5-2013

Site Location Modeling and Prehistoric Rock Shelter Selection on the Upper Cumberland Plateau of Tennessee

Lucinda M. Langston
East Tennessee State University

Site Location Modeling and Prehistoric Rock Shelter Selection on the Upper Cumberland Plateau of Tennessee

---

A thesis

presented to

the faculty of the Department of Geosciences

East Tennessee State University

In partial fulfillment

of the requirements for the degree

Master of Science in Geosciences

---

by

Lucinda M. Langston

May 2013

Jim Mead, Ph.D., Chair

Jay Franklin, Ph.D., Co-Chair

Eileen G. Ernenwein, Ph.D., Co-Chair

Ingrid Luffman, M.S.

Keywords: Predictive Modeling, GIS, Prehistory, Rock Shelters, Spatial Logistic Regression

ABSTRACT

Site Location Modeling and Prehistoric Rock Shelter Selection on the Upper Cumberland Plateau of Tennessee

by

Lucinda M. Langston

Using data collected from 2 archaeological surveys of the Upper Cumberland Plateau (UCP), Pogue Creek Gorge and East Obey, a site location model was developed for prehistoric rock shelter occupation in the region. Further, the UCP model was used to explore factors related to differential site selection of rock shelters. Different from traditional approaches such as those that use (aspatial) logistic regression, the UCP model was developed using spatial logistic regression. However, models were also generated using other regression-based approaches in an effort to demonstrate the need for a spatial approach to archaeological site location modeling. Based on the UCP model, proximity to the vegetation zones of Southern Red Oak and Hickory were the most influential factors in prehistoric site selection of rock shelters on the UCP.

DEDICATION

To Archer, because without you I would not have made it this far and because of you, I have the will and motivation to do so much more. You are my everything and my saving grace.

ACKNOWLEDGEMENTS

The development and fruition of this thesis project are due in large part to the contribution of many people both on a personal and professional level. This list is in no way exhaustive as there are many people to whom I owe thanks. For most of you, a simple acknowledgement here is not nearly enough to thank you for your contribution, support and guidance. So please know that I am grateful for all that each of you (and many others too numerous to name) have done for me.

First and foremost, I would like to thank my thesis committee as a whole for their support, tutelage, and patience: Jim Mead, Jay Franklin, Eileen Ernenwein, and Ingrid Luffman. I would like to acknowledge Jim Mead for his dedication to the Department of Geosciences and to his students. I am exceedingly grateful for the opportunity to have worked with and learned from him. In the last two years, Jim has played an integral role in my growth as a scholar, a teacher, and a researcher. I can honestly say that my experience as a graduate student has been a remarkable one and I owe much of that to Jim.

Next, I have to extend a multitude of thanks and my sincere gratitude to Jay Franklin. Not only did he directly contribute to this project by allowing me to work with data that he has dedicated 10 years of his career to collecting, but over the last seven years he has also served as a mentor and a friend--and now I hope that he is proud to call me a colleague. I would like to thank him for all of the opportunities and advice he afforded me as well as all of the knowledge and insight that he has bestowed upon me.

I would also like to thank Eileen Ernenwein for taking my initial thesis project idea and turning it into a multi-faceted project. Eileen introduced me to the concepts of archaeological site location modeling and taught me how to combine my knowledge of GIS and archaeology and use it to generate and address meaningful research questions in a more advanced way than I ever had before. Also, Eileen has become a great friend and mentor in the past two years and I consider myself lucky to have been her student.

The final member of my thesis committee, Ingrid Luffman, also deserves my greatest appreciation for sharing with me her vast knowledge of statistics, predictive modeling, and spatial analysis. Ingrid introduced me to spatial logistic regression and taught me how to use the statistical program R. By way of her knowledge and instruction, I was able to improve my overall methodology and produced a thesis that is truly interdisciplinary. Ingrid is a wonderful educator and I am very grateful for all the time she dedicated to making me a better scholar, writer, and individual.

I would also like to thank the ETSU Honors College, the Department of Sociology and Anthropology, and the Department of Geosciences for their ongoing financial support during both my undergraduate and graduate careers. Additionally, the Tennessee Historical Commission provided funding for the Pogue Creek State Natural Area archaeological survey by way of survey and planning grants. Further, I owe many thanks to the Tennessee Division of Archaeology, especially Mark Norton and Suzanne Hoyal, for providing the archaeological survey and testing permits and for processing all of the state site numbers.

I would especially like to thank the Estate of Bruno Gernt, Inc., and particularly Jerry Gernt, for allowing archaeological investigations on their landholdings and for their continued interest, encouragement, and support. Similarly, I would like to acknowledge the individuals at East Fork Stables for their hospitality and generosity. Without the graciousness of the above individuals, the archaeological visibility of the Upper Cumberland Plateau of Tennessee would be far less than it is today and this project would not have been possible.

I would also like to acknowledge Tennessee State Parks and Forests and, more specifically, Pickett State Park for providing the opportunity to conduct archaeological survey in Pogue Creek State Natural Area and thus gather a portion of the data needed for this project. On a more personal note, I have to thank Alan Wasik, Pickett State Park manager, for providing the resources needed to conduct the Pogue Creek archaeological survey and for his continued generosity and support. I would also like to thank Tennessee State Park employees Travis Bow, Brandon Taylor, and John Froeschauer for their assistance in conducting archaeological survey in and around Pogue Creek. Additionally, numerous others helped with the survey and thus deserve many thanks as well: Alan Cressler, Conor and Miller Franklin, Sierra Bow, Meagan Dennison, Jeff Navel, Jacob Wall, Andrew Dye, Andrew Hyder, Jessica Dalton, Carrie Welch, Sara Warfield, Michael Royston and Mathew Boehm.

I owe countless thanks to my friends who have supported me in this endeavor—especially Heather Davis, Meagan Dennison, Sierra and Travis Bow, and Christina Bolte. I am very grateful for the friendship that each of these individuals has blessed me with. And finally, but most importantly, I have to thank my family for all of their love, financial and moral support, encouragement, and tolerance over the past several years. Without the support system that my immediate family has provided, I would not have been able to get through these past few years with much success. My parents, Michael and Tammy, and my two sisters, Suzanna and Rebekah, are perhaps more to thank than anyone else. And lastly, I have to thank my little one, Archer. Although it will be several years before he can read and understand this, I would like to thank him for giving me big hugs, wet kisses, and a million reasons to smile.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF EQUATIONS

CHAPTER 1

INTRODUCTION

Because of their ability to provide ready-made shelters, rock shelter and cave environments in both the Old World and the Americas have played an important role in contributing to the archaeological record and defining prehistoric sequences (Straus 1990: 255). Compared to open-air sites, these enclosed cavities are important repositories for cultural material and thus provide a perfect opportunity to study culture change (Watson 2001). Worldwide, people have occupied both caves and rock shelters on a short- and long-term basis, yet, they were not uniformly favored for residential occupation and certain attributes influenced differential selection of such sites (Straus 1990: 260). This idea of differential site selection has been the focus of prehistoric settlement studies in archaeology since the 1960s.

In a region where thousands of rock shelters have formed and thus provided instant shelter, prehistoric hunter-gatherers could afford to be more selective in choosing where to locate residential sites. The Upper Cumberland Plateau of Tennessee—hereafter referred to as the UCP—is an example of such a unique landscape. Here, rock shelters are ubiquitous and are a part of both the natural and cultural landscape (Franklin 2002). Decades of archaeological survey conducted on the UCP have resulted in the documentation of more than 400 prehistoric rock shelter sites in the area (Franklin et al. 2013). In many cases shelters where no cultural material was recovered were recorded right next to or in close proximity to sites with cultural material (Franklin 2002; Langston and Franklin 2010; Langston et al. 2012). This raises questions about why certain rock shelters were selected for prehistoric occupation and others were not. The idea of differential selection of rock shelters on the UCP and adjacent regions is not a new concept.

Three studies addressing prehistoric rock shelter selection on or close to the UCP, led to the development of this thesis project (Figure 1). Each of these studies is briefly outlined.

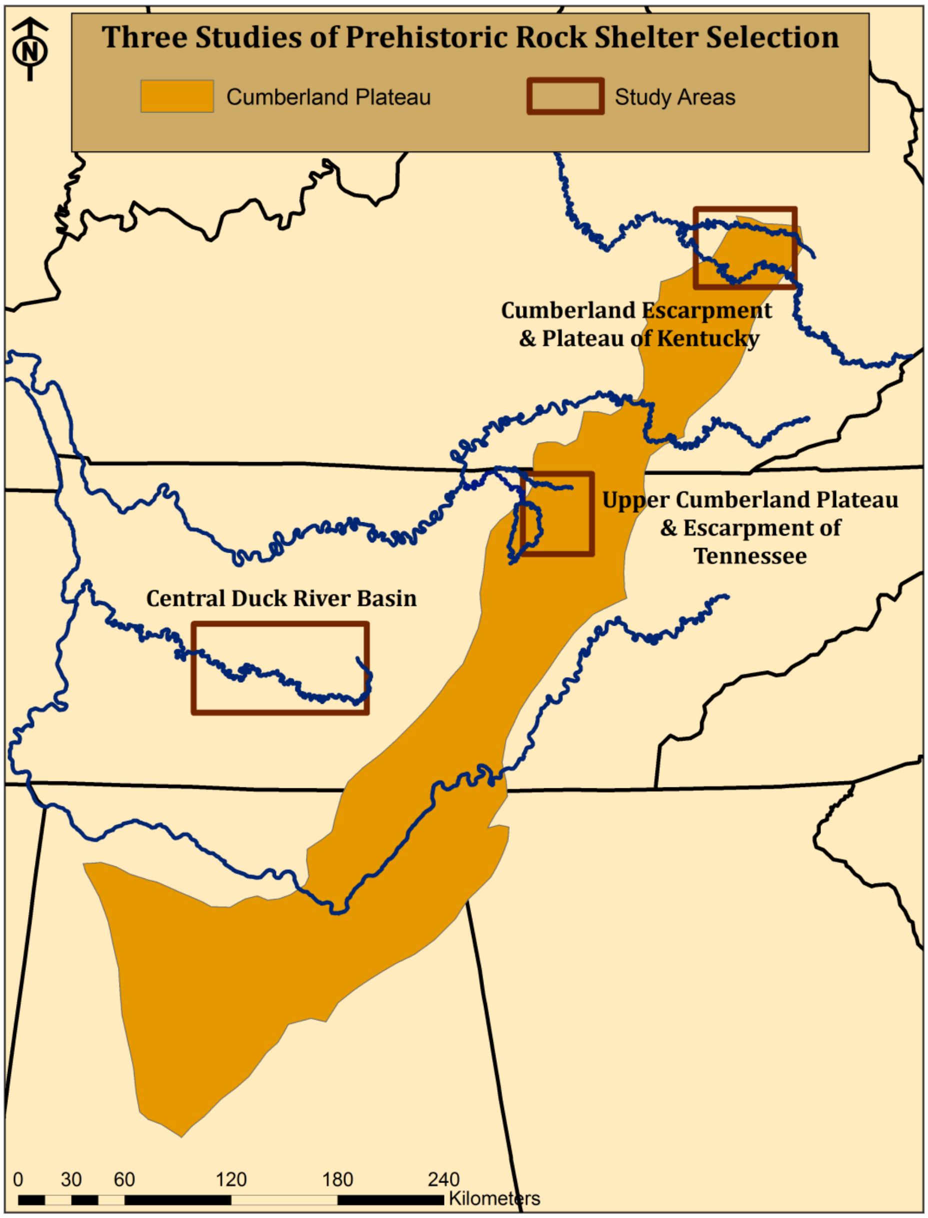


Figure 1: Locations of Three Prehistoric Rock Shelter Selection Studies on the Cumberland Plateau. The Cumberland Plateau is the most southern section of the Appalachian Plateaus Physiographic Province. This map shows the location of the project study area (the UCP) in relation to 2 other studies focusing on rock shelter selection—the Central Duck River Basin (Hall and Klippel 1988) and the Cumberland Escarpment and Plateau of Kentucky (Mickelson 2002).

In a study of rock shelters in the Central Duck River Basin, a lowland area proximal to the UCP, Hall and Klippel (1988) used a polythetic set of determinants including aspect, shelter size, and distance to water, to explain variation in shelter occupation. Using statistical tests and scores to evaluate each factor, Hall and Klippel (1988: 161) argued that shelter desirability was enhanced if the aspect provided protection from prevailing winter winds and/or it admitted abundant sun light; for the Southeastern United States, this suggests more southerly and easterly facing shelters. In addition, availability of water was expected to have affected suitability for prehistoric occupation (Hall and Klippel 1988: 161).

After conducting statistical analysis of 143 rock shelter locations, Hall and Klippel (1988: 168) concluded that shelters with cultural materials tended to have a more southerly orientation than those lacking cultural material. However, contra their assumptions, they found that shelters used prehistorically were further from water sources than those closer to water. A proposed explanation is that prehistoric peoples along the Duck River used shelters as protection from the threat of flooding, seemingly making them special purpose sites (Hall and Klippel 1988: 168).

Closer to this thesis's project area, Mickelson (2002) examined rock shelter distribution in the Cumberland Escarpment and Plateau region of eastern Kentucky; the study area is drained by the Red River, a tributary of the North Fork of the Kentucky River. Mickelson's (2002: 1) approach is based on a hypothetico-deductive method where the null hypothesis states that "…shifts in land use patterns consequent to changes in subsistence practices are not observable". An alternative hypothesis stated that observable fluctuations in space were temporally associated with changes in subsistence practices (Mickelson 2002: 2). In his study, Mickelson (2002: 23) looked at the distribution of rock shelter sites using 5 environmental coverages: elevation, aspect, slope, ecology, and distance to water.

Mickelson (2002: 81) argues that archaeologists often assume that aspect values indicating a more southerly site orientation means that the location receives more solar radiation and is therefore more appealing for occupation. He suggests that in mountainous terrain, south facing shelters might be selected more in order to locate gardens or fields. After analyzing aspect for 319 shelters, although Mickelson (2002: 87) was unable to document trends that he could verify statistically, he states that "Throughout prehistory, there appears to be a trend towards selecting southerly oriented landforms." In addition, he recognizes the problem that many seeps, springs, and small order streams escape being mapped, and therefore distance to water as a factor in shelter selection has not been addressed accurately (Mickelson 2002: 84).

More recently, GIS was used to conduct a preliminary investigation of site selection factors of prehistoric rock shelters on the UCP of Tennessee; factors including depth aspect and straight-line distance to blue-line streams were considered (Langston and Franklin 2010). This study showed that depth aspect was not a factor for rock shelter selection, a finding that distinguishes it from the adjacent lowland Duck River Basin according to Hall and Klippel (1988). Similar to Mickelson (2002), Langston and Franklin (2010) found that straight-line distance to blue-line streams was not a significant factor. This again raises the issue of intermittent and unmapped water sources that GIS analysis alone cannot reveal—especially in karstic regions such as the UCP of Tennessee where many seeps and springs are ubiquitous.

The above studies attempted to "model" or "quantify" patterns of human behavior by analyzing known settlement locations; when studying differential site selection, one is essentially analyzing behavioral practices. Though there are many ways to analyze and interpret prehistoric human behavior, one approach involves the development of site location models. Location, or predictive, models will be discussed further in Chapter 4; it is important to point out, however, that predictive models are not only useful in the context of Cultural Resource Management (CRM) but also for developing and addressing research questions related to differential site selection. By asking questions about where sites are located and why, archaeologists and geoscientists can gain a better understanding of human-land relations as well as human-human interactions within specific environments.

Research Objectives

Although more than 20 years of archaeological survey have been conducted on the Upper Cumberland Plateau of Tennessee, large parts of this region remain to be systematically surveyed (Ferguson et al. 1986; Franklin 2002; Langston and Franklin 2010; Langston et al. 2012). The development of a regional site location model for the UCP would greatly contribute to ongoing and future archaeological surveys of the region by increasing the potential for locating archaeological sites and improving survey methods. Also, the model could be used to investigate environmental and cultural factors that may have been a part of the decision-making process for prehistoric hunter-gatherers in choosing residential locations. Even though a predictive model cannot indicate each and every possible site location, it can increase the chances for locating sites when following basic settlement pattern principles. The theoretical basis is 2-fold: 1) human settlement behavior is non-random and 2) the distribution of resources within a particular environment strongly influences location choices of humans (Verhagen 2007: 13).

Data collected from 2 separate archaeological surveys are used to develop and evaluate a site location model for prehistoric rock shelter occupation on the Upper Cumberland Plateau of Tennessee (Figure 2). These 2 study areas are a good representation of the UCP as a whole because they include high and low altitude landforms at all aspects.

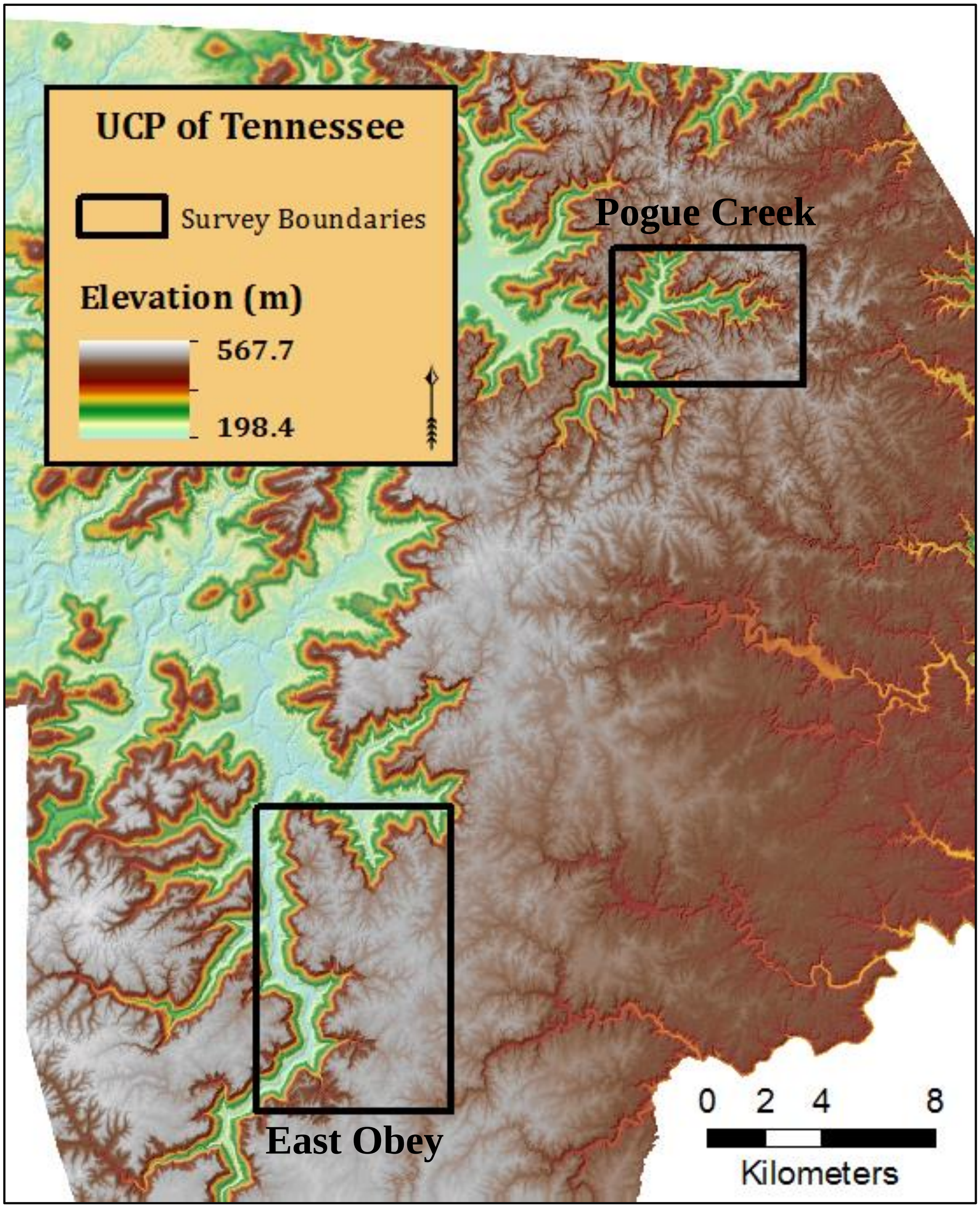


Figure 2: The Upper Cumberland Plateau of Tennessee. Data collected from archaeological surveys of the Pogue Creek State Natural Area and the East Obey are used to develop and test a site location model of prehistoric rock shelters in the region.

This project not only has great utility for practical application (e.g. locating areas with potential to yield archaeological material) but also for addressing specific research questions related to differential site and mobility patterns across prehistory. Though these kinds of research questions have been addressed through excavations on the UCP (Pace and Hays 1991; Franklin et al. 2010; Franklin et al. 2012), this project looks at these questions from a geospatial perspective at the landscape scale.

This research represents an interdisciplinary approach that demonstrates the practical, theoretical, and methodological diversity of archaeological predictive modeling; specific research objectives have been developed to address each of these facets. From these 3 research objectives, this thesis seeks to establish a baseline from which to develop predictive models for future archaeological survey in a way that not only accounts for the practical application but also for the analysis and interpretation of spatial patterning of prehistoric rock shelter selection. The first research objective is to determine if site location data from Pogue Creek and the East Obey can be used to develop and test a predictive model for other areas of the UCP and surrounding region that have yet to be surveyed. However, the primary goal of this research is to learn about prehistoric human spatial behavior and human-land relationships. Thus, the second objective is to use the model variables as a basis for determining what factors may have contributed to differential site selection of rock shelters on the UCP. Because human behavior is not usually the result of random processes, the analysis of such should incorporate methods designed to account for the nature of non-random, spatial relationships. Pertaining to site location modeling, traditional approaches have not addressed the issue of spatial dependence that is present in most archaeological datasets. Therefore, the third and final research objective is to determine if spatial logistic regression can be proposed as an alternative to modeling approaches using traditional (aspatial) statistical analysis.

This thesis focuses on the application of geospatial and statistical analysis in addressing specific archaeological research questions. Thus, the organization of this thesis reflects the interdisciplinary nature of archaeological site location modeling. First, the project area is discussed in terms of its environmental and cultural background. Chapter 2 focuses on the physiographic, topographic, and geologic setting of the UCP. In Chapter 3, summaries of the 4 prehistoric periods of the Southeast are provided with the main emphasis on the UCP of Tennessee specifically.

Following the environmental and cultural settings of the project area, Chapter 4 introduces the background, concepts, and methodological development of archaeological site location modeling. This chapter is divided into 4 sections: (1) a brief history of geographic information systems (GISs) and its applications in archaeology; (2) the development of predictive modeling; (3) an introduction to regression models with the focus on determining the most appropriate modeling method; and (4) a brief discussion on possible site selection factors and common variables used in modeling.

Chapter 5 outlines the methods used to generate the UCP site location model. The model variables (e.g. response and explanatory) are discussed in terms of data acquisition, compilation, and manipulation in a GIS environment. Also, the statistical process of building, running, and generating the model is detailed. All results of the preliminary and final statistical tests are provided in Chapter 6; this chapter also includes the graphical representation of the final UCP model. The final chapter, Chapter 7, includes a detailed discussion of each of the research objectives based on model results. Some general concluding remarks are also provided.

CHAPTER 2

ENVIRONMENTAL SETTING

Physiography

The Appalachian Plateaus is 1 of 7 physiographic provinces within the Appalachian Highlands physiographic region of the eastern United States as defined by Fenneman (1938). The Appalachian Plateaus province extends in an almost linear strip from New York to central Alabama and is further subdivided into 8 sections (Figure 3). Though all of these sections are geologically and topographically different they all consist of degrading plateaus (Fenneman 1938). The project area is situated on the most southern section of the Appalachian Plateaus, the Cumberland Plateau. The Cumberland Plateau is approximately 600 kilometers long and extends from the Kentucky River Drainage in southern Kentucky to the northern boundary of the Gulf Coastal Plain physiographic province in Central Alabama. Further, the Cumberland Plateau is drained by the Tennessee and Kentucky River systems.

For the purpose of this thesis, the Upper Cumberland Plateau of Tennessee corresponds to the parts of Fentress and Pickett Counties within the Mid and Northern sections of the Cumberland Plateau region, or as Hinkle (1989) refers to it, the Central Uplands of the Cumberland Plateau. Here, the elevation can range from approximately 900 feet above mean sea level in the floodplain to more than 1700 feet above the rim of the gorge. This region is generally characterized by rugged topography with steep sideslopes, and narrow to moderately broad valleys (Smalley 1986). More specifically, the western escarpment of the plateau is highly irregular with many incisions cut by westward draining streams (Sasowsky 1992: 5). The irregular topography is mostly a result of erosion of the horizontal and slightly dipping strata.

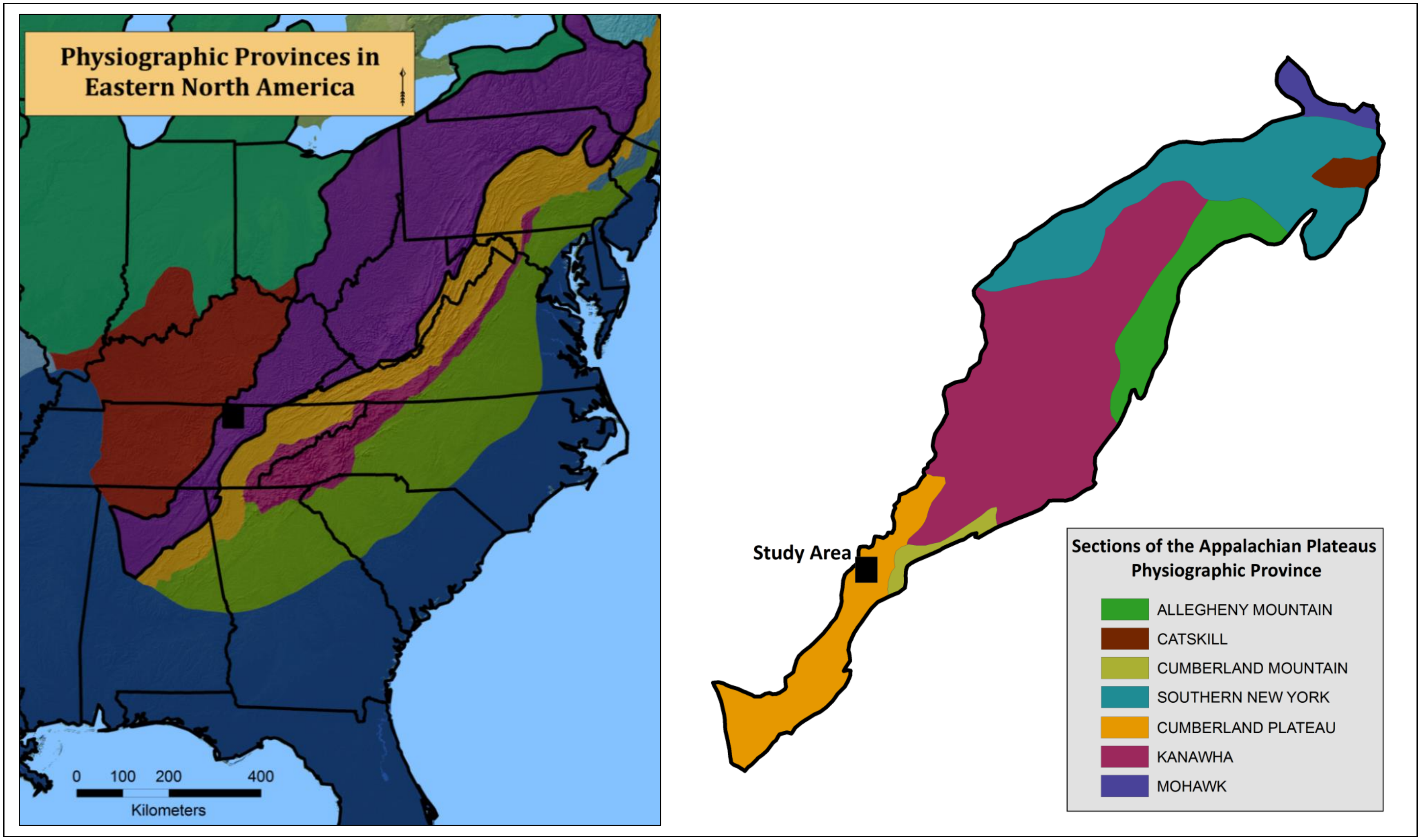


Figure 3: Appalachian Plateaus Physiographic Province. The study area—the Upper Cumberland Plateau of Tennessee--falls within the Cumberland Plateau section (right) of the Appalachian Plateaus physiographic province (left).

Differential weathering and erosion of the caprock has led to the formation of thousands of rock shelters in and around the gorges and on the plateau surface. Further, the UCP has a complex hydrology with an active underground drainage system; over millions of years, groundwater has eroded the softer rock strata beneath the more resistant sandstone caprock creating complex subsurface conduits and cave passages (Sasowsky 1992: 4).

Geology

Rocks forming the Cumberland Plateau were formed during the Upper Paleozoic from the deposition of marine and continental sedimentary deposits; the bedrock geology of the region includes Mississippian, Pennsylvania, and Permian-aged units (Hunt 1967: 19). In the UCP of Tennessee, thick, nearly continuous Pennsylvanian units lie almost completely horizontal atop Mississippian limestone, dolomite, sandstone, and shale (Smalley 1986). The caprock of the western escarpment on the UCP includes sandstones, shales, and conglomeratic units from the Crab Orchard Moutain and Gizzard Groups. The Mississippian limestone, dolomite, and shales form the less-resistant rock strata beneath the thick caprock. The oldest rocks exposed in the study area are within the Mississippian-aged Ft. Payne Formation which is composed of mixed shale, siltstone, and limestone. The Ft. Payne Formation is overlain by 200 meters of upper Mississippian formations that include the St. Louis Limestone & Warsaw Limestone, the Monteagle Limestone, the Bangor Limestone and Hartselle Formation, and the Pennington Formation.

In the mid to upper slopes, sandstone rock shelters dominate the landscape. However, caves and rock shelters can be found in some lower slopes and valley bottoms where deeply incised streams have eroded away the sandstone caprock into the underlying Mississippian-aged limestone.

Climate and Vegetation

The climate of the UCP is classified as a humid mesothermal though precipitation and temperature can vary locally based on topography (Thornthwaite 1948). Smalley (1986) describes the temperature of the region as long, moderately hot summers with short, mild to moderately cold winters. The region is generally humid with no distinct dry season and precipitation is well distributed throughout the year (Hart 2007: 35). However, the Cumberland Plateau is slightly wetter than the adjacent physiographic sections due to orographic precipitation.

The vegetation of the Cumberland Plateau section falls with Hunt's (1967) broad classification of a Central Hardwood Forest where mixed *Quercus* (oak) and *Castanea* (chestnut) species and *Liriodendron tulipifera* (yellow poplar) are the dominant tree types (Hunt 1967: 102). At a more regional scale, the UCP is characterized as a Mixed Mesophytic Forest region according to Braun's (1950) forest classification system. Common canopy species identified by Braun (1950) include *Quercus rubra* (red oak), *Quercus alba* (white oak), *Carya sp.* (hickory), *Tsuga canadensis* (hemlock), *Acer saccharum* (sugar maple), *Liriodendron tulipifera* (yellow poplar), *Tilia heterophylla* (white basswood), *Aesculus flava* (sweet buckeye), and *Castanea dentata* (American chestnut). Because Braun's (1950) forest classification system follows closely to Fenneman's (1938) physiographic provinces, the classifications are regionally-based and do not necessarily account for local variation. The topography of the Cumberland Plateau is highly variable; therefore, different forest communities exist within the region because forest composition is directly related to slope, aspect, and landform (Hinkle et al. 1993).

Following Braun's (1950) work, numerous vegetation studies were conducted on the Tennessee portion of the Cumberland Plateau in an effort to contribute to the knowledge and understanding of forest communities in the region. One such example includes Hinkle's (1989) summary of his dissertation work on the Cumberland Plateau (see Hinkle 1978) in which he classifies the vegetation of the region into 2 categories: the plateau uplands and the ravines and gorges. According to Hinkle (1989), the upland stands include (but are not limited to) *Acer rubrum* (red maple), *Betula nigra* (river birch), *Ilex opaca* (holly), *Quercus alba* (white oak), *Nyssa sylvatica* (black gum), and *Pinus virginiana* (Virginia Pine); slopes, however, are generally dominated by mixed *Quercus* species with White Oak being the most frequent (Hinkle 1989: 124–125). In contrast, the ravines and gorges (more characteristic of the western escarpment) are dominated by mixed *Quercus* species (e.g. *Q. alba*, *Q. prinus*, *Q. rubra*, and *Q. velutina*) at all slopes but with *Acer saccharum* stands at middle and lower slopes; *Tsuga canadensis* is mainly restricted to headwaters and along bedrock streams (Hinkle 1989: 125).

When comparing Hinkle's (1989) vegetation communities to Braun's (1950) classification, the ravines and gorges forest types were more representative of a Mixed Mesophytic Forest region than the upland communities. However, Hinkle (1989: 128) points out that many of the Mixed Mesophytic indicator species were secondary to oak and hickory species in the ravine and gorge areas. It is important to note here that the UCP has been subject to a long history of anthropogenic fires and exploitation of the landscape through mining and logging--these activities have most definitely altered (and continue to do so today) the composition of forest communities in the region.

CHAPTER 3

CULTURE HISTORY

This chapter provides a basic outline of the culture history of the UCP, though general trends for the Southeast as a whole are also included. Archaeological information and diagnostic artifacts recovered during archaeological survey and stratified excavations from rock shelter sites in the study area are used to discuss the 4 prehistoric cultural periods of the Southeast-- the Paleoindian, Archaic, Woodland, and Mississippian—although occupation of the UCP is best represented by the Archaic and Woodland cultures. All dates are presented using B.P. except for dates beginning Anno Domini (A.D.). However, when reporting specific dates from ceramics or other dated artifacts the date will be presented in B.C. or A.D. along with its error margin.

Paleoindian

Though there is still much debate about when the first Americans reached the Southeast, the most recent studies postulate that people arrived in the area sometime around or after the last glacial maximum at approximately 21,000 B.P. (Anderson and Sassaman 2012: 36). How these first peoples arrived—in both the Americas and the Southeast—has also been the focus of much debate over the years.

The Paleoindian period, refers to cultures older than approximately 10,000 B.P., which marks the transition into the latter Archaic period. Generally speaking, Paleoindians are characterized as highly mobile bands that engaged in periodic multi-band aggregation important for forming and maintaining networks and reinforcing social ties (Anderson and Sassaman 2012: 52). Evidence for hunting of extinct species such as mastodon at Kimmswick, MO (Graham et al. 1981) and Coates-Hines, TN (Breitburg et al. 1996) demonstrates that the earliest humans were initially big game hunters and gatherers.

The first unequivocal evidence for settlement of the Southeast dates back about 13,000 years ago and is marked by the appearance of a "readily identifiable diagnostic artifact category" (Anderson and Sassaman 2012: 47). Clovis points, commonly believed to be the earliest fluted projectile point type, are found all across the Southeast, though mostly as isolated finds not associated with other artifacts. Broad geographic trends and variation noted among Clovis points possibly represents either drifts in cultural transmission or temporal differences due to the movement and isolation of Clovis populations.

Around the onset of the Younger Dryas at approximately 12,800 B.P., new projectile point types appeared as the Clovis horizon comes to an end. The Late Paleoindian, or post-Clovis, saw broad changes in projectile point styles that occurred differentially in the Southeast. In Tennessee, examples of early fully fluted projectile points are Cumberland and Redstone. Later unfluted forms include Beaver Lake, Quad, and Dalton. As the Younger Dryas persisted, Dalton points and their subtypes (distinct geographic varieties) became quite common. It is possible that these changes in technology reflect the major changes in climate and biotic communities that occurred with the rapid cooling of the Younger Dryas (Anderson and Sassaman 2012: 58). As many large animals were becoming extinct, populations had to expand their diet to include more small game and plant foods. By the late Paleoindian, a wide-range of floral and faunal species was exploited and diverse subsistence strategies had been adopted. White-tailed deer, migratory birds, fish, and fruit and nut mast are some of the more common examples of the Late Paleoindian diet.

Evidence of Paleoindian occupation on the UCP is quite sparse. During his survey of the East Obey on the UCP, Franklin (2002) recorded 7 Paleoindian sites, all of which were based on surface finds in rock shelters. Clovis points were recovered at 2 sites, indicating an Early Paleoindian occupation; the Late Paleoindian was also represented by the presence of Beaver Lake, Quad, and Dalton projectile points (Franklin 2002: 215). More recently, a Late Paleoindian projectile point base was recovered from Red Spear Rock Shelter during archaeological survey of the Pogue Creek State Natural Area (Langston and Franklin 2010). On other parts of the UCP, Late Paleoindian artifacts have only been recovered from unprovenienced locations (Des Jean and Benthall 1994). So in total, at least 8 Paleoindian sites have been documented on the UCP of Tennessee. Franklin (2002: 215) suggests that comprehensive Paleoindian surveys similar to one conducted by Broster et al. (1996) would go a long way in helping to locate and document Paleoindian occupation in the region. Similarly, Anderson and Sassaman (2012: 65) point out that more work is needed to refine, and in a lot of cases, define Paleoindian culture sequences in the Southeast as a whole. Ongoing (and new) excavations at sites with Paleoindian components as well as examinations and analysis of assemblages have and are continuing to generate information on settlement patterns, subsistence strategies, and technological variations of the earliest Americans.

Archaic

Roughly coinciding with the Pleistocene/Holocene boundary, the Archaic (ca. 10,000-3000 B.P.) is the longest prehistoric period. Similar to their predecessors, Archaic peoples are generally defined as mobile groups of hunter-gatherers living in small bands that often aggregated throughout the year. Archaic diets consisted mainly of wild plant and animal foods. Also, some plant resources that were later domesticated were being intensively collected at this time (Anderson 2001: 157). Although populations were fairly high, evidence for residential structures is very limited. However, cultural features containing hearths, rock clusters, grinding slabs, and shallow pits have been discovered. These features were mainly used for either food preparation or cache pits. This may suggest that people lived in lightly constructed shelters rather than larger dwellings or that these settlements were smaller seasonal camps. Also, the first extensive use of cave and rock shelter sites is noted during this time, signifying changes in land use (Anderson and Sassaman 2012: 71). On the UCP specifically, evidence of Archaic Period cultures has been discovered in thousands of rock shelters possibly representing more seasonal occupation (Des Jean and Benthall 1994: 120).

The Archaic Period is commonly divided into 3 sub-periods: the Early Archaic (10,000-7500 B.P.), the Middle Archaic (7500-5000 B.P.), and the Late Archaic (5000-3000 B.P.). The divisions of the Archaic are as much based on climatic and environmental changes as on shifts in subsistence and technology.

Early Archaic

The beginning of the Archaic and thus the Early Archaic is marked by a sharp increase in global temperatures brought on by the onset of the Holocene Era. The early Holocene was warmer than the Pleistocene, though temperatures were still cooler and the overall climate still more humid than today. Many of the megafauna extinctions are believed to occurred during this time, possibly due to the warming climate or over-hunting by Paleoindians. Also, oak and hickory forests were gradually replacing grasslands and savannahs in the Southeast, causing major adaptations by prehistoric peoples (Delcourt and Delcourt 1987). Despite dramatic changes in the environment, much of the chipped stone tool assemblage of the Early Archaic was similar to that of Paleoindian times with some differences, however. Successive side- and corner-notched and bifurcate-based hafted bifaces characterize Early Archaic occupations (Anderson and Sassaman 2012: 72). Side-notched points including Big Sandy, Cache River, and Hardaway possibly extended from earlier Dalton forms. Following next was a sequence of corner-notched points (e.g. Kirk and Charleston) and bifurcate based points (e.g. St. Albans and LeCroy). The most obvious shift in tool technologies between the late Paleoindian and Early Archaic was the "gradual replacement of trianguloid endscrapers with a more varied (and less standardized) set of scraper forms" (Steponaitis 1986: 370–371). Other stone tools made during this time were mullers, grinding slabs, pitted cobbles, and polished slate celts. The formal toolkit of elaborately made scraping, cutting, and piercing stone tools was gradually replaced by a more expedient toolkit as lower quality raw materials were increasing used for manufacture (Anderson 2001: 157)

Through archaeological surveys on the UCP, the Early Archaic has been documented at at least 39 sites, almost all of which are rock shelters (Ferguson et al. 1986; Franklin 2002; Langston and Franklin 2010). This period is well-represented by the presence of side-notched (e.g. Big Sandy I), corner-notched (e.g. Kirk, Lost Lake, and Pine Tree), bifurcates (e.g. MacCorkle, St. Albans, and Lecroy), and (later) stemmed (e.g. Kirk Stemmed/Serrated) varieties (Franklin 2002: 216). The low numbers of the stemmed varieties possibly indicates the movement of peoples out and away from the UCP around the beginning of the Middle Archaic (Franklin 2002: 216–217)(Franklin 2002: 216).

The Early Archaic of the UCP has also been documented in stratigraphic context. Excavations at *Early Times Rock Shelter* revealed stratified Early and Late Archaic deposits. A late Paleoindian Quad biface was also recovered during general surface collection (Dye, Franklin, and Hays 2011).Two Early Archaic bifaces, a Lecroy and a MacCorkle Stemmed, were recovered in good stratigraphic context and an Early Archaic Kirk Stemmed biface was recovered during general surface collection (Dye, Franklin, and Hays 2011). Use-wear analysis revealed that both the late Paleoindian Quad biface and the Early Archaic MacCorkle Stemmed biface bore evidence of wood working (Dye et al. 2011:8). Further analysis of the tool assemblage and lithic material indicated that *Early Times Rock Shelter* served as a short-term situational camp for small task groups of Archaic hunter-gatherers (Dye et al. 2011). Both stratified excavations and survey data corroborate the occupation of the UCP during the Early Archaic. Though open-air ridge-top and terrace sites are not completely uncommon on the UCP, Early Archaic peoples seem to have favored rock shelter environments (Des Jean and Benthall 1994: 120; Franklin 2002: 217).

Middle Archaic

The Middle Archaic is marked by the beginning of the Hypsithermal, a Mid-Holocene climatic interval, when seasonal extremes in precipitation and temperature were greater than today (Anderson 2001: 158; Anderson and Sassaman 2012: 73). In the Midsouth, the Mid-Holocene climate was hotter and dryer than present conditions leading to reduced vegetation in upland environments. Delcourt and Delcourt (1987) also suggest a replacement of oak by the re-expanding pine forests. It has been postulated that the subsequent warming and drying trends made riverine and coastal areas more favorable for human occupation while the upland areas became less favorable (Brown and Vierra 1983; Brown 1985; Dye 1996). Whether a result of the changing climate or some other factors, the number of Middle Archaic sites is believed to be generally lower than in the Early Archaic. However, the distribution of Middle Archaic sites significantly varies throughout the Southeast and not all areas have a lower site density.

Middle Archaic subsistence patterns are similar to those of the Early Archaic with 2 notable additions: (1) the accumulation of shell middens dating to the Middle Archaic reflects an intensive exploitation of fresh water riverine resources (Griffin 1967: 178); and (2) curcurbit remains recovered at the Anderson Site (Dowd 1989) indicate the beginnings of plant domestication and horticulture in the Middle Archaic.

Overall, the Middle Archaic tool assemblage is characterized by the introduction of a stemmed biface technology believed to be derived from Early Archaic traditions (Anderson and Sassaman 2012: 73). In their study, Des Jean and Benthall (1994: 127) recognize Middle Archaic occupation of the UCP based on the recovery of lithics from Stanly, Big Sandy II, Morrow Mountain, and Guilford phases. In other parts of the UCP around the Obey River Drainage, Franklin (2002: 205) recovered Middle Archaic artifacts from 7 sites with tools representative of the Sykes/White Springs, Stanley Stemmed, and Eva clusters. However, in a more recent survey conducted on the UCP, no obvious Middle Archaic sites were recorded (Langston and Franklin 2010).

On the Cumberland Plateau, Des Jean and Benthall (1994: 123) note a decline in prehistoric population during the Middle Archaic based on the paucity of diagnostic materials. Franklin (2002: 212) also notes the lack of diagnostic Middle Archaic artifacts recovered from the region. However, based on radiocarbon assays attained from the UCP there appears to be a spike in Middle Archaic occupation around 5000 B.P. (Franklin 2002: 212). This does not support a general abandonment of the region during the Middle Archaic as is commonly believed. Langston and Franklin (2010) posit that the discrepancy between the artifactual and radiocarbon data highlights the dangers of interpreting prehistoric cultural components based on surface collections and so-called diagnostic artifacts; artifacts recovered from surficial and disturbed contexts may have been misidentified in certain cases. Franklin (2002: 218–219) suggests that more stratified excavations are needed to sort out and understand Middle Archaic occupation of the UCP.

Late Archaic

Around 5000 B.P. at the apex of the Hypsithermal, the climate began to stabilize and by 4000 B.P. conditions closely resembled those of today; the more stable environment provided support for large-scale, sustained occupation (Sassaman 2010: 23). Steponaitis (1986: 373) lists 4 trends that characterize the Late Archaic of the Southeastern United States: (1) the addition of cultivated plants to the diet; (2) the intensification of long-distance exchange networks; (3) the appearance of large, dense middens; and (4) the first use of containers and storage pits.

The increased importance of gathering wild and native plant foods led to an increase in sedentism in many areas during the Late Archaic; these shifts in subsistence and settlement patterns further facilitated the development and use of containers (Smith 1986). Some of the earliest container/vessel forms were made from modified gourds or carved out of steatite (soapstone) quarries. The earliest (clay) pottery vessels were tempered with vegetable (fiber) matter and made into bowls or pans (Steponaitis 1986: 373–374). More than likely, these early containers were used for processing, cooking, and/or storage purposes.

By the Late Archaic, a significant population increase and use of the UCP is evidenced by the increasing numbers of recorded components when compared to previous periods (Franklin 2002: 219; Langston and Franklin 2010). The tool assemblage of the Late Archaic on the UCP is quite diverse with numerous artifact types well-represented in the area (Franklin 2002: 219). The most commonly recovered Late Archaic artifact types are assymetrical/undifferentiated stemmed bifaces (e.g. Ledbetter and Iddins); Other Late Archaic biface types identified on the UCP include Damron, Perkiomen, Merom, and Saratoga (Franklin 2002: 219–220). Though more typical of the Middle-to-Late Archaic in the Kentucky, Ohio, and Illinois valleys, Matanzas bifaces are also a prevalent artifact type recovered on the UCP of Tennessee. Franklin (2002: 220) states that this “suggests frequent cultural interactions between the UCP of Tennessee and regions to the north.”

The very Late Archaic is represented on the UCP based on the high numbers of recovered Wade bifaces. Other very Late Archaic types include Adena Stemmed, Motley, Little Bear Creek, Brewerton, and Turkey-tail (Franklin 2002: 220–221). Some instances of exotic chert use (e.g. Burlington Chert from eastern Missouri and western Illinois) further supports Franklin’s (2002: 220) assertion of interaction between the UCP of Tennessee and cultures to the north.

The Late Archaic culture has also been identified through controlled stratigraphic excavations at rock shelters sites on the UCP of Tennessee. The previously discussed excavations at *Early Times Rock Shelter* also revealed a Late Archaic occupation; this is represented by the recovery of 2 diagnostic bifaces made from different chert types—a Table Rock or Cotaco Creek Cluster biface made from St. Louis chert and 1 asymmetrical stemmed type made from Monteagle chert (Dye et al. 2011). The entire lithic assemblage of *Early Times Rock Shelter* was analyzed in an effort to identify what types of activities were conducted on site. According to Magne’s (1989) approach, a lithic assemblage can indicate 4 different types of sites: a high number and greater diversity of tools but with low percentages of late stage debitage indicates a residential location; fewer tools, lower diversity, and low late stage flaking debris indicates a manufacturing site; a situational “emergency” camp is represented by fewer tools, low diversity, and higher late stage flaking debris; and a large number of tools with relatively high diversity and higher percentages of late stage flaking debris indicates a repeated logistical camp (Magne 1989; Dye et al. 2011). According to the lithic analysis at *Early Times Rock Shelter*, this site was a situational camp that was used as a temporary special purpose site where locally procured nodules of chert were reduced and occasionally, tools were produced and resharpened (Dye et al. 2011).

Within a few to several kilometers of *Early Times Rock Shelter* is *3rd Unnamed Cave*, a primary Monteagle Chert source location that was exploited by Late Archaic peoples (Franklin 1999, 2001; Franklin and Simek 2008; Simek et al., 1998). Only 2 stone tools were recovered from this site and late stage debitage made up less than 2% of the lithic assemblage—this coupled with the underground chert source strongly suggests that *3rd Unnamed Cave* was a quarry and manufacturing location. This clearly indicates that Late Archaic peoples were logistically mobile and exploiting their local resources (Franklin 1999, 2001; Franklin and Simek 2008; Simek et al. 1998).

During the Late Archaic, rock shelters were not only used as short-term, special purpose sites, but also as long-term repeated camps sites. Preliminary interpretations of archaeological testing at *Sachsen Cave Shelter* indicate repeated use of the site as a "residential base camp for small family groups over a long period of time" (Franklin et al. 2010: 447). Several lines of analysis (e.g. technological, use-wear, faunal, and archaeobotanical) indicate that multiple activities such as butchering, cooking, processing hides, nut processing, and wood working were conducted on site throughout the year.

Residential occupation of rock shelters on the UCP during the Late Archaic is evident from 4 summer excavations at *Eagle Drink Bluff Shelter*. Diagnostic artifacts recovered from the site along with radiometric age measures indicates an intermittent occupation of *Eagle Drink Bluff Shelter* from the Middle Archaic to the late Middle Woodland; the Late to Terminal

Archaic, however, appears to represent the most intensive occupation (Franklin 2008: 93; Franklin et al. 2012). Terminal Archaic Wade bifaces, Adena Bifaces, and steatite vessel fragments were recovered during excavation and sometimes in the same context as fabric-marked and cord-marked ceramics (Franklin et al. 2013). These associations demonstrate the difficulty in differentiating between the Late and Terminal Archaic and Early Woodland based on the presence of pottery alone.

Archaeological survey data coupled with recent excavations have provided a baseline from which the Late Archaic occupation/use of the UCP can be better understood. It is clear, however, that by the Late Archaic, hunters and gathered were intensely occupying the UCP. Further, Franklin (2002; 2006) and Dye et al. (2011) have hypothesized that by the Late Archaic, prehistoric peoples were using and occupying the UCP year round though shelters were possibly used for different purposes ranging from residential to logistical to situational. This is different from earlier periods where occupation of the UCP may have been more seasonally based. The recovery of steatite vessel sherds from *Sachsen Cave Shelter* and *Eagle Drink Bluff Shelter* indicates the existence of extensive trade networks—something that continues on into the Early Woodland (Franklin 2008; Franklin et al. 2010). Also, although pottery becomes a wholesale addition in the Early Woodland, recognizable Early Woodland pottery types have revealed dates coinciding with the Late and Terminal Archaic. One example is a sooted cross-mended Early Woodland Swannanoa vessel recovered from a rock shelter in Scott County that was dated to almost 3,000 B.P. (Franklin 2008: 95–96; Franklin et al. 2013). Lastly, Late Archaic peoples were both logistically and residentially mobile and were not constrained by the rugged terrain of the UCP, but instead were taking full advantage of its natural resources (Franklin 1991, 2001; Franklin and Simek 2008; Simek et al. 1998; Franklin et al. 2010; Dye et al. 2011).

Woodland

The Woodland Period (ca. 3000 B.P. – A.D. 900) is seen as a time of gradual change and an era of regionalism building on trends that first emerged in the Late Archaic (Steponaitis 1986: 378; Anderson and Sassaman 2012: 112). Distinct traditions evolved differentially throughout the Southeast during the Woodland Period though some broad trends have been proposed for the Southeast as a whole. Four major trends identified for the Woodland include the increasing importance of seeds for dietary purposes, increased sedentism, more elaborate mortuary rituals and burial mound complexes, and the widespread manufacture and use of pottery (Smith 1986; Steponaitis 1986; Jefferies 2004). Similar trends have been proposed by Chapman (1985) with the additions of bow and arrow technology and the rise of social stratification. The Woodland is typically divided into Early, Middle, and Late sub-periods.

Woodland peoples were broad-based hunter-gatherers who exploited the rich habitat diversity of coastal zones along the southern Atlantic and interior river valleys of the Southeast. Along the coast, these peoples represented a harvesting adaptation to marsh and swamp ecosystems with the addition of garden plots of squash and gourd (Smith 1986: 37–38). Small and medium sized semi-permanent to permanent villages occupied the interior riverine Southeast. Smith (1986: 39–41) notes that around these regions there was substantial house construction and simple “down-the-line” exchange networks. Also, the numerous cylindrical storage pits discovered indicate a more heavy reliance on nuts such as acorn, hickory, chestnut, and walnut (Smith 1986: 42).

Although ceramic technology had its origin in the Archaic, it was during the Woodland Period that pottery became a wholesale addition. Plant fibers as tempering agents were replaced with new tempering inclusions such as quartz, sand, grit, and limestone. In addition, twine and wooden paddles were used to decorate the clay-fired vessels. Common surface treatments include cord- and fabric-marked impressions. Because ceramics are both regionally and chronologically sensitive, archaeologists commonly use ceramic “phases” to identify and delineate cultural groups from the Early Woodland on instead of using projectile point types. However, the issue of delineating between Woodland ceramic phases has been a re-occurring theme in Southeastern archaeology (Faulkner 1968; Schroedl and Boyd, Jr 1991) and more specifically, on the UCP (Franklin and Bow 2008; Franklin et al. 2013). Because of this, the Upper Cumberland Plateau Archaeological Luminescence Dating Project was initiated in 2007 under the auspice that ceramics found in rock shelter contexts could be directly dated when there is no associated archaeological carbon (Franklin 2008a, Franklin and Bow 2008, 2009; Bow and Franklin 2009). This method is referred to as blue light optically stimulated luminescence (BOSL) dating and has been used to date pottery sherds collected during archaeological survey and stratigraphic excavations on the UCP (Wall 2013). Luminescence dates from controlled stratigraphic excavations are used to frame the ones recovered during archaeological surveys; thus far, results from stratigraphic and survey contexts have been consistent (Franklin 2008a; Franklin and Bow 2009). Twenty-two BOSL dates have been returned on pottery sherds recovered during the archaeological survey of the Pogue Creek State Natural Area (Franklin et al. 2013). Some of these dates are used to discuss Woodland occupation of the UCP below.

During the Pogue Creek Archaeological survey, 48% of sites where diagnostic artifacts were recovered indicated a Woodland occupation--clearly, Woodland peoples maintained a significant presence in the Pogue Creek area (Langston and Franklin 2010). This is similar to Franklin’s (2002:204) findings where the “Woodland Period appears to have been the time of most intensive use of the UCP”. However, the Early Woodland appears to be slightly less represented than the Middle and Late Woodland in Pogue Creek and other portions of the UCP as compared to the Big South Fork Area where occupation appears to drop off after the Early Woodland (Ferguson et al. 1986:93; Franklin 2002:204-207; Langston and Franklin 2010).

Early Woodland

The Early Woodland (ca. 2700 B.P. – A.D. 200) is represented on the UCP by diagnostic artifacts—ceramics and tools—recovered from archaeological survey and excavations. Early Woodland ceramics recovered on the UCP are generally typical of Early Woodland pottery (Franklin 2002; Franklin et al. 2013). The Early Woodland of the UCP includes largely grit and/or quartz tempered vessels that are either cord-marked or plain with limestone-tempered fabric-marked varieties increasing in number towards the eastern portion of the UCP (Franklin 2002: 223–226). Cord-marking appears to be the preferred method of surface treatment for the Early Woodland of the UCP, though fabric-marked and plain varieties have been recovered (Franklin 2002; Wall 2013). On the UCP, tools diagnostic of the Early Woodland include varieties of stemless triangular bifaces such as Greeneville (Lewis and Kneberg 1957) and McFarland (Faulkner 1988) types. Interregional interaction on the UCP during the Early Woodland is evidenced by the presence of the aforementioned Swannanoa vessel from nearby Scott County (Franklin 2008a: 95–96; Franklin et al. 2013) and by the recovery of 6 deeply cord-marked and incised limestone tempered body sherds from *Tevepaugh Rock Shelter* that are reminiscent of types from southern Illinois (Franklin 2002:42, 230; Franklin et al. 2013).

Several radiometric age determinations from sites such as *Eagle Drink Bluff Shelter*, *3rd Unnamed Cave*, *Pemberton Rock Shelter*, and *Calf Rock Cave* have indicated an intermittent but continuous occupation of the UCP during the Early Woodland Period (Franklin 2008a). Also, BOSL dates from Early Woodland ceramics have provided a wide temporal range for the period from circa 3150 B.P. to A.D. 600 (Franklin et al. 2013); these dates reveal overlap between the Early and Middle Woodland periods and further support Schroedl and Boyd's (1991:77-78, 85) assertion of the continuity of material culture between A.D. 400 and 900. A very early BOSL date of 1234 ± 339 B.C. (so possibly predating most of the other Early Woodland sherds that have been dated) was returned for a limestone tempered fabric-marked sherd from *Red Velvet Spider Rock Shelter*; this sherd is almost identical to one from *Eagle Drink Bluff Shelter* which returned a BOSL date of B.C. 1218 ± 115 (Franklin 2007; Franklin et al. 2013). Two other limestone tempered fabric marked sherds, recovered from *Gwinn Cove Rock Shelter* and *No Quarter Rock Shelter*, returned BOSL dates of A.D. 79 ± 209 (Wall 2013) and A.D. 648 ± 134 (Franklin et al. 2013), respectively. All of these dates combined demonstrate the persistence of this specific ceramic type for over a thousand years. Lastly, a quartz tempered fabric marked sherd, a ceramic type that usually precedes limestone tempered fabric marked wares in the adjacent Ridge and Valley, was recovered during excavation at *Hemlock Falls Rock House* returned a BOSL date of A.D. 552 ± 132 (Franklin et al. 2013).

Middle Woodland

During the Middle Woodland (ca. A.D 200-800), cord-marking continues to be the most common surface treatment found in UCP ceramic assemblages. Limestone tempered cord-marked wares account for almost 75% of the Middle Woodland assemblages on the UCP with limestone tempered plain wares accounting for almost all of the remaining 25% (Franklin 2006). Some simple stamped and check stamped varieties have also been recovered on the UCP (Franklin 2002: 229). For stone tools, McFarland bifaces (Faulkner 1988) continue into the Middle Woodland from the earlier period with the addition of types belonging to the Lowe Cluster (Justice 1987) of expanding stemmed bifaces and Copena types (Franklin 2002; Franklin and Bow 2009; Franklin et al. 2013). Intensive occupation of the UCP during the Middle Woodland is evident from excavations conducted at *York Palace* (Langston et al. 2010), *Hemlock Falls Rock House* (Dye et al. 2010), and *Indian Rock House* (Franklin et al. 2013).

The ceramic assemblage of *York Palace* includes mostly limestone tempered wares where cord marking is seemingly the most common surface treatment; 2 BOSL dates of A.D. 562 ± 84 and A.D. 498 ± 50 place this type in the Middle Woodland (Langston et al. 2010; Franklin et al. 2013). Some limestone tempered check-stamped wares were also recovered and are believed to be mostly from the same vessel; 1 sherd was BOSL dated and returned a date of A.D. 720 ±35 (Franklin et al. 2013). In addition to cord marking and check stamping, other surface treatments of limestone tempered wares recovered during excavation at *York Palace* include plain and simple-stamped. Though limestone tempering accounts for a majority of the *York Palace* assemblage, quartz and chalcedony are common tempering agents as well (Langston et al. 2010).

Similar to the *York Palace* ceramic assemblage, a majority of ceramics recovered from *Hemlock Falls Rock House* are limestone tempered cord-marked (Dye et al. 2010). One limestone tempered cord marked sherd returned a BOSL date of A.D. 678 ± 37 (Franklin et al. 2013). Though the limestone tempered cord-marked sherds account for 63% of the total assemblage, limestone tempered plain (8.4%) and siliceous stone tempered (5%) wares are also present but constitute a much smaller portion of the overall assemblage (Dye et al. 2010).

Consistent with *York Palace* and *Hemlock Falls Rock House,* the ceramic assemblage of *Indian Rock House* is dominated by limestone tempered cord marked wares. The remaining portion of the ceramic assemblage includes a variety of limestone tempered wares (plain, check stamped, and brushed), quartz tempered plain, and grit tempered cord marked. Two sherds were selected for BOSL dating, a grit tempered cord marked sherd and a limestone tempered check stamped sherd, and yielded dates of A.D. 680 and A.D. 584, respectively, firmly placing them in the Middle Woodland (Franklin et al. 2013).

Late Woodland

During the Late Woodland (ca. A.D. 800-1200) there was a continuation of hunting, gathering, and gardening economies. Settlements were still relatively small and dispersed, and sedentism increased in most areas of the Southeast. The diversity of foods that were hunted and gathered continued to increase as Late Woodland populations grew (Steponaitis 1986: 384). Other defining characteristics of this cultural period include a significant decrease in regional interaction in many locations, increased evidence for warfare, and the first unequivocal evidence for the bow and arrow (Anderson 2001: 163).

Late Woodland occupation of the UCP is represented by the presence of limestone tempered cord-marked (including smoothed-over cord-marked) and plain pottery (Franklin and Bow 2009: 148). Dates returned for limestone tempered cord marked types come from *Bobcat Arch* (A.D. 803 ± 40), *Mending Hole Rock Shelter* (A.D. 838 ± 101), *Hemlock Falls Rock House* (A.D. 877 ± 97), and *Abri Sous Massif Rock Shelter* (A.D. 887 ± 95) (Franklin et al. 2013). Similarly, BOSL dates were returned on 5 limestone tempered plain sherds from *York Palace* (A.D. 971 ± 97), *Mesa Gap Rock Shelter* (A.D. 1009 ± 34), *Simple Stamped Rock Shelter* (A.D. 1150 ± 92 and A.D. 1189 ± 81), and *Mending Hole Rock Shelter* (A.D. 1385 ± 97) demonstrating that ceramic types indicative of the Late Woodland continued to persist well into the later Mississippian period. Common biface types for the Late Woodland include Hamilton, Madison, and Jack’s Reef varieties (Franklin 2002: 236).

Though cord-marked and plain varieties are still the most prevalent, scraped, and knot-roughened varieties—all almost entirely limestone tempered--have been identified as well (Franklin 2002:238; Franklin 2006). Late Woodland ceramic assemblages from the Ridge and Valley and the UCP share a similar dominance of limestone tempered cord-marking (Franklin 2002: 240). In the Eastern Highland Rim, limestone tempering is minor in the Late Woodland compared to quartz and/or chert tempered wares. Also, knot-roughened and net impressed varieties are more present here than on the UCP (Franklin 2002: 238–239).

Stratified excavations at *Far View Gap Bluff Shelter* revealed a multi-component site with occupation ranging from the Late Paleoindian to the Late Woodland. The most intensive occupation, however, seems to have occurred during the Late Woodland as evidenced from a stratified midden deposit (Franklin 2008a: 91). Radiocarbon and luminescence dates of both limestone tempered plain and smoothed over cord-marked varieties (and a charcoal sooted sherd used for radiocarbon dating) provided a terminal Late Woodland age range for the midden (Franklin 2008a: 92). The recovery of Hamilton and Madison points in good stratigraphic context also corroborate the Late Woodland designation (Franklin 2008a: 91).

When comparing Archaic and Woodland use of the UCP, some differences in occupation and mobility strategies are noted. In other studies conducted on the UCP by Ferguson (1988) and, later, Pace and Hays (1991), different raw material procurement strategies and thus mobility patterns were suggested between Archaic and Woodland groups. Ferguson (1988: 21-32,166-172) proposed different strategies for the Archaic and Woodland on the UCP. Because lithic resources were comparatively scarce in the region, it is expected that most strategies were curated. Archaic hunter-gatherers are thought to have practiced curated technologies while Woodland groups seem to be more expedient.

Based on their work at Station Camp, Pace and Hays (1991) suggest that the differences between Archaic and Woodland patterns are due to under-representation of bifaces at Woodland sites. However, if flake tools are included in the technology, tool to flaking debris ratios for Woodland are comparable to the Archaic (Pace and Hays 1991:130). Pace and Hays (1991) also suggest that raw material use varied less during the Woodland on the UCP. Although Monteagle Chert is the most ubiquitous tool stone in the region other varieties of Mississippian-aged chert including Fort Payne and St. Louis are also available. Pace and Hays (1991: 132, 142) identified Archaic groups as using a wider array of raw materials whereas Woodland groups almost exclusively used local Monteagle Chert.

Franklin et al. (2013) used the previous studies conducted by Ferguson (1988) and Pace and Hays (1991) to frame their work and discussion of lithic technology and mobility within the Woodland on the UCP through excavations at sites such as Hemlock Falls Rock Shelter, York Palace, and Eagle Drink Rock Shelter. Of note here, are 2 important points. First, the sites where the most work has been conducted are all located on the western escarpment of the UCP where access to raw materials is not limited, and second, lithic use-wear analyses are included in these studies (Franklin et al. 2013). Lithic analyses from the above excavations revealed that the exploitation of different raw materials was no less variable in the Woodland than in the Archaic--likely meaning that mobility was high and far-ranging in both periods contra Ferguson (1988) and Pace and Hays (1991). Further, lithic use-wear analyses of stone tools recovered from 3 Woodland sites on the UCP indicate a variety of foraging activities were conducted on site. So, based on the lithic assemblages of Woodland sites on the western escarpment portion of the UCP, Woodland peoples appear to have practiced residential mobility strategies in contrast to the logistically organization seen during the Archaic period (Franklin et al. 2013). Also different from the Archaic period, Woodland people were exploiting dark-zone cave environments not just for chert but also for mineral resources such as gypsum (Franklin 2002, 2008b).

Faunal material recovered during excavations on the UCP of Tennessee reveal a broad subsistence range for the Woodland period that mainly included white-tailed deer and wild turkey though small-to-medium sized mammals such as squirrel, beaver, and fox were important resources as well (Franklin et al. 2013). In addition, variation in seasonal occupation and use of rock shelters on the UCP is evident from the recovery of fish, shellfish, and reptilian species—this coupled with the recovery of charred acorns and hickory nuts, suggests both warm and cold weather occupations.

Using multiple lines of evidence (analysis of lithic, faunal, and archaeobotanical material), Franklin et al. (2013) suggest that Woodland sites are not all simply special-purpose camps as was suggested by Pace and Hays (1991). A variety of activities were noted at several of the sites discussed above suggesting seasonal movement with the UCP by family groups. Unlike the Late Archaic, however, Woodland peoples were mainly residentially mobile hunter-gatherers that used rock shelters and caves for residential occupation, shelter, mineral extraction, burial, and artwork (Franklin et al. 2013).

Mississippian

Broadly speaking, the Mississippian Period (ca. A.D. 1200-1700) was a time of great changes in technology, subsistence, settlement patterns, sociopolitical integration, and ideology that in turn, produced societies far different than that of their predecessors. Some defining characteristics of the Mississippian Period include the construction of platform mounds that housed important religious or political structures, the arrangement of mounds or houses around central open plazas, dramatic population increases, the development of organized chiefdoms, increased conflict and warfare, the introduction of shell-tempered pottery, and the emergence of an elaborate ceremonial complex (Chapman 1985: 74; Steponaitis 1986: 387–388). More recently, however, Anderson and Sassaman (2012: 152-153) point out that the there is great variation in what “defines” the Mississippian period throughout the Southeast. This suggests that the traditional defining characteristics like those listed above are not enough to truly capture the geographical, temporal, and cultural variation seen during the Mississippian period.

Although there is little evidence of Mississippian peoples living in permanent nucleated villages, recovered artifacts, radiocarbon dates, and the presence of classic SECC iconography demonstrates their strong presence in the region (Franklin 2002: 244). Also, some mounds have been identified in the area though it is not clear yet whether these represent Woodland or Mississippian occupation (Franklin 2002; Franklin et al. 2013). Thus far, approximately 30 Mississippian components have been identified during archaeological surveys of the UCP of Tennessee (Franklin 2002; Langston and Franklin 2010). The high number of Mississippian Period sites on the UCP compared to adjacent regions (see Ferguson et al. 1986; Sussenbach 1990) is possibly explained by the inclusion of material and dates from dark zone cave environments in Franklin’s (2002) survey. It is clear that Mississippian peoples were at least occupying and or traversing the UCP based on BOSL dates from shell-tempered and limestone tempered plain ceramics (Franklin et al. 2013). One example includes a shell tempered plain sherd recovered during excavations at *Hemlock Falls Rock House* which was dated to A.D. 1497 ± 41.

The decline of the Mississippian culture began with the onset of the Little Ice Age (A.D. 1300) around the end of the Medieval Warm Period. During this time, Mississippian populations appear to have experienced times of increased warfare, settlement nucleation, and decreased long distance exchange (Anderson 2001: 166). European contact further facilitated the decline of the Mississippian culture complex. Disease and warfare brought on by the Europeans coupled with internal conflicts within chiefdoms eventually led to the ultimate demise of the Mississippian culture (Steponaitis 1986: 393).

CHAPTER 4

ARCHAEOLOGICAL SITE LOCATION MODELING

This chapter introduces the background and concepts of archaeological site location modeling, more commonly referred to as predictive modeling. In order to understand the methods detailed in the following chapter, a brief introduction to GIS and predictive modeling in archaeology is provided. Next, regression-based approaches used in modeling are reviewed with the goal demonstrating the need for the spatial logistic regression approach used in this thesis. Finally, factors believed to influence site selection are discussed with emphasis on determining model variables.

GIS and Archaeology

Archaeology deals with spatial data on a routine basis. In fact, almost all data recovered by archaeologists are spatial in nature (i.e. locations of sites, locations of artifacts within a site boundary, settlement and mobility patterns, distribution of cultural traits, *etc*.). As Wheatley and Gillings (2002: 3) state

> Artefacts, features, structures, and sites, whether monument complexes, chance finds or individual objects, scatters of ploughsoil material or rigorously excavated structural and artefactual, are all found *somewhere.* As well as the position of the feature or artifact itself there may also be a series of *relationships* between the locations of features and artefacts, revealed by significant patterns and arrangements relative to other features and things [emphasis in original].

The “other features and things” refer to either features of the environment, other archaeological features, or some cosmological phenomena. The underlying idea is that understanding spatial relationships is critical in constructing frameworks for studying and interpreting the archaeological past. Because archaeology is concerned with the interpretation of spatially (geographically) referenced material, spatial technologies can better facilitate archaeological research. Some examples of spatially-related technologies useful in archaeological analysis include Remote Sensing, Global Positioning Systems (GPS), and Geographical Information Systems (GIS)—the last of which is of interest here and will be discussed further.

Geographical Information Systems (GISs), broadly speaking, are computer-based applications concerning the acquisition, storage, or manipulation of spatial information. The spatial information can be modeled as either vector or raster data. Vector data (i.e., points, lines, and polygons) have discrete boundaries and are spatially independent. Examples of vector data used in archaeology include the location and boundary of a site, roads, water resources, and locations of technological resources. On the other hand, a raster (continuous surface made up of individual grid cells) represents data best visualized as a surface without discrete boundaries such as elevation, slope, aspect, temperature, or precipitation. The GIS interface provides archaeologists a way to combine and manage both vector and raster data, perform computationally intense calculations, and explore new avenues of analysis with unconventional data types (Kvamme 1989).

Development of GIS

Before the development of GIS, the spatial component of archaeological data was studied by simply viewing hand-plotted, flat maps for similarities or differences (Wheatley and Gillings 2002: 4–5). Around the early 1960s, the quantitative revolution and New (Processual)

Archaeology brought about major changes in how the spatial relationship of material culture was interpreted; previous practices were believed to be too subjective and descriptive without actually explaining spatial patterns (Wheatley and Gillings 2002: 5). During this time, archaeologists saw prehistoric behavior as identifiable and measurable patterns in space that could reveal the prime causal factors for changes in behavior. The shift to the Processual Archaeology school-of-thought was further facilitated by the application of new spatial analytic techniques and methods such as computer-aided cartography and GIS.

Though some cartographic computer programs are said to date as early as 1950, it was during the 1960s and 1970s that several computer programs were created for the sole purpose of making geographic maps from digital data (Coppock and Rhind 1991; Wheatley and Gillings 2002: 12). Similarly, the first recognizable GIS, the Canadian Geographic Information System (CGIS), was implemented in 1966 for managing and monitoring the country's natural resources; however, it took almost 3 years and over 566 technicians to overlay all of the Canada Land Inventory maps (Tomlinson 1988). The computational difficulties with the CGIS encouraged computer scientists to develop more efficient and automated approaches (Coppock and Rhind 1991: 23).

The significant developments in automated computer technology during the late 1960s and early 1970s are perhaps most attributable to activities within government departments and agencies. Some examples of systems implemented by federal and state agencies include the United States Geological Survey's (USGS) Geographical Information Retrieval and Analysis System (GIRAS) developed in 1973 and the Minnesota Land Management Information System (MLMIS) in 1976 (Coppock and Rhind 1991: 31). Around this time, there was also a shift in computer-automated cartography from the use and development only within government agencies to the commercial sector; the Environmental Systems Research Institute (ESRI) began selling its first vector-based GIS program in the early 1970s (Coppock and Rhind 1991; Wheatley and Gillings 2002: 14).

The USGS continued to play an important role in the development of gathering, analyzing, and displaying cartographic data; this began with the digitization of topographic maps and the collection of other digital land resource data in the mid to late 1970s. Then, in 1987, the USGS created and distributed one of the most widely-used types of spatial data—the digital elevation model, or DEM (Starr and Anderson 1991). By this time, GIS was on its way to becoming widely accepted as the number of programs, classes, facilities, and projects grew exponentially (Coppock and Rhind 1991: 33).

Archaeological Applications of GIS

Perhaps the first mention of GIS in the archaeological literature was by H.J. Pomerantz in 1981, though software for cartographic and spatial analysis had been in use for archaeological analyses since the 1970s (Kvamme 1998; Wheatley and Gillings 2002: 15). Although the beginnings of GIS in archaeology are not completely clear, by the late 1990s, GIS had become a wide-spread addition to the discipline of archaeology (for examples see K. L. Kvamme 1990; Gaffney and Stančič 1991; Lock and Moffett 1992; Andresen, Madsen, and Scollar 1993; Lock and Stančič 1995; Maschner 1996; Fisher et al. 1997). Kvamme (1998: 1) gives 3 main reasons for the growth of GIS in the field: the demand for state-mandated databases of cultural resources on government lands, the requirement of archaeological distribution models by CRM agencies, and the examination of sites with environmental data using computer technology.

Applications of GIS in archaeology have varied throughout the years with 3 typical applications: visualization, management, and predictive modeling (Church et al. 2000: 144).

Wheatley and Gillings (2002: 207) more broadly categorize current GIS applications in archaeology as either Management or Research. Under the Management category is Database Management and Cultural Resource Management (CRM); this category focuses on the storage, maintenance, and analysis of archaeological databases for the management and protection of archaeological (cultural) resources. The Research category is further subdivided into applications focusing on the regional landscape and intra-site spatial analysis—with landscape-based studies being the most common application of GIS in archaeology (Wheatley and Gillings 2002: 209). Regional landscape studies attempt to explain how prehistoric people interacted with their environment using the spatial statistical relationships between material culture, human alteration of the environment, and the natural environment. The application of landscape-based approaches in archaeology inevitably includes the development and application of predictive models.

Predictive Modeling

Background

As far back as Herodotus's *Histories* written in the fifth century BC, questions have been raised about the role of the environment in creating human diversity--this has been a reoccurring theme in both anthropology and geography over the centuries (Hodgen 1964). Throughout the development of the field of anthropology (and thus archaeology), several theories have focused on the environment and how it affects and influences culture and cultural change. Alfred Kroeber's (1939) work on the environmental relationships between native North American cultures and their culture areas (Wissler 1927) had a major influence on the study of environment and culture. Another prominent figure in anthropology at the time, Leslie White, also believed that humankind, and therefore culture, is dependent upon adjustment to the natural environment

(White 1949: 365). Following the work of Kroeber (1939) and White (1949) was the development of the concept of culture ecology by Julian Steward (1955); this concept focused on how the relationship between environmental resources, the tools and knowledge needed to exploit them, and the organization of work had a determinant effect on social practices. Further, Steward's (1955) work emphasized the interaction (and opposition) of humans with the environment. The study of archaeological settlement patterns developed mainly as a result of Julian Steward's work (Kohler 1988:30).

Following Julian Steward, Gordon Willey's (1953) work in the Viru Valley defined a new field of inquiry and pioneered the way for future settlement studies. Willey (1953: 1) defined the term "settlement pattern" as the "…way in which man disposed himself over the landscape in which he lived." Further, though he was more interested in social interaction and control and their effect on community patterns, Willey discussed the role of environmental, technological and demographic change on settlement patterns. Following his Viru Valley work, Willey (1956) put together an edited volume on prehistoric settlement patterns where authors investigated environmental, social, and political factors as determinants in the distribution of human populations. The study of archaeological settlement patterns continued for another decade as new determinants of site location (i.e., availability of natural resources, defense factors) were investigated (Trigger 1968).

During the 1970s, 2 major advances changed the nature of settlement pattern studies (Kohler 1988:31). First, a new analytical method for investigating determinants of site location was developed. Site catchment analysis, as it was termed, emphasized the importance of economic resources (the availability, abundance, spacing, and seasonality) in determining site location (Vita-Finzi and Higgs 1970; Roper 1979). The second important advance of the 1970s relates to the broader changes that were occurring in archaeology at the time. Much of the early settlement pattern studies follow what Kohler (1988:31) calls “an anecdotal form” because each mirrored Steward’s (1955) approach without any sense of progression. Then, with the shift to more quantitative methods in archaeology, formal statistical techniques were incorporated into settlement pattern analysis. This led to the development of statistical models used to predict site densities in areas yet to be surveyed by archaeologists (Verhagen and Whitley 2012: 51). This practice, termed “predictive modeling” became increasingly widespread throughout the 1970s.

The earliest works such as those by Plog and Hill (1971) and Green (1973) incorporated statistical procedures for predicting site locations. Green’s (1973) work in Belize was the first to apply multivariate statistics (e.g. multiple linear regression) to archaeological predictive modeling. However, some researchers did not support the application of predictive models to examine and explain prehistoric behaviors and proposed that they only be constructed for CRM purposes (Sullivan and Schiffer 1978). But even within a CRM context, some believed that predictive models did not provide reliable, hard data and there could be absolutely no substitution for intensive ground reconnaissance of the entire area of potential effect (Kohler 1988:34).

Still, the application of predictive models increased dramatically by the late 1970s and early 1980s in response to federal legislation such as the National Historic Preservation Act of 1966 (amended in 1976, 1980, and 1992) that required the identification of historical and archaeological resources. Because of the time required to complete comprehensive surveys of federal and state lands, agencies such as the Bureau of Land Management, Army Corps of Engineers, and the United States Forest Service began to fund archaeological surveys encouraging the creation and use of predictive models. Though many predictive models were produced at this time, Kohler (1988:35) states that "… (judging by the variability in techniques and products) no one was sure how prediction might be best accomplished."

As discussed previously, the archaeological applications of GIS soared during the 1990s with advancements in spatial technologies and computer programming. However, as GIS in archaeology was achieving heightened popularity and success, so was Post-Processual Archaeology. The processual approach to settlement studies focused more on the environmental factors that influenced the site selection process. In contrast, post-processualism emphasized the subjective nature of archaeology and argued that the use of GIS and predictive modeling encouraged ideas of environmental determinism (Gaffney and van Leusen 1995; Wheatley 1996; Wansleeben and Verhart 1997). Today, both sides continue to be argued and Processual and Post-Processual approaches to archaeological site location modeling are still employed.

Inductive vs. Deductive Models

Because the development of predictive modeling has both a theoretical (i.e., cultural ecology and settlement pattern analysis) and a quantitative (i.e., introduction of statistical techniques) background, 2 separate approaches to modeling emerged during the 1970s and 1980s. Though the approaches significantly differ in their underlying frameworks, they can often overlap and should not be considered mutually exclusive (Kamermans and Wansleeben 1999; Verhagen and Whitley 2012: 52). Early models developed by those such as Jochim (1976) and Bettinger (1980) were largely theoretical and did not include spatially quantitative evaluations. This type of theory-driven model, later called the "deductive" approach, is constructed using *a priori* knowledge of the archaeological record for a specific area; the model is then evaluated using known site locations (Kamermans and Wansleeben 1999: 225).

In contrast, an “inductive” model is constructed using correlations between known sites and their attributes (mostly environmental). This information is then used to predict potential site locations using some form of statistical analysis. Some of the earliest examples of the “inductive”, or data driven, approach include Kvamme’s (1984) model of prehistoric site location in Pinyon Canyon and Parker’s (1985) multivariate logistic approach to prehistoric settlements in the Sparta region of Arkansas. The data driven approach has been the most commonly applied method in the United States as evidenced by applications found in Judge and Sebastian (1988), Wescott and Brandon (2000), and Mehrer and Westcott (2006).

The Upper Cumberland Plateau model (developed herein) is a result of inductive and deductive approaches. Though the model was developed and tested using statistical techniques, the model variables were selected using what was already known about the region (geographically and archaeologically) and on theories of prehistoric hunter-gatherer behavior. The statistical and theoretical approaches to the Upper Cumberland Plateau model are the focus of the following sections.

Statistical Prediction Models

In inductive archaeological predictive modeling, several different statistical techniques have been used, both parametric and nonparametric. Both techniques are robust, with parametric models assuming a particular type of statistical distribution (i.e., multivariate normality) and nonparametric models making no assumptions about distributional form (Kvamme 1988: 364; K.L. Kvamme 1990). In practice, normality is a difficult condition to satisfy, especially with complex relationships involving human behavior and the environment. For this reason, nonparametric methods have been considered more appropriate for modeling complex, non-linear relationships (Parker 1985; Espa et al. 2006; Zhang et al. 2010). Whether a parametric or nonparametric method is employed, the chosen technique should be appropriate for addressing the model objectives and handling the type of data used.

Regression-based models are some of the most commonly used approaches in archaeological predictive modeling. The basic goal of regression analysis is to analyze the relationship between the dependent or response variable and one or more independent or explanatory variables. In general, site presence is the response variable, with a variety of environmental variables (e.g. distance to water, elevation, slope) used as explanatory variables. There are several types of regression analyses, each with associated strengths and weaknesses in producing archaeological predictive models (Wheatley and Gillings 2002: 152). Some of the more common types of regression analyses are outlined below with a focus on evaluating the appropriateness of each method for modeling the probability of a binary response variable (site presence vs. site absence) given a set of explanatory variables.

Linear Regression

Linear regression models the relationship between a scalar (continuous) response variable and one or more explanatory variables by fitting straight line to the set of observed data. The interpretation and analysis of linear regression is concerned with the effect of the explanatory variables on the response variable and the nature of the fit of the line (Rogerson 2010: 201). Simple regression involves a single explanatory variable, whereas multiple regression involves 2 or more explanatory variables. Linear regression, like other linear models, assumes there is a linear relationship between the response and explanatory variable(s) and the relationship is modeled through the error term, or residuals. The (multiple) linear regression model takes the form

$$y = \alpha + \beta_1 x_1 + \beta_2 x_2 + \cdots + \beta_n x_n + \varepsilon \qquad \text{Equation 1}$$

Where:

y= response variable

x= explanatory variables

α=intercept

β=regression coefficients

ε =residuals or error term

Linear models such as this require the unknown model parameters (β) to be estimated from the data in order to find the best-fitting straight line. Though there are many estimation techniques for linear regression, the most common method is ordinary least squares (OLS). This method fits a line to the data by minimizing the sum of the squared residuals. This is different from other methods which minimize the sum of the residuals and therefore cause the negative values to cancel out the positive values (Kahane 2008: 18–19).

Standard linear regression models make several assumptions about the relationship between the response and explanatory variables; if the assumptions are satisfied, then the estimated regression line represents the best possible fit (Kahane 2008: 31–33). The more formal assumptions include randomness, independence among the response variable, and normality. All of these assumptions (and others) apply to simple linear regression models. In the case of multiple linear regression, an additional assumption is required in that there should not exist any perfect linear relationship, or multicollinearity, between explanatory variables. Multicollinearity causes problems in a model because it does not allow for the subtle effects of 2 correlated variables to be clearly distinguished—the unique explanatory ability of one explanatory variable would be lost (Kahane 2008: 120).

In archaeological predictive modeling, linear regression methods are useful for predicting things such as artifact densities or site dimensions (K.L. Kvamme 1990: 270; Wheatley and Gillings 2002: 154). However, when the response variable is categorical—site or no site--methods such as OLS are not appropriate. In addition, if linear regression is used to predict the probability of a dichotomous outcome, the predicted values are not necessarily restricted to the 0 to 1 interval; this will severely complicate model interpretation and analysis (Parker 1985: 176). For these reasons, standard linear regression and OLS are not suitable methods for predicting archaeological site locations in the form of "site presence" or "site absence".

Logistic Regression

Unlike standard linear regression models, logistic regression can properly handle a categorical response variable and does not assume that the explanatory variables are normally distributed. Similarly, given a set of values for the explanatory variables, logistic regression predicts the probability of a positive response variable (Parker 1985: 176). There are 2 types of logistic regression: binomial (or binary) and multinomial. In binomial logistic regression, only 2 possible outcomes are modeled (e.g. "yes" vs. "no", "site presence" vs. "site absence"); the codes "0" and "1" are generally used for this method. Multinomial logistic regression is applied to cases where 3 or more possible categorical outcomes (i.e., artifact classes, site types, or time periods) are modeled.

Logistic regression uses the logit transform to convert the standard regression equation into a probability of a case by restricting the output between 0 and 1. The probability of the event occurring increases as the predicted value gets closer to 1. In the case of binomial logistic regression, the resulting equation (Eq. 2) yields the probability of a positive response for each unit of analysis (Parker 1985: 177).

$$p(Y) = 1 / (1 + Exp\{-z\}) \qquad \text{Equation 2}$$

Where:

p(Y) = the probability of the event occurring

$z = \alpha+\beta_1 x_1+\beta_2 x_2 \ldots.+\beta_n x_n$

α= constant, or intercept

β= regression coefficients

x=explanatory variables

From this equation, the probability of occurrence is modeled from a binary response where, in the case of archaeological predictive modeling, "1"can represent site presence and site absence equals "0". The results can then be interpreted as the probability of archaeological site presence given a set of values for the independent variables.

Because it can handle different data types and operates under fewer assumptions about the form of the independent variables, logistic regression has become increasingly popular in archaeological modeling (Kvamme 1990: 275). However, one main issue related to logistic regression (and many other traditional statistical approaches) is the assumption of spatial independence of the response variable without considering its spatial nature (Espa et al. 2006: 148). According to Tobler's (1970) first law in geography, everything is related to everything else; that is, phenomena distributed in space are related by their proximity to each other. This concept, spatial autocorrelation, "means a dependency exists between values of a variable in neighboring or proximal locations, or a systematic pattern in values of a variable across the locations on a map due to underlying common factors" (Griffith 2009: 1). Traditional logistic regression assumes the data are spatially independent and the output can be misleading if the data are, in fact, spatially autocorrelated. Spatial statistical tests, though based on conventional statistics, incorporate the spatial aspect of data and can provide more robust findings (Schwarz and Mount 2006: 155). One way to address the issue of spatially autocorrelated data is to use a spatial model in lieu of traditional methods such as logistic regression.

Spatial Dependence Models

A traditional logistic regression model is not appropriate for handling spatial data when spatial autocorrelation is present in a dataset. When a value of a variable at one location depends on its value at neighboring locations, there is spatial dependence, or spatial autocorrelation. Positive spatial autocorrelation exists when values tend to be more similar the closer they are together (e.g. high values near high, low values near low); this type of spatial autocorrelation is common in many environmental datasets such as elevation, temperature, and rainfall (Conolly and Lake 2006: 158). Conversely, when dissimilar values are located closer together (e.g. high values near low values), negative spatial autocorrelation is present. For a dataset with significant positive or negative spatial autocorrelation, a spatial statistical model should be employed; if spatial dependence is ignored, the real variance in a dataset can be underestimated. There are 2 types of spatial dependence models that can handle spatially autocorrelated data: spatial lag and spatial error. These are alternative ways of running a linear regression but with a spatial component—this is the reason for their discussion here. Both models operate under the same assumptions: 1) normality in the dependent variable; 2) spatial autocorrelation; and 3) a linear relationship between inputs and outputs. The difference between the 2 models is how spatial autocorrelation is handled—as either substance or nuisance (Ward and Gleditsch 2007: 30).

Spatial Lag Model. A spatial lag model accounts for spatial autocorrelation in the response variable that can be explained by the explanatory variables. This model considers spatial association an important feature that can reveal something about the relationship between the response and explanatory variables. The spatial lag model is represented by

$$y = \alpha + \beta_n x_n + \rho W y_i + \varepsilon \qquad \text{Equation 3}$$

Where:

$\alpha =$ constant, or intercept

β_n = regression coefficients

x_n= explanatory variables

$\rho =$ spatial autoregressive parameter

W = Spatial Weights Matrix

y_i= lagged predictions at nearby points

$\varepsilon =$ random error term

Spatial Error Model. In contrast to a spatial lag model, a spatial error model captures spatial autocorrelation in the error term. This model is primarily used when it is believed that there is some spatial pattern that will be reflected in the error terms but no assumptions can be made about the origin of the error (Ward and Gleditsch 2007: 59). This means that the explanatory variables do not fully capture (or explain) the spatial dependence and therefore, it is mostly ignored. The spatial error model is represented by

$$y = \alpha + \beta_n x_n + \lambda W \xi_i + \varepsilon \qquad \text{Equation 4}$$

Where:

$\alpha =$ y-intercept

x_n = explanatory variables

$\beta_n =$ coefficient of explanatory variables

$\lambda =$ coefficient of lagged autoregressive errors

W = Spatial Weights Matrix

$\xi =$ error term associated with nearby points

$\varepsilon =$ random error term

Simply put, a spatial lag model assumes that “neighboring values of the response variable exert a direct effect on the value of the response variable itself”, while a spatial error model assumes that the errors of a model are spatially correlated and “disregards the possibility that the observed correlation may reflect something meaningful about the data generation process” (Ward and Gleditsch 2007: 55). Though both of these models can account for spatial dependence in a dataset, they are parametric methods with strict statistical assumptions and model continuous response variables; these are not suitable for this project because the response variable is dichotomous and normality cannot be assumed. A statistical method that is spatial and can handle a categorical response variable is ideal for this study.

Spatial Logistic Regression

Though traditional (e.g. aspatial) logistic regression has been one of the preferred statistical techniques in archaeological predictive modeling, it does not account for the spatial nature of many archaeological phenomena. In recent years, the incorporation of spatial statistical methods in archaeological predictive modeling has been strongly encouraged in order to generate more accurate and valid models (Schwarz and Mount 2006: 172). Spatial logistic regression is preferred over traditional logistic regression in archaeological modeling because it has a built-in spatial function and does not ignore spatial autocorrelation.

Geographic Information System (GIS) programs have facilitated the application of archaeological predictive models as new visual and analytical tools have been developed. Using a combination of GIS and statistical programs, spatial logistic regression can be applied to a study area divided into evenly-spaced grid cells (or pixels). Each cell represents either site presence or absence, according to a database of archaeological sites. Spatial logistic regression can then be used to predict the presence of a site based on values of the explanatory variables at the known "site presence" locations. This method is referred to as pixel-based spatial logistic regression and has been equated to a Poisson point process model for the original data points (Baddeley et al. 2010: 1155). The spatial logistic regression formula (Equation 5) takes a similar form as traditional logistic regression, but with an offset term equal to the log of pixel area (Baddeley et al. 2010: 1173).

$$p(Y) = 1/(1 + Exp(\{-z\}\ )$$ Equation 5

Where:

$p(Y_j)$=the probability of a case for a given cell or pixel

z = $log\ \alpha +\ \beta^T z_j$

α= pixel area

β^T= regression coefficients for corresponding explanatory variable

z_j= values for each explanatory variable associated with a pixel

The concept of spatial logistic regression was originally developed in geology to predict potential metallic deposits for mineral exploration in Western Australia (Agterberg 1974). This study demonstrated that the predicted probabilities of a traditional logistic regression are significantly influenced by the size of the spatial unit (i.e., grid cell or pixel) under consideration (Baddeley et al. 2010: 1156). Most spatial datasets are aggregated into zones (i.e. arbitrary boundaries for a study area or site); the placement and geographic scale of a zone can influence the interpretation of statistical analysis where different zoning systems can produce different results. This concept is known as the “modifiable area unit problem” (Rogerson 2010: 16). Spatial logistic regression attempts to minimize this problem by incorporating the size of a “zone” as a new model term.

With the exception of a few studies (Agterberg 1974; Scholtz 1981; Hasenstab 1983; Kvamme 1995), there seems to be very little literature addressing spatial logistic regression directly. Not only is the method more complex than traditional logistic regression, but it is not an option in most commonly used spatial statistic software packages. Spatial programs such as GeoDa (Anselin et al. 2006) and ArcGIS (ESRI 2011) have the capabilities to perform different types of linear regression such as OLS and Geographically Weighted Regression (GWR) but not logistic regression. Also, traditional statistical packages like SPSS (IBM Corp 2011) can be used to perform logistic regression but treat the data as if they are non-spatial. The statistical and graphical R environment (R Core Team 2012) is seemingly one of very few statistical systems that has the capability of performing a spatial logistic regression.

Spatial logistic regression is the most statistically robust approach to archaeological predictive modeling and therefore merits heavy consideration as a methodological approach. Because it is the only method that satisfies the requirements of a binary response variable and accounts for spatial autocorrelation within a dataset, spatial logistic regression is used to generate the Upper Cumberland Plateau predictive model.

Site Selection Factors

Besides choosing a modeling approach, it is necessary to identify what factors might have influenced site selection in order to generate relevant model variables. The choice of model variables largely depends on the availability of data. In this case, the availability of existing spatial data has a major impact on what can be used to generate a predictive model using GIS. This is a common and often criticized problem in predictive modeling. Though the specific variables used to generate the UCP model will be discussed in the following chapters, this section provides some background on prehistoric site selection and the types of variables commonly used in archaeological predictive modeling.

In one of the earlier works on predictive modeling, Jochim (1976) developed a model specifically addressing hunter-gatherer settlement and subsistence patterns and how hunter-gatherer settlement locations can be viewed as the result of the decision-making process. From Jochim's (1976: 50) seminal work, 3 primary goals guiding hunter-gatherer settlement placement have been used in predictive modeling studies as a basis for analyzing and interpreting the location of prehistoric hunter-gatherer settlements: the proximity of economic resources, shelter, and view. Though Jochim (1976) believed that subsistence-related activities were the primary factors influencing settlement locations, critics point out that models should also incorporate variables that describe social factors as well.

Common Variables

Environmental variables such as elevation, slope, aspect, and measures of topographic relief are some of the most common variables used in archaeological modeling (Kohler and

Parker 1986; Warren and Asch 2000; Altschul et al. 2004; Ridges 2006). Similarly, modeling studies often employ variables related to geologic and geomorphic changes within an area; some examples include different measures of terrain roughness, topographic position, geology, vegetation, soil series, and soil-related properties such as drainage class or erosion (Kvamme 1988; Duncan and Beckman 2000; Warren and Asch 2000; Altschul et al. 2004; Lock and Harris 2006; Mink II et al. 2006; Ridges 2006; Veljanovski and Stančič 2006; Finke et al. 2008). Measures of solar radiation and viewshed have also appeared in modeling studies, though they are much less common than other environmental variables (Duncan and Beckman 2000; Krist Jr. 2006; Madry et al. 2006; Veljanovski and Stančič 2006). Lastly, the availability or proximity to water resources is a common variable used in archaeological predictive modeling. Though most basic models include straight-line distance to water sources, variables incorporating cost-distance analysis are becoming more popular (Madry et al. 2006; Ridges 2006). All of these variables are useful in archaeological site location modeling because they are related to fundamental utilitarian needs of humans.

The correlation of the natural environment and the distribution of hunter-gatherer settlements was a well-established concept by the early 1980s (Jochim 1981; Ebert and Kohler 1988). However, environmental variables are not entirely sufficient to explain the variation in settlement patterns (Gaffney and van Leusen 1995). Factors beyond those that are strictly related to the environment must be considered in order to understand the full range of prehistoric site location variability. Rock shelters pose a problem in that they are fixed places on the landscape and dictated purely by environmental variables. The presence of a prehistoric rock shelter site, however, is a combination of environmental restrictions and selection by prehistoric peoples. Variables that introduce some degree of decision-making by prehistoric peoples in the site selection process can also be used to generate a site location model of prehistoric rock shelters. Though such variables (e.g. proximity to resources, solar radiation, and viewshed) are directly related to the environment, they can be used to investigate human behavior and associated land-use patterns. So although it is impossible to completely understand the adopted beliefs and strategies of prehistoric peoples, modeling attempts should incorporate variables that most accurately reflect the environmental setting and the archaeological record of the area under study. With this in mind, the UCP model was developed using explanatory variables that incorporate factors related to the physical environment and human behavior.

CHAPTER 5

MODEL DEVELOPMENT

This chapter details the methods used to develop and test the UCP site location model. Data required for building the model include known rock shelter locations and model variables generated by GIS data layers. These 2 sets of data will hereafter be referred to as the response and explanatory variables respectively. ArcGIS 10.0 (ESRI 2011) was used to create, process, and store all GIS data layers for the UCP model and the statistical and graphical R environment (R Core Team 2012), hereafter referred to as R, was used to run the spatial logistic regression model. All GIS data were projected using the North American Datum 1983 State Plane of Tennessee.

Response Variable

In terms of a statistical model, the response variable is predicted from a set of explanatory variables. The known locations of prehistoric rock shelter sites are used as the initial response variable to identify the unique characteristics that identify them and to find where other not-yet-discovered rock shelters are likely to exist. Data collected from 2 archaeological surveys, the East Obey and Pogue Creek State Natural Area, are used to develop and analyze the UCP site location model (Figure 4).

The first long-term, systematic archaeological survey on the UCP of Tennessee focused on the southern portion of the Western Escarpment (Franklin 2002). The overall purpose of the survey was to identify archaeological sites that could define the cultural history of the region. One hundred forty-five new sites were identified—77 of which were selected for this study (Franklin 2002: 245,249).

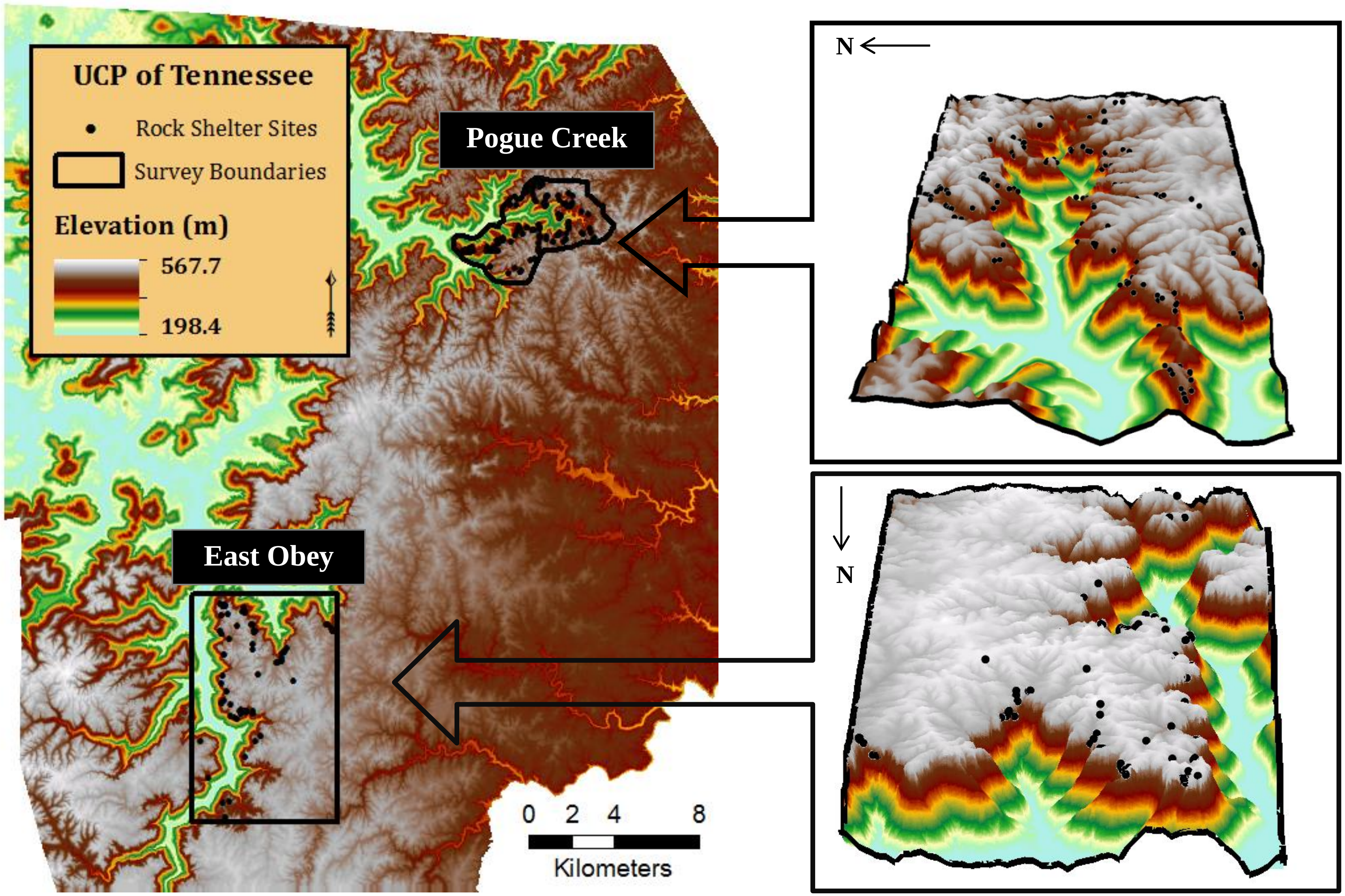


Figure 4: Pseudo-3D Renderings of the Pogue Creek and East Obey Survey Areas. The renderings have been rotated in a way that best demonstrates the topographic locations of the known prehistoric rock shelter sites used to develop and test the site location model.

In the summer of 2006, Franklin entered into a long-term Memorandum of Understanding with Tennessee State Parks to conduct archaeological survey of the newly acquired Pogue Creek State Natural Area (Langston and Franklin 2010). The land was purchased by the Tennessee Chapter of Nature Conservancy to protect it from development and the State of Tennessee subsequently purchased the property. The Pogue Creek State Natural Area archaeological survey was completed in 2010; 135 archaeological sites were recorded over the course of 4 short winter survey seasons, of which 127 were prehistoric rock shelters sites (Langston and Franklin 2010).

Two rock shelter databases were created for the Pogue Creek and East Obey survey areas. The databases included both geographic location and archaeological information recorded during survey. Point shapefiles were generated for each database in ArcMap 10.0 (ESRI 2011). After eliminating spatial outliers, 125 known rock shelter locations in the Pogue Creek State Natural Area were used to develop the UCP site location model. Because the East Obey rock shelter sites (n=77) are relatively close to Pogue Creek and the topography of the Western Escarpment of the UCP is very similar, the East Obey dataset was used to test the model.

Explanatory Variables

The explanatory variables in a statistical model are the inputs used to predict an event or response. The explanatory variables used in the UCP model attempt to address both the environmental restrictions of rock shelter locations and other factors that may have influenced site selection by prehistoric hunter-gatherers. Explanatory variables were chosen in an effort to isolate and satisfy the above conditions for locating a prehistoric rock shelter site. Two separate models were developed and then combined to generate the final UCP model; the explanatory variables were assigned to 1 of the 2 models. The following section introduces the 2 different types of models used to create the final model. Then, the explanatory variables are discussed in terms of creation and incorporation within their respective model groups.

Data Acquisition

A GIS was developed for the UCP using several sources of geospatial data in ArcMap 10.0 (ESRI 2011). Table 1 is a list of the original data sources used in this study.

Table 1: Sources of Geospatial Data for the UCP Model. Four geospatial datasets were used to generate the UCP site location model. The scale, download source, and original source are listed for each of the 4 datasets needed for this project.

Data Type & Scale	Data Download Source	Original Source
Elevation (10 m horizontal resolution)	Tennessee Data Spatial Server, Data Collections, Digital Elevation Models (DEM) http://www.tngis.org/	United States Geological Survey, National Elevation Dataset http://ned.usgs.gov/
Soil (1:24,000)	United States Department of Agriculture, Natural Resources Conservation Science, Soil Data Mart http://SoilDataMart.nrcs.usda.gov/	Soil Survey of Fentress and Pickett Counties Area, Tennessee, 1995; Soil Survey of Big South Fork National River and Recreation Area, Kentucky and Tennessee, 2008
Geology (1:250,000)	Tennessee Data Spatial Server, Data Collections, Geology of Tennessee http://www.tngis.org/	Hardeman, W.D. (1966). Geologic map of Tennessee: State of Tennessee Department of Conservation, Division of Geology, 4 sheets, scale 1:250,000.Digitized in 2000 by the U.S. Geological Survey Water Resources Office in Tennessee.
Hydrography (1:24,000)	United States Department of Agriculture, Natural Resources Conservation Science, Geospatial Data Gateway http://datagateway.nrcs.usda.gov/	United States Geological Survey, National Hydrography Dataset http://nhd.usgs.gov/

Static and Dynamic Variables

Two factors dictate the location of prehistoric rock shelter sites: 1, where rock shelters are located based on where they naturally form and, 2, selection by prehistoric people based on some set of preferential conditions. In order to capture both conditions, the preliminary explanatory variables were divided into 2 groups using static and dynamic factors (Zhang, Zhang, and Zhou 2010: 389). The 2 groups of variables were used to generate separate models. The static (P1) and dynamic (P2) models were then combined (by multiplication) to generate the final UCP model. The first group (P1) represents the physical attributes of the landscape more likely to produce a rock shelter location and thus the static factors. Theoretically, the P1 model could be used by itself to identify areas with the potential to yield any rock shelter—site or non-site. Thus the second group (P2) includes dynamic factors that may have been important to prehistoric peoples for selecting residential sites. Explanatory variables are discussed within the context of these 2 model groups.

Preliminary Explanatory Variables

A total of 27 preliminary explanatory variables were generated for the UCP model (Table 2). Because the model was run in the statistical and graphical R environment (R Core Team 2012) using the spatial logistic regression model (slrm) function (Baddeley et al. 2010), all explanatory variables had to be scalar, or continuous image files (e.g. TIFFs). Each variable is discussed in terms of its relevance in developing the UCP site location model. Also, a brief summary of each explanatory variable is provided along with a graphic illustration of its raster surface (for descriptive statistics see Appendix A). Raster surfaces for 3 of the explanatory variables (*Curvature*, *Northness,* and *Eastness*) are not provided because they are not visually useful.

Table 2: Preliminary Explanatory Variables for the UCP Model. Twenty-seven preliminary explanatory variables were identified for this study. The 27 preliminary variables are listed under their respective model groups; measurement units are also provided along with abbreviations that will be used frequently throughout this thesis.

P1(Static) Model Variables	Abbreviation	Measurement Unit
Elevation	ELE	Meters
Slope	Slope	Degrees (0-90°)
Earth Curvature	Curv	1/100th of a Degree
Percent of Bangor Limestone & Hartselle Formation	PerMbh	Percentage (0-100%)
Percent of Monteagle Limestone	PerMm	Percentage (0-100%)
Percent of Pennington Formation	PerMp	Percentage (0-100%)
Percent of Fentress Formation	PerPf	Percentage (0-100%)
Percent of Rockcastle Conglomerate	PerPf	Percentage (0-100%)
Soil Thickness	SoilThick	Inches
Soil Erosion	Erosion	t ha h ha^{-1} MJ^{-1} mm^{-1}
P2 (Dynamic) Model Variables	**Abbreviation**	**Measurement Unit**
Average Potential Volume of Wood Fiber	VolWood	ft^{3}/ac
Annual Solar Radiation	Solar	Wh/m^{2}
Direct Duration of Solar Radiation	DirDur	hrs/yr
Eastness	East	Unitless (range from -1 to1)
Northness	North	Unitless (range from -1 to1)
Shelter Index at 100meters	SI100	m^{3}
Shelter Index at 300meters	SI300	m^{3}
Shelter Index at 1000meters	SI1000	m^{3}
Terrain Texture	TerTex	m^{2}
Cost Distance to Chestnut Oak	CDChest	Minutes
Cost Distance to Northern Red Oak	CDNred	Minutes
Cost Distance to Southern Red Oak	CDSred	Minutes
Cost Distance to Scarlett Oak	CDScar	Minutes
Cost Distance to White Oak	CDWhite	Minutes
Cost Distance to Hickory	CDHick	Minutes
Cost Distance to Walnut	CDWalnut	Minutes
Cost Distance to Water	CDWater	Minutes

Elevation. Elevation was included as a preliminary P1 model variable because rock shelters on the UCP of Tennessee are commonly found within the same elevation ranges. The study area lies within 9 topographic quadrangles of Fentress and Pickett counties, Tennessee: Burrville, Grimsley, Jamestown, Moody, Pall Mall, Riverton, Sharp Place, Stockton, and Wilder. Ten meter Digital Elevation Models (DEMs) were downloaded for each quadrangle and mosaicked together to make a single continuous elevation surface (Figure 6).

Slope and Curvature. Two other P2 model variables include slope and curvature. These variables were included because the locations of rock shelters exhibit specific characteristics of the landscape. Gorge shelters (instead of upland shelters) like the ones in this study are commonly found in areas with a higher degree of slope than the rest of the landscape. Further, it is possible that a specific type of landform curvature (convex vs. concave surfaces) would help identify where rock shelters naturally form. Slope and Curvature tools available in ArcMap 10 (ESRI 2011) were used to generate raster surfaces from the mosaicked DEMs. Both tools calculate values on a cell-by-cell basis using the 8 surrounding cells (a 9-by-9 rectangle neighborhood). The Slope tool calculates the rate of change in elevation values for a given surface, either in degrees or percent rise (Figure 7). Curvature is calculated by taking the second derivative of the surface, or the slope-of-the-slope. A positive value indicates an upwardly convex surface (e.g. a hill or mound), and a negative value indicates an upwardly concave surface (e.g. a depression). The curvature units are expressed as one hundredth (1/100) of the corresponding z-unit—in this case, the z-unit is a degree.

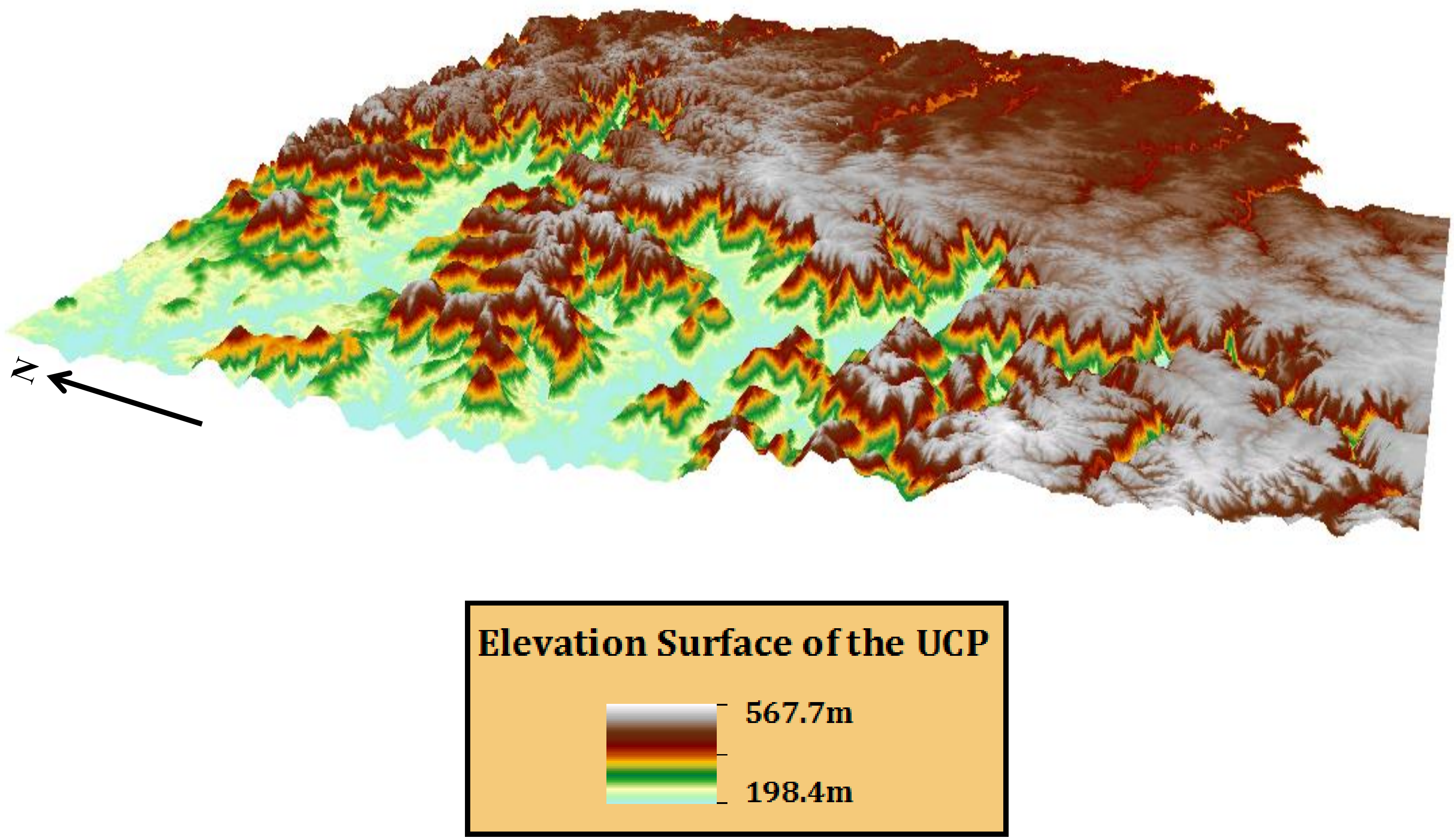
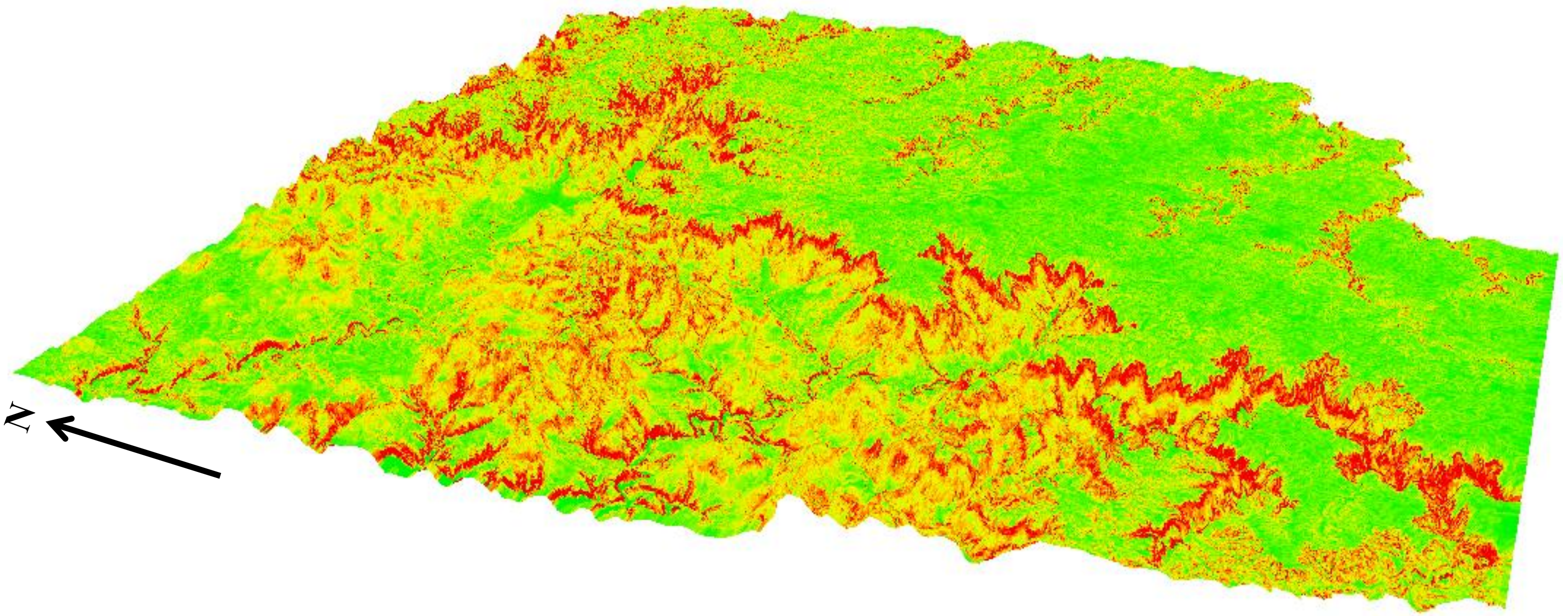


Figure 5: Raster Surface of the Elevation P1 Model Variable. A mosaicked elevation surface for the study area using digital elevation models from 9 topographic quadrangles in Fentress and Pickett Counties, Tennessee—tilted and rotated with a vertical exaggeration of 3 applied to show relief. The study area is approximately 34km wide and 41km long (see Figure 4 for scale).



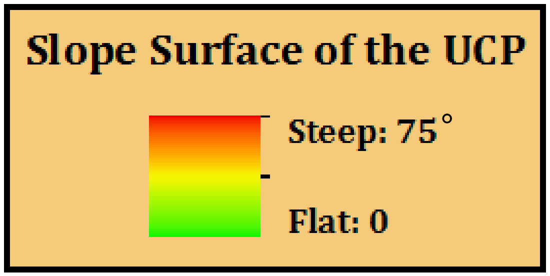


Figure 6: Raster Surface of the Slope P1 Model Variable. The mosaicked elevation surface (see Figure 5) was used to generate a slope surface for the study area. The areas with the highest degree of slope (in red) are where the plateau surface drops off into the deep gorges and ravines; this is characteristic of the western escarpment portion of the UCP and where a majority of rock shelters are found.

Geology. Five of the preliminary P2 model variables relate to the geology of the UCP of Tennessee. Rock shelters generally occur in specific geologic units and these variables will most likely be powerful predictors in isolating where rock shelters (site or non-site) might be located. Geologic formations on the UCP range from sandstone conglomerates to shale to limestone (Table 3). Most of the rock shelters in the study area occur in the sandstone conglomerate types, though some are found in shale and limestone. The relationship between rock shelter occurrence and geologic formation is of interest here.

The Tennessee geology polygon layer was clipped in order to isolate only the study area. Then the polygon layer was converted to a raster using the formation name as the ID for each cell. This categorical layer would normally be included in a predictive model as is since it represents classes or categories of a specific geologic formation (Figure 7). However, the spatial logistic regression function in R (R Core Team 2012) is unable to handle categorical rasters. To convert categorical rasters into usable variables, percentage rasters were created for each class. Using the Reclassify tool, a Boolean raster was made for each geologic formation where 1 equaled the formation of interest and 0 equaled the other formations. The raster was then multiplied by 100 so that each raster would represent a percentage. The Focal Statistics tool was used to calculate the mean of a 3-by-3 rectangle neighborhood around each cell. The resulting raster represented the percent of a specific geologic formation found in each cell using a 3-by-3 neighborhood (Figure 8). This method best represents the original vector data and uses the same cell resolution as the other data sets.

Table 3: Descriptions of the Common Geologic Units present on the UCP. The descriptions were taken from the East-Central sheet of the geologic map of Tennessee ((Hardeman, Miller, and Swingle 1966). Additional information specific to the project area was added from (Wilson, Jr., Jewell, and Luther 1956).

Formation Name	Brief Description
Rockcastle Conglomerate (Pr)	Conglomeratic sandstone and sandstone, gray to brown, fine-coarse-grained. Thin coal bearing shale locally present near middle. Thickness 150-220 feet.
Fentress Formation (Pf)	Mostly dark-gray to light-brown shale, with minor siltstone and sandstone. Wilder Coal near middle. Laterally equivalent to entire Gizzard Group and all of Crab Orchard Mountains Group below Rockcastle Conglomerate. Thickness as much as 340 feet. The name “Fentress Formation” is used only where the Sewanee Conglomerate and other recognizable constituent formations are not mappable—for the UCP, this means the northwestern portion of the study area.
Sewanee Conglomerate (Pco)	Conglomeratic sandstone and sandstone, gray to brown, fine- to coarse-grained. Thickness as much as 200 feet, average about 100 feet. One of the most consistent units of the Fentress Formation on the Cumberland Plateau (except in the northwest where it is almost completely absent).
Pennington Formation (Mp)	Reddish and greenish shale and siltstone; fine-grained dolomite; dark-gray limestone; and thin-bedded sandstone. Persistent dolomite bed at base. Thickness 150-400 feet.
Bangor Limestone & Hartselle Formation (Mbh)	Bangor Limestone: Dark brownish-gray limestone, thick-bedded. Thickness 70-400 feet. Hartselle Formation: Thin-bedded, fine-grained sandstone interbedded with gray shale; with oolitic and coarse-grained limestone beds locally. Thickness 0-80 feet.
Monteagle Limestone (Mm)	Mainly fragmental and oolitic, light-gray limestone; blocky bryozoan chert weathers from base. Thickness 180-300 feet.
St. Louis Limestone & Warsaw Limestone (Msw)	St. Louis Limestone: Fine-grained, brownish-gray limestone, dolomitic and cherty. Thickness 80-160 feet. Warsaw Limestone: Mainly medium- to coarse-grained, gray limestone, crossbedded. Includes much calcareous sandstone and shale to the north. Thickness 100-130 feet.
Fort Payne Formation (Mfp)	Calcareous and dolomitic silicastone; contains bedded chert, cherty limestone, and shale: scattered crinoidal limestone lenses. Thin green shale (Maury) at base. Thickness 100-275 feet.

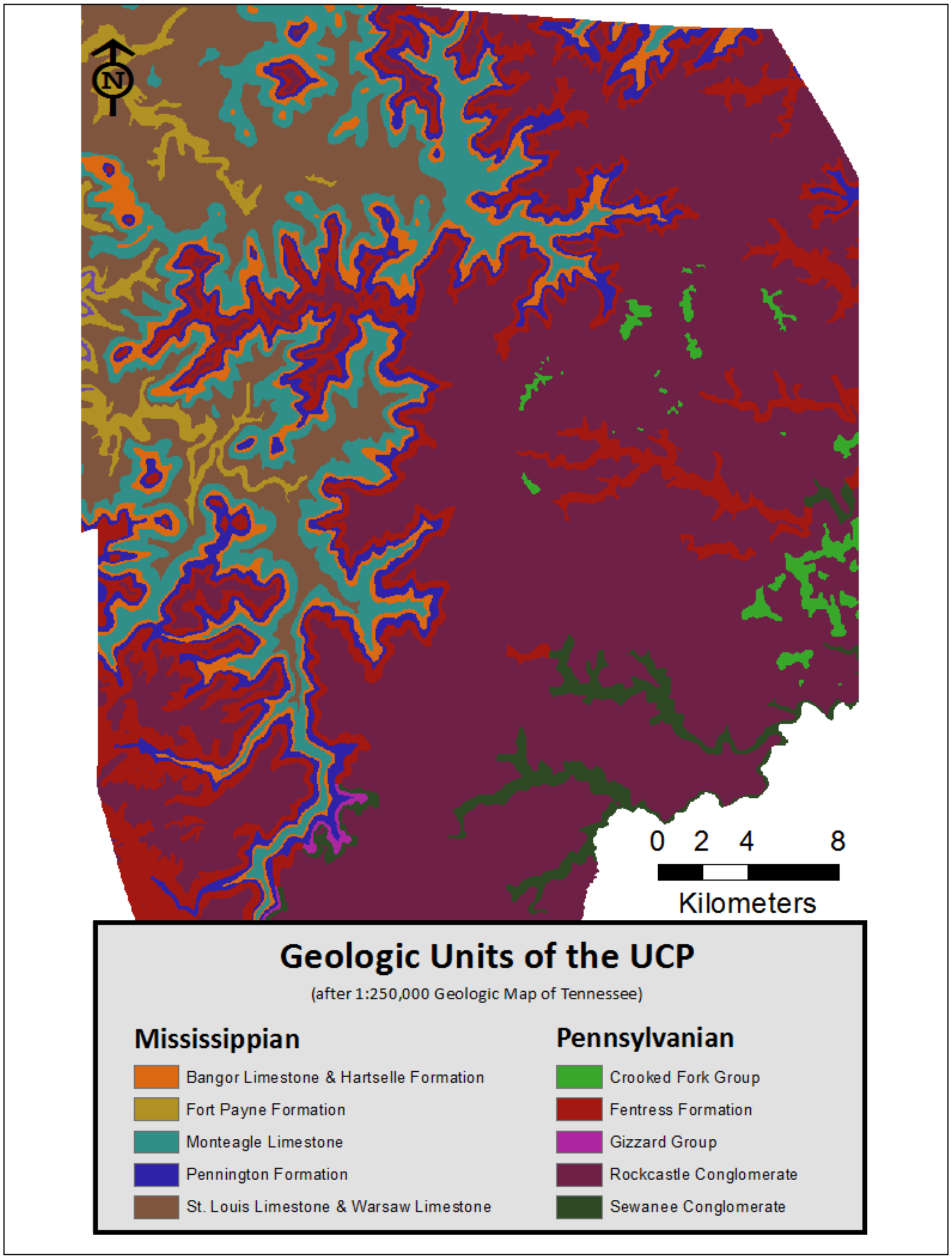


Figure 7: Geology of the UCP of Tennessee. The information in this map is based solely on the GIS data. Rock shelters mainly occur in the Rockcastle Conglomerate and Fentress Formation though the Sewanee Conglomerate is somewhat exposed in the southern portion of the study area (modified after Hardeman, Miller, and Swingle 1966).

0	0	0	0	1
0	0	0	1	1
0	0	0	1	1
1	1	1	1	1
1	1	1	1	1

A

0	0	0	0	100
0	0	0	100	100
0	0	0	100	100
100	100	100	100	100
100	100	100	100	100

B

0%	22%	55%
33%	55%	77%
66%	77%	88%

C

Figure 8: Converting Boolean Rasters into Percentage Surfaces Using Geologic Formations. The geologic variables could not be included in the model as categorical variables and were thus converted to percentage rasters. This process involves 3 main steps. First, a binary raster was created for each geologic formation where 1 equaled the formation of interest and 0 equaled the other formations (A). Then the raster was multiplied by 100 using the Raster Calculator (B). Finally, the Focal Statistics tool was used to calculate the mean of a 3-by-3 rectangle neighborhood around each cell creating a raster that represents the percent of a specific geologic formation found in each cell (C). In a percentage raster such as this, most cells equal either 100% or 0%. However, the boundaries of each formation are captured by increasing and decreasing percentage values as see in C.

This process was executed for each geologic formation. Most of the cells in each raster equaled either 0% or 100% (indicating complete absence or complete coverage). However, the formation boundaries were captured by decreasing and increasing percentages. A total of 5 variables were created using the process outlined above: *Percent of Bangor Limestone and Hartselle Formation (Mbh)*, *Percent of Monteagle Limestone (Mm)*, *Percent of Pennington Formation (Mp)*, *Percent of Fentress Formation (Pf)*, and *Percent of Rockcastle Conglomerate (Pr)*.Figure 9 shows the raster surfaces for all 5 geologic variables; although each looks binary, they are continuous surfaces as demonstrated by Figure 8.

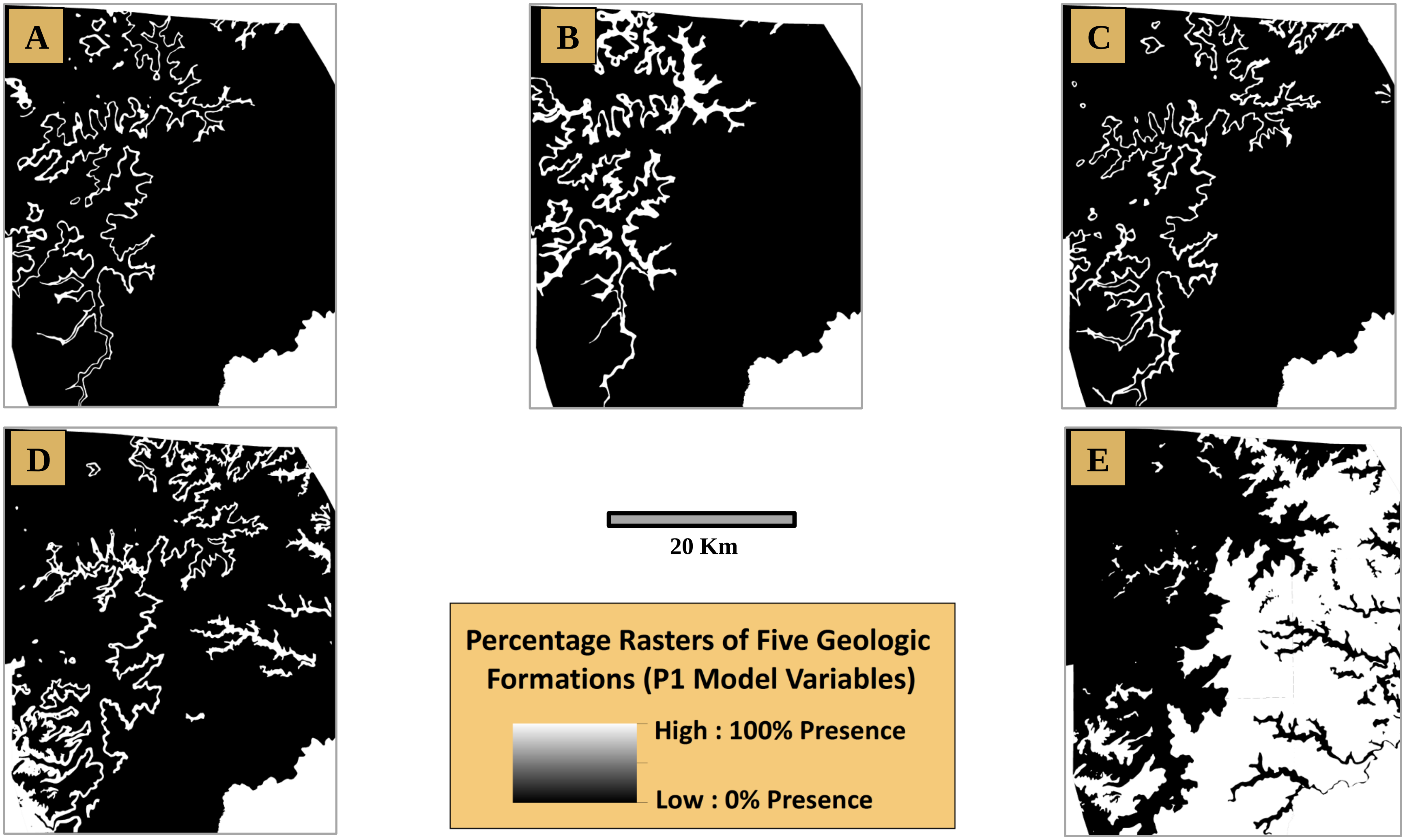


Figure 9: Raster Surfaces of the Five Geologic P1 Model Variables. (A) Bangor Limestone & Hartselle Formation; (B) Monteagle Limestone; (C) Pennington Formation; (D) Fentress Formation; and (E) Rockcastle Conglomerate. Though these surfaces are not binary, they can be viewed as such—the white represents where the formation is present and the black represents the presence of other geologic units.

Soil. Two soil surveys have been conducted in the study area and information from these were acquired from Soil Survey Geographic (SSURGO) Databases (Soil Survey Staff 2009, 2011). Both spatial and tabular data were downloaded: the soil survey polygon layers and the accompanying National Soil Information System relational databases. The 2 soil polygon layers were first merged to create 1 shapefile, creating a GIS layer with 6,848 polygons representing 74 different soils series. The accompanying databases provided information on the mapped soil series and their various properties. For this project, soil data were used as a proxy for generating model variables that might be important for isolating where rock shelters naturally form and for identifying resources that might have been important in prehistoric rock shelter selection. Tables for physical soil properties and forestland productivity were used to generate 2 P1 model variables, *Average Soil Thickness* and *Potential for Soil Erosion*, and 1 P2 model variable, *Average Potential Volume of Wood Fiber*.

The physical soil properties table includes measurements of soil depth and erosion. Soil depth is indicated by the upper (surface of the layer) and lower (restrictive layer or bedrock) boundaries of each soil series. The thickness of a soil series may indicate where rock shelters would be located because a thinner series indicates near-surface or exposed bedrock. The erosion factor Kw indicates the erodibility of the soil; the estimated Kw values range from 0.02 to 0.69 where the higher values indicate increased vulnerability to erosion by water (Soil Survey Staff 2009, 2011). This indicates that rock shelters might tend to occur in areas with less potential for soil erosion because of the absence of floodplains or terraces.

The Forestland Productivity table is meant to aid forestland owners and managers by reporting the estimated potential productivity of each soil for wood crops (Soil Survey Staff 2009, 2011).The potential volume of wood fiber for each soil is based on the “important” tree species present and is expressed as cubic feet per acre per year. Because the number of tree species varies by soil, the average volume of wood fiber was estimated. The potential for wood fiber in an area might have been important to prehistoric peoples as a resource for gathering wood.

New fields for soil thickness, soil erosion (Kw factor), and average potential volume of wood fiber were added to the attribute table of the soil polygon layer; values for each were added by soil series. A raster surface was created for each of the 3 new fields using the Polygon to Raster tool. Thus 3 more variables were created: *Soil Thickness* (Figure 10), *Soil Erosion,* (Figure 11), and *Average Potential Volume of Wood Fiber* (Figure 12).

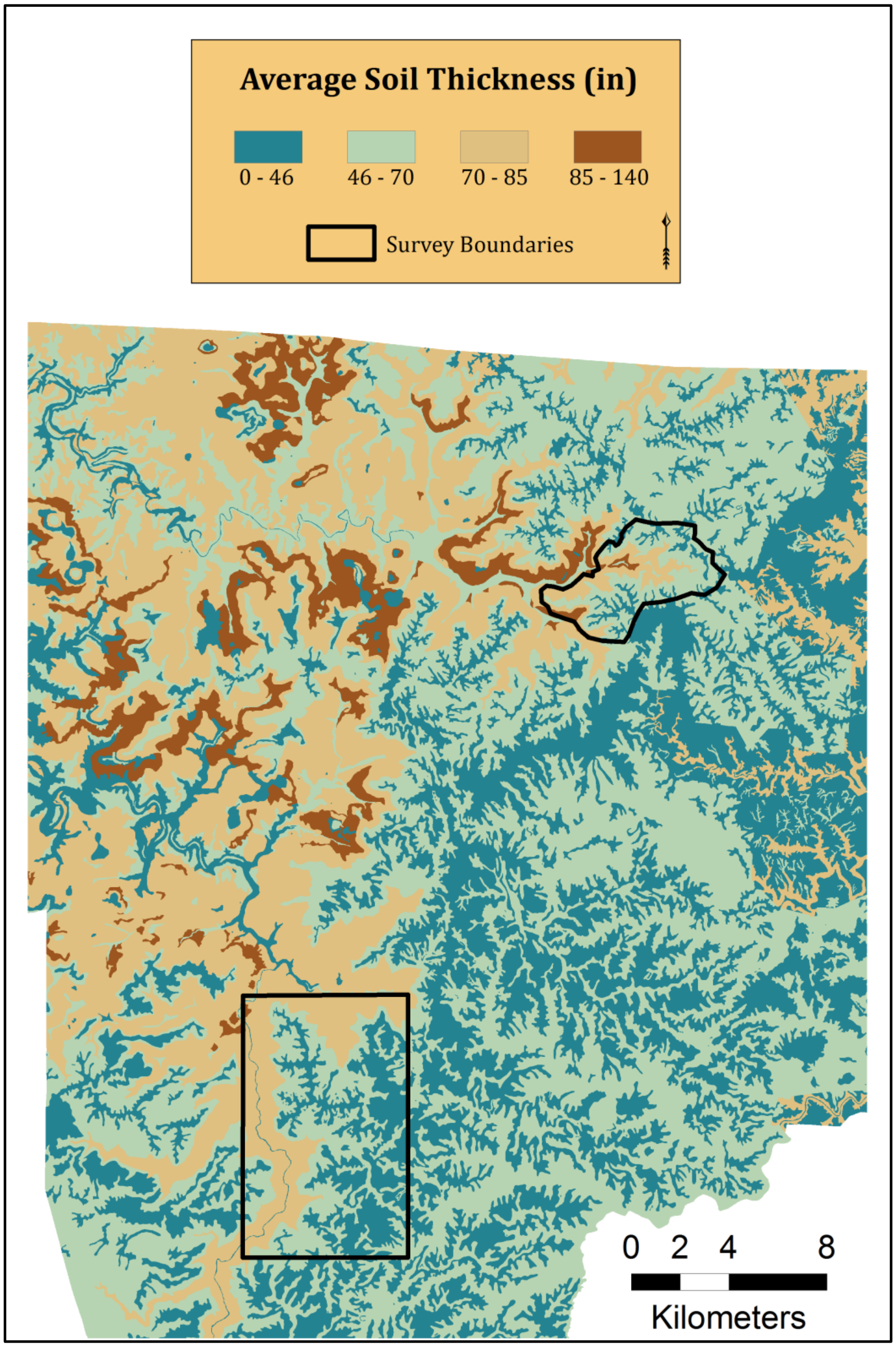


Figure 10: Raster Surface of the Average Soil Thickness P1 Model Variable.This raster surface represents the average soil thickness of 74 different soil series on the UCP of Tennessee. The thickness series are located in the bottom of the ravines/gorges close to river terraces (though the rivers appear blue in this raster surface because water has a average thickness of 0). The thinnest series then are located on the top of the plateau where bedrock may be near surface or exposed.

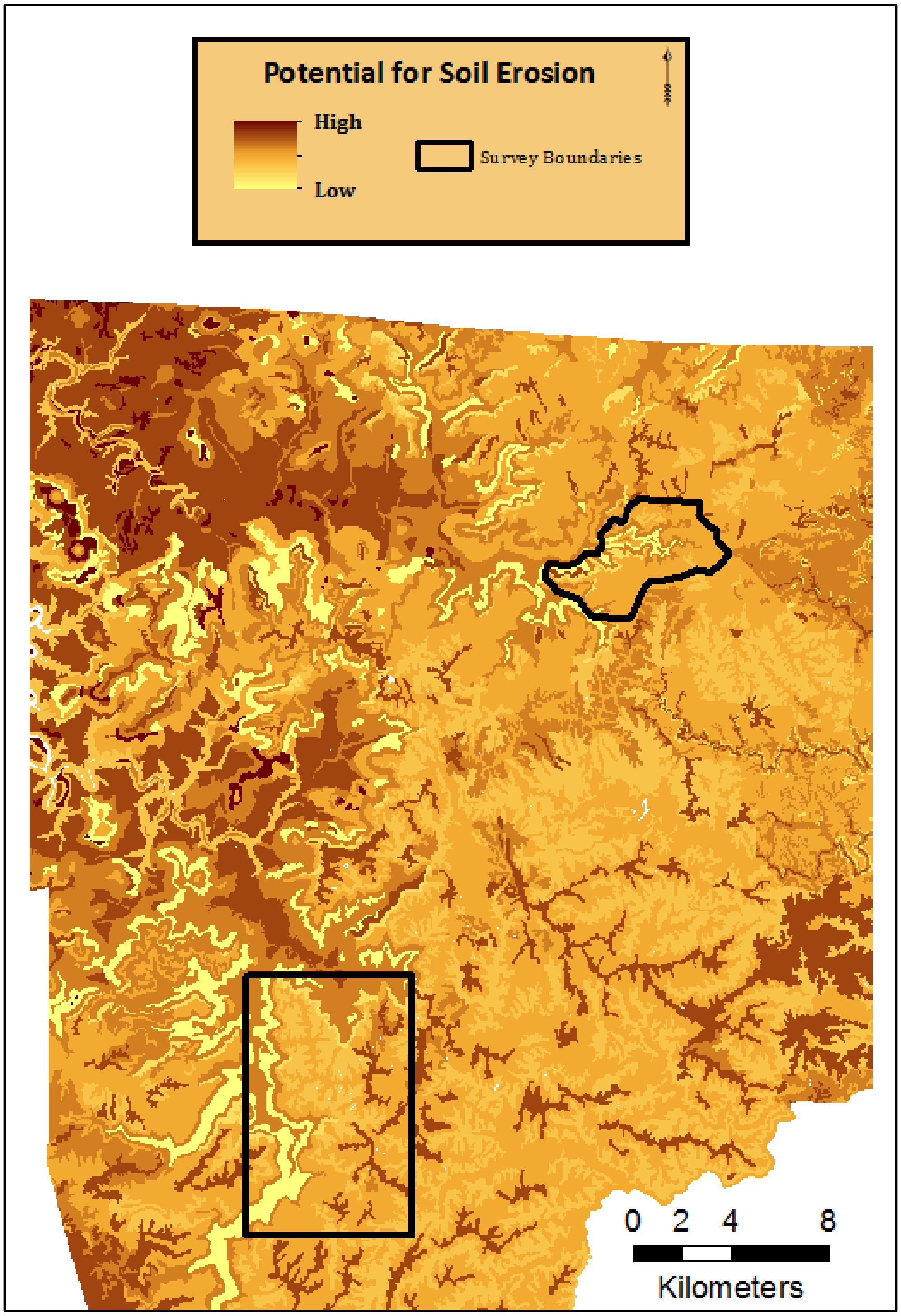


Figure 11: Raster Surface of the Potential for Soil Erosion P1 Model Variable. The potential for soil erosion is highest in floodplain/river terrraces such as those found in the bottom of the gorge and on the upper portions of plateau. Areas around the bluff lines in the gorges have the lowest potential for soil erosion. Areas of "no data" are displayed in white.

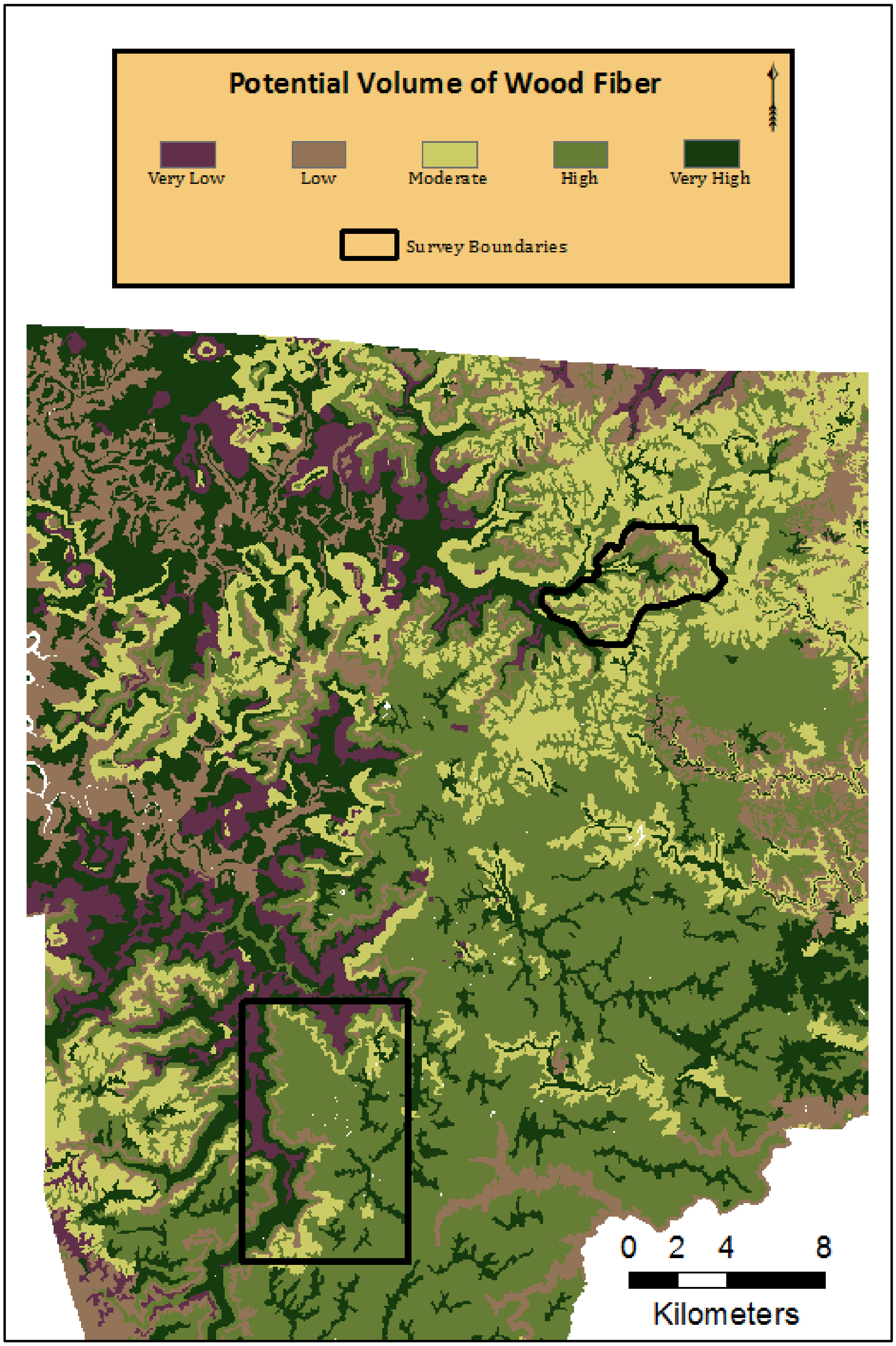


Figure 12: Raster Surface of the Average Potential Volume of Wood Fiber P2 Model Variable. The highest potential volume of wood fiber occurs in and around the top of the gorges and on the top of the plateau. Areas along the bottom of the gorge (though not the river terraces) have the lowest potential. Areas of "no data" are displayed in white.

Solar Radiation. As discussed in the introduction of this thesis, the amount of solar radiation a location receives has been suggested as a possible factor contributing to differential site selection. Therefore, it is important to incorporate variables reflecting solar radiation into the model. The Solar Radiation toolset provides tools for performing solar radiation analysis over a geographic area for specified time periods or increments. The Area Solar Radiation tool produces insolation maps for a geographic area by calculating the insolation across an entire elevation surface (ESRI 2011). Several time configuration options are available (i.e., within a day, multiple days in a year, whole year). Also, additional surfaces can be generated such as a Direct Duration raster surface; this raster represents the total duration, in hours, of direct incoming solar radiation. Two variables were created using the Area Solar Radiation tool: *Annual Solar Radiation* and *Direct Duration of Incoming Solar Radiation*.

The *Annual Solar Radiation* variable was generated using the Area Solar Radiation tool in ArcMap 10.0 (ESRI 2011). This tool uses the DEM surface (e.g. the elevation raster surface shown in Figure 5) to calculate the amount of solar radiation a location receives based on geographic location (latitude).The resulting raster surface represents the amount of solar radiation a location receives within a year. The Solar Radiation toolset was also used to generate another raster surface representing the total hours per year that a location receives direct incoming solar radiation and thus the *Direct Duration of Incoming Solar Radiation* variable. This variable was generated in the same way and using the same input data as the *Annual Solar Radiation* variable. These variables reflect 2 ways in which solar radiation can be measured—in energy or time. Both variables were initially included in the model in order to see which might be significant in the P2 model. The rasters surfaces for both solar radiation variables are shown in Figure 13.

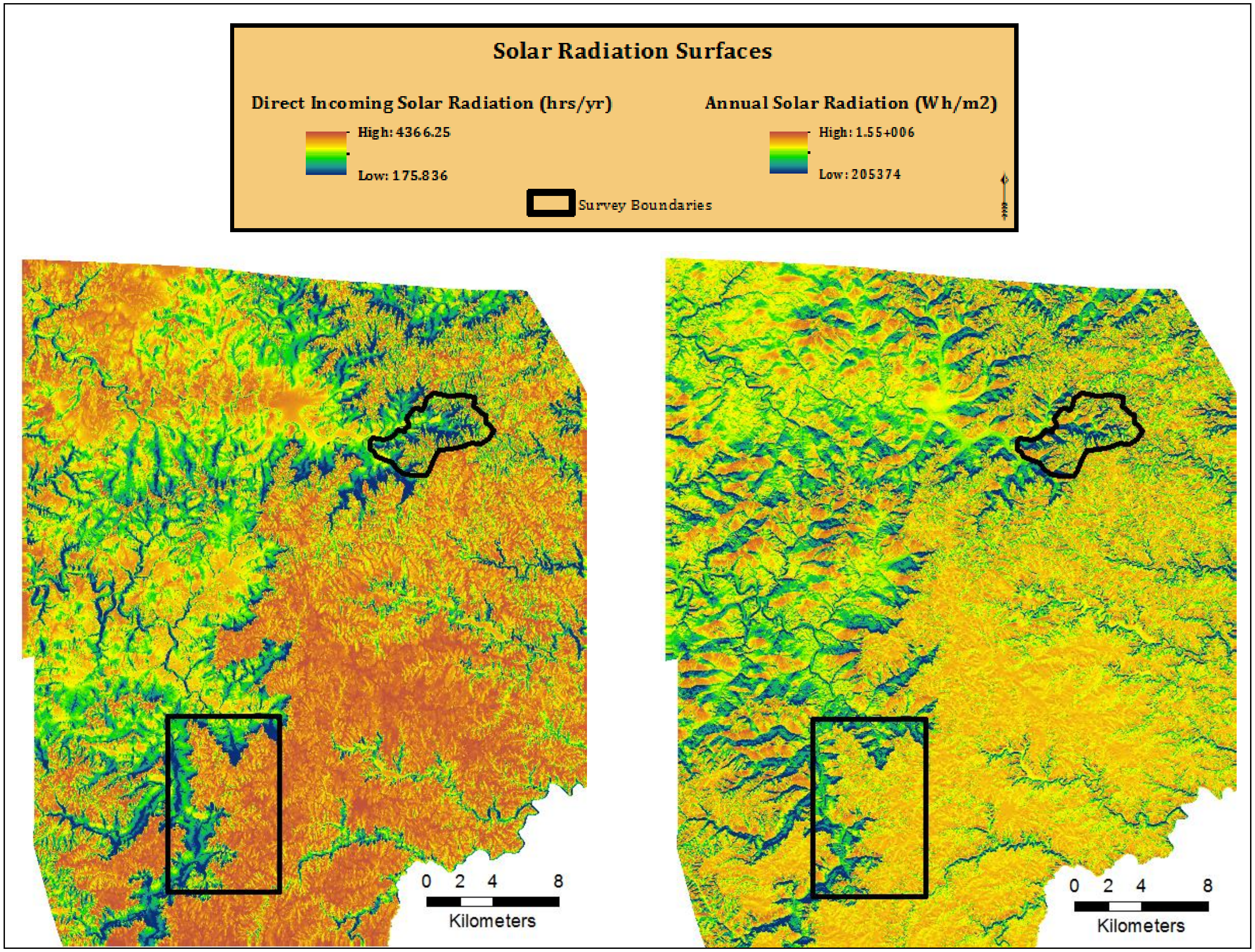


Figure 13: Raster Surfaces of the Solar Radiation P2 Model Variables. The *Annual Solar Radiation* (left) and *Direct Duration of Incoming Solar Radiation* (right) variables both measure the amount of solar radiation a location receives based on its elevation and geographic location (latitude).

Aspect. Another common variable used in archaeological site location modeling is aspect; this variable has also been explored in other studies of prehistoric rock shelter selection (Hall and Klippel 1988; Mickelson 2002; Langston and Franklin 2010). Aspect is the compass direction of the slope and is considered circular data because large values are next to low values (i.e., 359 degrees and 1 degree both represent approximately north). For this reason, aspect values need to be transformed to a linear scale. Aspect can be transformed to an aspect value using trigonometric functions (Hartung and Lloyd 1969: 180; Roberts 1986: 125). Using the elevation surface of the project area, an aspect map was generated using the Aspect tool in the Surface toolset (ESRI 2011). Two aspect value variables were created to measure the amount of "northness" (Equation 6) and the amount of "eastness" (Equation 7) of each location in the project area.

$$\text{Northness} = \cos(\text{aspect angle}) \qquad \text{Equation 6}$$

$$\text{Eastness} = \sin(\text{aspect angle}) \qquad \text{Equation 7}$$

For "northness", values close to 1 represent aspects generally northward, values close to -1 represent southward aspects, and values close to 0 represent either east or west. "Eastness" is very similar with values close to 1 indicating more east-facing slopes, values close to -1 indicating more west-facing slopes, and values close to 0 represent either north or south. The Raster Calculator was used to take the cosine and the sine of the aspect surface in order to create 2 new rasters for the variables of *Northness* and *Eastness*.

Shelter. In an effort to identify cliff dwellings in the southwestern region of the US, Kvamme (1984: 354; 1988: 335–337) developed an index to measure the shelter or exposure of a location. The index (known as the rim, exposure, or shelter index) is generated by passing an imaginary cylinder over an elevation surface, where the height is set at 20 meters above the ground surface and the radius depends on the study area (Kvamme 1988: 335–337). The computed volume of the cylinder provides an index for measuring shelter (Figure 14). For example, a site located on an exposed ridge (Figure 14B) would increase the height and therefore the volume of the cylinder. On the other hand, a site located in a horseshoe-shaped canyon (like those found on the UCP) or a valley (Figure 14A) would decrease the height and volume of the cylinder.

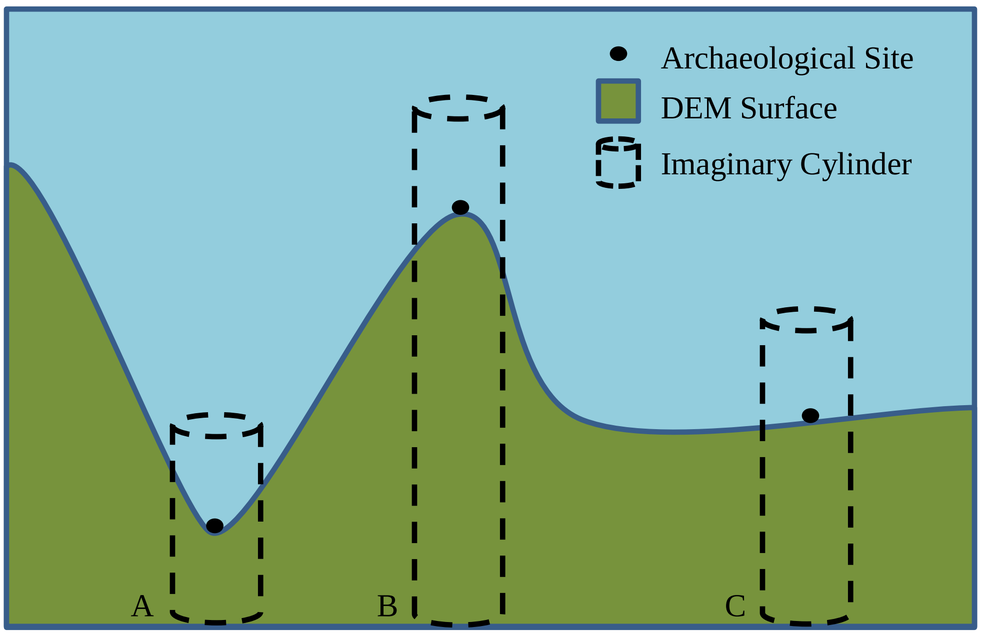


Figure 14: Measuring the Amount of "Shelter" using an Imaginary Cylinder. This figure is an example of how the amount of shelter varies for different topographic positions: (A) an archaeological site located in a valley; (B) an archaeological site located on a hilltop; and (C) an archaeological site located in a flat, open area. The amount of shelter/exposure of these locations is measured by first calculating the volume of an imaginary cylinder over each of the locations. Then, using a digital elevation model (DEM), the volume of the DEM within the cylinder (the green area) is calculated. Because the cylinder is set at a constant height above each of the locations, the amount of shelter/exposure (the blue area) is calculated by subtracting the volume of the DEM within the cylinder (the green area) from the volume of the entire cylinder. (A) Sheltered; (B) Intermediate; and (C) Exposed.

For the Pogue Creek Model, 3 shelter indices were created using 100, 300, and 1000 meter radii to explore the effects of a range of scales from local to regional. Figure 15 outlines the steps executed using the Raster Calculator and Focal Statistics tool to generate the 3 shelter indices: *Shelter Index at 100m, 300m, and 1000m* (Figure 16).

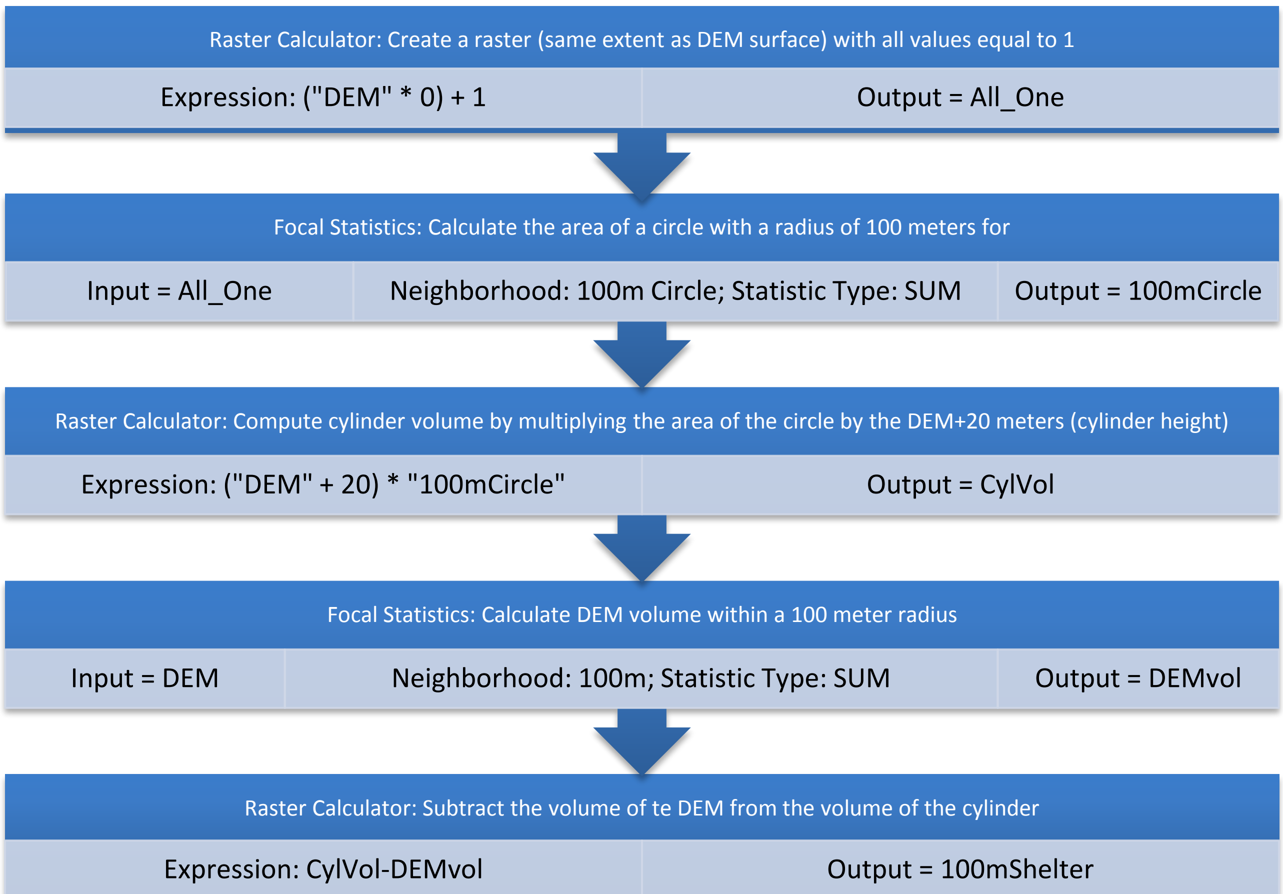


Figure 15: Flowchart for Generating a Shelter Index at 100 meters. This flowchart details the workflow process for generating a Shelter Index with a 100m radius in ArcMap 10 (ESRI 2011). This process was modified after Campbell (2006: 55).

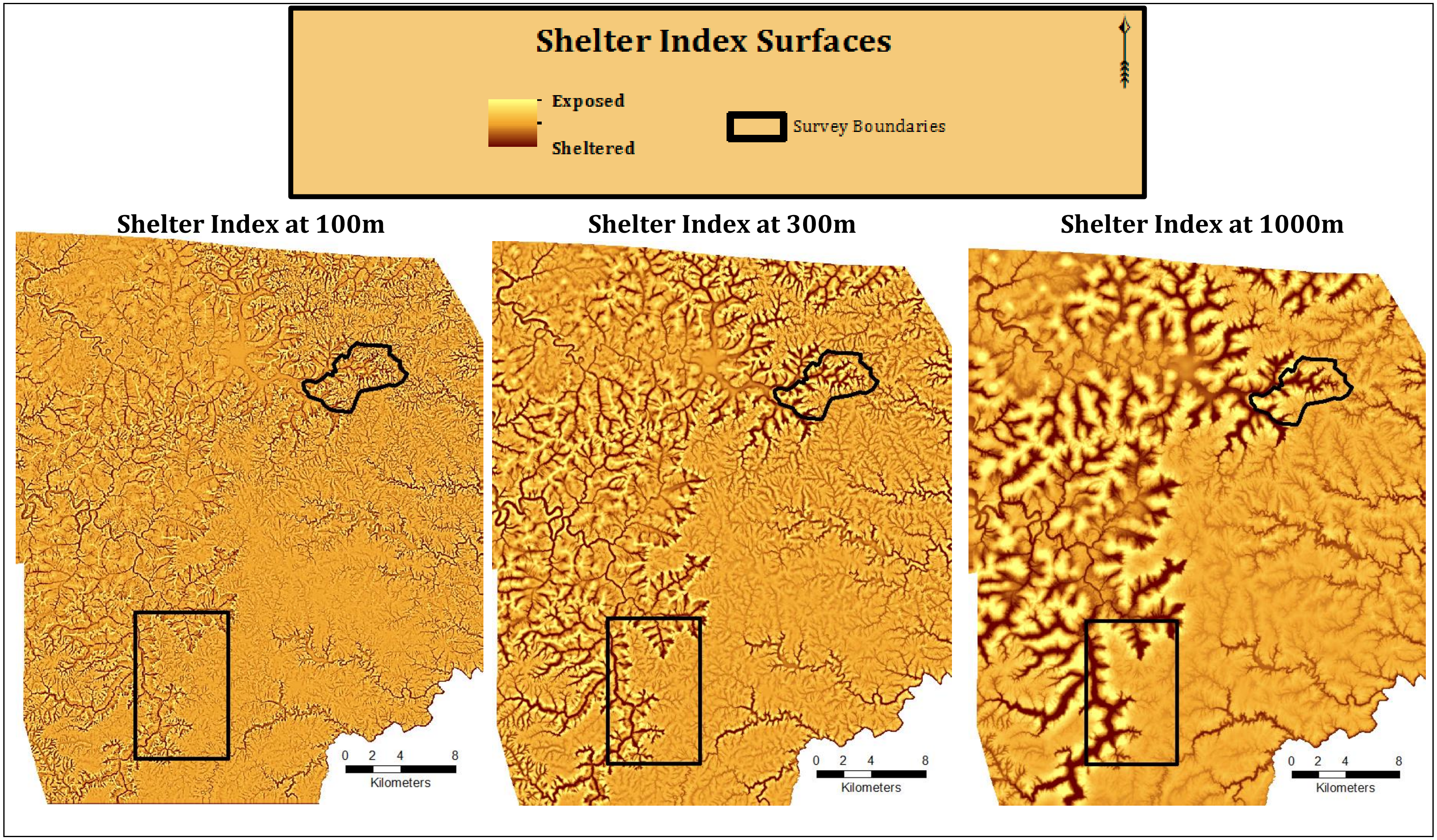


Figure 16: Raster Surfaces of the Three Shelter P2 Model Variables. Three raster surfaces were generated that indicate whether a location is sheltered/exposed when compared to other locations within a given radius. Generating shelter indices using different radii demonstrates the difference in assessing shelter/exposure of a location on a local, intermediate, or regional scale.

Terrain Texture. The variance (σ^2) of elevation within a specified neighborhood can suggest whether a terrain is variable and dissected or if it is more smooth and level (Kvamme 1988: 333–334). High values indicate more variation in the terrain roughness while low values indicate more smooth terrain. Using the study area DEM surface, the Focal Statistics tool was used to calculate the standard deviation of elevation values within a 3-by-3 rectangle neighborhood. The Raster Calculator was then used to square the standard deviation raster and produce an elevation variance surface. The final elevation variance surface represents a measure of terrain texture for every cell in the study area. The raster surface for the *Terrain Texture* variable is shown in Figure 17.

Cost Surface Calculation. Several model variables were generated to represent the "cost" of travelling from one location to another on foot. On the UCP, one of the main factors affecting mobility across the landscape is slope. Prehistoric hunter-gatherers would have needed to traverse the gorges and plateaus on a daily basis and may have chosen where to live based on ease of access to available resources (e.g. water, food, trails leading out of the gorges). Modeling the effect of slope using cost functions provides a more accurate analysis of the time/distance traveled from one location to another than using Euclidean (straight line) distance alone. The Cost Distance tool calls for a cost raster and source feature layer. The source feature layer is the resource (such as a streams polyline layer) for which the accumulated cost distance is calculated. The cost raster represents the cell-by-cell cost of moving through or past that cell.

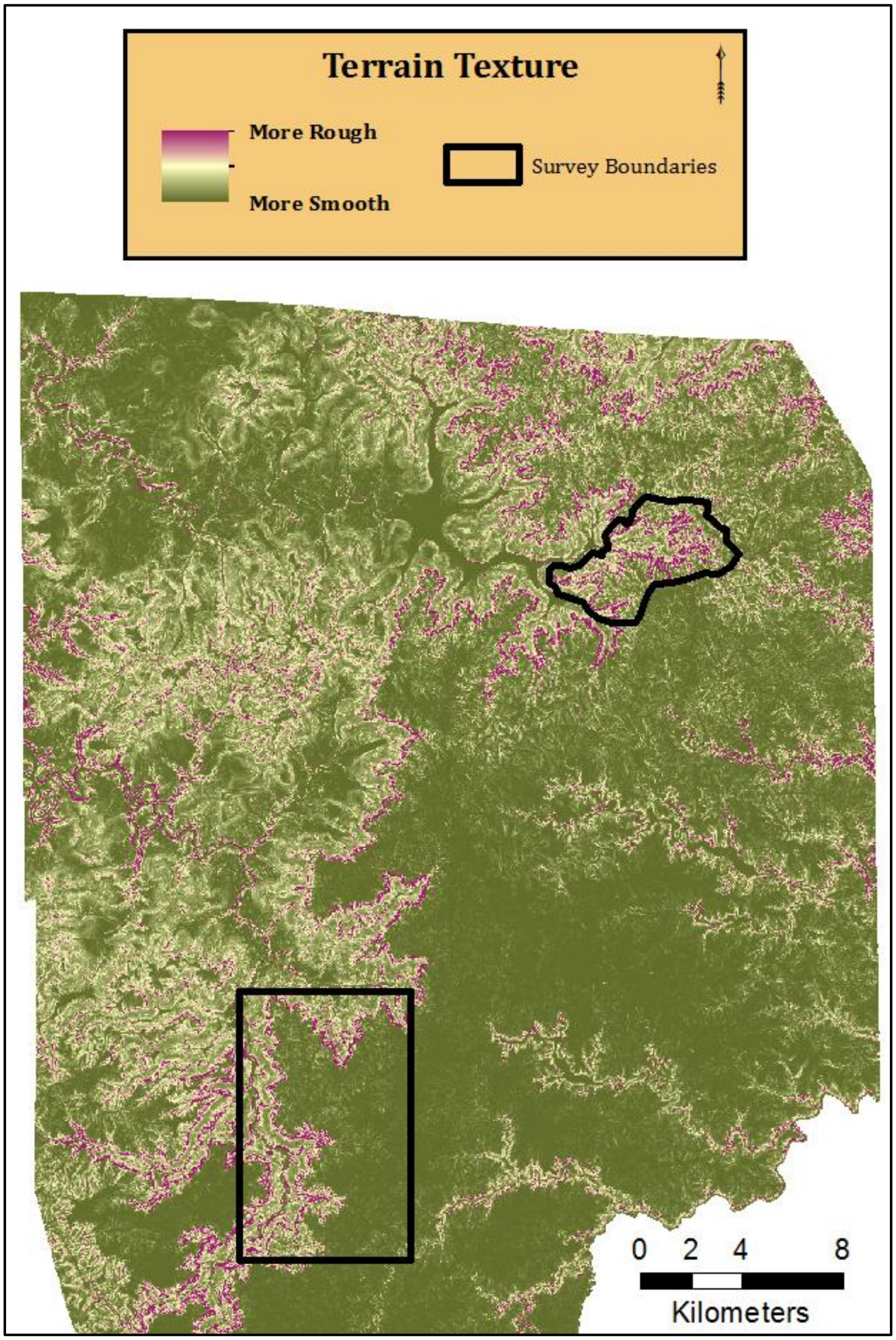


Figure 17: Raster Surface of the Terrain Texture P2 Model Variable. This raster represents the variation in terrain roughness on the UCP based on elevation and extreme changes in relief. While a majority of the study area is indicative of a more smooth terrain, “rough” areas exist around the top of gorge and the edges of the steep escarpment as well as along some of the stream/river channels and drainages.

The slope raster previously created was converted to a cost raster using Gorenflo and Gale's (1990) equation for modeling the effect of slope on the speed of walking (Equation 8).

$$v= 6 \exp \{-3.5 \text{ abs } |S + 0.05|\} \qquad \text{Equation 8}$$

Where:

v= walking speed in km/hr

S= slope (in degrees)

According to Tobler (1993), the estimated average walking velocity for on-path travel is 5km/hr. Off-path travel is calculated by multiplying the walking velocity by $^{3}/_{5}$ (= 0.6). The Raster Calculator was used to insert the slope raster into the above equation. The Raster Calculator was used again to multiply the walking velocity raster by 0.06 which is the conversion rate for kilometers per hour to minutes per meter (1 m/min = .06 km/hr). This was done so that the final cost distance variables would represent the amount of time in minutes required for travel to the source features. The initial slope cost raster however, indicates the walking velocity associated with travelling through that cell (location) given the effect of slope in mountainous terrain; this raster was used to generate cost surfaces for 8 model variables (see Table 2, page 66)

Proximity to Vegetation Zones. Tables listing tree and plant species commonly found in a given soil class were also available in the National Soil Information Databases (Soil Survey Staff 2009, 2011). Modern soil surveys can be used as a proxy for determining food sources that might have been present in prehistoric times. Of relevance to this project are nut and fruit-bearing trees/plants that would have served as food resources for humans and/or animals alike. Depending on their properties and features, different soils can support different tree and plant species. Three genera were identified as potentially significant food resources: *Quercus* (oak), *Carya* (hickory), and *Juglans* (walnut). Five *Quercus* species were present in the study area:

Chestnut Oak, Northern Red Oak, Southern Red Oak, White Oak, and Scarlett Oak. Two species of *Juglans*, *Juglans nigra* (Black walnut) and *Juglans cinerea* (Butternut) occur in the area, though they are not widespread. Lastly, *Carya* was mainly identified at the genus level. "Supporting zones" were determined for the different vegetation types by creating polygon layers for each of the 5 individual *Quercus* species, 1 for *Juglans* species, and 1 for *Carya* species using the soil classes where they commonly occur as a proxy. After creating the polygon layers, cost distance surfaces were generated using the previously discussed slope cost raster. The final cost distance rasters represent the time required to access supporting zones of different species of oak, walnut, and hickory. These zones have the potential to represent a direct (i.e., gathering nuts for human consumption) or indirect (i.e., to hunt game) food resource for prehistoric hunter-gatherers. Thus, 7 more variables were added: *Cost Distance to Supporting Zones of Chestnut Oak* (Figure 19), *Northern Red Oak* (Figure 20), *Southern Red Oak* (Figure 21), *White Oak* (Figure 22), *Scarlett Oak* (Figure 23), *Walnut* (Figure 24), and *Hickory* (Figure 25).

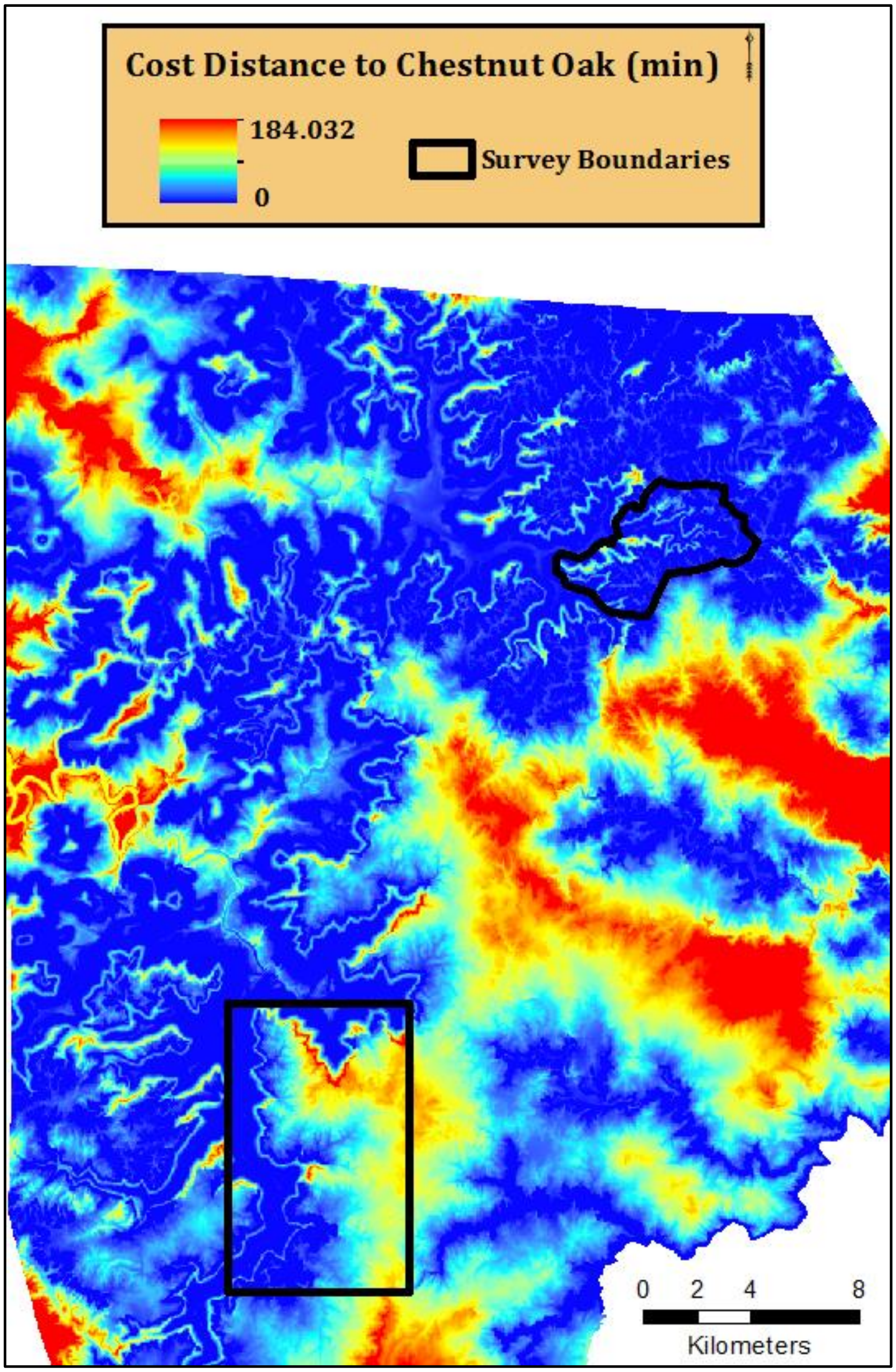


Figure 18: Raster Surface of the Cost Distance to Chestnut Oak P2 Model Variable. This raster surface represents the amount of time it takes to reach a "supporting zone" of Chestnut Oak. The dark blue areas are where the original "supporting zones" are located; as the color ramp progresses from dark blue to yellow to red, the time (and thus cost distance) to reach the boundaries of the "supporting zones" increases.

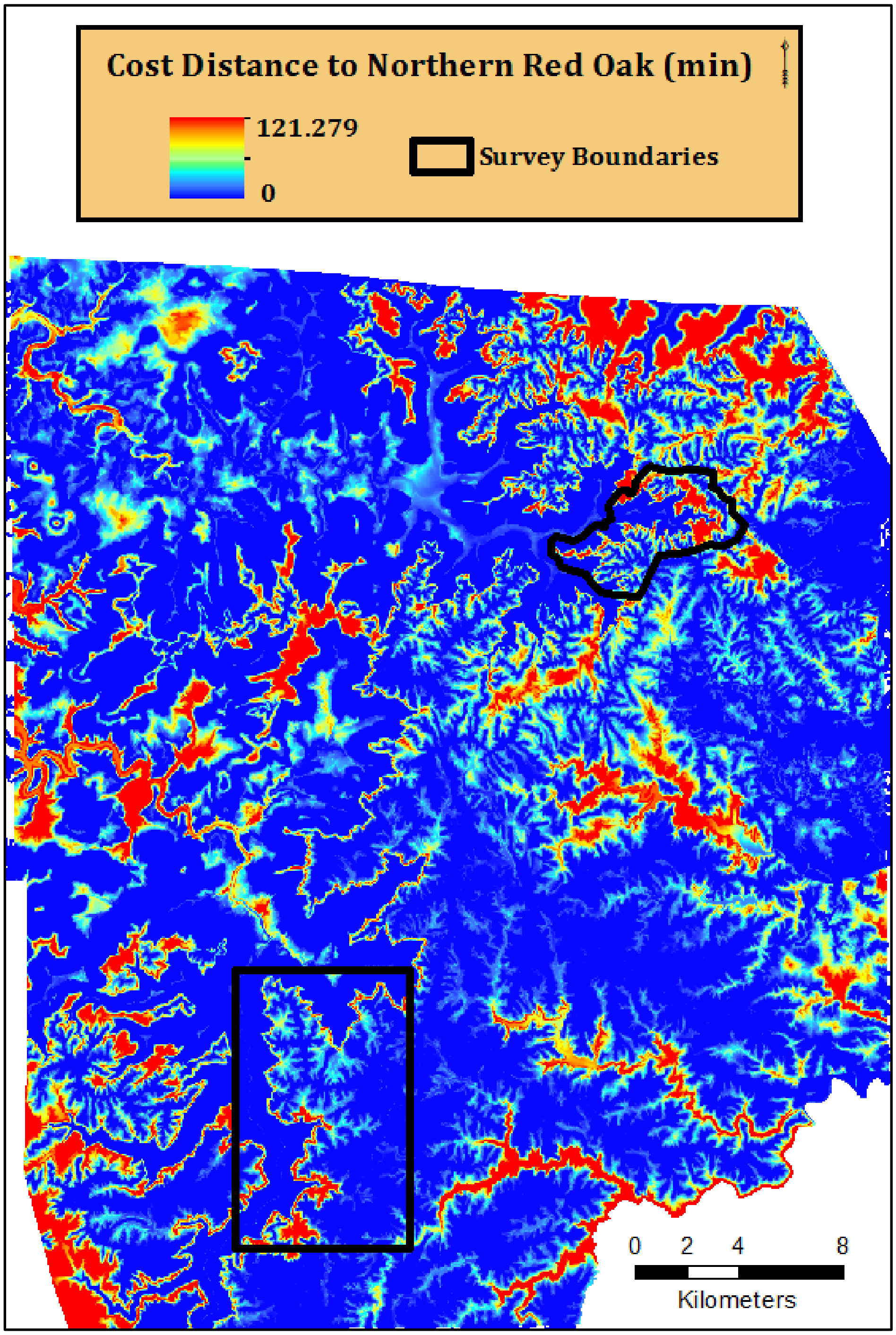


Figure 19: Raster Surface of the Cost Distance to Northern Red Oak P2 Model Variable. This raster surface represents the amount of time it takes to reach a "supporting zone" of Northern Red Oak. The dark blue areas are where the original "supporting zones" are located; as the color ramp progresses from dark blue to yellow to red, the time (and thus cost distance) to reach the boundaries of the "supporting zones" increases. As the raster surface indicates, Northern Red Oak is widespread in the study area.

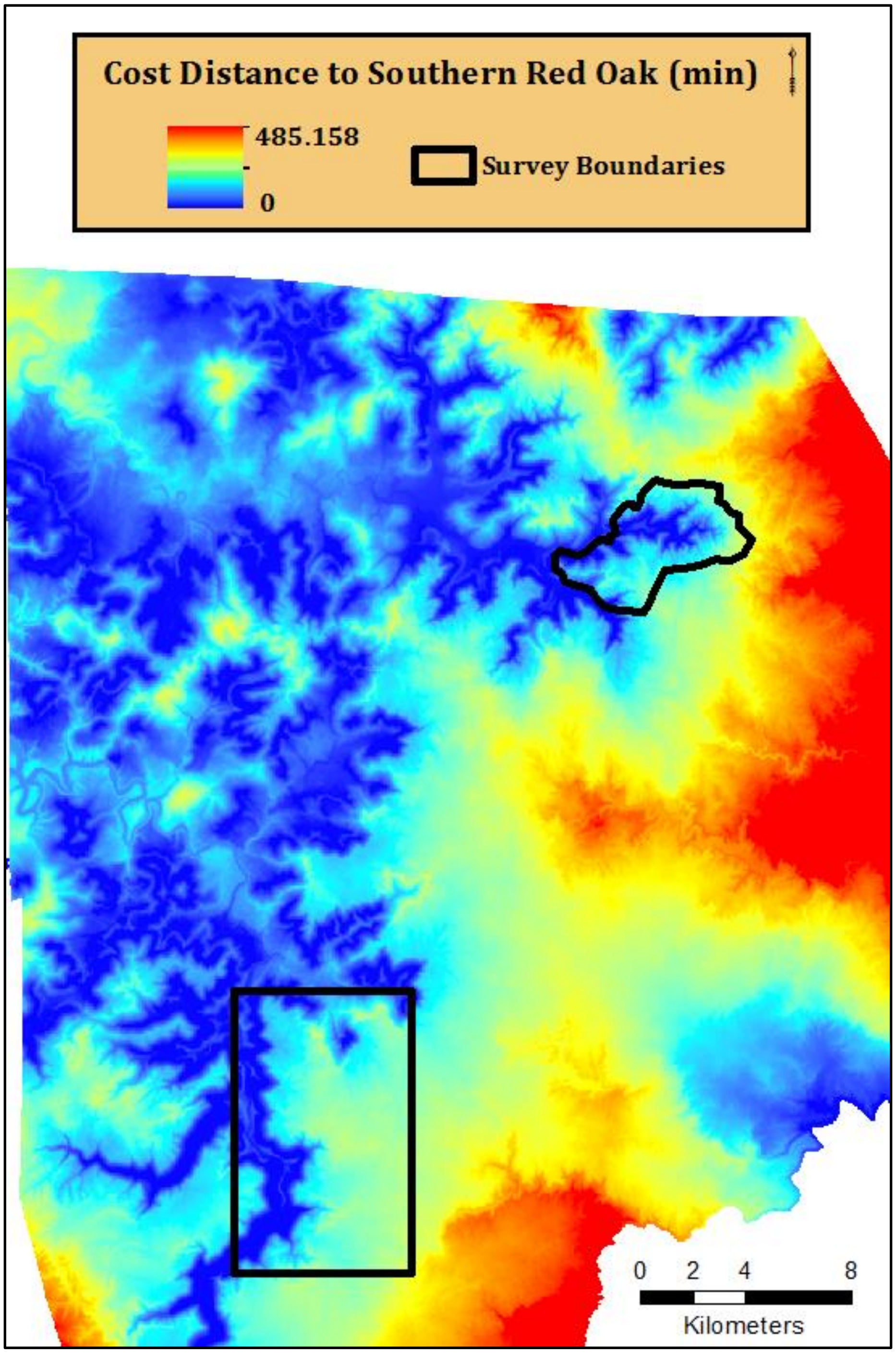


Figure 20: Raster Surface of the Cost Distance to Southern Red Oak P2 Model Variable. This raster surface represents the amount of time it takes to reach a "supporting zone" of Southern Red Oak. The dark blue areas are where the original "supporting zones" are located; as the color ramp progresses from dark blue to yellow to red, the time (and thus cost distance) to reach the boundaries of the "supporting zones" increases. Based on this raster surface, Southern Red Oak appears to be limited to the gorge/ravine bottoms and some portions of the upper plateau area.

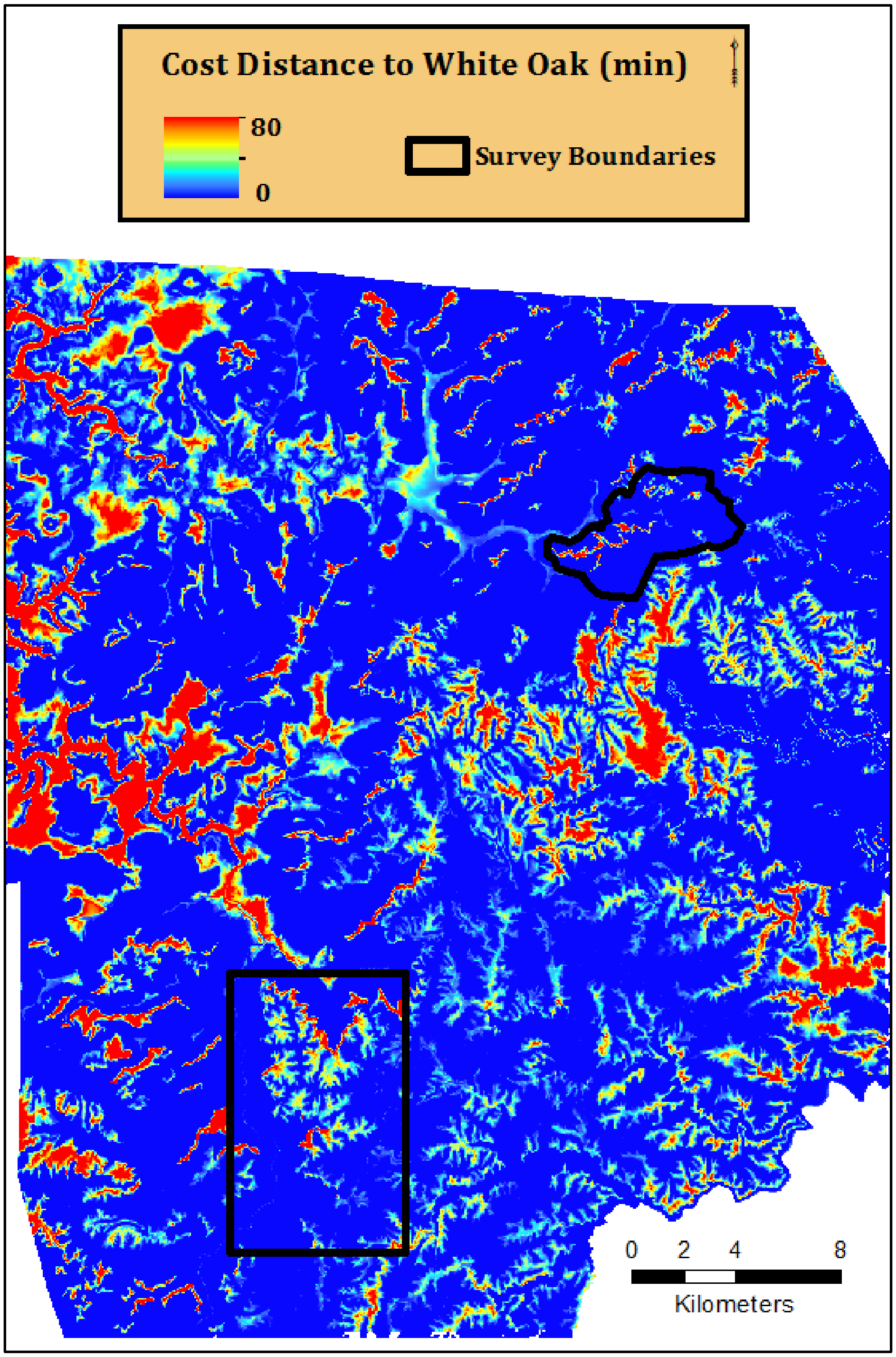


Figure 21: Raster Surface of the Cost Distance to White Oak P2 Model Variable. This raster surface represents the amount of time it takes to reach a “supporting zone” of White Oak. The dark blue areas are where the original “supporting zones” are located; as the color ramp progresses from dark blue to yellow to red, the time (and thus cost distance) to reach the boundaries of the “supporting zones” increases. White Oak is perhaps the most common and widespread Oak species found in the study area, as evident from this raster surface.

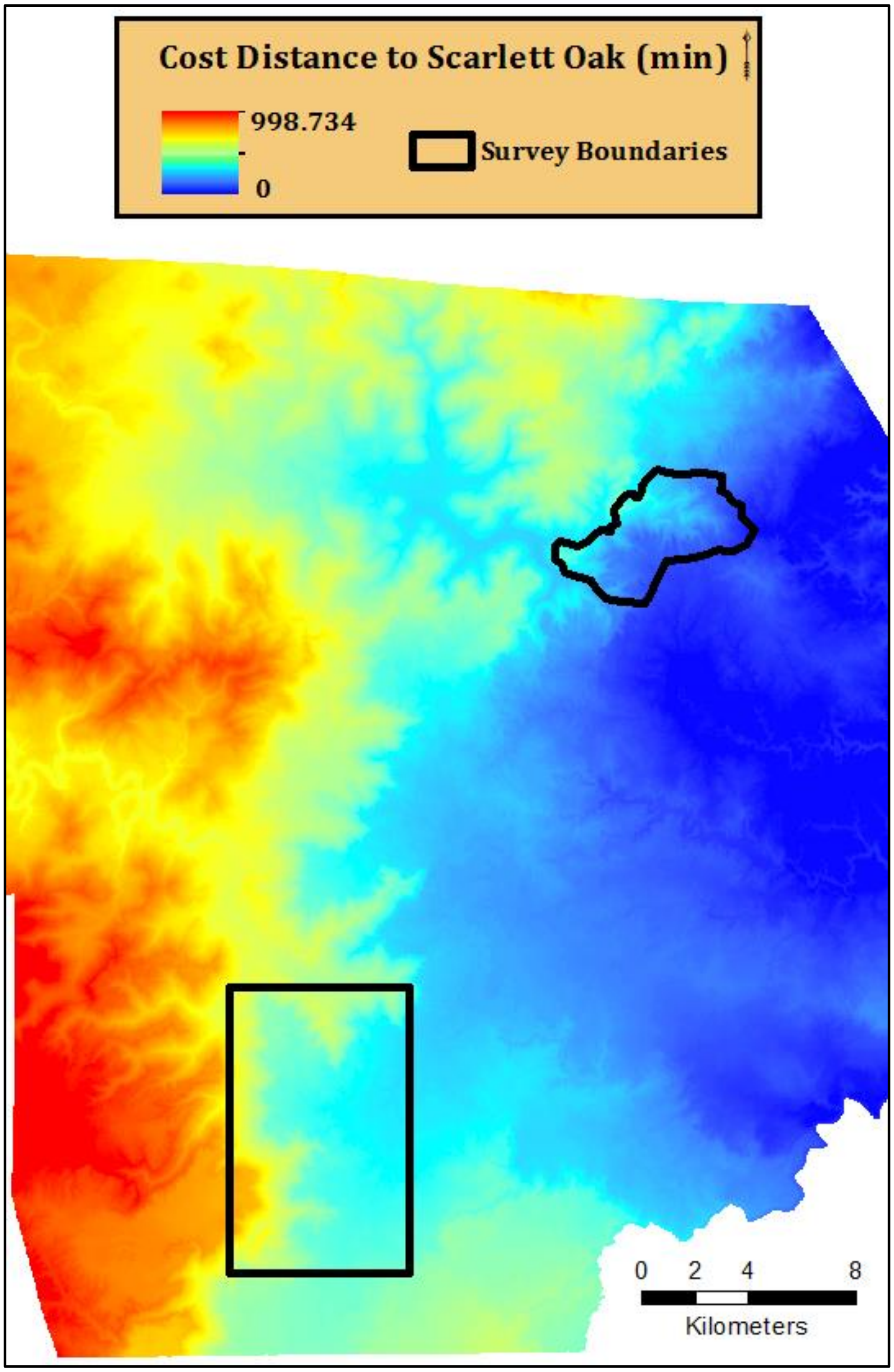


Figure 22: Raster Surface of the Cost Distance to Scarlett Oak P2 Model Variable. This raster surface represents the amount of time it takes to reach a "supporting zone" of Scarlett Oak. The dark blue areas are where the original "supporting zones" are located; as the color ramp progresses from dark blue to yellow to red, the time (and thus cost distance) to reach the boundaries of the "supporting zones" increases. Scarlett Oak appears to be restricted to the eastern portion of the study area in the highest elevations.

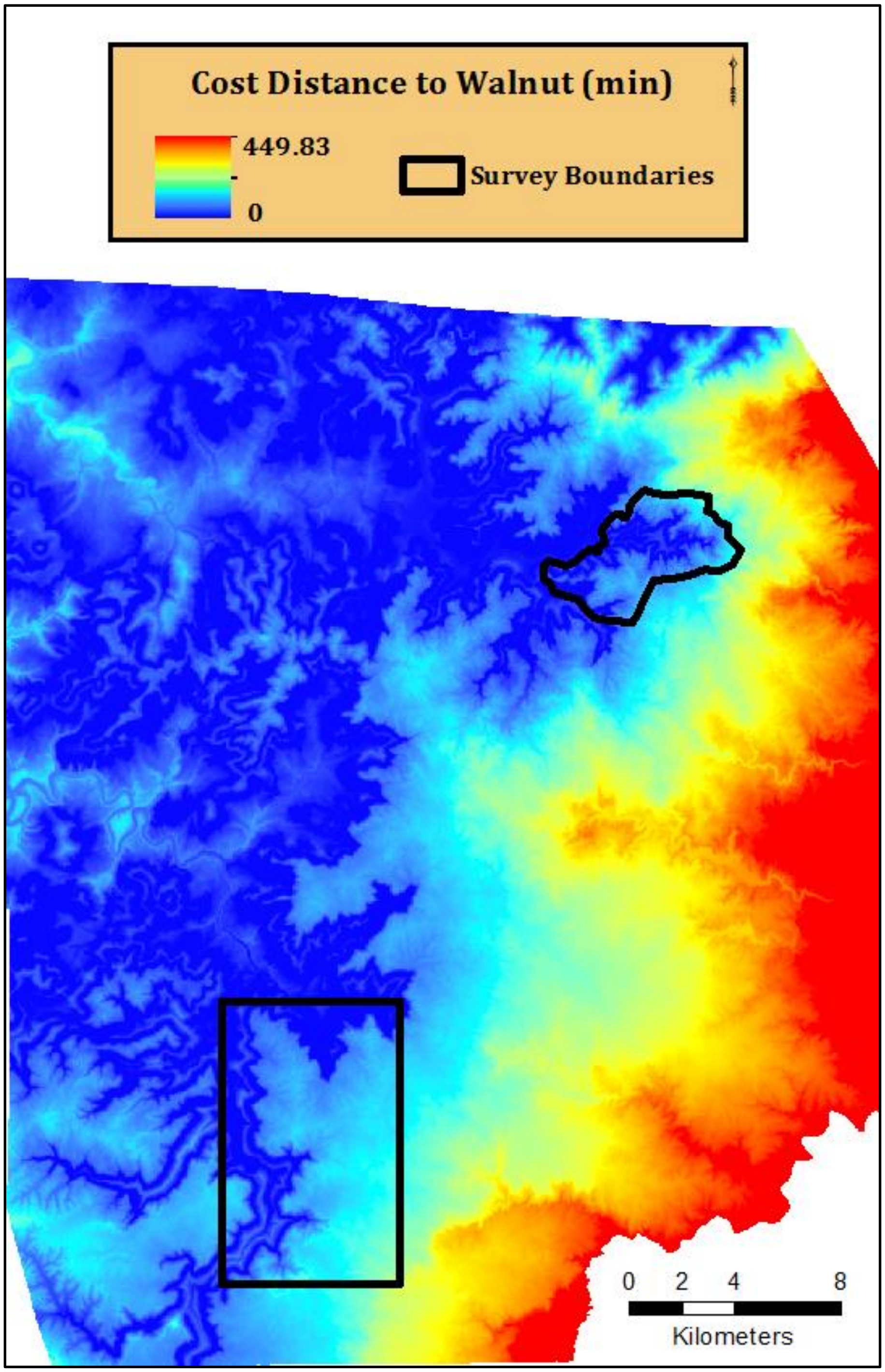


Figure 23: Raster Surface of the Cost Distance to Walnut P2 Model Variable. This raster surface represents the amount of time it takes to reach a "supporting zone" of Walnut. The dark blue areas are where the original "supporting zones" are located; as the color ramp progresses from dark blue to yellow to red, the time (and thus cost distance) to reach the boundaries of the "supporting zones" increases. Walnut is widespread in the western portion of the study area around the Cumberland Escarpment.

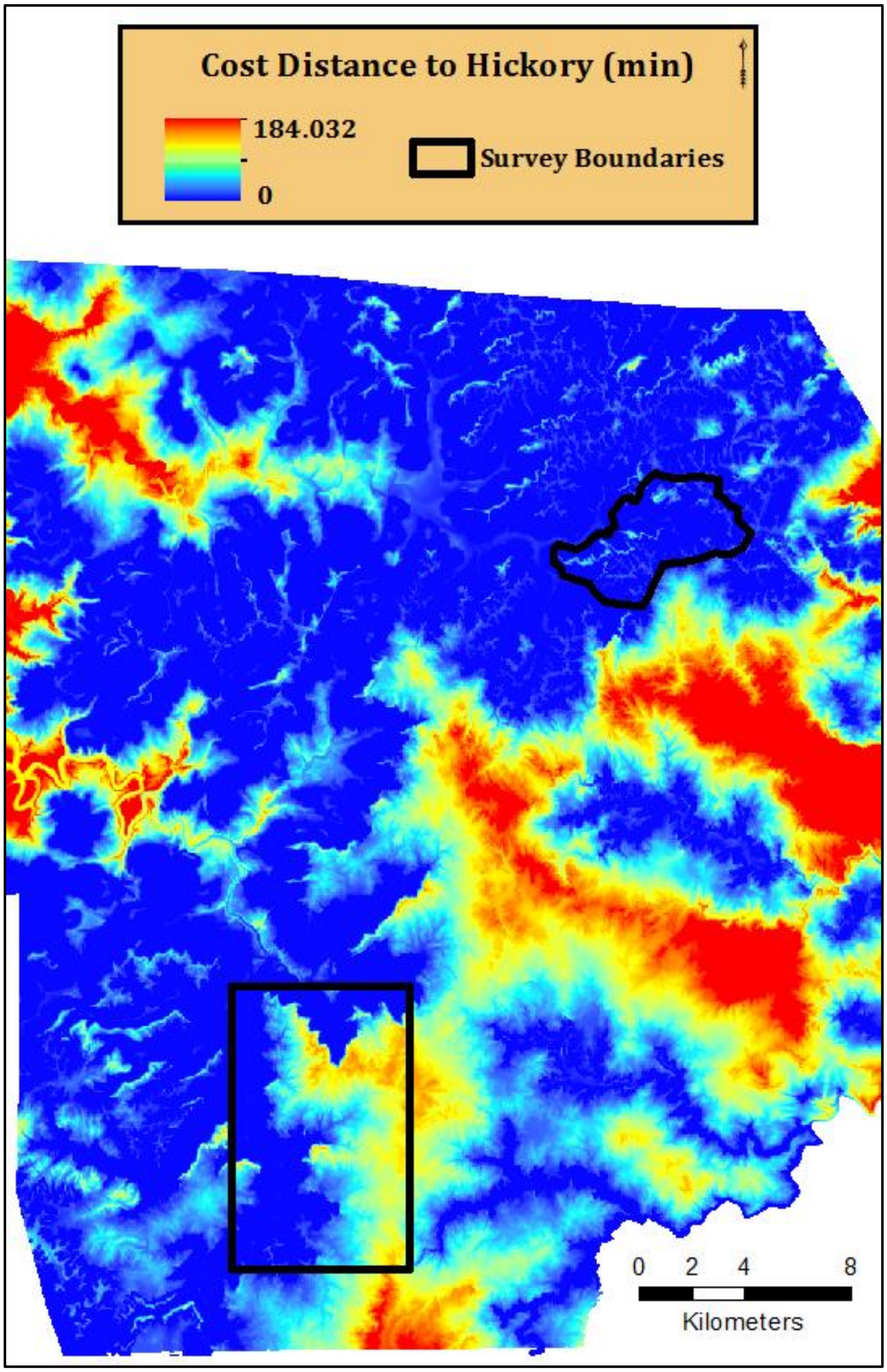


Figure 24: Raster Surface of the Cost Distance to Hickory P2 Model Variable. This raster surface represents the amount of time it takes to reach a "supporting zone" of Hickory. The dark blue areas are where the original "supporting zones" are located; as the color ramp progresses from dark blue to yellow to red, the time (and thus cost distance) to reach the boundaries of the "supporting zones" increases. Hickory is common in the Escarpment portion of the UCP and it does not appear to occur in some of the eastern portion of the uplands.

Proximity to Water Sources. The availability of water was and continues to be an important resource for humans. Using the slope cost raster and hydrography data, a cost distance raster was created to indicate the amount of time in minutes it would take to reach a viable (in this case, perennial stream) water source (Figure 26). However, the resulting calculations are not completely accurate due to the many unmapped seeps, springs, waterfalls, and intermittent (seasonal) streams in the region.

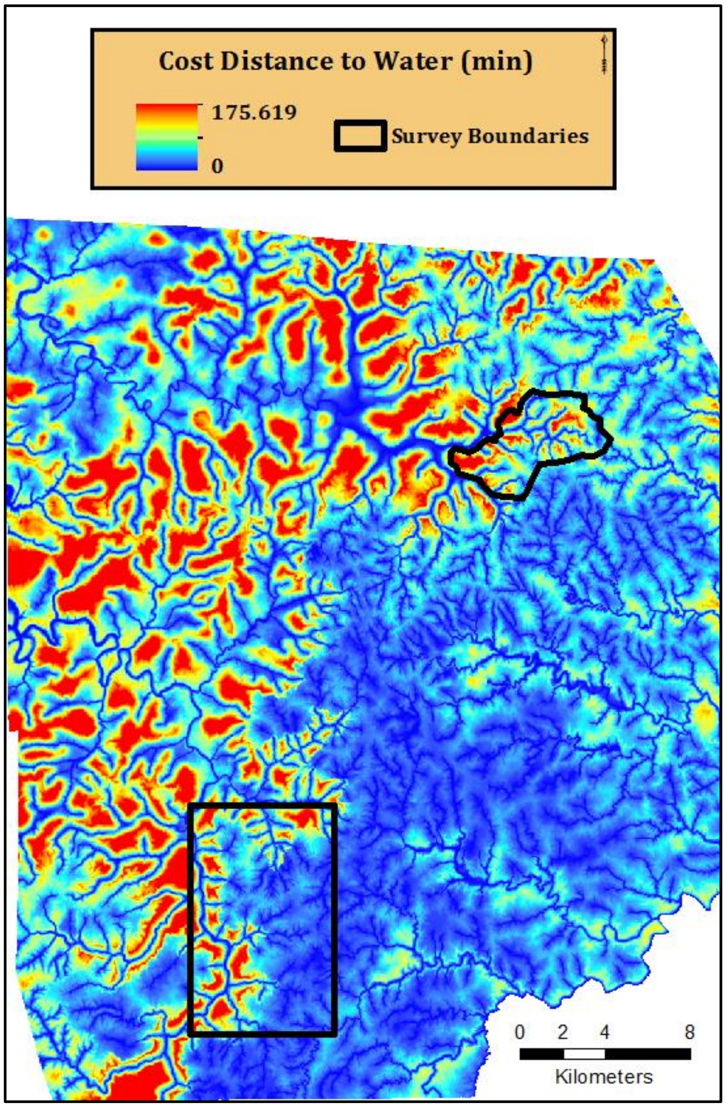


Figure 25: Raster Surface of the Cost Distance to Water P2 Model Variable. The dark blue lines represent the actual blue-line streams. As the color ramp progresses from dark blue to yellow to red, the time (and thus cost distance) to reach a water resource increases.

Raster Extraction

Each of the 27 raster surfaces generated as explanatory variables covered all 9 topographic quadrangles of the UCP. Because the model will be developed using the Pogue Creek data and tested using the East Obey data, rasters for each variable were extracted for each survey area. Two vector data layers were created to represent the survey boundaries of Pogue Creek and the East Obey (see Figure 4, page 74). The Pogue Creek survey boundary layer was created by digitizing the general outline of the proposed Pogue Creek State Natural Area (Langston and Franklin 2010). For the East Obey, there was no pre-defined survey area so an arbitrary survey boundary was assigned for the Wilder and Grimsley quadrangles (Franklin 2002). The 2 boundaries were used as masks to extract only the raster values for the corresponding survey area. Altogether there are 3 datasets representing the 28 preliminary model variables for a total of 84 raster surfaces: the UCP, Pogue Creek, and East Obey.

Data Standardization

All rasters for each dataset (the UCP, Pogue Creek, and the East Obey) were individually standardized on a scale of 0 to 1 using the Raster Calculator. With 27 possible model variables, there are many different measurement units and all on different numerical scales (see Table 2, page 78); standardizing the rasters made them unitless and all on the same scale. More importantly, standardizing the rasters allowed a direct comparison of regression coefficients for an individual study area. This was important for discussing the possible significance of variables in relationship to the site selection by prehistoric peoples. However, the standardized rasters and regression coefficients cannot be directly compared for the 2 separate study areas.

Preliminary Statistical Analysis

After the raster surfaces were created for the explanatory variables, some preliminary statistical analyses were performed to determine the final candidate variables for running the spatial logistic regression. The Pogue Creek data (rasters and point data) were used to conduct the preliminary statistical analysis and to develop the UCP model, while the East Obey data were used to evaluate model performance.

Goodness-of-fit

A goodness-of-fit test establishes whether or not an observed distribution differs from a theoretical distribution. The Kolmogorov-Smirnov test is one example of a goodness-of-fit test with a null hypothesis that samples are drawn from the same distributions. Tests such as this are appropriate for determining whether a variable should be a candidate for a model because an explanatory variable with similar distributions for sites and non-sites would not be a good predictor of potential site locations. Similar to traditional statistical programs, the Kolmogorov-Smirnov test is available in R (R Core Team 2012), though as a pixel-based function (Berman 1986; Baddeley et al. 2005). The `kstest.ppm` function is executed using 4 (internal) steps: (1) the original data points (e.g., sites) are extracted from the model and the observed distribution is determined by collecting the values of the covariate at those points; (2) the predicted distribution is computed by evaluating the values of the covariate at all locations and putting them together in a cumulative distribution function; (3) the observed distribution is transformed on a scale of 0 to 1 using the cumulative distribution function; (4) the null hypothesis is rejected if the transformed numbers are not independent and identically distributed (i.i.d.) uniform random numbers (Baddeley and Turner 2005; Baddeley and Turner 2012: 416). The code used to execute the `kstest.ppm` function is shown below

```
>kstest(X, covariate)
>plot(kstest(X, covariate)
```

where “X” is a point pattern file (i.e., site presence data) and the “covariate” is a spatially-referenced pixel image (i.e., rasters representing the explanatory variables). The first command returns the basic results of the test such as the p-value while the second command plots the observed and predicted distributions (Baddeley and Turner 2005; Baddeley and Turner 2012). The `kstest.ppm` (and many other functions in R) requires that the explanatory variables (or covariates) be converted into an image file. The raster surfaces for the candidate explanatory variables were converted to TIFFs in ArcMap 10.0 (ESRI 2011); those were subsequently added to the workspace in R (R Core Team 2012) and converted to image files. Following conversion, all 27 preliminary explanatory variables were tested using the `kstest.ppm` function; p-values and plots were generated for each.

Multicollinearity

Following the goodness-of-fit tests, the remaining explanatory variables were tested for multicollinearity. When two or more variables are exact or near exact linear functions of each other, multicollinearity is present in the dataset. Multicollinearity in a regression equation can produce inaccurate regression coefficients because highly correlated variables cause redundancy in the model. Explanatory variables were checked for correlation within each model group (P1 and P2) using the Band Collection Statistics tool in ArcMap 10.0 (ESRI 2011).

Candidate Variables

After performing goodness-of-fit tests and checking for multicollinearity, the remaining variables are considered candidate variables for the spatial logistic regression model. To assess model stability and consistency, traditional logistic regression and spatial dependence models were developed and compared to the spatial logistic regression model.

Site Absence Data

Both site (presence) and non-site (absence) data are needed to conduct the final steps of preliminary statistical analysis. The site presence data, the 125 Pogue Creek rock shelters, were discussed at the beginning of this chapter. Ideally, site absence data would include recorded rock shelter locations where no cultural material was identified. However, sterile shelters were not recorded on a routine or systematic basis during the Pogue Creek survey; shelters recorded as non-sites were not always shovel-tested to see if cultural materials lay beneath the surface. Because these data could not be verified with any certainty, site absence data (n=125 points) were generated using a random point generator. The site presence and absence layers were merged together to make a single shapefile. The values of the 27 standardized raster surfaces (the preliminary explanatory variables) for Pogue Creek were extracted to the site presence and absence point locations. The attribute tables for the site presence/absence data were exported from ArcMap 10.0 (ESRI 2011) and used to test for spatial autocorrelation. They were later used to run a logistic regression in SPSS (IBM Corp 2011) and a spatial error model in GeoDa (Anselin et al. 2006) as a means of comparison with spatial logistic regression.

Spatial Autocorrelation

In the previous chapter, some common types of regression-based approaches used in site location modeling were discussed with emphasis on determining an appropriate model for the UCP dataset. It was determined that a spatial logistic regression model would be the best approach because of the categorical response variable (site presence vs. site absence) and because it would capture the underlying spatial dependence present in most archaeological datasets. The presence of spatial dependence was determined by testing the Pogue Creek data for spatial autocorrelation; this is usually the first step in choosing whether or not a spatial model is needed in place of an aspatial model such as traditional logistic regression. Spatial autocorrelation, in this case, would mean that the location of a known rock shelter site is dependent on the location of other nearby sites—the observations (sites) are not spatially independent of each other. If a dataset is spatially autocorrelated, the regression assumption of independence of observations is violated; an aspatial regression approach could then lead to inaccurate coefficients and unreliable results. A common way to test for spatial autocorrelation is to examine the residuals of a linear regression such as OLS (Ward and Gleditsch 2008). The Pogue Creek site presence and absence data were tested for spatial autocorrelation using both ArcMap 10.0 (ESRI 2011) and the open source program GeoDa (Anselin et al. 2006). The local Moran's I value of 0.0349 was significant at p-value = 0.003. Even though this indicates a low degree of spatial autocorrelation, it is still significant. After determining that the Pogue Creek data were spatially autocorrelated, a spatial approach was adopted and the model development process was modified accordingly.

Spatial Logistic Regression Model

Because the Pogue Creek data were spatially autocorrelated and the response variable is categorical, neither traditional logistic regression nor spatial dependence models were appropriate for generating the UCP model. Therefore, spatial logistic regression was used to develop and test the UCP site location model using the `slrm.ppm` function (Baddeley et al. 2010) in the statistical program R (R Core Team 2012). The `slrm.ppm` function requires 2 types of inputs: the geographic locations of the site presence data and image files for each explanatory variable. Three functions were used to run the spatial logistic regression model (SLRM) and generate the regression coefficients and significance values:

```
1. >slrm(PresData ~1 + Variable1 + Variable2 + …)
2. >print(P1Model)
3. >anova(P1Model, test=”Chi”)
```

The first function uses the site presence data and the image files for each explanatory variable to run a binary logistic regression. The second function prints the regression coefficients and the third function generates the significance values for each explanatory value. The P1 and P2 models were run separately; regression coefficients and significance values were generated for each model. The explanatory variables and corresponding SLRM coefficients were entered into the Raster Calculator using the spatial logistic regression equation (see Equation 5, page 68). Three potential surfaces were generated for the UCP: the P1 static model, the P2 dynamic model, and finally, the P3 cumulative model. The geometric interval classification method (ESRI 2011) was then used to classify the raster values into 5 categories of archaeological potential: very low, low, moderate, high, and very high.

Comparing Model Approaches

Though spatial logistic regression was used to generate the UCP model, it is important to empirically demonstrate the advantages of using spatial logistic regression over more traditional approaches. The candidate variables were used to run a logistic regression in SPSS (IBM Corp 2011) and a spatial error model in GeoDa (Anselin et al. 2006) so that regression coefficients and significance values could be compared for all 3 approaches.

CHAPTER 6

RESULTS

This chapter provides the results of both the preliminary statistical analysis and the models produced in R (R Core Team 2012), SPSS (IBM Corp 2011), and GeoDa (Anselin et al. 2006). The graphic representation (map) of the UCP site location model is also provided. Model results will be discussed in the following chapter; only basic results are presented here.

Preliminary Statistical Analysis

Goodness-of-fit

The Kolmogorov-Smirnov goodness-of-fit tests were run in R (R Core Team 2012) using the kstest.ppm function (Berman 1986; Baddeley et al. 2005). Graphs comparing the observed and predicted distributions were generated for all 27 variables (Appendix B). Five explanatory variables were removed from the model because the observed and predicted distributions were not significantly different: *Percent of Monteagle Limestone (Mm)*, *Soil Thickness*, *Cost Distance to Chestnut Oak*, *Cost Distance to Scarlett Oak*, and *Shelter Index at 300m*.

Multicollinearity

The variables were tested for correlation within each model group. If 2 or more variables were positively or negatively correlated above 0.6, at least 1 variable was removed. The p-values of the Kolmogorov-Smirnov tests were used to help decide which variables would be eliminated in the event of high correlation. Correlation matrices were generated using the Band Collection Statistics tool (Appendix C). Table 4 shows correlations above a 0.6 for both P1 and P2 model groups.

Table 4: Correlation in the P1 and P2 Models. The Band Collection Statistics tool in ArcMap 10 (ESRI 2011) was used to check the raster surfaces of the explanatory variables for correlation. Correlations above a 0.6 that indicate cases of high correlation are listed.

Correlation of Model Variables	
P1 Variables	
Elevation & Percent of Rockcastle Conglomerate (Pr)	0.82
P2 Variables	
Solar Radiation & Direct Duration	0.71
(Cost Distance) Walnut & Southern Red Oak	0.89
(Cost Distance) Hickory & Walnut	0.87

For the P1 Model, the variables *Elevation* and *Percent of Rockcastle Formation (Pr)* were correlated at a 0.82. Because the variable *Elevation* had a lower Kolmogorov-Smirnov p-value AND because other geologic formations were retained as candidate variables, the *Percent of Rockcastle Formation (Pr)* variable was removed from the P1 model. For the P2 model group, there were several cases of high correlation between variables. The *Direct Duration of Solar Radiation* variable had a more significant Kolmogorov-Smirnov p-value than *Annual Solar Radiation*, so the latter was removed from the model. Also, because *Cost Distance to Walnut* is correlated with 2 other variables, it was removed from the model.

Candidate Variables

After removing variables based on preliminary statistical tests, 19 variables were considered candidate variables for inclusion in the UCP model (Table 5).

Table 5: Candidate Variables for the UCP Model. After preliminary statistical testing, 7 P1 variables and 12 P2 variables remain as candidate variables for the UCP model.

P1 Static Variables	P2 Dynamic Variables
Elevation	Direct Duration
Earth Curvature	Eastness
Slope	Northness
Soil Erosion	Cost Distance to Northern Red Oak
Percent of Bangor Limestone & Hartselle Formation (Mbh)	Cost Distance to Southern Red Oak
Percent of Pennington Formation (Mp)	Cost Distance to White Oak
Percent of the Fentress Formation (Pf)	Cost Distance to Hickory
	Cost Distance to Water
	Potential Volume of Wood Fiber
	100m Shelter Index
	1000m Shelter Index
	Terrain Texture

Spatial Logistic Regression

The `slrm.ppm` function (Baddeley et al. 2010) was used to run a spatial logistic regression in R (R Core Team 2012). The codes used to run the P1 and P2 models are provided in Appendix D.

SLRM Results

The results of the spatial logistic regression are divided into sections showing the significance values and regression coefficients for the final variables. Seven explanatory variables were used to generate the P1, or static, model. By itself (without the dynamic model) this model represents the best attempt to identify where any rock shelter (not necessarily a prehistoric site) could be located. The P2, or dynamic, model represents factors that might have influenced the site selection process by prehistoric peoples. For the P1 model, the variable

Percent of Fentress Formation was not captured as significant in predicting site presence/absence. Also, for the P2 model, the *Eastness* and *Northness* variables were not significant and were therefore removed from the dynamic model. Table 6 lists the significance values for the final P1 and P2 variables.

Table 6: SLRM Significance Values for P1 and P2 Variables. The significance values of the final explanatory variables are listed by model group.

P1 Variables	Significance (p-value)	P2 Variables	Significance (p-value)
Elevation	0.004093	Direct Duration	<2.2e-16
Curvature	1.494e-09	CD Northern Red Oak	4.525e-14
Slope	<2.2e-16	CD Southern Red Oak	0.0028734
Soil Erosion	0.010503	CD White Oak	9.621e-06
PerMbh	0.007311	CD Hickory	0.0397648
PerMp	7.051e-06	CD Water	3.696e-05
		Potential Vol. Wood	3.703e-06
		100 m Shelter Index	0.0398552
		1000m Shelter Index	0.0002019
		Terrain Texture	0.0042453

In archaeological site location modeling (and many other applications of predictive modeling), the regression coefficients for each explanatory variable are used to generate the graphic, or visual model. Also, because the explanatory variables were standardized on a scale of 0 to 1, their regression coefficients can be compared to discuss possible links to differential site selection of rock shelters on the UCP. Positive regression coefficients mean that both the explanatory and response variable change in value in the same direction, whereas negative coefficients represent a change in opposite directions. Similarly, the absolute values of the regression coefficients (for the standardized variables only) can be used to directly compare the contribution of each variable to the prediction of site presence; high absolute values indicate a stronger relationship and vice versa. Table 7 shows the SLRM coefficients for the P1 and P2 variables.

Table 7: SLRM Coefficients for P1 and P2 Variables. The regression coefficients of the final explanatory variables are listed by model group. The P1 and P2 Equation columns indicate how each variable is included in the (multiple) regression equation used to generate the UCP model.

P1 Variables	Regression Coefficient	P1 Equation	P2 Variables	Regression Coefficient	P2 Equation
Elevation	0.4430579	β_1	Direct Duration	-3.5815320	β_1
Curvature	-4.4996610	β_2	CD Northern Red Oak	0.7000523	β_2
Slope	5.0355586	β_3	CD Southern Red Oak	2.0731965	β_3
Soil Erosion	-2.5550436	β_4	CD White Oak	9.1120472	β_4
PerMbh	-97.4354671	β_5	CD Hickory	-7.0056506	β_5
PerMp	-3.2240448	β_6	CD Water	0.5960222	β_6
			Potential Vol. Wood	-3.9423587	β_7
			100m Shelter Index	-4.9597100	β_8
			1000m Shelter Index	4.1100522	β_9
			Terrain Texture	3.3754449	β_{10}

Archaeological Potential Surfaces

The SLRM regression coefficients and the raster surfaces for each of the P1 and P2 explanatory variables were entered in the Raster Calculator to produce 3 probability surfaces (Table 8). After generating the initial model surfaces, the geometric interval classification method was used to re-classify the probability surfaces into archaeological "potential" surfaces. Altogether, 3 archaeological potential surfaces were generated for the UCP model: the P1 static model (Figure 26), the P2 dynamic model (Figure 27), and finally, the P3 cumulative model (Figure 28).

Table 8: Equations for Generating Archaeological Potential Surfaces. This table shows how the explanatory variables and their spatial logistic regression coefficients were used to generate the raster surfaces for the UCP model. The equations were executed using the Raster Calculator in ArcMap 10 (ESRI 2011).

P1 (Static) Model Equation $$\frac{1}{1 + \text{Exp} -(\log(100) + (\text{Elevation} * \beta_1) + (\text{Curvature} * \beta_2) + (\text{Slope} * \beta_3) + (\text{Potential Soil Erosion} * \beta_4) + (\text{Percent of Mbh} * \beta_5) + (\text{Percent of Mp} * \beta_6))}$$
P2 (Dynamic) Model Equation $$\frac{1}{1 + \text{Exp} -(\log(100) + (\text{Direct Duration} * \beta_1) + (\text{CD Northern Red Oak} * \beta_2) + (\text{CD Southern Red Oak} * \beta_3) + (\text{CD White Oak} * \beta_4) + (\text{CD Hickory} * \beta_5) + (\text{CD Water} * \beta_6) + (\text{Potential Volume Wood} * \beta_7) + (\text{100m Shelter Index} * \beta_8) + (\text{1000m Shelter Index} * \beta_9) + (\text{Terrain Texture} * \beta_{10}))}$$
P3 (Cumulative) Model Equation P1 * P2

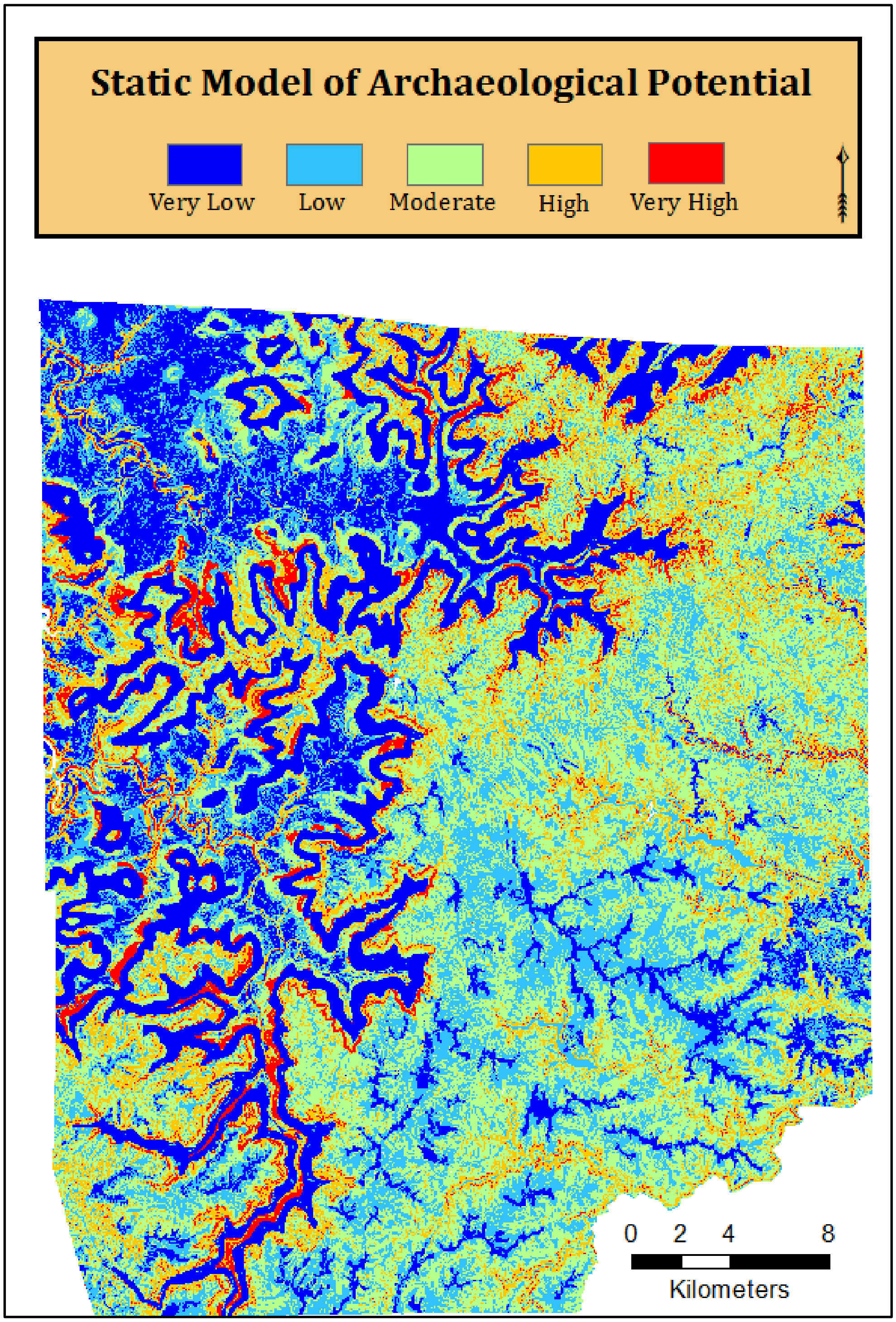


Figure 26: P1 (Static) Model of Archaeological Potential for the UCP. This raster surface was generated using the final P1 static variables and represents the potential for locating any rock shelter, site or non-site. It is important to point out here that only gorge rock shelter locations were modeled, and this surface does not indicate where upland shelters (e.g. on top of the plateau) would be located.

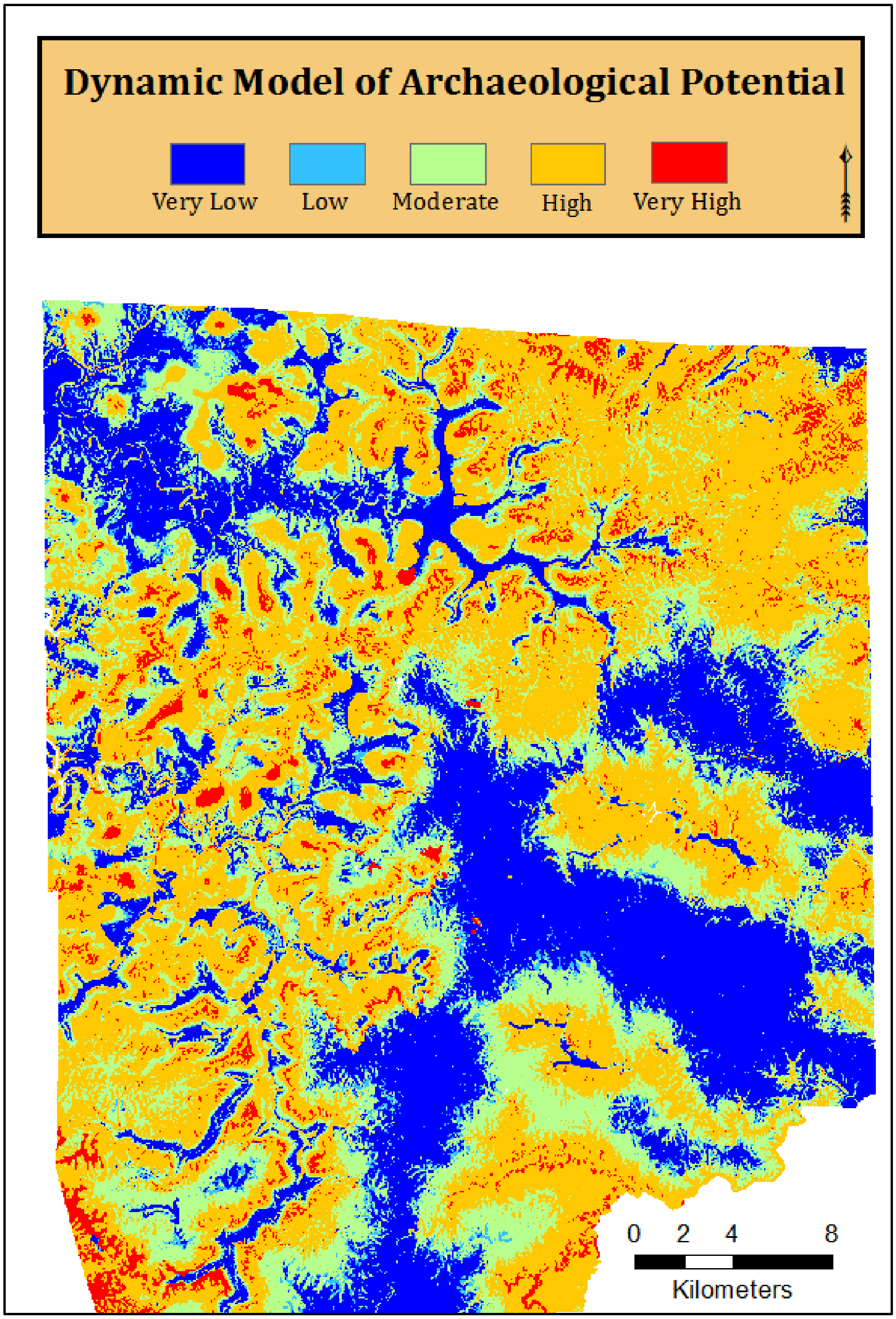


Figure 27: P2 (Dynamic) Model of Archaeological Potential of the UCP. This raster surface was generated using the final P2 dynamic variables and represents areas with the potential of finding archaeological sites based on factors that may have been important to prehistoric peoples for locating residential sites.

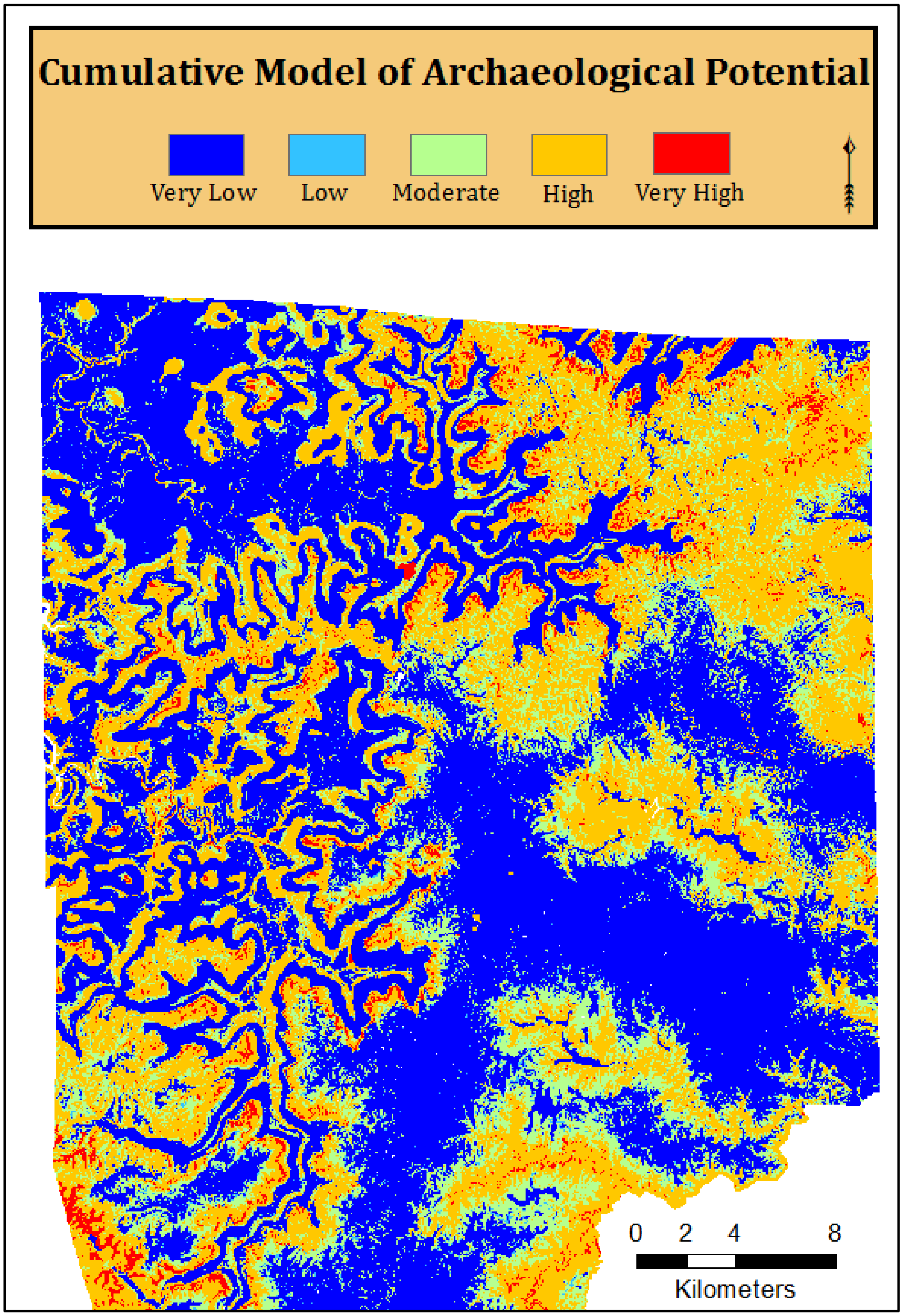


Figure 28: P3 (Cumulative) Model of Archaeological Potential on the UCP. This raster surface was generated by multiplying the raster surfaces of the P1 and P2 models and represents the potential for locating prehistoric (gorge) rock shelter sites. This model only applies to rock shelters that are located in gorges and along bluff lines but not on the upland portion of the UCP.

Model Performance

After generating the archaeological potential maps for the UCP, the locations of the Pogue Creek and East Obey rock shelters were used evaluate the models performance based on the “potential” categories. Also, the percent of total land area within each “potential” category was calculated. Ideally, the high or extremely high categories should cover a relatively small portion of the study area. Table 9 is a summary of the UCP model performance for Pogue Creek and the East Obey. The Pogue Creek data were used to construct the model and the East Obey data were used to test model performance. Eighty-three percent of the East Obey sites were correctly classified as falling in the high and very high potential areas which cover 35% of the total land area of the UCP. This indicates a model with high performance. Figure 29 shows the locations of the Pogue Creek and East Obey rock shelter sites in the potential categories.

Table 9: UCP Model Performance. This table shows the number of known prehistoric rock shelter sites from 2 archaeological surveys that fell within each of the archaeological potential categories of the UCP site location model

Archaeological Potential	# of Pogue Creek Rock Shelters (n=125)	# of East Obey Rock Shelters (n=77)	Percentage of total known sites (n=202)	Percentage of total area (UCP)
Very Low	1	3	2%	48%
Low	0	0	0%	1%
Moderate	0	10	5%	16%
High	49	30	39%	32%
Very High	75	34	54%	3%

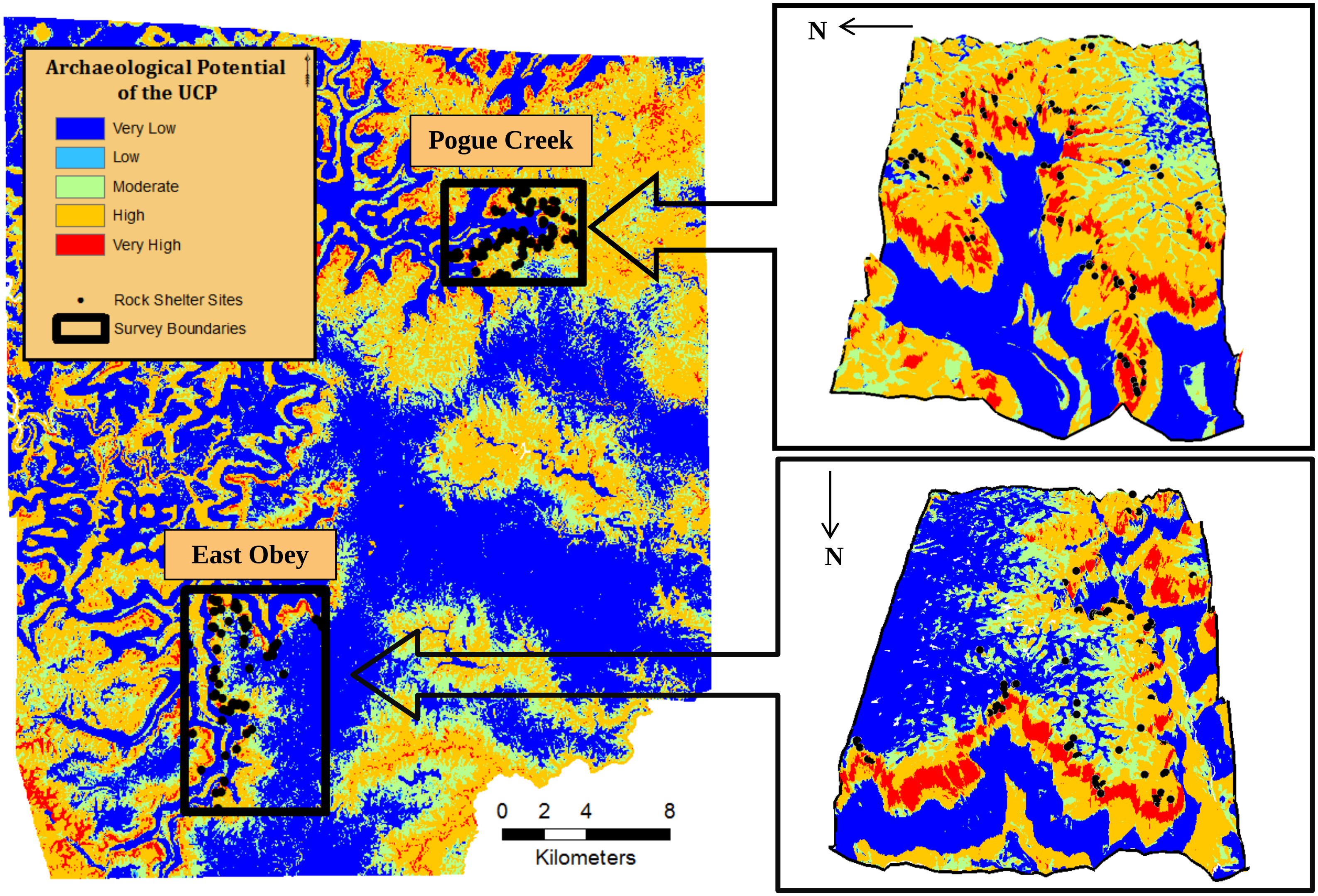


Figure 29: Terrain Surfaces of the Pogue Creek and East Obey Survey Areas in the P3 Model. The known prehistoric rock shelter sites in the Pogue Creek and East Obey survey areas are shown based on the archaeological potential categories of the final P3 site location model.

Other Model Approaches

In addition to the spatial logistic regression model, a binary logistic regression was run in SPSS (IBM Corp 2011) and a spatial error model was used to run an OLS regression in GeoDa (Anselin, Syabri, and Kho 2006). The significance values and regression coefficients for each explanatory variable were compared for the spatial logistic regression (SLRM), traditional logistic regression (TLR), and spatial error model (SEM). Full reports of the logistic regression and spatial error models are provided in Appendix E and F, respectively.

Significance Values

The significance levels of the final 17 explanatory values used to generate the UCP model are provided in Table 10 for comparison purposes; significance values for each model approach are in Appendix G. Eight of the 16 explanatory variables were not significant at p = 0.05 when the traditional logistic regression (TLR) approach was used: *Elevation*, *Percent of Bangor Limestone and Hartselle Formation*, *Direct Duration of Incoming Solar Radiation*, *Cost Distance to Northern Red Oak*, *Cost Distance to White Oak*, *Cost Distance to Hickory*, *Cost Distance to Water,* and *Terrain Texture*. Similarly, the spatial error model (SEM) did not find 8 variables as significant when compared to the spatial logistic regression model (SLRM): *Elevation*, *Curvature*, *Cost Distance to Northern Red Oak*, *Cost Distance to White Oak*, *Cost Distance to Hickory*, *Cost Distance to Water*, *Average Potential Volume of Wood Fiber*, and *Terrain Texture.* Except with 4 variables—*Direct Duration of Incoming Solar Radiation*, *Curvature*, *Percent of Bangor Limestone and Hartselle Formation*, *Potential Volume of Wood Fiber*--the TLR and SEM approaches agreed on the significance (or insignificance in this case) of the model variables. These results will be discussed further in the next chapter.

Table 10: Comparison of Significance Levels by Model Approach. The significance level of each explanatory variable within the traditional logistic regression model (TLR), spatial error model (SEM), and the spatial logistic regression model (SLRM) is denoted by the number of asterisks: 1 asterisk indicates that the variable was significant at p=0.05; 2 asterisks indicates significance at p=0.01; and 3 asterisks indicates significance at p=0.001.

P1 Variables	TLR	SEM	SLRM
Elevation	-	-	**
Curvature	*	-	***
Slope	***	***	***
Soil Erosion	***	***	**
PerMbh	-	*	**
PerMp	**	**	***
P2 Variables	**TLR**	**SEM**	**SLRM**
Direct Duration	-	***	***
CD Northern Red Oak	-	-	***
CD Southern Red Oak	**	*	**
CD White Oak	-	-	***
CD Hickory	-	-	*
CD Water	-	-	***
Potential Vol. Wood	**	-	***
100 m Shelter Index	*	**	*
1000m Shelter Index	*	**	***
Terrain Texture	-	-	**

"-" p value > 0.05   *p = 0.05   **p = 0.01   ***p=0.001

Regression Coefficients

When comparing regression coefficients, 2 things should be considered: the sign (positive or negative) and the absolute value. The sign of a regression coefficient corresponds to the relationship between the explanatory and response variable and whether or not their values increase or decrease together. Comparisons can also be made based on the absolute value of regression coefficients—as the absolute value of the coefficient increases, so does the strength of the relationship between the explanatory and response variable (and vice versa). Though the absolute value of regression coefficients can change with model approach (and are better for comparisons within a model instead of between approaches), the signs should be consistent. The regression coefficients for the 3 model approaches are compared in Table 11.

Table 11: Comparison of Regression Coefficients by Model Approach. This table lists the regression coefficients from the traditional logistic regression model (TLR), spatial error model (SEM), and the spatial logistic regression model (SLRM). The reason for this comparison is to look for differences in the coefficient sign (+ or -) between the 3 approaches. Only 1 difference was noted and this was for the *Cost Distance to Water* variable. Parentheses indicate that a variable was not significant (at $p < 0.05$).

P1 Variables	TLR	SEM	SLRM
Elevation	(2.508)	(0.2262229)	0.4430579
Curvature	-8.473	(-0.7701032)	-4.4996610
Slope	9.212	1.391825	5.0355586
Soil Erosion	-6.513	-0.7874126	-2.5550436
PerMbh	(-49.645)	-0.5639221	-97.4354671
PerMp	-3.286	-0.3743277	-3.2240448
P2 Variables	**TLR**	**SEM**	**SLRM**
Direct Duration	(-4.915)	-0.9321794	-3.5815320
CD N. Red Oak	(1.802)	(0.3230226)	0.7000523
CD S. Red Oak	4.348	0.374139	2.0731965
CD White Oak	(16.203)	(1.783052)	9.1120472
CD Hickory	(-11.512)	(-1.102964)	-7.0056506
CD Water ±	(-3.65)	(-0.04108933)	0.5960222
Pot. Vol. Wood	-4.450	(-0.655538)	-3.9423587
100 m SI	-6.882	-0.6138773	-4.9597100
1000m SI	5.423	0.7538912	4.1100522
Terrain Texture	(41.356)	(0.6904599)	3.3754449

± difference in sign between model approaches

CHAPTER 7

DISCUSSION AND CONCLUSIONS

In the introductory chapter, 3 research objectives specific to this project were identified:

1. To determine if the Pogue Creek and East Obey survey data could be used to develop and test a predictive model for unsurveyed areas of the UCP;
2. To determine the possible factors contributing to prehistoric rock shelter selection on the UCP; and
3. To determine whether spatial logistic regression can be proposed as a better alternative than traditional statistical models for developing archaeological predictive models

This chapter re-visits each of the 3 research objectives by reviewing the results presented in the previous chapter. The discussion of model results focuses on the practical, theoretical, and methodological facets of the Upper Cumberland Plateau site location model. First, the different "potential" categories (very high, high, moderate, low, and very low) of the UCP model will be described using the model variables. In the second section, a few model variables are used to discuss site selection factors of the Pogue Creek and East Obey rock shelters. Finally, a comparison of the different model approaches is offered along with a discussion on the advantages of using spatial logistic regression.

Practical

The graphic representation of the UCP model was presented in the previous chapter (see Figure 28, page 124). Now, the different categories of archaeological potential will be discussed in terms of the explanatory variables within each model group (P1 and P2); a brief summary of each category is also provided. This discussion focuses on the range and/or average values of the explanatory variables in each category (Appendix G). This model "narrative" will offer further insight into differential site selection which will be covered in the following section.

Model Description

Very High Potential. The "very high potential (VHP)" area of the UCP comprises approximately 2.2% of the total overall area. Out of the 202 known rock shelters sites in the Pogue Creek and East Obey survey areas, 109 (54%) fall in the VHP areas. This category is characterized by an average elevation of 455 meters, though it can range from 198-550 meters. The curvature of the landforms are both negative (concave) and positive (convex), though concave areas are more common; this is most likely due to topographic depressions associated with rock shelter formation. This category has the highest average slope (29°), though there are known prehistoric rock shelters in areas with 74° slopes in this category. Overall, the potential for soil erosion is lower here than in any of the other categories—this is probably because there is very little soil in these areas to begin with. The Bangor Limestone and Hartselle Formation (MBH) is not present in this category, and the Pennington Formation (Mp) appears in less than 1% of the total area inside the VHP category. The main geologic formations present in the VHP areas are the Fentress Formation (Pf) and the Rockcastle Conglomerate (Pr), accounting for 50% and 49% of the total area, respectively.

The VHP areas receive fewer hours of sunlight per year on average (3,225 hours) than any of the other categories. In relation to travel time to supporting zones of different oak species, the VHP potential areas are further, on average, from Northern Red Oak, Southern Red Oak, and White Oak than in the other "potential" areas. In contrast, supporting zones of Hickory are closer and take less time to access. Travel time to water sources from VHP areas is greater than in the other categories with an average of 46 minutes. The average volume of wood fiber in the VHP areas is 57 ft3/acre—the lowest of the 5 “potential” categories. Lastly, the VHP areas are extremely sheltered locally, though when compared to areas within a kilometer, they are regionally exposed surfaces.

In summary, the VHP area coincide with the bluff lines-- the upper slopes of the gorge. These areas occur in the highest slopes (up to 74°) at varying elevations between approximately 198-550 meters above sea level (masl). These areas are extremely rugged, with minimal soil erosion, and lightly forested. Also, these areas are very sheltered within 100 meters—this coupled with the high slopes means the least amount of average direct incoming solar radiation per year (3,225 hours). As far as geology, the VHP areas mainly occur in the Fentress Formation and Rockcastle Conglomerate. Lastly, it takes more time to reach sources of water and zones of oak species from the VHP areas than the other 4 categories.

High Potential. Seventy-nine, or 39%, of the known prehistoric rock shelter sites from Pogue Creek and East Obey fall within the “high potential (HP)” area. This category covers 31.8% of the total study area and has a lower average elevation (442 masl) than the VHP category (457 masl). Concave (negative curvature) landforms are still more common than convex (positive curvature) areas. There is little difference in measures of soil erosion between the VHP and HP categories. However, the average slope decreases from 28° to 16° in the HP category. Geologically, the high and very high potential areas are similar except that the presence of the Fentress Formation decreases significantly as the Rockcastle Conglomerate becomes more prevalent.

On average, the HP areas receive more direct insolation (3,808 hours) than the VHP potential areas. Also, from the HP areas, less travel time is required to access the supporting zones of each of the 5 tree species. However, there is little difference between the VHP and HP areas for access to zones of Southern Red Oak and Hickory. Water sources are closer as well, with an average access time of 31.7 minutes. For potential volume of wood fiber, the HP areas average about 73.5 cubic feet per acre compared to the 57 cubic feet per acre in the VHP areas. Most definitely, the increasing potential volume of food fiber is related to the closer proximity of vegetation zones of oak species. In addition, the HP areas are much less exposed on a regional scale than the VHP areas, though similar to the VHP areas, they are still rather sheltered locally; the HP areas have the greatest range in both local and regional shelter compared to the rest of the potential categories. Finally, there was a significant decrease in Terrain Texture from the VHP category indicating that the terrain of the HP areas is less rugged.

To summarize, the HP category is characterized by slopes up to 70° within the mid-elevation ranges below the VHP areas—so the mid-to upper slopes of the gorge. The geology is still predominantly sandstone or sandstone conglomerates though there is less of the Fentress Formation and more Rockcastle Conglomerate. The travel time to water sources and zones of oak species is less than in VHP areas. These areas are not as sheltered as the previous category and they are significantly smoother and more level.

Moderate Potential. The “moderate potential (MP)” area comprises approximately 16.8% of the total project area and has a slightly higher average elevation than the HP category--though lower than in VHP areas. Out of the 202 known archaeological sites in Pogue Creek and East Obey, 10 sites (5%) are in the MP areas. It is in this category that convex landforms become more widespread than convex surfaces representing a shift towards flatter surfaces such as the top of the plateau instead of the concave slopes. Measures of soil erosion are similar to VHP and HP areas. Similarly, the Bangor Limestone and Hartselle Formation (Mbh) is absent. However, percentages of the Pennington Formation (Mp) have increased as the presence of the Fentress Formation (Pf) and Rockcastle Conglomerate (Pr) continue to decrease; this category captures the transition from Pennsylvanian-aged sandstone, shale, and siltstone to outcrops of Mississippian-aged carbonates.

As the potential for archaeological sites decreases, increases are noted in the total number of hours of direct incoming solar insolation (4041 hours in MP areas vs. 3808 hours in the HP areas). Also while access times for zones of Northern Red Oak, Southern Red Oak, and White Oak are less, the time required to access zones of Hickory is greater than for the VHP and HP areas. This inverse relationship demonstrates the importance of hickory zones in predicting site presence. MP areas are closer to water sources (average access time of 20.8mins) and have higher potential wood fiber volumes (average of 79 cubic feet per meter) than the previous categories. Additionally, the MP areas are more exposed than any other areas on a local scale. The terrain texture continues to decrease with archaeological potential.

The MP areas can be summarized as having considerably lower slope angles but higher elevations than the VHP and HP areas. Potential for soil erosion is consistent with previous categories though MP areas are more forested. These areas generally overlap supporting zones of Northern Red Oak and White Oak, though they are further away from zones of Hickory. For geology, the Pennington Formation (Mp) is slightly more common than in the HP areas, and the Rockcastle Conglomerate (Pr) remains dominant with some areas in the Fentress Formation (Pf). Lastly, these areas are the most locally exposed.

Low Potential. The “low potential (LP)” category has the highest average elevation of 458 meters and slightly higher potential for soil erosion than the previous categories. This category accounts for the smallest portion of the study area at only 1.5% and none of the Pogue

Creek of East Obey sites fall within this category. Though the average curvature is slightly lower (so slightly more concave) than the MP category, this is the lowest range of curvature values. This possibly indicates a trend towards flatter surfaces and fewer extremes (either extremely convex or concave). The average slope is about 8° with a maximum of 58°. The same geologic trends are visible in the LP category as with the MP areas: average of 3.8% of the Pennington Formation (Mp) and 9.8% for the Fentress Formation (Pf). A variety of other geologic formations, ranging from sandstone conglomerates to limestone, account for the remaining percentages.

The LP areas receive the most hours of direct insolation and are closer (in time and distance) to all 3 zones of oak species and to water sources than any other category. The potential volume of wood fiber is relatively high compared to the VHP, HP, and MP areas. However, the LP areas are more sheltered than MP and HP areas, though less so than the VHP areas. In this area, there is little degree of terrain roughness and the LP areas are the smoothest.

In summary, the LP areas have the highest average elevation and the flattest surfaces with an average slope of about 8°—these areas occur on the top of the plateau with some areas at the bottom of the gorges. The Rockcastle Conglomerate is the dominant geologic rock unit with some occurrences of the Fentress Formation and the Pennington Formation. These areas receive the highest average solar insolation per year—they are very exposed areas surrounded by White and Northern Red Oak. Also, these areas are the closest to streams that appear on USGS topographic maps.

<u>Very Low Potential</u>. Areas classified as "very low potential (VLP)" cover 48% of the total survey area; 4 rock shelter sites (2% of total sites) from the Pogue Creek and East Obey survey areas fall within this category. These areas have the overall lowest average elevation at

414 meters but with the overall largest range (because they cover almost half of the study area). Most of the areas in this category are flat or convex surfaces with the highest potential for soil erosion. The slope is the lowest in this category with an average of 7.94°. Only in this category does the Bangor Limestone and Hartselle Formation appear and with an average of 11.9%. The Pennington Formation is present at similar percentages with an average of 11.7%. The Fentress Formation is much less prevalent (average of 3.2%) as with the Rockcastle Conglomerate.

The VLP areas, second to the LP category, receive an average of 4041 hours of direct solar insolation yearly. Zones of Northern Red Oak and White Oak co-occur in both the LP and VLP areas, so access times are minimal. However, zones of Hickory are found closer to VHP areas, so access time to supporting zones of Hickory species from VLP areas average 27 minutes—this is still minimal compared to the average time it takes to access Southern Red Oak (98 minutes). Little difference is noted in proximity to water, potential volume of wood fiber, and local measures of shelter between the VLP and LP areas. One notable exception is a significant increase in shelter on a regional scale; the VLP areas are the most sheltered regionally.

To summarize, the VLP areas are the only areas where the Bangor Limestone and Hartselle Formation is present. Other limestone and sandstone formations are also present, though less so in any other category. In addition, these areas are very flat and have the lowest degree slopes with the highest potential for soil erosion. White Oak and Northern Red Oak occur in these areas though the presence of Hickory is rare. The defining characteristics of the VLP category significantly vary across the UCP because this category covers the highest percentage of land in the study area.

Conclusions

The UCP site location model was developed using the Pogue Creek survey data (n=125 sites) and tested using the East Obey survey data (n=77), all of which are rock shelter sites with some prehistoric component. The model performed extremely well on the Pogue Creek data—not surprising because the model was developed using this dataset. Only 1 Pogue Creek site fell in the VLP category. This shelter is located at a lower elevation than all of the other shelters and occurs in the Pennington Formation (Mississippian-aged) instead of a Pennsylvanian formation. For the East Obey data, 3 of the East Obey sites were classified as VLP sites. For these shelters, their low potential is the result of differences in the P1 and P2 models from the Pogue Creek shelters. Even for study areas within the same county, there can be significant differences in geologic units, soil conditions, vegetation, etc. This means that models have to be developed for individual study areas based on the environmental conditions and archaeological resources unique to that area. In this case, a model developed using the Pogue Creek data would need to be adjusted and refined to fit other survey areas. The concept of a single Upper Cumberland Plateau model is not necessarily realistic if the ultimate goal is to have a model that most accurately reflects the relationship between the archaeological record and the environmental setting. However, if based solely on the model's performance on the East Obey dataset where 83% of the known prehistoric rock shelter sites fell within the high and very high potential areas, the UCP model developed herein can be described as highly successful. This is one of the first archaeological site location modes that focuses on modeling rock shelter locations and sites. For this reason, this model is extremely unique and has great implications in both upland archaeology and geospatial analysis. This model demonstrates the usefulness and application of GIS studies in archaeology, especially in a CRM context.

Theoretical

Now that a model has been generated for the UCP of Tennessee and the "potential" categories described in terms of model variables, the significant variables can be used to discuss possible factors contributing to differential site selection. Using the 3 primary goals proposed by Jochim (1976: 50) as guiding hunter-gatherer settlement practices, variables relating to the proximity of resources, shelter, and view will be discussed

Proximity to Resources

Five variables used to generate the UCP model are related to resources that may have been important to prehistoric hunter-gatherers: *Cost Distance to Northern Red Oak*, *Cost Distance to Southern Red Oak*, *Cost Distance to Hickory*, and *Cost Distance to Water*. Oak and Hickory species are both important sources of food for humans and wildlife so it is foreseeable that prehistoric hunter-gatherers would have situated themselves close to areas where food sources (both for gathering nuts and hunting wildlife) were plentiful. Of the oak species, White Oak is the most widespread on the UCP occurring on upper and lower slopes and at almost every elevation. The Pogue Creek and East Obey shelters are, on average, farther from supporting zones of White Oak than other areas but only by about 3 minutes—this is not a big enough difference to consider access to supporting zones of White Oak as a site selection factor. However, the *Cost Distance to Northern Red Oak* variable was the second most significant variable in the P2 model. On average, the Pogue Creek and East Obey rock shelters are about 10 to 20 minutes away from areas likely to support Northern Red Oak. This is similar to the UCP as a whole. However, rock shelter sites are closer to areas of Southern Red Oak than the rest of the UCP. This indicates a trend towards locating sites closer to areas that support Southern Red Oak. Interestingly, sites are also situated closer to supporting zones of Hickory than any other vegetation type used in this study. It can be inferred then that Hickory was a more important resource than the oak species for prehistoric hunter-gatherers. Out of the 4 vegetation types, it is possible that the proximity of a rock shelter to Southern Red Oak and Hickory influenced site selection choices.

According to the UCP model, prehistoric rock shelter sites are, on average, farther (in both time and distance) away from water sources than non-sites. However, this variable does not take into account unmapped seeps and springs that are myriad on the UCP. Water was and continues to be a very important resource, and access to such a resource is critical to the maintenance and development of human populations. Most likely, access to water resources was an important factor influencing choices made by prehistoric hunter-gatherers in locating residential sites. However, this variable does not accurately reflect the availability of water sources on the UCP. Instead of using blue-line streams or even flow accumulation rasters, the locations of intermittent streams, seeps, springs, and waterfalls need to be better documented in archaeological surveys. This is not to say that variables using cost distance or even straight-line (Euclidean) distance to water sources cannot be used to develop reliable predictive models; variables such as this can be very useful (especially in arid landscapes) only if the discrepancies between the (real) environment and the mappable data that represent the environment are understood. It is therefore possible that a majority of rock shelters on the UCP are much closer to water sources than this variable is able to reflect.

Shelter and View

Though the rock shelters provide shelter in the sense that they are ready-made structures there are varying degrees of exposure related to the surrounding landscape. The Pogue Creek and East Obey sites are located in extremely sheltered local areas (within a 100 meter radius) relative to the UCP as a whole. These areas would have provided protection from winter winds, abundant solar insolation, and rain/snow. However, at a 1000 meter radius (regional scale), the rock shelter locations are very exposed. This is not a contradiction and instead indicates that though the sites are sheltered locally, when compared to the rest of the plateau, they are situated higher and offer better views of the overall landscape. So whereas the steep gorges offer some protection from the natural elements (at least more so than being on top of the plateau), those locations also offer prime viewing locations—similar to vantage points. However, it is important to point out that some of the most sheltered sites are in horseshoe-shaped gorges and do not provide wide views of the landscape. Because view and shelter are not always related, variables should be incorporated to address both characteristics individually. A variable relating to view was not used in the UCP model and therefore view cannot be addressed independently. It appears, however, that the amount of shelter a location provides (beyond the rock shelter itself) was a contributing factor in rock shelter selection on the UCP.

Though it does not specifically relate to the 3 goals proposed by Jochim (1976: 50), the *Direct Duration of Solar Radiation* variable was the most significant variable in the dynamic model, meaning that it was the best overall predictor of site presence. However, the regression coefficient was negative, indicating that sites receive less solar insolation than non-site areas. The amount of solar insolation a location receives is generally related to the location’s aspect, and for the Southeastern United States, southerly and easterly facing landforms potentially receive more direct sunlight than northerly and westerly facing landforms. The variables of *Eastness* and *Northness* were removed from the final model because neither were statistically significant in predicting site presence. So in contrast to the findings of Hall and Klippel (1988) and Mickelson (2002) who propose aspect as a “trend” for prehistoric site selection, in this study, insolation was negatively correlated to site presence and aspect was not a significant factor. These results agree more closely with the conclusions of Langston and Franklin (2010). The significance of the *Direct Duration of Solar Radiation* variable in predicting site presence does not necessarily indicate that prehistoric peoples *chose* locations that received less sunlight. It is more likely that this variable helped to narrow down locations where rock shelters naturally form and perhaps belonged in the P1 static model. Rock shelters crop-out in eroded bluff lines around steep slopes; many occur in horseshoe-shaped canyons that do not receive a lot of sunlight. Another issue to consider is that this variable measured the amount of yearly sunlight locations receive. It is possible that many of these shelters were occupied on a seasonal basis or seasonal rounds were made between shelters, and the amount of solar insolation a location receives varies with sun angle throughout the year—especially in a landscape characterized by steep gorges and high plateaus. Perhaps a better way to investigate the relationship between solar insolation and site selection is to examine seasonal variability in solar insolation at locations where sites have been documented. By doing this, it might be possible to determine which shelters may have been used in warmer months and which ones might have been used for winter occupation.

Conclusions

As has been pointed out in this thesis, archaeological site location modeling can be extremely useful in the context of CRM. However, this thesis also demonstrates the application of archaeological site location modeling in exploring patterns of human behavior as related to differential site selection. It is impossible to definitely know why prehistoric hunter-gatherers choose to live in certain rock shelters and why others remained unoccupied for over 10,000 years. However, the variables that were significant in predicting known prehistoric rock shelter locations can be used to develop hypotheses about factors relevant to differential site selection on the UCP. From the model developed herein, the close proximity to Southern Red Oak and Hickory appear to have been important to prehistoric peoples. Also, sheltered areas were chosen for protection from the natural elements. Most likely, the availability of water was not a factor--only because there is no shortage of intermittent streams, springs, and seeps (sometimes coming out of the back of the rock shelter itself) on the UCP. Finally, the amount of sunlight cannot be identified as a contributing factor to site selection based on the model results; however, as pointed out previously, this may be due to a disconnect between the temporal scale of study and variable importance.

Methodological

The explanatory variables used to generate the UCP model were also used to run a traditional logistic regression (TLR) and a spatial error model (SEM); the 2 approaches are compared to the spatial logistic regression model (SLRM) using the significance values and regression coefficients of the final model variables.

Significance Values

For the UCP model, the 3 most significant variables in predicting site presence are *Slope*, *Direct Duration of Incoming Solar Radiation*, and *Cost Distance to Northern Red Oak*. Though both the TLR and SEM approaches captured *Slope* as significant (and the most significant variable), *Cost Distance to Northern Red Oak* was not significant (p-value > 0.1) in the TLR model. However, both *Direct Duration of Incoming Solar Radiation* and *Cost Distance to Northern Red Oak* were significant (p-value < 0.1) in the SEM approach. This means that for the *Cost Distance to Northern Red Oak* variable, there may be an important underlying spatial process that the TLR model was unable to capture because it assumes independence of the explanatory variables. Similarly, the SEM and SLRM capture the *Percent of Bangor Limestone* *and Hartselle Formation* variable significant in predicting site presence while the (non-spatial) TLR model did not. Except for *Cost Distance to Southern Red Oak*, neither the SEM nor TLR approaches found the vegetation variables as significant. This is interesting because these are some of the most significant variables in the UCP model when the SRLM approach was used. It is not surprising that the *Percent of Fentress Formation* variable was not significant in the SEM or TLR approaches because it was the least significant variable in the SLRM. Similarly, *Cost Distance to Hickory* was also one of the lesser significant variables in the SLRM and was not significant in both the TLR and SEM approaches. However, the second least significant variable in the SLRM, the *Shelter Index at 100m*, was more significant in the other 2 approaches; different from the *Cost Distance to Northern Red Oak* variable, the SLRM and SEM approaches revealed that the *Shelter Index at 100m* variable is not as significant because of its spatial relationship with the site presence data.

There are 2 possible explanations for why some variables were not captured as significant in the TLR and SEM approaches. The first is an issue of spatial dependence or spatial autocorrelation in the dataset. Most likely, a majority of these variables are the result of underlying spatial processes or relationships that have to be accounted for, or handled properly, in order to produce an accurate and reliable model. Traditional logistic regression does not have a spatial component and is not equipped to appropriately handle spatially autocorrelated data. However, the spatial error model does account for spatial dependence. The assumption then is that the SEM approach would have captured the spatial dependence and the same variables would have been significant as in the SLRM approach. However, a spatial error model is a linear regression and thus, its assumptions are violated because of the categorical response variable and because it assumes the relationship between the response and explanatory variables is linear. So just because several variables did not emerge as significant in the SEM approach, this does not mean that there are not important underlying spatial processes that need to be captured in order to produce an accurate model. Though it has the “spatial” component, the spatial error model is still not appropriate for modeling site presence/absence data. Thus, significance values and regression coefficients of the SEM approach can only be loosely interpreted.

The second possible explanation for discrepancies between the different approaches relates to the way in which the models are run. For the TLR and SEM approaches, a random sample of points representing site absence data was generated so that there would be equal numbers of site presence and absence data (n=125 points each). The SLRM model, which is a pixel-based approach, requires site presence data only--the locations of which are compared to every other location in the study area and not just 125 other absence locations as with the TLR and SEM approaches. Even though a random sample of points was used for site absence data in the TLR and SEM approaches, it is likely that not all of the variation in the explanatory variables was captured by the randomly sampled absence data. The pixel-based SLRM approach includes values of the explanatory variables throughout the study area and compares those to the values at each known site. This is one major advantage of using pixel-based approaches for modeling site locations.

Regression Coefficients

Only one major difference was noted when the regression coefficients for the 3 different model approaches were compared. One variable, *Cost Distance to Water*, had a different sign for the TLR and SEM approaches than for the SLRM. It is interesting that the TLR and SEM approaches had a negative coefficient for this variable, because this indicates that as the cost distance (time) to water sources decreased, the likelihood of site presence increased—so sites should be closer to water. However, as noted by Langston and Franklin (2010) in a previous GIS analysis of the Pogue Creek rock shelters, sites were seemingly farther away from water sources. As pointed out in the previous comparisons of significance values, the TLR and SEM approaches used randomly sampled data; this means that the randomly generated site absence data points had lower values (meaning less access time to water sources) for the *Cost Distance to Water* variable than the site presence data. This was not the case in the SLRM approach where the site presence locations (pixels) had higher values of *Cost Distance to Water* compared to all the other locations or pixels in the Pogue Creek study area. The SLRM approach more accurately reflects the relationship between the site presence data and the *Cost Distance to Water* variable because it includes absence data from throughout the entire study area.

Conclusions

Traditional logistic regression is the most common modeling approach used in archaeological predictive modeling today. More than likely, spatial autocorrelation is present in most datasets used to generate site location models. Though it may be common practice to use an aspatial approach to analyze what are essentially spatial patterns of behavior, it is possible that important information regarding prehistoric settlement patterns is being overlooked or masked. For the UCP site location model, the spatial logistic regression model was able to capture important spatial relationships between the response and explanatory variables that would have been missed if a traditional logistic regression was used. The methodological considerations that go into developing archaeological models are just as important as the theoretical basis for developing them in the first place. If the goal is to analyze and interpret patterns of prehistoric hunter-gatherer behavior, then the methodology should accurately reflect the spatial relationships inherent to human behavior.

Final Thoughts

The UCP site location model can be considered successful based on its validation using the East Obey survey data. Though it performed well with the test set, its efficiency and performance as a "working" model will be evaluated during ongoing and future archaeological surveys on the Upper Cumberland Plateau. In a Cultural Resources Management (CRM) context, archaeological predictive models can help facilitate decisions about identifying, evaluating, and monitoring archaeological resources. For models to be sufficient and successful in this endeavor there has to be a better understanding of how prehistoric peoples were using and occupying the landscape and how that can be conceptualized in a GIS environment. Also, models have to be refined and updated as new information is gathered and/or as better methods are developed. The model developed herein is no exception as there are things that can already be improved upon. However, this does not refute the validity of the UCP site location model in having the potential to predict archaeological (rock shelter) sites. The real test of model performance can only take place in the field and the opportunity to do so is rare, to say the least. Prehistoric occupation and use of the Upper Cumberland Plateau of Tennessee is only beginning to be understood. This is a landscape that is both culturally and naturally unique and archaeological investigations of the region have only scratched the surface. Ongoing and future archaeological surveys will go a long way in not only protecting archaeological resources but in better understanding prehistoric lifeways in the region.

REFFERENCES

Agterberg FP. 1974. Automatic Contouring of Geological Maps to Detect Target Areas for Mineral Exploration. Journal of the International Association for Mathematical Geology 6:373–395.

Altschul JH, Sebastian L, Heidelberg K. 2004. Predictive Modeling in the Military. Similar Goals, Divergent Paths. Rio Rancho: SRI Foundation.

Anderson DG. 2001. Climate and Culture Change in Prehistoric and Early Historic Eastern North America. Archaeology of Eastern North America 29:143–186.

Anderson DG, Sassaman KE. 2012. Recent Developments in Southeastern Archaeology: From Colonization to Complexity. Washington, D.C.: Society for American Archaeology Press.

Andresen J, Madsen T, Scollar I eds. 1993. Computing the Past: Computer Applications and Quantitative Methods in Archaeology. Aarhus: Aarhus University Press.

Anselin L, Syabri I, Kho Y. 2006. GeoDa: An Introduction to Spatial Data Analysis. Geographical Analysis 38:5–22.

Baddeley A, Berman M, Fisher NI, Hardegen A, Milne RK, Schuhmacher D, Shah R, Turner R. 2010. Spatial Logistic Regression and Change-of-Support in Poisson Point Processes. Electronic Journal of Statistics 4:1151–1201.

Baddeley A, Turner R, Moller J, Hazelton M. 2005. Residual Analysis for Spatial Point Processes. Journal of the Royal Statistical Society 67:617–666.

Baddeley A, Turner R. 2005. Spatstat: An R Package for Analyzing Spatial Point Patterns. Journal of Statistical Software 12:1–42.

Baddeley A, Turner R. 2012. Package “spatstat”: Spatial Point Pattern Analysis, Model-fitting, Simulation, Tests. Available from: http://www.spatstat.org/

Berman M. 1986. Testing for Spatial Association Between a Point Process and Another Stochastic Process. Journal of the Royal Statistical Society 35:54–62.

Bettinger RL. 1980. Obsidian Hydration Dates for Owens Valley Settlement Categories. Journal of California and Great Basin Anthropology 2:286–292.

Bow SM, Franklin JD. 2009. Luminescence Dating and the Pogue Creek Archaeological Survey. Paper Presented at the 21st Annual Current Research in Tennessee Archaeology Meeting, Nashville, Tennessee.

Braun EL. 1950. Eastern Deciduous Forests of North America. Philadelphia, PA: Blakiston.

Breitburg E, Broster JB, Reesman AL, Strearns RG. 1996. The Coats-Hines site: Tennessee’s first Paleoindian-Mastodon Association. Current Research in the Pleistocene 13:6–8.

Broster JB, Norton MR, Stanford DJ, Haynes, Jr. CV, Jodry MA. 1996. Stratified Fluted Point Deposits in the Western Valley of Tennessee. In: Walling R, Wharey C, Stanley C, editors. Proceedings of the 14th Annual Mid-South Archaeological Conference. Special Publication No. 1. Memphis: Panamerican Consultants, Inc. p. 1–11.

Brown JA. 1985. Long-term Trends to Sedentism and the Emergence of Complexity in the American Midwest. In: Price TD, Brown JA, editors. Prehistoric Hunter-Gatherers: The Emergence of Cultural Complexity. Orlando, Florida: Academic Press. p. 201–231.

Brown JA, Vierra RK. 1983. What Happened in the Middle Archaic? Introduction to An Ecological Approach to Koster Site Archaeology. In: Phillips JA, Brown JA, editors. Archaic Hunters and Gatherers in the American Midwest. New York: Academic Press. p. 165–196.

Campbell JS. 2006. Archaeological Predictive Model of Southwestern Kansas. Master's Thesis, University of Kansas.

Chapman J. 1985. Tellico Archaeology: 12,000 years of Native American History. University of Tennessee Department of Anthropology Report of Investigations No. 43 and the Tennessee Valley Authority Publications in Anthropology No. 41. Knoxville, Tennessee: The University of Tennessee.

Church T, Brandon RJ, Burgett GR. 2000. GIS Applications in Archaeology: Method in Search of Theory. In: Wescott KL, Brandon RJ, editors. Practical Applications of GIS for Archaeologists: A Predictive Modeling Toolkit. London: Taylor and Francis. p. 144–166.

Conolly J, Lake M. 2006. Geographical Information Systems in Archaeology. New York: Cambridge University Press.

Coppock JT, Rhind DW. 1991. The History of GIS. In: Maguire DJ, Goodchild MF, Rhind DW, editors. Geographical Information Systems: Principles and Applications. Longman Scientific and Technical. p. 21–43.

Delcourt PA, Delcourt HR. 1987. Long-term Forest Dynamics of the Temperate Zone. A Case Study of Late-Quaternary Forests in Eastern North America. Springer-Verlag.

Des Jean T, Benthall JL. 1994. A Lithic-Based Prehistoric Cultural Chronology of the Upper Cumberland Plateau. Tennessee Anthropologist 19:115–147.

Dowd JT. 1989. The Anderson Site: Middle Archaic Adaptation in Tennessee's Central Basin. Miscellaneous Paper No. 13. Knoxville, Tennessee: Tennessee Anthropological Association.

Duncan RB, Beckman KA. 2000. The Application of GIS Predictive Site Location Models within Pennsylvania and West Virginia. In: Wescott KL, Brandon RJ, editors. Practical

Applications of GIS for Archaeologists: A Predictive Modeling Toolkit. London: Taylor and Francis. p. 37-61.

Dye AD, Franklin JD, Hays MA. 2011. Lithic Technology and Site Function at Early Times Rock Shelter, Upper Cumberland Plateau, Tennessee. Paper Presented at the 23rd Annual Current Research in Tennessee Archaeology Meeting, Nashville, Tennessee.

Dye AD, Navel J, Dennison ME, Franklin JD. 2010. Archaeological Testing of Hemlock Falls Rock House (40Fn239), Pogue Creek State Natural Area. Paper Presented at the 22nd Annual Current Research in Tennessee Archaeology Meeting, Nashville, Tennessee.

Dye DH. 1996. Initial Riverine Adaptation in the Midsouth: An Examination of Three Middle Holocene Shell Middens. In: Carstens KC, Watson PJ, editors. Of Caves and Shell Mounds. Tuscaloosa, Alabama: University of Alabama Press. p. 140–158.

Ebert JI, Kohler TA. 1988. The Theoretical Basis of Archaeological Predictive Modeling and a Consideration of Appropriate Data-Collection Methods. In: Judge WJ, Sebastian L, editors. Quantifying the Present and Predicting the Past: Theory, Method, and Application of Archaeological Predictive Modeling. Denver, Colorado: US Department of the Interior, Bureau of Land Management. p. 97–172.

Espa G, Benedetti R, De Meo A, Ricci U, Espa S. 2006. GIS Based Models and Estimation Methods for the Probability of Archaeological Site Location. Journal of Cultural Heritage 7:147–155.

ESRI. 2011. ArcGIS Desktop. Redlands, CA: Environmental Systems Research Institute.

Faulkner CH. 1968. A Review of Pottery Types in the Eastern Tennessee Valley. Proceedings of the Southeastern Archaeological Conference Bulletin 8:23–35.

Faulkner CH. 1988. Middle Woodland Community and Settlement Patterns on the Eastern

Highland Rim, Tennessee. In: Mainfort, Jr. RC, editor. Middle Woodland Settlement and Ceremonialism in the Mid-South and Lower Mississippi Valley. Archaeological Report No. 22. Jackson, Mississippi: Mississippi Department of Archives and History. p. 75–98.

Fenneman NM. 1938. Physiography of Eastern United States. New York: McGraw-Hill.

Ferguson TA. 1988. Lithic Analysis and the Discovery of Prehistoric Man-Land Relationships in the Uplands of the Big South Fork of the Tennessee Cumberland Plateau. Doctoral Dissertation, Department of Anthropology, The University of Tennessee, Knoxville.

Ferguson TA, Pace RA, Gardner JW, Hoffman RW. 1986. Final Report of the Big South Fork Archaeological Project: Survey, Testing, and Recommendations. Knoxville, Tennessee: The University Tennessee, Department of Anthropology.

Finke PA, Meylemans E, Van de Wauw J. 2008. Mapping the Possible Occurrence of Archaeological Sites by Bayesian Inference. Journal of Archaeological Science 35:2786–2796.

Fisher PF, Farrelly C, Maddocks A, Ruggles CLN. 1997. Spatial Analysis of Visible Areas from the Bronze Age Cairns of Mull. Journal of Archaeological Science 24:331–347.

Franklin JD. 1999. The Rime of the Ancient Miners, Master's Thesis, Department of Anthropology, The University of Tennessee, Knoxville.

Franklin JD. 2001. Excavating and Analyzing Prehistoric Lithic Quarries: An Example from 3rd Unnamed Cave, Tennessee. Midcontinental Journal of Archaeology 26 (2):199-217.

Franklin JD. 2002. The Prehistory of Fentress County, Tennessee: An Archaeological Survey. Doctoral Dissertation, Department of Anthropology, The University of Tennessee, Knoxville. Ann Arbor, Michigan: University Microfilms, University of Michigan.

Franklin JD. 2006. Prehistoric Culture Chronology on the Upper Cumberland Plateau of Tennessee. Paper Presented at the 63rd Annual Southeastern Archaeological Conference, Little Rock, Arkansas.

Franklin JD. 2007. Cave and Rock Shelter Excavations on the Tennessee River, Roane County, Tennessee. Paper presented at the 64th Annual Southeastern Archaeological Conference, Knoxville, Tennessee.

Franklin JD. 2008a. Luminescence Dates and Woodland Ceramics from Rock Shelters on the Upper Cumberland Plateau of Tennessee. Tennessee Archaeology 3:87–100.

Franklin JD. 2008b. Big Cave Archaeology in the East Fork Obey River Gorge. In: Dye DH, editor. Cave Archaeology of the Eastern Woodlands: Essays in Honor of Patty Jo Watson. Knoxville, Tennessee: University of Tennessee Press. p. 141-155.

Franklin JD, Bow SM. 2008. The Upper Cumberland Plateau Archaeological Luminescence Dating Project. Paper Presented at the 65th Annual Southeastern Archaeological Conference, Charlotte, North Carolina.

Franklin JD, Bow SM. 2009. Archaeological Exploration of Workshop Rock Shelter, Upper Cumberland Plateau, Tennessee. In: Baumann T, Groover M, editors. Colleague, Mentor, and Friend: Papers in Honor of Charles H. Faulkner. Special Edition of Tennessee Archaeology 4(1-2):145–162.

Franklin JD, Dennison ME, Hays MA, Navel J, Dye AD. 2013. The Early and Middle Woodland of the Upper Cumberland Plateau, Tennessee. In: Wright A, Henry E, editors. Landscapes and People of the Early and Middle Woodland Southeast, in press.

Franklin JD, Hays MA, Sherwood SC, Langston LM. 2012. An Integrated Approach: Lithic Analyses and Site Function, Eagle Drink Bluff Shelter, Upper Cumberland Plateau, Tennessee. In: Carr PJ, Bradbury AP, Price SP, editors. Contemporary Lithic Analysis in the Southeast: Problems, Solutions, and Interpretations. Tuscaloosa, Alabama: University Alabama Press. p. 128–145.

Franklin JD, Simek JF. 2008. Core Refitting and the Accuracy of Techniques for Aggregate Lithic Analyses: The Case of 3rd Unnamed Cave, Tennessee. Southeastern Archaeology 27(1):108-121.

Franklin JD, Walker R, Hays MA, Beck CW. 2010. Late Archaic Site Use at Sachsen Cave Shelter, Upper Cumberland Plateau, Tennessee. North American Archaeologist 31:447–479.

Gaffney V, van Leusen M. 1995. Postscript—GIS, Environmental Determinism and Archaeology: A Parallel Text. In: Lock GR, Stančič Z, editors. Archaeology and Geographic Information Systems: A European Perspective. London: Taylor and Francis. p. 367-382.

Gaffney V, Stančič Z. 1991. GIS Approaches to Regional Analysis: A Case Study of the Island of Hvar. Ljubljana: Znanstveni institut Filozofske fakultete.

Gorenflo LJ, Gale N. 1990. Mapping Regional Settlement in Information Space. Journal of Anthropological Archaeology 9:240–274.

Graham RW, Haynes CV, Johnson DL, Kay M. 1981. Kimmswick: a Clovis-Mastodon Association in Eastern Missouri. Science 213:1115–1117.

Green E. 1973. Locational Analysis of Prehistoric Maya Sites in Northern British Honduras. American Antiquity 38:279–293.

Griffin JB. 1967. Eastern North American Archaeology: A Summary. Science 156:175–191.

Hall CL, Klippel WE. 1988. A Polythetic-Satisficer Approach to Prehistoric Natural Shelter Selection in Middle Tennessee. Midcontinental Journal of Archaeology 3(2):159–186.

Hardeman WD, Miller RA, Swingle GD. 1966. Geologic Map of Tennessee. Nashville, Tennessee: State of Tennessee, Department of Conservation, Division of Geology.

Hart JL. 2007. A Multi-Scale Analysis of Disturbance Dynamics in Hardwood Forest Communities on the Cumberland Plateau, USA. Doctoral Dissertation, University of Tennessee, Knoxville.

Hartung RE, Lloyd WJ. 1969. Influence of Aspects on Forests of the Clarksville Soils in Dent County, Missouri. Journal of Forestry 67:178–182.

Hasenstab RJ. 1983. A Preliminary Cultural Resource Sensitivity Analysis for the Proposed Flood Control Facilities Construction in the Passaic River Basin of New Jersey. Report submitted to the Passaic River Basin Special Studies Branch, Department of the Corps of Engineers, Soil Systems, Inc., Marietta, Georgia.

Hinkle CR, McComb WC, Safely Jr. JM, Schmalzer PA. 1993. Mixed Mesophytic Forests. In: Martin WH, Boyce SG, Echternacht AC, editors. Biodiversity of the Southeastern United States: Upland Terrestrial Communities. New York: John Wiley and Sons. p. 203–253.

Hinkle CR. 1989. Forest Communities of the Cumberland Plateau of Tennessee. Journal of the Tennessee Academy of Science 64:123–129.

Hodgen M. 1964. Early Anthropology in the Sixteenth and Seventeenth Centuries. University of Pennsylvania Press.

Hunt CB. 1967. Physiography of the United States. W. H. Freeman and Company.

IBM Corp. 2011. IBM SPSS Statistics for Windows. Armonk, NY: IBM Corp.

Jefferies RW. 2004. Regional Cultures, 700 B.C.-A.D. 1000. In: Fogelson RD, editor. Smithsonian Handbook of North American Indians, Volume 14, The Southeast. Washington, D.C.: Smithsonian Institution. p. 115–127.

Jochim MA. 1976. Hunter-Gatherer Subsistence and Settlement: A Predictive Model. Academic Press New York.

Jochim MA. 1981. Strategies for Survival: Cultural Behavior in an Ecological Context. Academic Press New York.

Judge WJ, Sebastian L eds. 1988. Quantifying the Present and Predicting the Past: Theory, Method, and Application of Archaeological Predictive Modeling. Denver, Colorado: US Department of the Interior, Bureau of Land Management.

Justice ND. 1987. Stone Age Spear and Arrow Points of the Midcontinental and Eastern United States. Bloomington, Indiana: Indiana University Press.

Kahane L. 2008. Regression Basics. Sage Publications, Inc.

Kamermans H, Wansleeben M. 1999. Predictive Modelling in Dutch Archaeology, Joining Forces. BAR International Series 757:225–229.

Kohler TA, Parker SC. 1986. Predictive Models for Archaeological Resource Location. Advances in Archaeological Method and Theory 9:397–452.

Krist Jr. FJ. 2006. Multicriteria/Multiobjective Predictive Modeling: A Tool for Simulating Hunter Gatherer Decision Making and Behavior. In: Mehrer MW, Wescott KL, editors. GIS and Archaeological Site Location Modeling. Taylor and Francis. p. 310–325.

Kroeber AL. 1939. Cultural and Natural Areas of Native North America. University of California Press.

Kvamme KL. 1984. Models of Prehistoric Site Location Near Pinon Canyon, Colorado. In: Condie, CJ, editor. Papers of the Philmont Conference on the Archaeology of Northeastern New Mexico, Proceedings 6. Albuquerque: New Mexico Archaeological Council. p. 347–370.

Kvamme KL. 1988. Development and Testing of Quantitative Models. In: Judge WJ, Sebastian L, editors. Quantifying the Present and Predicting the Past: Theory, Method, and Application of Archaeological Predictive Modeling. Washington, DC: US Bureau of Land Management, Department of Interior. p. 325–427.

Kvamme KL. 1989. Geographic Information Systems in Regional Archaeological Research and Data Management. Archaeological Method and Theory 1:139–203.

Kvamme KL. 1990. One-sample Tests in Regional Archaeological Analysis: New Possibilities Through Computer Technology. American Antiquity 55:367–381.

Kvamme KL. 1990. The Fundamental Principles and Practice of Predictive Archaeological Modeling. In: Voorrips A, editor. Mathematics and Information Science in Archaeology: A Flexible Framework. Studies in Modern Archaeology, Volume 3. Bonn: Holos-Verlag. p. 257–295.

Kvamme KL. 1995. A View from Across the Water: the North American Experience in Archaeological GIS. In: Lock GGR, Stančič Z, editors. Archaeology and Geographical Information Systems: A European Perspective. p. 1–14.

Kvamme KL. 1998. GIS in North American Archaeology: A Summary of Activity for the CAERE Project. Archeologia e calcolatori 9:127–146.

Langston LM, Dennison ME, Franklin JD. 2010. Archaeological Testing at York Palace (40Fn220), Pogue Creek State Natural Area. Paper Presented at the 22nd Annual Current Research in Tennessee Archaeology Meeting, Nashville, Tennessee.

Langston LM, Franklin JD. 2010. Archaeological Survey of Pogue Creek State Natural Area: A GIS Perspective. Paper Presented at the 67th Annual Meeting of the Southeastern Archaeological Conference, Lexington, Kentucky.

Langston LM, Navel J, Franklin JD. 2012. Archaeological Survey in Pickett State Forest, 2010-12. Paper Presented at the 24th Annual Current Research in Tennessee Archaeology Meeting, Nashville, Tennessee.

Lewis TMN, Kneberg MD. 1957. The Camp Creek Site. Tennessee Archaeology 13:1–48.

Lock G, Harris. 2006. Enhancing Predictive Archaeological Modeling: Integrating Location, Landscape, and Culture. In: Mehrer MW, Wescott KL, editors. GIS and Archaeological Site Location Modeling. Taylor and Francis. p. 115–134.

Lock G, Moffett J eds. 1992. Computer Applications and Quantitative Methods in Archaeology 1991. Oxford: Tempus Reparatum.

Lock G, Stančič Z eds. 1995. Archaeology and Geographical Information Systems. London: Taylor & Francis.

Madry S, Cole M, Gould S, Resnick B, Seibel S, Wilkerson M. 2006. A GIS-based Archaeological Predictive Model and Decision Support System for the North Carolina Department of Transportation. In: Mehrer MW, Wescott KL, editors. GIS and Archaeological Site Location Modeling. Boca Raton, Florida: Taylor & Francis. p. 292–309.

Magne MPR. 1989. Lithic Reduction Stages and Assemblage Formation Processes. In: Amick DS, Mauldin RP, editors. Experiments in Lithic Technology. Oxford: BAR International Series 528, British Archaeological Reports. p. 15-31.

Maschner HDG ed. 1996. New Methods, Old Problems: Geographic Information Systems in Modern Archaeological Research. Carbondale, Illinois: Southern Illinois University Center for Archaeological Investigations.

Mehrer MW, Wescott KL eds. 2006. GIS and Archaeological Site Location Modeling. Taylor and Francis.

Mickelson AM. 2002. Changes in Prehistoric Settlement Patterns as a Result of Shifts in Subsistence Practices in Eastern Kentucky. Doctoral Dissertation, Ohio State University.

Mink II PB, Stokes BJ, Pollack D. 2006. Points vs. Polygons: A Test Case Using a Statewide Geographic Information System. In: Mehrer MW, Wescott KL, editors. GIS and Archaeological Site Location Modeling. Taylor and Francis. p. 200–219.

Pace RA, Hays CT. 1991. Perspectives on Prehistoric Settlement in the Cumberland Plateau: The View from Station Camp. Tennessee Anthropologist 16:115–149.

Parker SC. 1985. Predictive Modelling of Site Settlement Systems Using Multivariate Logistics. In: Carr C, editor. For Concordance in Archaeological Analysis. Bridging Data Structure, Quantitative Technique, and Theory. Kansas City, Kansas: Westport Publishers. p. 173–207.

Plog F, Hill JN. 1971. Explaining Variability in the Distribution of Sites. In: Gumerman GJ, editor. The Distribution of Prehistoric Population Aggregates. Prescott: Prescott College Press. p. 7–36.

R Core Team. 2012. R: A language and environment for statistical computing. Vienna, Austria: R Foundation for Statistical Computing. Available from: http://www.R-project.org/

Ridges M. 2006. Regional Dynamics of Hunting and Gathering: An Australian Case Study Using Archaeological Predictive Modeling. In: Mehrer MW, Wescott KL, editors. GIS and Archaeological Site Location Modeling. Taylor and Francis. p. 115–134.

Roberts DW. 1986. Ordination on the Basis of Fuzzy Set Theory. Vegetatio 66:123–131.

Rogerson PA. 2010. Statistical Methods for Geography: A Student's Guide. Sage Publications Limited.

Roper DC. 1979. Archaeological Survey and Settlement Pattern Models in Central Illinois. Midcontinental Journal of Archaeology Special Paper 2.

Sasowsky ID. 1992. Evolution of the Appalachian Highlands: Geochemistry, Hydrogeology, Cave-sediment Magnetostratigraphy, and Historical Geomorphology of the East Fork Obey River, Fentress County, Tennessee. Doctoral Dissertation, Pennsylvania State University.

Sassaman KE. 2010. The Eastern Archaic, Historicized. Lanham, Maryland: AltaMira Press.

Scholtz SC. 1981. Location Choice Models in Sparta. In: Lafferty III R, Ottinger JL, Scholtz SC, Limp WF, Watkins B, Jones RD, editors. Settlement Predictions in Sparta: A Locational Analysis and Cultural Resource Assessment on the Uplands of Calhoun County, Arkansas. Arkansas Archaeological Survey Research Series. Fayetteville, Arkansas: Arkansas Archaeological Survey. p. 207–222.

Schroedl GF, Boyd, Jr CC. 1991. Late Woodland Period Culture in East Tennessee. In: Nassaney MS, Robb CR, editors. Stability, Transformation, and Variation: The Late Woodland Southeast. Springer. p. 69–90.

Schwarz KR, Mount J. 2006. There and Back Again: Revisiting Archaeological Locational Modeling. In: Mehrer MW, Wescott KL, editors. GIS and Archaeological Site Location Modeling. Boca Raton, Florida: Taylor and Francis. p. 154–175.

Simek JF, Franklin JD, Sherwood SC. 1998. The Context of Early Southeaster Prehistoric Cave Art: A Report on the Archaeology of 3rd Unnamed Cave. American Antiquity 63(4):663-677.

Smalley GW. 1986. Classification and Evaluation of Forest Sites on the Northern Cumberland Plateau. US Department of Agriculture, Forest Service, Southern Forest Experiment Station.

Smith BD. 1986. Archaeology of the Southeastern United States: From Dalton to De Soto 10,500-500 B.P. Advances in World Archaeology 5:1–92.

Soil Survey Staff. 2009. Soil Survey Geographic (SSURGO) Database for Big South Fork National River and Recreation Park, Kentucky and Tennessee. Natural Resources Conservation Science, United States Department of Agriculture. Available online at http://soildatamart.nrcs.usda.gov.

Soil Survey Staff. 2011. Soil Survey Geographic (SSURGO) Database for Fentress and Pickett Counties, Tennessee. Natural Resources Conservation Science, United States Department of Agriculture. Available online at http://soildatamart.nrcs.usda.gov.

Starr LE, Anderson KE. 1991. A USGS Perspective on GIS. In: Maguire DJ, Goodchild, MF, Rhind DW, editors. Geographical Information Systems: Principles and Applications. Longman Scientific & Technical. p. 11–22.

Steponaitis VP. 1986. Prehistoric Archaeology in the Southeastern United States, 1970-1985. Annual Review of Anthropology 15:363–404.

Steward J. 1955. The Concept and Method of Cultural Ecology. Bobbs-Merrill.

Straus LG. 1990. Underground Archaeology: Perspectives on Caves and Rock Shelters. Archaeological Method and Theory 2:255–304.

Sullivan AP, Schiffer MB. 1978. A Critical Examination of SARG. In: Euler RC, Gumerman GJ, editors. Investigations of the Southwestern Anthropological Research Group: The Proceedings of the 1976 Conference. Flagstaff, Arizona: Museum of Northern Arizona. p. 168–175.

Thornthwaite CW. 1948. An Approach Toward a Rational Classification of Climate. Geographical Review 38:55–94.

Tobler WR. 1970. A Computer Movie Simulating Urban Growth in the Detroit Region. Economic Geography 46:234–240.

Tomlinson RF. 1988. The Impact of the Transition from Analogue to Digital Cartographic Representation. The American Cartographer 15:249–262.

Trigger BG. 1968. The Determinants of Settlement Patterns. In: Chang KC, editor. Settlement Archaeology. Palo Alto. p. 53-78.

Veljanovski T ana, Stančič Z. 2006. Regional Dynamics of Hunting and Gathering: An Australian Case Study Using Archaeological Predictive Modeling. In: Mehrer MW, Wescott KL, editors. GIS and Archaeological Site Location Modeling. Taylor and Francis. p. 362–379.

Verhagen P, Whitley TG. 2012. Integrating Archaeological Theory and Predictive Modeling: a Live Report from the Scene. Journal of Archaeological Method and Theory 19:49–100.

Verhagen P. 2007. Case Studies in Archaeological Predictive Modeling. Leiden University Press.

Vita-Finzi C, Higgs ES. 1970. Prehistoric Economy in the Mount Carmel Area of Palestine: Site Catchment Analysis. Proceedings of the Prehistoric Society 36:1–37.

Wall J. 2013. Initial Investigation of Ceramic Assemblage from Gwinn Cove Rock Shelter, Upper Cumberland Plateau, Tennessee. Paper Presented at the 25th Annual Current Research in Tennessee Archaeology Meeting, Nashville, Tennessee.

Wansleeben M, Verhart L. 1997. Geographical Information Systems. Archaeological Dialogues 4:53–64.

Ward MD, Gleditsch KS. 2007. An Introduction to Spatial Regression Models in the Social Sciences. Available online at http://www.dces.wisc.edu/documents/articles/curtis/cesoc977-12/W4_W7_WardGleditsch.pdf.

Ward MD, Gleditsch KS. 2008. Spatial Regression Models. Sage Publications, Inc.

Warren RE, Asch DL. 2000. A Predictive Model of Archaeological Site Location in the Eastern Prairie Peninsula. In: Westcott KL, Brandon RJ, editors. Practical Applications of GIS for Archaeologists: A Predictive Modelling Toolkit. Philadelphia, Pennsylvania: Taylor and Francis p. 6–36.

Watson PJ. 2001. Ridges, Rises, and Rocks; Caves, Coves, Terraces, and Hollows: Appalachian Archaeology at the Millennium. In: Sullivan LP, Prezzano SC, editors. Archaeology of the Appalachian Highlands. Knoxville, Tennessee: The University of Tennessee Press. p. 319–322.

Wescott KL, Brandon RJ eds. 2000. Practical Applications of GIS for Archaeologists: A Predictive Modeling Toolkit. Philadelphia, Pennyslvania: Taylor and Francis.

Wheatley D, Gillings M. 2002. Spatial Technology and Archaeology: the Archaeological Applications of GIS. New York: Taylor and Francis.

Wheatley D. 1996. Between the Lines: The Role of GIS-based Predictive Modelling in the Interpretation of Extensive Survey Data. In: Kamermans H, Fennema K, editors. Interfacing the Past: Computer Applications and Quantitative Methods in Archaeology CAA95. Analecta Praehistorica Leidensia. p. 275–292.

White L. 1949. The Science of Culture: A Study of Man and Civilization. New York: Grove Press.

Willey GR. 1953. Prehistoric Settlement Patterns in the Viru Valley, Peru. Bureau of American Ethnology Bulletin 155:1–453.

Willey GR. 1956. Prehistoric Settlement Patterns in the New World. Greenwood Press.

Wilson, Jr. CW, Jewell JW, Luther ET. 1956. Pennsylvanian Geology of the Cumberland Plateau. Nashville, Tennessee: State of Tennessee, Department of Conservation, Division of Geology.

Wissler C. 1927. The Culture-Area Concept in Social Anthropology. American Journal of Sociology 32:881–891.

Zhang ZX, Zhang H., Zhou DW. 2010. Using GIS Spatial Analysis and Logistic Regression to Predict the Probabilities of Human-caused Grassland Fires. Journal of Arid Environments 74:386–393.

APPENDICES

Appendix A

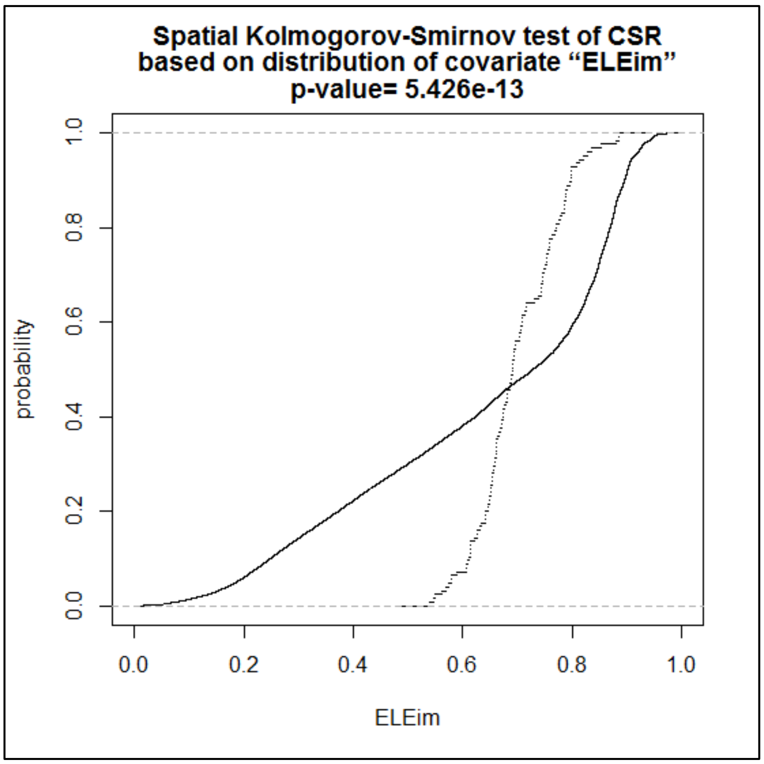
Descriptive Statistics for Preliminary Explanatory Variables

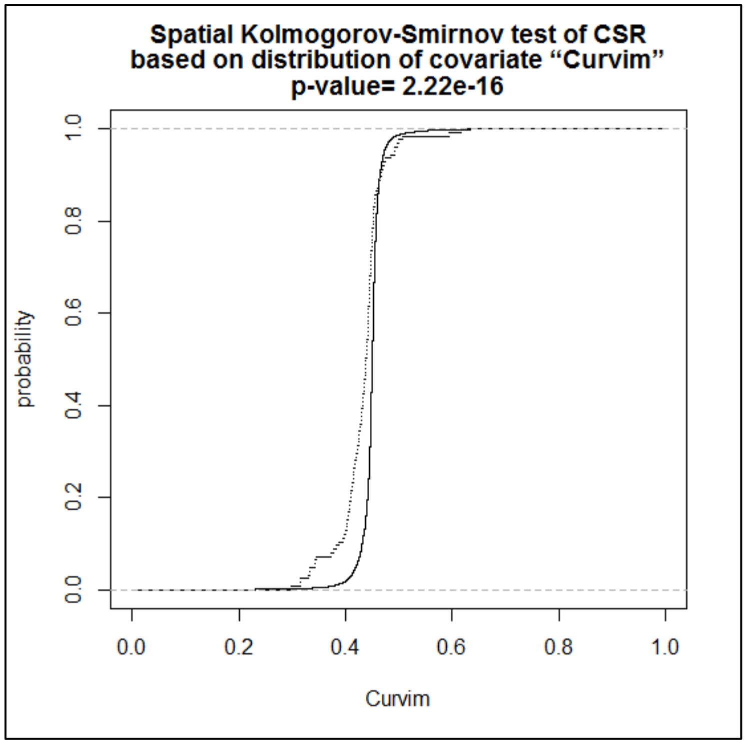
Variable (Abb.)	Unit	Minimum	Maximum	Mean	Standard Deviation
Solar	Wh/m^2	205,373.94	1,552,976.5	1,356,525.53	113,592.83
Eastness	unitless	-1	1	-0.0036	0.71
Northness	unitless	-1	1	0.0054	0.71
CDChest	min	0	184.03	19.58	23.6
CDNred	min	0	121.23	5.32	9.94
CDSred	min	0	483.26	115.39	87.87
CDScar	min	0	998.73	373.67	250.64
CDWhite	min	0	79.69	2.23	5.33
CDHick	min	0	184.03	18.21	24.01
CDWalnut	min	0	449.83	91.66	95.48
CDWater	min	0	175.62	24.17	23.11
Curv	$1/100^{th}$ °	-87.1	84.9	-0.33	1.63
DirDur	hrs/yr	175.84	4,366.25	9,948.87	337.01
ELE	m	198.4	567.7	429.1	87.63
PerMbh	%	0	100	5.04	21.26
PerMm	%	0	100	8.7	27.76
PerMp	%	0	100	5.7	22.33
PerPf	%	0	100	11.06	30.75
PerPr	%	0	00	42.35	48.97
VolWood	ft^3 /ac	0	114	77.8	11.31
SI100	m^3	-7,310.09	25,832.11	6,260.72	1,491.02
SI300	m^3	-140,798	331,900.75	55,114.86	35,892.58
SI1000	m^3	-4,705,172	5,580,292	595,394	944,090.89
Slope	°	0	75.0	10.81	8.55
Erosion	---	0	0.43	0.29	0.07
SoilThick	in	0	140	62.01	19.47
TerTex	m^2	0	903.56	3.49	9.28

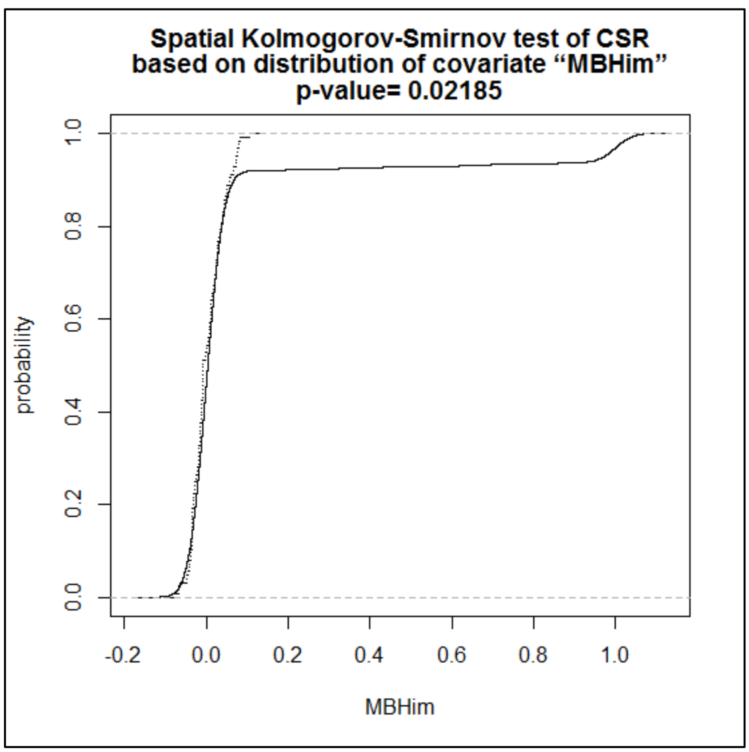
Appendix B

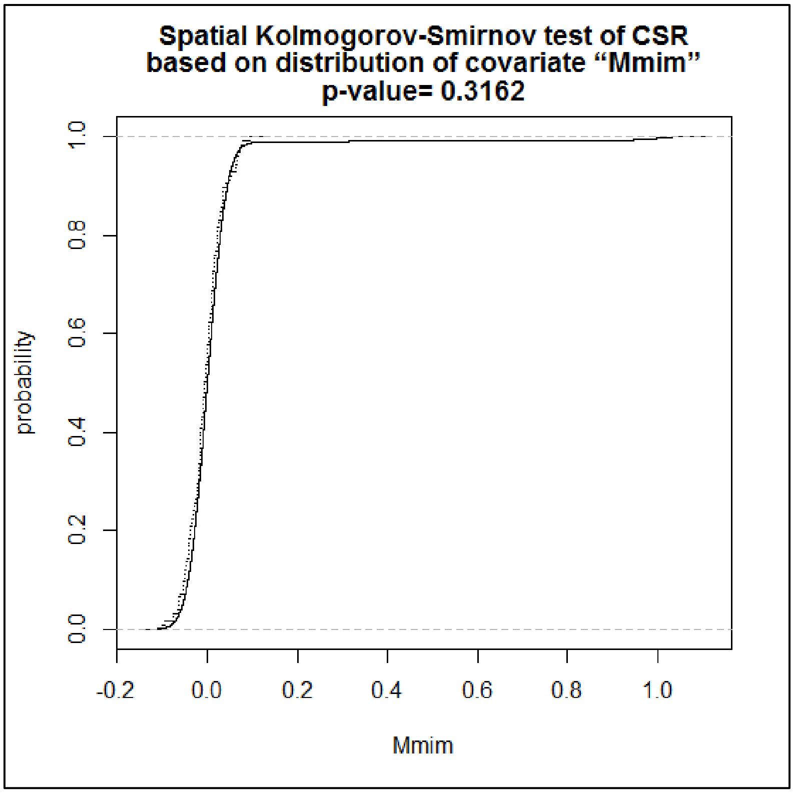
Kolmogorov-Smirnov Graphs Generated in R

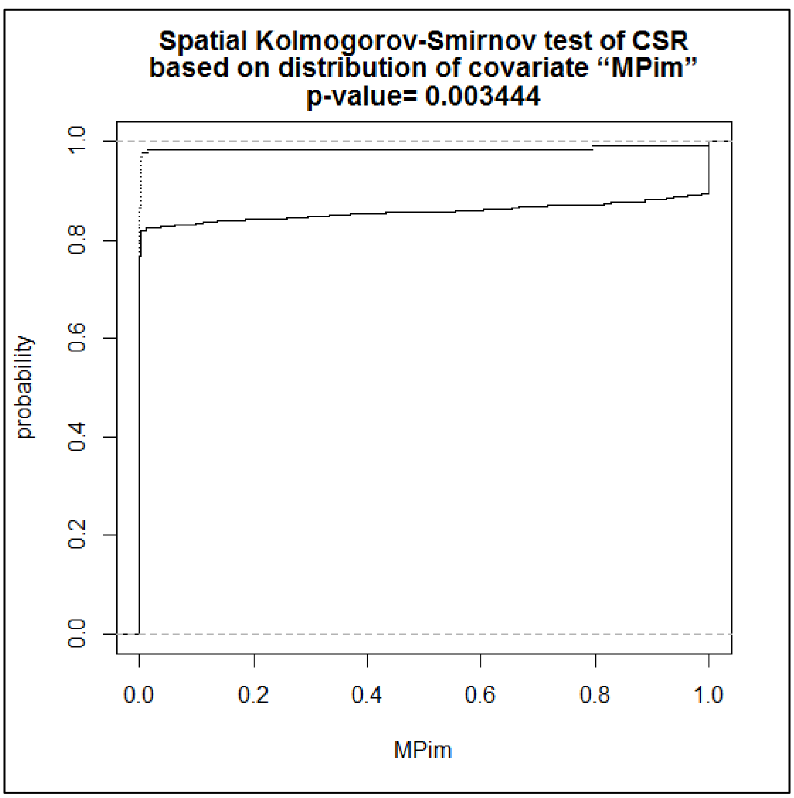
Below are the graphical results of the Kolmogorov-Smirnov (K-S) test for each explanatory variable that were generated in R (R Core Team 2012). The K-S test graphs show the predicted and observed (site presence) distributions. The predicted distribution is the smooth, dark line and the observed distribution is the lighter, jagged line. The distributions are significantly different at p-value < 0.05.

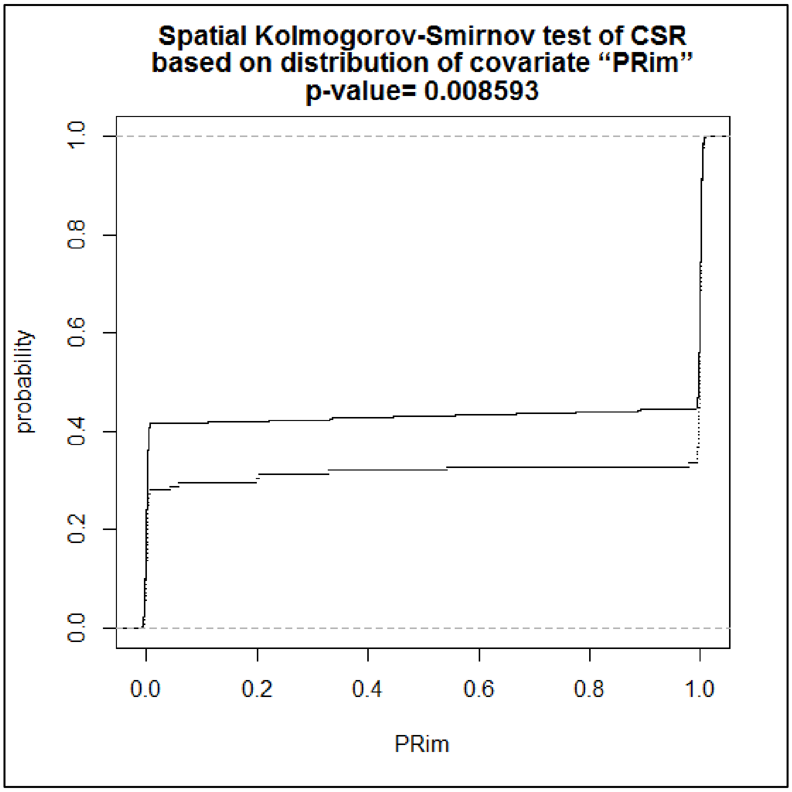


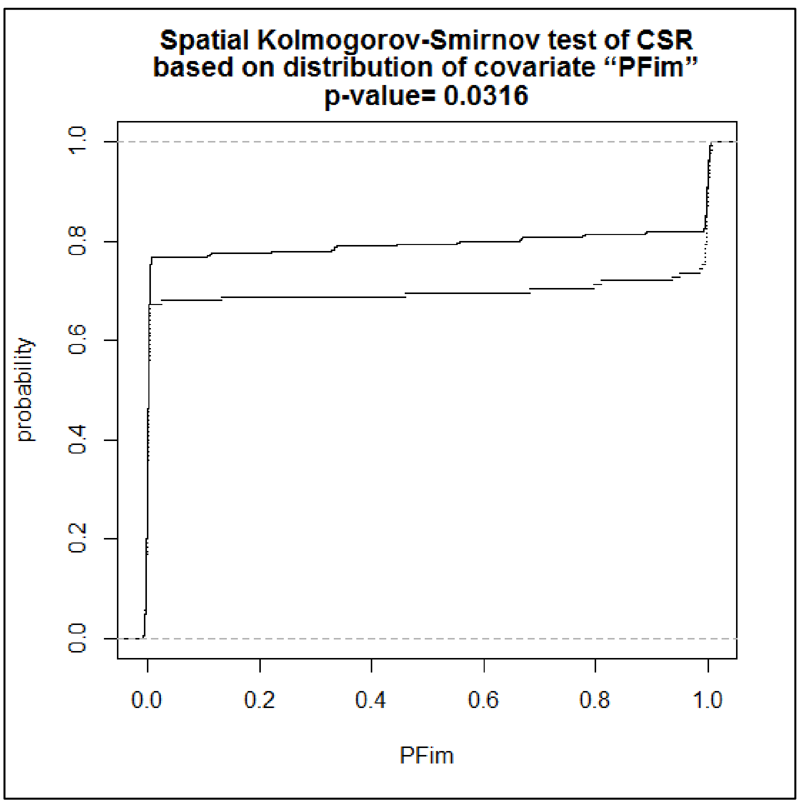

Spatial Kolmogorov-Smirnov test of CSR
based on distribution of covariate "Curvim"
p-value= 2.22e-16
probability
0.0
0.2
0.4
0.6
0.8
1.0
Curvim

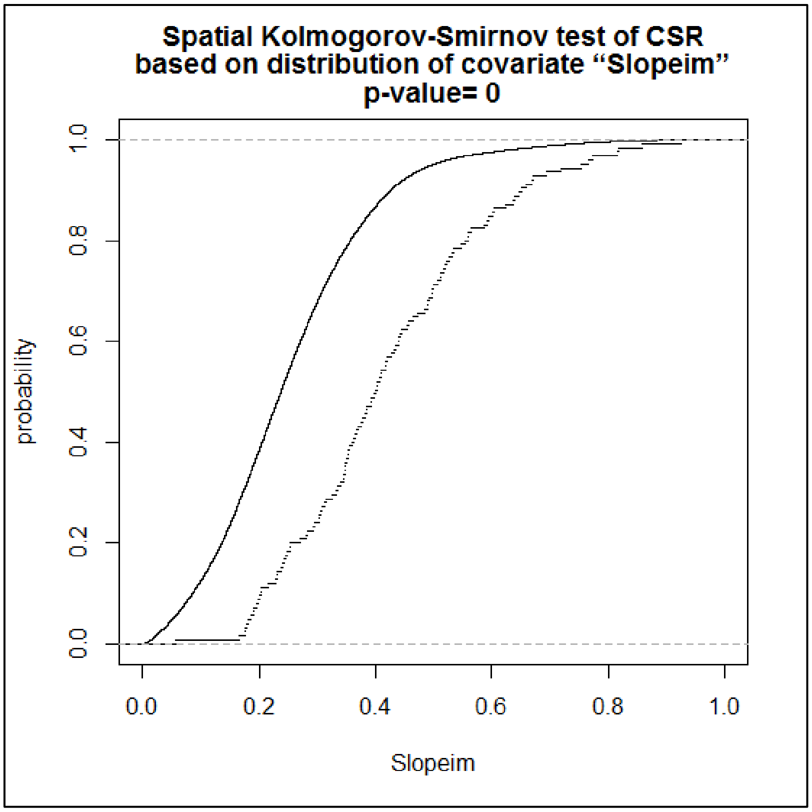

Spatial Kolmogorov-Smirnov test of CSR
based on distribution of covariate "MBHim"
p-value= 0.02185
probability
-0.2
0.0
0.2
0.4
0.6
0.8
1.0
MBHim

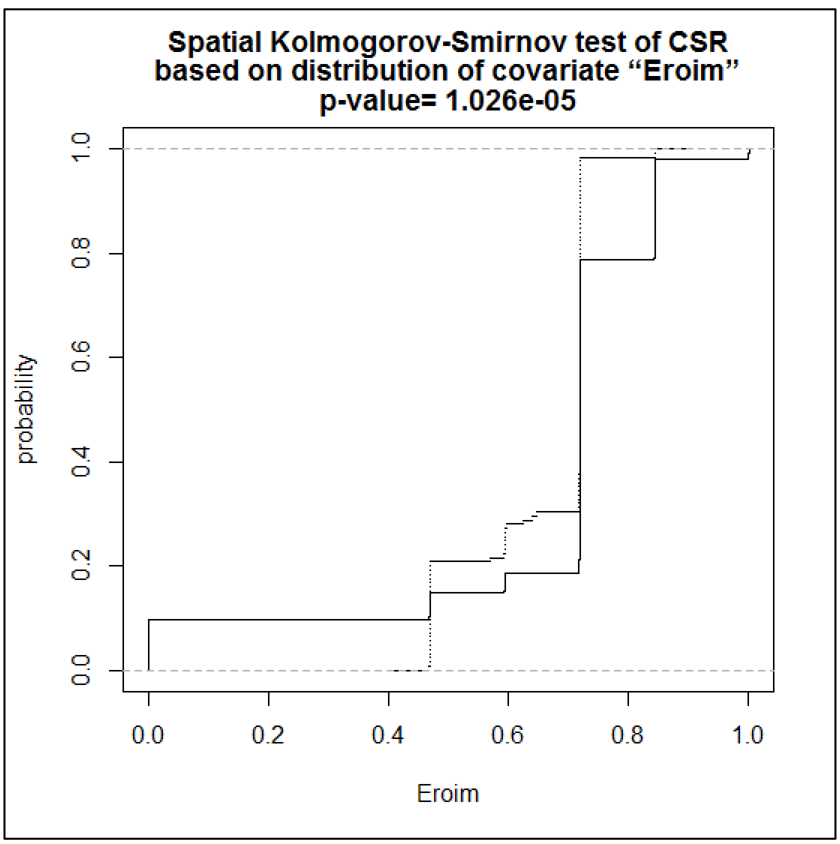

Spatial Kolmogorov-Smirnov test of CSR
based on distribution of covariate “Mmim”
p-value= 0.3162
probability
Mmim
0.0
0.2
0.4
0.6
0.8
1.0
-0.2

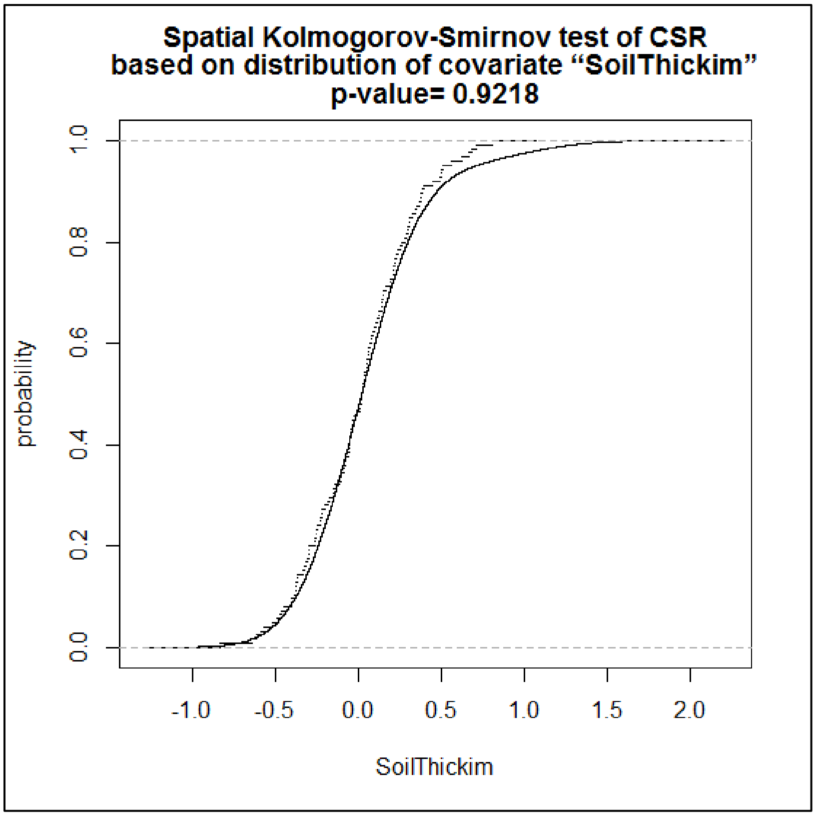

Spatial Kolmogorov-Smirnov test of CSR
based on distribution of covariate “MPim”
p-value= 0.003444
probability
MPim
0.0
0.2
0.4
0.6
0.8
1.0

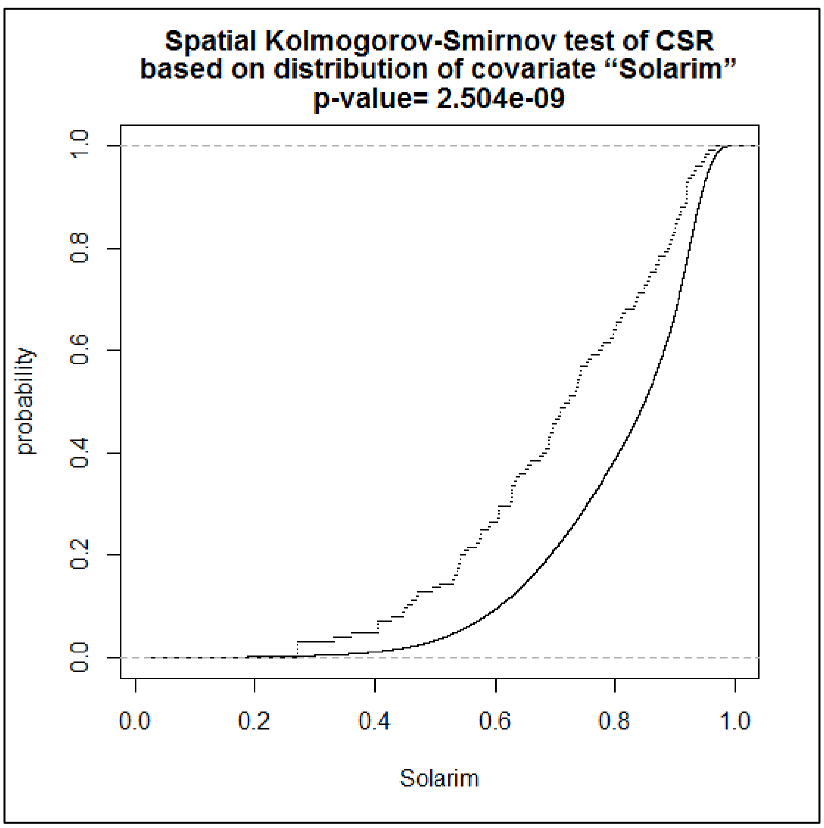

Spatial Kolmogorov-Smirnov test of CSR
based on distribution of covariate "PRim"
p-value= 0.008593
probability
0.0
0.2
0.4
0.6
0.8
1.0
PRim

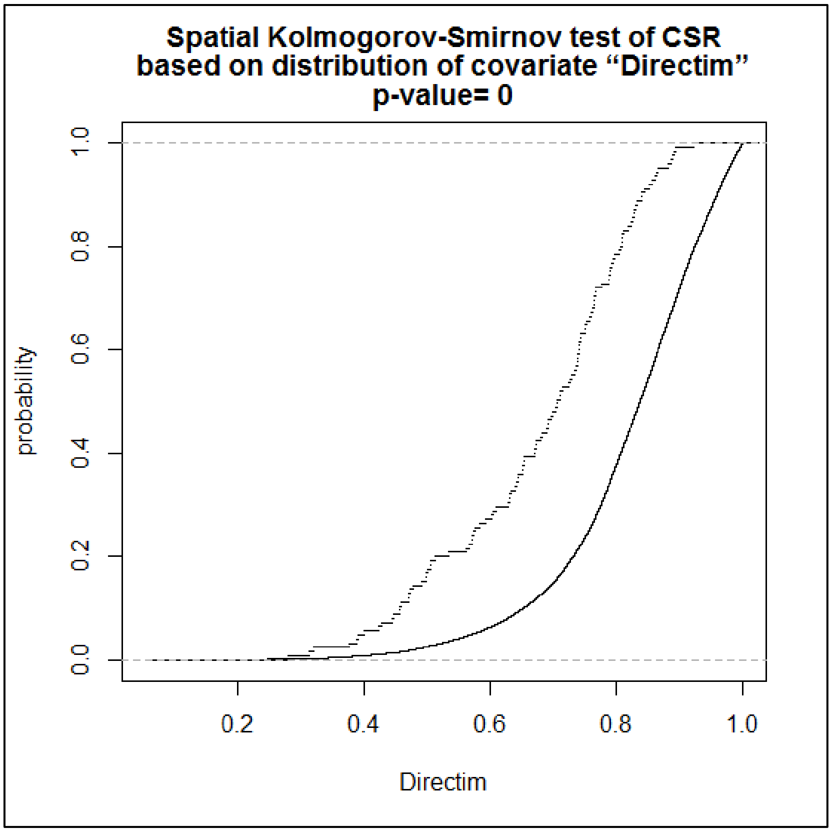

Spatial Kolmogorov-Smirnov test of CSR
based on distribution of covariate "PFim"
p-value= 0.0316
probability
0.0
0.2
0.4
0.6
0.8
1.0
PFim

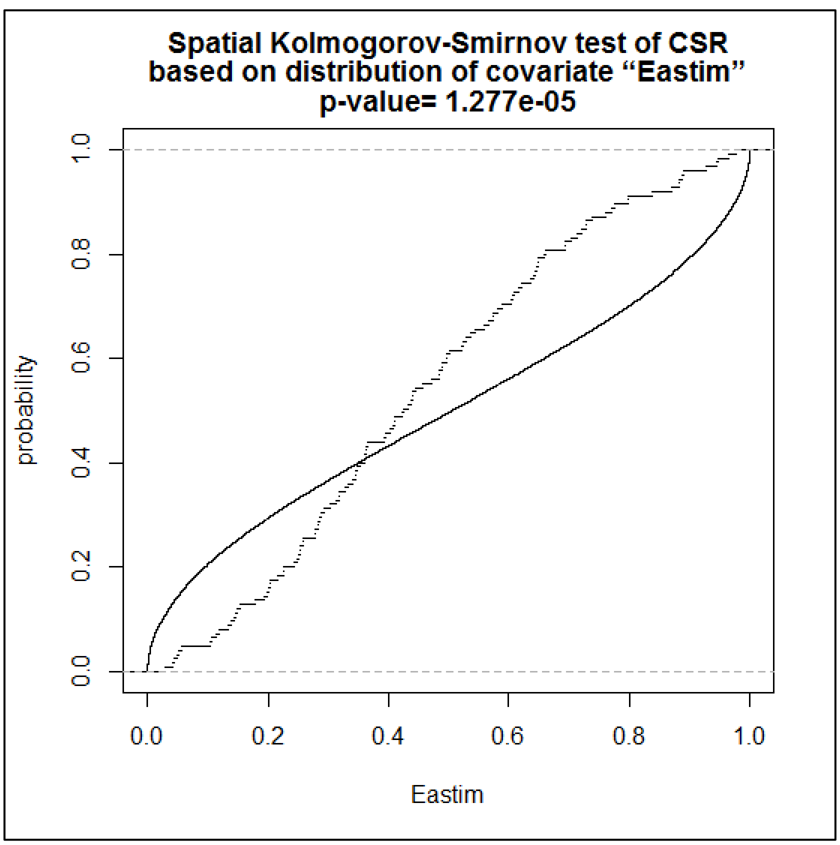

Spatial Kolmogorov-Smirnov test of CSR
based on distribution of covariate "Slopeim"
p-value= 0
probability
0.0
0.2
0.4
0.6
0.8
1.0
Slopeim

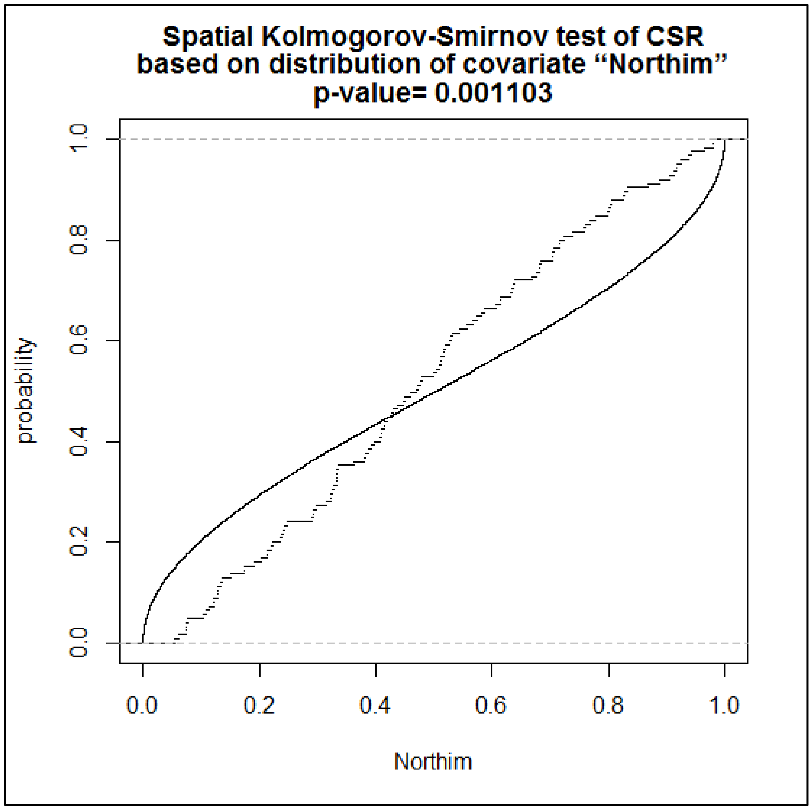

Spatial Kolmogorov-Smirnov test of CSR
based on distribution of covariate "Eroim"
p-value= 1.026e-05
probability
0.0
0.2
0.4
0.6
0.8
1.0
Eroim

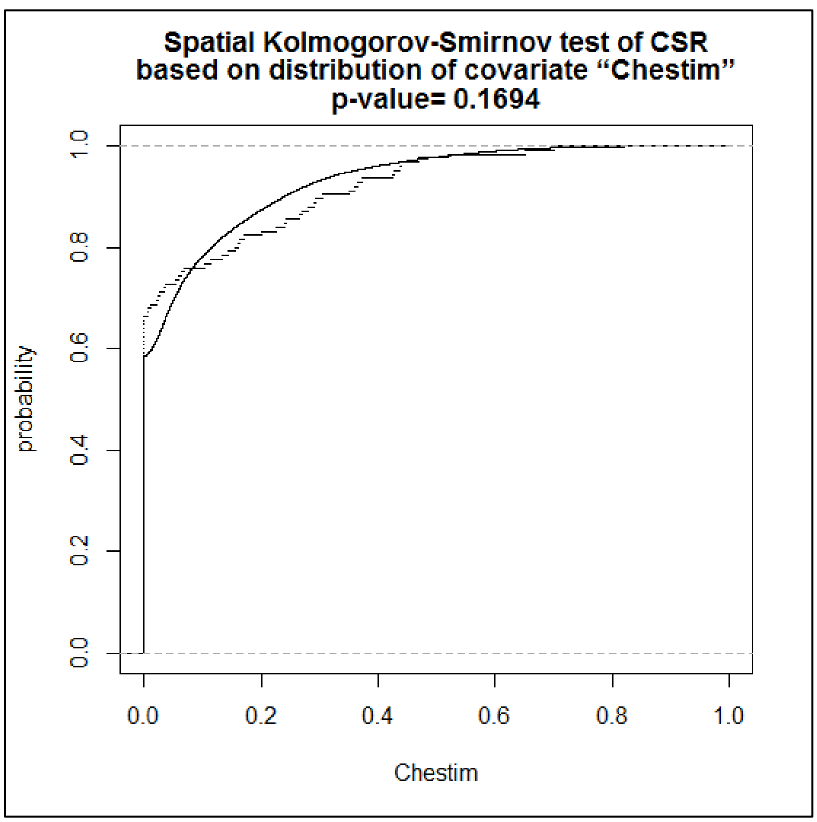

Spatial Kolmogorov-Smirnov test of CSR
based on distribution of covariate “SoilThickim”
p-value= 0.9218
probability
0.0
0.2
0.4
0.6
0.8
1.0
-1.0
-0.5
0.0
0.5
1.0
1.5
2.0
SoilThickim

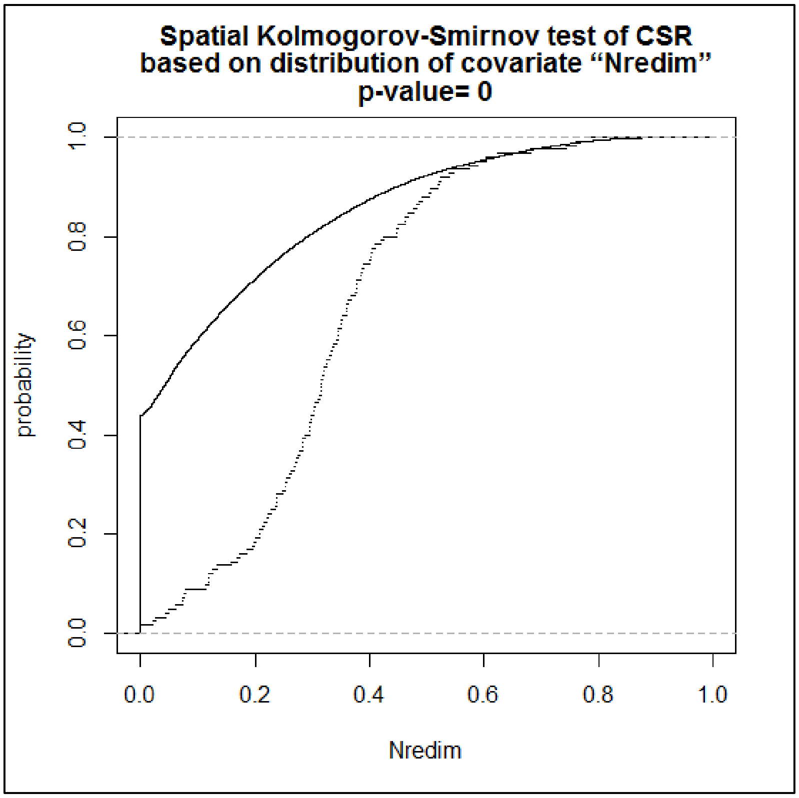

Spatial Kolmogorov-Smirnov test of CSR
based on distribution of covariate “Solarim”
p-value= 2.504e-09
probability
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.2
0.4
0.6
0.8
1.0
Solarim

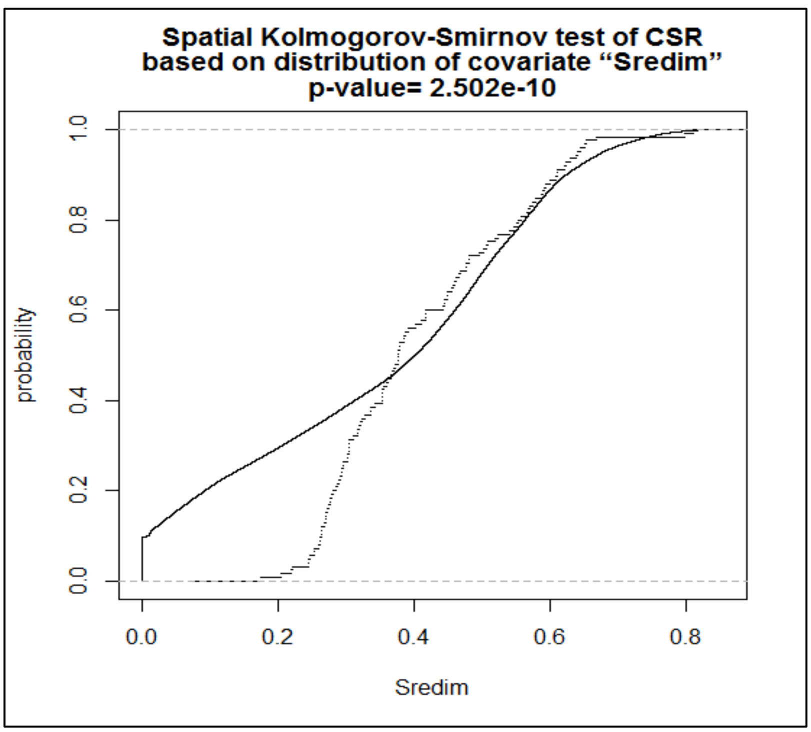

Spatial Kolmogorov-Smirnov test of CSR
based on distribution of covariate "Directim"
p-value= 0
probability
0.0
0.2
0.4
0.6
0.8
1.0
0.2
0.4
0.6
0.8
1.0
Directim

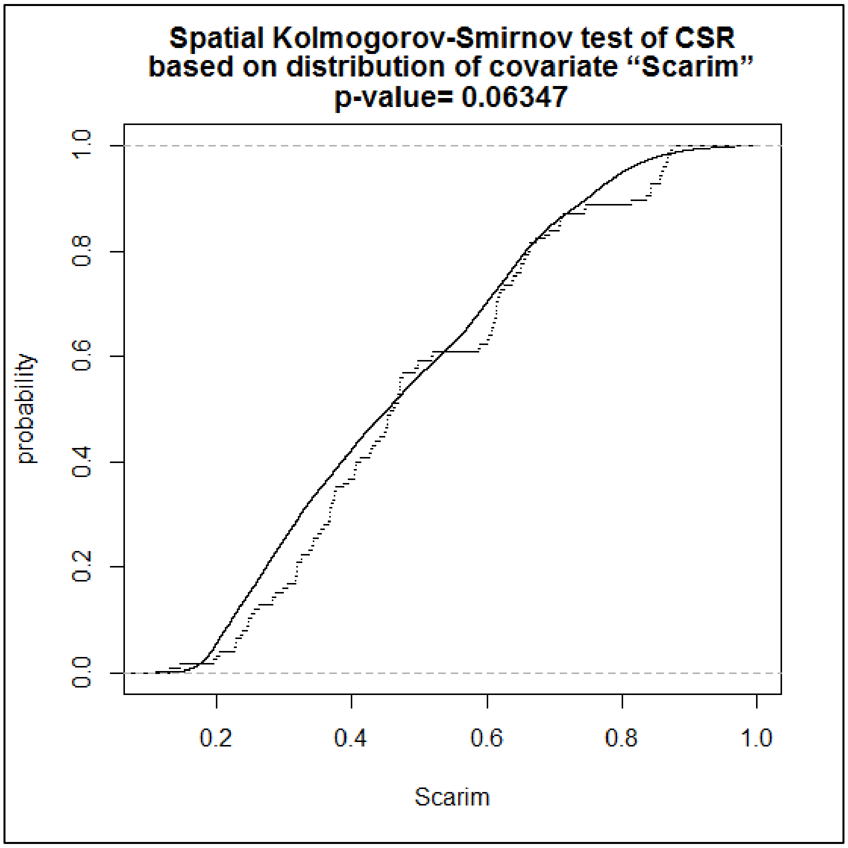

Spatial Kolmogorov-Smirnov test of CSR
based on distribution of covariate "Eastim"
p-value= 1.277e-05
probability
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.2
0.4
0.6
0.8
1.0
Eastim

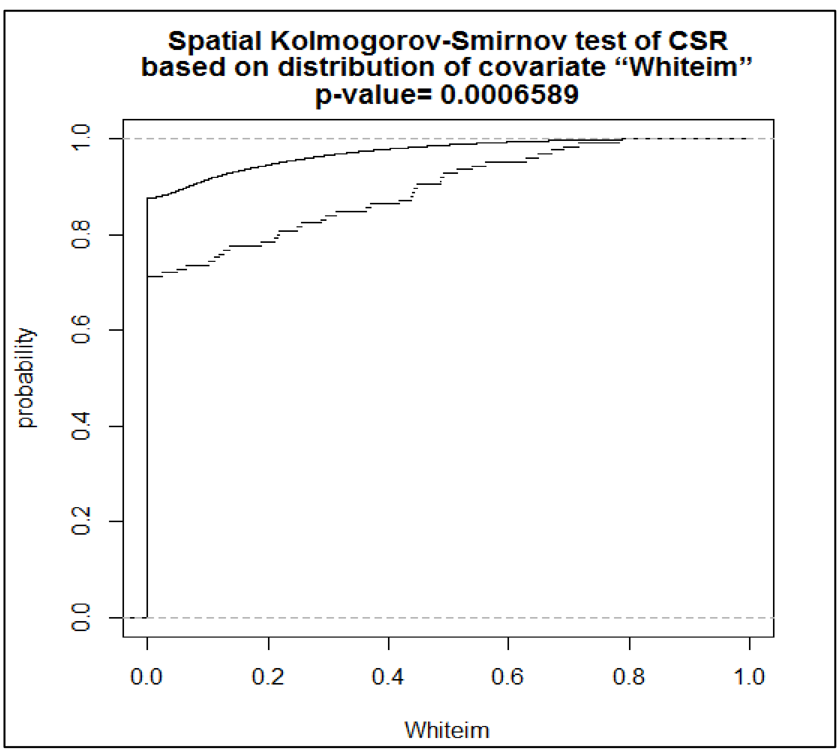

Spatial Kolmogorov-Smirnov test of CSR
based on distribution of covariate “Northim”
p-value= 0.001103
probability
0.0
0.2
0.4
0.6
0.8
1.0
Northim

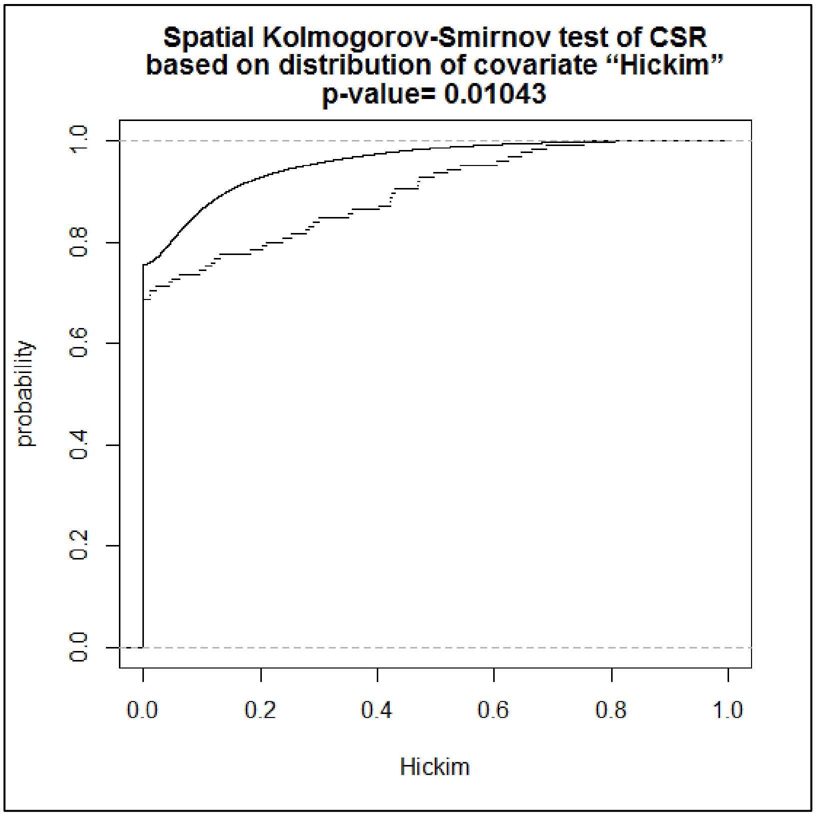

Spatial Kolmogorov-Smirnov test of CSR
based on distribution of covariate “Chestim”
p-value= 0.1694
probability
0.0
0.2
0.4
0.6
0.8
1.0
Chestim

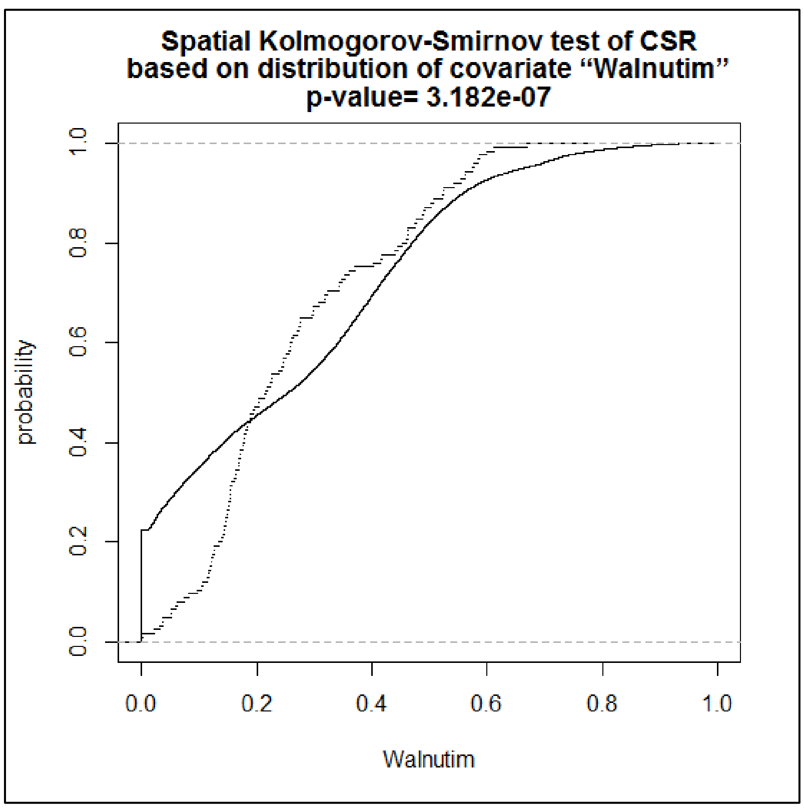

Spatial Kolmogorov-Smirnov test of CSR
based on distribution of covariate "Nredim"
p-value= 0
probability
0.0
0.2
0.4
0.6
0.8
1.0
Nredim

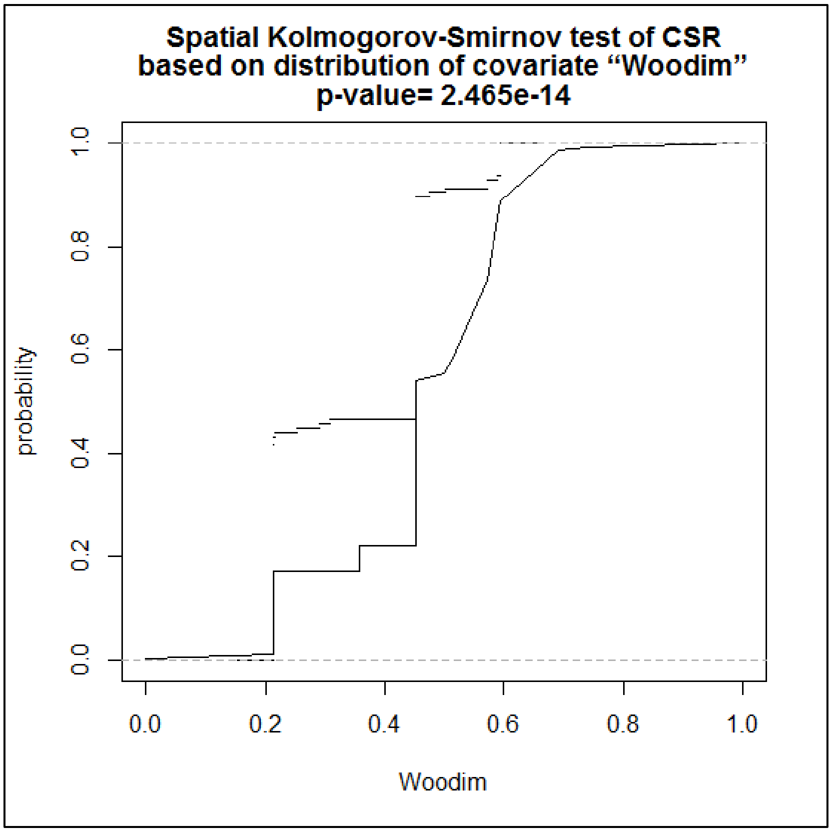

Spatial Kolmogorov-Smirnov test of CSR
based on distribution of covariate "Sredim"
p-value= 2.502e-10
probability
0.0
0.2
0.4
0.6
0.8
1.0
Sredim

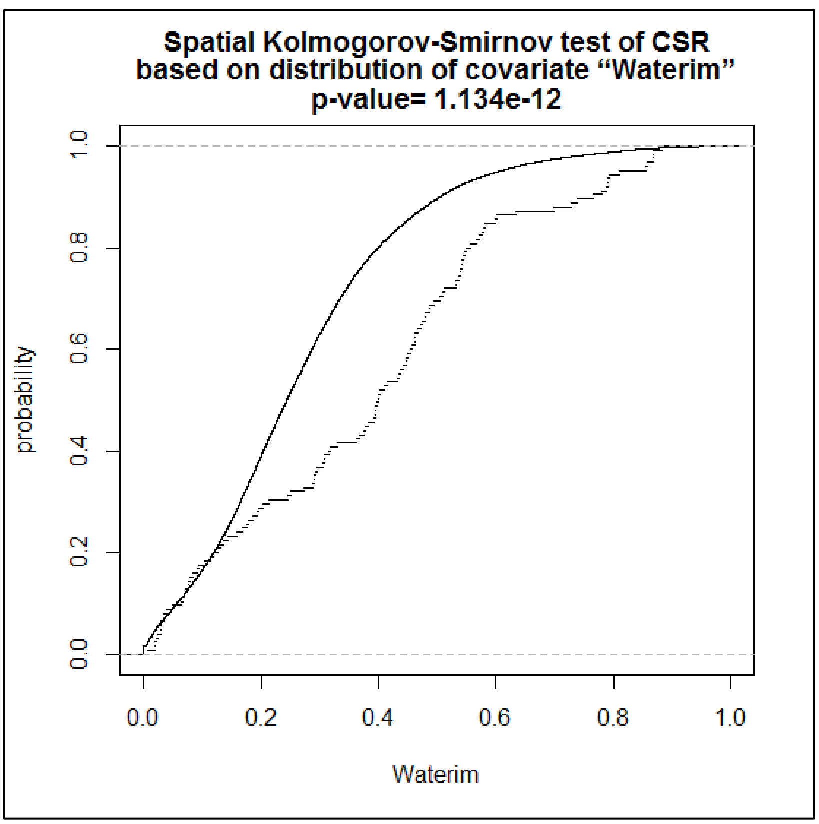

Spatial Kolmogorov-Smirnov test of CSR
based on distribution of covariate "Scarim"
p-value= 0.06347
probability
Scarim
0.0
0.2
0.4
0.6
0.8
1.0

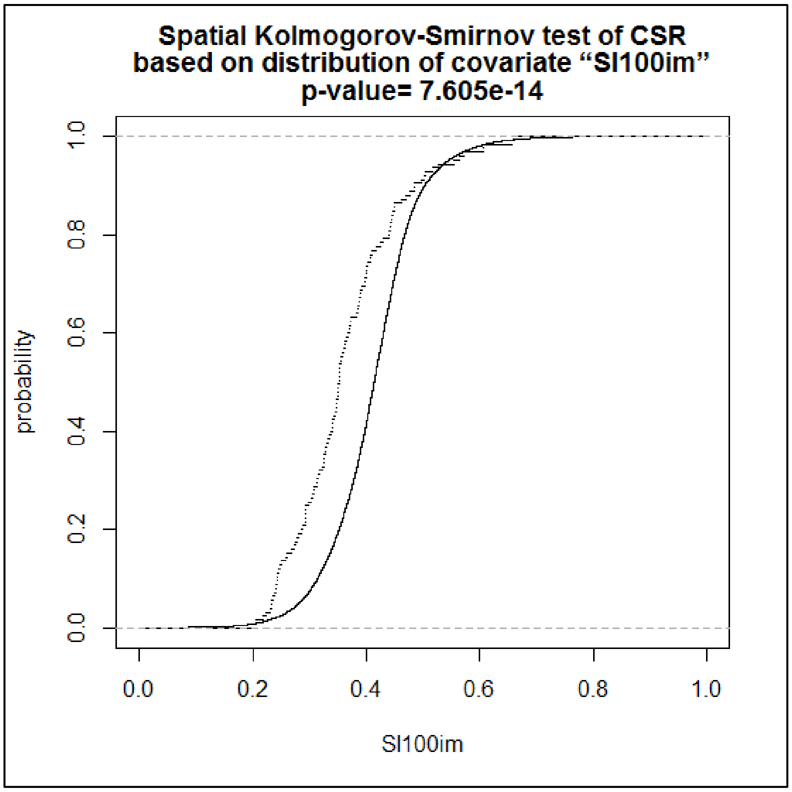

Spatial Kolmogorov-Smirnov test of CSR
based on distribution of covariate "Whiteim"
p-value= 0.0006589
probability
Whiteim
0.0
0.2
0.4
0.6
0.8
1.0

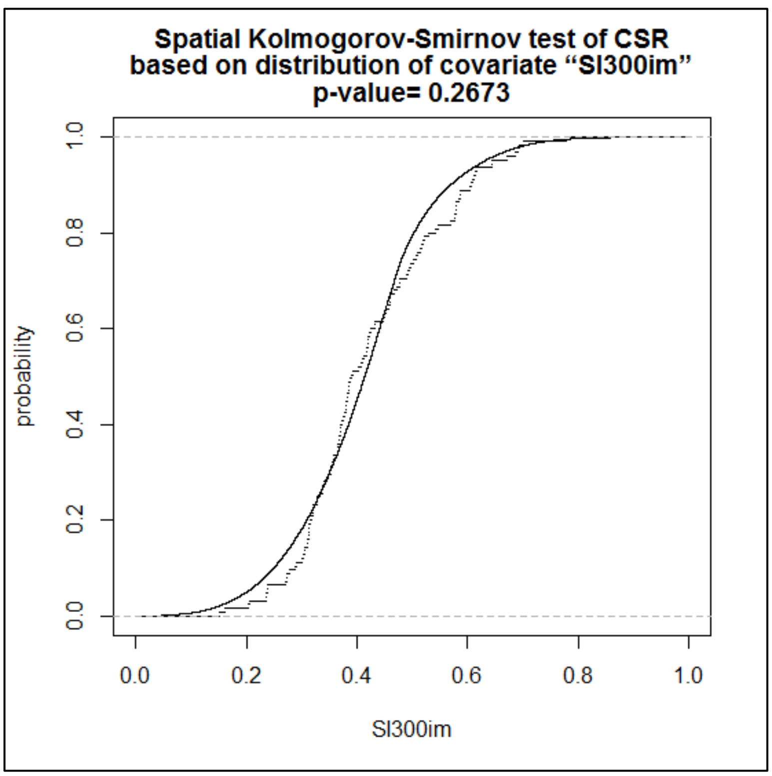

Spatial Kolmogorov-Smirnov test of CSR
based on distribution of covariate "Hickim"
p-value= 0.01043
probability
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.2
0.4
0.6
0.8
1.0
Hickim

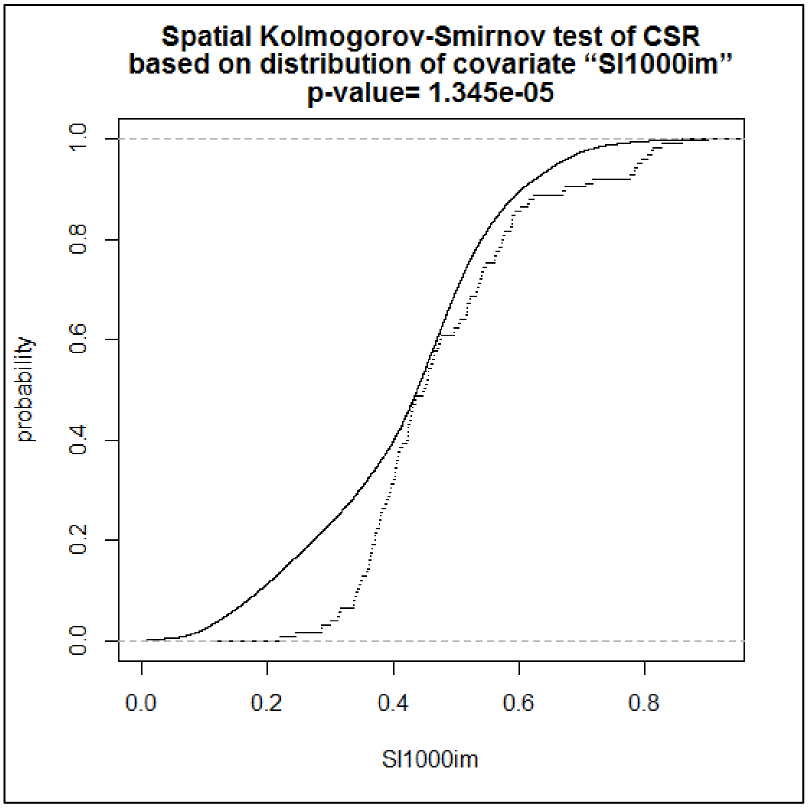

Spatial Kolmogorov-Smirnov test of CSR
based on distribution of covariate "Walnutim"
p-value= 3.182e-07
probability
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.2
0.4
0.6
0.8
1.0
Walnutim

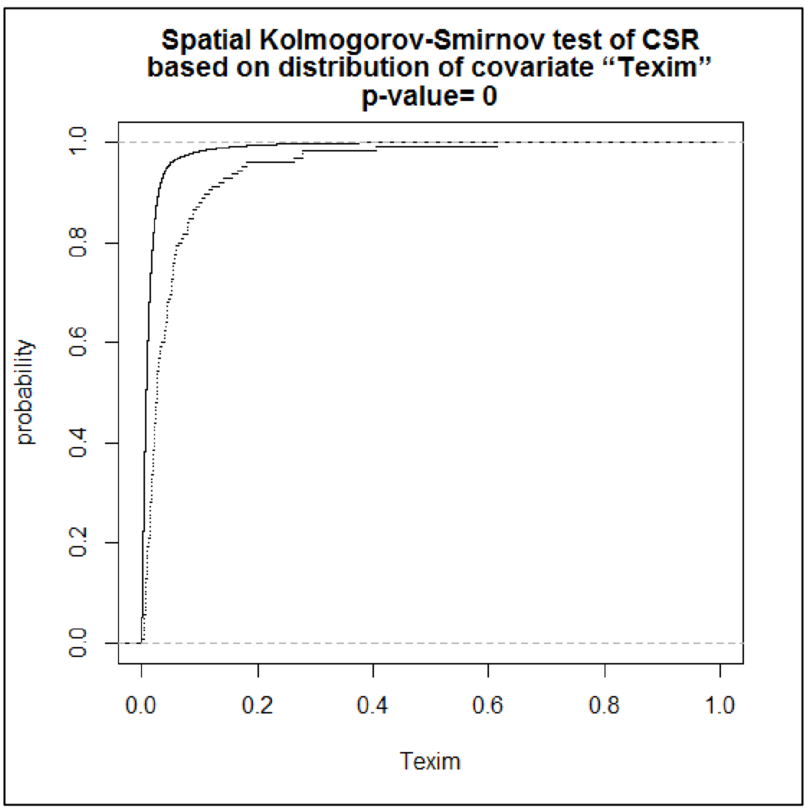

Spatial Kolmogorov-Smirnov test of CSR
based on distribution of covariate "Woodim"
p-value= 2.465e-14
probability
0.0
0.2
0.4
0.6
0.8
1.0
Woodim


Spatial Kolmogorov-Smirnov test of CSR
based on distribution of covariate "Waterim"
p-value= 1.134e-12
probability
0.0
0.2
0.4
0.6
0.8
1.0
Waterim


Spatial Kolmogorov-Smirnov test of CSR
based on distribution of covariate "SI100im"
p-value= 7.605e-14
probability
0.0
0.2
0.4
0.6
0.8
1.0
SI100im


Spatial Kolmogorov-Smirnov test of CSR
based on distribution of covariate "SI300im"
p-value= 0.2673
probability
0.0
0.2
0.4
0.6
0.8
1.0
SI300im


Spatial Kolmogorov-Smirnov test of CSR
based on distribution of covariate “SI1000im”
p-value= 1.345e-05
probability
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.2
0.4
0.6
0.8
SI1000im


Spatial Kolmogorov-Smirnov test of CSR
based on distribution of covariate “Texim”
p-value= 0
probability
0.0
0.2
0.4
0.6
0.8
1.0
0.0
0.2
0.4
0.6
0.8
1.0
Texim

Appendix C

Correlation Matrices Generated from Band Collection Statistics

P1 Model

	Slope	PerMP	PerMBH	Curv	Erosion	Elevation	PerPF	PerPr
Slope	1.00	0.04	-0.05	0.03	0.03	-0.20	0.36	-0.28
PerMP	0.04	1.00	-0.07	-0.01	-0.29	-0.53	-0.17	-0.50
PerMBH	-0.05	-0.07	1.00	-0.01	-0.29	-0.50	-0.14	-0.33
Curv	0.03	-0.01	-0.01	1.00	0.00	0.10	0.00	0.01
Erosion	0.03	-0.29	-0.29	0.00	1.00	0.38	0.16	0.26
Elevation	-0.20	-0.53	-0.50	0.10	0.38	1.00	-0.21	0.82
PerPF	0.36	-0.17	-0.14	0.00	0.16	-0.21	1.00	-0.59
PerPr	-0.28	-0.50	-0.33	0.01	0.26	0.82	-0.59	1.00

P2 Model

	SolarRad	Eastness	Northness	CDHick	CDWalnut	CDWater	DirDur	VolWood	SI100m	SI1000m	TerTex	CDWhite	CDNred	CDSred
Solar Rad	1.00	-0.01	-0.01	-0.03	0.25	-0.06	0.71	0.10	0.18	0.16	-0.43	-0.12	-0.10	0.23
Eastness	-0.01	1.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
Northness	-0.01	0.00	1.00	0.01	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.01	0.00	0.00
CDHick	-0.03	0.00	0.01	1.00	0.10	0.22	0.05	0.15	0.13	0.32	0.02	0.87	0.11	0.12
CDWalnut	0.25	0.00	0.00	0.10	1.00	-0.12	0.41	-0.08	0.15	0.41	-0.14	-0.03	0.21	0.89
CDWater	-0.06	0.00	0.00	0.22	-0.12	1.00	0.00	-0.18	0.26	0.62	0.15	0.30	0.34	-0.05
DirDur	0.71	0.00	0.00	0.05	0.41	0.00	1.00	0.17	0.56	0.38	-0.42	-0.08	-0.13	0.33
VolWood	0.10	0.00	0.00	0.15	-0.08	-0.18	0.17	1.00	-0.04	-0.16	-0.24	0.04	-0.47	-0.23
SI100m	0.18	0.00	0.00	0.13	0.15	0.26	0.56	-0.04	1.00	0.47	0.04	0.10	0.10	0.14
SI1000m	0.16	0.00	0.00	0.32	0.41	0.62	0.38	-0.16	0.47	1.00	0.02	0.26	0.32	0.49
TerTex	-0.43	0.00	0.00	0.02	-0.14	0.15	-0.42	-0.24	0.04	0.02	1.00	0.08	0.21	-0.07
CDWhite	-0.12	0.00	0.01	0.87	-0.03	0.30	-0.08	0.04	0.10	0.26	0.08	1.00	0.24	-0.01
CDNred	-0.10	0.00	0.00	0.11	0.21	0.34	-0.13	-0.47	0.10	0.32	0.21	0.24	1.00	0.27
CDSred	0.23	0.00	0.00	0.12	0.89	-0.05	0.33	-0.23	0.14	0.49	-0.07	-0.01	0.27	1.00

Appendix D

R Code for Running SLRM Function

Red Text = Input and Blue Text = Output

P1 Model

```
> P1run1 <-slrm(PresData ~ 1 + ELEim + Curvim + Slopeim + Eroim
+ MBHim + MPim + PFim)

> print(P1run1)

Fitted spatial logistic regression model

Formula: PresData ~ 1 + ELEim + Curvim + Slopeim + Eroim + MBHim
+ MPim + PFim


Fitted coefficients:

(Intercept) ELEim     Curvim     Slopeim    Eroim      MBHim
-9.6875546 0.4430579 -4.4996610 5.0355586 -2.5550436 -97.4354671
   MPim        PFim
 -3.2240448 -0.4171968


> anova(P1run1, test="Chi")

Analysis of Deviance Table
Model: binomial, link: logit
Response: PresData
Terms added sequentially (first to last)

        Df Deviance Resid. Df Resid. Dev  Pr(>Chi)
NULL                   173927     2059.4
ELEim    1    8.242    173926     2051.2  0.004093 **
Curvim   1   36.542    173925     2014.7 1.494e-09 ***
Slopeim  1  118.378    173924     1896.3 < 2.2e-16 ***
Eroim    1    6.548    173923     1889.7  0.010503 *
MBHim    1    7.195    173922     1882.5  0.007311 **
MPim     1   20.179    173921     1862.3 7.051e-06 ***
PFim     1    3.246    173920     1859.1  0.071580 .
---
Signif. codes:  0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1
```

P2 Model

```
> P2run2 <-slrm(PresData ~ 1 + Directim + Nredim + Sredim +
Whiteim + Hickim + Waterim + Woodim + SI100im + SI1000im +
Texim)

> print(P2run2)

Fitted spatial logistic regression model

Formula: PresData ~ 1 + Directim + Nredim + Sredim + Whiteim +
Hickim + Waterim + Woodim + SI100im + SI1000im + Texim


Fitted coefficients:

(Intercept) Directim    Nredim    Sredim     Whiteim     Hickim
-8.8224012 -3.5815320  0.7000523 2.0731965 9.1120472  -7.0056506
Waterim      Woodim      SI100im    SI1000im       Texim
0.5960222  -3.9423587  -4.9597100   4.1100522   3.3754449

> anova(P2run2, test="Chi")

Analysis of Deviance Table
Model: binomial, link: logit
Response: PresData
Terms added sequentially (first to last)

         Df Deviance Resid. Df Resid. Dev  Pr(>Chi)
NULL                    173927     2059.4
Directim  1  114.068    173926     1945.4 < 2.2e-16 ***
Nredim    1   56.926    173925     1888.4 4.525e-14 ***
Sredim    1    8.886    173924     1879.5 0.0028734 **
Whiteim   1   19.585    173923     1860.0 9.621e-06 ***
Hickim    1    4.228    173922     1855.7 0.0397648 *
Waterim   1   17.022    173921     1838.7 3.696e-05 ***
Woodim    1   21.413    173920     1817.3 3.703e-06 ***
SI100im   1    4.224    173919     1813.1 0.0398552 *
SI1000im  1   13.813    173918     1799.3 0.0002019 ***
Texim     1    8.176    173917     1791.1 0.0042453 **
---
Signif. codes:  0 ‘***’ 0.001 ‘**’ 0.01 ‘*’ 0.05 ‘.’ 0.1 ‘ ’ 1
```

Appendix E

Logistic Regression Output from SPSS

P1 Model

Case Processing Summary

Unweighted Cases[a]		N	Percent
Selected Cases	Included in Analysis	250	100.0
	Missing Cases	0	.0
	Total	250	100.0
Unselected Cases		0	.0
Total		250	100.0

a. If weight is in effect, see classification table for the total number of cases.

Dependent Variable Encoding

Original Value	Internal Value
0	0
1	1

Block 0: Beginning Block

Classification Table[a,b]

Observed			Predicted		
			PresAb		Percentage Correct
			0	1	
Step 0	PresAb	0	0	125	.0
		1	0	125	100.0
	Overall Percentage				50.0

a. Constant is included in the model.

b. The cut value is .500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 0	Constant	.000	.126	.000	1	1.000	1.000

Variables not in the Equation

			Score	df	Sig.
Step 0	Variables	PC_EleSTD	2.840	1	.092
		PC_CurvSTD	5.843	1	.016
		PC_SlopeST	63.484	1	.000
		PC_EroSTD	10.977	1	.001
		PC_MbhSTD	4.481	1	.034
		PC_MpSTD	11.172	1	.001
		PC_PfSTD	1.120	1	.290
	Overall Statistics		91.934	7	.000

Block 1: Method = Enter

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	121.856	7	.000
	Block	121.856	7	.000
	Model	121.856	7	.000

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	224.717[a]	.386	.514

a. Estimation terminated at iteration number 20 because maximum iterations has been reached. Final solution cannot be found.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	7.812	8	.452

Contingency Table for Hosmer and Lemeshow Test

	PresAb = 0	PresAb = 1	Total

		Observed	Expected	Observed	Expected	
Step 1	1	24	24.289	1	.711	25
	2	23	21.809	2	3.191	25
	3	23	19.641	2	5.359	25
	4	14	17.174	11	7.826	25
	5	14	14.689	11	10.311	25
	6	11	11.701	14	13.299	25
	7	7	7.998	18	17.002	25
	8	7	5.054	18	19.946	25
	9	1	2.184	24	22.816	25
	10	1	.460	24	24.540	25

Classification Table[a]

Observed			Predicted		
			PresAb		Percentage Correct
			0	1	
Step 1	PresAb	0	102	23	81.6
		1	30	95	76.0
	Overall Percentage				78.8

a. The cut value is .500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	PC_EleSTD	2.508	1.800	1.941	1	.164	12.275
	PC_CurvSTD	-8.473	4.088	4.296	1	.038	.000
	PC_SlopeST	9.212	1.434	41.264	1	.000	10020.148
	PC_EroSTD	-6.513	1.699	14.685	1	.000	.001
	PC_MbhSTD	-49.645	24196.751	.000	1	.998	.000
	PC_MpSTD	-3.286	1.560	4.440	1	.035	.037
	PC_PfSTD	-.299	.511	.342	1	.558	.741
	Constant	3.904	2.457	2.524	1	.112	49.601

a. Variable(s) entered on step 1: PC_EleSTD, PC_CurvSTD, PC_SlopeST, PC_EroSTD, PC_MbhSTD, PC_MpSTD, PC_PfSTD.

P2 Model

Case Processing Summary

Unweighted Cases[a]		N	Percent
Selected Cases	Included in Analysis	250	100.0
	Missing Cases	0	.0
	Total	250	100.0
Unselected Cases		0	.0
Total		250	100.0

a. If weight is in effect, see classification table for the total number of cases.

Dependent Variable Encoding

Original Value	Internal Value
0	0
1	1

Block 0: Beginning Block

Classification Table[a,b]

Observed			Predicted		
			PresAb		Percentage Correct
			0	1	
Step 0	PresAb	0	0	125	.0
		1	0	125	100.0
	Overall Percentage				50.0

a. Constant is included in the model.

b. The cut value is .500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 0	Constant	.000	.126	.000	1	1.000	1.000

Variables not in the Equation

			Score	df	Sig.
Step 0	Variables	PC_DirSTD	49.609	1	.000
		EastSTD	.161	1	.689
		NorthSTD	.341	1	.559
		PC_NredSTD	63.509	1	.000
		PC_SredSTD	2.809	1	.094

PC_WhiteST	17.219	1	.000
PC_HickSTD	13.504	1	.000
PC_WaterST	14.967	1	.000
PC_WoodSTD	36.430	1	.000
PC_100siST	14.117	1	.000
PC_1000siS	7.299	1	.007
PC_TexSTD	22.036	1	.000
Overall Statistics	113.874	12	.000

Block 1: Method = Enter

Omnibus Tests of Model Coefficients

		Chi-square	df	Sig.
Step 1	Step	163.668	12	.000
	Block	163.668	12	.000
	Model	163.668	12	.000

Model Summary

Step	-2 Log likelihood	Cox & Snell R Square	Nagelkerke R Square
1	182.906[a]	.480	.641

a. Estimation terminated at iteration number 7 because parameter estimates changed by less than .001.

Hosmer and Lemeshow Test

Step	Chi-square	df	Sig.
1	44.222	8	.000

Contingency Table for Hosmer and Lemeshow Test

		PresAb = 0		PresAb = 1		Total
		Observed	Expected	Observed	Expected	
Step 1	1	25	24.685	0	.315	25
	2	24	23.824	1	1.176	25
	3	23	21.810	2	3.190	25
	4	18	18.310	7	6.690	25
	5	18	15.269	7	9.731	25

6	9	10.839	16	14.161	25
7	3	6.167	22	18.833	25
8	2	2.852	23	22.148	25
9	1	1.152	24	23.848	25
10	2	.093	23	24.907	25

Classification Table[a]

Observed			Predicted		
			PresAb		Percentage Correct
			0	1	
Step 1	PresAb	0	109	16	87.2
		1	20	105	84.0
	Overall Percentage				85.6

a. The cut value is .500

Variables in the Equation

		B	S.E.	Wald	df	Sig.	Exp(B)
Step 1[a]	PC_DirSTD	-4.915	2.652	3.434	1	.064	.007
	EastSTD	-.475	.551	.743	1	.389	.622
	NorthSTD	-.811	.542	2.240	1	.134	.445
	PC_NredSTD	1.802	1.238	2.118	1	.146	6.062
	PC_SredSTD	4.348	1.502	8.378	1	.004	77.349
	PC_WhiteST	16.203	10.915	2.204	1	.138	10881401.319
	PC_HickSTD	-11.512	11.018	1.092	1	.296	.000
	PC_WaterST	-.365	1.371	.071	1	.790	.694
	PC_WoodSTD	-4.450	1.648	7.293	1	.007	.012
	PC_100siST	-6.882	3.445	3.991	1	.046	.001
	PC_1000siS	5.423	2.517	4.643	1	.031	226.603
	PC_TexSTD	41.356	12.884	10.303	1	.001	9133373061835 30620.000
	Constant	3.073	1.719	3.195	1	.074	21.600

a. Variable(s) entered on step 1: PC_DirSTD, EastSTD, NorthSTD, PC_NredSTD, PC_SredSTD, PC_WhiteST, PC_HickSTD, PC_WaterST, PC_WoodSTD, PC_100siST, PC_1000siS, PC_TexSTD.

Appendix F

Spatial Error Model Output from GeoDa

A distance-based spatial weights matrix was used to run an OLS regression in GeoDa (Anselin, Syabri, and Kho 2006)—a distance of 1410 meters was chosen after examination of a semivariogramThe Lagrange Multiplier (LM) is used to indicate whether a spatial error or spatial lag model is needed. Based on the output of the OLS, the LM was significant for both the lag and error terms. In this type of scenario, the value of the LM can be used to choose which spatial dependence model is best to use. Because the value of the LM error was the lowest, a spatial error model was used.

P1 Model

SUMMARY OF OUTPUT: SPATIAL ERROR MODEL - MAXIMUM LIKELIHOOD

ESTIMATION

Data set : PC_PresAb

Spatial Weight : 1410weights.gwt

Dependent Variable : PRESAB Number of Observations: 250

Mean dependent var : 0.500000 Number of Variables : 8

S.D. dependent var : 0.500000 Degrees of Freedom : 242

Lag coeff. (Lambda) : 0.108393

R-squared : 0.368035 R-squared (BUSE) : -

Sq. Correlation : - Log likelihood : -124.105182

Sigma-square : 0.157991 Akaike info criterion : 264.21

S.E of regression : 0.397481 Schwarz criterion : 292.382

Variable Coefficient Std.Error z-value Probability

CONSTANT	0.8305167	0.3303648	2.513938	0.0119392
PC_ELESTD	0.2262229	0.2511118	0.9008853	0.3676492
PC_CURVSTD	-0.7701032	0.5326202	-1.445877	0.1482119
PC_SLOPEST	1.391825	0.1519199	9.161573	0.0000000
PC_EROSTD	-0.7874126	0.1854922	-4.244989	0.0000219
PC_MBHSTD	-0.5639221	0.2573206	-2.191516	0.0284144
PC_MPSTD	-0.3743277	0.1503449	-2.489794	0.0127818
PC_PFSTD	-0.1053578	0.07536567	-1.397955	0.1621266
LAMBDA	0.1083928	0.3355448	0.3230352	0.7466687

REGRESSION DIAGNOSTICS

DIAGNOSTICS FOR HETEROSKEDASTICITY

RANDOM COEFFICIENTS

TEST	DF	VALUE	PROB
Breusch-Pagan test	7	7.615032	0.3677598

DIAGNOSTICS FOR SPATIAL DEPENDENCE

SPATIAL ERROR DEPENDENCE FOR WEIGHT MATRIX : 1410weights.gwt

TEST	DF	VALUE	PROB
Likelihood Ratio Test	1	0.07381001	0.7858681

======================== END OF REPORT

P2 Model

SUMMARY OF OUTPUT: SPATIAL ERROR MODEL - MAXIMUM LIKELIHOOD ESTIMATION

Data set : PC_PresAb

Spatial Weight : 1410weights.gwt

Dependent Variable : PRESAB Number of Observations: 250

Mean dependent var : 0.500000 Number of Variables : 13

S.D. dependent var : 0.500000 Degrees of Freedom : 237

Lag coeff. (Lambda) : -0.530098

R-squared : 0.458494 R-squared (BUSE) : -

Sq. Correlation : - Log likelihood : -105.185539

Sigma-square : 0.135377 Akaike info criterion : 236.371

S.E of regression : 0.367936 Schwarz criterion : 282.15

Variable	Coefficient	Std.Error	z-value	Probability
CONSTANT	1.158673	0.1840727	6.294652	0.0000000
PC_DIRSTD	-0.9321794	0.2530409	-3.683909	0.0002297
EASTSTD	-0.04683539	0.06982069	-0.6707953	0.5023508
NORTHSTD	-0.09782797	0.06620999	-1.477541	0.1395307
PC_NREDSTD	0.3230226	0.1709217	1.889886	0.0587731
PC_SREDSTD	0.374139	0.1746226	2.142558	0.0321485

PC_WHITEST	1.783052	1.129268	1.578946	0.1143485
PC_HICKSTD	-1.102964	1.164162	-0.9474317	0.3434188
PC_WATERST	-0.04108933	0.1711472	-0.2400818	0.8102670
PC_WOODSTD	-0.655538	0.2116072	-3.0979	0.0019491
PC_100SIST	-0.6138773	0.3849613	-1.594647	0.1107912
PC_1000SIS	0.7538912	0.290461	2.595499	0.0094454
PC_TEXSTD	0.6904599	0.393624	1.75411	0.0794116
LAMBDA	-0.5300982	0.488859	-1.084358	0.2782061

REGRESSION DIAGNOSTICS

DIAGNOSTICS FOR HETEROSKEDASTICITY

RANDOM COEFFICIENTS

TEST	DF	VALUE	PROB
Breusch-Pagan test	12	8.760904	0.7232054

DIAGNOSTICS FOR SPATIAL DEPENDENCE

SPATIAL ERROR DEPENDENCE FOR WEIGHT MATRIX : 1410weights.gwt

TEST	DF	VALUE	PROB
Likelihood Ratio Test	1	0.553799	0.4567696

======================== END OF REPORT

Appendix G

Comparison of Significance Values by Model Approach

P1 Variables	TLR	SEM	SLRM
Elevation	.164	0.3676492	0.004093
Curvature	.038	0.1482119	1.494e-09
Slope	.000	.000	<2.2e-16
Soil Erosion	.000	0.0000219	0.010503
PerMbh	.998	0.0284144	0.007311
PerMp	.035	0.0127818	7.051e-06

P2 Variables	TLR	SEM	SLRM
Direct Duration	.064	0.0002297	<2.2e-16
CD Northern Red Oak	.146	0.0587731	4.525e-14
CD Southern Red Oak	.004	0.0321485	0.0028734
CD White Oak	.138	0.1143485	9.621e-06
CD Hickory	.296	0.3434188	0.0397648
CD Water	.79	0.8102670	3.696e-05
Potential Vol. Wood	.007	0.1107912	3.703e-06
100 m Shelter Index	.046	0.0094454	0.0398552
1000m Shelter Index	.031	0.0094454	0.0002019
Terrain Texture	.074	0.0794116	0.0042453

p value > 0.05 (not significant)

Appendix H

Zonal Statistics for Cumulative UCP Model

	Potential	MIN	MAX	RANGE	MEAN	STD
Elevation (m)	Very Low	198.4	567.7	369.3	413.2979	93.58974
	Low	198.4	567.1	368.7	458.1239	83.12529
	Moderate	198.4	567.2	368.8	452.4184	80.94522
	High	198.4	563.9	365.5	442.2866	75.34406
	Very High	198.4	551.5	353.1	457.437	56.26306

	Potential	MIN	MAX	RANGE	MEAN	STD
Earth Curvature	Very Low	-59.1	51.10001	110.2	0.068434	1.038539
	Low	-13.7	32.49997	46.19995	0.077809	1.164348
	Moderate	-21.4001	64.29996	85.70001	0.131055	1.250927
	High	-51.2999	84.89996	136.1999	-0.05093	2.008461
	Very High	-87.1	72	159.1	-1.42765	4.733228

	Potential	MIN	MAX	RANGE	MEAN	STD
Slope (°)	Very Low	0	63.12169	63.12169	7.89603	6.718626
	Low	0	57.99117	57.99117	7.740254	6.336202
	Moderate	0	64.89958	64.89958	8.922302	6.590149
	High	0	70.22811	70.22811	15.68087	8.110688
	Very High	0	75.0257	75.0257	28.11991	10.74406

	Potential	MIN	MAX	RANGE	MEAN	STD
Kw Factor for Potential Soil Erosion	Very Low	0	0.43	0.43	0.303169	0.078175
	Low	0.05	0.43	0.379999	0.28364	0.054306
	Moderate	0.05	0.43	0.379999	0.272435	0.058212
	High	0	0.43	0.43	0.272712	0.058515
	Very High	0	0.43	0.43	0.269997	0.050841

	Potential	MIN	MAX	RANGE	MEAN	STD
Percent of Bangor Limestone & Hartselle Formation	Very Low	0	100	100	12.02074	31.59688
	Low	0	0	0	0	0
	Moderate	0	0	0	0	0
	High	0	0	0	0	0
	Very High	0	0	0	0	0

Percent of Pennington Formation	Potential	MIN	MAX	RANGE	MEAN	STD
	Very Low	0	100	100	11.75483	31.25431
	Low	0	100	100	3.856592	18.48858
	Moderate	0	100	100	2.647881	15.17728
	High	0	100	100	1.236625	9.833561
	Very High	0	100	100	0.785245	6.724248

Direct Duration of Incoming Solar Radiation (hrs/yr)	Potential	MIN	MAX	RANGE	MEAN	STD
	Very Low	1103.865	4366.245	3262.38	4042.093	279.4386
	Low	2072.047	4366.245	2294.198	4078.83	254.6976
	Moderate	1642.032	4366.245	2724.213	4041.543	267.8388
	High	1049.788	4366.245	3316.457	3808.853	329.5793
	Very High	175.8363	4352.755	4176.919	3225.272	482.3553

Cost Distance to Supporting Zones of Northern Red Oak (min)	Potential	MIN	MAX	RANGE	MEAN	STD
	Very Low	0	113.2907	113.2907	2.378269	5.238659
	Low	0	103.2941	103.2941	2.873141	6.160607
	Moderate	0	120.6932	120.6932	3.508843	7.179666
	High	0	121.2794	121.2794	9.779979	12.88304
	Very High	0	121.1599	121.1599	22.90339	15.98167

Cost Distance to Supporting Zones of Southern Red Oak (min)	Potential	MIN	MAX	RANGE	MEAN	STD
	Very Low	0	485.1584	485.1584	96.96442	78.7146
	Low	0	472.3633	472.3633	135.4598	78.73288
	Moderate	0	481.0388	481.0388	134.2686	86.68871
	High	0	475.6359	475.6359	134.3481	95.7395
	Very High	0	419.4442	419.4442	132.1754	87.68176

Cost Distance to Supporting Zones of White Oak (min)	Potential	MIN	MAX	RANGE	MEAN	STD
	Very Low	0	72.39667	72.39667	1.847887	4.148276
	Low	0	72.89396	72.89396	2.2017	4.903661
	Moderate	0	78.49101	78.49101	2.230272	5.008158
	High	0	80.00523	80.00523	2.654863	6.367507
	Very High	0	76.54548	76.54548	5.664675	9.807933

Cost Distance to Supporting Zones of Hickory (min)	Potential	MIN	MAX	RANGE	MEAN	STD
	Very Low	0	184.0318	184.0318	27.28857	27.70941
	Low	0	154.0818	154.0818	22.57851	19.76546
	Moderate	0	165.9545	165.9545	15.7714	18.51354
	High	0	160.1159	160.1159	7.193517	14.4102
	Very High	0	141.5042	141.5042	7.139106	12.70974

Cost Distance to Nearest Water Source (min)	Potential	MIN	MAX	RANGE	MEAN	STD
	Very Low	0	154.4588	154.4588	20.74524	21.27078
	Low	0	152.3128	152.3128	17.18684	17.20014
	Moderate	0	162.6823	162.6823	20.01039	17.64831
	High	0	166.5214	166.5214	31.30694	24.76869
	Very High	0	175.6186	175.6186	46.71612	33.85068

Average Potential Volume of Wood Fiber (ft^3 /ac)	Potential	MIN	MAX	RANGE	MEAN	STD
	Very Low	0	114	114	81.46689	9.959917
	Low	0	114	114	80.39365	7.112513
	Moderate	0	114	114	78.99126	7.639501
	High	0	114	114	73.87251	9.839685
	Very High	0	114	114	57.67532	26.49762

100m Shelter Index	Potential	MIN	MAX	RANGE	MEAN	STD
	Very Low	-2974.3	25603.52	28577.81	6434.688	1137.058
	Low	-2510.8	20248.61	22759.41	6494.506	1253.519
	Moderate	-3376.49	25218.89	28595.38	6592.334	1374.113
	High	-4614.69	25832.11	30446.8	6194.252	1840.936
	Very High	-7310.09	21087.2	28397.3	4997.708	2442.179

1000m Shelter Index	Potential	MIN	MAX	RANGE	MEAN	STD
	Very Low	-4075172	5252821	9327993	494832.1	894974.6
	Low	-3711468	5113608	8825076	660271.7	783922.8
	Moderate	-3974394	5187558	9161952	641442.3	896419.9
	High	-3987755	5580292	9568047	778984	1062584
	Very High	-3228353	5482441	8710794	1095411	1289830

Terrain Texture	Potential	MIN	MAX	RANGE	MEAN	STD
	Very Low	0	289.3893	289.3893	2.455691	4.352556
	Low	0	192.4216	192.4216	2.345312	4.62612
	Moderate	0	321.9063	321.9063	2.881358	5.131256
	High	0	611.2885	611.2885	7.576442	9.836127
	Very High	0	903.56	903.56	27.72758	34.86284

VITA

LUCINDA M. LANGSTON

Personal: Date of Birth: May 19, 1986

Place of Birth: Eglin Air Force Base, Florida

Education: South Walton High School, Santa Rosa Beach, FL. Valedictorian, 2004

B.S. Geography, Anthropology Minor, East Tennessee State University, Johnson City, TN, 2009

M.S. Geosciences, Geospatial Concentration, East Tennessee State University, Johnson City, TN, 2013

Professional: Teaching Associate, Introduction to Geographic Information Systems, East Tennessee State University, Department of Geosciences, 2012-2013

GIS Analyst and Graduate Research Assistant, Pickett State Forest Archaeological Survey, 2010-2013

GIS Analyst, Pogue Creek State Natural Area Archaeological Survey, 2006-2010

Publications: Langston, LM. 2013. A Preliminary Predictive Model of Prehistoric

Rock Shelter Selection on the Upper Cumberland Plateau of Tennessee, Paper submitted to Upland Archaeology in the East—Archaeology of the Blue Ridge and Beyond: Symposium XI (accepted, in revision).

Franklin JD, Hays MA, Sherwood SC, Langston LM. 2012. An Integrated Approach: Lithic Analyses and Site Function, Eagle Drink Bluff Shelter, Upper Cumberland Plateau, Tennessee. In: Carr PJ, Bradbury AP, Price SP, editors. Contemporary Lithic Analysis in the Southeast: Problems, Solutions, and Interpretations. Tuscaloosa: University Alabama Press. p. 128–145.

Franklin JD, Price EK, Langston LM. 2010. Mortuary Practices at the Holliston Mills Site, a Mississippian Town in Upper East Tennessee. In: Sullivan LP, Mainfort, Jr. RC, editors. Mississippian Mortuary Practices: Beyond Hierarchy and the Representationist Perspective. Gainesville: University of Florida Press. p. 325-350.

Grants & Awards: Graduate Research Grant, Tennessee Council for Professional Archaeology, 2013

Graduate Student Paper Award (1st Place), Upland Archaeology in the East—Archaeology of the Blue Ridge and Beyond: Symposium XI, 2012