single-station DGPS systems, including coverage of large, inaccessible regions using fewer reference stations. The most useful WADGPS correction systems for remote accessibility are provided by OMNISTAR and RACAL LandStar. These systems use satellite data link, with OMNISTAR operating in the C-band of the frequency spectrum, while the LandStar service operates in the L-band. To access either service, a subscriber needs the system data receiver to receive and decode the DGPS corrections. The data receiver must be interfaced to a differential-
ready GPS receiver to obtain the corrected position. Accuracy of the order of a sub meter to a few meters can be obtained, depending mainly on the GPS receiver type.

Advances in commercially available GPS equipment are also improving the use of Global Positioning for precise positioning applications. Many older receivers relied exclusively on the use of the coarse/acquisition (C/A) pseudo-random code signals transmitted by the SVs. Many of the newer models, however, have integrated carrier-phase receivers. In older models, the receiver was designed to monitor a smaller number of channels, sequencing through every visible satellite to obtain positioning information. Multichannel GPS units can track several satellites simultaneously, allowing the receiver to monitor carrier phase signals and calculate accurate positions at a much faster rate. In simple navigation applications, input from a single receiver is adequate. However, the more precise-positioning demands of archaeological mapping and surveying require a much greater degree of accuracy and thus must depend on carrier-phase observations taken at by at least two receivers at regular intervals. The measurement accuracy of the carrier phase is about 1/100 of a cycle, which amounts to 2 mm distance for the 19 cm L1 carrier. Measurements of phase on the L-band carrier signals thus have millimeter random error, while pseudo-range measurements made with the aid of the time signals modulated on the carrier waves are between 100 and 1000 times noisier. It is the resolution of this carrier wave cycle that permits high degrees of accuracy in archaeological applications of global positioning for the purposes of recording positional information and for extensive survey use.

There are several advantages that archaeologists may find when utilizing GPS satellite surveying techniques for archaeological site mapping and general survey.
Perhaps most importantly, intervisibility between stations is not necessary. This advantage cannot be overstated enough, especially in cases such as the Upper Xingu project, involving vast areas of unsurveyed area and extremely thick undergrowth. Secondly, because GPS uses radio frequencies to transmit the signals, the system is independent of weather conditions. Additionally, much of the hardware is weather resistant, so the process of data collection can continue under a variety of climatic conditions. Because of the generally homogeneous accuracy of GPS surveying, the traditional task of planning a network of intervisible transit stations is no longer relevant. Points can be taken where they are required and need not be located at evenly distributed sites. Because intervisibility of stations is not requisite, and conventional network design strategies can be cast aside, GPS surveying is a more efficient, more flexible, and less time consuming method of mapping. In addition, because GPS is in operation 24-hours a day, 7 days a week, data collection can be done at any time. Finally, high accuracy can be achieved with relatively little effort, unlike conventional survey techniques. However, a few disadvantages need to be taken into consideration.

Because station intervisibility is not necessary, GPS survey methods are especially attractive to researchers working in rugged terrain, or covering exceptionally large regions. The relative ease of use, however, is often offset by the logistical problems of transporting and supporting both the technological side, the technical support to run the system. If additional units are put into operation, costs and logistical issues could rise even dramatically. Perhaps most important of all, GPS requires that there be no obstruction to the antenna of the receiver. Thus, a clear view of the overhead sky must be maintained, meaning that overhanging branches or structures have to be removed (though
the antenna can be raised above the obstruction). The benefit gained through the georeferenced positioning of GPS (GPS coordinates are provided in the earth-centered, earth-fixed coordinate system defined by the GPS satellite ephemerides) means that any positions collected need to be transformed into a local geodetic system before they can be integrated with results from conventional surveys. However, GPS technologies are becoming more and more accurate all the time, vertical measurements retain significant error, thus limiting the ability of GPS to be used in acquiring accurate three-dimensional maps (GPS vertical positions also must be reduced to the employed geoid). Finally, GPS requires extensive training to be an effective replacement. It involves a significant investment of both time and financial resources, and requires the development of new procedures and strategies for planning, field operation, and data analysis. The inherent accuracy of GPS may be the largest barrier to its widespread acceptance, since; under proper survey conditions, the measurements obtained via GPS are often more accurate than surrounding control marks established by traditional survey methods, and integration of the two methods requires manipulation of one coordinate system on another in most cases.

**GPS Use in the Upper Xingu**

The Upper Xingu Project initially employed a Trimble Pathfinder GPS 12-channel Pro XRS receiver attached to a TSC1 data logger. This unit is capable of receiving C/A code with carrier-phase filtering and has instantaneous full wavelength carrier-phase measurement. The data logger was connected to compact dome antenna and was held in the operator’s hand. The data logger was attached to the GPS receiver by a short cable, and information was displayed on a small LCD screen. The interface of the data logger consisted of an easy-to-use menu-driven program. Users input data through a keyboard,
numeric pad, directional arrows, enter keys, and a range of function keys. Customized database "libraries" can be loaded into the data logger to make data acquisition easier, but we found the "generic library" to be so simple, flexible, and easy to use that it was unnecessary to modify it. The GPS unit is easily operated by one person in the field, although two people were often used to ensure safety of the operator in remote locations. Pathfinder Office differential correction software was loaded onto a laptop computer resident at the base camp. The Pro XRS model allows differential correction to be done in real time using either a control beacon or via a satellite signal provided by one of several satellite companies.

The Upper Xingu Project leader chose a subscription to the Omnistar satellite differential correction service due to the remote location of the field sites. The GPS rover data were downloaded from the data logger daily. Since data were collected using DGPS, no post-processing was needed to make use of the data instantly in the field (though the data was subjected to post-processing once downloaded onto systems back in the United States. Maps were generated each night on the laptop, both to identify errors in the data collection and to guide further research efforts. This collection of equipment was used to rapidly complete the survey and mapping of surface archaeological features in the defined research area. Figures 3-1 and 3-2 illustrate the effectiveness of GPS survey for rapidly detailing overall site extent and providing good coverage of the overall feature space, including a delineation of the presence of circular central plazas, radiating causeways, and defensive berms, often noted as principle landscape features of prehistoric and historic Xinguano settlement patterns (Heckenberger et al. 2003; Heckenberger 2005).
Figures 4-4 through 4-8 detail different areas of the Upper Xingu research project. These figures are for illustrative purposes, meant to showcase the results of ongoing work in fully surveying multiple sites throughout the area. A 5, 4, 3 color composite of a Landsat 7 image (August 2002) is used as a backdrop for the GPS layer overlay. Figure
Figure 4-3 is an expanded version of this composite backdrop, complete with a generalized legend explaining how to interpret different land-cover classes.
Figure 4-4. MTFX-6 and MTFX-13 survey data integrated into a GIS with a 5, 4, 3 color composite from Landsat 7 imagery
Figure 4-5. GPS survey map of MTFX-18 integrated into a GIS with a 5, 4, 3 color composite from Landsat 7 imagery.
Figure 4-6. GPS survey maps of MTFX-19 and MTFX 20 integrated into a GIS with a 5, 4, 3 color composite from Landsat 7 imagery
Figure 4-7. Composite GPS survey map of the Nokugu cluster of sites integrated into a GIS with a 5,4,3 color composite from Landsat 7 imagery.
Figure 4-8. Composite GPS survey map of Kuikugu cluster of sites integrated into a GIS with a 5, 4, 3 color composite from Landsat 7 imagery

The most significant source of operator error occurred when the number of satellites visible to the local base station dropped to a level where it was difficult to establish three-dimensional positions. Differential knowledge of the operation of the
GPS unit meant that some of the less-capable operators believed such issues could be overcome by selecting the two-dimensional option on the local base station and rover receivers, requiring one less satellite than the 3-D option. Unfortunately, this option requires an accurate elevation to be determined and entered into the local base station receiver. The error was apparent once the data was differentially corrected. It was immediately apparent that the points collected in this manner were significantly offset in comparison to the data collected using full 3-D mode. These and other instances of operator error were resolved by forcing less-knowledgeable operators to re-read portions of the manuals that came with the equipment, and retracing those features collected using this form of mapping. Once our initial problems were resolved, the equipment proved easy to use. A major advantage of GPS survey is having the ability to make plan maps of archaeological sites and features with the instrument. This capability is not generally appreciated, but advancements in GPS technology now make it possible to create feature maps with comparable accuracy to detailed tape and compass maps.

In most cases the level of accuracy provided by the GPS receiver (20 to 70 cm) is the same or greater than the discrepancy introduced by the value judgment made by the archaeologist as to where to place the GPS antenna to record a point. Increased acceptance of GPS use in archaeological mapping and survey seems assured, particularly as the cost of GPS systems drops and new higher productivity techniques are developed. However, for all its technical advantages, there remain a number of substantial differences between GPS mapping techniques and the traditional transit survey still utilized by most of the archaeological world. To actually make GPS survey and mapping a viable endeavor to most researchers, techniques will have to be developed to reduce the
significant post-processing time still required to permit GPS positional data to complement more traditional survey methodologies (and hence maintain compatibility with the geodetic framework already established in most regions under investigation). To their credit, the primary manufacturers of GPS receiver equipment continue to improve upon the ergonomics of their models, striving to make equipment that is ever more "user-friendly." Truly, this is where the battle will be won, when even the most technologically inept can effectively put GPS equipment into the field and expect to see a significant return on their investment.
CHAPTER 5
REMOTE SENSING

Brief Overview

As mentioned earlier, there is nothing new about the process of utilizing remotely sensed information to aid in archaeological investigations. There have been numerous efforts among researchers to bring the inherent value of remotely sensed data (and specifically the widely-available Landsat imagery) to bear on archaeological investigations. The vast majority of cases have seen the implementation of Landsat data sets to delineate environmental “zones” that can be correlated in a general way with archaeological site distributions (Schalk and Lyons 1976). The principle applications of Landsat imagery by archaeologists who work with remote sensing techniques typically revolve around attempts to identify very specific landforms or anthropogenic signatures (crop marks, shell scatters, or structural remains) to locate archaeological sites (Ebert and Lyons 1976). However, the 30-meter spatial resolution of even the more current Landsat imagery is still far too coarse to identify any but the most gargantuan of individual features. The solution to this problem is not to bemoan the resolution of the Landsat imagery, but rather identify the proper spatial scale at which to conduct research. Instead of trying to identify archaeological features (many of which are surely smaller than 30 meters on a side) on the ground by analyzing the satellite imagery, perhaps researchers would be better served by attempting to identify larger signatures of human activity.

One approach to this problem is to identify those environmental features that would be represented at the scale of Landsat imagery that might have some bearing on site
placement. The issue of how to link environmental variables (determined by prior predictive models to be critical determinants of site locations) with specific on-the-ground locations is often poorly addressed. Most summaries of predictive modeling in archaeology do not consider this rather critical problem (e.g., Bettinger 1977; Jochim 1976; Johnson 1977) and predictive studies (e.g., Gunn 1979; Jochim 1976; Wood 1978) usually deal with this problem in an impressionistic manner.

“Remote sensing” is simply a catchall term encompassing a number of different techniques of acquiring and processing information gathered by distant sensor platforms. For the purposes of this study, remotely sensed data will specifically refer to information gathered from satellite sensor platforms. In the past few decades, there has been a growing interest in the application of such data sets to answering fundamental questions in social science research. The growing appreciation of scientists for the usefulness of these types of data and the ever-increasing availability of such data are equally culpable for the expanded use of remotely sensed information. The real key to the usefulness of these data sets lies in the ability of RS information to provide a means of measuring numerous dependent variables that may have an impact on human settlement and distribution patterns. The Landsat series of sensors has provided some of the most widely distributed RS data to date, encompassing a series of satellites specifically engineered to provide high-resolution imagery of Earth’s land surfaces. Originally designed to provide unclassified data for use in land and water resource assessment, the uses of data from Landsat have been expanded to address everything from visualizing population movement and migration, to monitoring global deforestation and fire damage, to estimating soil moisture and land-cover classifications. NASA recognized the
potential of using space technology to study the Earth's environment as early as the 1960’s. A concerted effort was made to advance a program for developing a remote sensing platform designed to analyze both land and water based features from space. Earth Resources Technology Satellite (ERTS), later re-named Landsat 1, was launched by NASA in July, 1972.

Landsat 1 carried both a television camera a sensor package called the Multi-Spectral Scanner (MSS, which provided swath-like scans of the Earth’s surface in different portions of the electromagnetic spectrum. The MSS proved a valuable sensor package, and was included on the next four Landsat satellites. By the launch of Landsat 4 in July, 1982, a newly improved multispectral sensor called the Thematic Mapper was incorporated into the sensor platform. The Thematic Mapper added three new spectral bands and provided improved spatial resolution of 30 meters (compared to the MSS resolution of 80 meters). Landsat 7 served as the basis for an improved sensor array called the Enhanced Thematic Mapper Plus (ETM+). The ETM+ provides seven channels in the visible, near, mid, and thermal infrared channels, along with a 15-meter spatial resolution panchromatic sensor. Additionally, Landsat 7 was designed with increased storage and data transmission capabilities, dramatically increasing the ability of the platform to acquire and transmit images within a given timeframe. The data for a single scene taken by the Landsat 7 ETM+ includes image data for each of the bands and for the panchromatic sensor and is stored as image data. The images are usually corrected for radiometric and geometric distortions and are then made available as a complete set of raw data.
The spectral response of Band 1 is in the visible portion of the electromagnetic spectrum that corresponds with blue-green light. Energy at this portion of the electromagnetic spectrum is easily scattered by particles in the atmosphere, often giving images in this band a hazy appearance. This band is capable of being transmitted through water and is especially sensitive to particles suspended in water (such as sediments and algae). Data from this band can be used with bands 2 and 3 to create "true" color composite images, which most closely approximate how the scene would appear to the human eye. The spectral response of Band 2 is in the visible portion of the electromagnetic spectrum that corresponds with green light. It can be used with bands 1 and 3 to create "true" color composite images. The spectral response of Band 3 is in the visible portion of the electromagnetic spectrum that corresponds with red light. It is also one of the three component bands used to create "true" color composite images.

The spectral response of Band 4 is in the Near Infrared (NIR) portion of the electromagnetic spectrum. This form of infrared sits just outside the visible red light portion of the electromagnetic spectrum. This form of radiation is reflected to a high degree off leafy vegetation since chlorophyll (the green pigment in green vegetation) reflects much of the NIR that reaches it (it has a high albedo in this band). The spectral response of Band 5 is in the Middle Infrared (Mid-IR) portion of the electromagnetic spectrum. This portion of the spectrum is sensitive to variations in water content in both leafy vegetation and soil moisture. The spectral response of Band 6 is in the Thermal Infrared portion of the electromagnetic spectrum. Thermal infrared is radiation that is detected as heat energy; therefore, the thermal IR band effectively measures the temperature of the surfaces it scans. Band 6 on the ETM+ sensor can distinguish
temperature difference of about 0.6 Celsius, which allows it to detect relatively small
differences in land and water surface temperatures. The cooling effect of water
evaporating from vegetation can be detected, assisting in efforts to map land use
characteristics of a region. Spectral response of Band 7 is in the Middle Infrared (Mid-
IR) portion of the electromagnetic spectrum. This portion of the electromagnetic
spectrum is sensitive to moisture and thus responds to the moisture contents in soils and
vegetation. This band is useful in detecting moisture levels in leafy vegetation and thus
provides a means to monitor productivity and identify areas under cultivation. The
panchromatic band is composed of a black and white sensor with a 15 m spatial
resolution. The higher resolution of this data assists land-use researchers by making
identification of smaller objects easier.

**Processing Imagery for Analysis**

Image processing is the first step in assembling remotely sensed data into a usable
product. The raw data is taken in its un-processed form, manipulated using a combination
of raw computer hardware power and a variety of software modules, and rendered into a
final form. There have been many texts written about this topic, including those by
Campbell (1996), Lillesand and Kiefer (1994), Jensen (1996), and Cracknell and Hayes
(1991). The authors may differ on the finer points of how one goes about processing
images, however, there remains a consistent thread or basic agreement on the major steps
involved. There are at least four primary stages of image processing that must be
undertaken at the onset of any analysis of remotely sensed imagery: preprocessing,
classifying, post processing, and an assessment of accuracy for the final product.

Preprocessing is preparing raw digital data for the main analysis, usually
classification. According to Campbell (1996), preprocessing can be separated into three
functions: feature extraction, radiometric corrections, and geometric corrections. Feature extraction is a process of determining which areas of the raw data imagery to utilize. Multiple bands of information are contained in Landsat TM and other reflectance-based Earth observation data. It is a precarious balancing act, deciding which bands may be kept and which may be discarded, but necessary as each information band requires storage space and exponentially increases processing time. The combinations of bands chosen for processing are dictated by the types of feature extraction required. It is this process of feature extraction, determined by the types of classification being undertaken, that effectively determines which bands are critical to the analysis.

In some cases, research goals may dictate that a greater amount of information (read: more bands) is of more import than the efficiency of the processing. A common compromise between information detail and efficiency of the analysis of the data is to utilize principal component analysis (PCA). PCA creates a few new assimilated bands with most of the variation of the original data, but without the redundancy and noise. PCA achieves thus data reduction by looking for correlations between bands and uses these correlations to reduce any redundancy. Davis (1986) and Gould (1967) are excellent resources for more detailed discussions of principle component analysis.

Another strategy in feature extraction is arithmetic operations. These operations can take the form of simple band ratios, vegetation indices, the tasseled cap transformation (Crist 1983), and even multitemporal combinations.

After a determination of which imagery bands are of the greatest value to the analysis, the bands themselves must be subjected to a process of radiometric correction. Radiometric correction is an algorithmic process of altering the brightness values of the
original data to resolve sensor specific-noise or to compensate for various atmospheric effects such as haze or surface shadows (like those found on slopes). Sensor-specific noise can show up as striping, which can be attributed to errors in data transmission or collection, or to actual sensor malfunction (Lillesand and Kiefer 1994). Whereas sensor noise can be eliminated mathematically with relative ease, atmospheric interference, due primarily to the augmentation of measured radiation as it passes from the sun through the atmosphere to the Earth’s surface and then back through the atmosphere to be recorded on the Landsat instrument. The most common culprits of atmospheric interference are innumerable particulates in the atmosphere which can cause problems ranging from total obscuring of ground features (e.g., clouds or smoke), or by altering, to varying degrees, the spectral values (Lavreau 1991). Many researchers have sought methods of correcting or minimizing these effects (Chavez 1996; Hill and Sturm 1991; Lavreau 1991).

Additional factors that can influence recorded brightness values include reflectance of the target, angle of the sensor, solar elevation angle, and slope and aspect of the target in relation to the solar direction. The reflectance of the target is dependent on localized land cover, and thus is typically the variable actually sought by the analysis. Sensor angle and solar elevation angle essentially remain constants and are not factors in interpreting images due to the sun-synchronous orbits of Earth observation satellites (Campbell 1996). Topographic effects, or the influence of the slope and aspect of localized elevation values on reflectance values, while an important factor in some studies, especially in mountainous environments, has no bearing on these discussions due to the lack of elevation within the study area under question.
 Following radiometric correction, the next step in the process of data preparation before a classification can be performed requires that the satellite data be registered with a consistent georeferencing system. Such referencing allows all geospatial data to be used readily in a GIS. First, a standard correction is made to adjust the effects of the Earth’s rotation on the satellite image, and then the image is georeferenced to a users defined map projection and to a coordinate system. While it is technically true that geometric corrections can actually be performed at any stage in image processing, it is imperative the data geometrically corrected before ancillary data are brought in to aid in radiometric correction, classification, or accuracy assessment. Geometric corrections as well as the other steps in preprocessing must be done with caution. Most of these processes alter the pixel reflectance values in some way. Alteration of these values may affect classification accuracy, thus, it is usually best to alter these values as little as possible (Campbell 1996).

 Raw data streams from the Landsat platform are of little use initially. After being transmitted to a ground-based station, the data streams must be processed and converted into a usable format, often times into a form of imagery that provides a visualization of the data collected by the sensor. Satellite image data is sent from the satellite to the ground station in a raw digital format (essentially a stream of numerical data). Each byte in the data stream corresponds to a single pixel element. The numerical value of the pixel, known as its Digital Number (DN), is translated into a gray-shade. These pixels, when aggregated and then arranged in proper sequence form an image wherein varying shades of gray represent the discrete energy levels detected on within a measured area. Since a satellite image is actually a collection of numeric data, the underlying dataset can
be manipulated using algorithms (mathematical equations) that correct for errors (like atmospheric interference), georeference or register the data set to a specific geographical reference point, or extract information that may not be readily apparent to create a limitless number of derived products generated simply by performing calculations on the raw numerical data.

In terms of optical perception, the human eye can only perceive, on average, about 16 shades of gray. The inability of the human eye to differentiate gray-shaded images necessitates the conversion of gray-scale satellite images into color derivatives by assigning a specific digital number (DN) value (or ranges of DN values) to specific colors, thereby increasing the contrast of particular DN values with the surrounding pixels in an image.

The real advantage of utilizing digital imagery is that it allows the researcher to exercise some control over the image information through manipulation of the digital pixel values in an image. Many of the images distributed to end users have often times already been subjected to a number of different forms of image enhancement (radiometric corrections for illumination, atmospheric influences, and sensor characteristics may be done prior to distribution of data to the user, for example). However, the image may still not be optimized for visual interpretation. Satellite-based sensor platforms must be designed to cope with levels of target/background energy typically encountered during routine image acquisition. Since the spectral variation across bands may vary widely depending on the type of land cover under study, no generic radiometric correction could optimally account for and display the optimum brightness range and contrast for all targets. Thus, a much more targeted enhancement of specific images often is necessary.
Pixel values within each band for Landsat imagery can range in value (DN) from 0 to 255. The most basic form of image enhancement, termed contrast enhancement, involves changing the original values so that more of the available range is used, thereby increasing the contrast between targets and their backgrounds. By manipulating the range of digital values in an image, graphically represented by its histogram, it is possible to increase the fidelity of specific features in relation to background information within an image. The simplest type of enhancement is a linear contrast stretch. This involves identifying lower and upper bounds from the histogram (usually the minimum and maximum brightness values in the image) and applying a transformation to stretch this target range to fill the full range of histogram values. This type of contrast enhancement reassigns pixel contrast so that light toned areas appear lighter and dark areas appear darker, with the net effect of making visual interpretation a much easier enterprise.

A linear stretch creates a uniform distribution of the input range of values across the full range, but this may not always be the most appropriate way to enhance an image (this is especially true if the original input range is itself not uniformly distributed). To create a contrast enhancement of a less evenly distributed input range, a more targeted histogram stretch can be utilized to assign more display values (range) to the frequently occurring portions of the histogram. In this way, the detail in these areas retains greater fidelity relative to those areas of the original histogram where values occur less frequently. By expanding various portions of the original target spectral histogram utilizing a number of different stretching techniques, researchers can achieve better contrast of desired target areas.
One important advantage of Landsat Thematic Mapper (TM) is its ‘synoptic view’, which makes it possible to study large portions of the Earth’s surface at a relatively low cost. Sensor systems of higher spatial and spectral resolution do not yet provide this large area view at a low cost. For this reason, improvement of classification methods of Landsat sensor data should focus on problems faced by large-area projects. Instead, this type of research has often been limited to small image subsets, sometimes comparable in size to that of only a few aerial photographs. Consequently, innovative classification methods that produce more accurate results, but require more software, hardware or ground data, have rarely been applied to large areas.

The most critical (and the most error-prone) step in the process of image analysis is the actual classification of individual pixel elements. The intent of the classification process is to assign each pixel in a digital image to one of several land cover classes. This categorized data may then be used to produce thematic maps of the land cover present in an image. Simple visual classification relies on the analyst's ability to use visual elements (tone, contrast, shape, etc) to classify an image. This method may seem the easiest to employ, but, in fact, such methods tend to have too great an error margin built into the analysis to be useful for anything more than simple “first-run” classification or verification of image analysis via other means.

The two primary methods of image classification involve either spectral pattern or spatial pattern recognition (Lillesand and Kiefer 1994). Spatial pattern classification is produced through informed relationships between neighboring pixels based on statistical analysis of certain elements of each cell. Spectral pattern classification relies on statistics of the brightness values for each pixel of the image in each band. Spectral pattern
classification is the most traditional approach, and the approach that will be sued in this project. Rather than relying on the ability of the human eye to accurately perceive differences, these software based solutions instead use the actual DN of each pixel to more accurately assign image elements to specific categories. The result of a classification is that all pixels in an image are assigned to particular classes or themes resulting in a classified image that is essentially a thematic map of the original image. Various themes, then, actually represent categorized pixels sorted into a number of spectral classes. Spectral classes are groups of pixels that have nearly uniform spectral characteristics. It is the task of the analyst to observe and understand the various spectral classes and convert them into information classes, the various themes or groups the analyst is attempting to identify in an image. Information classes may include such classes as varied types of forest, various agricultural crop types, inland bodies of water, or urbanized areas. Thus, the real task of the human element of image classification is matching the spectral classes in the data to the information classes of interest.

While image analysis can be performed on a single spectral band, it is simple not the most effective means of deriving accurate classification thresholds. One band classification is usually very difficult to classify since it is entirely within the realm of likelihood that a number of different surface types may exhibit the same spectral value within any given band. Thus, any spectral classes in a single band classification will likely contain several information classes, and distinguishing between them would be difficult. Multispectral classification, utilizing two or more bands of information, allows for the combination of digital numbers to identify in a much more specific way the spectral signatures of the spectral classes present in the image. A greater number of
bands nearly ensure a more accurate differentiation of different cover classes. Normally, multispectral data are used to perform the classification, with the spectral pattern present within the data for each pixel is used as the numerical basis for categorization (Lillesand and Kiefer 1994).

Common classification procedures can be broken down into two broad subdivisions based on the method used: supervised classification and unsupervised classification. With supervised classification, the researcher locates areas on the unmodified image for which he knows the type of land cover, defines a polygon around the known area, and assigns that land cover class to the pixels within the polygon. Thus, a supervised classification requires the input of known ground locations with known cover types. This process is continued until a statistically significant number of pixels exist for each class in the classification scheme. Then, multispectral data from the pixels in the sample polygons are used to train a classification algorithm. Once trained, the algorithm can then be applied to the entire image and a final classified image is obtained. The numerical information in all spectral bands for the pixels comprising these areas is used to "train" the computer to recognize spectrally similar areas for each class. Supervised classification thus utilizes acquired (ground-truthed) knowledge of different classes in a scene to aid in identifying representative samples of different surface cover types termed training sites (areas that are designed to identify the spectral characteristics of each class of interest). These are referred to as “training areas.” Once training sites have been established, the numerical information in all of the image's spectral bands is used to define the spectral "signature" of each class. Image analysis software is coded to recognize these signatures and attribute then to specific classes, which the software then
compares to every other pixel in the image. On-screen digitizing (or other methods) can then be used to designate these areas as training sites. The computer then analyzes these sites and computes statistics about the reflectance values of each category, according to a specific algorithm. Next all of the pixels are examined and, based on their brightness values, placed in the category with the closest (stated statistical bounds) training values. This method requires a large amount of ancillary data (ground data) for training sites, prior to classification. Thus, in a supervised classification, the analyst starts with information classes and uses these to define spectral classes. Each pixel in the image is then assigned to the class that it most closely resembles.

Problems that could reduce thematic accuracy in supervised classification include: lack of training data for every land cover type, poor quality of ground data, low quantity of training sites, or a poor distribution of training sites. A paucity of adequate training data could cause large areas not to be classified or to be labeled as an incorrect class, depending on the classification method utilized. Inaccurate or imprecise training data will inevitably have a dramatic affect the training statistics, consequently reducing the accuracy of the classification procedure. If training sites are distributed poorly (for example, training sites for each class are clustered rather than being evenly spread) then spatial autocorrelation problems may occur (Campbell 1981). If the number of training areas for each category is low, classification statistics may be skewed, depending on the classifier.

Training sites are used to form different classes, and the software calculates the mean spectral signature for each class based on the Digital Numbers corresponding to each pixel element in each band involved in the classification procedure. When the
classifier is trained and applied to the image, the Euclidean distance is calculated between
the spectral signature of each pixel and the mean spectral signature of each class, and
pixels are assigned to the class to which the distance is the lowest. Minimum distance
classifies image data on a database file using a set of 256 possible class signature
segments as specified by the signature parameter. Only the mean vector in each class
signature segment is used. The chief drawback to this method is that every pixel within
the image is relegated to a class. While this may be the intent of the classification, it does
not necessarily result in the most accurate output. Under this method of classification, it
is entirely possible to have outlying pixels (elements that in no way resemble any of the
designated classes) be lumped into a class.

The Parallelepiped classification scheme improves on the Minimum Distance to
Means algorithm by calculating the mean spectral signature for each class, and then
generating an additional parallelepiped image consisting of as many dimensions as there
are bands used in the classification (in the case of a multiple band classification utilizing
3 bands, the resulting parallelepiped image would be a 3-dimensional cube). The
boundaries of the generated image are one standard deviation away from the mean.
When the classification scheme is applied to the image, those pixels falling within a
particular class's parallelepiped are assigned to that class, with outliers assigned to an
unclassified category. Obviously, this method is more accurate than the first since it
restricts the assignment of classes to only those pixels that fall within one standard
deviation of the class mean; however, the final result produces a number of unclassified
pixels.
Maximum Likelihood operates under the assumption that values for each pixel element within a training site are distributed normally. This procedure calculates the mean spectral signature and a covariance matrix for all spectral bands within a specific class’s training data. When applied to the image, the Maximum Likelihood algorithm uses each pixel's spectral signature and class statistics to calculate the probability of each class belonging to that pixel. Whichever class has the highest probability of affinity for that particular pixel is then assigned. Maximum Likelihood is one of the most commonly used supervised classifiers and generally outputs class maps with high classification accuracy, and, given the speed of modern computers, is relatively efficient in its computational demands.

In essence, the process of unsupervised classification works exactly opposite that of supervised classification. Unsupervised classification, sometimes referred to as clustering, does not require large amounts of initial input (Jensen 1996). The basic premise is that values within a given cover type should be close together in the measurement space (i.e. have similar gray levels), whereas data in different classes should be comparatively well separated (Lillesand and Kiefer 1994). Although these clusters are not always equivalent to actual classes of land cover, this method can be used without having prior knowledge of the ground cover in the study site. Thus, the computer looks for logical clustering of reflectance values. The analyst may enter initial input like the number of desired thematic or information classes, threshold standard deviations for separating clusters, or the number of passes the algorithm makes across the image. The pixels in an image are examined, without user input, by the image analysis software, and classified into spectral classes. Spectral classes are grouped first, based
solely on the numerical information in the data, and are then matched by the analyst to
information classes (if possible). Programs, called clustering algorithms, are used to
determine the natural (statistical) groupings or structures in the data. Usually, the analyst
specifies how many groups or clusters are to be looked for in the data. In addition to
specifying the desired number of classes, the analyst may also specify parameters related
to the separation distance among the clusters and the variation within each cluster. The
culmination of this iterative clustering process may result in some clusters that the analyst
will want to subsequently combine, or clusters that should be broken down further - each
of these requiring a further application of the clustering algorithm. It is then up to the
analyst to translate these clusters into informational classes. According to Campbell
(1996), there are several advantages and disadvantages to this type of classification. The
advantages are: low amount of \textit{a priori} knowledge of the area is required, potential for
human error is reduced, additional relevant classes may be identified which may have
otherwise been overlooked, and the output classes are spectrally consistent.
Disadvantages include lack of analyst control and the possibility of the spectrally
clustered output being inconsistent with the desired informational classes. There are
other disadvantages. Spectral relationships change over seasons and years. These
changes limit applications to a single time period. There is also a need for detailed
ground data when translating the clustered output from spectral classes to information
classes. The identity of the spectral class is not be initially known, and must be compared
to classified data or some other form of reference data to ground truth the spectral
classes. The timing of required ground data is just the opposite of supervised
classification (Lillesand and Kiefer 1994). Thus, in a supervised classification, the
observer first defines classes based on informed knowledge, and then attempts to quantify their spectral separability. In an unsupervised approach, the software determines spectrally separable class, but does so without any basis for the formation of specific categories.

Post-processing of classified imagery involves the preparation of the thematic layer for accuracy assessment, map output, and further analysis. A common procedure in post-processing is smoothing of the classified image, which removes the “salt and pepper” appearance on the map caused by spectral variability in pixel-based classification algorithms (Lillesand and Kiefer 1994). Smoothing can be accomplished using various filters. Other types of post-processing may involve conversion of the image data from raster to vector, alteration of pixel size, creating a new color palette, or forming mosaics of separate images to cover an entire study area.

The final procedure in image analysis should always be an assessment of the accuracy of the analysis. Accuracy at this phase refers to how close the classified pixels are to the actual land cells. Due to the arithmetic nature of GIS overlays, errors are propagated through each step of a geographic analysis, and so it is essential to keep track of error in individual map layers (Janssen and van der Wel 1994). There has been much attention given to the accuracy assessment of land cover maps in the last decade (Verbyla 1995). Typical accuracy reports include overall accuracy, user’s and producer’s accuracy for each category, an error matrix, and the Kappa statistic. Overall accuracy is a simple measure derived by taking the number of total reference cells and dividing it by the number of correctly classified cells. The user’s accuracy refers to the number of reference points within a class that were actually classified as that class. Producer’s
accuracy refers to the number of pixels that were labeled a certain class that were verified by the reference data. Both of these statistics can be calculated using the error matrix. The error matrix is a table with the reference data on the horizontal axis and the classified data on vertical axis. The Kappa statistic is a measure of how much better the results of classification are versus random pixel assignment (Congalton and Mead 1983). The Kappa unit expresses the percent improvement of the classification results over random classification. This statistic can also be calculated from the error matrix. Congalton (1991) and Congalton (1996) provided thorough reviews of this accuracy assessment research and of the current standards.

Vegetation is the most dynamic element of the landscape from a remote-sensing perspective. Diachronic analysis of vegetative cover must take into account seasonal changes and their effects on the overall spectrographic signature of any given region. To remove this sort of bias form the analysis, it is often preferable to use images that are recorded at times when the development of vegetation is at an identical or very similar stage. By breaking the vegetation down into component species, mapping of biogenic emissions can be greatly improved. This can be performed using supervised classifications and will allow vegetation inventories to be produced quickly and efficiently. Vegetation mapping derived from the use of remotely sensed imagery is primarily classified though the cross correlation of various bands and the application of band ratios. Numerous ratios have been tested; however, Tucker (1979) using Landsat Multi Spectral Scanner (MSS) data, found the red and near-infra-red ratio was seven to 14 percent better at detection of vegetation than the green/red band combination methods previously used for vegetation detection. The main vegetation component that the band
combinations detect is the green leaf biomass or green leaf area. These ratios are primarily used for vegetation density measurement. There are a number of other properties of vegetation evidenced in spectra visible to satellite imaging equipment that can be used to classify flora into various categories. Several studies have measured the total biomass of vegetation and attempted to measure and predict evaporation of water to investigate atmospheric interactions (Gholz et al. 1997; Montieth 1976; Gholz 1982) and correlating leaf biochemistry to the spectral properties of vegetation for the purposes of remote sensing. Internal leaf pigments and cell structure, as well as reflection or scattering of light, have been addressed (Buschmann and Nagel 1991). Ultraviolet light effects on vegetation are also important for remote sensing of vegetation due to absorption and reflection characteristics (Mazzinghi et al. 1994). One form of a vegetation index is a band ratio of the red band with the near infrared band. An abundance of indices are available for detection of vegetation from remote sensing (Elvidge and Chen 1995; Jackson 1983). In its simplest form, the division of the near infrared band by the red band correlates to vegetation density and health (greenness). Other more advanced forms of indices include corrections for the influence of soil and atmosphere. Generally, most vegetation indices are ratios that eliminate shadowing effects through highlighting the difference in reflectance between two image bands. Removal of shadow and albedo effects from vegetation indices can offer improvements in classification (Qi et al. 1995).

Vegetation cover maps are composed of polygons with a content, structure, and composition matching a type description. In a Landsat TM image with a nominal pixel size of 30m, a given vegetation type may occupy anywhere from a few, to several
hundred, or even thousands of pixels. Spectral reflectance values of pixels in close spatial proximity within a given stand tend to be similar. Conversely, the prevalence of spatially contiguous vegetation improves the likelihood that adjoining image pixels belong to the same cover type class. Contextual classification exploits these relationships among neighboring pixels as opposed to a per-pixel classification that derives a cover type from the information linked to a single pixel. Stable, consistent, and predictable relationships among neighboring pixels can be quantified and used to improve the classification accuracy. Further refinement of methodology has led to a substantial number of contextual methods derived from Markov random fields, spatial statistics, Bayesian methods, fuzzy logic, segmentation, texture, or neural nets.

In the context of this research, remote sensing analysis relates specifically to a human dimension of landscape formation in that it captures the outcomes of human actions writ large in the vegetation itself. Such analysis provides both spatial and temporal information for land use and land cover analysis. For the purposes of this research, pixels components of the satellite images are assigned to land cover classes and classification maps are created. This provides a means of quickly assessing areas of similarity and dissimilarity to our classification schema, and allows for rapid aggregation of spectral values within a specific threshold of similarity.

An integrated approach to remote sensing is not a new concept. Many researchers in the Amazon have used a combination of remotely sensed data in combination with other technologies. Utilizing maps based on a visual interpretation of Landsat TM imagery, Alves (1999) was able to analyze geographical patterns of deforestation for states, municipalities, and road buffers. Skole and Tucker (1993) using Landsat TM
images and GIS integration, mapped land use/land change for the entire Brazilian
Amazon. Deforestation, fragmented forest, defined as areas smaller than 100 km$^2$
surrounded by deforestation and edge effects within 1 km into forest from adjacent areas
of deforestation - were measured for 1978 and 1988. The findings supported analyses on
the effects of human colonization within the region, specifically tracing rates of
deforestation (Skole et al. 1994). Eva and Lambin (1998) outline a number of more
recent initiatives utilizing an integrative approach of combining multiple sensor platforms
to aid in estimating biomass burning at a regional scale. Wood and Skole (1998) have
linked satellite, census, and field observation data to analyze trends in deforestation.
Several other research initiatives using remote sensing and GIS techniques have taken
place at distinct sites and more detailed scales in the Amazon. A number of studies
relating to physical, biological, and social processes have aided researchers in
understanding how human decisions affect local and regional land use (Mausel et al.
recent publications adequately discuss the progression of sensor technologies and
techniques to monitor land use and land change, specifically within the Amazon (Adams
CHAPTER 6
METHODS

Increasing complexity in instrumentation and enhanced capabilities in data production and methods of analysis in current analysis programs have led to new approaches and to a more integrative vision about Land Use/Land Change (LULC) within and across research sites (Burrough and Frank 1995). The use of integrated approaches to available technology has allowed the burgeoning of a new era in Amazonian research. The evolution of these tools has caused a fundamental shift in the way we approach investigative strategy and implementation. The possibility of testing spatial models using georeferenced databases and algorithms to measure spatial heterogeneity has opened new pathways to research issues of archaeological investigations. A new era of ecosystems spatial dynamics studies necessitate new quantitative methods capable of analyzing patterns, determining the importance of spatial processes, and developing models about landscapes (Gardner and Turner 1991; Fortin 1999). Therefore, many ecological studies have described features in the landscape by number, diversity, distribution, complexity, and dispersion of spatial components (Jurdant 1977; Domon et al. 1989).

Advanced airborne and satellite technologies, image processing and analysis, and the ever-increasing capabilities of advances in technology to provide more efficient means of analyzing spatial data through GIS and associated software have rapidly increased the development and testing of new quantitative methods of spatial assessment (Goodchild et al. 1992, 1993; Sample 1994; Burrough and McDonell 1998). The variety of aerial and orbital data in distinct spatial, temporal, and spectral resolutions have
required the generation of digital image processing techniques in applications related to
the characterization and management of natural resources (Johannsen and Sanders 1982;

The exclusive use of Landsat TM imagery in this particular study is purposeful.
There are a number of exceptionally good reasons to use Landsat imagery for studies of
the Brazilian Amazon, and especially concerning the human dimensions of
environmental change. Both types of Landsat imagery (earlier MSS and later TM data)
cover a broad spatial extent. Each individual scene acquired by either platform covers is
approximately 185 kilometers wide, based on Landsat’s large instantaneous field of view.
Perhaps more importantly, however, the Landsat platform acquires scenes of all the
Earth’s terrestrial surface between 81º N and 81º S latitude at regular intervals (Campbell
1996: 162), with the archives dating back to 1972. While a mere 30 or so odd years is
decidely small in comparison to the history of human occupation of the landscape, these
regular observations provide a substantial temporal range to capture many of the human-
induced changes that have occurred in Amazonian landscapes. Perhaps most importantly
for the purposes of these investigations, Landsat TM instrumentation provides a relatively
high degree of spectral resolution in comparison to competing platforms.

Traditional Analysis Techniques

Techniques of image analysis have already been thoroughly defined as varied
methods for displaying and interpreting band-to-band variations of multispectral satellite
images. The most common approaches to imagery analysis comprise numerous variants
of single-band analyses, color composite generation, band ratioing, vegetation indices,
principal component analysis (PCA), and classification. Several of these techniques were
used in the course of this study, and most will be thoroughly explained in coming text, so
at this point simple descriptions seem to be the most expedient way of dealing with this subject.

A single-band analysis typically involves a simple display of individual bands. Color compositing techniques superimpose three bands together, displaying each band of information using three primary light colors: blue, green, and red. The limitations of simple color composite analysis is that three, and only three, bands may be used at any one time.

Band ratioing is the use of a variety of band mathematic functions to compare various bands comprising an image. Vegetation indices is a catch-all term referring to a set of some of the more popular band-ratioing techniques specifically geared towards vegetation separability and classification (and thus of particular interest to this study). One of the limitations of ratios is that they generally do not take advantage of all the available information contained in multispectral images, usually using only two bands (Adams et al. 1995). Additionally, band ratio techniques specifically geared towards vegetation analysis can be influenced by many factors not associated with vegetation itself (e.g., soil background and sensor differences) (Campbell 1996).

Principle Component analysis will be explained in more detail later in this chapter. PCA is often used to remove interband correlations that typically exist within multispectral image data. PCA identifies linear combinations of the original band data of an image to produce component images representing the axes of maximum variation (Campbell, 1996). An assessment of variance among multispectral image components allows the researcher to “compress” the data by utilizing PC bands in the place of the original band
data (Lillesand and Kiefer 2000). The use of PCA distinguishes itself from previous methods by utilizing all the information contained in a multispectral image.

At its core, a classification analysis utilizes a classification algorithm (chosen from among a sizable number of different options) which then assigns individual pixels of a multispectral image to discrete categories. The goal of this type of analysis is to greatly simplify continuous image data (7 bands of data for TM) using quantitative techniques for identification of spectrally similar land-cover classes within the image (Lillesand and Kiefer 2000). As previously mentioned, classification of satellite imagery falls into two basic camps, arguably with a third “hybridized” variant. Unsupervised classification procedures group or cluster the multispectral values of the image into distinct classes (e.g., water, soil, and vegetation) based solely on the image statistics and produce a new raster displaying the class designations within the image. Supervised classification involves the use of ground-truthed data sets to “train” the classifier algorithm to assign appropriate clusters of data to certain land-cover classes. Many Amazonian researchers, in particular, are using image classification techniques to aid in the understanding of the human dimensions of environmental change (Brondizio et al. 1994; Lee and Marsh 1995; Moran et al. 1994). Again, one of the strengths of these methods is that all available band information is included in the analysis. As to the primary weaknesses, each pixel may belong to only one land-cover class and most classifications are extremely sensitive to errors induced by surface reflectant anomalies, as well as non-surface sources of variability (sensor calibration, illumination, and atmospheric differences).

**Realization of an Integrated Approach**

The term integrated GIS (or IGIS) has been used extensively within the GIS and remote sensing research community to more accurately describe the methodological
inclusion of both RS and GIS information, and the integration between these two analysis platforms into a singular analytical technique (Faust et al. 1991, Lauer et al. 1991, Star et al. 1991).

Many researchers have already made use of a number of different image analysis systems and GIS platforms within individual projects in order to reap the maximum advantage of available functionality. The recent trend has been towards multi-system use of multi-format spatial databases. In practice, the actual transfer of spatial data between the two platforms is complicated, primarily owing to the unique way each system handles storage and processing, resulting in a number of cross-platform issues that need to be addressed. Firstly, there can be subtle differences in spatial models even within the same generic format (geo-relation feature versus object-oriented vector models) which can be hamper information exchange at the attribute level. Additionally, the conversion of data between formats can lead to some generalization and loss of accuracy. One is dependant on the export utility of each platform in its translation of native data to other formats. In order to minimize any loss of quality the data should be left in its native format, but proprietary formatting is the life-blood of many of these systems (and so the motivation to truly integrate is not present). Finally, at the most fundamental level (i.e. actual machine code), these systems are still either vector-based or raster-based, placing inherent limits on the functionality of applications to data outside their primary domain. Even in the case of ENVI, which does have built-in functionality to access and display vector data from a GIS system, the system still has limited vector GIS functionality for spatial decision analysis. The reverse is just as true, with current GIS software containing
a fundamental weakness as components of integrated systems, requiring conversion into a native format to achieve interoperability.

There are many case studies involving substantial manual and digital cartographic interaction of both types of platforms, highlighting the issues with interoperability, as well as studies wherein satellite images were classified and then vectorised in order to port the feature classes into a GIS module. This is a tested methodology that has been successfully implemented many times over, but the processing steps must be well defined in advance. Kontoes et al. (1993) used GIS derived data to post-process a classified image utilizing data from both a raster based image processing system and a vector GIS. By using data resulting from digitized maps, the researchers were able to use co-registered raster forms that could be ported into the image processing software and be integrated into the classification methodology. Utilization of data interchange format polygons in combination with the layered attribute data stored in a GIS, allows the polygons, as well as the attributes attached to those polygons, to be transferred into an image analysis system. There, such information can be used in, for example, image segmentation, to aid selection of training sites in supervised classification, or in image enhancement tasks. The issue, however, is that once transferred out of their native GIS, many GIS functions can no longer be applied to the polygons (e.g., selection/editing, attribute query, topological queries). Johansen et al. (1994) have presented GIS as a tool for the integrated analysis and interpretation of remote sensing based maps with georeferenced in situ or model environmental information. In their analysis, involving both unsupervised and supervised classifications, the authors used a vector-based GIS, a raster-based image processing system, and a data visualization program. The
“integrated” nature of such a methodology is in the passing of data between the various platforms. This approach to integration certainly is wrought with problems, but is necessary until fully integrated systems with capabilities for both raster and vector analyses become more prevalent.

The value of GIS and associated software is in their ability to provide a data structure to efficiently store and manage ecosystems data for large areas, as well as enable both the aggregation and disaggregation of data between multiple scales, support spatial statistical analysis, improve information-extraction for remotely sensed imagery, and provide input data and parameters for various forms of modeling (Haines-Young et al. 1996). This has led to an increasing number of applications for synthesized technological approaches.

I have established the relationship between these various technological realms. The issues I shall attempt to resolve is the following:

1. Are there specific reflective signatures useful in delineating vegetative types within the study area?

2. Is there a significant difference in reflective values between vegetation within the localized boundaries of known historical occupation sites derived from GPS survey and vegetation surrounding those sites?

3. Can we use the selected signatures of vegetation present within anthrosols to extrapolate out to regional-scale models of unknown occupational site locations?

4. Can such a process be applied to a predictive model at an even broader scale, using analytical principles under discussion?
Figure 6-1. Methodological flowchart
Remote sensing, GIS, GPS survey techniques, and spatial analysis must play a central role in providing elements for this discussion. We must first define our units of analysis. Units of analysis for the region in question must be centered on biophysical characteristics, socioeconomic context, and the spatial-temporal arrangements of occupation.

Pre-classification Techniques

Several pre-processing techniques were carried out prior to classification. The first step was to correct geometric distortions present in the raw Landsat TM images. Geometric rectification is the process of image adjustment to a pre-established coordinate system (Lillesand and Kiefer 2000). Before remotely sensed data can be effectively used, the imagery must be geometrically corrected. It is often the case that imagery that has already been geometrically corrected is available to the user. However, researchers working in remote areas will likely find that geometrically corrected data are not available for the study area. Fortunately, many software packages provide geometric correction modules.

The process of geometric correction (commonly referred to as georeferencing) involves match points in the imagery with known coordinates collected in the field with GPS or recorded from map data known as ground control points (GCPs). By using GCPs in combination with identified pixels in the map imagery, one can apply a coordinate system to the data according to a transformation. More dispersed and more numerous GCP's will improve the accuracy of georeferencing (Jensen 1996). The accuracy of image georeferencing determines the spatial integrity of information derived from image processing performed on that scene. If the data are not properly georeferenced then
identification of features whose locations were not known a priori remains in significant doubt. Imagery used in this analysis was registered based on identifiable control points.

Radiometric correction of satellite imagery was utilized to account for path radiance (the random entry of energy into a detector's field of view at a given pixel through the process of scattering, causes additive offsets). It is assumed that atmospheric factors creating additive offsets are fairly homogenous and increase pixel values evenly throughout an entire scene. To correct for additive path radiance effects, pixel digital number minimums are subtracted from each band of the data so that the DN minimum for all bands is equal to zero. The RS imagery package from RSI called ENVI has a built-in function for radiometric correction of Landsat ETM+ images.

Species classification and vegetation separability of tropical forest satellite coverages is usually dependant on supervised and/or unsupervised classification techniques. The spectral complexity of tropical forest classes has further led to numerous suggestions for procedures and techniques to improve classifications including for example stratification by ecological zone (Thenkabail 1999; Helmer et al. 2000), topographic normalization (Colby and Keating 1998), spatial filtering (Hill and Foody 1994), image segmentation (Hill 1999), object-oriented classifications (Foody et al. 1996), vegetation indices (Boyd et al. 1996; Helmer et al. 2000) and multi-temporal image data (Lucas et al. 1993). However, no standardized classification approach has been developed for tropical forest mapping as the approaches vary according to objectives and scale of study, environmental settings and software abilities (Thenkabail 1999).
The prospect of separability of tropical forest types using Landsat Imagery is quite poor in tropical environments owing primarily to the rapid regrowth of vegetation, as well as the consistently high level of greenness and density of the vegetation canopy (Salas and Brunner 1998). Add to this equation the complexity of the reflectance patterns due to variegation within canopy types (Hill 1999) and the 30m² grain of Landsat imagery, and the result is a challenging task to define and classify vegetation types. The textural complexity becomes especially evident in higher resolution imagery of tropical environments as the in-class spectral difference is significant relative to the between-class spectral variation (Thenkabail 1999; Hill 1999). Two near cloud-free Landsat 7 ETM+ images, from 19 May 2003 and 04 August 2002, were acquired and co-registered using GPS surveyed ground control points to rectify the August image to a 2003 image using a nearest neighbor resampling routine that maintained the original 30m spatial resolution. All visible and infrared bands were available for the analyses. However, the thermal infrared bands were excluded due to their lower spatial resolution.

**Classification**

The explicit reason for an adoption of an integrative approach for the purposes of this research is combining vector information with image classification in the selection of training areas. Critical to the realization of an integrated solution for classification purposes would be maintaining the flow of information between the two platforms, and maintaining links between the raster image and the vector dataset. Both unsupervised and supervised classification algorithms were utilized. Thirty information classes form the basis of the supervised image classifications used in this research. Training and reference data were extracted based on GPS survey points and polygons (ground truthed information), as well as known areas of discrete land-cover types that would have been
difficult to gain access into for the purposes of collecting GPS referenced data (river/lake regions, recently burned ranching areas, and the like). Training data displayed an intentionally high degree of spectral separability overall based on statistics extracted for the radiometrically calibrated imagery. However, the various vegetation classes had a great deal of spectral overlap, which was certainly to be expected. Generally, it is preferable to select unimodal training data. This way an analyst is sure to select training data from a single spectral class. The hazard of utilizing multi-modal training data is the ever-present threat of a conflation of multiple spectral groups into a single training class. However, as this project aims to find broad classes relevant to Xinguano settlement patterns, some multi-modal spectral groups were necessary. Four classes selected from image data contained multi-modal training data: cultural clearings, manioc fields, bare soils, and water bodies. Cultural clearings are heterogeneous spaces, and these training sites were as restricted as possible to prevent contamination of the spectral signature from surrounding vegetation. Rather than take broad swaths of coverage from singular sites, small areas were selected from a number of sites to insure that, although the individual training site samples in each known archaeological site were small, overall, a large sample of the specific spectra in question was represented for classification purposes. Gardens and fields were much more problematic to identify tight signatures for as they are in various stages of re-growth and canopy development at different stages of cultivation. Garden sites compounded the issues of identifying spectra unique to regrowth in anthropogenic soils, as Xinguanos place a premium on the richness of *terra preta* soils, and thus tend to select such sites for cultivation. This brought up an
interesting dilemma, since training areas designated as known site locations were often also representative of known cultivated locations.

**Results**

The results of these investigations were highly mixed. The calibrated imagery offered too little separation of vegetative classes to be of any use at all. The analysis portion of this research hinged on the ability of the classification algorithms to be able to distinguish between different vegetated classes within an expanse of highly vegetated land-cover. One of the primary issues to contend with when attempting to classify the complex vegetative features found throughout the study area is related, in part, to the spatial configuration of agricultural fields, the re-use of archaeological sites as garden plots, and different stages of secondary succession within both agriculturally active and inactive regions. The relatively small size of each of these types of land cover, and the mixed spectral responses of pixels representing their classes, are responsible for difficulty in proper classification, and misclassifications in both supervised and unsupervised approaches. Several studies have shown that these problems can be overcome utilizing data with higher spatial resolution, the integration of detailed field data to support the classification process (Mausel et al. 1993, Li et al. 1994, Brondizio et al. 1996), the use of spectral mixture analysis (Adams et al. 1995), object-based classifiers (Foody et al. 1996), indices (Steininger 1996), and hybrid techniques.

The difficulty I initially faced can certainly be attributed to a lack of distinction between the classes under consideration resulting in a misclassification of the images. Simple supervised classification of the imagery was completely unsuccessful due to the inability of the imagery to encompass enough variability among the vegetative components of the landscape to be able to accurately distinguish between vegetation
growing within known site locations and vegetation growing outside of archeological sites. A maximum likelihood algorithm was chosen, but the error in the classification was simply at an unacceptable level. Relatively higher accuracy values were found for the 2002 classifications versus the 2003 classifications. The higher accuracy in 2002 is certainly due to the timing of the imagery during the middle of the dry season for the region.

Unsupervised classifications yielded poorer results yet, even when flexibility was given to the number of classes that might be formed. Overall, I was wholly unimpressed with the ability of any of ENVI’s built-in supervised or unsupervised classification schemes to delineate different types of vegetation, let alone discriminate vegetative materials reclaiming areas laden with anthropogenic soils. If the approach adopted was unable to distinguish between known sites and “natural” vegetative growth, it was readily apparent that the primary goal of the research, to provide a successful methodological approach to integrating GIS, GPS, and Remote Sensed imagery into a meaningful predictive modeling tool for past occupation sites, was simply untenable.

A new approach was needed, one that specifically addressed and emphasized the variability within the land-cover classes selected. Thus, I incorporated a number of different indices and transformations in an attempt to provide some measure of contrast between the vegetative components of the imagery.

**Data Transformations**

Utilizing the IR bands in combination with visible wavelength bands to study vegetation is broadly termed as the Vegetation Indices (VI) approach to analysis. The derivation of vegetation indices is loosely defined as "mathematical transformations designed to assess the spectral contribution of vegetation to multispectral observation"
The basic premise behind the use of vegetation indices rests on an assumption that selected algebraic combinations of remotely sensed spectral bands could correlate to the presence of green vegetation in the pixels scene. The essential characteristics of most vegetation spectra provide for chlorophyll pigment absorption in the red (R) visible bands contrasted against the high reflectivity of plant materials in other spectrum. Jordan (1969) is credited with first combining near infrared and red spectral responses into a ratio that was then shown to correlate highly with leaf-area index. In the interim, a vast number of spectral band combinations have been studied as measures of vegetation, resulting in a vast number of publications that discuss R and NIR use of the different indices to estimate vegetation variables such as percent green cover, leaf area index (LAI), absorbed photosynthetically active radiation and others either for general vegetation studies or related to forestry (Fisher 1994; Huete et al. 1994; Myneni and Williams 1994; Spanner et al. 1994). Various vegetation indices have been proposed, modified, analyzed theoretically, compared, summarized, categorized, and criticized. R and NIR combinations are typically expressed as a ratio, a slope, or other formulation that can generally be separated into three categories: intrinsic indices, soil-line related indices, and atmospheric-corrected indices (Rondeaux et al. 1996). The first types, ratio-based indices, generally address the characteristic chlorophyll absorption by vegetation in the red portion of the spectrum and high reflectance by vegetation in the near-infrared portion (Tucker 1979). Ratio-based indices include the simple ratio or simple vegetation index (variously SR or SVI in the common literature) developed by Jordan (1969), the normalized difference vegetation index (NDVI) developed by Rouse et al. (1973), and various modified versions of NDVI designed to address its sensitivity to factors such as
soil variability and atmospheric conditions. A second type of index is termed soil-line based or an orthogonal index, centered on a line in spectral space (assuming two dimensions, a plane in three dimensions, or a hyperplane in higher dimensions) along which bare soils of differing brightness will lie. Vegetation increases perpendicularly to the soil line. Kauth and Thomas (1976) developed their “Tasseled Cap” transformation for Landsat MSS data, the second component of which has become known as the greenness index, which is sometimes called the green vegetation index (GVI). Crist and Cicone (1984) have extended the analysis to six bands of Landsat Thematic Mapper (TM) data (excluding the thermal infrared band), and it is this variation that has been used during the course of this analysis. The following section describes the extent to which various transformations were applied to the data. Transformations do not require training and test data or even the construction of information classes.

Transformations are operations that place the data onto more readily interpretable axes. Supervised classification training and test data were used in this study to extract summary statistics and multivariate descriptors of analytically important land cover classes in transformed data. Transformations can be powerful sources of information for image interpretation particularly for historic imagery where reconstruction of training data may not be possible. However, assumptions supporting transformation operations must be carefully considered.

**Normalized Differential Vegetation Index (NDVI)**

Born out of the need to produce accurate information on vegetation at a global scale, remote sensing scientists have developed a number of different transformations of multispectral data that are broadly described as vegetation indices. These transformations of multispectral remotely sensed data are designed specifically to derive information on
canopy characteristics relating to total biomass, overall productivity, leaf area index (LAI), amount of photosynthetically active vegetation (PAR), and percent of vegetated ground cover (Jensen 1996). First, it should be understood that a great many of the hundreds of variations of vegetation indices available often contain redundant information, stemming from the fact that discrimination of vegetative attributes often takes place within the same narrow parts of the electromagnetic spectrum. Rather than compare the results of different vegetation indices, this research identifies one of the most commonly employed vegetation indices and applies it to the problem of locating historic and prehistoric settlement areas and forest modifications.

The normalized differential vegetation index (NDVI) is defined by the equation:

\[
NDVI = \frac{(TM_4 - TM_3)}{(TM_4 + TM_3)}
\]

Results of laboratory and field studies indicate that NDVI is strongly correlated with fractions of active photoabsorbant vegetation and leaf area index (LAI) (Teillet et al 1997; Walter-Shea 1997). As such, NDVI has become a valuable analysis tool for research into everything from global carbon cycling models to commercial crop studies. A high NDVI value is an indicator of high fractions of photoabsorbant vegetation, as well as a high LAI within observed pixel elements. Low NDVI values are usually indicative of a relative paucity of photosynthetically active vegetation, as well as a low leaf area index. If NDVI values are positively correlated with the amount of photoabsorbant vegetation or LAI contained in a pixel, then NDVI should show differences between disturbed canopies and climax forest cover. One would expect recent clearings to have very low NDVI values due to the high composition of bare soils and dead desiccating vegetation exposed after such activities have taken place relative to surrounding healthy climax canopy. By that reasoning, any culturally active
areas, including circular-plaza villages or recently cleared rocas, should produce low NDVI values due to the existence of, in the case of villages, large central plazas of bare soil surrounded by thatched-roof housing, or bare soils in the case of recent clearing for planting or other activities. Only after habitation locations have been abandoned and vegetation allowed to recolonize should former residential occupation sites or other culturally active areas begin to return relatively high NDIV values.

One would expect bodies of water to have very low NDVI values since IR is absorbed by water and these land cover types contain almost no photoabsorbant matter. Cultural clearings and bare soils (including recently burned ranching/agricultural fields immediately surrounding the park region) exhibit mean NDVI values that are one standard deviation below any of the vegetated land cover classes due to sparse vegetation and the quantity of exposed soil. It is interesting to note that the category defined as generalized soil actually has a higher NDVI value than the cultural clearing class. This is likely due to grass cover that quickly covers any undisturbed ground. Both soil and cultural clearings have large standard deviations. The cultural clearing class has large spread because the entities comprising the group are heterogeneous. Soil on the other hand should be more homogenous. Different kinds of soils reflect and absorb energy in unique ways, but the variability observed in these data is probably better explained by the presence of vegetation in pixels otherwise consistent with the mean values for a general soils classification.

Net primary productivity of Neotropical forests tends to decrease with the age of the stand. Thus, younger regenerating patches of forest should have higher NDVI. Complex limb architecture of climax forests should produce the most self-shading of any
vegetated pixel, while older gardens and actively cultivated gardens have less complex
limb architecture thereby creating a less obstructed path for radiant energy from the sun
to travel between photoabsorbant leaf surface and the space borne detector. Visual
inspection of NDVI transformed data demonstrated adequate separability of habitation
settlements (Figures 6-2 and 6-3).

![Figure 6-2. Detail of the Upper Xingu study region (2002 NDVI transform)](image)

Upon initial inspection, it was possible to discriminate and identify known
settlements as well as places known to contain areas under active cultivation. Piecewise
linear contrast enhancement provided a good way to isolate various land cover types.
Adjustments made using linear contrast enhancement drew out riparian corridors and recent meanders of major river systems. This gradient of productivity becomes apparent in NDVI transformed data, with these corridors appearing more productive because nutrients are deposited in these areas during flooding events.

Figure 6-3. Detail of the Upper Xingu study region (2003 NDVI transform)

NDVI may be helpful in defining edges of some features once they have been identified, but as a settlement discovery technique, it is woefully inadequate given the circumstance of this particular region. In this particular case, NDVI values for both riparian corridors and regrowth material in known archaeological sites were close enough
that it was difficult to separate these two classes within a single dimension of data. Shape and context play significant roles in finding past settlement sites and discriminating them from riparian corridors. Typically, Xinguano settlements are located near rivers, as are riparian corridors. However, the circular plaza construction of settlement areas contrasts to the riparian corridors’ far more linear shape.

The NDVI image is a single layer of 8 bit data rather than seven layers of raw data or any of the other multi-layered image transformations. Since NDVI creates single band 8 bit grayscale data, it has limited value for automated class extraction on its own. Suspecting the data reduction that takes place during the calculation of NDVI was responsible for its poor discriminating abilities, the methodological procedure turned to producing a number of different vegetative indices (VI) for comparison to see which had the most separability of vegetative classes.

**Transformed Normalized Differential Vegetation Index (TNDVI)**

Transformed Normalized Difference Vegetation index (TNDVI) is the square root of the NDVI. It has higher coefficient of determination for the same variable and this is the difference between TNDVI and NDVI. The formula of TNDVI is designed to produce positive values and the variances of the ratio are proportional to mean values. TNDVI indicates a relation between the amount of green biomass found in a pixel (Senseman et al. 1996). The values from the Transform normalized vegetation index range from 0 to 1.0 and can be used to find the leaf area index. Values near one indicate plant vigor and low vegetation cover near zero. The equation for the TNDVI is \( \sqrt{\left( \frac{TM4 - TM3}{TM4 + TM3} \right) + 0.5} \). The 0.5 is added to the division to keep the value positive.
Transform Vegetation Index (TVI)

The transform vegetation index is nothing more than square root of NDVI. At times this particular index is shown with the same formula used in TNDVI transformations. It is important to note, however, that in this study, the two differ, if only slightly.

Simple subtraction Vegetation Index (SVI)

This is really just a simple band math expression (SVI = TM4-TM3)

Devised Band Ratios

In an effort to attempt to discriminate vegetative elements as much as possible, I introduced two additional devised ratios. Ratio 1 was dropped from subsequent analysis, however, due to its poor discrimination abilities.

Ratio 1: TM2/TM4*TM3

Ratio 2: TM4*TM3/TM2 (an expansion on simple green ratio)

Soil Adjusted Vegetation Index (SAVI)

SAVI is the Soil Adjusted Vegetation Index that was introduced by Huete (1988). This index attempts to bridge ratio-based indices and perpendicular indices by acknowledging that the isovegetation lines are not parallel, and that they do not all converge at a single point. The initial construction of this index was based on measurements of cotton and range grass canopies with dark and light soil backgrounds, and the adjustment factor L was found by trial and error until a factor that gave equal vegetation index results for the dark and light soils was found. The result is a ratio-based index where the point of convergence is not the origin. The convergence point ends up being in the quadrant of negative NIR and red values, which causes the
isov egetation lines to be more parallel in the region of positive NIR and red values than is
the case for RVI, NDVI, and IPVI.

Huete (1988) does present a theoretical basis for this index based on simple
radiative transfer, so SAVI probably has one of the better theoretical backgrounds of the
vegetation indices. However, the theoretical development gives a significantly different
correction factor for a leaf area index of 1 (0.5) than resulted from the empirical
development for the same leaf area index (0.75). The correction factor was found to vary
between zero for very high densities to one for very low densities. The standard value
typically used in most applications is 0.5 that is for intermediate vegetation densities.

**Modified Soil Adjusted Vegetation Index**

MSAVI2 is the second Modified Soil Adjusted Vegetation Index that was
developed by Qi et al. (1994) as a recursion of MSAVI. Basically, they use an iterative
process and substitute 1-MSAVI (n-1) as the L factor in MSAVI (n). They then
inductively solve the iteration where MSAVI (n) =MSAVI (n-1). In the process, the need
to precalculate WDVI, NDVI, and the need to find the soil line are eliminated, hence the
formula is a much easier index to both code and to implement.

\[
MSAVI2 = (1/2)*(2(NIR+1)-sqrt((2*NIR+1)^2-8(NIR-red)))
\]

**Tassel Cap Transformation**

The Tassel-Cap transformation was primarily developed for and tested in
agricultural applications of remote sensing data. Kauth and Thomas (1976) developed
the tassel cap transformation originally for application to Landsat Multispectral Scanner
(MSS) data. However, Crist and Cicone (1984) developed a similar transformation to
TM data occupying the same spectral regions that Kauth and Thomas (1976) examined
(Jensen 1996: 183). Given its utility in other agricultural settings, Tassel-Cap should be
an effective means of transforming multi-spectral TM data into information more readily applicable to developing settlement modeling data.

The tassel cap transformation can be described as a vegetation index, but mathematically it is a factor analysis. TM data are highly correlated permitting band ratio transformations like NDVI, but because of this, the effective dimensionality of TM data may be less than the total number of bands recorded (Crist and Cicone 1984: 334). Knowing that high correlations exist within the data indicates that factor transformations may be particularly effective at reducing dimensionality while maintaining variability. Tassel-Cap is derived from a rotation of principle components. However, the axes are rotated according to a set of coefficients. Standard uncalibrated coefficients are applied in this use of tassel cap (Crist and Cicone 1984, Jensen 1996: 185).

If data are distributed into two perpendicular planes, then PCA may not be effective at defining the actual planes of variation along which data resides (Crist and Cicone 1984: 345). If this is the case and factors are rotated, they may intersect more meaningful axes of variation describing the data. Tassel cap factors may not necessarily be perfectly orthogonal (Crist and Cicone 1984). Laboratory and field studies of agricultural canopies indicate that TM data is distributed along two roughly perpendicular planes with a transitional zone spanning between these two (Crist and Cicone 1984: 46). Fully vegetated test plots define a plane of vegetation while bare test plots define a plane of soil. The third transition plane, roughly forming a right triangle between vegetation and soil in feature space, was defined by data from partially vegetated plots containing both vegetation and soil.
The plane of vegetation can be defined along two rotated axes of variation: greenness and brightness. The Greenness axis accounts for contrast that exists between near infrared and visible bands, while Brightness is a partial sum of data in all bands. Crist and Cicone's (1984: 347) experimental data indicate that over time re-vegetation of plots from "maximum vegetative development (high Greenness) to maturity" produces movement primarily along the plane of vegetation rather than the transition zone.

Given previous agricultural applications of the Tassel-Cap transformation, one can make several predictions regarding Xinguano settlements. One would expect recently cleared plots of land to exhibit high values in the plane of soil. However, the relative brightness in this plane may be influenced by soil moisture. Recently planted gardens should fall along the transition plane, until the garden canopy begins to close. Once garden canopy has closed blocking the soil substraight, pixels covering these kinds of spaces should exhibit very high greenness and brightness values. On the other hand, climax forest should be characterized by pixels with low values in the plane of soil with high greenness value but the self-shading of climax forest should produce relatively low brightness values compared to recent clearings. This should permit a quick estimation about land usage and relative age of vegetation within the study area, in addition to providing a means of discriminating different vegetation types across the region.

**Decorrelation**

Decorrelation stretching enhances the color separation of an image with significant band-band correlation. The exaggerated colors improve visual interpretation and make feature discrimination easier. Decorrelation stretch is based on a principal component transformation of correlated multispectral image data. In general, highly correlated image channels, such as the Red, Green and Blue (RGB) channels in Thematic Mapper
images, show subtle differences well, but these colors are not clearly related to the
different surface types. One can utilize simple contrast exaggeration to expand the range
of intensities of highly correlated images, but contrast exaggeration does little to expand
the range of colors. To enhance the color in highly correlated images requires a selective
exaggeration the least correlated portion of the spectral data (that is, one must decrease
the correlation). Decreasing the correlation of spectral data corresponds to exaggerating
the color saturation without changing the distribution of hues (or relative color
composition).

The decorrelation stretch process involves three fundamental steps: First, a
principal-component transformation is applied with the rows and columns of the
eigenvector matrix transposed. Second, contrast equalization is applied by a Gaussian
stretch, so that histograms of all principal components approximate a Gaussian
distribution of a specified variance. Third, a coordinate transformation that is the inverse
of the principal component rotation is applied so that the data are projected in their
original spectral channels, using eigenvectors as weightings for each principal
component. This inverse operation maximizes the spectral separability of different
surface types in the restored spectral channels. The decorrelation stretched images that
are created by this process can also be used as components for making color composites.

Principal Components Analysis

The fact that data are strongly correlated in more than one band (Sabins 1987:
261) allows for application of ratio transformations like the NDVI. However, these
strong correlations, either positive or negative, create redundancy in multivariate data.
Principal components transformations are a common means of reducing dimensionality in
multivariate data. PCA has been widely applied and well discussed in the analysis of
multispectral satellite data as it permits the capture of variability in multivariate measurements while at the same time reducing dimensionality.

Defining new composite axes, PCA provides a basis for investigating the primary sources of input variation in multivariate or multispectral data (Baxter 1994: 48; Jensen 1996: 172). In this application of PCA eigenvalues and eigenvectors for the components

Table 6-1. The PCA statistics for 2002 (August) transformation

<table>
<thead>
<tr>
<th>Band</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.014</td>
<td>0.439</td>
<td>0.059</td>
<td>0.050</td>
</tr>
<tr>
<td>2</td>
<td>-0.016</td>
<td>0.489</td>
<td>0.046</td>
<td>0.044</td>
</tr>
<tr>
<td>3</td>
<td>-0.015</td>
<td>0.443</td>
<td>0.036</td>
<td>0.043</td>
</tr>
<tr>
<td>4</td>
<td>-0.022</td>
<td>1.039</td>
<td>0.159</td>
<td>0.127</td>
</tr>
<tr>
<td>5</td>
<td>-0.020</td>
<td>0.617</td>
<td>0.101</td>
<td>0.099</td>
</tr>
<tr>
<td>6</td>
<td>-0.019</td>
<td>0.605</td>
<td>0.043</td>
<td>0.061</td>
</tr>
</tbody>
</table>

Eigen Values (% variance)

<table>
<thead>
<tr>
<th></th>
<th>1</th>
<th>2</th>
<th>3</th>
<th>4</th>
<th>5</th>
<th>6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.030</td>
<td>0.006</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
</tr>
<tr>
<td>2</td>
<td>0.041</td>
<td>-0.022</td>
<td>-0.210</td>
<td>0.674</td>
<td>-0.550</td>
<td>-0.500</td>
</tr>
<tr>
<td>3</td>
<td>0.597</td>
<td>0.484</td>
<td>0.405</td>
<td>-0.229</td>
<td>-0.439</td>
<td>0.032</td>
</tr>
<tr>
<td>4</td>
<td>0.449</td>
<td>0.104</td>
<td>-0.341</td>
<td>-0.189</td>
<td>0.462</td>
<td>-0.650</td>
</tr>
<tr>
<td>5</td>
<td>0.438</td>
<td>-0.142</td>
<td>-0.690</td>
<td>-0.012</td>
<td>-0.153</td>
<td>0.537</td>
</tr>
<tr>
<td>6</td>
<td>0.415</td>
<td>-0.821</td>
<td>0.387</td>
<td>0.027</td>
<td>-0.003</td>
<td>-0.058</td>
</tr>
</tbody>
</table>

are calculated from a correlation matrix rather than a covariance matrix. Using a correlation matrix produces standardized PC transformations that are more readily applied to discrimination and change detection problems (Jensen 1997: 179; Siljestroem Ribed and Lopez 1995). By creating new composite axes (PC), transformation of raw multispectral data may produce results that are interpreted more easily than raw imagery
(Jensen 1996: 172.) For TM data, PC1 generally describes the vast majority of variability in the measurements, with each band successively describing less variability between input bands. Examination of eigen values and eigen vectors demonstrates that component one explains an enormous amount of the variation in the 2002 image (72.5%) with component 2 adding an additional 13.5% and PC3 a mere 1%. In PC1, eigen vectors for all the bands of the image data are positively correlated. PC2 shows negative correlation in the eigen vectors for bands 2, 3, 5, and 6.

Table 6-2. The PCA statistics for 2003 (May) transformation.

<table>
<thead>
<tr>
<th>Band</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>SD</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>-0.014</td>
<td>0.436</td>
<td>0.055</td>
<td>0.046</td>
</tr>
<tr>
<td>2</td>
<td>-0.016</td>
<td>0.479</td>
<td>0.042</td>
<td>0.040</td>
</tr>
<tr>
<td>3</td>
<td>-0.015</td>
<td>0.441</td>
<td>0.030</td>
<td>0.034</td>
</tr>
<tr>
<td>4</td>
<td>-0.022</td>
<td>1.033</td>
<td>0.156</td>
<td>0.121</td>
</tr>
<tr>
<td>5</td>
<td>-0.020</td>
<td>0.613</td>
<td>0.093</td>
<td>0.086</td>
</tr>
<tr>
<td>6</td>
<td>-0.020</td>
<td>0.601</td>
<td>0.037</td>
<td>0.049</td>
</tr>
</tbody>
</table>

Eigen Values (% variance)

<p>| | |</p>
<table>
<thead>
<tr>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.027</td>
</tr>
<tr>
<td>2</td>
<td>0.002</td>
</tr>
<tr>
<td>3</td>
<td>0.000</td>
</tr>
<tr>
<td>4</td>
<td>0.000</td>
</tr>
<tr>
<td>5</td>
<td>0.000</td>
</tr>
<tr>
<td>6</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Eigenvectors

<table>
<thead>
<tr>
<th>Eigenvec</th>
<th>Band 1</th>
<th>Band 2</th>
<th>Band 3</th>
<th>Band 4</th>
<th>Band 5</th>
<th>Band 6</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0.269</td>
<td>0.237</td>
<td>0.194</td>
<td>0.722</td>
<td>0.494</td>
<td>0.259</td>
</tr>
<tr>
<td>2</td>
<td>0.083</td>
<td>-0.014</td>
<td>-0.197</td>
<td>0.606</td>
<td>-0.593</td>
<td>-0.485</td>
</tr>
<tr>
<td>3</td>
<td>0.639</td>
<td>0.486</td>
<td>0.387</td>
<td>-0.269</td>
<td>-0.362</td>
<td>0.044</td>
</tr>
<tr>
<td>4</td>
<td>0.451</td>
<td>0.049</td>
<td>-0.420</td>
<td>-0.192</td>
<td>0.486</td>
<td>-0.587</td>
</tr>
<tr>
<td>5</td>
<td>0.427</td>
<td>-0.262</td>
<td>-0.612</td>
<td>0.014</td>
<td>-0.192</td>
<td>0.581</td>
</tr>
<tr>
<td>6</td>
<td>0.355</td>
<td>-0.798</td>
<td>0.472</td>
<td>0.023</td>
<td>0.029</td>
<td>-0.115</td>
</tr>
</tbody>
</table>

This correlation means that when measured reflectance is high in, for example, band 3, it is very likely to be low in band 4. Likewise, when band 4 registers a high digital number band 3 is likely to be low. If band 3 records reflectance wavelengths corresponding to red light (known to be an important absorption band for green
vegetation), and band 4 is attributed to a portion of the near-infrared spectrum (also highly reflective in vegetation), then PC2 should be showing primarily differences in vegetation coverage.

PC3 eigen vectors show that TM bands 4 and 5 are negatively correlated to all other bands. Bands 5 (eigen vector = -0.439) and 1 (eigen vector = 0.597) have the strongest negative correlations. Band 1 records reflectance in a section of blue wavelengths of light corresponding to peak transmittance frequencies for water. Therefore, rivers and atmospheric moisture should be highly reflective in band 1. Conversely, band 5 is sensitive to the amount of moisture in green plants. These relationships between bands 1 and 5 indicate that PC3 is primarily relating information of differences of moisture content.

Principle component analysis of the 2003 image likewise revealed that the majority of the variance between the various bands is explained primarily by the first three PC bands (90%, 7.7%, and 1.3% respectively). Again, PC1 possessed positive correlation across all bands. PC2 once more demonstrated negative correlation in bands 2, 3, 5, and 6. Finally, PC3 once again showed negative correlation in bands 4 and 5, with bands 1 and 5 representing the greatest negative correlation.

**Revised Methodology**

After combing the available literature for a possible solution, I decided that some adjustments were necessary to achieve a higher accuracy for the final LULC classifications. First, it was obvious that the spectral similarity of the classes I had chosen would not permit an accurate classification of the original calibrated imagery. Second, the emphasis on vegetative separability provided by the NDVI and Tassled Cap transforms came at the cost of data reduction, and thus had a direct impact on the ability
of the classifier to make fine distinctions between different classes. What was needed was an expansion of the number of data sets that provided information about the minute differences between the various types of vegetation.

First, I made a distinction between the five vegetation indices without a soil reflectance control (NDVI, Tassled Cap Greeness, TNDVI, SVI, and TVI) and the two indices with a control for soil (SAVI and MSAVI). The first five were stacked together to form VI composite images for 2002 and for 2003. I then produced a PCA transform of these 5 indices to reduce the amount of data. SAVI and MSAVI were layered and transformed by PCA into two bands as wells.

As previously mentioned, PCA is often used to “compress” data sets and reduce the amount of data that must be considered. I chose PC1 and PC2 from the vegetation indices composite, as well as PC1 of the soil-adjusted indices (for this study the first PC acquired from the 2 soil indices contains spectral information adequate for the classification to normalize the effects that emerge due to the different soil types of the areas with low canopy of vegetation). I layered these bands with PC2 and PC3 from the original calibrated images. As previously explained, an examination of principal components eigenvector loadings was undertaken to determine which PC possessed information that could be related directly to the spectral signatures of vegetation. Eigenvector loadings for PC2 indicated that PC2 described the difference between the visible channels TM2 and TM3, and the infrared (IR) channels TM5 and TM6. Eigenvector loadings for PC3 indicated that TM3 and TM4 were negatively correlated owing to absorption of chlorophyll in TM3 and high reflectance in TM4. Operating under an informed assumption that PC2 and PC3 were, in fact, describing differences in
vegetative cover, PC2 and PC3 were selected as for inclusion in the final analysis composite. Decorrelation stretch results for bands 3 and 4 from the original calibrated images rounded out the final two bands of information used in the revised classification. Even though these bands still show the properties of the original bands, the color separation of these bands are enhanced with significant band to band correlation. A decrease in correlation of spectral data will result in an exaggeration of the color saturation without changing the distribution of hues (or relative color composition) (Gillespie et al. 1987). In the final analysis of the data, an assumption was made that by selecting PC1 and PC2 of vegetation indices, PC1 of soil indices, PC2 and PC3 of raw bands and DC3 and DC4 as the primary bands for analysis and classification, one could effectively remove much of the redundant data among the multivariate datasets, maximizing the potential of the classification algorithm to delineate between vegetative cover classes with more accuracy.

Landscape attributes observed during data collection provided much of the information used in refining the classification, producing “theme” layers re-incorporated into the GIS, which was then used to produce a landscape attribute informed descriptive model of the sample area. From this descriptive template, areas matching specific spectral profiles are extracted, allow us to model where uninvestigated or unknown habitation sites might be located. These areas are treated as potentially containing highly visible archaeological remains. The importance of such efforts should be apparent when compared to traditional sampling strategies, especially in reference to the relative efficiency of traditional methods in data collection over such large spatial extents. Developed to predict archaeological site yield potential on non-surveyed areas, this
approach first derives descriptive spectral attributes of known “sites” and their
encapsulating context, then extrapolates and the template to surrounding areas sharing the
same attributes.
CHAPTER 7
RESULTS

Post-Classification Procedures and GIS Manipulation

Supervised image classification is essentially a three-stage process. First, a number of training pixels, representative of their respective classes, are located in the image under consideration. These training pixels are used to calculate descriptive statistics for each class, thus defining a generalized spectral signature corresponding to each selected category. Second, based on the derived class descriptions, each pixel is then allocated to the class with which it has the greatest similarity, as assessed relative to the classifier’s decision rules. For supervised classifications using a standard maximum likelihood classification algorithm, this process labels each pixel as belonging to the class with which it has the highest posterior probability of membership (Lillesand and Kiefer 2000, Campbell 1996). Finally, the accuracy of the classified image is assessed with respect to a set of pixels for which reference or ground data on class membership is available.

Testing of the classification is usually based on an error matrix, demonstrating the correspondence between the predicted and the actual classes of membership for an independent testing set, and from which it is possible to derive a range of quantitative measures of classification accuracy. The end result of the classification process is effectively a thematic map depicting the spatial distribution of the selected classes accompanied by an accuracy statement. The accuracy of the classification is dependant on a number of variables, related to both the methodological approach selected and to the nature of the classes and remotely sensed data itself (Campbell 1996).
The supervised classification used in this study remains conventional in all aspects, only the methodological approach differs from tradition. Thus, the output for each pixel subjected to classification comprises only the code of the class with which it has the highest strength of membership. Often referred to as a “hard” or “crisp” form of classification (based on conventional crisp set theory), each pixel has a single membership in a mutually exclusive classification scheme (full membership to the named class, and zero membership to other classes). The training stage of the classification should, ideally, also be representative of conventional crisp set theory, with training pixels selected such that they belong (or are assumed to belong) to their named classes with full membership and have zero membership to other classes. However, in the case of this study, the nature of the classes under investigation (primarily those that are cultural in nature) have no firm, defined boundaries, and thus are not easily incorporated into such a theoretical mode. There are other methodologies for dealing with pixel membership (including fuzzy-set theory and logistic regression), but this dissertation did not address the application of these methods to this specific problem (leaving these other treatments as objectives in future work). Instead, GPS survey coupled with ethnographic information were used to select training sites, at times unavoidably containing mixed spectra.

Conventional classification methodologies of remotely sensed imagery assume that the area under investigation is composed of a number of unique, internally homogeneous classes that remain mutually exclusive, thus any classification based on remotely sensed data and ancillary data can be used to identify these classes with the aid of ground data (Townshend 1981, Lillesand and Kiefer 2000). It is certainly the case that such
assumptions are often invalid in areas where the classes exist as continua rather than as a mosaic of discrete classes (for example, vegetation coverages, which are rarely internally homogeneous and mutually exclusive). As a result, the classes overlap and are not separated by sharp boundaries (Wood and Foody 1993, Kent et al. 1997). Allowances must also be made for the complex relationship between spectral responses recorded by a remote sensor and the corresponding ground situations. Similar entities at different locations may possibly exhibit varied spectral responses, and, conversely, completely dissimilar entities may exhibit very similar spectral responses (Forster 1983).

To perform a conventional classification, then, requires an acknowledgement of a number of fundamental sources of error. As addressed later in this chapter, it is a balancing act between the accuracy of the classification, and the quality of the classification, necessitating that allowances be made in some cases for the sake of cost, expediency, ease of interpretation or replicability, and always mindful of what goals the end-user of the final classification product has in mind. In that spirit, no sub-pixel classification was undertaken, though it is definitely a technique that would be quite valuable to future offshoots of this research. For the scope of this dissertation, however, pixels were assumed to have belonged exclusively to one class. It should be re-iterated that membership to a specific classificatory category was further simplified by the combining of classes into more useful entities for aiding in the discussion of probable site locations, and, more broadly, the distribution of anthropogenic soils resulting from long-term human occupation and interaction within the region of interest.

Each of the classes was then converted into a vector format by the ENVI software package. Each class represented a single layer of data, and each layer was ported into a
GIS. The layers were then stacked on top of one another to duplicate the original classified image, but with each class broken into its own dataset.

Figure 7-1. Detail of the 2002 (August) supervised classification
Figure 7-2. Detail of the 2003 (May) supervised classification
Table 7-3. The 2002 (August) accuracy results for combined classes

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Water Body</td>
<td>99.64</td>
<td>99.94</td>
<td>10577/10615</td>
<td>10577/10583</td>
</tr>
<tr>
<td>Forested</td>
<td>70.24</td>
<td>62.86</td>
<td>347/494</td>
<td>347/552</td>
</tr>
<tr>
<td>Savannah</td>
<td>91.44</td>
<td>85.85</td>
<td>267/292</td>
<td>267/311</td>
</tr>
<tr>
<td>Inlet Regions</td>
<td>22.30</td>
<td>35.98</td>
<td>68/305</td>
<td>68/189</td>
</tr>
<tr>
<td>Transitional Vegetation</td>
<td>41.74</td>
<td>46.60</td>
<td>48/115</td>
<td>48/103</td>
</tr>
<tr>
<td>Bare Soils</td>
<td>96.15</td>
<td>43.86</td>
<td>25/26</td>
<td>25/57</td>
</tr>
<tr>
<td>Aldea</td>
<td>96.36</td>
<td>92.98</td>
<td>53/55</td>
<td>53/57</td>
</tr>
<tr>
<td>Pasture Outside Park</td>
<td>99.18</td>
<td>100.00</td>
<td>2528/2549</td>
<td>2528/2528</td>
</tr>
<tr>
<td>Culturally Active/Rosas</td>
<td>86.07</td>
<td>97.19</td>
<td>346/402</td>
<td>346/356</td>
</tr>
<tr>
<td>Anthropogenic vegetation</td>
<td>78.45</td>
<td>40.63</td>
<td>91/116</td>
<td>91/224</td>
</tr>
<tr>
<td>Actively cultivated sites</td>
<td>40.85</td>
<td>61.70</td>
<td>58/142</td>
<td>58/94</td>
</tr>
</tbody>
</table>

Table 7-4. The 2003 (May) accuracy results for combined classes

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Water Body</td>
<td>99.16</td>
<td>99.73</td>
<td>10526/10615</td>
<td>10526/10555</td>
</tr>
<tr>
<td>Forested</td>
<td>67.81</td>
<td>56.49</td>
<td>335/494</td>
<td>335/593</td>
</tr>
<tr>
<td>Savannah</td>
<td>77.05</td>
<td>87.89</td>
<td>225/292</td>
<td>225/256</td>
</tr>
<tr>
<td>Inlet Regions</td>
<td>38.36</td>
<td>68.02</td>
<td>117/305</td>
<td>117/172</td>
</tr>
<tr>
<td>Transitional Vegetation</td>
<td>41.74</td>
<td>30.77</td>
<td>48/115</td>
<td>48/156</td>
</tr>
<tr>
<td>Bare Soils</td>
<td>100.00</td>
<td>42.62</td>
<td>26/26</td>
<td>26/61</td>
</tr>
<tr>
<td>Aldea</td>
<td>96.77</td>
<td>21.13</td>
<td>30/31</td>
<td>30/142</td>
</tr>
<tr>
<td>Pasture Outside Park</td>
<td>78.82</td>
<td>99.74</td>
<td>774/982</td>
<td>774/776</td>
</tr>
<tr>
<td>Culturally Active/Rosas</td>
<td>81.59</td>
<td>98.20</td>
<td>328/402</td>
<td>328/334</td>
</tr>
<tr>
<td>Anthropogenic vegetation</td>
<td>75.86</td>
<td>30.14</td>
<td>88/116</td>
<td>88/292</td>
</tr>
<tr>
<td>Actively cultivated sites</td>
<td>6.34</td>
<td>75.00</td>
<td>9/142</td>
<td>9/12</td>
</tr>
</tbody>
</table>

Utilizing the maximum likelihood classifier once again in conjunction with the ground-truthed data set, I was able to produce classified images with much higher
accuracy overall than in any previous iteration. For 2002, the confusion matrix reported an overall accuracy of 86.1326% with a Kappa Coefficient of 0.7949. For 2003, the confusion matrix reported an overall accuracy of 82.7279% and a Kappa Coefficient of 0.7457.

The original 30 classes were then combined into more manageable categories. The final combined classification of the 2002 and 2003 images consisted of 12 categories. The confusion matrix for the combined categories of the 2002 classification reported an overall accuracy of 95.3478% with a Kappa Coefficient of 0.9025. For the 2003 combined classification categories, the confusion matrix reported an overall accuracy of 92.500% with a Kappa Coefficient of 0.8024.

Perhaps more interesting than the strength of the model based on measurements of overall accuracy are the achievements of the model within these specific classes denoting probable site locations. This dissertation is specifically interested in vegetative signatures of anthropogenic soils in the region of study. To that end, the classes of most interest are “Anthropogenic vegetation” and “Actively cultivated sites.” The combination of new, naturally occurring successional species, in addition to the Xinguano cultural practice of reusing past occupational sites as garden (due primarily to the richness of the anthrosols contained in such areas), the recent burning than accompanies new plantings, and the presence of older, possibly anthropogenic, vegetative classes would create an overall spectral target that would be difficult, at best, to classify into a discrete category, thus it was also deemed appropriate to maintain separate classifications of probable site locations based upon the level of contemporary modification of those regions. The active cultivation of former occupational sites is bound to produce highly mixed spectral
responses, and thus represents the underlying cause of the relatively low accuracy exemplified by the “Actively cultivated sites” category. The results for the category of “Anthropogenic vegetation” are much more promising, demonstrating a better than 75% accuracy on the producer side of the classification. Based upon these results, I feel confident in stating that the methodology used in this study has a great deal of potential for extracting probable historic and prehistoric Xinguano occupational sites from the surrounding vegetation, so long as they are not under active cultivation or otherwise in the process of cultural modification (burning, clearing, etc.).

The rise in accuracy utilizing composite images derived from the application of the methodology described above suggests that this process, undertaken to maximize the information contained in the imagery while minimizing redundant data, allowed for the expression of data substantive enough that the classifier could successfully discriminate between vegetation classes that initially possessed very similar spectral reflectance values. The results of the classifications suggest that Principle Component transformations of both raw imagery bands and a sequence of vegetation indices can extract valuable vegetation coverage information, and distill it into a concentrated form by creating a new variable set, eliminating much of the interband correlation, and greatly reducing the dimensionality of the data.

The two classes of particular interest to this study (Actively cultivated sites, Anthropogenic vegetation), representing vegetation located in the anthropogenic soils of former occupation sites, were unified into single “predicted sites” layers by year, and exported as shapefiles. The shapefiles were then overlain on the original images, and
layered with the GPS surveyed inventory of known site locations and their extents for comparison.
This “tethering” of GPS-survey data to geo-rectified satellite images allows a researcher to extract the exact location of classified features, as well as providing the ability at some future point to expand on this study and perhaps do further analysis of the relationships of features to other environmental and cultural variables.

Due to the relatively low predictive value of the “Actively cultivated sites” layer, it was discarded from a final assessment of predicted site locations. Instead, this study relied upon the “Anthropogenic vegetation” layers for 2002 and 2003. These layers were overlain on top of one another. An intersection of the two layers was performed (Figure 7-3), insuring that only those regions that had been categorized as “Anthropogenic vegetation” in both the 2002 and 2003 scenes were utilized in the last stage of analysis. Analysis of this final layer suggested that there are approximately 1800km² of vegetation within this region that exhibit spectral values similar to that of the vegetation within known occupation sites that are not currently under cultivation. By utilizing this final layer, one can readily assess areas of particular interest for future archaeological exploration based upon the size, density, and pattern of vegetation within these regions of interest, and, given the fact that these regions are geo-rectified, one can readily navigate to them using GPS. This final unification in GIS enabled a visualization of probable site locations in relation to surveyed feature components, and thus, a ready identification of those areas that hold the most potential of containing residual occupational or cultural modified soil components.

The predictive value of this methodology relies on the differential organic components of anthropogenic soils, and the manifestation of those differences in the
types of vegetation that grow in those soils. These culturally modified soils promote
distinctive species of vegetation to thrive in successional stages, distinguishing these
areas from surrounding vegetative communities. This, coupled with the Xinguano
practice of transplanting or cultivating culturally significant species in and around
occupational sites (a practice that continues today as it did in the past), allows researchers
to characterize these areas as vastly different from surrounding environments. While this
is not an absolute model in any sense of the word, the course of these investigations has
laid open a number of interesting possibilities. A unified approach, utilizing image
processing, GIS, and GPS survey is a viable one. Each of these components has
complemented the capabilities of the others in measurable ways. These investigations
have also shown that it is possible, with some degree of accuracy, to separate out
distinctive vegetative signatures over extremely large spatial extents. While both the
methodological approach and the results of these analyses are hardly groundbreaking
within geographical sciences or within the remote sensing community, the implications of
this study for an integrative approach available to archaeologists are of substantial
import. This study has demonstrated that, not only is GPS capable of accurate ground
survey, it may actually be a preferred method of collecting survey data if those data are
then to be processed within a GIS or image processing environment. The use of
numerous vegetative indices in combination with decorrelation and principle component
analysis is an approach that allows us to specifically target variations within vegetation,
and provide a sound platform upon which to perform informed classification based upon
data collected in the field. This has far-reaching implications within the archaeological
community, especially in expanding the methodological approaches to designing
predictive models, which seem to remain tied to geomorphological characteristics (proximity to water, elevation, and soil types) that themselves are highly subject to temporal transformation and are dependent, often times, on the accuracy of the original classifier, as well as the digitizer of these information layers. The procedures laid out in this study show that, rather than remain dependant on scarce or expensive data sets, there are viable alternatives to gaining an understanding of the human dimension as it plays out across a landscape in a timely, cost-effective way.

**Discussion of the Classification Process**

Map products derived from remote sensing are usually critical components of a GIS. Remote sensing is an important technique to study both spatial and temporal phenomena (monitoring). Through the analysis of remotely sensed data, one can derive different types of information that can be combined with other spatial data within a GIS. The integration of the two technologies creates a synergy in which the GIS improves the ability to extract information from remotely sensed data, and remote sensing in turn keeps the GIS up-to-date with actual environmental information. As a result, large amounts of spatial data can now be integrated and analyzed. This allows for better understanding of ecological processes and better insight in the effect of human activities. GIS and remote sensing can thus help people to take informed decisions about their environment. Like in all models, however, both maps and thematic data are abstractions or simplifications of the real world. Therefore, GIS and remote sensing can complement but never completely replace field observations. GPS survey data with was exported to a GIS for thematic development. Without direct observations in the field, the classification categories extracted from the imagery using regions of interest derived from thematic classes designated in the GIS environment would not have been possible. The lack of overall
separation of vegetative classes in the study area necessitated extensive field survey and
categorization by both visual inspection and with the help of indigenous informants.

Again, this requires redressing the issue of classificatory accuracy. In this analysis,
(as is the case in most current research), a confusion matrix was employed to estimate the
efficacy of the classifier algorithm in approximating real-world distinctions between
various coverage classes. Using a confusion matrix to estimate accuracy is a sound
approach. However, obtaining a reliable confusion matrix continues to prove the most
critical part of accuracy assessment. Potential problems are the subjectivity inevitably
induced by the choice of the classification scheme (labels), the training samples (in the
case of supervised classification), and the reference data sampling size and strategy.

When the classes are well separated in a feature space and there is no overlap
between the distributions of the categories, most classifiers should return the same result,
which will hardly depend on the choice of training sample. When the choice of classes
produces overlapping spectra or spectra that are multimodal, the classifiers may disagree,
depending on the a priori information that is incorporated into the classification models.
Moreover, even a given classifier can produce different results when trained with a
different data set.

In its purest form, a classifier is simply an automated method of pattern
recognition. Pattern recognition is defined as a process wherein a form of decision rule is
applied to produce a categorical identification sequence that assigns each object to a
discrete class (I use pattern identification here interchangeably with pattern classification,
or, more simply classification). In image processing software platforms, such as the one
used in this analysis, the process of classification utilizes the cosen algorithm to label each pixel composing the image as representing particular ground cover types, or classes.

There are inherent issues when dealing with classifications of remotely sensed images. The fact remains that the real-world components of the satellite image does not, generally consist of neatly spaced, homogeneous parcels of land. One must also accept that no natural or man-made environment is constructed to fit the raster model of data collection used by many sensor platforms. Seldom can we expect true ground conditions to be accurately modeled by remote sensing platforms. The implication of these fundamental truths of image processing necessitates an acknowledgement on the part of the researcher that, to some degree, there must be a corruption of the class signatures. Rigorous methodologies have been applied to these issues, resulting in claims that these effects can be accounted for, modeled, or reduced using various procedures (linear mixture models, fuzzy membership functions). As observed earlier, however, some level of generalization is necessary, and for the development of this analysis, it is assumed that each pixel belongs to one class only, accepting a certain degree of error, and realizing that this may have implications for the applicability of coarse data (e.g., Landsat data).

As mentioned previously, selected features should be chosen by optimizing a criterion, estimated from the ground-truthed training data. For each category used in the classification, it is critical to question how well the data represent the class overall. In essence, do the selected regions adequately sample the feature space for each class? The issues of mislabeling of pixel elements are coupled to the “representativeness” of the training data and are a direct result of issues related to mixed pixels (class overlap), transition zones, dynamic zones, within-class variability (covariance), limited training
data, resolution issues (as well as any number of uncounted variables) since classification is directly impacted by affected by the scale as well as the spatial and spectral characteristics of the image data. In terms of spatial resolution, if the size of a specific type of object imaged is smaller than the instantaneous field of view (IFOV) of the sensor, the resulting pixel will contain other types of objects as well, thus giving rise to the problem of mixed pixels. The separability issues I encountered are, in part, a result of this mixed pixel effect, as well as the other issues mentioned above. In terms of spectral characteristics, energy transfer from neighboring pixels results in mixed spectral signatures within a single pixel. Land use/cover datasets are, therefore, spectrally ambiguous within the measurable spectral resolution.

Remote sensing research has focused considerable time and energy towards finding empirical relationships between vegetation indices measured \textit{in situ} and local spectral reflectance in the hopes of defining the relationship between observable remote sensed data and actual biophysical properties of vegetative matter. By realizing an empirical relationship between vegetation indices and spectral reflectance, scientists have provided a means to extrapolate these relationships to whole landscapes. For example, NDVI has been correlated to vegetation health, biomass, LAI, productivity, and fractional ground cover among many physical properties. Unfortunately, all of these derived relationships are site specific. Often times, other aspects of the observed landscape (soil reflectance, reflectance of a thick vegetation canopy, and the attenuation coefficient for radiation in the canopy that affect a given predicted reflectance), directly impact measured reflectance values, making these empirical relationships difficult to extrapolate across regions. Perhaps this, then, is the greatest difficulty faced in this analysis. Fieldwork was
restricted to a narrowly defined geographic portion of the region. The ground-truthed data collected within this relatively small portion of the overall scenes’ extents means that narrowly defined spectra signature information was extrapolated across a much larger spatial extent. While certainly not the optimal method of ensuring a highly accurate land cover classification, this approach was necessitated by the larger goals of this project.

The most difficult part of dealing with these very issues is that they are so difficult to quantify. We can attempt to reduce classification error by utilizing carefully defined classes and increasing the number of classes. It is well established that there is a strong relationship between the number of classes, the optimum number of features and the size of the training set. When the ratio of the number of training samples to the number of feature measurements is small, the estimates of the discriminant functions are inevitably reduced, which may, in turn, influence the quality of the result. We often times bandy about the term “accuracy,” as if one could objectively determine the relation between the spectral signature of a single element in a remote-sensed image and a specific category. This is valid only in select cases in which the physical parameters used to describe the training data directly correspond to the physical parameters sensed by the remote-sensing instrument. In most cases this relation is, to say the least, tenuous. It would be more profitable to speak in terms of map signature consistency rather than speak of “accuracy” would be more appropriate to talk about the relative accuracy of the classification. The use of ground-truthed data for both the training and validation of the classification remains one of the most critical aspects of the methodology presented in this research,
and should be seen as an integral part of the accuracy assessment of an image classification method.

A common method for classification accuracy assessment is the error matrix. The error matrix compares the relationships between ground-truth data (reference data) and classified results category-by-category. From the error matrix, some important measures can be derived, such as overall accuracy, producer's accuracy, and user's accuracy. Many works have provided the meanings and calculation methods for these measures (Congalton 1991; Richards 1993; Janssen and Wel 1994; Campbell 1996; Jensen 1996). Another method to interpret the classification accuracy is to calculate Kappa coefficients (Ma and Redmond 1995; Jensen 1996; Kalkahan et al. 1997). It measures the difference between the agreement between reference data and classification results and the chance of agreement between the reference data and a random classifier. The Kappa coefficient is computed as:

\[
\kappa = \frac{N \sum_{i=1}^{r} x_{ii} - \sum_{i=1}^{r} (x_{i+} \times x_{+i})}{N^2 - \sum_{i=1}^{r} (x_{i+} \times x_{+i})}
\]

where \( r \) is the number of rows in the error matrix, \( N \) is the number of observations in row \( i \) and column \( i \) in the error matrix (i.e., the corrected classified number), and \( x_{i+} \) and \( x_{+i} \) are the marginal total in row \( i \) and column \( i \) respectively, and \( N \) is the total number of observations included in the error matrix.

I have provided the confusion matrices, in their entirety, in the Appendix of this study. The confusion matrix is a square array of numbers, laid out in rows and columns,
which express the number of sample units assigned to a particular category relative to the actual category as verified by ground-truthed information or reference data set. The importance of the confusion matrix is that it permits both descriptive and analytical statistics to be calculated. Descriptive techniques are relatively simple and include computation of the overall accuracy (division of the total correct by the total number of units) and the individual class accuracy. The latter measure can be expressed in two ways: by calculating the producer’s accuracy, which is based on the reference data (error of omission), or the user’s accuracy, based on the total number of pixels classified within a specific category (commission error). Analytical statistical calculations are useful for comparing different classification methods. Previously mentioned, the most common type of discrete multi-variate techniques for statistical tests on the classification accuracy is termed “Kappa analysis.” The estimate of kappa (also called the KHAT statistic) gives a measure that indicates if the confusion matrix is significantly different from a random result. The kappa analysis also can be used to compare different matrices from different classifiers and to determine if one result is significantly better than the other.

As has been previously stated and totally contrary to logic, the term “accuracy” is not only hard to quantify, but also hard to define, and the same holds for quality. To some, the capability of the methodology presented here to accurately identify areas where Xinguano occupation sites lie might seem lackluster. The relatively high accuracy of the classifier for finding pixels belonging to the “Anthropogenic vegetation” may indeed be offset by the extremely low predictive value of the process for finding pixels in the “Actively cultivated sites” category. However, the quality of the classification must be defined in a completely different light. The difficulty of the task at hand, considering the
relatively small region from which samples were taken, and the spatial extent to which those training data were then extrapolated certainly had dramatic effects on the efficacy of this analysis. However, the methodology by which those results were obtained highlights several key features of interest to archaeologists in particular. First, the computational complexity of the methodological process, while high, is by no means extreme. The indices utilized in this study, while simple, remain tried and test as effective means of discriminating among the types of coverage classes under scrutiny. Second, the descriptional complexity of the generated results has been kept to a minimum. Thus, I tether the term accuracy to quality, well aware that I cannot do this except by creating a link between the accuracy, and the objectives and requirement of the end users of such a methodology. Make careful note that the term “users” does not necessarily need relate to those who perform image processing and manipulation, but rather, the end user of the final product, be they field archaeologists, cultural resource managers, ethnologists, historians, or representative of any number of other disciplines. Thus, the quality of the classification is rather subjective, tied to a specific user’s objective and their requirements.

Remote-sensing image interpretation has a wide range of uses across numerous disciplines, and in any discussion of the results of this process and these analyses, it would be more appropriate to indicate how the results come to bear on a specific problem. This methodology was designed with a Neotropical archaeologist in mind, one who has a limited budget and must quickly target areas of interest for the sake of cost and expediency. With these goals in mind, I have developed a system that 1) has shown an overall accuracy over the course of two years of better than 90% for the combined land
cover classes chosen for the study; 2) there is a relatively high probability that areas
identified through this classification as being past occupational sites of Xinguano peoples
are, in fact, actual site locations; 3) that the classification has been effectively applied
over a large spatial extent, leaving open the question of how far the results of this
methodology might be extrapolated or if it might have import in other geographic
regions; 4) that the process replicated by other researches and produce the same results;
5) that the classes chosen for this study permit vegetation and other land cover classes to
be used as surrogates for other types of activity (deforestation studies, comparison of
Xinguano cultivation and its impacts on the regenerative capacity of cleared field areas
versus the impacts of ranchers on the periphery of the park); 6) that the classification
system can be used during both wet and dry seasons (with drier periods producing
slightly better overall results); 7) that the categories may be aggregated into a more
generalized classification scheme or separated into more distinct categories; and 8) that
multiple types of land uses can be recognized.

Predictive modeling has become critical as a means of identifying landscape
variables that are consistently correlated with known site distributions. By identifying
these correlates, researchers are better able to identifying uninvestigated localities that
have a high probability of containing sites based upon their geographic similarity to
known sites. There is a danger in this, however. Simply identifying new site locations
based upon the attributes of known site locations is not really making progress in the
investigation of the unknown. Instead, we are simply investigating more of the same
sorts of sites, with the added bonus of identifying areas that have not yet been surveyed
(we are merely modeling existing assumptions and expectations). This is not to say that
such undertakings are any less important than other pursuits, but I submit that we, as
archaeologists, can take it one step farther. For us to make predictions of the unknown,
we must step outside what is “expected” and employ a modeling rationale that does not
build exclusivity into its results. Thus, the underlying flaw in correlation models is
exposed. Such models are exceedingly good at illustrating the probable location of any
number of like sites based upon an approximate “type,” but without a more substantial
theoretical foundation, they cannot be expected to produce information governing “why”
or “how” such sites came to be. Here is where the integration of real, ground-truthed data
and true ethno archaeological research really comes into play. In the case of the Upper
Xingu project, the Kuikuru themselves have been instrumental in helping us to identify
the locations of past habitation areas. Rather than simply seeking correlations of
brightness values of remote sensed images to specific environmental features, this work
aims to establish a deeper understanding of the sorts of variables critical in settlement
location planning of current resident populations to make informed inferences about the
activities of related groups in the past.
CHAPTER 8
DISCUSSION

This study has focused on a few key themes within the Upper Xingu region, beginning with a concentration on the structure of the region landscape, and some insight into how those structures may have changed under the tenure of Xinguano settlement. It attempts to specifically develop a means that can be utilized to predict site locations, and, thus, primary areas of anthropological alteration. The ecological concept of landscape as a product of human/environment interaction has already been discussed at some length. This research has focused on the Xinguano landscape as a complicated mosaic of patches or ecosystems relevant to the phenomenon of distribution of prehistoric and historic Xinguano settlements. Building on findings described in the last chapter, land cover classification is used as a proxy for landscape transformation to understand how this region has evolved from a complex mosaic of human-induced transformations of vegetative content in AD 1500 to a less disturbed environment covered by closed forests. Landscape structure is defined by the spatial relationships among ecosystems, and landscape function is related to the interactions among the spatial elements (i.e., flows of energy, materials, and species). Landscape change is the alteration in the structure and function of the ecological mosaic over time (Forman and Godron 1986). This research has not focused on the function of the Xinguano landscape or most of the vegetative classes defined for the study area, although its findings can be used for this purpose in the future.
Summary and Conclusions of Methods

Scale is the temporal or spatial dimension of an object or process, characterized by both grain and extent. Resolution is the precision of measurement (grain size, if spatial). Grain is the finest level of spatial resolution possible with a given data set (pixel size for raster data). Extent is the size of the study area or the duration of time under consideration. Each of these parameters was kept under control to allow for an accurate analysis of landscape structure. In general, spatial aggregation tends to reduce the variation in spatial mosaics (Burt and Barber 1996). The qualitative and quantitative changes in measurements across spatial scales differ depending on how scale is defined. Therefore, measurements carried out at different scales may not be comparable. In addition, the exact relationship between classes varies across landscapes, creating difficulties in extrapolating from one region to another (Meentemeyer and Box 1987).

Diversity, for example, will decrease with increasing grain size, with the more rare classes becoming lost as grain becomes coarser, and dispersed classes dropping off at a much faster rate than more aggregated ones. Each coverage was classified through the lens of Landsat TM images, using geometrically and radiometrically corrected images from the same geographic region, separated temporally by a mere nine months. Thus, grain size is equivalent. The extent of each landscape was defined by the extent of the scene, as also mentioned before.

This analysis hinges on determining a quantitative methodology that can be utilized to analyze and describe the structure of landscapes. Since much of predictive modeling is still dominated by empirical approaches and case studies, there exists a pressing need to create some sort of standardized approach. A multi-temporal approach was used to characterize landscape in the study area, owing to the two years over which survey and
other fieldwork took place. The recoded classes were based on the main processes occurring in the study area and affecting landscape transformation, primarily vegetation recovery through succession, and land occupation through agriculture conversion. The delimitation of landscape boundaries for calculation of metrics is central, however, as has been discussed, a major issue involved with the calculation of any metric associated with theses classifications would need to address some fundamental sources of error, the most important of which the subjective nature of deciding where settlement boundaries can be delineated. Survey would seem to suggest that no hard and fast boundaries apply in the Xinguano cases. Transitions of vegetative classes are much more gradual, clines rather than abrupt shifts in composition.

The aim of this research was to develop a methodology useful for rapidly assessing a given area for Xinguano settlements using multispectral remotely sensed data. While this model could certainly be applied to locating current settlements, archaeologists will likely find more utility in the predictive aspects of the model and its ability to accurately delineate prehistoric or historic settlements that have been abandoned. Cultural features on the landscape are surrounded by various types of climax forest and are often located in close proximity to water bodies. Of particular interest to this research is a separation of forest that has been modified recently or in the past from areas that would not seem to hold as high a probability of containing an archaeological site. Stated another way, feature discrimination and edge detection require the separation of cultural entities from a surrounding unmodified forest baseline. However, species composition creates conditions that produce a variable and heterogeneous baseline. Additionally, as culturally modified clearings age they absorb and reflect light in different ways.
Literature on anthropological applications of space borne multispectral data like Landsat emphasizes classical set theory approaches to analysis. Discussion tends to focus on using the signal of known locations to produce classifications of image data. To date, this approach has been extremely successful, and thus, commonly employed as a means of extracting information, most especially in the Neotropics. In most cases, however, only portions of scenes are utilized. The dispersed clustering settlement pattern noted among the Xinguanos may encompass even greater range than can be captured in a single Landsat scene, but the complications encountered during the course of this study would certainly be exponentially increased in one were to utilize multiple scenes in the classification process. The larger the scale of the analysis, the more difficulty signature extension would pose.

The use of transformations in concert with image interpretation, as well as manual feature extraction through the use of ground-truthed data in combination with GPS for accurate position measurements is an under exploited methodology in anthropological and archaeological applications of remote sensing. The importance of image transformation cannot be stressed enough. While it is true that, in and of themselves, image transformations do not automate feature definition, they do have the ability to dramatically increase the level of information available for interpretation across space and through time. This is especially true of research focused on vegetation, as the baseline may be quite heterogeneous. Actively cultivated land, whether on the site of previous occupational debris or in newly cleared areas, as well as other cultural features, may indeed be distinct from the surrounding vegetation. However, at issue is not appearance, but if, in fact, these cultural features are statistically distinct from other kinds
of vegetation coverage within the limits of the data collection abilities of the remotely
sensed imagery platform, or even within all contexts.

Recalling that, in large part, multispectral remote sensing of vegetation relies on
differences in the pattern of canopy shading, it stands to reason that while multispectral
data are quite useful for discriminating material differences in land cover, particularly in
the context of geology, medium resolution detectors like Landsat may not discriminate
differences in vegetated land cover through material specific absorption bands. Landsat
has sufficient spectral resolution to discriminate between vegetation and soil based on
variation in energy absorption, but the detector may not have sufficient spectral
resolution to detect differences in sub-classes of vegetation without the addition of other
analysis tools to either increase the sensitivity of the data to variations between
reflectance measurements, or by utilizing methodologies that seek to push analysis to the
sub-pixel level (e.g. Spectral Mixture Analysis). The Landsat platform can, however,
distinguish structural differences in vegetated canopies that are consistently related to
land cover classes, and it is these structural differences that separate cultural features
from climax forest. Thus, context is absolutely critical to feature identification. Simply
examining a 5-4-3 color composite with a piecewise linear contrast enhancement or
histogram equalization applied to unprocessed data can generate an enormous amount of
information through visual inspection. Pattern recognition capabilities make any human
observer likely to be able to distinguish archaeological sites from surrounding vegetation
with little instruction. The regular patterning of Xinguano archaeological sites stands, at
times, in sharp contrast to the seemingly irregular patterns of many vegetative features. It
is important to note that exploration of raw geometrically and radiometrically corrected
data does have significant merit in attempts to develop settlement information, and predict the location of archaeological sites. However, 8-bit data displayed on a cathode ray tube can only represent 256 values in three bands: red, green, and blue. Multispectral TM data is collected in seven bands each with 8-bit quantization. Visual exploration of raw imagery excludes four dimensions of data variability. The strength of digital multispectral data is the ability to examine multivariate relationships, and the ability to quantify what patterns actually make a feature cultural rather than natural, as well as allow for replicability. The point of this exercise is to demonstrate that there does exist a methodology that provides for objective assessment of the potential of a specific region to contain archaeological sites derived from an inductive model formed by an amalgamation of different anthropological data sources, some traditional and some more technologically informed.

Classifications and transformations build relationships between variables producing an image that contains information rather than only brightness values in a given portion of the electromagnetic spectrum. Image processing creates a scene composed of values that can be related more directly to real world entities in the study area. By knowing how radiant energy interacts with matter, by intelligently sampling multiple absorption bands of energy, and by examining the relationships between variable outcomes in many wavelengths, one can see that remote sensing is a powerful tool for analysis of spatial phenomena.

Land-use/land-cover classification in Amazonia is, at best, an extremely complex, arduous task, and its degree of difficulty increases with the number of classes one wants to distinguish. The spectral signatures detected by Landsat TM sensors produce images
that, more often than not, contain mixed responses for the heterogeneous tropical
environment. What would appear to be a straightforward case of assessing and defining
the separability of vegetative cover classes becomes complicated by the fact that distinct
vegetation types simply do not exist. Cover classes, instead, are forced to encompass
scenarios within the complex dynamics of vegetation clearing or recovering, and
assigning these cover classes is a tricky enterprise, attempting to tease out minute
variations in the spectral signatures of vegetation at various stages of successional
growth, within a variety of soil environments. The area under question is remote and
quite large, with a lack of ready access. The paucity of data with direct bearing on this
research, such as soil maps and extensive archaeological investigations, also complicates
the process of classification.

A number of issues creep into the equation. Some problems faced are simply
operational. For example, the choice of an August image from 2002 and a May image
from 2003 might, at a glance, seem problematic. August comes at the end of a very dry
period in the Xingu. Vegetation seems, on visual inspection, more clearly delineated in
May. On the other had, comes at the tail end of the wet season, thus vegetation is
suffused with moisture and at perhaps it greenest. Optimally, both images under analysis
would have been from the August time period, the loss of Landsat 7’s scan-line corrector
(SLC) in late May of 2003 meant that any imagery collected for August would have been
SLC-off data, vastly complicating the analysis.

Other questions are more methodological, such as the reasonable implementation of
a classification system designed to describe the variance of vegetation class dynamics at a
scale and complexity that would permit the separation of vegetation located in
anthropogenic soils from surrounding cover. It would seem obvious that the best way to overcome the difficulty of discriminating vegetation types would be an increase in fieldwork efforts to gather more ground-truth data. In the case of the Upper Xingu, the spatial extent is extensive. As the efforts of this study were to provide for predictive modeling of site locations outside the park boundaries, the entire scene was utilized, providing enough area to include multiple scenarios within the distinct landscapes and classificatory outcomes. During fieldwork, data was collected as GPS survey was being conducted. Indigenous informants often accompanied the survey team, providing an incredible resource for identification of different vegetation types, as well as assisting in the location of archaeological sites.

Accuracy assessment of the classification process brought attention to some critical issues. The original classification schema seemed to have a great deal of difficulty distinguishing between vegetation within archaeological sites and vegetation comprising the surrounding environment. Perhaps the distinction between these two areas has been exaggerated more than is warranted. It is entirely conceivable that the anthropogenic nature of the soil content within archaeological sites has less bearing on the flora composition of those areas than does the legacy of a managed landscape. Perhaps the similarity between the vegetation internal to any given site and external to it cannot be completely delineated. One way of reducing the risk of misclassification would be to group classes such as archaeological site, and the areas surrounding the sites into a single class of “culturally modified.” While this would likely have made quite a bit more sense from some perspectives (not the least of which would be ease and accuracy of the
classification itself), it would come at the cost of generalizing coverage classes to prevent the predictive model from localizing areas with a high probability of being “sites.”

Additionally, there was no real distinction at all between culturally cleared areas within the park boundaries (resulting from indigenous land-use techniques) and burned or cleared pasture areas outside the park boundaries, save for the massive size of the external clearings, as well as their more regular shape. This point is important enough to be driven home again. The overall land management practiced by the indigenous populations within the park showed small, dispersed, cleared areas, with plenty of climax forest between these regions. Pastureland outside the park boundaries indicated a wholesale clear-cutting of what were obviously recently forested areas. Perhaps even more distressing was the proximity of these cleared pastures to the park boundaries. Nearly the entire southern boundary of the park had cleared pasture all the way up to the boundary itself. In some cases, there appeared to even be some encroachment into the park itself, a possibility that, while outside the scope of these investigations, certainly warrants further investigation at a later time. Based purely on reflectance, however, the similarity of values for indigenous clearings and external pasturelands, explained by the greater contribution of soil spectral response to the signature of sparsely covered grassy vegetation.

The rapid and aggressive regrowth of secondary vegetation in tropical areas has long been a topic of discussion. Different regeneration patterns occur depending on land management practices following deforestation. Perhaps one important question not covered by this research is the role of species composition within the different stages of regrowth, and within different regions cultivated by different cultural groups. Studies
have shown that disturbance from slash-and-burn agriculture affects species composition much more than stand structure and biomass (Uhl 1987). Although such an issue is of central relevance to the maintenance of local and regional biodiversity, the utilization of remotely sensed data to classify distinct vegetative communities on the basis of species composition remains little more than a distant possibility. At best, current applications using imagery at present resolutions limits the scale at which researchers may conduct investigation. Presently, one can only hope to recognize different structural patterns and processes.

Even so, the enhanced capabilities provided by an ever increasing repertoire of applications and sensor technologies are changing the face of analysis for Earth surface feature information, spawning new approaches and moving researchers towards a more integrative vision about LULC change within and across research sites (Burrough and Frank 1995). The expansion, refinement, and integration of remote sensing and geoprocessing techniques provide for the manipulation of spatial data at several scales, further highlighting the mutual benefits of closer links between Geographic Information Systems (GIS) and methods of spatial data analysis. There are quite a number of major hurdles to integrating cross-platform spatial datasets, and though these issues have been well documented in general GIS literature (see Ehlers 1992; Goodchild et al. 1992), most have received little attention overall. The primary impediments to a fully integrated system of image processing and spatial analysis stem from a) The different spatial resolutions of the various data sets; b) the prevalence of errors in individual data sets compounded by integration; c) physical constraints in integrating raster- and vector-based data sets effectively; and d) matching data sets captured over appropriate time intervals.
To be truly successful, integration of satellite imagery, GPS data sets, and GIS analysis, the researcher must consider possible constraints caused by the differences in their various spatial resolutions. I refer not only to the value of the actual datasets themselves, but also to the degree of compatibility between each of the data sets. Specifically, one should question whether the varied resolutions of the source data actually compliment each other. Can survey data collected from differential GPS instrumentation, providing accuracy down to a sub-meter level, be fully integrated into the 20-30 meter scale of mid-resolution satellite imagery, combined with the scale of DEM maps or LULC classification digitized into GIS data layers? They certainly can compliment one another, so long as the scalar issues are addressed (in the case of this study by limiting the scope of the analysis to the sensor platform with the lowest resolution).

It is a simple fact that some form of error will likely be encountered in any given spatial data set. These errors may result during data collection, manipulation, interpretation, or presentation (e.g. image capture for remote sensing data and the digitizing of data for input into GIS). These types of error can, for the most part, be mitigated through judicious use of data quality assessment, carefully noting throughout the process potential error sources and incorporating error estimates. The implication of the raster/vector data representation of spatial data is an on-going issue of relevance to many GIS based investigations. Ehlers (1992) has suggested that a fully integrated single system is the only effective way of resolving the different representations of space as given currently by raster- and vector-based data sets.

One of the most critical aspects of integrating data is to insure that data capture occurs either at precisely the same time or within a reasonable span (such as within the
same season). If one lacks a good temporal match, it can frequently result in a degraded data synthesis, introducing yet another source of error into the equation. There seems to be a real lack of recognition or failure to document how data integration problems have been overcome in the current literature, making the comparison of results between archaeological survey studies, LULC classification and inventory studies, and overall the approaches taken to the spatial analysis of processed imagery problematic and, in some cases, impossible. In the same way that archaeologists have established tried-and-tested methods for the analysis and comparison of traditional field mapping techniques, there is a need for the GIS/RS community (specifically of archaeologists who wish to avail themselves of these technologies) to develop a similar approach towards an established methodology or, at the very least, a strategy for the integrated use of multi-source digital data sets for feature detection, classification, and modeling.

I propose a multisource strategy that relies first on the use of GPS in fieldwork survey situations. Not only has it been shown that GPS is an extremely effective way to collect survey data (especially in inhospitable, or, in some cases, impassible environments), but GPS can allow for targeting of specific ground-truthed control point for georectification of imagery data, as well as precise collection of regions of interest (ROIs) for later classification. Researchers can avail themselves of common remote sensing algorithms, such as filtering and textural enhancement procedures enabling feature detection, as well as classification (especially in conjunction with ground-truthed data sets), resulting in raster data that can provide detailed information over large spatial extents, at regular intervals, and for relatively little cost. It is though a GIS that these various data sets (vector and raster) can be fused together into unified spatial data.
The potential for using digital remotely sensed data as a primary source of data for GIS has received considerable attention in recent years, particularly in applications related to natural resources, as well as research conducted at large spatial scales. Data remains the most expensive component of a GIS, and the need for inexpensive, accurate, and current GIS data has created a huge demand for remotely sensed imagery of all types. The fundamental issue at hand, however, is the integration of a vector data structure into a raster model, and thus GIS-based data sets remain separated from remotely sensed data at a fundamental level. The vector processing capabilities of GIS can easily be used to delineate areas in an image for processing. The real power of image processing platforms derives from their ability to utilize pattern recognition, edge extraction, and segmentation algorithms, processes not yet available at equivalent functionality levels in commercial GIS applications. In a perfectly integrated system, one can see the utility of vector polygons extracted from images using edge detection methods that can then be modified according to rule sets based on the value of the image pixels. The argument has been made that most image processing platforms have adequate vector conversion utilities such that the burden of integrated design should fall on the shoulders of the GIS community. However, close inspection of raster-derived lines and boundaries is that they contain a stepped appearance betraying their raster origin. On the other hand, one would expect GIS functionality to incorporate some form of pixel interpolation to produce smooth boundaries, especially considering that both vector and raster data are often used within the confines of this one platform. Perhaps the terminology, then, is at fault. The integrated approaches discussed here (and the approach taken in this study) should more likely be deemed an “interfacing” of the technologies rather than a true “integration.”
Consider, though, that the most important concern is to have GIS users using remotely sensed data, and remote sensing scientists using GIS in combination to maximize data extraction, and not to quibble over the semantics of the situation. It could be well argued that the best way to achieve this goal is to provide a software environment wherein a researcher could use the best format for their specific data and have the software able to provide a level of functionality facilitating cross platform integration.

Remote sensing studies over the past few years have steadily moved from empirically based image classifications, mapping, and land-use/land-cover inventories to more deterministic modeling of scene characteristics including more knowledge-based image interpretation. GIS researchers have expanded from simple map overlay and relational models to spatially distributed simulation modeling, veiwshed analyses, and informed three-dimensional analytical models. To progress, researchers in remote sensing and GIS will need to possess hardware and software that can aid in the integration of these platforms, including improved interfaces between image processing, GIS, database management, and statistical software packages. It will be difficult going, and the reluctance of commercial software developers to take the lead in these efforts suggests that such a product will have to rise from open-source efforts. The potential, though, is far to tantalizing to not work towards this common goal. The possibility of testing spatial models through the use of georeferenced databases and algorithms to measure spatial heterogeneity, to fully integrate remotely sensed imagery with GIS applications, to effectively turn the established methodological approaches to endeavors like survey, mapping, classification, and multispectral, multi-temporal analyses on their
head lays bear a fundamentally new way in which researchers can attempt to understand land-use, as well as human impacts, in the Amazon.

**Land Use/Land Change in Amazonia**

**Perspectives**

The search for some means of quantitatively analyzing and describing the structure of landscapes has become a high priority in both social and environmental sciences. In the midst of contemporary archaeological methodologies still dominated by empirical approaches and case studies, perhaps it is time to explore different means of information gathering and modeling. This dissertation represents an effort to develop a more integrative scientific vision addressing the complex interactions between people and the environment, carried out through a merging of the old and the new, traditional approaches to data collection and utilization of new technological advances to augment and refine a new model of human-driven landscape formation.

Approaches may differ slightly in the examination of questions of human-environmental interaction at multiple temporal and spatial scales. Human ecological studies form the foundation of new methodologies, like historical ecology and landscape ecology, investigating the science of human alteration of landscapes. Historical ecology in particular has provided new schema for researching human ecosystems. However, these new ways of looking at anthropological problems have not been fully integrated into anthropological theory. In addition, there remains a relative paucity of methodological procedure and standardization in research. This makes the study of the human dimensions of landscape change a real challenge.

In the current scientific environment, priorities for research have seen a decided shift towards an inclusion of the study of processes across temporal, spatial, and socio-
cultural axis (Arrow et al. 1995; Stern et al. 1992; Turner et al. 1990). Issues of land use/land change are central to these new lines of inquiry, due to their intricate dynamics and its consequences in landscape structure and function (Lambin 1997; Turner et al. 1995; Turner 1997). This is particularly relevant in the Upper Xingu because of the landscape transformation evidenced in archaeological investigations, and, more specifically, its implications for shifting the view of prehistoric human-environmental interaction from a stagnant, limiting, and reactionary subsistence to a dynamic interplay of long-term wide-scale landscape alteration. In this sense, this investigative research into alternate methodological approaches to detecting these changes can offer an opportunity to broaden the discussion of the human dimensions of landscape pattern and process. This dissertation has sought to create a hybrid approach of traditional techniques and an integrative use of widely available technologies to create a new means of answering these questions of landscape transformation in a more efficient, expedient way, and over a larger spatial scale than would be feasible using full-scale excavation. The landscape characterizations presented in this study permit us to attain a broader perspective regarding the larger debate about the role of humans in the formation of landscapes, and the impacts of long-term human occupations on the environment in non-western tropical settings.

As knowledge about these processes increases through studies at several scales and approaches, there is growing agreement on the need to investigate culture and context, before testing general hypotheses about human-induced ecological outcomes (Lambin 1997). Thus, this study provides a methodological approach as an exploratory step towards a broader understanding of these issues. Through this base-line extraction of
features from image data, we can set the groundwork for future research into issues of
settlement pattern analysis and a deeper understanding of landscape transformation
through time. It should be remembered however, that while the techniques described in
this dissertation may be of use elsewhere, when dealing with the complex relationships
between people and the environment, different environmental outcomes may result from
similar socioeconomic dynamics, and distinct socioeconomic dynamics may produce
similar environmental outcomes. Simply put, while this dissertation may provide a useful
alternative approach to research in the Upper Xingu, its usefulness may be limited. Only
further testing in other regions will bear out the utility of this specific methodology
beyond the cultural and geographic confines of this particular study area.

Beyond the elaboration of a methodological strategy based on multi-disciplinary
integration for the study of human-environmental interaction and landscape
transformation in the Upper Xingu, the results of this dissertation can contribute to the
body of work regarding the use of new sensor platforms and processing techniques to
improve the classification of different types of forest, specifically relating to
anthropogenic alterations of vegetation structure. Achieving more accurate information
on these types of classificatory categories represents a crucial step to inform ongoing
research into settlement pattern analysis and the size and nature of Pre-Columbian human
occupation of the Upper Xingu, as well as the human dimensions of environmental
change.

Although the majority of the Upper Xingu region is composed of closed tropical forest,
the landscape retains a degree of distinctive ecological zones. Acting as clines, these
zones blend into one another (again, the probably source of much of the error in the
classification schema presented in this dissertation) making distinct boundaries difficult to define. During the course of the Upper Xingu Project, headed by Michael Heckenberger, a small sample of this region was discovered to have undergone massive alteration by human agents (Heckenberger et al. 2003). Based on his findings, Heckenberger (2005) hypothesized a significant portion, perhaps as much as fifty percent or more, of the upland (non-inundated or *terra firme*) forests and adjacent wetlands, are the result of human transformation of the landscape. Balee (1989) has also suggested that one of the most striking features of the region is the degree to which it has been altered by the Xinguanos, creating vast anthropogenic landscapes. Heckenberger (2005: 32) has suggested that “the conclusion that much of the landscape was not only anthropogenic in origin but intentionally constructed and managed is inescapable the more the scale of ancient settlements and their ‘monuments’ (e.g., plaza and causeway peripheral mounds and massive ditches) are investigated.” The pattern of anthropogenic vegetation uncovered through this investigation adds to a mounting body of evidence that would seem to reaffirm this hypothesis.

Today, I would not accept any part of the forest to be “pristine” without a detailed examination on the ground, at least in the territory of the Kuikuro (Xinguano Carib) community where the present study largely took place. In place of small paths in the forest and minor openings related to plaza villages and gardens, I now envision tree lined causeways, well maintained and broad roads, large, patchy tracts of agricultural fields, leading out from the towns and villages that make up the skeleton of Xinguano history, and an equally constructed wetland environment, including major transportation canals, managed ponds, reservoirs to improve fishing, drinking and bathing reservoirs, raised causeways, wells, raised fields, road systems, among other features (Heckenberger 2005: 33).

**Future Directions**

Within anthropological study, we often find during the course of our research that as we attempt to answer a handful of questions, we often succeed in raising dozens of
new questions. The findings presented in Chapters 6 and 7 shed a great deal of light on
the issues posited in Chapter 1 with regard to the nature of the anthropogenic
transformation of much of the Upper Xingu region. This research has made important
steps to refuting the hypothesis of static, small-scale communities, instead building on the
hypothesis of large-scale, interconnected communities dramatically altering their
environment to suit their needs. At the same time, however, in keeping with the tone of
this new methodological approach to answering such questions, it is readily apparent that
further studies will be necessary to uncover other socioecological processes that affected
the trajectories of landscape transformation. The findings of this research can be built on
to delve further into the Xingu region addressing still more issues including:

• How can other remote-sensing platforms (i.e. soil-resistivity, ground penetrating
  radar) improve our understanding of the formation of *terra preta*?

• How do the patterns uncovered by this classification fit add to our knowledge of
  Xinguano settlement patterning?

• What are the implications do the presence of Pre-Columbian large-scale sedentary
  populations in the Upper Xingu have for finally putting to rest the Tropical Forest
  Tribe concept?

• Does the dramatic alteration of landscape on such a large scale and the apparent
  interconnectedness of communities add to the growing body of research into
  possibility of chiefdom formation in Amazonia?

These questions should be addressed in a multi-disciplinary fashion, with social
sciences and environmental sciences helping to identify driving forces affecting LULC
dynamics and landscape change. If nothing else, the research presented here should
highlight the value of integrating new forms of data collection into the methodological
approaches to such questions. In this sense, various initiatives are appropriate to
maximize our data collection efforts including:
• Increasing temporal resolution (i.e., other satellite image dates) to have a better control over LULC dynamics
• Testing new sensors to scale the analysis down and up (e.g., IKONOS and MODIS)
• Testing new geoprocessing techniques to improve the accuracy of LULC classifications (e.g., sub-pixel classification, linear mixture models, multivariate analysis, spatial-spectral classifiers)
• Increasing the use of ancillary data for the study of landscape structure and change (specifically soil studies)
• Increasing fieldwork efforts to gather data on recent trends regarding actors' land-use decisions

We should prioritize further research into ADE, more extensive archaeological investigations of the probable site locations suggested by the classifications presented here (in essence, further ground-truthing of findings), and the additional integration of ethnographic information into the archaeology of the region, combined with the use of GPS survey, GIS, and additional LULC assessments to aid in our achieving a finer understanding of Xinguano's decisions regarding land use, both in the past and in the present.
Table A-1. The 2002 (August) combined classes confusion matrix

Overall Accuracy = \( \frac{14408}{15111} \) 95.3478%
Kappa Coefficient = 0.9025

<table>
<thead>
<tr>
<th>Class</th>
<th>Water body</th>
<th>Unclassified</th>
<th>Forested</th>
<th>Savannah</th>
<th>Inlet areas</th>
<th>Culturally active</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unclassified</td>
<td>38</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2</td>
</tr>
<tr>
<td>Water body</td>
<td>10577</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>6</td>
<td>0</td>
</tr>
<tr>
<td>Forested</td>
<td>0</td>
<td>347</td>
<td>11</td>
<td>135</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Savannah</td>
<td>0</td>
<td>6</td>
<td>267</td>
<td>14</td>
<td>15</td>
<td>0</td>
</tr>
<tr>
<td>Inlet areas</td>
<td>0</td>
<td>97</td>
<td>0</td>
<td>68</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Culturally active</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>1</td>
<td>346</td>
<td>0</td>
</tr>
<tr>
<td>Transitional</td>
<td>0</td>
<td>6</td>
<td>14</td>
<td>13</td>
<td>5</td>
<td>0</td>
</tr>
<tr>
<td>Bare soil</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>31</td>
<td>0</td>
</tr>
<tr>
<td>Anthropogenic vegetation</td>
<td>0</td>
<td>36</td>
<td>0</td>
<td>46</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Pasture outside park</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Aldea</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Actively cultivated site</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>22</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>10615</td>
<td>494</td>
<td>292</td>
<td>305</td>
<td>402</td>
<td></td>
</tr>
</tbody>
</table>

Ground Truth (Pixels)

<table>
<thead>
<tr>
<th>Class</th>
<th>Transitional</th>
<th>Bare soil</th>
<th>Anthropogenic vegetation</th>
<th>Pasture outside park</th>
<th>Aldea</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unclassified</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>17</td>
<td>0</td>
</tr>
<tr>
<td>Water body</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Forested</td>
<td>39</td>
<td>0</td>
<td>8</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Savannah</td>
<td>4</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Inlet areas</td>
<td>2</td>
<td>0</td>
<td>17</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Culturally active</td>
<td>2</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>1</td>
</tr>
<tr>
<td>Transitional</td>
<td>48</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Bare soil</td>
<td>0</td>
<td>25</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Anthropogenic vegetation</td>
<td>11</td>
<td>0</td>
<td>91</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Pasture outside park</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>2528</td>
<td>0</td>
</tr>
<tr>
<td>Aldea</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>53</td>
</tr>
<tr>
<td>Actively cultivated site</td>
<td>9</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td><strong>Total</strong></td>
<td>115</td>
<td>26</td>
<td>116</td>
<td>2549</td>
<td>55</td>
</tr>
<tr>
<td>Ground Truth (Pixels)</td>
<td>Actively cultivated site</td>
<td>Total</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------------</td>
<td>--------------------------</td>
<td>-------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unclassified</td>
<td>0</td>
<td>57</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Water body</td>
<td>0</td>
<td>10583</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Forested</td>
<td>11</td>
<td>552</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Savannah</td>
<td>5</td>
<td>311</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inlet areas</td>
<td>5</td>
<td>189</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Culturally active</td>
<td>5</td>
<td>356</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transitional</td>
<td>17</td>
<td>103</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bare soil</td>
<td>1</td>
<td>57</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anthropogenic vegetation</td>
<td>40</td>
<td>224</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pasture outside park</td>
<td>0</td>
<td>2528</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aldea</td>
<td>0</td>
<td>57</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Actively cultivated site</td>
<td>58</td>
<td>94</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>142</td>
<td>15111</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ground Truth (Percent)</th>
<th>Class</th>
<th>Water body</th>
<th>Forested</th>
<th>Savannah</th>
<th>Inlet areas</th>
<th>Culturally active</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unclassified</td>
<td>0.36</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.50</td>
</tr>
<tr>
<td>Water body</td>
<td>99.64</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Forested</td>
<td>0.00</td>
<td>70.24</td>
<td>3.77</td>
<td>44.26</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Savannah</td>
<td>0.00</td>
<td>1.21</td>
<td>91.44</td>
<td>4.59</td>
<td>3.73</td>
<td>0.00</td>
</tr>
<tr>
<td>Inlet areas</td>
<td>0.00</td>
<td>19.64</td>
<td>0.00</td>
<td>22.30</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Culturally active</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.33</td>
<td>86.07</td>
<td>0.00</td>
</tr>
<tr>
<td>Transitional</td>
<td>0.00</td>
<td>1.21</td>
<td>4.79</td>
<td>4.26</td>
<td>1.24</td>
<td>0.00</td>
</tr>
<tr>
<td>Bare soil</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>7.71</td>
<td>0.00</td>
</tr>
<tr>
<td>Anthropogenic vegetation</td>
<td>0.00</td>
<td>7.29</td>
<td>0.00</td>
<td>15.08</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Pasture outside park</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Aldea</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Actively cultivated site</td>
<td>0.00</td>
<td>0.40</td>
<td>0.00</td>
<td>7.21</td>
<td>0.75</td>
<td>0.00</td>
</tr>
<tr>
<td>Total</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Ground Truth (Percent)</th>
<th>Class</th>
<th>Transitional</th>
<th>Bare soil</th>
<th>Anthropogenic vegetation</th>
<th>Pasture outside park</th>
<th>Aldea</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unclassified</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.67</td>
<td>0.00</td>
</tr>
<tr>
<td>Water body</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Forested type</td>
<td>33.91</td>
<td>0.00</td>
<td>0.00</td>
<td>6.90</td>
<td>0.00</td>
<td>1.82</td>
</tr>
<tr>
<td>Savannah</td>
<td>3.48</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Inlet areas</td>
<td>1.74</td>
<td>0.00</td>
<td>14.66</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Culturally active</td>
<td>1.74</td>
<td>3.85</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>1.82</td>
</tr>
<tr>
<td>Transitional</td>
<td>41.74</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Bare soil</td>
<td>0.00</td>
<td>96.15</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Class</td>
<td>Actively cultivated site</td>
<td>Total</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>-----------------------------</td>
<td>--------------------------</td>
<td>-------</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unclassified</td>
<td>0.00</td>
<td>0.38</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Water body</td>
<td>0.00</td>
<td>70.04</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Forested type</td>
<td>7.75</td>
<td>3.65</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Savannah</td>
<td>3.52</td>
<td>2.06</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inlet areas</td>
<td>3.52</td>
<td>1.25</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Culturally active</td>
<td>3.52</td>
<td>2.36</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transitional</td>
<td>11.97</td>
<td>0.68</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bare soil</td>
<td>0.70</td>
<td>0.38</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anthropogenic vegetation</td>
<td>28.17</td>
<td>1.48</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pasture outside park</td>
<td>0.00</td>
<td>16.73</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aldea</td>
<td>0.00</td>
<td>0.38</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Actively cultivated site</td>
<td>40.85</td>
<td>0.62</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>100.00</td>
<td>100.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Class</th>
<th>Commission (Percent)</th>
<th>Omission (Percent)</th>
<th>Commission (Pixels)</th>
<th>Omission (Pixels)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water body</td>
<td>0.06</td>
<td>0.36</td>
<td>6/10583</td>
<td>38/10615</td>
</tr>
<tr>
<td>Forested type</td>
<td>37.14</td>
<td>29.76</td>
<td>205/552</td>
<td>147/494</td>
</tr>
<tr>
<td>Savannah</td>
<td>14.15</td>
<td>8.56</td>
<td>44/311</td>
<td>25/292</td>
</tr>
<tr>
<td>Inlet areas</td>
<td>64.02</td>
<td>77.70</td>
<td>121/189</td>
<td>237/305</td>
</tr>
<tr>
<td>Culturally active</td>
<td>2.81</td>
<td>13.93</td>
<td>10/356</td>
<td>56/402</td>
</tr>
<tr>
<td>Transitional</td>
<td>53.40</td>
<td>58.26</td>
<td>55/103</td>
<td>67/115</td>
</tr>
<tr>
<td>Bare soil</td>
<td>56.14</td>
<td>3.85</td>
<td>32/57</td>
<td>1/26</td>
</tr>
<tr>
<td>Anthropogenic vegetation</td>
<td>59.38</td>
<td>21.55</td>
<td>133/224</td>
<td>25/116</td>
</tr>
<tr>
<td>Pasture outside park</td>
<td>0.00</td>
<td>0.82</td>
<td>0/2528</td>
<td>21/2549</td>
</tr>
<tr>
<td>Aldea</td>
<td>7.02</td>
<td>3.64</td>
<td>4/57</td>
<td>2/55</td>
</tr>
<tr>
<td>Actively cultivated site</td>
<td>38.30</td>
<td>59.15</td>
<td>36/94</td>
<td>84/142</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Water body</td>
<td>99.64</td>
<td>99.94</td>
<td>10577/10615</td>
<td>10577/10583</td>
</tr>
<tr>
<td>Forested type</td>
<td>70.24</td>
<td>62.86</td>
<td>347/494</td>
<td>347/552</td>
</tr>
<tr>
<td>Savannah</td>
<td>91.44</td>
<td>85.85</td>
<td>267/292</td>
<td>267/311</td>
</tr>
<tr>
<td>Inlet areas</td>
<td>22.30</td>
<td>35.98</td>
<td>68/305</td>
<td>68/189</td>
</tr>
<tr>
<td>Culturally active</td>
<td>86.07</td>
<td>97.19</td>
<td>346/402</td>
<td>346/356</td>
</tr>
</tbody>
</table>
Table A-2. The 2003 (May) combined classes confusion matrix

<table>
<thead>
<tr>
<th>Class</th>
<th>Water body</th>
<th>Forested</th>
<th>Savannah</th>
<th>Inlet areas</th>
<th>Culturally active</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unclassified</td>
<td>89</td>
<td>0</td>
<td>1</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Water body</td>
<td>10526</td>
<td>10</td>
<td>0</td>
<td>15</td>
<td>0</td>
</tr>
<tr>
<td>Forested</td>
<td>0</td>
<td>335</td>
<td>0</td>
<td>106</td>
<td>12</td>
</tr>
<tr>
<td>Savannah</td>
<td>0</td>
<td>0</td>
<td>225</td>
<td>9</td>
<td>13</td>
</tr>
<tr>
<td>Inlet areas</td>
<td>0</td>
<td>10</td>
<td>0</td>
<td>117</td>
<td>0</td>
</tr>
<tr>
<td>Culturally active</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>2</td>
<td>328</td>
</tr>
<tr>
<td>Transitional</td>
<td>0</td>
<td>3</td>
<td>62</td>
<td>12</td>
<td>14</td>
</tr>
<tr>
<td>Bare soil</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Anthropogenic vegetation</td>
<td>0</td>
<td>136</td>
<td>0</td>
<td>44</td>
<td>0</td>
</tr>
<tr>
<td>Pasture outside park</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Aldea</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Actively cultivated site</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>10615</td>
<td>494</td>
<td>292</td>
<td>305</td>
<td>402</td>
</tr>
</tbody>
</table>

Ground Truth (Pixels)

<table>
<thead>
<tr>
<th>Class</th>
<th>Transition</th>
<th>Bare soil</th>
<th>Anthropogenic vegetation</th>
<th>Pasture outside park</th>
<th>Aldea</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unclassified</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>80</td>
<td>0</td>
</tr>
<tr>
<td>Water body</td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Forested</td>
<td>42</td>
<td>0</td>
<td>22</td>
<td>13</td>
<td>0</td>
</tr>
<tr>
<td>Savannah</td>
<td>6</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Inlet areas</td>
<td>8</td>
<td>0</td>
<td>2</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Culturally active</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Transitional</td>
<td>48</td>
<td>0</td>
<td>0</td>
<td>3</td>
<td>0</td>
</tr>
<tr>
<td>Bare soil</td>
<td>0</td>
<td>26</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Anthropogenic vegetation</td>
<td>7</td>
<td>0</td>
<td>88</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Pasture outside park</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>774</td>
<td>1</td>
</tr>
<tr>
<td>Aldea</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>112</td>
<td>30</td>
</tr>
<tr>
<td>Actively cultivated site</td>
<td>3</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>Total</td>
<td>115</td>
<td>26</td>
<td>116</td>
<td>982</td>
<td>31</td>
</tr>
<tr>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
<td>---</td>
</tr>
</tbody>
</table>

Ground Truth (Pixels)

<table>
<thead>
<tr>
<th>Class</th>
<th>Actively cultivated site</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unclassified</td>
<td>1</td>
<td>171</td>
</tr>
<tr>
<td>Water body</td>
<td>0</td>
<td>10555</td>
</tr>
<tr>
<td>Forested</td>
<td>63</td>
<td>593</td>
</tr>
<tr>
<td>Savannah</td>
<td>3</td>
<td>256</td>
</tr>
<tr>
<td>Inlet areas</td>
<td>35</td>
<td>172</td>
</tr>
<tr>
<td>Culturally active</td>
<td>0</td>
<td>334</td>
</tr>
<tr>
<td>Transitional</td>
<td>14</td>
<td>156</td>
</tr>
<tr>
<td>Bare soil</td>
<td>0</td>
<td>61</td>
</tr>
<tr>
<td>Anthropogenic vegetation</td>
<td>17</td>
<td>292</td>
</tr>
<tr>
<td>Pasture outside park</td>
<td>0</td>
<td>776</td>
</tr>
<tr>
<td>Aldea</td>
<td>0</td>
<td>142</td>
</tr>
<tr>
<td>Actively cultivated site</td>
<td>9</td>
<td>12</td>
</tr>
<tr>
<td>Total</td>
<td>142</td>
<td>13520</td>
</tr>
</tbody>
</table>

Ground Truth (Percent)

<table>
<thead>
<tr>
<th>Class</th>
<th>Water body</th>
<th>Forested</th>
<th>Savannah</th>
<th>Inlet areas</th>
<th>Culturally active</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unclassified</td>
<td>0.84</td>
<td>0.00</td>
<td>0.34</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Water body</td>
<td>99.16</td>
<td>2.02</td>
<td>0.00</td>
<td>4.92</td>
<td>0.00</td>
</tr>
<tr>
<td>Forested</td>
<td>0.00</td>
<td>67.81</td>
<td>0.00</td>
<td>34.75</td>
<td>2.99</td>
</tr>
<tr>
<td>Savannah</td>
<td>0.00</td>
<td>0.00</td>
<td>77.05</td>
<td>2.95</td>
<td>3.23</td>
</tr>
<tr>
<td>Inlet areas</td>
<td>0.00</td>
<td>2.02</td>
<td>0.00</td>
<td>38.36</td>
<td>0.00</td>
</tr>
<tr>
<td>Culturally active</td>
<td>0.00</td>
<td>0.00</td>
<td>1.37</td>
<td>0.66</td>
<td>81.59</td>
</tr>
<tr>
<td>Transitional</td>
<td>0.00</td>
<td>0.61</td>
<td>21.23</td>
<td>3.93</td>
<td>3.48</td>
</tr>
<tr>
<td>Bare soil</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>8.71</td>
</tr>
<tr>
<td>Anthropogenic vegetation</td>
<td>0.00</td>
<td>27.53</td>
<td>0.00</td>
<td>14.43</td>
<td>0.00</td>
</tr>
<tr>
<td>Pasture outside park</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>8.15</td>
</tr>
<tr>
<td>Aldea</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Actively cultivated site</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Total</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
<td>100.00</td>
</tr>
</tbody>
</table>

Ground Truth (Percent)

<table>
<thead>
<tr>
<th>Class</th>
<th>Transitional</th>
<th>Bare soil</th>
<th>Anthropogenic vegetation</th>
<th>Pasture outside park</th>
<th>Aldea</th>
</tr>
</thead>
<tbody>
<tr>
<td>Unclassified</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>8.15</td>
<td>0.00</td>
</tr>
<tr>
<td>Water body</td>
<td>0.00</td>
<td>0.00</td>
<td>3.45</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Forested type</td>
<td>36.52</td>
<td>0.00</td>
<td>18.97</td>
<td>1.32</td>
<td>0.00</td>
</tr>
<tr>
<td>Savannah</td>
<td>5.22</td>
<td>0.00</td>
<td>1.72</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Inlet areas</td>
<td>6.96</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Culturally active</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Class</td>
<td>Actively cultivated site</td>
<td>Total</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>----------------------------</td>
<td>--------------------------</td>
<td>-------</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Unclassified</td>
<td>0.70</td>
<td>1.26</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Water body</td>
<td>0.00</td>
<td>78.07</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Forested type</td>
<td>44.37</td>
<td>4.39</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Savannah</td>
<td>2.11</td>
<td>1.89</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inlet areas</td>
<td>24.65</td>
<td>1.27</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Culturally active</td>
<td>0.00</td>
<td>2.47</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transitional</td>
<td>9.86</td>
<td>1.15</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bare soil</td>
<td>0.00</td>
<td>0.45</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anthropogenic vegetation</td>
<td>11.97</td>
<td>2.16</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pasture outside park</td>
<td>0.00</td>
<td>5.74</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aldea</td>
<td>0.00</td>
<td>1.05</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Actively cultivated site</td>
<td>6.34</td>
<td>0.09</td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Total</td>
<td>100.00</td>
<td>100.00</td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Class</th>
<th>Commission (Percent)</th>
<th>Omission (Percent)</th>
<th>Commission (Pixels)</th>
<th>Omission (Pixels)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water body</td>
<td>0.27</td>
<td>0.84</td>
<td>29/10555</td>
<td>89/10615</td>
</tr>
<tr>
<td>Forested type</td>
<td>43.51</td>
<td>32.19</td>
<td>258/593</td>
<td>159/494</td>
</tr>
<tr>
<td>Savannah</td>
<td>12.11</td>
<td>22.95</td>
<td>31/256</td>
<td>67/292</td>
</tr>
<tr>
<td>Inlet areas</td>
<td>31.98</td>
<td>61.64</td>
<td>55/172</td>
<td>188/305</td>
</tr>
<tr>
<td>Culturally active</td>
<td>1.80</td>
<td>18.41</td>
<td>6/334</td>
<td>74/402</td>
</tr>
<tr>
<td>Transitional</td>
<td>69.23</td>
<td>58.26</td>
<td>108/156</td>
<td>67/115</td>
</tr>
<tr>
<td>Bare soil</td>
<td>57.38</td>
<td>0.00</td>
<td>35/61</td>
<td>0/26</td>
</tr>
<tr>
<td>Anthropogenic vegetation</td>
<td>69.86</td>
<td>24.14</td>
<td>204/292</td>
<td>28/116</td>
</tr>
<tr>
<td>Pasture outside park</td>
<td>0.26</td>
<td>21.18</td>
<td>2/776</td>
<td>208/982</td>
</tr>
<tr>
<td>Aldea</td>
<td>78.87</td>
<td>3.23</td>
<td>112/142</td>
<td>1/31</td>
</tr>
<tr>
<td>Actively cultivated site</td>
<td>25.00</td>
<td>93.66</td>
<td>3/12</td>
<td>133/142</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Water body</td>
<td>99.16</td>
<td>99.73</td>
<td>10526/10615</td>
<td>10526/10555</td>
</tr>
<tr>
<td>Forested type</td>
<td>67.81</td>
<td>56.49</td>
<td>335/494</td>
<td>335/593</td>
</tr>
</tbody>
</table>
Table A-3. The spectral profiles for the 2002 (August) combined classification

<table>
<thead>
<tr>
<th>Area</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Stdev</th>
<th>Num</th>
<th>Eigenvalue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Savannah</td>
<td>77.05</td>
<td>87.89</td>
<td>225/292</td>
<td>225/256</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Inlet areas</td>
<td>38.36</td>
<td>68.02</td>
<td>117/305</td>
<td>117/172</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Culturally active</td>
<td>81.59</td>
<td>98.20</td>
<td>328/402</td>
<td>328/334</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Transitional</td>
<td>41.74</td>
<td>30.77</td>
<td>48/115</td>
<td>48/156</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Bare soil</td>
<td>100.00</td>
<td>42.62</td>
<td>26/26</td>
<td>26/61</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Anthropogenic vegetation</td>
<td>75.86</td>
<td>30.14</td>
<td>88/116</td>
<td>88/292</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Pasture outside park</td>
<td>78.82</td>
<td>99.74</td>
<td>774/982</td>
<td>774/776</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Aldea</td>
<td>96.77</td>
<td>21.13</td>
<td>30/31</td>
<td>30/142</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Actively cultivated site</td>
<td>6.37</td>
<td>75.00</td>
<td>9/142</td>
<td>9/12</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Area</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Stdev</th>
<th>Num</th>
<th>Eigenvalue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bare Soils</td>
<td>-0.064</td>
<td>0.024</td>
<td>-0.032</td>
<td>0.023</td>
<td>1.000</td>
<td>0.082</td>
</tr>
<tr>
<td>Band1</td>
<td>-0.210</td>
<td>-0.014</td>
<td>-0.122</td>
<td>0.048</td>
<td>2.000</td>
<td>0.008</td>
</tr>
<tr>
<td>Band2</td>
<td>-0.107</td>
<td>-0.059</td>
<td>-0.077</td>
<td>0.012</td>
<td>3.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band3</td>
<td>-0.160</td>
<td>0.152</td>
<td>-0.057</td>
<td>0.082</td>
<td>4.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band4</td>
<td>0.030</td>
<td>1.189</td>
<td>0.454</td>
<td>0.283</td>
<td>5.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band5</td>
<td>-0.100</td>
<td>-0.060</td>
<td>-0.085</td>
<td>0.010</td>
<td>6.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band7</td>
<td>-0.007</td>
<td>0.010</td>
<td>0.001</td>
<td>0.003</td>
<td>7.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>
### Culturally active

<table>
<thead>
<tr>
<th>Band</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Stdev</th>
<th>Num</th>
<th>Eigenvalue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Band1</td>
<td>-0.083</td>
<td>0.280</td>
<td>0.037</td>
<td>0.078</td>
<td>1.000</td>
<td>0.250</td>
</tr>
<tr>
<td>Band2</td>
<td>-0.249</td>
<td>0.334</td>
<td>-0.021</td>
<td>0.132</td>
<td>2.000</td>
<td>0.048</td>
</tr>
<tr>
<td>Band3</td>
<td>-0.198</td>
<td>-0.048</td>
<td>-0.092</td>
<td>0.015</td>
<td>3.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band4</td>
<td>-0.172</td>
<td>1.314</td>
<td>0.293</td>
<td>0.307</td>
<td>4.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band5</td>
<td>-0.627</td>
<td>1.859</td>
<td>0.642</td>
<td>0.424</td>
<td>5.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band6</td>
<td>-0.111</td>
<td>0.061</td>
<td>-0.051</td>
<td>0.036</td>
<td>6.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band7</td>
<td>-0.025</td>
<td>0.030</td>
<td>0.003</td>
<td>0.007</td>
<td>7.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

### Non-anthropogenic forest

<table>
<thead>
<tr>
<th>Band</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Stdev</th>
<th>Num</th>
<th>Eigenvalue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Band1</td>
<td>0.109</td>
<td>0.271</td>
<td>0.192</td>
<td>0.024</td>
<td>1.000</td>
<td>0.043</td>
</tr>
<tr>
<td>Band2</td>
<td>0.215</td>
<td>0.439</td>
<td>0.327</td>
<td>0.030</td>
<td>2.000</td>
<td>0.009</td>
</tr>
<tr>
<td>Band3</td>
<td>-0.074</td>
<td>0.018</td>
<td>-0.024</td>
<td>0.013</td>
<td>3.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band4</td>
<td>0.529</td>
<td>1.143</td>
<td>0.827</td>
<td>0.096</td>
<td>4.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band5</td>
<td>-0.881</td>
<td>0.652</td>
<td>-0.090</td>
<td>0.205</td>
<td>5.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band6</td>
<td>0.018</td>
<td>0.062</td>
<td>0.041</td>
<td>0.008</td>
<td>6.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band7</td>
<td>-0.008</td>
<td>0.017</td>
<td>0.004</td>
<td>0.004</td>
<td>7.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>
### Inlet Areas

<table>
<thead>
<tr>
<th>Basic Stats</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Stdev</th>
<th>Num</th>
<th>Eigenvalue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Band1</td>
<td>0.057</td>
<td>0.314</td>
<td>0.199</td>
<td>0.056</td>
<td>1.000</td>
<td>0.105</td>
</tr>
<tr>
<td>Band2</td>
<td>0.072</td>
<td>0.444</td>
<td>0.311</td>
<td>0.075</td>
<td>2.000</td>
<td>0.031</td>
</tr>
<tr>
<td>Band3</td>
<td>-0.086</td>
<td>0.003</td>
<td>-0.035</td>
<td>0.015</td>
<td>3.000</td>
<td>0.001</td>
</tr>
<tr>
<td>Band4</td>
<td>0.285</td>
<td>1.383</td>
<td>0.907</td>
<td>0.228</td>
<td>4.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band5</td>
<td>-0.641</td>
<td>0.907</td>
<td>0.099</td>
<td>0.275</td>
<td>5.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band6</td>
<td>-0.031</td>
<td>0.087</td>
<td>0.046</td>
<td>0.022</td>
<td>6.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band7</td>
<td>-0.012</td>
<td>0.023</td>
<td>0.006</td>
<td>0.007</td>
<td>7.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

#### Eigenvector

<table>
<thead>
<tr>
<th>Band1</th>
<th>Band2</th>
<th>Band3</th>
<th>Band4</th>
<th>Band5</th>
<th>Band6</th>
<th>Band7</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.157</td>
<td>-0.200</td>
<td>0.010</td>
<td>-0.574</td>
<td>0.777</td>
<td>-0.045</td>
<td>0.020</td>
</tr>
<tr>
<td>-0.138</td>
<td>-0.166</td>
<td>0.028</td>
<td>-0.744</td>
<td>-0.625</td>
<td>-0.087</td>
<td>-0.004</td>
</tr>
<tr>
<td>0.037</td>
<td>0.816</td>
<td>0.477</td>
<td>-0.234</td>
<td>0.046</td>
<td>0.197</td>
<td>0.098</td>
</tr>
<tr>
<td>0.000</td>
<td>0.455</td>
<td>-0.837</td>
<td>-0.158</td>
<td>0.022</td>
<td>0.074</td>
<td>-0.248</td>
</tr>
<tr>
<td>0.013</td>
<td>0.226</td>
<td>-0.044</td>
<td>0.032</td>
<td>0.025</td>
<td>-0.935</td>
<td>0.267</td>
</tr>
<tr>
<td>0.056</td>
<td>0.021</td>
<td>0.261</td>
<td>0.008</td>
<td>0.027</td>
<td>-0.269</td>
<td>-0.925</td>
</tr>
<tr>
<td>-0.976</td>
<td>0.091</td>
<td>0.027</td>
<td>0.189</td>
<td>-0.033</td>
<td>-0.001</td>
<td>-0.049</td>
</tr>
</tbody>
</table>

### Water Bodies

<table>
<thead>
<tr>
<th>Basic Stats</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Stdev</th>
<th>Num</th>
<th>Eigenvalue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Band1</td>
<td>-0.479</td>
<td>-0.089</td>
<td>-0.420</td>
<td>0.019</td>
<td>1.000</td>
<td>0.291</td>
</tr>
<tr>
<td>Band2</td>
<td>-1.581</td>
<td>-0.196</td>
<td>-1.336</td>
<td>0.072</td>
<td>2.000</td>
<td>0.005</td>
</tr>
<tr>
<td>Band3</td>
<td>-0.030</td>
<td>0.083</td>
<td>0.038</td>
<td>0.005</td>
<td>3.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band4</td>
<td>-0.919</td>
<td>0.015</td>
<td>-0.849</td>
<td>0.026</td>
<td>4.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band5</td>
<td>2.238</td>
<td>6.126</td>
<td>4.144</td>
<td>0.537</td>
<td>5.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band6</td>
<td>-0.027</td>
<td>-0.022</td>
<td>-0.024</td>
<td>0.001</td>
<td>6.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band7</td>
<td>0.057</td>
<td>0.161</td>
<td>0.138</td>
<td>0.012</td>
<td>7.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

#### Eigenvector

<table>
<thead>
<tr>
<th>Band1</th>
<th>Band2</th>
<th>Band3</th>
<th>Band4</th>
<th>Band5</th>
<th>Band6</th>
<th>Band7</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.031</td>
<td>-0.067</td>
<td>0.003</td>
<td>0.006</td>
<td>0.997</td>
<td>0.000</td>
<td>0.019</td>
</tr>
<tr>
<td>0.118</td>
<td>0.917</td>
<td>-0.069</td>
<td>0.367</td>
<td>0.064</td>
<td>0.009</td>
<td>-0.024</td>
</tr>
<tr>
<td>0.202</td>
<td>-0.379</td>
<td>0.043</td>
<td>0.898</td>
<td>-0.027</td>
<td>0.029</td>
<td>0.079</td>
</tr>
<tr>
<td>0.039</td>
<td>-0.072</td>
<td>-0.242</td>
<td>0.058</td>
<td>0.015</td>
<td>-0.024</td>
<td>-0.965</td>
</tr>
<tr>
<td>-0.058</td>
<td>-0.062</td>
<td>-0.964</td>
<td>0.010</td>
<td>-0.008</td>
<td>0.069</td>
<td>0.243</td>
</tr>
<tr>
<td>-0.266</td>
<td>0.015</td>
<td>0.072</td>
<td>0.036</td>
<td>-0.007</td>
<td>0.959</td>
<td>-0.052</td>
</tr>
<tr>
<td>Band7</td>
<td>0.932</td>
<td>-0.033</td>
<td>-0.031</td>
<td>-0.233</td>
<td>0.028</td>
<td>0.271</td>
</tr>
<tr>
<td>-------</td>
<td>-------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>-------</td>
<td>-------</td>
</tr>
</tbody>
</table>

Pasture outside park

<table>
<thead>
<tr>
<th>Basic Stats</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Stdev</th>
<th>Num</th>
<th>Eigenvalue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Band1</td>
<td>-0.188</td>
<td>-0.056</td>
<td>-0.122</td>
<td>0.016</td>
<td>1.000</td>
<td>0.194</td>
</tr>
<tr>
<td>Band2</td>
<td>-0.511</td>
<td>-0.255</td>
<td>-0.376</td>
<td>0.030</td>
<td>2.000</td>
<td>0.008</td>
</tr>
<tr>
<td>Band3</td>
<td>-0.274</td>
<td>-0.118</td>
<td>-0.169</td>
<td>0.023</td>
<td>3.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band4</td>
<td>-0.589</td>
<td>0.092</td>
<td>-0.229</td>
<td>0.101</td>
<td>4.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band5</td>
<td>0.690</td>
<td>4.610</td>
<td>2.172</td>
<td>0.436</td>
<td>5.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band6</td>
<td>-0.275</td>
<td>-0.108</td>
<td>-0.157</td>
<td>0.020</td>
<td>6.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band7</td>
<td>-0.027</td>
<td>0.026</td>
<td>0.003</td>
<td>0.006</td>
<td>7.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Eigenvector</th>
<th>Band1</th>
<th>Band2</th>
<th>Band3</th>
<th>Band4</th>
<th>Band5</th>
<th>Band6</th>
<th>Band7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Band1</td>
<td>-0.001</td>
<td>-0.019</td>
<td>-0.038</td>
<td>0.141</td>
<td>0.989</td>
<td>0.003</td>
<td>0.007</td>
</tr>
<tr>
<td>Band2</td>
<td>-0.183</td>
<td>-0.323</td>
<td>0.028</td>
<td>-0.909</td>
<td>0.125</td>
<td>-0.142</td>
<td>-0.002</td>
</tr>
<tr>
<td>Band3</td>
<td>-0.116</td>
<td>0.118</td>
<td>0.672</td>
<td>-0.100</td>
<td>0.039</td>
<td>0.684</td>
<td>0.203</td>
</tr>
<tr>
<td>Band4</td>
<td>-0.094</td>
<td>0.036</td>
<td>-0.660</td>
<td>-0.122</td>
<td>-0.007</td>
<td>0.686</td>
<td>-0.262</td>
</tr>
<tr>
<td>Band5</td>
<td>-0.008</td>
<td>-0.933</td>
<td>0.022</td>
<td>0.301</td>
<td>-0.061</td>
<td>0.158</td>
<td>0.094</td>
</tr>
<tr>
<td>Band6</td>
<td>0.060</td>
<td>-0.081</td>
<td>0.333</td>
<td>0.034</td>
<td>0.013</td>
<td>-0.020</td>
<td>-0.937</td>
</tr>
<tr>
<td>Band7</td>
<td>0.970</td>
<td>-0.046</td>
<td>0.002</td>
<td>-0.194</td>
<td>0.027</td>
<td>0.124</td>
<td>0.057</td>
</tr>
</tbody>
</table>

Savannah

<table>
<thead>
<tr>
<th>Basic Stats</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Stdev</th>
<th>Num</th>
<th>Eigenvalue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Band1</td>
<td>-0.288</td>
<td>-0.119</td>
<td>-0.211</td>
<td>0.028</td>
<td>1.000</td>
<td>0.106</td>
</tr>
<tr>
<td>Band2</td>
<td>-0.665</td>
<td>-0.164</td>
<td>-0.424</td>
<td>0.080</td>
<td>2.000</td>
<td>0.011</td>
</tr>
<tr>
<td>Band3</td>
<td>-0.022</td>
<td>0.044</td>
<td>0.012</td>
<td>0.013</td>
<td>3.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band4</td>
<td>-0.801</td>
<td>-0.300</td>
<td>-0.575</td>
<td>0.085</td>
<td>4.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band5</td>
<td>0.767</td>
<td>2.576</td>
<td>1.514</td>
<td>0.321</td>
<td>5.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band6</td>
<td>-0.113</td>
<td>-0.047</td>
<td>-0.082</td>
<td>0.013</td>
<td>6.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band7</td>
<td>0.035</td>
<td>0.072</td>
<td>0.054</td>
<td>0.007</td>
<td>7.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Eigenvector</th>
<th>Band1</th>
<th>Band2</th>
<th>Band3</th>
<th>Band4</th>
<th>Band5</th>
<th>Band6</th>
<th>Band7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Band1</td>
<td>-0.051</td>
<td>-0.167</td>
<td>-0.014</td>
<td>-0.037</td>
<td>0.984</td>
<td>0.006</td>
<td>0.017</td>
</tr>
<tr>
<td>Band2</td>
<td>0.213</td>
<td>0.551</td>
<td>-0.039</td>
<td>0.793</td>
<td>0.134</td>
<td>0.047</td>
<td>-0.015</td>
</tr>
<tr>
<td>Band3</td>
<td>-0.244</td>
<td>0.338</td>
<td>0.633</td>
<td>-0.175</td>
<td>0.040</td>
<td>0.582</td>
<td>0.232</td>
</tr>
<tr>
<td>Band4</td>
<td>0.033</td>
<td>0.718</td>
<td>-0.078</td>
<td>-0.504</td>
<td>0.107</td>
<td>-0.459</td>
<td>-0.044</td>
</tr>
<tr>
<td>Band5</td>
<td>0.037</td>
<td>-0.172</td>
<td>0.530</td>
<td>0.184</td>
<td>-0.018</td>
<td>-0.614</td>
<td>0.527</td>
</tr>
<tr>
<td>Band6</td>
<td>-0.113</td>
<td>-0.062</td>
<td>0.514</td>
<td>0.095</td>
<td>0.010</td>
<td>-0.217</td>
<td>-0.814</td>
</tr>
<tr>
<td>Band7</td>
<td>0.937</td>
<td>-0.072</td>
<td>0.217</td>
<td>-0.206</td>
<td>0.031</td>
<td>0.156</td>
<td>-0.053</td>
</tr>
</tbody>
</table>
### Actively cultivated site

<table>
<thead>
<tr>
<th>Basic Stats</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Stdev</th>
<th>Num</th>
<th>Eigenvalue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Band1</td>
<td>0.074</td>
<td>0.239</td>
<td>0.159</td>
<td>0.034</td>
<td>1.000</td>
<td>0.049</td>
</tr>
<tr>
<td>Band2</td>
<td>0.154</td>
<td>0.364</td>
<td>0.257</td>
<td>0.046</td>
<td>2.000</td>
<td>0.020</td>
</tr>
<tr>
<td>Band3</td>
<td>-0.077</td>
<td>-0.006</td>
<td>-0.035</td>
<td>0.013</td>
<td>3.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band4</td>
<td>0.366</td>
<td>1.051</td>
<td>0.716</td>
<td>0.141</td>
<td>4.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band5</td>
<td>-0.498</td>
<td>0.559</td>
<td>0.046</td>
<td>0.215</td>
<td>5.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band6</td>
<td>-0.003</td>
<td>0.057</td>
<td>0.024</td>
<td>0.014</td>
<td>6.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band7</td>
<td>-0.011</td>
<td>0.017</td>
<td>0.006</td>
<td>0.005</td>
<td>7.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Eigenvector</th>
<th>Band1</th>
<th>Band2</th>
<th>Band3</th>
<th>Band4</th>
<th>Band5</th>
<th>Band6</th>
<th>Band7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Band1</td>
<td>-0.096</td>
<td>-0.147</td>
<td>-0.011</td>
<td>-0.254</td>
<td>0.951</td>
<td>-0.014</td>
<td>0.018</td>
</tr>
<tr>
<td>Band2</td>
<td>0.187</td>
<td>0.207</td>
<td>-0.047</td>
<td>0.909</td>
<td>0.294</td>
<td>0.089</td>
<td>-0.007</td>
</tr>
<tr>
<td>Band3</td>
<td>0.022</td>
<td>0.749</td>
<td>0.582</td>
<td>-0.188</td>
<td>0.075</td>
<td>0.210</td>
<td>0.121</td>
</tr>
<tr>
<td>Band4</td>
<td>-0.004</td>
<td>0.539</td>
<td>-0.752</td>
<td>-0.188</td>
<td>0.032</td>
<td>0.150</td>
<td>-0.293</td>
</tr>
<tr>
<td>Band5</td>
<td>-0.057</td>
<td>-0.260</td>
<td>-0.023</td>
<td>-0.011</td>
<td>-0.037</td>
<td>0.960</td>
<td>0.076</td>
</tr>
<tr>
<td>Band6</td>
<td>-0.024</td>
<td>0.098</td>
<td>-0.305</td>
<td>-0.016</td>
<td>-0.014</td>
<td>-0.058</td>
<td>0.945</td>
</tr>
<tr>
<td>Band7</td>
<td>-0.975</td>
<td>0.081</td>
<td>0.017</td>
<td>0.197</td>
<td>-0.033</td>
<td>-0.032</td>
<td>-0.027</td>
</tr>
</tbody>
</table>

### Anthropogenic vegetation

<table>
<thead>
<tr>
<th>Basic Stats</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Stdev</th>
<th>Num</th>
<th>Eigenvalue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Band1</td>
<td>0.120</td>
<td>0.227</td>
<td>0.168</td>
<td>0.022</td>
<td>1.000</td>
<td>0.038</td>
</tr>
<tr>
<td>Band2</td>
<td>0.197</td>
<td>0.363</td>
<td>0.279</td>
<td>0.033</td>
<td>2.000</td>
<td>0.008</td>
</tr>
<tr>
<td>Band3</td>
<td>-0.057</td>
<td>-0.005</td>
<td>-0.028</td>
<td>0.011</td>
<td>3.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band4</td>
<td>0.594</td>
<td>0.979</td>
<td>0.759</td>
<td>0.085</td>
<td>4.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band5</td>
<td>-0.342</td>
<td>0.556</td>
<td>0.082</td>
<td>0.193</td>
<td>5.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band6</td>
<td>0.014</td>
<td>0.052</td>
<td>0.033</td>
<td>0.008</td>
<td>6.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band7</td>
<td>-0.001</td>
<td>0.016</td>
<td>0.007</td>
<td>0.003</td>
<td>7.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Eigenvector</th>
<th>Band1</th>
<th>Band2</th>
<th>Band3</th>
<th>Band4</th>
<th>Band5</th>
<th>Band6</th>
<th>Band7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Band1</td>
<td>-0.064</td>
<td>-0.129</td>
<td>-0.038</td>
<td>-0.091</td>
<td>0.985</td>
<td>-0.004</td>
<td>0.011</td>
</tr>
<tr>
<td>Band2</td>
<td>0.200</td>
<td>0.218</td>
<td>-0.046</td>
<td>0.941</td>
<td>0.127</td>
<td>0.086</td>
<td>-0.010</td>
</tr>
<tr>
<td>Band3</td>
<td>0.025</td>
<td>0.798</td>
<td>0.514</td>
<td>-0.195</td>
<td>0.107</td>
<td>0.182</td>
<td>0.122</td>
</tr>
<tr>
<td>Band4</td>
<td>-0.003</td>
<td>0.477</td>
<td>-0.792</td>
<td>-0.169</td>
<td>0.020</td>
<td>0.156</td>
<td>-0.303</td>
</tr>
<tr>
<td>Band5</td>
<td>0.030</td>
<td>0.254</td>
<td>-0.118</td>
<td>0.014</td>
<td>0.025</td>
<td>-0.933</td>
<td>0.224</td>
</tr>
<tr>
<td>Band6</td>
<td>0.031</td>
<td>0.005</td>
<td>0.301</td>
<td>0.017</td>
<td>0.025</td>
<td>-0.255</td>
<td>-0.918</td>
</tr>
<tr>
<td>Band7</td>
<td>0.976</td>
<td>-0.080</td>
<td>-0.015</td>
<td>-0.196</td>
<td>0.034</td>
<td>0.014</td>
<td>0.021</td>
</tr>
</tbody>
</table>

### Transitional

<table>
<thead>
<tr>
<th>Basic Stats</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Stdev</th>
<th>Num</th>
<th>Eigenvalue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Band1</td>
<td>0.083</td>
<td>0.293</td>
<td>0.209</td>
<td>0.045</td>
<td>1.000</td>
<td>0.062</td>
</tr>
<tr>
<td>Band</td>
<td>Min</td>
<td>Max</td>
<td>Mean</td>
<td>Stdev</td>
<td>Num</td>
<td>Eigenvalue</td>
</tr>
<tr>
<td>-------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>-----</td>
<td>------------</td>
</tr>
<tr>
<td>Band1</td>
<td>-0.145</td>
<td>-0.036</td>
<td>-0.082</td>
<td>0.023</td>
<td>1.00</td>
<td>0.165</td>
</tr>
<tr>
<td>Band2</td>
<td>-0.215</td>
<td>-0.020</td>
<td>-0.124</td>
<td>0.041</td>
<td>2.00</td>
<td>0.020</td>
</tr>
<tr>
<td>Band3</td>
<td>-0.002</td>
<td>0.160</td>
<td>0.084</td>
<td>0.041</td>
<td>3.00</td>
<td>0.001</td>
</tr>
<tr>
<td>Band4</td>
<td>-0.503</td>
<td>0.144</td>
<td>-0.125</td>
<td>0.142</td>
<td>4.00</td>
<td>0.000</td>
</tr>
<tr>
<td>Band5</td>
<td>1.298</td>
<td>3.029</td>
<td>2.209</td>
<td>0.402</td>
<td>5.00</td>
<td>0.000</td>
</tr>
<tr>
<td>Band6</td>
<td>-0.077</td>
<td>0.017</td>
<td>-0.019</td>
<td>0.021</td>
<td>6.00</td>
<td>0.000</td>
</tr>
<tr>
<td>Band7</td>
<td>0.020</td>
<td>0.061</td>
<td>0.047</td>
<td>0.009</td>
<td>7.00</td>
<td>0.000</td>
</tr>
</tbody>
</table>

Table A-4. The spectral profiles for the 2003 (May) combined classification

**Aldea**

<table>
<thead>
<tr>
<th>Basic Stats</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Stdev</th>
<th>Num</th>
<th>Eigenvalue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Band1</td>
<td>-0.131</td>
<td>-0.186</td>
<td>-0.066</td>
<td>-0.436</td>
<td>0.870</td>
<td>-0.037</td>
</tr>
<tr>
<td>Band2</td>
<td>0.160</td>
<td>0.175</td>
<td>-0.046</td>
<td>0.836</td>
<td>0.484</td>
<td>0.090</td>
</tr>
<tr>
<td>Band3</td>
<td>0.037</td>
<td>0.808</td>
<td>0.483</td>
<td>-0.219</td>
<td>0.080</td>
<td>0.218</td>
</tr>
<tr>
<td>Band4</td>
<td>-0.033</td>
<td>0.435</td>
<td>-0.833</td>
<td>-0.159</td>
<td>0.015</td>
<td>0.186</td>
</tr>
<tr>
<td>Band5</td>
<td>0.062</td>
<td>0.291</td>
<td>-0.079</td>
<td>0.006</td>
<td>0.032</td>
<td>-0.949</td>
</tr>
<tr>
<td>Band6</td>
<td>0.089</td>
<td>-0.009</td>
<td>0.248</td>
<td>-0.009</td>
<td>0.025</td>
<td>-0.079</td>
</tr>
<tr>
<td>Band7</td>
<td>0.971</td>
<td>-0.088</td>
<td>-0.058</td>
<td>-0.193</td>
<td>0.031</td>
<td>0.046</td>
</tr>
</tbody>
</table>

**Bare soils**

<table>
<thead>
<tr>
<th>Basic Stats</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Stdev</th>
<th>Num</th>
<th>Eigenvalue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Band1</td>
<td>-0.123</td>
<td>-0.015</td>
<td>-0.019</td>
<td>0.029</td>
<td>1.00</td>
<td>0.043</td>
</tr>
<tr>
<td>Band2</td>
<td>-0.203</td>
<td>0.008</td>
<td>-0.035</td>
<td>0.044</td>
<td>2.00</td>
<td>0.010</td>
</tr>
<tr>
<td>Band3</td>
<td>0.028</td>
<td>0.118</td>
<td>0.050</td>
<td>0.019</td>
<td>3.00</td>
<td>0.000</td>
</tr>
<tr>
<td>Band4</td>
<td>-0.586</td>
<td>0.039</td>
<td>-0.133</td>
<td>0.133</td>
<td>4.000</td>
<td>0.000</td>
</tr>
<tr>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
<td>--------</td>
</tr>
<tr>
<td>Band5</td>
<td>0.841</td>
<td>1.668</td>
<td>1.106</td>
<td>0.178</td>
<td>5.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band6</td>
<td>-0.065</td>
<td>-0.030</td>
<td>-0.038</td>
<td>0.007</td>
<td>6.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band7</td>
<td>0.016</td>
<td>0.037</td>
<td>0.024</td>
<td>0.005</td>
<td>7.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Eigenvector</th>
<th>Band1</th>
<th>Band2</th>
<th>Band3</th>
<th>Band4</th>
<th>Band5</th>
<th>Band6</th>
<th>Band7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Band1</td>
<td>-0.125</td>
<td>-0.202</td>
<td>0.085</td>
<td>-0.519</td>
<td>0.816</td>
<td>-0.028</td>
<td>0.018</td>
</tr>
<tr>
<td>Band2</td>
<td>0.131</td>
<td>0.150</td>
<td>0.019</td>
<td>0.800</td>
<td>0.565</td>
<td>0.035</td>
<td>-0.015</td>
</tr>
<tr>
<td>Band3</td>
<td>-0.254</td>
<td>0.040</td>
<td>-0.918</td>
<td>0.003</td>
<td>0.070</td>
<td>0.187</td>
<td>0.226</td>
</tr>
<tr>
<td>Band4</td>
<td>0.022</td>
<td>0.767</td>
<td>0.024</td>
<td>-0.225</td>
<td>0.071</td>
<td>0.442</td>
<td>-0.400</td>
</tr>
<tr>
<td>Band5</td>
<td>0.447</td>
<td>0.371</td>
<td>0.083</td>
<td>-0.153</td>
<td>0.035</td>
<td>-0.037</td>
<td>0.793</td>
</tr>
<tr>
<td>Band6</td>
<td>0.365</td>
<td>0.227</td>
<td>-0.325</td>
<td>-0.109</td>
<td>0.058</td>
<td>-0.762</td>
<td>-0.337</td>
</tr>
<tr>
<td>Band7</td>
<td>-0.754</td>
<td>0.398</td>
<td>0.191</td>
<td>0.074</td>
<td>-0.009</td>
<td>-0.430</td>
<td>0.214</td>
</tr>
</tbody>
</table>

**Culturally active**

<table>
<thead>
<tr>
<th>Basic Stats</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Stdev</th>
<th>Num</th>
<th>Eigenvalue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Band1</td>
<td>-0.113</td>
<td>0.150</td>
<td>0.025</td>
<td>0.066</td>
<td>1.000</td>
<td>0.217</td>
</tr>
<tr>
<td>Band2</td>
<td>-0.186</td>
<td>0.162</td>
<td>0.012</td>
<td>0.082</td>
<td>2.000</td>
<td>0.055</td>
</tr>
<tr>
<td>Band3</td>
<td>-0.048</td>
<td>0.141</td>
<td>0.047</td>
<td>0.031</td>
<td>3.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band4</td>
<td>-0.596</td>
<td>0.807</td>
<td>0.141</td>
<td>0.304</td>
<td>4.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band5</td>
<td>0.014</td>
<td>2.740</td>
<td>1.177</td>
<td>0.409</td>
<td>5.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band6</td>
<td>-0.068</td>
<td>0.007</td>
<td>-0.024</td>
<td>0.015</td>
<td>6.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band7</td>
<td>0.003</td>
<td>0.056</td>
<td>0.022</td>
<td>0.010</td>
<td>7.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Eigenvector</th>
<th>Band1</th>
<th>Band2</th>
<th>Band3</th>
<th>Band4</th>
<th>Band5</th>
<th>Band6</th>
<th>Band7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Band1</td>
<td>-0.128</td>
<td>-0.164</td>
<td>0.056</td>
<td>-0.507</td>
<td>0.835</td>
<td>-0.016</td>
<td>0.019</td>
</tr>
<tr>
<td>Band2</td>
<td>0.116</td>
<td>0.129</td>
<td>0.001</td>
<td>0.821</td>
<td>0.542</td>
<td>0.047</td>
<td>-0.004</td>
</tr>
<tr>
<td>Band3</td>
<td>-0.381</td>
<td>0.034</td>
<td>-0.856</td>
<td>0.020</td>
<td>0.020</td>
<td>0.303</td>
<td>0.172</td>
</tr>
<tr>
<td>Band4</td>
<td>0.280</td>
<td>0.905</td>
<td>-0.141</td>
<td>-0.236</td>
<td>0.088</td>
<td>-0.071</td>
<td>-0.120</td>
</tr>
<tr>
<td>Band5</td>
<td>0.543</td>
<td>-0.251</td>
<td>-0.413</td>
<td>-0.020</td>
<td>0.030</td>
<td>-0.621</td>
<td>0.289</td>
</tr>
<tr>
<td>Band6</td>
<td>-0.365</td>
<td>-0.033</td>
<td>-0.174</td>
<td>0.081</td>
<td>0.005</td>
<td>-0.536</td>
<td>-0.736</td>
</tr>
<tr>
<td>Band7</td>
<td>-0.564</td>
<td>0.269</td>
<td>0.209</td>
<td>0.081</td>
<td>-0.021</td>
<td>-0.477</td>
<td>0.575</td>
</tr>
</tbody>
</table>

**Non-anthropogenic forest**

<table>
<thead>
<tr>
<th>Basic Stats</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Stdev</th>
<th>Num</th>
<th>Eigenvalue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Band1</td>
<td>0.140</td>
<td>0.270</td>
<td>0.190</td>
<td>0.020</td>
<td>1.000</td>
<td>0.043</td>
</tr>
<tr>
<td>Band2</td>
<td>0.203</td>
<td>0.354</td>
<td>0.271</td>
<td>0.021</td>
<td>2.000</td>
<td>0.010</td>
</tr>
<tr>
<td>Band3</td>
<td>-0.168</td>
<td>-0.063</td>
<td>-0.109</td>
<td>0.018</td>
<td>3.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band4</td>
<td>0.640</td>
<td>1.264</td>
<td>0.912</td>
<td>0.102</td>
<td>4.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band5</td>
<td>-0.843</td>
<td>0.299</td>
<td>-0.241</td>
<td>0.203</td>
<td>5.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band6</td>
<td>0.006</td>
<td>0.044</td>
<td>0.029</td>
<td>0.006</td>
<td>6.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band7</td>
<td>-0.015</td>
<td>0.028</td>
<td>-0.001</td>
<td>0.004</td>
<td>7.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>
### Inlet areas

<table>
<thead>
<tr>
<th>Eigenvector</th>
<th>Band1</th>
<th>Band2</th>
<th>Band3</th>
<th>Band4</th>
<th>Band5</th>
<th>Band6</th>
<th>Band7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Band1</td>
<td>0.066</td>
<td>-0.081</td>
<td>0.066</td>
<td>-0.146</td>
<td>0.981</td>
<td>0.000</td>
<td>0.012</td>
</tr>
<tr>
<td>Band2</td>
<td>0.145</td>
<td>0.114</td>
<td>0.017</td>
<td>0.968</td>
<td>0.162</td>
<td>0.050</td>
<td>-0.011</td>
</tr>
<tr>
<td>Band3</td>
<td>-0.162</td>
<td>0.244</td>
<td>-0.925</td>
<td>-0.006</td>
<td>0.069</td>
<td>0.170</td>
<td>0.157</td>
</tr>
<tr>
<td>Band4</td>
<td>-0.155</td>
<td>0.820</td>
<td>0.261</td>
<td>-0.104</td>
<td>0.028</td>
<td>0.362</td>
<td>-0.306</td>
</tr>
<tr>
<td>Band5</td>
<td>0.032</td>
<td>-0.109</td>
<td>0.206</td>
<td>-0.018</td>
<td>-0.032</td>
<td>0.688</td>
<td>0.685</td>
</tr>
<tr>
<td>Band6</td>
<td>0.138</td>
<td>-0.435</td>
<td>-0.138</td>
<td>-0.004</td>
<td>-0.010</td>
<td>0.603</td>
<td>-0.640</td>
</tr>
<tr>
<td>Band7</td>
<td>0.951</td>
<td>0.219</td>
<td>-0.100</td>
<td>-0.175</td>
<td>0.062</td>
<td>-0.030</td>
<td>0.050</td>
</tr>
</tbody>
</table>

### Basic Stats

<table>
<thead>
<tr>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Stdev</th>
<th>Num</th>
<th>Eigenvalue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Band1</td>
<td>0.147</td>
<td>0.374</td>
<td>0.219</td>
<td>0.033</td>
<td>1.000</td>
</tr>
<tr>
<td>Band2</td>
<td>0.219</td>
<td>0.414</td>
<td>0.298</td>
<td>0.030</td>
<td>2.000</td>
</tr>
<tr>
<td>Band3</td>
<td>-0.171</td>
<td>-0.049</td>
<td>-0.113</td>
<td>0.020</td>
<td>3.000</td>
</tr>
<tr>
<td>Band4</td>
<td>0.674</td>
<td>1.901</td>
<td>1.095</td>
<td>0.176</td>
<td>4.000</td>
</tr>
<tr>
<td>Band5</td>
<td>-1.082</td>
<td>0.429</td>
<td>-0.287</td>
<td>0.245</td>
<td>5.000</td>
</tr>
<tr>
<td>Band6</td>
<td>0.015</td>
<td>0.075</td>
<td>0.039</td>
<td>0.009</td>
<td>6.000</td>
</tr>
<tr>
<td>Band7</td>
<td>-0.022</td>
<td>0.005</td>
<td>-0.004</td>
<td>0.005</td>
<td>7.000</td>
</tr>
</tbody>
</table>

### Water Bodies

<table>
<thead>
<tr>
<th>Eigenvector</th>
<th>Band1</th>
<th>Band2</th>
<th>Band3</th>
<th>Band4</th>
<th>Band5</th>
<th>Band6</th>
<th>Band7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Band1</td>
<td>-0.107</td>
<td>-0.102</td>
<td>0.047</td>
<td>-0.441</td>
<td>0.884</td>
<td>-0.015</td>
<td>0.016</td>
</tr>
<tr>
<td>Band2</td>
<td>0.112</td>
<td>0.082</td>
<td>0.033</td>
<td>0.875</td>
<td>0.459</td>
<td>0.050</td>
<td>-0.006</td>
</tr>
<tr>
<td>Band3</td>
<td>-0.138</td>
<td>0.267</td>
<td>-0.925</td>
<td>-0.012</td>
<td>0.058</td>
<td>0.174</td>
<td>0.145</td>
</tr>
<tr>
<td>Band4</td>
<td>-0.089</td>
<td>0.855</td>
<td>0.274</td>
<td>-0.112</td>
<td>0.027</td>
<td>0.298</td>
<td>-0.289</td>
</tr>
<tr>
<td>Band5</td>
<td>0.052</td>
<td>0.329</td>
<td>-0.101</td>
<td>0.000</td>
<td>0.036</td>
<td>-0.932</td>
<td>-0.094</td>
</tr>
<tr>
<td>Band6</td>
<td>-0.006</td>
<td>-0.257</td>
<td>-0.215</td>
<td>0.020</td>
<td>0.008</td>
<td>0.027</td>
<td>-0.942</td>
</tr>
<tr>
<td>Band7</td>
<td>0.973</td>
<td>0.076</td>
<td>-0.101</td>
<td>-0.161</td>
<td>0.053</td>
<td>0.094</td>
<td>-0.004</td>
</tr>
</tbody>
</table>

### Basic Stats

<table>
<thead>
<tr>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Stdev</th>
<th>Num</th>
<th>Eigenvalue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Band1</td>
<td>-0.436</td>
<td>0.305</td>
<td>-0.277</td>
<td>0.182</td>
<td>1.000</td>
</tr>
<tr>
<td>Band2</td>
<td>-1.406</td>
<td>0.380</td>
<td>-0.761</td>
<td>0.501</td>
<td>2.000</td>
</tr>
<tr>
<td>Band3</td>
<td>-0.174</td>
<td>0.696</td>
<td>0.376</td>
<td>0.251</td>
<td>3.000</td>
</tr>
<tr>
<td>Band4</td>
<td>-1.306</td>
<td>1.594</td>
<td>-0.695</td>
<td>0.586</td>
<td>4.000</td>
</tr>
<tr>
<td>Band5</td>
<td>-1.365</td>
<td>8.820</td>
<td>2.453</td>
<td>1.110</td>
<td>5.000</td>
</tr>
<tr>
<td>Band6</td>
<td>-0.199</td>
<td>0.075</td>
<td>-0.013</td>
<td>0.025</td>
<td>6.000</td>
</tr>
<tr>
<td>Band7</td>
<td>-0.021</td>
<td>0.129</td>
<td>0.084</td>
<td>0.037</td>
<td>7.000</td>
</tr>
</tbody>
</table>

### Eigenvector

<table>
<thead>
<tr>
<th>Band1</th>
<th>Band2</th>
<th>Band3</th>
<th>Band4</th>
<th>Band5</th>
<th>Band6</th>
<th>Band7</th>
</tr>
</thead>
<tbody>
<tr>
<td>-0.132</td>
<td>-0.357</td>
<td>0.176</td>
<td>-0.385</td>
<td>0.822</td>
<td>-0.003</td>
<td>0.027</td>
</tr>
</tbody>
</table>
### Pasture outside park

<table>
<thead>
<tr>
<th>Band</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Stdev</th>
<th>Num</th>
<th>Eigenvalue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Band1</td>
<td>-0.169</td>
<td>0.293</td>
<td>0.082</td>
<td>0.066</td>
<td>1.000</td>
<td>0.540</td>
</tr>
<tr>
<td>Band2</td>
<td>-0.388</td>
<td>0.365</td>
<td>0.026</td>
<td>0.097</td>
<td>2.000</td>
<td>0.093</td>
</tr>
<tr>
<td>Band3</td>
<td>-0.131</td>
<td>0.415</td>
<td>0.144</td>
<td>0.065</td>
<td>3.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band4</td>
<td>-1.428</td>
<td>1.541</td>
<td>0.206</td>
<td>0.422</td>
<td>4.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band5</td>
<td>-0.759</td>
<td>5.552</td>
<td>1.213</td>
<td>0.660</td>
<td>5.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band6</td>
<td>-0.297</td>
<td>0.060</td>
<td>-0.069</td>
<td>0.039</td>
<td>6.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band7</td>
<td>-0.020</td>
<td>0.045</td>
<td>0.001</td>
<td>0.007</td>
<td>7.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

### Savannah

<table>
<thead>
<tr>
<th>Band</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Stdev</th>
<th>Num</th>
<th>Eigenvalue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Band1</td>
<td>-0.086</td>
<td>0.102</td>
<td>0.004</td>
<td>0.044</td>
<td>1.000</td>
<td>0.156</td>
</tr>
<tr>
<td>Band2</td>
<td>-0.146</td>
<td>0.116</td>
<td>-0.015</td>
<td>0.059</td>
<td>2.000</td>
<td>0.052</td>
</tr>
<tr>
<td>Band3</td>
<td>-0.009</td>
<td>0.168</td>
<td>0.057</td>
<td>0.030</td>
<td>3.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band4</td>
<td>-0.510</td>
<td>0.546</td>
<td>0.014</td>
<td>0.224</td>
<td>4.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band5</td>
<td>0.460</td>
<td>3.224</td>
<td>1.239</td>
<td>0.388</td>
<td>5.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band6</td>
<td>-0.065</td>
<td>0.011</td>
<td>-0.031</td>
<td>0.012</td>
<td>6.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band7</td>
<td>0.008</td>
<td>0.048</td>
<td>0.023</td>
<td>0.007</td>
<td>7.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Band</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Stdev</th>
<th>Num</th>
<th>Eigenvalue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Band1</td>
<td>-0.063</td>
<td>-0.102</td>
<td>0.067</td>
<td>-0.152</td>
<td>0.979</td>
<td>-0.004</td>
</tr>
<tr>
<td>Band2</td>
<td>0.158</td>
<td>0.190</td>
<td>-0.025</td>
<td>0.950</td>
<td>0.180</td>
<td>0.048</td>
</tr>
<tr>
<td>Band3</td>
<td>-0.384</td>
<td>-0.109</td>
<td>-0.842</td>
<td>0.047</td>
<td>0.027</td>
<td>0.295</td>
</tr>
<tr>
<td>Band4</td>
<td>0.178</td>
<td>0.906</td>
<td>-0.251</td>
<td>-0.235</td>
<td>0.088</td>
<td>-0.020</td>
</tr>
<tr>
<td>Band5</td>
<td>0.527</td>
<td>-0.190</td>
<td>-0.384</td>
<td>-0.028</td>
<td>0.030</td>
<td>-0.675</td>
</tr>
</tbody>
</table>

### Eigenvector

#### Pasture outside park

<table>
<thead>
<tr>
<th>Band</th>
<th>Band1</th>
<th>Band2</th>
<th>Band3</th>
<th>Band4</th>
<th>Band5</th>
<th>Band6</th>
<th>Band7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Band1</td>
<td>-0.080</td>
<td>-0.122</td>
<td>0.084</td>
<td>-0.449</td>
<td>0.877</td>
<td>-0.042</td>
<td>0.009</td>
</tr>
<tr>
<td>Band2</td>
<td>-0.096</td>
<td>-0.123</td>
<td>0.028</td>
<td>-0.863</td>
<td>-0.474</td>
<td>-0.070</td>
<td>-0.002</td>
</tr>
<tr>
<td>Band3</td>
<td>-0.348</td>
<td>-0.012</td>
<td>-0.744</td>
<td>-0.050</td>
<td>0.038</td>
<td>0.558</td>
<td>0.100</td>
</tr>
<tr>
<td>Band4</td>
<td>-0.365</td>
<td>-0.300</td>
<td>0.616</td>
<td>0.086</td>
<td>-0.060</td>
<td>0.615</td>
<td>-0.094</td>
</tr>
<tr>
<td>Band5</td>
<td>-0.064</td>
<td>-0.875</td>
<td>-0.151</td>
<td>0.173</td>
<td>-0.041</td>
<td>-0.296</td>
<td>0.297</td>
</tr>
<tr>
<td>Band6</td>
<td>0.474</td>
<td>-0.335</td>
<td>-0.179</td>
<td>-0.031</td>
<td>0.014</td>
<td>0.184</td>
<td>-0.772</td>
</tr>
<tr>
<td>Band7</td>
<td>-0.708</td>
<td>0.046</td>
<td>-0.071</td>
<td>0.111</td>
<td>-0.009</td>
<td>-0.427</td>
<td>-0.545</td>
</tr>
</tbody>
</table>

#### Savannah

<table>
<thead>
<tr>
<th>Band</th>
<th>Band1</th>
<th>Band2</th>
<th>Band3</th>
<th>Band4</th>
<th>Band5</th>
<th>Band6</th>
<th>Band7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Band1</td>
<td>-0.063</td>
<td>-0.102</td>
<td>0.067</td>
<td>-0.152</td>
<td>0.979</td>
<td>-0.004</td>
<td>0.013</td>
</tr>
<tr>
<td>Band2</td>
<td>0.158</td>
<td>0.190</td>
<td>-0.025</td>
<td>0.950</td>
<td>0.180</td>
<td>0.048</td>
<td>-0.012</td>
</tr>
<tr>
<td>Band3</td>
<td>-0.384</td>
<td>-0.109</td>
<td>-0.842</td>
<td>0.047</td>
<td>0.027</td>
<td>0.295</td>
<td>0.205</td>
</tr>
<tr>
<td>Band4</td>
<td>0.178</td>
<td>0.906</td>
<td>-0.251</td>
<td>-0.235</td>
<td>0.088</td>
<td>-0.020</td>
<td>-0.144</td>
</tr>
<tr>
<td>Band5</td>
<td>0.527</td>
<td>-0.190</td>
<td>-0.384</td>
<td>-0.028</td>
<td>0.030</td>
<td>-0.675</td>
<td>0.285</td>
</tr>
<tr>
<td>Band6</td>
<td>-0.326</td>
<td>-0.103</td>
<td>-0.187</td>
<td>0.083</td>
<td>0.002</td>
<td>-0.458</td>
<td>-0.794</td>
</tr>
<tr>
<td>Band7</td>
<td>-0.638</td>
<td>0.272</td>
<td>0.202</td>
<td>0.092</td>
<td>-0.021</td>
<td>-0.494</td>
<td>0.474</td>
</tr>
</tbody>
</table>

**Actively cultivated site**

<table>
<thead>
<tr>
<th>Basic Stats</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Stdev</th>
<th>Num</th>
<th>Eigenvalue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Band1</td>
<td>-0.263</td>
<td>0.038</td>
<td>-0.101</td>
<td>0.084</td>
<td>1.000</td>
<td>0.506</td>
</tr>
<tr>
<td>Band2</td>
<td>-0.452</td>
<td>0.050</td>
<td>-0.158</td>
<td>0.128</td>
<td>2.000</td>
<td>0.048</td>
</tr>
<tr>
<td>Band3</td>
<td>-0.002</td>
<td>0.198</td>
<td>0.087</td>
<td>0.041</td>
<td>3.000</td>
<td>0.001</td>
</tr>
<tr>
<td>Band4</td>
<td>-0.677</td>
<td>0.222</td>
<td>-0.324</td>
<td>0.234</td>
<td>4.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band5</td>
<td>0.749</td>
<td>3.608</td>
<td>1.727</td>
<td>0.689</td>
<td>5.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band6</td>
<td>-0.068</td>
<td>-0.004</td>
<td>-0.032</td>
<td>0.015</td>
<td>6.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band7</td>
<td>0.015</td>
<td>0.079</td>
<td>0.042</td>
<td>0.019</td>
<td>7.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Eigenvector</th>
<th>Band1</th>
<th>Band2</th>
<th>Band3</th>
<th>Band4</th>
<th>Band5</th>
<th>Band6</th>
<th>Band7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Band1</td>
<td>-0.101</td>
<td>-0.160</td>
<td>0.053</td>
<td>-0.169</td>
<td>0.965</td>
<td>0.014</td>
<td>0.025</td>
</tr>
<tr>
<td>Band2</td>
<td>0.185</td>
<td>0.259</td>
<td>-0.050</td>
<td>0.919</td>
<td>0.226</td>
<td>0.033</td>
<td>-0.016</td>
</tr>
<tr>
<td>Band3</td>
<td>-0.496</td>
<td>-0.717</td>
<td>-0.127</td>
<td>0.315</td>
<td>-0.117</td>
<td>0.300</td>
<td>0.142</td>
</tr>
<tr>
<td>Band4</td>
<td>0.280</td>
<td>-0.843</td>
<td>0.114</td>
<td>-0.054</td>
<td>-0.163</td>
<td>-0.151</td>
<td>0.396</td>
</tr>
<tr>
<td>Band5</td>
<td>0.437</td>
<td>-0.395</td>
<td>0.030</td>
<td>0.005</td>
<td>-0.597</td>
<td>0.723</td>
<td>0.300</td>
</tr>
<tr>
<td>Band6</td>
<td>0.602</td>
<td>-0.012</td>
<td>-0.036</td>
<td>-0.117</td>
<td>0.010</td>
<td>0.723</td>
<td>0.300</td>
</tr>
<tr>
<td>Band7</td>
<td>0.279</td>
<td>-0.284</td>
<td>-0.358</td>
<td>-0.017</td>
<td>0.020</td>
<td>0.056</td>
<td>-0.842</td>
</tr>
</tbody>
</table>

**Anthropogenic vegetation**

<table>
<thead>
<tr>
<th>Basic Stats</th>
<th>Min</th>
<th>Max</th>
<th>Mean</th>
<th>Stdev</th>
<th>Num</th>
<th>Eigenvalue</th>
</tr>
</thead>
<tbody>
<tr>
<td>Band1</td>
<td>-0.366</td>
<td>0.223</td>
<td>0.038</td>
<td>0.206</td>
<td>1.000</td>
<td>1.676</td>
</tr>
<tr>
<td>Band2</td>
<td>-0.995</td>
<td>0.301</td>
<td>0.017</td>
<td>0.389</td>
<td>2.000</td>
<td>0.038</td>
</tr>
<tr>
<td>Band3</td>
<td>-0.152</td>
<td>0.496</td>
<td>-0.008</td>
<td>0.174</td>
<td>3.000</td>
<td>0.007</td>
</tr>
<tr>
<td>Band4</td>
<td>-0.971</td>
<td>1.111</td>
<td>0.384</td>
<td>0.707</td>
<td>4.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band5</td>
<td>-0.547</td>
<td>2.650</td>
<td>0.589</td>
<td>0.998</td>
<td>5.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band6</td>
<td>-0.014</td>
<td>0.081</td>
<td>0.017</td>
<td>0.016</td>
<td>6.000</td>
<td>0.000</td>
</tr>
<tr>
<td>Band7</td>
<td>-0.007</td>
<td>0.084</td>
<td>0.024</td>
<td>0.032</td>
<td>7.000</td>
<td>0.000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Eigenvector</th>
<th>Band1</th>
<th>Band2</th>
<th>Band3</th>
<th>Band4</th>
<th>Band5</th>
<th>Band6</th>
<th>Band7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Band1</td>
<td>-0.158</td>
<td>-0.291</td>
<td>0.126</td>
<td>-0.537</td>
<td>0.765</td>
<td>-0.011</td>
<td>0.024</td>
</tr>
<tr>
<td>Band2</td>
<td>0.117</td>
<td>0.409</td>
<td>-0.209</td>
<td>0.605</td>
<td>0.639</td>
<td>0.030</td>
<td>-0.003</td>
</tr>
<tr>
<td>Band3</td>
<td>-0.045</td>
<td>0.652</td>
<td>-0.520</td>
<td>-0.546</td>
<td>-0.061</td>
<td>-0.024</td>
<td>0.018</td>
</tr>
<tr>
<td>Band4</td>
<td>0.506</td>
<td>0.402</td>
<td>0.642</td>
<td>-0.170</td>
<td>0.034</td>
<td>-0.311</td>
<td>-0.208</td>
</tr>
<tr>
<td>Band5</td>
<td>0.783</td>
<td>-0.381</td>
<td>-0.458</td>
<td>-0.079</td>
<td>0.033</td>
<td>-0.148</td>
<td>0.047</td>
</tr>
<tr>
<td>Band6</td>
<td>0.275</td>
<td>0.064</td>
<td>0.120</td>
<td>-0.107</td>
<td>0.004</td>
<td>0.934</td>
<td>-0.150</td>
</tr>
<tr>
<td>Band7</td>
<td>0.119</td>
<td>0.111</td>
<td>0.185</td>
<td>-0.024</td>
<td>-0.009</td>
<td>0.086</td>
<td>0.965</td>
</tr>
</tbody>
</table>

**Transition**

<table>
<thead>
<tr>
<th>Basic Stats</th>
<th>Max</th>
<th>Stdev</th>
<th>Num</th>
<th>Eigenvalue</th>
</tr>
</thead>
</table>
Table A-5. The separability values for the 2002 (August) classes
(Jeffries-Matusita, Transformed Divergence)

<table>
<thead>
<tr>
<th></th>
<th>Min</th>
<th>Mean</th>
<th></th>
<th>Min</th>
<th>Mean</th>
<th></th>
<th></th>
<th>Mean</th>
<th></th>
<th>Mean</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>Band1</td>
<td>-0.081</td>
<td>0.135</td>
<td>0.025</td>
<td>0.063</td>
<td>1.000</td>
<td>0.259</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Band2</td>
<td>-0.151</td>
<td>0.143</td>
<td>0.018</td>
<td>0.078</td>
<td>2.000</td>
<td>0.043</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Band3</td>
<td>-0.022</td>
<td>0.141</td>
<td>0.042</td>
<td>0.035</td>
<td>3.000</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Band4</td>
<td>-0.412</td>
<td>0.763</td>
<td>0.185</td>
<td>0.267</td>
<td>4.000</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Band5</td>
<td>0.426</td>
<td>2.661</td>
<td>1.252</td>
<td>0.469</td>
<td>5.000</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Band6</td>
<td>-0.047</td>
<td>0.007</td>
<td>-0.018</td>
<td>0.011</td>
<td>6.000</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Band7</td>
<td>0.008</td>
<td>0.048</td>
<td>0.023</td>
<td>0.010</td>
<td>7.000</td>
<td>0.000</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Eigenvector</th>
<th>Band1</th>
<th>Band2</th>
<th>Band3</th>
<th>Band4</th>
<th>Band5</th>
<th>Band6</th>
<th>Band7</th>
</tr>
</thead>
<tbody>
<tr>
<td>Band1</td>
<td>-0.107</td>
<td>-0.141</td>
<td>0.059</td>
<td>-0.382</td>
<td>0.905</td>
<td>-0.011</td>
<td>0.018</td>
</tr>
<tr>
<td>Band2</td>
<td>0.148</td>
<td>0.142</td>
<td>0.026</td>
<td>0.886</td>
<td>0.413</td>
<td>0.033</td>
<td>-0.012</td>
</tr>
<tr>
<td>Band3</td>
<td>-0.399</td>
<td>0.136</td>
<td>-0.848</td>
<td>0.040</td>
<td>0.046</td>
<td>0.278</td>
<td>0.149</td>
</tr>
<tr>
<td>Band4</td>
<td>0.337</td>
<td>0.890</td>
<td>-0.078</td>
<td>-0.235</td>
<td>0.086</td>
<td>-0.103</td>
<td>-0.126</td>
</tr>
<tr>
<td>Band5</td>
<td>0.556</td>
<td>-0.294</td>
<td>-0.459</td>
<td>-0.021</td>
<td>0.029</td>
<td>-0.584</td>
<td>0.230</td>
</tr>
<tr>
<td>Band6</td>
<td>-0.528</td>
<td>0.248</td>
<td>0.195</td>
<td>-0.081</td>
<td>-0.019</td>
<td>-0.576</td>
<td>0.531</td>
</tr>
<tr>
<td>Band7</td>
<td>-0.326</td>
<td>-0.040</td>
<td>-0.148</td>
<td>0.071</td>
<td>0.004</td>
<td>-0.488</td>
<td>-0.792</td>
</tr>
</tbody>
</table>

Water Bodies [Yellow] 10615 points:
- Non-anthropogenic forest [Maroon] 494 points: (2.00000000 2.00000000)
- Savannah [Purple] 292 points: (2.00000000 2.00000000)
- Inlet areas [Blue1] 305 points: (2.00000000 2.00000000)
- Culturally active [Blue2] 402 points: (2.00000000 2.00000000)
- Transitional [Blue3] 115 points: (2.00000000 2.00000000)
- Bare soil [Yellow3] 26 points: (2.00000000 2.00000000)
- Anthropogenic vegetation [Green] 116 points: (2.00000000 2.00000000)
- Pasture outside park boundary [Cyan] 2549 points: (2.00000000 2.00000000)
- Aldea [Green1] 55 points: (2.00000000 2.00000000)
- Actively cultivated site [Red] 142 points: (2.00000000 2.00000000)

Non-anthropogenic forest [Maroon] 494 points:
- Water Bodies [Yellow] 10615 points: (2.00000000 2.00000000)
- Savannah [Purple] 292 points: (2.00000000 2.00000000)
- Inlet areas [Blue1] 305 points: (1.45663532 1.87789118)
- Culturally active [Blue2] 402 points: (1.99995519 2.00000000)
- Transitional [Blue3] 115 points: (1.56012337 1.94761477)
- Bare soil [Yellow3] 26 points: (2.00000000 2.00000000)
- Anthropogenic vegetation [Green] 116 points: (1.26343197 1.37078161)
- Pasture outside park boundary [Cyan] 2549 points: (2.00000000 2.00000000)
Aldea [Green1] 55 points: (2.00000000 2.00000000)
Actively cultivated site [Red] 142 points: (1.70913325 1.99990950)

Savannah [Purple] 292 points:
Water Bodies [Yellow] 10615 points: (2.00000000 2.00000000)
Non-anthropogenic forest [Maroon] 494 points: (2.00000000 2.00000000)
Inlet areas [Blue1] 305 points: (2.00000000 2.00000000)
Culturally active [Blue2] 402 points: (1.99999997 2.00000000)
Transitional [Blue3] 115 points: (2.00000000 2.00000000)
Bare soil [Yellow3] 26 points: (2.00000000 2.00000000)
Anthropogenic vegetation [Green] 116 points: (2.00000000 2.00000000)
Pasture outside park boundary [Cyan] 2549 points: (2.00000000 2.00000000)
Aldea [Green1] 55 points: (1.99999960 2.00000000)
Actively cultivated site [Red] 142 points: (1.9999988 2.00000000)

Inlet areas [Blue1] 305 points:
Water Bodies [Yellow] 10615 points: (2.00000000 2.00000000)
Non-anthropogenic forest [Maroon] 494 points: (1.45663532 1.87789118)
Savannah [Purple] 292 points: (2.00000000 2.00000000)
Culturally active [Blue2] 402 points: (1.98013285 2.00000000)
Transitional [Blue3] 115 points: (0.92378415 0.99379191)
Bare soil [Yellow3] 26 points: (1.9999959 2.00000000)
Anthropogenic vegetation [Green] 116 points: (1.35877650 1.92027862)
Pasture outside park boundary [Cyan] 2549 points: (2.00000000 2.00000000)
Aldea [Green1] 55 points: (2.00000000 2.00000000)
Actively cultivated site [Red] 142 points: (1.14340612 1.43507853)

Culturally active [Blue2] 402 points:
Water Bodies [Yellow] 10615 points: (2.00000000 2.00000000)
Non-anthropogenic forest [Maroon] 494 points: (1.99995519 2.00000000)
Savannah [Purple] 292 points: (1.9999997 2.00000000)
Inlet areas [Blue1] 305 points: (1.98013285 2.00000000)
Transitional [Blue3] 115 points: (1.95837189 1.9999983)
Bare soil [Yellow3] 26 points: (1.90475146 2.00000000)
Anthropogenic vegetation [Green] 116 points: (1.99976972 2.00000000)
Pasture outside park boundary [Cyan] 2549 points: (1.99998766 2.00000000)
Aldea [Green1] 55 points: (2.00000000 2.00000000)
Actively cultivated site [Red] 142 points: (1.88989777 1.99238816)

Transitional [Blue3] 115 points:
Water Bodies [Yellow] 10615 points: (2.00000000 2.00000000)
Non-anthropogenic forest [Maroon] 494 points: (1.56012337 1.94761477)
Savannah [Purple] 292 points: (2.00000000 2.00000000)
Inlet areas [Blue1] 305 points: (0.92378415 0.99379191)
Culturally active [Blue2] 402 points: (1.95837189 1.99999983)
Bare soil [Yellow3] 26 points: (1.99999778 2.00000000)
Anthropogenic vegetation [Green] 116 points: (1.69947247 1.99275216)
Pasture outside park boundary [Cyan] 2549 points: (2.00000000 2.00000000)
Aldea [Green1] 55 points: (2.00000000 2.00000000)
Actively cultivated site [Red] 142 points: (1.20400230 1.52254514)

Bare soil [Yellow3] 26 points:
  Water Bodies [Yellow] 10615 points: (2.00000000 2.00000000)
  Non-anthropogenic forest [Maroon] 494 points: (2.00000000 2.00000000)
  Savannah [Purple] 292 points: (2.00000000 2.00000000)
  Inlet areas [Blue1] 305 points: (1.99999959 2.00000000)
  Culturally active [Blue2] 402 points: (1.90475146 2.00000000)
  Transitional [Blue3] 115 points: (1.99999778 2.00000000)
  Anthropogenic vegetation [Green] 116 points: (2.00000000 2.00000000)
  Pasture outside park boundary [Cyan] 2549 points: (2.00000000 2.00000000)
  Aldea [Green1] 55 points: (1.99999642 2.00000000)
  Actively cultivated site [Red] 142 points: (1.99412867 2.00000000)

Anthropogenic vegetation [Green] 116 points:
  Water Bodies [Yellow] 10615 points: (2.00000000 2.00000000)
  Non-anthropogenic forest [Maroon] 494 points: (1.26343197 1.37078161)
  Savannah [Purple] 292 points: (2.00000000 2.00000000)
  Inlet areas [Blue1] 305 points: (1.35877650 1.92027862)
  Culturally active [Blue2] 402 points: (1.99976972 2.00000000)
  Transitional [Blue3] 115 points: (1.69947247 1.99275216)
  Bare soil [Yellow3] 26 points: (2.00000000 2.00000000)
  Pasture outside park boundary [Cyan] 2549 points: (2.00000000 2.00000000)
  Aldea [Green1] 55 points: (2.00000000 2.00000000)
  Actively cultivated site [Red] 142 points: (1.57018403 1.99927390)

Pasture outside park boundary [Cyan] 2549 points:
  Water Bodies [Yellow] 10615 points: (2.00000000 2.00000000)
  Non-anthropogenic forest [Maroon] 494 points: (2.00000000 2.00000000)
  Savannah [Purple] 292 points: (2.00000000 2.00000000)
  Inlet areas [Blue1] 305 points: (2.00000000 2.00000000)
  Culturally active [Blue2] 402 points: (1.99998766 2.00000000)
  Transitional [Blue3] 115 points: (2.00000000 2.00000000)
  Bare soil [Yellow3] 26 points: (2.00000000 2.00000000)
  Anthropogenic vegetation [Green] 116 points: (1.99986922 1.99999504)
  Aldea [Green1] 55 points: (2.00000000 2.00000000)
  Actively cultivated site [Red] 142 points: (2.00000000 2.00000000)

Aldea [Green1] 55 points:
  Water Bodies [Yellow] 10615 points: (2.00000000 2.00000000)
  Non-anthropogenic forest [Maroon] 494 points: (2.00000000 2.00000000)
  Savannah [Purple] 292 points: (1.99999960 2.00000000)
  Inlet areas [Blue1] 305 points: (2.00000000 2.00000000)
Table A-6. The separability values for the 2003 (May) classes (Jeffries-Matusita, Transformed Divergence)

<table>
<thead>
<tr>
<th>Class</th>
<th>Separability Value</th>
</tr>
</thead>
<tbody>
<tr>
<td>Water Bodies [Yellow]</td>
<td>10615 points: (2.00000000 2.00000000)</td>
</tr>
<tr>
<td>Non-anthropogenic forest [Maroon]</td>
<td>494 points: (1.99815024 2.00000000)</td>
</tr>
<tr>
<td>Savannah [Purple]</td>
<td>292 points: (1.99179107 2.00000000)</td>
</tr>
<tr>
<td>Inlet areas [Blue1]</td>
<td>305 points: (1.99762303 2.00000000)</td>
</tr>
<tr>
<td>Culturally active [Blue2]</td>
<td>402 points: (1.98447784 2.00000000)</td>
</tr>
<tr>
<td>Transitional [Blue3]</td>
<td>115 points: (1.98329705 2.00000000)</td>
</tr>
<tr>
<td>Bare soil [Yellow3]</td>
<td>26 points: (1.99736425 2.00000000)</td>
</tr>
<tr>
<td>Anthropogenic vegetation [Green]</td>
<td>116 points: (1.79684895 1.99998233)</td>
</tr>
<tr>
<td>Pasture outside park boundary [Cyan]</td>
<td>2549 points: (2.00000000 2.00000000)</td>
</tr>
<tr>
<td>Aldea [Green1]</td>
<td>55 points: (1.99052248 2.00000000)</td>
</tr>
<tr>
<td>Actively cultivated site [Red]</td>
<td>142 points: (1.94201223 2.00000000)</td>
</tr>
</tbody>
</table>

Non-anthropogenic forest [Maroon] 494 points:
Water Bodies [Yellow] 10615 points: (1.99815024 2.00000000)
Savannah [Purple] 292 points: (1.99999346 2.00000000)
Inlet areas [Blue1] 305 points: (1.07920289 1.76202451)
Culturally active [Blue2] 402 points: (1.99996033 2.00000000)
Transitional [Blue3] 115 points: (1.99953199 2.00000000)
Bare soil [Yellow3] 26 points: (2.00000000 2.00000000)
Anthropogenic vegetation [Green] 116 points: (1.94755274 2.00000000)
Pasture outside park boundary [Cyan] 2549 points: (1.99999960 2.00000000)
Aldea [Green1] 55 points: (2.00000000 2.00000000)
Actively cultivated site [Red] 142 points: (1.99999997 2.00000000)

Savannah [Purple] 292 points:
Water Bodies [Yellow] 10615 points: (1.99179107 2.00000000)
Non-anthropogenic forest [Maroon] 494 points: (1.99999346 2.00000000)
Inlet areas [Blue1] 305 points: (1.99999997 2.00000000)
Culturally active [Blue2] 402 points: (0.59405983 0.70226112)
Transitional [Blue3] 115 points: (0.93992958 1.14855702)
Bare soil [Yellow3] 26 points: (1.55669068 1.96453161)
Anthropogenic vegetation [Green] 116 points: (1.99993939 2.00000000)
Pasture outside park boundary [Cyan] 2549 points: (1.99973826 2.00000000)
Aldea [Green1] 55 points: (1.86154996 1.99999970)
Actively cultivated site [Red] 142 points: (1.61879949 1.99999989)

Inlet areas [Blue1] 305 points:
Water Bodies [Yellow] 10615 points: (1.99762303 2.00000000)
Non-anthropogenic forest [Maroon] 494 points: (1.07920289 1.76202451)
Savannah [Purple] 292 points: (1.99999539 2.00000000)
Culturally active [Blue2] 402 points: (1.9997316 2.00000000)
Transitional [Blue3] 115 points: (1.99969323 2.00000000)
Bare soil [Yellow3] 26 points: (2.00000000 2.00000000)
Anthropogenic vegetation [Green] 116 points: (1.93948949 2.00000000)
Pasture outside park boundary [Cyan] 2549 points: (1.99997083 2.00000000)
Aldea [Green1] 55 points: (2.00000000 2.00000000)
Actively cultivated site [Red] 142 points: (1.99999987 2.00000000)

Culturally active [Blue2] 402 points:
Water Bodies [Yellow] 10615 points: (1.98447784 2.00000000)
Non-anthropogenic forest [Maroon] 494 points: (1.9996033 2.00000000)
Savannah [Purple] 292 points: (0.59405983 0.70226112)
Inlet areas [Blue1] 305 points: (1.99997316 2.00000000)
Transitional [Blue3] 115 points: (0.44881983 0.50181532)
Bare soil [Yellow3] 26 points: (1.69078470 1.99996587)
Anthropogenic vegetation [Green] 116 points: (1.99959838 2.00000000)
Pasture outside park boundary [Cyan] 2549 points: (1.99488084 1.9999996)
Aldea [Green1] 55 points: (1.83189435 2.00000000)
Actively cultivated site [Red] 142 points: (1.99773774 2.00000000)

Transitional [Blue3] 115 points:
Water Bodies [Yellow] 10615 points: (1.98329705 2.00000000)
Non-anthropogenic forest [Maroon] 494 points: (1.99953199 2.00000000)
Savannah [Purple] 292 points: (0.93992958 1.14855702)
Inlet areas [Blue1] 305 points: (1.99999346 2.00000000)
Culturally active [Blue2] 402 points: (0.44881983 0.50181532)
Bare soil [Yellow3] 26 points: (1.86435250 1.99999531)
Anthropogenic vegetation [Green] 116 points: (1.99737630 2.00000000)
Pasture outside park boundary [Cyan] 2549 points: (1.99554694 2.00000000)
Aldea [Green1] 55 points: (1.86699267 2.00000000)
Actively cultivated site [Red] 142 points: (1.72551744 1.99999777)

Bare soil [Yellow3] 26 points:
  Water Bodies [Yellow] 10615 points: (1.99736425 2.00000000)
  Non-anthropogenic forest [Maroon] 494 points: (2.00000000 2.00000000)
  Savannah [Purple] 292 points: (1.55669068 1.96453161)
  Inlet areas [Blue1] 305 points: (2.00000000 2.00000000)
  Culturally active [Blue2] 402 points: (1.69078470 1.99996587)
  Transitional [Blue3] 115 points: (1.86435250 1.9999531)
  Anthropogenic vegetation [Green] 116 points: (1.99999994 2.00000000)
  Pasture outside park boundary [Cyan] 2549 points: (2.00000000 2.00000000)
  Aldea [Green1] 55 points: (1.97474179 2.00000000)
  Actively cultivated site [Red] 142 points: (1.80953818 2.00000000)

Anthropogenic vegetation [Green] 116 points:
  Water Bodies [Yellow] 10615 points: (1.79684895 1.99998233)
  Non-anthropogenic forest [Maroon] 494 points: (1.94755274 2.00000000)
  Savannah [Purple] 292 points: (1.99993939 2.00000000)
  Inlet areas [Blue1] 305 points: (1.93948949 2.00000000)
  Culturally active [Blue2] 402 points: (1.99959838 2.00000000)
  Transitional [Blue3] 115 points: (1.99737630 2.00000000)
  Bare soil [Yellow3] 26 points: (1.99999994 2.00000000)
  Pasture outside park boundary [Cyan] 2549 points: (1.99989408 2.00000000)
  Aldea [Green1] 55 points: (1.99947366 2.00000000)
  Actively cultivated site [Red] 142 points: (1.99580144 1.9999993)

Pasture outside park boundary [Cyan] 2549 points:
  Water Bodies [Yellow] 10615 points: (1.99989850 2.00000000)
  Non-anthropogenic forest [Maroon] 494 points: (1.99999960 2.00000000)
  Savannah [Purple] 292 points: (1.99973826 2.00000000)
  Inlet areas [Blue1] 305 points: (1.99997083 2.00000000)
  Culturally active [Blue2] 402 points: (1.99488084 1.99999996)
  Transitional [Blue3] 115 points: (1.99554694 2.00000000)
  Bare soil [Yellow3] 26 points: (2.00000000 2.00000000)
  Anthropogenic vegetation [Green] 116 points: (1.99989408 2.00000000)
  Aldea [Green1] 55 points: (1.9990138 2.00000000)
  Actively cultivated site [Red] 142 points: (1.99992405 2.00000000)

Aldea [Green1] 55 points:
  Water Bodies [Yellow] 10615 points: (1.99052248 2.00000000)
  Non-anthropogenic forest [Maroon] 494 points: (2.00000000 2.00000000)
  Savannah [Purple] 292 points: (1.86154996 1.9999970)
Inlet areas [Blue1] 305 points: (2.00000000 2.00000000)
Culturally active [Blue2] 402 points: (1.83189435 2.00000000)
Transitional [Blue3] 115 points: (1.86699267 2.00000000)
Bare soil [Yellow3] 26 points: (1.97474179 2.00000000)
Anthropogenic vegetation [Green] 116 points: (1.99947366 2.00000000)
Pasture outside park boundary [Cyan] 2549 points: (1.99990138 2.00000000)
Actively cultivated site [Red] 142 points: (1.74285492 1.99994819)

Actively cultivated site [Red] 142 points:
   Water Bodies [Yellow] 10615 points: (1.94201223 2.00000000)
   Non-anthropogenic forest [Maroon] 494 points: (1.9999997 2.00000000)
   Savannah [Purple] 292 points: (1.61879949 1.99999989)
   Inlet areas [Blue1] 305 points: (1.99999987 2.00000000)
   Culturally active [Blue2] 402 points: (1.60733774 1.99976504)
   Transitional [Blue3] 115 points: (1.72551744 1.99999777)
   Bare soil [Yellow3] 26 points: (1.80953818 2.00000000)
   Anthropogenic vegetation [Green] 116 points: (1.99580144 1.99999993)
   Pasture outside park boundary [Cyan] 2549 points: (1.99992405 2.00000000)
   Aldea [Green] 55 points: (1.74285492 1.99994819)
LIST OF REFERENCES

Achard, F., H. Eva and P. Mayaux

Adams, J., D. Sabol, V. Kapos, R. Almeida, D. Roberts, M. Smith and A. Gillespie

Agrawal, A. and C. Gibson

Aldenderfer, M. S. and H. D. G. Maschner

Allen, K. M. S., S. W. Green and E. B. W. Zubrow

Allen, K. M. S., S. W. Green and E. B. Zubrow

Allen, T. F. H. and T. B. Starr

Altschul, J. H.

Alves, D. S.
Alves, D. S., J. Pereira, C. de Sousa, J. Soares and F. Yamaguchi

Anderson, A. B.


Anderson, J. R.

Anselin, L.

Arrow, K., B. Bolin, R. Costanza, P. Dasgupta and C. Folke

Ashmore, W. and A. B. Knapp

Asrar, G., R. B. Myneni and B. J. Choudhury

Baerwald, T. J.

Baker, A. R. H. and G. Biger
Baker, W. L.

Balee, W.

Balee, W. and D. Campbell

Balee, W. L.

Balick, M.

Basso, K. H.

Bates, M.

Baxter, J. M.

Becker, B. K.

Beckerman, S.
Behrens, C. A.
1992  A Formal Justification for the Application of GIS to the Culture
Ecological Analysis of Land-use Intensification and Deforestation in the Amazon.
In The Anthropology of Human Behavior through Geographic Information
Analysis, edited by M. S. Aldenderfer and H. Maschner, pp. 16. University of
California, Santa Barbara.


Behrens, C. A., M. G. Baksh and M. Mothes
1994  A Regional-Analysis of Bari Land-Use Intensification and Its Impact on

Bellehumeur, C. and P. Legendre
1998  Multiscale Sources of Variation in Ecological Variables: Modeling Spatial

Bender, B.
Berg, Providence.

Berkes, F., C. Folke and J. Colding
Mechanisms for Building Resilience. Cambridge University Press, Cambridge,
New York, NY.

Berry, B.
1964  Approaches to Regional Analysis. Annals of the Association of American
Geographers 54:2-11.

Berry, J. K.
1995  Spatial Reasoning for Effective GIS. GIS World Books, Fort Collins,
Colorado.

Bettinger, R.
1977  Predicting the Archaeological Potential of the Inyo-Mono Region of

Bian, L. and R. Butler
1999  Comparing Effects of Aggregation Methods on Statistical and Spatial
Properties of Simulated Spatial Data. Photogrammetric Engineering and Remote
Sensing 65(1):12.
Bilsborrow, R. and D. Hogan (editor)

Binswanger, H. P.

Boas, F.

Boaz, J. and E. Uleberg

Boserup, E.


Bradley, J. E., W. R. Killam, G. R. Burns and M. A. Martorano

Brandt, R., B. Groenewoudt and K. Kvamme

Brondizio, E. S.

Brondizio, E. S., E. F. Moran, P. Mausel and Y. Wu

Brondizio, E., E. Moran, P. Mausel and Y. Wu
Brondizio, E., S. McCracken, E. Moran, A. Siqueira, D. Nelson and C. Rodrigues-Pedraza

Browder, J. O.

Browder, J. O. and B. J. Godfrey

Brown, M. K.

Brown, S. and A. Lugo

Bunker, S. G.

Burkey, T. V.

Burrough, P. A.


Burrough, P. A. and A. U. Frank

Burrough, P. A., R. McDonnell and P. A. Burrough
Burt, J. E., G. M. Barber and G. M. Barber  

Buschmann, C. and E. Nagel  

Bush, M., D. Piperno and P. Colinvaux  

Butzer, K. W.  


Campbell, D., G. Prance, U. Maciel  

Campbell, J. B.  


Campbell, J. B. and J. O. Browder  

Cao, C. and N. Lam  

Carmichael, D. L.  

Carneiro, R.


Carr, C. and University of Arkansas Fayetteville (Institute for Quantitative Archaeology)

Carr, T. and M. Turner

Castro, Eduardo V. d.


Castro, Eduardo V. d., M. C. d. Cunha and S. Dreyfus
1993 *Amazônia: Etnologia e História Indígena*. Núcleo de História Indígena e do Indigenismo: Fundação de Amparo à Pesquisa do Estado de São Paulo, São Paulo, Brazil.

Cecchi, G., P. Mazzinghi, L. Pantani, R. Valentini, D. Tirelli and P. Deangelis

Chavez, P. S.

Chen, R.

Chernela, J. M.

Cherry, J. F., J. L. Davis and E. Mantzourani
1991 Landscape Archaeology as Long-Term History: Northern Keos in the Cycladic Islands from Earliest Settlement until Modern Times. UCLA Institute of Archaeology, Los Angeles.

Chomentowski, W., B. Salas and D. Skole

Choudhury, B. J.

Christaller, W. and C. W. Baskin
1966 *Central places in Southern Germany*. Prentice-Hall, Englewood Cliffs, N.J.

Church, T., R. Brandon and G. Burgett
Clarke, D. L.


Clarke, D. L. and B. Chapman

Clastres, P.

Clay, J. W.

Clements, F. E.


Cleveland, W. and R. McGill

Cochrane, T. and P. Sanchez

Colby, J. D. and P. L. Keating

Conant, F. P.
Conese, C. and F. Maselli

Congalton, R.

Congalton, R. and K. Green

Congalton, R. G.

Congalton, R. G. and R. A. Mead

Congalton, R. G., R. G. Odeerwald and R. A. Mead

Cooke, R. and D. Harris

Cooney, G.

Cordell, L. S.

Corlett, R. T.
Costanza, R., L. Wainger, C. Folke and K.-G. Maler

Coulson, R., C. Lovelady, R. Flamm, S. Spradling and M. Saunders

Cowgill, G.

Cox, C.

Cracknell, A. P. and L. Hayes

Crist, E. P. and R. C. Cicone

Crist, E. P. and R. J. Kauth

Crumley, C.

Crumley, C. and W. Marquardt

Custer, J., T. Eveleigh, V. Klemas, I. Wells
Dale, V. and S. Pearson  

Dale, V. H., R. V. O'Neill, M. Pedlowski and F. Southworth  

Dale, V. H., R. V. O'Neill, F. Southworth and M. Pedlowski  

Dale, V. H., S. M. Pearson, H. L. Offerman and R. V. O'Neill  

Dalla, B. L.  


Dangermond, J.  


Davis, J. C.  

de Castro, F., M. C. Silva-Forsberg, W. Wilson, E. Brondâizâio and E. Moran  
Dean, J. S.

Decker, D.

DeMers, M. N.

Denevan, W.


Denevan, W. M. and C. Padoch

Dickinson, R. E. and United Nations University.

Domon, G., M. Gariépy and A. Bouchard

Douglas, I.
Dunnell, R. and W. Dancey

Ebert, J. I.


Ebert, J. I. and T. Kohler

Ebert, J. I. and T. Lyons

Ebert, J. I., E.L. Camilli and M.J. Berman

Effland, R.

Ehlers, M.

Eidt, R.
Elvidge, C. D. and Chen, Z.

Erickson, C.

Eva, H. and E. F. Lambin

Falesi, I. C.

Faust, N. L., W. Anderson and J. Star

Fedje, D. and T. Christensen

Fish, S. K. and S. A. Kowalewski

Fisher, A.

Flannery, K. V.

Flowers, N.
Foody, G. M. and P. J. Curran


Foresman, T. W.

Forman, R. T. T. and M. Godron

Forster, B.

Fortin, M. J.

Franchetto, B. and M. Heckenberger

Freeman, A.

Frohn, R. C.

Fullagar, R. and L. Head
Fuller, R., J. Sheail and C. Barr

Gaffney, V. and M. van Leusen

Gaffney, V., Z. Stancic and H. Watson


Gardner, R. H. and M. G. Turner

Gardner, W. M.
1978 Comparison of Ridge and Valley, Blue Ridge, Piedmont and Coastal Plain Archaic Period Site Distribution: An Idealized Transect (Preliminary Model). In *Middle Atlantic Archaeological Conference*, Rehoboth Beach, Delaware.

Geoghegan, J.

Gholz, H. L.

Gholz, H. L., Nakane, K. and Shimoda, H.
Gillespie, A. R., Kahle, A.B. and Walker, R.E.  

Gillings, M.  

Glaser B. E., Haumaier L., Guggenberger G., Zech W  

Gonzalo, A.  

Goodchild, M., R. Haining and S. Wise and 12 others  

Goodchild, M. F.  


Goodchild, M. F. and National Center for Geographic Information & Analysis (U.S.)  

Goodchild, M. F., B. O. Parks and L. T. Steyaert  

Goodland, R. J. A. and H. S. Irwin  
Gould, P.

Gow, P.

Grant, L.

Green, E. L.

Green, S. W.

Gregor, T.

Gross, D.


Gumerman, G. J. and T. Lyons

Gunn, J.

Guyot, G. and X. F. Gu
Haffer, J.  

Haines-Young, R., D. R. Green and S. H. Cousins (editor)  

Hames, R.  

Hames, R. and W. Vikers (editor)  

Harris, T. and G. Lock  

Hasenstab, R. and B. Resnick  

Hecht, S. B. and A. Cockburn  

Heckenberger, M. J.  


Heckenberger, M., J. Petersen and E. G. Neves  
2001  Of Lost Civilizations and Primitive Tribes, Amazonia: Reply to Meggers.  

Heckenberger, M., J. Petersen, E. Neves  
1999  Village Permanence in Amazonia: Two Archaeological Examples from Brazil.  

2003  Amazonia 1492: Pristine Forest or Cultural Parkland?  

Helmer, E. H., S. Brown and W. B. Cohen  
2000  Mapping Montane Tropical Forest Successional Stage and Land Use With Multi-Date Landsat Imagery.  

Herrera, L. F., L. Cavelier, C. Rodriguesz and S. Mora  
1992  The Technical Transformation of an Agricultural System in the Colombian Amazon.  

Hill, J. D.  

Hill, J. D. and F. Santos-Granero  

Hill, J. D. and B. Sturm  
1991  Radiometric Correction of Multitemporal Thematic Mapper Data for Use in Agricultural Land-Cover Classification and Vegetation Monitoring.  

Hill, R. A.  
1999  Image Segmentation for Humid Tropical Forest Classification in Landsat TM Data.  

Hill, R. A. and G. M. Foody  
1994  Separability of Tropical Rain-Forest Types in the Tambopata-Candamo Reserved Zone, Peru.  

Hofmann-Wellenhof, B., H. Lichtenegger and J. Collins  
Hord, R. M. and W. Brooner

Huete, A., C. Justice, and H. Liu

Huete, A.

Hunt, E. D.

Jackson, R. D.

Jacoli, M. and A. Carrara

James, S. E. and R. Knudson

Janssen, L. and F. van der Wel

Jensen, J. R.


Jensen, R. R., G. Yu, P. W. Mausel, V. Lulla, E. Brondizio and E. Moran
Jochim, M. A.


Johansen, M. E., Tommervik, H., Guneriussen, T. and Pedersen, J. P.

Johnson, A. W. and T. K. Earle

Johnson, G.


Jordan, C. F.


Joseph, J. W.

Judge, J. and L. Sebastian (editor)
Judge, W. J.

Jurdant, M.

Kalkahan, M. A., R. M. Reich and R. L. Czaplewski

Kaplan, E. D.

Kauth, R. J. and G. Thomas

Keene, A. S.

Kent, M., W. J. Gill, R. E. Weaver and R. P. Armitage

Kern, D. C. and M. Costa

Kern, D. C. and N. Kämpf


Koussoulakou, A. and E. Stylianidis

Kowalewski, S. A.


Kowalewski, S. A. and S. Fish

Kruckman, L.

Kvamme, K. L.


Labovitz, M. L. and J. W. Marvin

Ladefoged, T., S. McLachlan, S. Ross, P. Sheppard and D. Sutton

Lambin, E. F.

Langley, R. B.

1998 RTK GPS. GPS World 9(9):70-76.

Lathrap, D. W.


1977 Our Father the Cayman, Our Mother the Gourd: Spinden Revisited, or a Unitary Model for the Emergence of Agriculture in the New World. In Origins of Agriculture, edited by C. Reed, pp. 713-751.

Lauer, D., Estes, J., Jensen, J., Greenlee, D.

Lavreau, J.

Law, B. E. and R. H. Waring

Layton, R. and P. Ucko

Lee, C. T. and S. E. Marsh

Lillesand, T. M. and R. W. Kiefer


Limp, W. F.

Lock, G. R. and Z. Stanicic

Lookabill, Anna B.
Lu, D., P. Mausel, E. Brondizio and E. Moran


Ma, Z. and R. L. Redmond

Madry, S. and C. Crumley

Madry, S. and L. Rakos

Maleczewski, J.

Mann, C. C.

Maracci, G., Schmuck, G., Hosgood, B. and Andreoli, G.
Marble, D. F.

Marble, D. F., H. Calkins, and D. Peuguet

Mark, D. M. and A. Frank

Marquardt, W. and C. Crumley

Marsh, G. P.

Martin, P.

Martin, W. A.

Maschner, H. D. G.

Mausel, P., Y. Wu, Y. Li, E. F. Moran and E. S. Brondizio

Maybury-Lewis, D. and J. Bamberger
McCann J.M. and D.W. Meyer  

McNeil, M.  

Medina, J. T., G. Carvajal and H. C. Heaton  

Meentemeyer, V. and E. O. Box  

Meggers, B. J.  


Meggers, B. J. and C. Evans  

Meggers, B. J. and E. Brondizio  

Montieth, J. L.

Mora, S., Herrera, L.F., Cavelier, I., & Rodríguez, C.

Morain, S. A.

Moran, E.
1981 *Developing the Amazon*. Indiana University Press, Bloomington.


Moran, E., E. Brondizio, P. Mausel and Y. Wu

Moran, E., E. Brondizio, J. Tucker, M. da Silva-Forsberg, S. McCracken and I. Falesi

Moran, E. and R. Herrera

Moran, E., A. Packer, E. Brondizio and J. Tucker
Mosteller, F. and J. W. Tukey

Mudar, K.

Muller, J.-P.

Mumford, L.

Myers, T. P.

Myneni, R., S. Maggion, J. Iaquinto, J. Privette, N. Gobron, B. Pinty, D. Kimes, M. Verstraete and D. Williams

Myneni, R. and D. Williams

Nance, C. R., H. Holstein and D. C. Hurst

Neves, E.
2003 Historical and Socio-cultural Origins of Amazonian Dark Earths. In
*Amazonian Dark Earths: Origin, Properties, Management*, edited by J. Lehmann,
Netherlands.

Ngheim, S. V., T. Letoan, J. A. Kong, H. C. Han and M. Borgeaud
1993 Layer Model with Random Spheroidal Scatterers for Remote-Sensing of

Nicholaides, J., D. Bandy, P. Sanchez, J. Villachica, A. Coutu and C. Valverde
1984 Continuous Cropping Potential in the Upper Amazon Basin. In *Frontier
University of Florida Press, Gainesville.

Nicolson, M.
1987 Alexander von Humboldt, Humboldtian Science and the Origins of the

Nunez, M., A. Vikkula and T. Kirkinen
1995 Perceiving Time and Space in an Isostatically Rising Region. In *Archaeology and geographical information systems: a European perspective*,

Nyerges, A. and G. Green
2000 The Ethnography of Landscape: GIS and Remote Sensing in the Study of
Forest Change in West African Guinea Savanna. *American Anthropologist*
102:271-289.

Oliver, J.
2001 The Archaeology of Forest Foraging and Agricultural Production in
Amazonia. In *Unknown Amazon*, edited by C. McEwan, C. Barreto and E. Neves,

Oliveria, A. E.
1994 The Evidence for the Nature of the Process of Indigenous Deculturation
and Destabilization in the Brazilian Amazon in the Last Three-hundred Years. In *Amazonian Indians from Prehistory to the Present: Anthropological Perspectives*,

Olsson, G.
1965 *Distance and Human Interaction: A Review and Bibliography*. Regional
Science Research Institute, Philadelphia.
Orton, C.  

Parker, S.  


Parrington, M.  

Parsons  

Pashley, E. F.  
1966 Structure and Stratigraphy of the Central, Northern And Eastern Parts of the Tucson Basin, Arizona.

Paul, C. and A. Mascarenhas  

Paynter, R.  

Perkins, Philip  

Petersen, J. and M. Heckenberger  
Peuquet, D. J.  

Peuquet, D. J. and D. F. Marble  


Pickles, J.  

Pinty, B. and M. M. Verstraete  

Pires, J. M.  

Pires, J. M. and G. T. Prance  

Plog, F.  

Pope, K. and B. Dahlin  

Porro, A.  
Posey, D. A.


Posey, D. A. and W. L. Balée


Prance, G. T. and T. E. Lovejoy

Price, J. C.

Price, J. C. and W. C. Bausch

Qi, J., Chehbouni, A., Huete, A.R., Kerr, Y.H. and Sorooshian, S.

Qi, J., Moran, M.S., Cabot, F. and Dedieu, G.
Ramenofsky, A. F. and A. Steffen
University of Utah Press, Salt Lake City.

Rapp, G. R.

Rapp, G. R. and C. L. Hill
Yale University Press, New Haven.

Renfrew, C., M. J. Rowlands and B. A. Segraves,

Reynolds, R.

Ribed, P. S. and A. M. Lopez

Richards, J.

Richards, J. A.
Springer-Verlag, Berlin; New York.

Richardson, A. J. and C. L. Wiegand

Rondeaux, G., Steven, M. and Baret, F.

Roosevelt, A. C.


Roosevelt, A. and B. J. Meggers

Roper, D. C.

Rose, M. R. and J. H. Altschul

Ross, E. B.

Rowe, J.

Ruggles, C. and D. Medyckyj-Scott

Ruggles, C., D. Medyckyj-Scott and A. Gruffyd

Saatchi, S. S., J. V. Soares and D. S. Alves

Sabins, F.

Sader, S. A., T. Sever, J. C. Smoot and M. Richards

Saffirio, J. and R. Hames

Salas, B. and J. Brunner

Salthe, S.

Sample, V. A.
Sánchez, P. A.

Sanchez, P. A. and North Carolina Agricultural Experiment Station.

Sauer, C.

Saunders, N.

Savage, S.

Sayce, R.

Schaan, D.


Schalf, R. and T. Lyons

Schmink, M. and C. H. Wood
Schubart, H. and E. Salati

Sebastian, L. and W.J. Judge

Senseman, G. M., C.F. Bagley, S.A. Tweddale

Sever, T.

Shermer, S. J. and J. A. Tiffany

Siljestroem Ribed, P. and A. Moreno Lopez

Singh, A.

Sioli, H.

Skole, D. and C. J. Tucker

Smith, C. A.


Smith, N. J. H.


Sombroek, W. G.
1966 Amazon Soils: A Reconnaissance of the Soils of the Brazilian Amazon Region. Centre for Agricultural Publications and Documentation, Wageningen.


Sponsel, L. E.

Stahl, P. W.

Star, J., Estes, J. and Davis, F.

Steiniger, M.
Stern, P. C., O. R. Young and D. Druckman

Steward, J.

Steward, J. H. and L. C. Faron

Stewart, D. I.

Stine, R. and D. Lanter

Teillet, P. M., K. Staenz and D. J. Williams

Thenkabail, P. S.

Theodoraus, D. and F. LaPena

Thomas, D. H.
Tilley, C. Y.  

Tipps, B. and A. Schroedl  

Tobler, W.  

Tomlinson, R.  

Tosta, N.  

Tottrup, C.  

Townshend, J.  

Trimble, N. L.  
1997  Characterizing Accuracy of the Trimble Pathfinder Mapping Receivers.  

Tucker, C. J.  

Tucker, J. M., E. S. Brondizio and E. F. Moran  

Turkey, J.  
1977  *Exploratory Data Analysis*. Addison-Wesley, Massachusetts.
Turner, B. L.

Turner, B., D. Skole, S. Sanderson, G. Fischer, L. Fresco and R. Leemans (editor)


Turner, M.

Uhl, C.

van Leusen, M.

Verbyla, D. L.

Vinta-Finzi, C. and E. Higgs

Wallace, A. F. C.

Walschburger, T. and von Hildebrand, P.
Walter-Shea, E., J. Privette, D. Cornel, M. Mesarch, and C. Hays
1997 Relations Between Directional Spectral Vegetation Indices and Leaf Area

Wandsnider, L.
1998 Regional Scale Processes and Archaeological Landscape Units. In Unit
University of Utah Press, Salt Lake City.

Warren, P., D. Gordon, W. Philpot and M. Duggin
1989 Field Reflectance Calibration with Grey Standard Reflectors. International

Warren, R.
1990 Predictive Modeling of Archaeological Site Location. In Interpreting
Space: GIS and Archaeology, edited by K. M. S. Allen, S.W. Green and E.

Waters, M.
University of Arizona Press, Tuscon.
1988 The Impact of the Fluvial Process and Landscape Evolution on
Archaeological Sites and Settlement Patterns Along the San Xavier Reach of the

Waters, M. R. and D. D. Kuehn
1996 The geoarchaeology of place: The effect of geological processes on the
preservation and interpretation of the archaeological record. American Antiquity
61(3):483-497.

Weill, L. R.

Wells, D. and N. Beck
Canada.

Whalen, M.
1990 Sampling Versus Full-coverage Survey. In The Archaeology of Regions: A
Case for Full-Coverage Survey, edited by S. K. Fish and S. Kowalewski, pp. 219-
236. Smithsonian Institution Press, Washington, D.C.
Wheatley, D.


Wheatley, D and M. Gillings
2002 *Spatial Technology and Archaeology*. Taylor & Francis, NY.

Whitehead, N.

Whitmore, T. C.

Wildesen, L. E.


Willey, G. R.

Winterhalder, B.

Wise, Alicia L.
Witcher, R. E.

Wobst, H. M.

Wood, C. H. and D. Skole

Wood, J. J.

Wood, T. F. and G. Foody

Woods, W. I.

Woods, W. I., and J. McCann

Wright, D., M. Goodchild, and J. Proctor

Wrigley, N.
Wust, I.

Yanasse, C., S. Sant’Anna, A. Frery, C. Rennó, J. Soares and A. Luckman

Yentsch, A. E.

Zech, W., L. Haumaier and R. Hempfling

Zeidler, J.

Zubrow, E. B. W. and J. W. Harbaugh
BIOGRAPHICAL SKETCH

Joseph C. Russell was born in Gainesville, Florida. After completing his master’s in anthropology at Florida State University, he returned to his hometown to complete his doctoral program at the University of Florida. Not specifically tied to the state of Florida, he is currently enjoying both the weather and the academic climate of the University, with no immediate plans to abandon either.