5-1-2011

Mapping current range with respect to abiotic site factors of selected southern oaks (Quercus spp) in Mississippi

Laura Kim Hobbs

MAPPING CURRENT RANGE WITH RESPECT TO ABIOTIC SITE FACTORS OF

SELECTED SOUTHERN OAKS (QUERCUS SPP) IN MISSISSIPPI

By

Laura Kim Hobbs

A Thesis
Submitted to the Faculty of
Mississippi State University
in Partial Fulfillment of the Requirements
for the Degree of Master of Science
in Forestry
in the Department of Forestry

Mississippi State, Mississippi

April 2011

MAPPING CURRENT RANGE WITH RESPECT TO ABIOTIC SITE FACTORS OF SELECTED SOUTHERN OAKS (QUERCUS SPP) IN MISSISSIPPI

By

Laura Kim Hobbs

Approved:

David L. Evans
Professor of Forestry
(Director of Thesis)

Gary N. Ervin
Associate Professor of Biological Sciences
(Committee Member)

Randall J. Rousseau
Associate Extension/Research Professor of Forestry
(Committee Member)

Andrew W. Ezell
Head and Graduate Coordinator of the Department of Forestry

George M. Hopper
Dean of the College of Forest Resources

Name: Laura Kim Hobbs

Date of Degree: April 29, 2011

Institution: Mississippi State University

Major Field: Forestry

Major Professor: Dr. David L. Evans

Title of Study: MAPPING CURRENT RANGE WITH RESPECT TO ABIOTIC SITE FACTORS OF SELECTED SOUTHERN OAKS (QUERCUS SPP) IN MISSISSIPPI

Pages in Study: 94

Candidate for Degree of Master of Science

Physical characteristics of a site that plant species inhabit may be useful in deriving the range of the species. Current range maps for tree species of the United States were originally developed by Elbert Little. These range maps were based primarily on observations. The purpose of this study was to update Little's (Little, 1971) range maps of select southern oak species in Mississippi by calculating the topological, soil, and climatic features of sites using a Geographic Information System (GIS) to analyze environmental variables associated with species distributions. Data collected from databases were input into ArcMap and site data extracted using Hawth's Analyst Tools. Stepwise logistic regression performed with site variables yielded the parameters used in predictive models to generate probability maps for each species across Mississippi. These probability maps demonstrate the potential to efficiently manage forests by giving a more encompassing view of species occurrence.

DEDICATION

To my mother, Jenny Hobbs, whom I miss daily and who is now watching over me in Heaven. This is for you, for always encouraging me and seeing endless potential. May I always make you proud.

ACKNOWLEDGEMENTS

. There is a verse from the Bible, Romans 8:28 that says: "And we know that all things work together for good to them that love God, to them who are the called according to [his] purpose." This verse was quoted many times to me during one of the darkest times in my life. It is for this reason and many more that I want to first thank God. He has strengthened me and helped me even when I wanted to give up and fought Him every step of the way. I thank God for His constant love, mercy, and guidance, without Him, I can do nothing. Thank you for working this to the benefit of myself and may I one day use it to bring you honor and glory. Graduate school, life, and this study have shown this verse to be true in many ways. Especially if I had been told a month before applying to graduate school that one day I would be graduating with a Masters degree, I would have laughed. Had things worked the way I had planned them I never would have met and known the people who made this part of my life so great.

I would like to thank Dr. David Evans and my committee members, Dr. Gary Ervin and Dr. Randy Rousseau. Your help and guidance enabled me to retain some sense of sanity through the soil issues and statistical nightmares of convergence. I can honestly say without you, this would not have been possible. Thank you for your willingness to answer questions and allowing me the freedom to try to figure it out first on my own.

I would also like to thank my family. Thank you, Daryl Hobbs, my dad for helping me with encouragement, with realistic words to keep me grounded and money when I needed it. To my sisters, Shannon, Vicky, Krystalyne and their families, Thank you for providing me with numerous homes to crash at when I needed a break away from school and strengthening me and encouraging during this time. Thank you for letting me vent and making me laugh when I so readily needed one. Thank you to my nephews and nieces who have made life so much brighter for me. Thank you, Kevin for being goofy and making me laugh when I was stressed and for being there when I needed to talk. Thank you all for believing in me.

I would like to also thank my fellow graduate students in my lab and surrounding offices. David Wilkinson and Michael Crosby for their antics that made the time in the office so enjoyable and for answering all my questions. Mark Garrigues for tolerating my incisive talking when I could not figure something out and for helping me figure it out no matter what it was whether it was statistics or Canadian Mounties. To Mohan Tiruveedhula and Rida Sherif for allowing me become familiar with their cultures.

I would like to thank Xiana Santos and Johnny Rance Tomlinson, two great friends that made me better for having known them and who strengthened me and helped me more than I can ever say.

Lastly, and not any least important to my friends that God blessed me with before I came to Starkville and since coming to Starkville, who have provided me we numerous words of encouragement and many a fun night of laughter, relaxation and games. Thank you all so much for enriching my life and making my time here at Mississippi State so memorable. God truly has blessed me without measure.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER I

INTRODUCTION

Heightened interest in the changing availability and distribution of natural resources has fueled interest and debate ranging from oil resources, climate change, deforestation, carbon build-up, and forest utilization. This concern is widespread, as a general call to become "green" in our daily lives is being sent up by governments and manufacturers alike.

According to the Forest Resources of the United States, 2002, published by the U.S Forest Service, it is estimated that 749 million acres of land in the United States are forests, with 368 billion cubic feet in hardwood (Smith et al., 2004). The south-central portion of the United States composed of Alabama, Arkansas, Kentucky, Louisiana, Mississippi, Oklahoma, Tennessee, and Texas is responsible for 88,703 million cubic feet of the hardwood growing stock in the United States. Of this growing stock, 40,306 million cubic feet is oak, Quercus spp. (Smith et al., 2004). Oaks constitute 48 percent of all hardwood lumber produced (Cassens, 2007).

There are multiple species in the Quercus genus, which is usually broken into two sections, red oaks and white oaks. While there are several reasons for this delineation among oaks, the main difference is due to the structure of pores within the wood. White oaks have characteristic pores clogged with membranous growths called tyloses. Red oaks have open pores thus restricting their use to products such as veneer, pulp and paper, railroad ties, pallets and lumber, whereas white oaks can also be used for wet or "tight" cooperage (e.g. whiskey barrels) due to their closed pores which makes the wood impervious to liquids.

The use of the tree as lumber is dictated by three things: the quality of lumber yielded, and the purpose for that lumber and the range of the tree. Lumber quality is partly contingent upon the site conditions and by the intensity of competition that the tree experiences while growing. Competition on a site dictates the number of limbs by influencing the epicormic branching of a tree as well as growth rate. Aside from competition, optimal growth is attained when the species is correctly matched to the environmental characteristics of the site. Site characteristics, aside from genetics and human influences like management techniques, play a part in the growth of the tree, the number of limbs, and the color of the wood. All of these characteristics play an important role in the value of the tree. Specific site characteristics include soil texture, moisture availability, possible competition, slope, aspect, and elevation. Baker and Broadfoot (1977) recognized this and made a field guide of site evaluation technique that examines site characteristics to estimate tree growth and site suitability at a site-specific level.

Range influences the utilization of the tree through to the economical practicality of using a wood source close to the mills. Oak species are rarely hauled out of their current range due to the economical cost of hauling to the mills. This cost is usually only overlooked if the tree species has unique characteristics that are key in a specific industry that is unable to meet this need with another species in the immediate area. Each species of oak has a specific range based upon specific site characteristics and climatic tolerance. This range is dictated by factors such as day length during the growing season, minimum temperature, and soil characteristics. The U.S Department of Agriculture recognized these characteristics of the forests and decided it was important for the range of each tree species to be determined (Little, 1971). Similarly, Baker and Broadfoot (1979) also reported later on the sites and ranges of trees.

The ranges of U.S. tree species were first mapped as early as 1905 (Little 1971). Often, the ranges of those species were based upon old forest cover reports made as far back as the 19th century. The maps and information were reviewed and updated in 1951 by Elbert Little (Little 1971). Botanical lists and forest surveys, along with other unpublished field notes and herbarium specimens, were used to compile the range maps for most of the economically important trees including oaks. The range maps by Little show the generalized distribution of the trees based primarily on observations, rather than an analysis of site characteristics.

Today, Little's (1971) maps are being paired with other scientific information of micro-climatic characteristics to depict the range of many tree species found in the literature, e.g. Harlow and Harrar's Textbook of Dendrology (1969). It is imperative for foresters, silviculturists, landowners and land mangers to be aware of what specific species are found or can be grown in their area for both forest management and possible marketable aspects. Authors of current textbooks on dendrology and silviculture are including these adaptations of Little's (Little, 1971) maps in their textbooks to depict the range of tree species and to show the interactions between species and their respective ecosystems.

Unfortunately, there are problems resulting from using Little's (1971) range maps to show distribution. One is that the information is not current, representing findings that can be up to a hundred years old. When these maps were initially developed, it is unknown if there were clear establishment of rules for data collection or validation of species identification, which could be a problem since much of the information was based upon field observations. Little (1971), incorporated his observations with those of others to estimate the ranges. According to the Atlas of United States Trees: Volume 1. conifers and important hardwoods, “these maps do not show where a species grows outside the natural range after having been introduced directly or indirectly by man, whether planted, escaped, adventives, or naturalized” (Little, 1971). This resulted in incompleteness of species range maps, leading to misunderstandings by foresters and landowners alike.

It has become somewhat common for dendrologists to find some tree species thriving and reproducing in abundance in areas that, according to Little’s (1971) range maps, should not have these species present. An example of this is pin oak (Quercus palustris Muenchh) Most textbooks such as Harlow and Harrar’s Textbook of Dendrology (1969), show the current range of pin oak reaching only to south central Tennessee, but it has been documented through observation farther south in northern Mississippi.

In order to better manage and utilize our forests, we must first have an accurate and current understanding of the tree species present. This study represents the initial process of providing for the clarification to range maps of ten economically and ecologically important oaks in Mississippi. This was accomplished through spatial analysis of sites on which these trees are found and predictive modeling of these oak species distributions in Mississippi.

Literature Review

A number of recent studies have correlated site characteristics with current and historic range and distribution of tree species. Many of the studies, like Ohmann and

Gregory (2002), recognized the need to know the composition and forest structure at a regional spatial scale to determine natural resource policies. Ohmann and Gregory utilized Landsat imagery, climate, topography, geology and location to yield a gradient nearest neighbor model for predicting the spatial distribution of vegetation in coastal Oregon for ecological research. Shostak *et al.* (2004) used topographic position, but also paired it with related site index to determine if there were common variables in the relationships of various ecological factors that influenced oak distribution in the Cumberland Mountain-Plateau region of Alabama. Other studies, such as Iverson and Prasad (2001), Lister *et al.* (2000), and Iverson and Prasad (1998), looked at FIA (Forest Inventory Analysis) data to determine tree species distribution. Iverson and Prasad (2001) used FIA to calculate importance values of trees in the eastern United States based upon their relative density and relative basal area in both the overstory and understory to yield maps of distribution, abundance and relative frequency of oaks in the western and eastern United States. Iverson and Prasad (2001) then utilized this information for the Hadley and Canadian Climate Center, to find potential future suitable habitats using scenarios via the empirical model DISTRIB and the SHIFT model, a tree migration model. DISTRIB uses a tree regression analysis approach combined with SHIFT model to incorporate historical migration rates and fragmentation of habitats.

Iverson and Prasad (1998) utilized the FIA and classification and regression tree (CART) analysis previously to map tree species based upon climate change. They found that there were some limitations to this regression technique, since they were unable to capture environmental factors that influenced the trees on a finer scale of slope at the tree location. Nevertheless, Iverson and Prasad (1998) believed that their model could be used to analyze datasets with many possibly interacting variables, and that it is suitable for determining between regional and site-specific effects.

Lister *et al.* (2000), on the other hand, used FIA data and site characteristics such as elevation and slope to compare geostatistical procedures to improve predictions of species distribution. They found that their methods of using exploratory analysis and least squares regression led to improved predictions over those made using regression procedures alone, as well as those only utilizing kriging and sequential Gaussian conditional simulation (SGCS).

The finding posed in Listers *et al.* (2000) study may explain some of the confusion that Shostak *et al.* (2004) saw in their study. Shostak *et al.* (2004) used a stepwise logistic regression procedure to look at a data set of variables thought to influence the outcome of the occurrence of oak. They also ran analysis on the maximum likelihood estimates for individual oak stem success for topographic position and predicted oaks to be 80% less successful on upper slopes than on lower slopes. This seemed contradictory to the whole oak composition analysis that found that the greatest numbers of oaks were found on upper slope positions than that of the middle and much more than the lower. Shostak et al. (2004) believed the cause of this was that the upper slope positions in their study at pre-harvest had a larger number of stems. They believed the density at this topographic position caused a decrease of not only oaks but also all trees.

Other studies have looked at climate as a factor influencing the distribution and range of tree species both now and in the past. Felicisimo *et al.* (2002), who came up with a model for the potential distribution of forests in Spain, looked at climate as one of the variables influencing the current distribution of trees. Felicisimo *et al.* (2002) utilized a logistic model based upon the presence/absence of the tree species to construct maps in which they were able to identify the places of greatest probability of species occurrence. Their model used a geographic information system (GIS) and data from vegetation maps, climatic, and lithological information.

Hall *et al.* (2002), used early land surveys, 19th century maps of forest cover, and contemporary agricultural censuses, to look at land use and land cover for 300 years in Massachusetts. They found that forest compositions correlated strongly with environmental conditions, especially variation in climate. Hall *et al.* (2002) used GIS to compare changes in climate, geology and land-use to that of both the historical and modern forest composition.

Historical data also were used by both He *et al.* (2006) and Abrams (2003) to show historical components of the forests and the changes that have occurred. He *et al.* (2006) used survey records of the Midwestern United States and hierarchical Bayesian model to look at historical forest composition and tree species distribution. The Bayesian approach combined species and environmental relationships and explicit spatial dependence to map data. The Bayesian model was used to identify the seven most significant covariates for terrain and soil. The seven chosen were based upon their influence on the basic needs of the tree. The seven were "elevation, slope, aspect, soil water capacity, soil organic matter, soil depth and depth of bedrock."

He *et al.* (2006) narrowed their study to three different classification groups. One was black oak, as their individual species, because of its abundance in Missouri. For the genus group, they looked at the "bottomland oaks (*Quercus* Spp.) that included primarily pin oak, white oak and red oak." Their last group was a functional group that incorporated hickories and oaks. He *et al.* (2006) found that bottomland oaks might have occurred where there were wide floodplains with islands of high elevation. They were also able to determine that black oak would have been found on most upland sites examined in this study due to intolerance of flooding.

Lastly, Abrams (2003) looked at and analyzed historical writings such as early land surveys and witness trees to determine the forest makeup before European settlement in the eastern United States. Abrams (2003) found that areas in the United States, currently dominated by red oak (*Quercus rubra*) and chestnut oak (*Quercus prinus)*, once were predominately white oak (*Quercus alba)*. He found that white oak is a more versatile oak than any other eastern oak species. It is unspecialized in its growing range, causing it to lose out in competing for nutrients with oaks that are more site-specific, such as red oak, which grows on rocky ridges and chestnut oak that is "more xerophytic, fire resistant, and tolerant of nutrient-poor soils". The previously mentioned site factors along with human interventions have made the forest composition as we see it today.

All of the studies agreed that oak composition of a site is influenced by the climatic, topographic, and soil characteristics of a site. The researchers of these studies considered different aspects of site characteristics and evaluated them in various ways but all agreed that knowing the interaction of a tree species with site conditions would aid in determining distribution of trees. It is the aim of this study to follow in the ideas of previous investigators to derive predicted distributions of selected oak species. .

Objectives

The first objective of this study was to determine the absence\presence and abundance of ten *Quercus* species in Mississippi that are utilized in the hardwood industry and by wildlife.

- *Q. alba* L. (white oak)
- *Q. falcata* Michx. (southern red oak)
- *Q. marilandica* Muenchh. (blackjack oak)
- *Q. michauxii* Nutt. (swamp chestnut oak)
- *Q. nuttallii* Palmer (Nuttall oak)
- *Q. pagoda* Raf. (cherrybark oak)
- *Q. phellos\nigra* L. (water/willow oak)
- *Q. shumardii* Buckl. (Shumard oak)
- *Q. stellata* Wangenh. (post oak)
- *Q. velutina* Lam (black oak).

This information was used to make maps showing the predicted distribution of the chosen oak species in Mississippi that are more up-to-date and site specific than Little's range maps. This will help foresters and landowners identify and manage the trees on their lands in Mississippi more successfully.

The second objective of this study was to use the abundance information to determine the physical characteristics of the sites that these oaks occupy. The physical characteristics of the sites was determined by looking at precipitation, temperature, flow accumulation, elevation, slope position, aspect, and soil texture and drainage class of each plot. This information was used to make a species suitability model derived from the parameters of stepwise logistic regression. The logistic parameters were used to model site preferences across the state to determine the potential distribution of each species in Mississippi. The outputs were maps showing probable distribution of species across Mississippi. Those maps can potentially serve as management tools to be used by landowners and foresters to determine if a tree will grow on a specific site if planted there, so more economical and ecological pairing of oak species to a site can occur.

CHAPTER II

METHODS

Data for this study came from publically available databases. All secondary data had to be downloaded from- various sites and adapted to fit the specific nature of this study. The tree species information came from Mississippi Institute for Forest Inventory (MIFI) database. The data used to determine the physical characteristics of each site were gathered from several sources. Aspect, slope position, elevation, along with flow accumulation were derived from a 10-meter Digital Elevation Model (DEM) that was acquired through contact with the Mississippi Automated Resource Information System (MARIS) Technical Center. The Natural Resources Conservation Service (NRCS's) Soil Survey Geographic (SSURGO) Database website yielded the data needed to determine the soil characteristics of the sites. Lastly, the precipitation and temperature data used were derived from the Parameter-elevation Regression on Independent Slopes Model (PRISM) data put out by Oregon State University (Daly, et al., 2002b).

Study Area

The study area for this project encompasses the whole state of Mississippi excluding three counties (Figure 1). Only 79 counties were used out of the 82 total counties of Mississippi due to inconsistencies or unavailability in the SSURGO soil data. At the time of this study, the soil data for Scott and Greene counties had yet to be made available to the public in digital format. Wilkinson county data were not used because of a lack of consistency in the delineation of the horizon layers with those of the other counties due to a differing of nomenclature used to describe soil types. This made matching horizon information to that for other counties impossible. Therefore, all analysis of plots in these areas was dropped from further consideration.

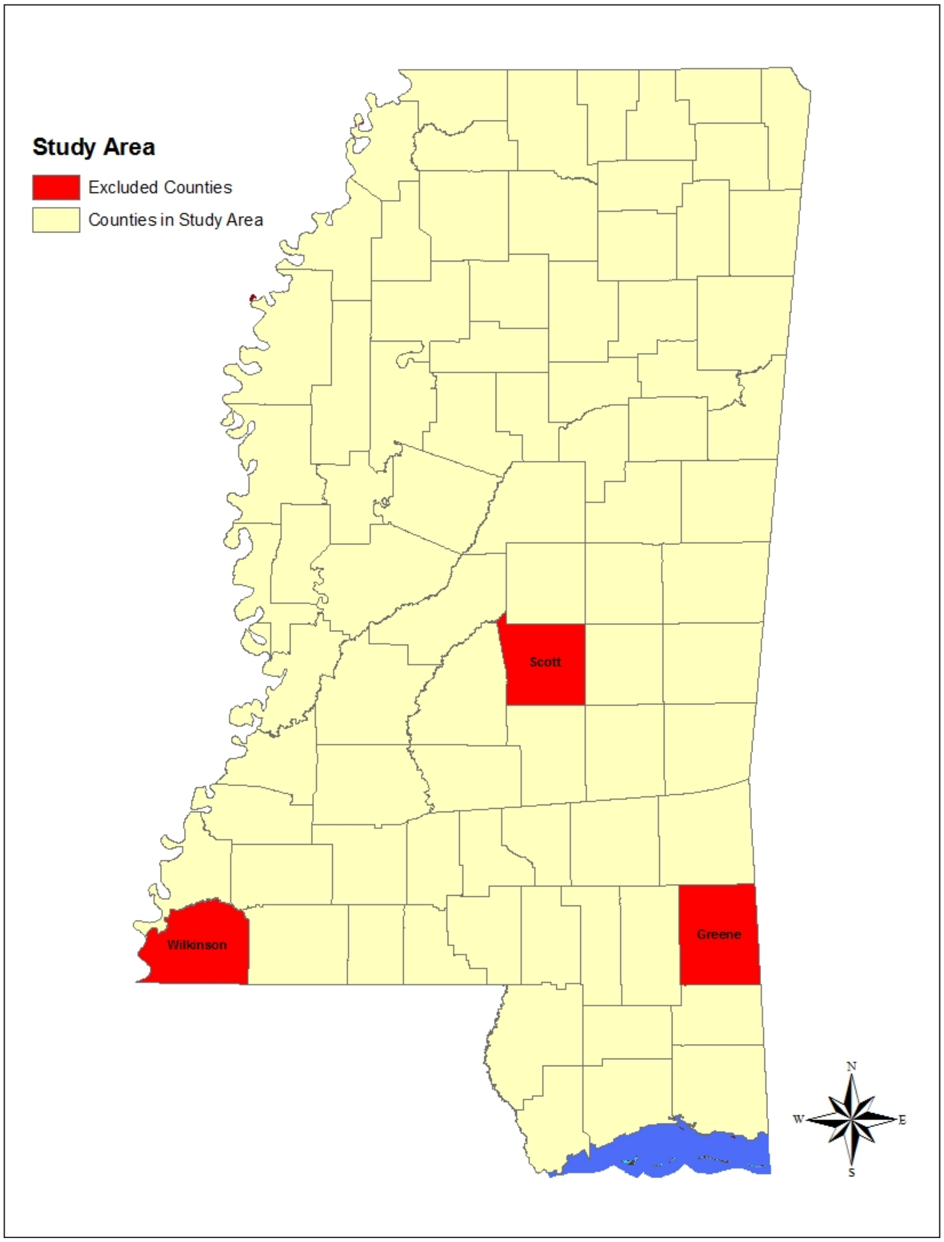

Figure 1 Study area map showing the counties in Mississippi used in this study. Greene, Scott, and Wilkinson shown in red were excluded from the study.

Tree Species Data

MIFI plot information was gathered from 2004 to 2009. MIFI divided the state into five survey regions (Southwest, Southeast, Central, North, and the Delta) and collected forest inventory data on one region per year. MIFI collected over one hundred plots per county. The goal of MIFI is to "develop and implement a continuous, statewide forest resource inventory necessary for the sustainable forest based economy" (Tucker, 2010).

The information generated by MIFI used to determine tree density and distribution for this study came from both the plot- and tree-level tables. Presence and abundance data for the ten target species were extracted from the MIFI database and compiled into data tables. These data include the latitude and longitude of the plot centers, the plot number and a tally of all tree species found on the plots. The latitudes and longitudes were entered into ESRI® ArcMAP™ 9.3.1 to yield maps in point shapefile format showing the currently recorded Mississippi range and distribution of each economically or ecologically important oak. The other data were incorporated into the attribute tables of the plot data in binary form as a 1 or 0 for the presence or absence of each particular species on that plot. Then from each set of plot data, a validation set of 20 % of the total plots that showed occurrence of each species of oak were taken to check the accuracy of the range maps. Table 1 shows the total plots for each species and the number of plots showing occurrence of the species. The number of total plots considered in the stepwise logistic regression differed from the total number of plots collected by the MIFI field crews, due to removal of the 20% validation set of occurrence for each species. This was done prior to running SAS on the species and with each validation set being randomly selected, some plots that would have eventually be eliminated due to inhabiting sites in which environmental data were not available when the regression was run could have be selected out. The validation sets and then the elimination of plots due to missing site characteristic data caused the total number of plots on which the regression was run to vary for each species. The validation set was held out to check the validity of the models once the maps were yielded.

Table 1 The total number of MIFI plots used in the logistic regression and the validation set with the number of presence and absence by oak species used to derive the logistic parameter for the models

Oak Species	20% validation set	Number of Plots with Presence(1)	Absence(0)	Total plots
black oak	14	49	5731	5780
blackjack	66	247	5481	5728
cherrybark	129	381	5313	5694
Nuttall	30	106	5662	5768
post	129	445	5226	5671
Shumard	18	62	5719	5781
southern red	272	898	4642	5540
swamp chestnut	27	104	5658	5762
water/willow	651	2002	3293	5295
white	270	873	4687	5560

In order to run the species suitability analysis, the physical characteristics of each site were determined. Environmental factors used in these analyses were topographic features (aspect, slope, and elevation), temperature, precipitation and the soil texture and drainage class of the site along with percent clay and total representative horizon depths.

Elevation and Derived Data

For this study, a DEM of the whole state that had already been mosaicked together was requested from the MARIS Technical Center. The DEM is at a 10-meter horizontal resolution and in the Mississippi Transverse Mercator (MSTM) projection. The 10-meter DEM of Mississippi was derived from U.S. Geological Survey (USGS) Digital Line Graph (DLG) hypsographic and hydrographic source data. The DEM was imported into ArcMap Version 9.3.1. In the ArcMAP ArcToolbox, Spatial Analysis tools for hydrology and surface were utilized to calculate (or estimate) aspect, percent slope and flow accumulation of each pixel from the Mississippi DEM. The process to compute the aspect was to input the DEM into the Aspect tool under the Surface tool in the Arc Toolbox. The output was a raster showing the azimuth direction that each pixel in the Mississippi DEM faces. ArcMAP classifies the ranges of azimuth direction into narrow ranges of nine classes (flat, North, Northeast, East, Southeast, South, Southwest, West, and Northwest) (Appendix-Table 15). These categories were further reclassified into four variables of North, South, East, and West to generalize variation in the data and aid in the convergence of the logistic regression. Flat areas were eliminated from the study since visual inspection when compared to the elevation raster indicated these are typically water.

The process for determining the percent of slope was equally as simple, except a conversion factor of 0.3048 was applied to the "z" units of the DEM in the slope calculator toolbox tool, since the "x"(easting) and "y" (northing) units were in meters while the "z" (elevation above mean sea level) units were in feet. This allows the z units to be converted into meters since 1 foot equals 0.3048 meters. The ending result was a raster file showing the percent slope of each pixel from the Mississippi statewide DEM.

In order to determine the potential wetness of the pixels in the 10 meter DEM, the flow accumulation was calculated. The flow accumulation of each pixel was determined using the hydrology tools under the Spatial Analyst Tools in the ArcMap Toolbox. The raster was first run through the fill tool in which all the small imperfections in the data were filled to resemble the elevation of the surrounding pixels. The filled raster was then used to calculate the flow direction in which the movement of water from a cell to the steepest adjacent cell is shown. The flow direction raster was used as the input raster to determine the flow accumulation of each cell by calculating the contribution of the flow into each cell from surrounding up-slope cells. To aid in analysis in SAS, the log-plus-one transformation was applied to the flow accumulation raster to scale the range of values to those of the ranges of the other variables to aid in convergence. Figures 2-5 show maps of the variables derived from the MARIS DEM.

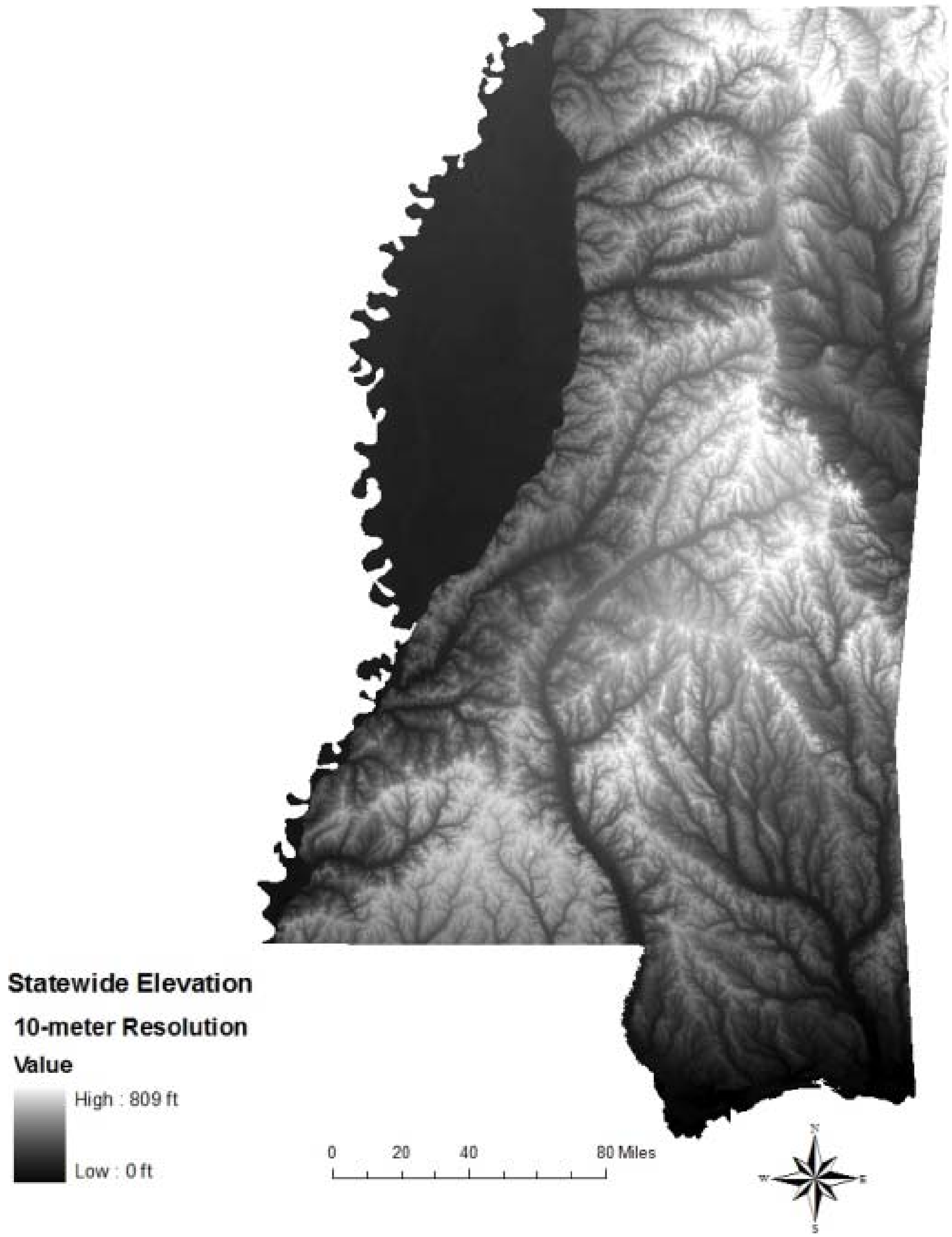

Figure 2 Digital Elevation Model (DEM) of Mississippi based upon horizontal samples of elevation at 10-meter intervals derived from U.S. Geological Survey (USGS) Digital Line Graph (DLG) hypsographic and hydrographic source data

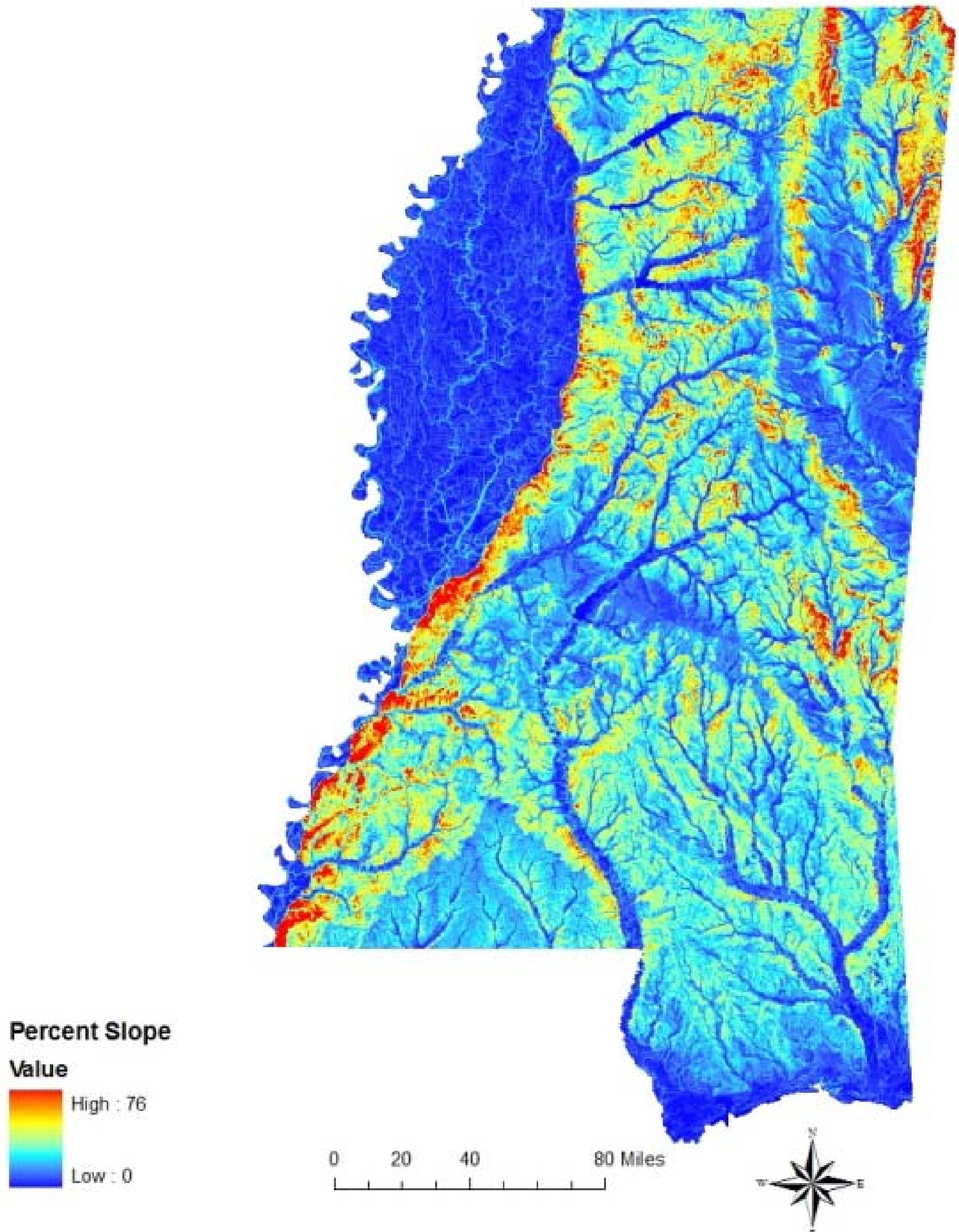

Figure 3 Percent slope for Mississippi derived from 10-meter Digital Elevation Model (DEM).

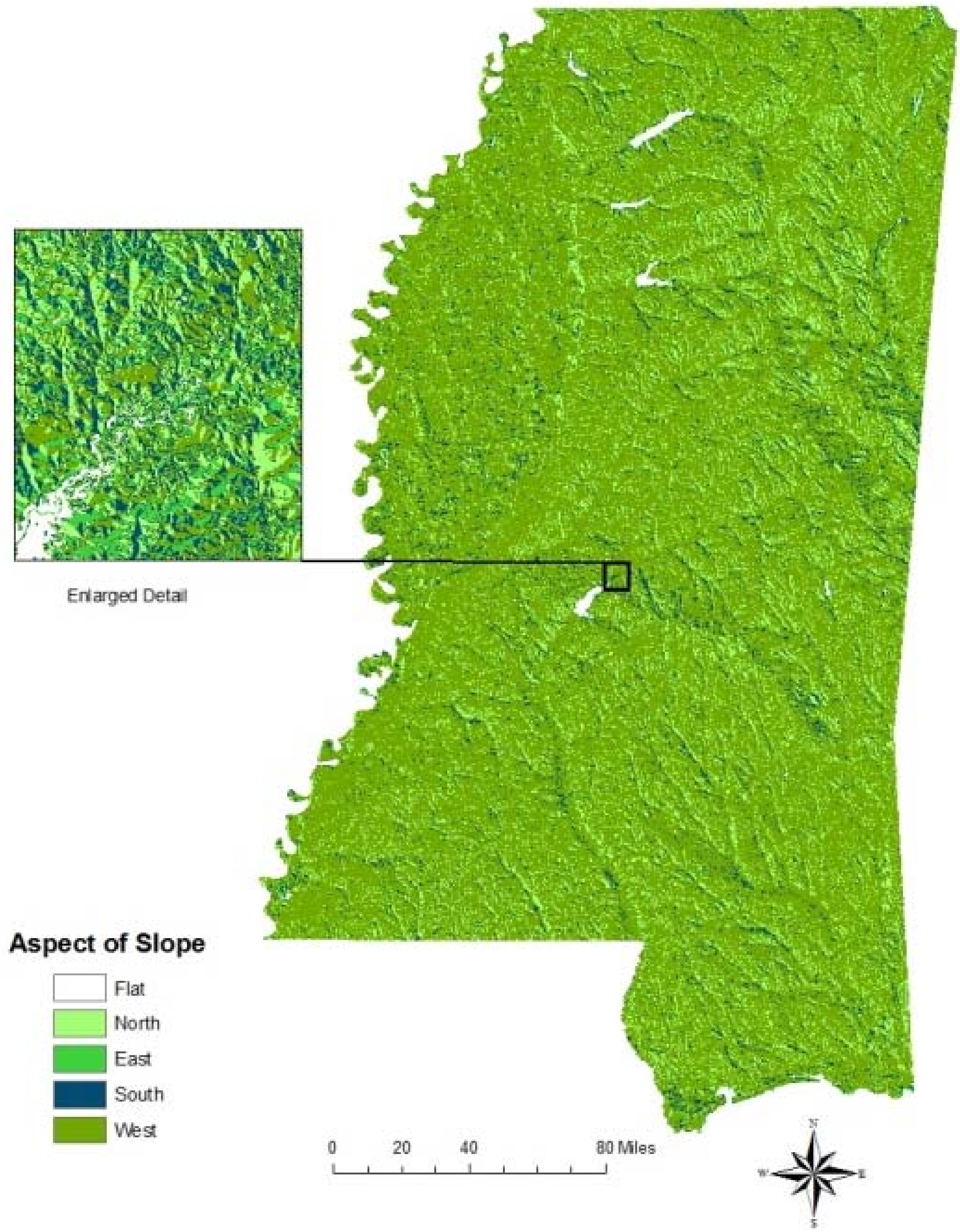

Figure 4 Aspect for Mississippi derived from 10-meter Digital Elevation Model (DEM).

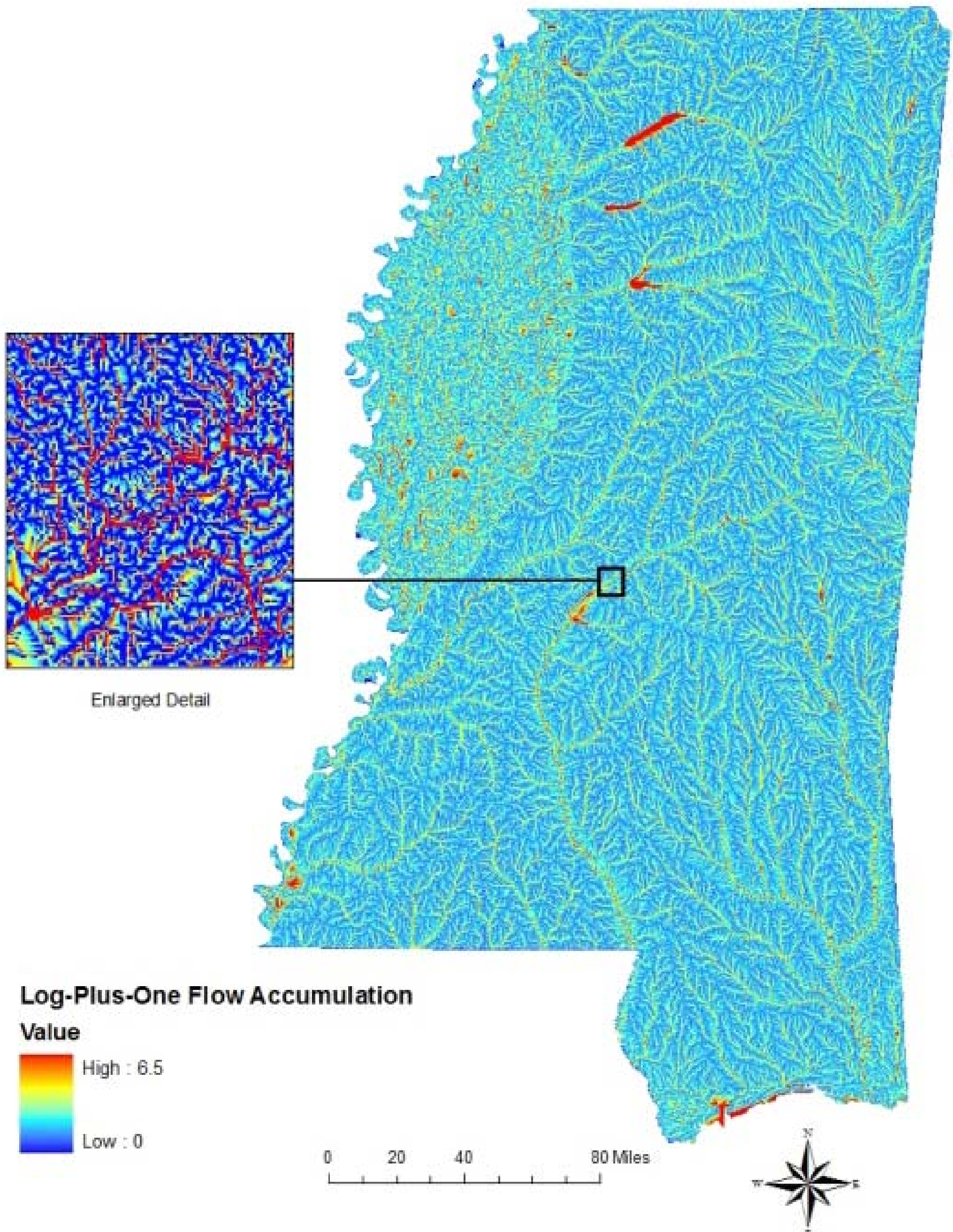

Figure 5 Flow Accumulation for Mississippi derived from 10-meter Digital Elevation Model (DEM).

Climatic Features

PRISM Data (Parameter-elevation Regression on Independent Slopes Model) is a climate analysis system that generates gridded estimates of annual, monthly and eventbased climatic parameters using point data, DEM and other available spatial datasets (Daly *et al.,* 1997). The stated accuracy of the original DEMs were 130 m circular error with 90% probability (Daly *et al.*, 2002). The maximum temperature data for the conterminous United States were downloaded for 1958-2008 and were combined using the model maker in ERDAS Imagine to yield one raster showing the average maximum 50-year temperature for the conterminous U.S. It was then resampled to 10-meter resolution. The raster was then clipped to the boundary of Mississippi and converted from degrees Celsius to Fahrenheit for ease of interpretation. The same technique was then applied to maximum annual precipitation, which after being combined, was converted from millimeters to inches for ease of interpretation of results. The outputs for the maximum annual temperature and precipitation for the state can be seen in Figure 6 and 7.

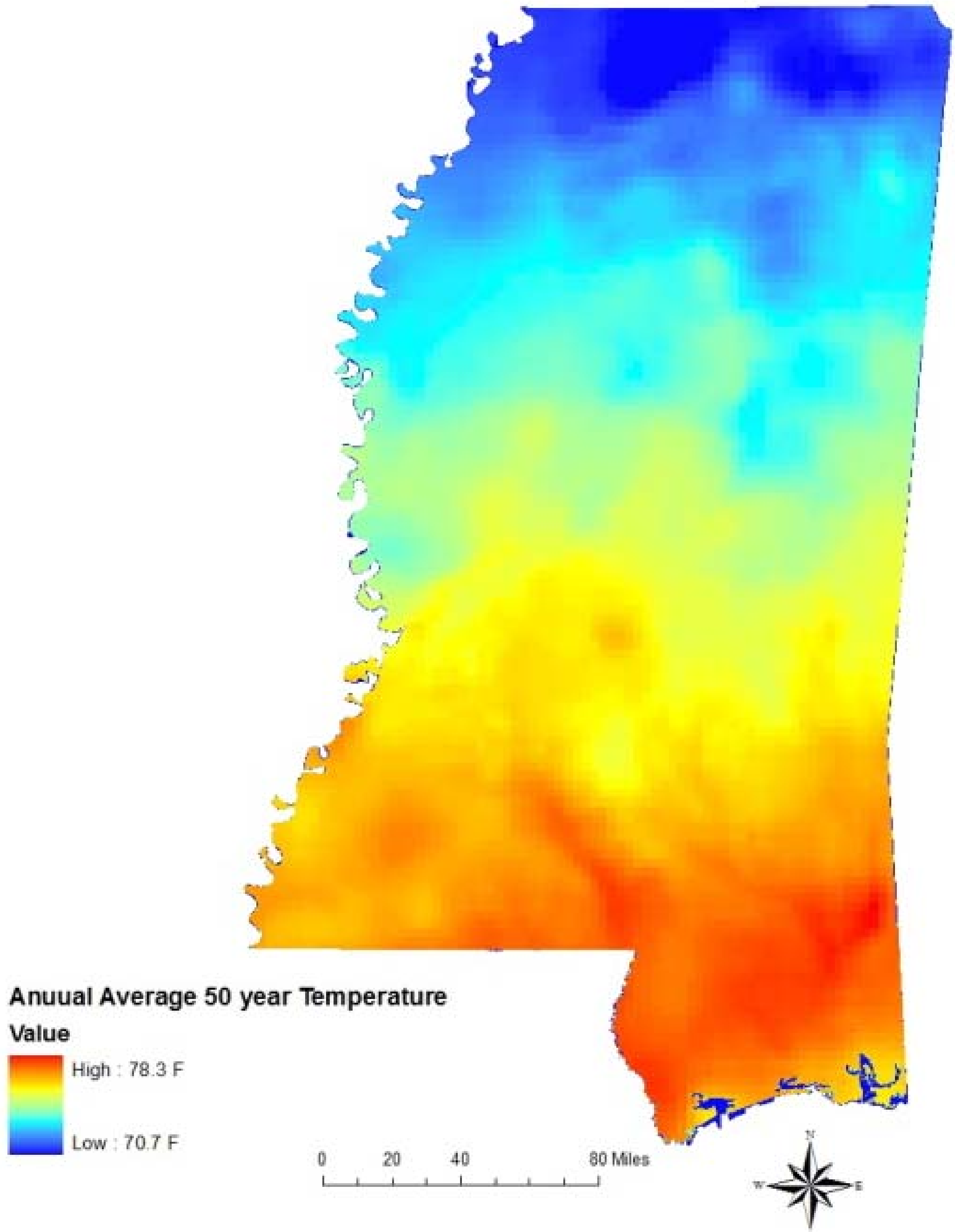

Figure 6 Annual average temperature for Mississippi derived from 1958-2008 annual temperature PRISM data

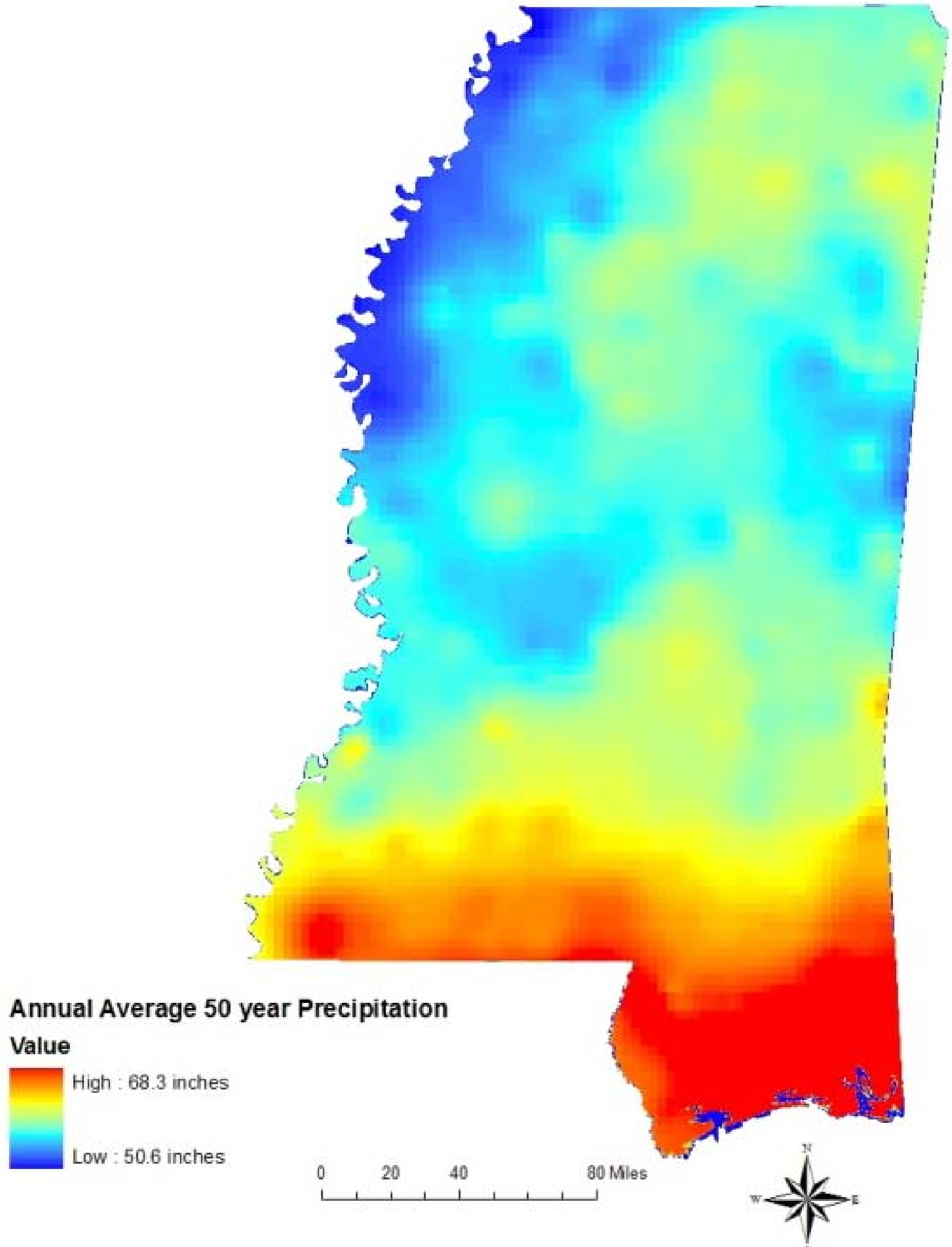

Figure 7 Annual average precipitation for Mississippi derived from 1958-2008 annual precipitation PRISM data.

Soils Data

The soils data for this study were retrieved from the SSURGO database. They were downloaded by county for Mississippi. The scales of these maps generally range from 1:12,000 to 1:63,360. The SSURGO mapunits consist of one to three components each. The attributes table of the mapunits tells the proportional extent of the component soils and gives the properties of the mapunits, which are tied to the National Soil Information System (NASIS) attribute database (NRCS, 2009).

At the time of this study, Scott and Greene counties had yet to be put into digital format, so these counties were eliminated from the study. Wilkinson County was also eliminated from the study do to a change in horizon classification that had taken place when this soil information was collected. The change in the nomenclature of the soil horizons in Wilkinson county was a result in a change in policy coming into effect during the year that these data were being processed; this changed made it incompatible with the other county layers. The downloaded county soil databases were input into Microsoft Access using the template provide by SSURGO and tied to the spatial representation of the map unit following the technique laid out in *Downloading SSURGO Soil Data from Internet* by Merwade (2008). Once the tables had been compiled into the template, the tables were sorted for information concerning horizon depths, textures, drainage classes, and percent clay. The tables that contained this information and their linking fields are shown in Figure 8. Tables- chorizon, chtexturegrp and chtexture contain information on a horizon level such as the percent clay, texture and depth (NRCS, 2009). Component table delineates the soil overall make up and tells the drainage class. Tables mapunit and muaggatt contain the soil units names, size and the unique key field, mukey, needed to tie all tabular data to the polygonal spatial representation of the soils.

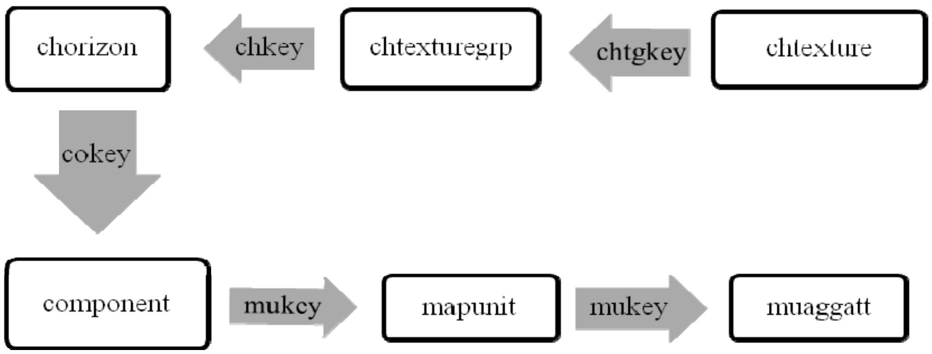

Figure 8 Natural Resources Conservation Service (NRCS)'s Soil Survey Geographic (SSURGO) database tables with soil variables used and the key fields in the tables. The key fields in the arrows are names of columns that have unique identifiers used to link the tables to one another (NRCS, 2009).

The tables that had information important in this study were then input into JMP and joined based upon their common key fields. Before the tables could be joined, the entries had to be reordered and some fields deleted to ensure that each soil map unit would have only one unique identifier. One field that had to be sorted and every "no" entry deleted was the RV indicator field, which tells, "if a soil structure is representative for the horizon" (NRCS, 2009). The second column in which entries were deleted was the majcompflag column that "indicates whether or not a component is a major component in the map unit" (NRCS, 2009). All entries of "no" were deleted and not taken into consideration for this study. Nonessential columns, those that did not hold data of concern for this study, were then deleted from the selected tables. The chorizon table was then reconstructed to segregate the information of concern by horizon layers, which SSURGO designated as H1, H2, H3, H4, H5, Cr, R, Oa (NRCS, 2009). The consolidated tables were joined into one table showing total representative horizon depth, texture of each horizon, average percent clay of each horizon and the overall drainage class for the soil type. The information was tied spatially in ArcMap to the soil mapunit polygons. This process was repeated for each of the 79 counties of the study and then merged into five groups similar to the MIFI regions. The ranges in texture classes were so significantly different between counties that a generalization of reclassification had to be applied to narrow overall variation. The soil texture classes were reclassified to fit in to four classes: sandy, silty, clay, and rock. The drainage classes were reclassed into seven classes also for this reason. (See Appendix Tables 13 and 14 for classifications) Figures 9-12 show the output for the texture classes for the three main horizons and the drainage classes of the soil map units.

Figure 9 Generalized soil texture of H1 for map units across Mississippi derived from the Natural Resources Conservation Service (NRCS)'s Soil Survey Geographic (SSURGO) database.

Figure 10 Generalized soil texture of H2 for map units across Mississippi derived from the Natural Resources Conservation Service (NRCS)'s Soil Survey Geographic (SSURGO) database

Figure 11 Generalized soil texture of H3 for map units across Mississippi derived from the Natural Resources Conservation Service (NRCS)'s Soil Survey Geographic (SSURGO) database.

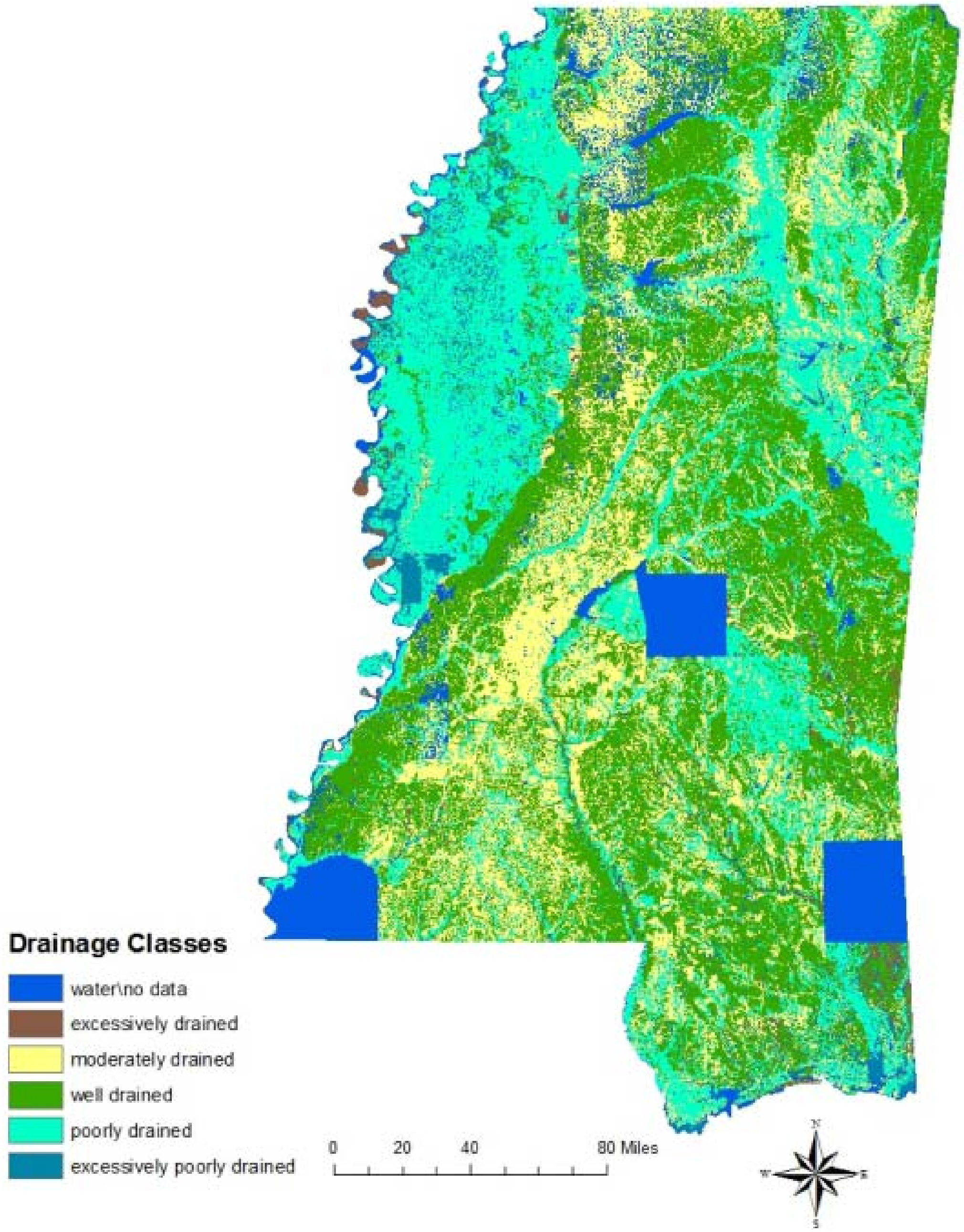

Figure 12 Generalized soil drainage class for map units across Mississippi derived from the Natural Resources Conservation Service (NRCS)'s Soil Survey Geographic (SSURGO) database

Statistical Analysis

This study used stepwise logistic regression and GIS, similarly to the approach used by Felicisimo *et al.* (2002) and Shostak *et al.* (2004). All the variables were joined based upon spatial location to the soil polygons in ArcMap. Tree plot information was then geospatially combined with the soils and other variable's datasets. Using Hawth's Analysis Tools in ArcMap, the polygon information of the soils were extracted to the attribute tables of the tree data. This resulted in a complete database of the oak composition of the plots along with the slope, aspect, flow accumulation, elevation, climate information and soil type information for each plot.

The data were input into SAS and analyzed using the stepwise logistic regression. SAS code was written to eliminate from consideration plots that had no data, missing data, or zeros in the fields to aid in model convergence. The variables for -aspect, textures for each of the three horizons, and drainage class were treated as categorical variables and dummy variables were assigned to each of the possible classes for each of these variables (Table 2). The soil texture and drainage class were dummy coded using sandy class and water class as the reference groups. An alpha of 0.05 was used to determine variable significance in the entry and exit of variables into the stepwise logistic regression. The predictability of the logistic regression estimates was evaluated by the percent concordant and percent discordant values in SAS. The percent concordant is the percent of comparisons between species presence to species absence, in which the predicted probability of species presence is higher than species absence.

Table 2 Categorical variables of aspect, drainage, and texture classes and the dummy variable classifications with the dichotomous variable design assigned to represent the variables in the stepwise logistic regression assigned by SAS

Variable	Category	Value	Design				
aspect	North	1	*0	0	0		
	East	2	1	0	0		
	South	3	0	1	0		
	West	4	0	0	1		
drainage class	water\other	0	*0	0	0	0	0
	Excessively drained	1	1	0	0	0	0
	moderately drained	2	0	1	0	0	0
	well drained	3	0	0	1	0	0
	poorly drained	4	0	0	0	1	0
	excessively poor drained	5	0	0	0	0	1
H1 texture class	sandy	1	*0	0			
	silty	2	1	0			
	clayey	3	0	1			
H2 texture class	sandy	1	*0	0			
	silty	2	1	0			
	clayey	3	0	1			
H3 texture class	sandy	1	*0	0			
	silty	2	1	0			
	clayey	3	0	1			

* represents the reference group for the dummy variables

The parameter estimates of variables that proved significant through the stepwise logistic regression in SAS were then input into the logistic regression equation (1) using a model built in ERDAS Imagine.

$$P(Y=1) = \frac{e^{\alpha+\beta_1 X_1+\beta_2 X_2+\cdots+\beta_k X_k}}{1+e^{\alpha+\beta_1 X_1+\beta_2 X_2+\cdots+\beta_k X_k}} \quad (1)$$

where:

P(Y=1) = probability of oak occurrence on site e = exponentiation

α = intercept from SAS output

β = parameter estimate

X = significant variable

This yielded eight rasters showing the probability of occurrence for each of the eight oaks. Once the probabilities were calculated, the rasters were input back into ArcMap and the validation sets of the oak presence point data were projected onto them to check the accuracy of the probabilities.

According to King and Zeng (2001), when binary dependent variables (e.g., presence and absence data) are unequally represented, as is the case with some of these oak species, logistic regression can underestimate the probability of occurrence. The presence data are typically underestimated while the larger pool of absence data is overestimated. In order to address the underestimation resulting from logistic regression on these data, 1000 test regression iterations were executed and small random samples were selected out of the full dataset with each sample composed of all presences and an equal amount of absences. SAS code was written to randomly run 1000 iterations of these equal samples and conduct stepwise logistic regression on the smaller datasets so that it could be determined if the same variables were significant without the overestimation of absences. The results of the iterations for each species can be seen in the Appendix in Table 16.

CHAPTER III

RESULTS

Soils Database

The initial statistical analysis of the independent variables proved to be unsuccessful due to quasi-complete and complete separation issues occurring within the data configuration. These issues stemmed from too many independent variables and too much variation within the variables from the SSURGO database. With the validity of the model in question due to these issues, the questionable variables were investigated and were either eliminated or reclassified. It was determined that since only horizon levels, H1, H2, and H3, based upon change in color, were consistently delineated similarly across the counties, the independent variables found in these layers would be the only variables taken into consideration for this study. Also percent clay and texture class appeared to be correlated. Therefore, the variable percent clay was dropped to aid in the resolution of the quasi-complete separation issue, since percent clay failed to consider differences in sand and silt fractions of the soil. After making these changes, texture for the three horizons and drainage class were the only variables used from the SSURGO database. In order to achieve a workable reliable model, stepwise logistic regression was executed for each species. Convergence and a reliable model were obtained for all species of concern except for black oak and swamp chestnut oak. These oaks, therefore, were removed from further consideration in the study.

Statistical Models

The predictive ability of the full models was assessed by looking at the percent concordant and percent discordant numbers (Table 3) as well as the c value which according to Kutner, et al. (2004) are useful measures of a model's predictive power. These numbers were obtained from the Association of Predicted Probabilities and Observed Responses of the SAS output. The c value indicates the area under the receiver operating characteristic (ROC) curve, which provides gives the concordance index for the models (Kutner, et al., 2004). The predictability of each species of oaks' stepwise logistic regression models showed satisfactory predictability with all percent concordant and c being 0.64 or greater. The model with the highest predictability was Nuttall oak with a concordance of 85.50 and the poorest predictability was with the post oak model with a concordance of 64.00.

Table 3 Statistical comparison of predictability and fit for the stepwise logistic regression model parameters. The percent concordant is the percent of comparisons between species presence to species absence, in which the predicted probability of species presence is higher than species absence.

Oak species	Concordant (%)	Discordant (%)	Tied (%)	Pairs	c
Shumard oak	68.60	24.50	6.90	354578	0.720
post oak	64.00	34.60	1.40	2325570	0.647
white oak	70.90	28.50	0.50	4091751	0.712
southern red oak	67.60	31.80	0.60	4168516	0.679
blackjack oak	84.00	15.00	1.10	1353807	0.845
water/willow oak	66.60	33.00	0.40	6592586	0.668
Nuttall oak	85.50	10.50	3.90	600172	0.875
cherrybark oak	72.10	26.80	1.10	2024253	0.727

The variables that proved to be significant for each of the oak species varied so much that each species is best discussed separately than as a group. The number of species in which the variable was significant in the occurrences of that species from the logistic regression is graphically shown in Figure 13. Tables 4 through Table 11 in the sections that follow show the logistic regression coefficient, the standard error, odds ratio and p-values for the significant variables of the final and best model for each of the oaks.

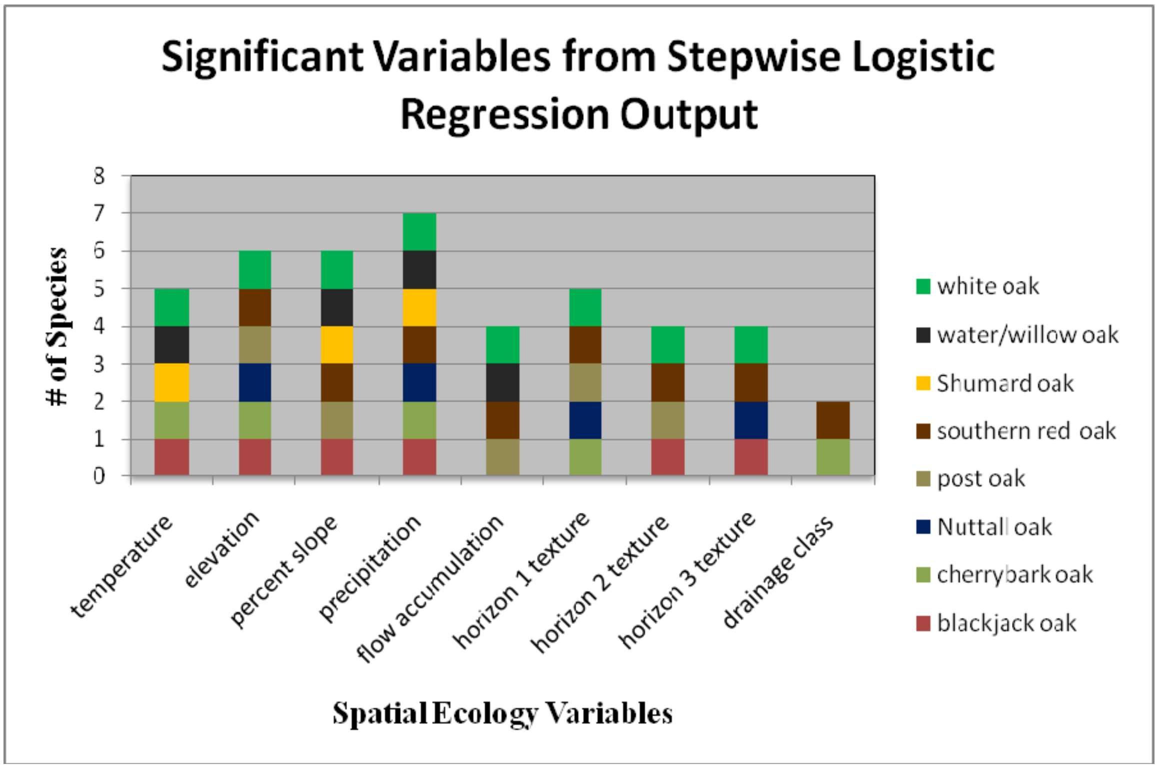

Figure 13 Summary of variables significant in predicting occurrence for each species derived from the stepwise logistic regression output.

White oak

Elevation, precipitation, temperature, percent slope, flow accumulation and texture class of the soils were all significant predictors for the occurrence of white oak ($P< 0.05$; Table 4, Figure 14) The test regression iterations revealed that only one variable, flow accumulation, was not significant over 50 % of the iterations (Appendix-Table 16). The odds ratio for flow accumulation indicates that when holding all other significant variables at a constant, an increase of flow accumulation by one unit will increase the likelihood of white oak occurring by 1.11 times. More clay and silt occurring in H2 increases the likelihood of white oak occurring and has the most significant influence with odds ratios of 1.9 and greater. Increases of precipitation or silt or clay in H1 and H3, reduce the probability of white oak occurrence (Table 4).

Table 4 Stepwise logistic regression coefficients and fit statistics for significant variables correlated with occurrence of white oak in Mississippi

Variable	Coefficients	Standard Error	Odds Ratio	Lower Confidence Level	Upper Confidence Level	P-value
Intercept	-0.28388	1.74641				0.87087
Climate variables						
precipitation (in.)	-0.14869	0.02009	0.86184	0.82856	0.89645	0.00000
mean annual temperature (F)	0.08229	0.02715	1.08577	1.02950	1.14511	0.00244
Topographic variables						
elevation (ft)	0.00122	0.00041	1.00122	1.00042	1.00202	0.00271
percent slope	0.09292	0.01016	1.09737	1.07574	1.11944	0.00000
flow accumulation	0.10681	0.05442	1.11273	1.00014	1.23798	0.04970
Soil variables						
H1 silty*	-0.48385	0.09456	0.61641	0.51213	0.74192	0.00000
H1 clay*	-3.11858	0.52458	0.04422	0.01582	0.12364	0.00000
H2 silty	0.64217	0.18449	1.90059	1.32388	2.72854	0.00050
H2 clay	0.73620	0.18466	2.08799	1.45394	2.99855	0.00007
H3 silty	-0.26927	0.11466	0.76394	0.61018	0.95644	0.01886
H3 clay	-0.30541	0.12315	0.73682	0.57881	0.93798	0.01314

*H=horizon silty= soil having a particle size between 0.002-0.05 clay=soil having a particle size <0.002

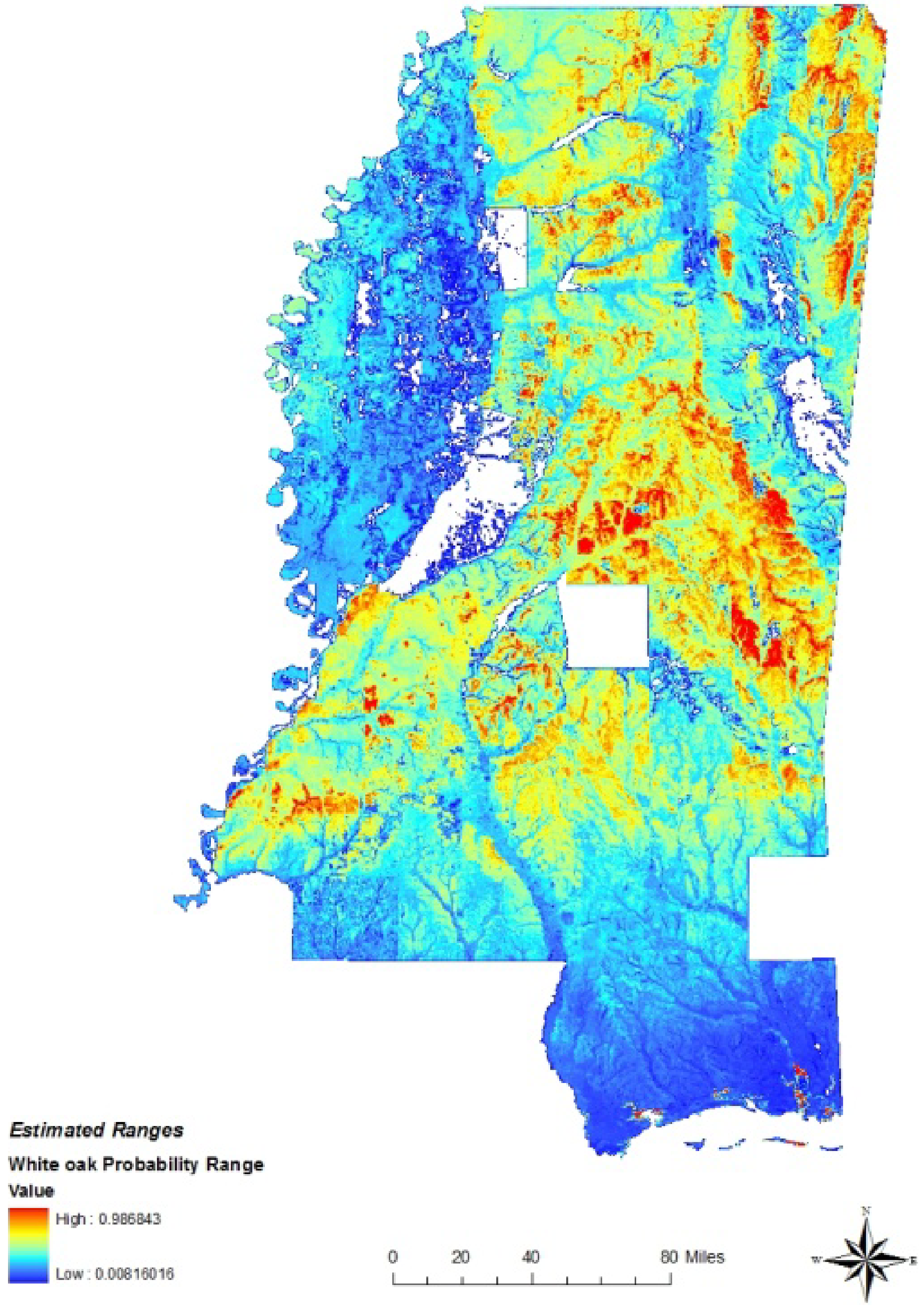

Figure 14	Probability map showing the predicted range of white oak (*Q. alba*) derived from the stepwise logistic regression parameters of the significant variables: elevation, precipitation, temperature, percent slope, flow accumulation, H1, H2, and H3 texture classes.

Southern red oak

Elevation, precipitation, percent slope, flow accumulation, drainage class, and texture class of the soils were all significant in explaining the occurrence of southern red oak ($P < 0.05$; Table 5, Figure 15). The test regression iterations revealed that the variables flow accumulation, and H2 and H3 texture class, were not significant over 50 % of the iterations (Appendix-Table 16). Slope was significant 100% of the iterations. Although elevation was significant, it had an odds ratio of 1.00074, which shows that there is only a little more than random association between elevation and the presence of southern red oak. Soil textures play a significant role in the presence of southern red oak with most texture classes being negatively correlated to the occurrence of southern red oak with the exception of H2 (Table 5). The more clay found in H2, the higher the probability of southern red oak occurring (Table 5).

Drainage classes would appear to have the absolute highest influence on the occurrence of southern red oak although it is they are more of an indicator of where southern red oaks would not occur. Excessively drained and very poorly drained sites have the highest significance when looking at the individual p-values but it would appear the southern red oak tends to favor well-drained sites more than any drainage class with an odds ratio of 0.82462 and with the overall drainage class has a p-value of 0.0138 this is still applicable.

Table 5 Stepwise logistic regression coefficients and fit statistics for significant variables correlated with occurrence of southern red oak in Mississippi

Variable	Coefficients	Standard Error	Odds Ratio	Lower Confidence Level	Upper Confidence Level	P-value
Intercept	2.50966	0.99458				0.01162
Climate Variables						
elevation	0.00074	0.00037	1.00074	1.00001	1.00147	0.04581
precipitation (in.)	-0.07635	0.01592	0.92649	0.89804	0.95585	0.00000
Topographic Variables						
percent slope	0.05516	0.01080	1.05671	1.03458	1.07931	0.00000
flow accumulation	-0.12746	0.06061	0.88033	0.78173	0.99138	0.03548
Soil Variables						
excess. drain	-1.34901	0.75602	0.25950	0.05897	1.14197	0.07436
mod. well drain	-0.24894	0.19382	0.77963	0.53323	1.13989	0.19900
well drain	-0.19283	0.17839	0.82462	0.58132	1.16976	0.27971
poorly drain	-0.64614	0.21646	0.52406	0.34287	0.80101	0.00284
very poorly drain	-1.08772	0.76775	0.33699	0.07483	1.51748	0.15655
H1 silty*	-0.33796	0.10240	0.71322	0.58353	0.87174	0.00097
H1 clay*	-1.26370	0.31403	0.28261	0.15271	0.52298	0.00006
H2 silty	0.50606	0.17712	1.65875	1.17224	2.34717	0.00427
H2 clay	0.56003	0.17903	1.75073	1.23262	2.48662	0.00176
H3 silty	-0.22762	0.11365	0.79642	0.63739	0.99513	0.04519
H3 clay	-0.29870	0.12506	0.74178	0.58053	0.94781	0.01692

*H=horizon silty= soil having a particle size between 0.002-0.05 clay=soil having a particle size <0.002

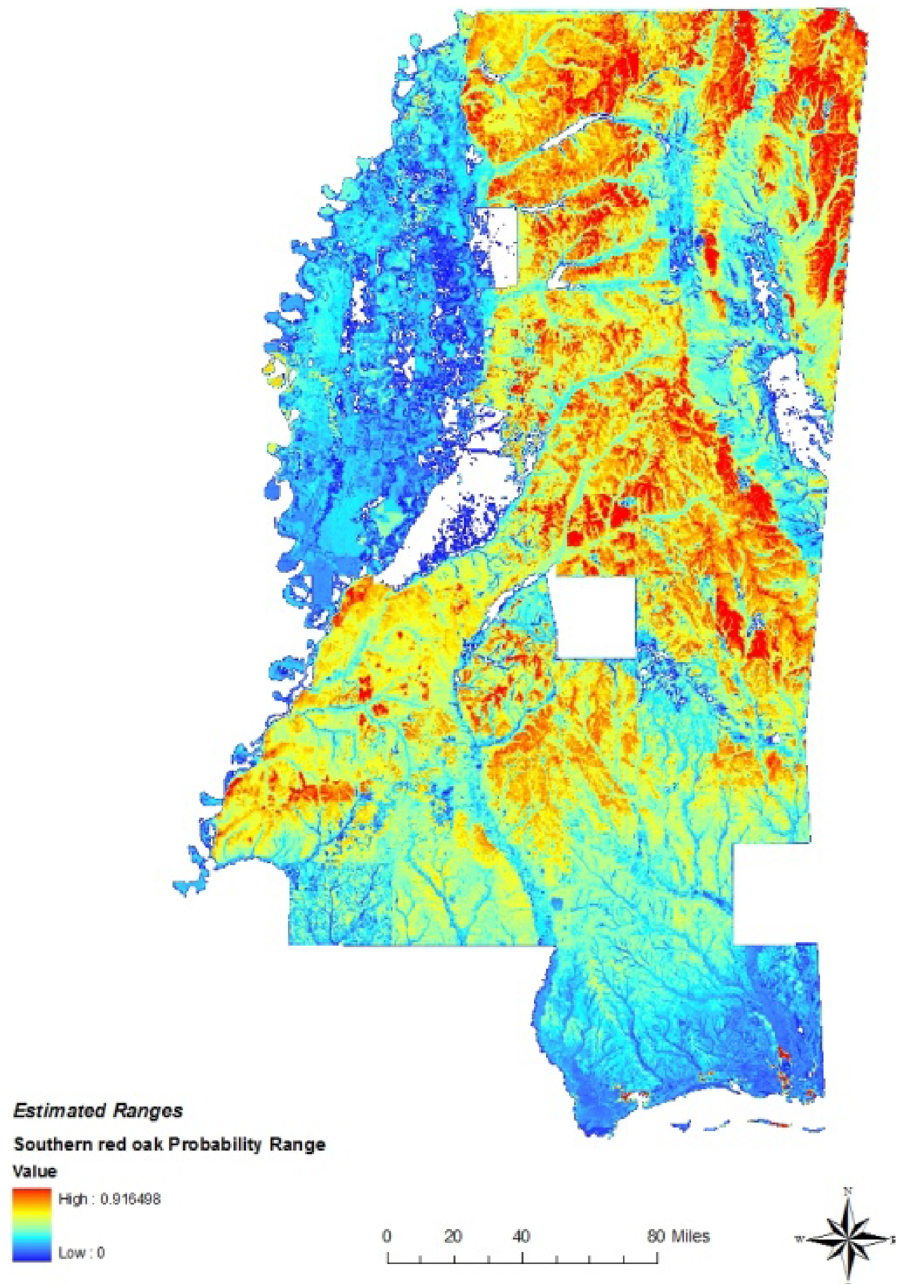

Figure 15 Probability map showing the predicted range of southern red oak (*Q. falcata*) derived from the stepwise logistic regression parameters of the significant variables: elevation, precipitation, percent slope, flow accumulation, drainage class, H1, H2, and H3 texture classes.

Blackjack oak

Elevation, precipitation, temperature, percent slope, and texture class of the soils in the second and third horizons were all significant in the occurrence of blackjack oak ($P< 0.05$; Table 6, Figure 16). The test regression iterations revealed that only one variable, H3 texture class, was not significant over 50 % of the iterations (Appendix-Table 16). Clay in soil H2 and higher temperatures negatively influence the occurrence of blackjack oak; whereas clay in H3 is the highest predictor in blackjack occurrence with an odds ratio of 2.66806 and is positively correlated with occurrence. Sandy texture in H2 appears to be significant for blackjack oak since the p-values of silty and clayey textures in this layers are both greater than 0.05 but the overall p-value is <0.0001, meaning that the reference group, sand, must be the significant variable.

Table 6 Stepwise logistic regression coefficients and fit statistics for significant variables correlated with occurrence of blackjack oak in Mississippi

Variable	Coefficients	Standard Error	Odds Ratio	Lower Confidence Level	Upper Confidence Level	P-value
Intercept	19.59757	3.54891				0.00000
Climate variables						
precipitation (in.)	0.24871	0.04202	1.28236	1.18098	1.39245	0.00000
mean annual temperature (F)	-0.54073	0.05564	0.58232	0.52216	0.64941	0.00000
Topographic variables						
elevation (ft)	0.00460	0.00075	1.00462	1.00313	1.00610	0.00000
percent slope	0.09993	0.01688	1.10509	1.06913	1.14226	0.00000
Soil variables						
H2 silty*	0.66081	0.53740	1.93636	0.67538	5.55170	0.21883
H2 clay*	-0.24963	0.55619	0.77909	0.26192	2.31748	0.65356
H3 silty	0.59286	0.23671	1.80916	1.13759	2.87717	0.01226
H3 clay	0.98135	0.25194	2.66806	1.62832	4.37171	0.00010

*H=horizon silty= soil having a particle size between 0.002-0.05 clay=soil having a particle size <0.002

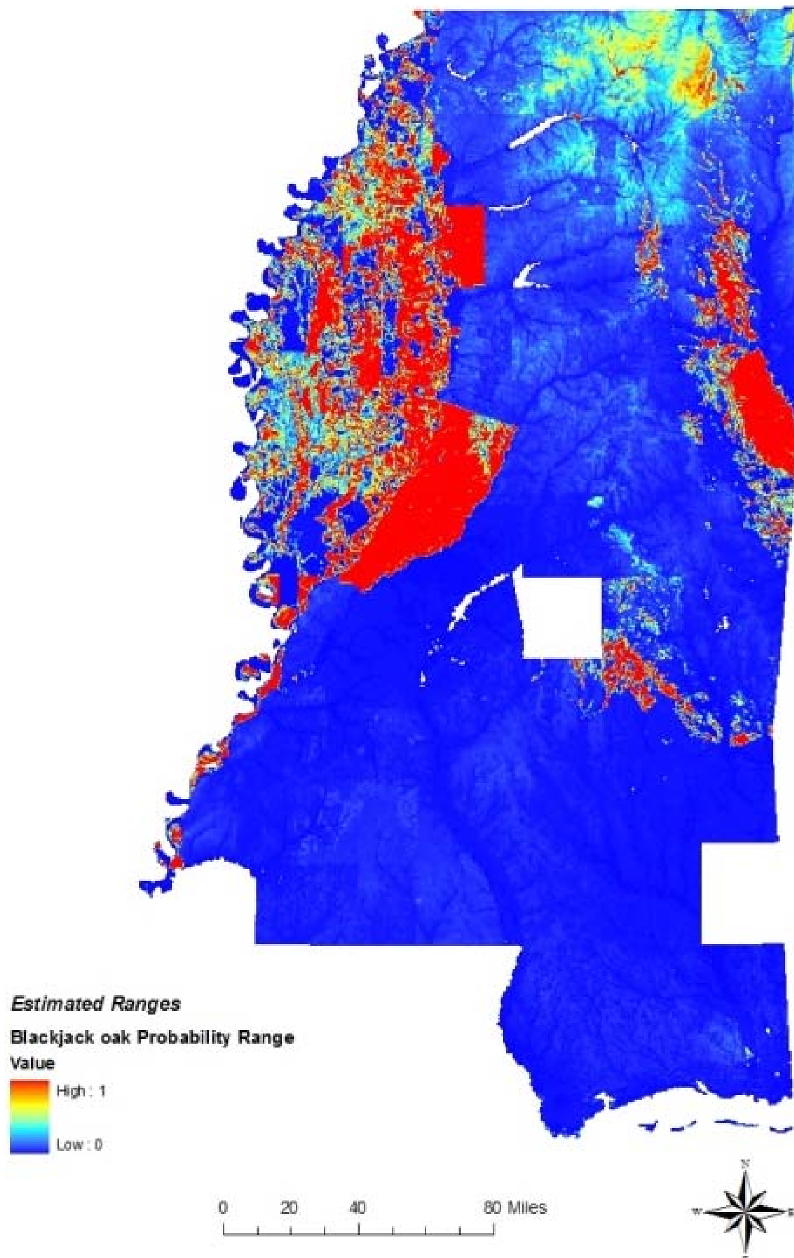

Figure 16 Probability map showing the predicted range of blackjack oak (*Q. marilandica*) derived from the stepwise logistic regression parameters of the significant variables: elevation, precipitation, temperature, percent slope, H2 and H3 texture classes.

Nuttall oak

Elevation, precipitation, and texture class of the soils in H1 and H3 were all significant in predicting the occurrence of Nuttall oak (P< 0.05; Table 7, Figure 17). The test regression iterations revealed that only one variable, H1 texture class, was significant over 50 % of the iterations no other variables that were significant in the full model were found to be consistently significant in the iterations (Appendix-Table 16). Clay in H1 and soil texture of H3 were the only variables that positively influenced the presence of Nuttall oak, all with very high odds ratios of 1.99667 or greater. On soils that were classified as having clay as their texture class in H3, the odds of finding Nuttall oak on these sites were 5.52 more times than any other variable when all other variables are held at a constant.

Table 7 Stepwise logistic regression coefficients and fit statistics for significant variables correlated with occurrence of Nuttall oak in Mississippi.

Variable	Coefficients	Standard Error	Odds Ratio	Lower Confidence Level	Upper Confidence Level	P-value
Intercept	6.68185	3.41819				0.05061
Climatic variable						
precipitation (in.)	-0.18945	0.05589	0.82741	0.74156	0.92320	0.00070
Topographic variable						
elevation (ft)	-0.00641	0.00151	0.99361	0.99067	0.99656	0.00002
Soil variable						
H1 silty*	-0.17479	0.42505	0.83963	0.36499	1.93149	0.68091
H1 clay*	1.35600	0.52419	3.88066	1.38905	10.84158	0.00968
H3 silty	0.69148	0.67417	1.99667	0.53266	7.48452	0.30505
H3 clay	1.70906	0.66414	5.52378	1.50286	20.30273	0.01007

*H=horizon silty= soil having a particle size between 0.002-0.05 clay=soil having a particle size <0.002

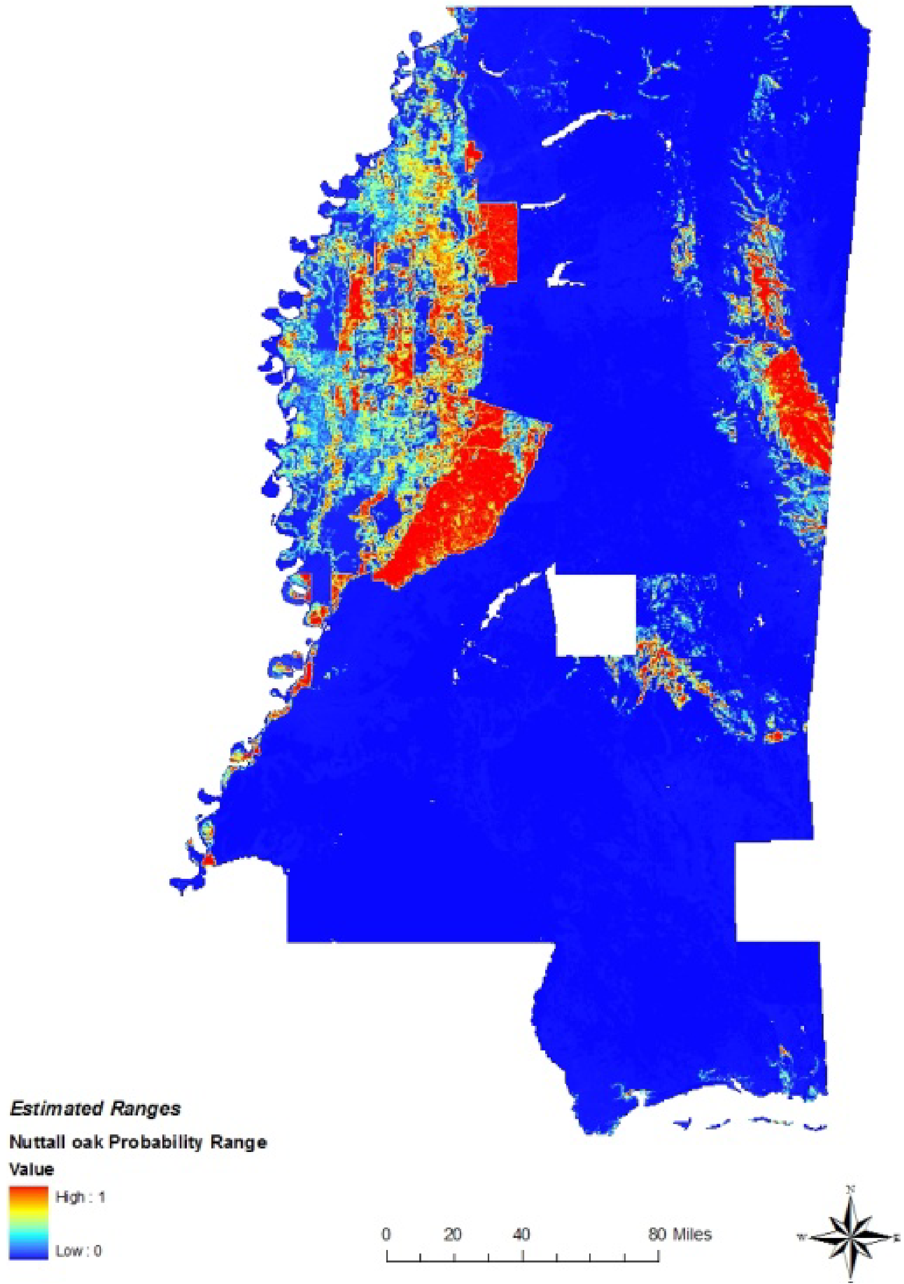

Figure 17 Probability map showing the predicted range of Nuttall oak (*Q. nuttallii*) derived from the stepwise logistic regression parameters of the significant variables: elevation, precipitation, H1 and H3 texture classes

Cherrybark oak

Elevation, precipitation, temperature, drainage class, and texture class of the soil H1 were all significant in the occurrence of cherrybark oak (P< 0.05; Table 8, Figure 18.

The test regression iterations revealed that only variables, temperature and drainage class, were not significant over 50 % of the iterations (Appendix-Table 16). High temperature and H1 silty texture were the only variables that were positively correlated to the presence of cherrybark oak. Silty texture in H1 was the most influential with an odds ratio of 1.77293 meaning that cherrybark oaks are 1.77 times more likely to occur on soils with a silty texture in the first horizon than on any other soil texture type in that horizon. From the results of the regression, it is obvious that moderately well drained sites are the most significant in the presence of cherrybark oak since moderately well drained is the only drainage class with a p-value less than 0.05 (Table 8).

Table 8 Stepwise logistic regression coefficients and fit statistics for significant variables correlated with occurrence of cherrybark oak in Mississippi.

Variable	Coefficients	Standard Error	Odds Ratio	Lower Confidence Level	Upper Confidence Level	P-value
Intercept	3.43127	2.58899				0.18506
Climate variable						
precipitation (in.)	-0.25937	0.02892	0.77153	0.72901	0.81654	0.00000
mean annual temperature (F)	0.12976	0.04278	1.13855	1.04698	1.23814	0.00242
Topographic variable						
elevation (ft)	-0.00250	0.00058	0.99750	0.99637	0.99864	0.00002
Soil variable						
excess. drain	-1.62470	1.05075	0.19697	0.02512	1.54457	0.12205
mod. well drain	-0.84144	0.26581	0.43109	0.25604	0.72582	0.00155
well drain	-0.39625	0.25049	0.67284	0.41181	1.09933	0.11368
poorly drain	-0.33220	0.25545	0.71735	0.43480	1.18350	0.19345
very poor drain	-0.79947	0.59080	0.44957	0.14122	1.43116	0.17599
H1 silty*	0.57263	0.13523	1.77293	1.36014	2.31100	0.00002
H1 clay*	-0.90714	0.27758	0.40368	0.23429	0.69553	0.00108

*H=horizon silty= soil having a particle size between 0.002-0.05 clay=soil having a particle size <0.002

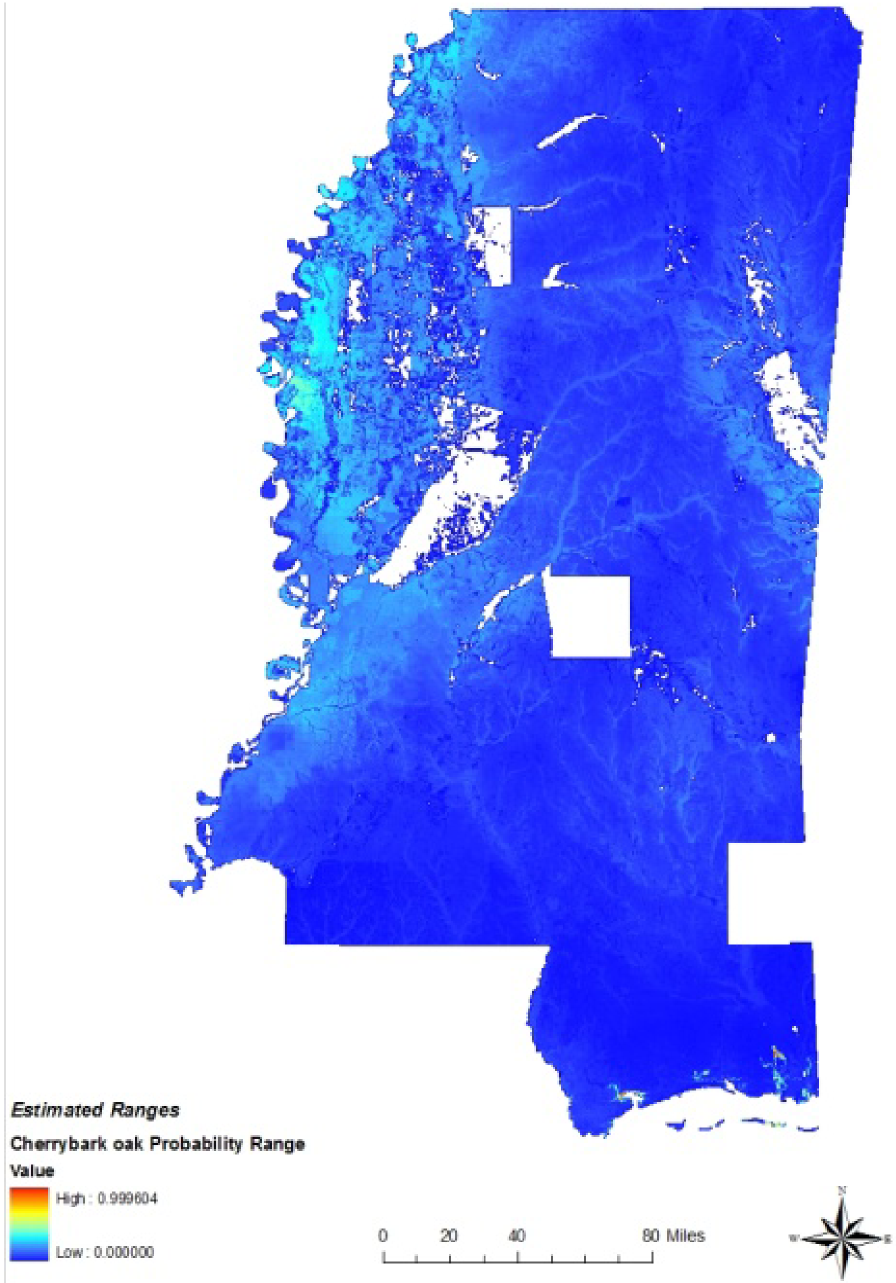

Figure 18 Probability map showing the predicted range of cherrybark oak (*Q. pagoda*) derived from the stepwise logistic regression parameters of the significant variables: elevation, precipitation, temperature, drainage class and H1 texture class.

Water/willow oak

Precipitation, temperature, percent slope, and flow accumulation were all significant in the occurrence of water/willow oak (P< 0.05; Table 9, Figure 19). The test regression iterations revealed that all variables that proved to be significant in the full model were also significant in over 80 percent of the iterations (Appendix-Table 16). Precipitation and percent slope are both negatively correlated with the occurrence of water/willow oak and both are similar in the significance in the role they play on presences with odds ratios around 0.95. High temperature is the most positively significant variable to the presence of water/willow oak by increasing the probability of occurrence by a factor of 1.38383 when all other variable are held at a constant.

Table 9 Stepwise logistic regression coefficients and fit statistics for significant variables correlated with occurrence of water/willow oak in Mississippi

Variable	Coefficients	Standard Error	Odds Ratio	Lower Confidence Level	Upper Confidence Level	P-value
Intercept	-22.68681	1.27884				0
Climate variable						
precipitation (in.)	-0.03700	0.01202	0.96368	0.94125	0.98665	0.00208
temperature (F)	0.32485	0.02211	1.38383	1.32513	1.44513	0.00000
Topographic variable						
percent slope	-0.04942	0.00893	0.95178	0.93527	0.96859	0.00000
flow accumulation	0.16372	0.03961	1.17789	1.08989	1.27298	0.00004

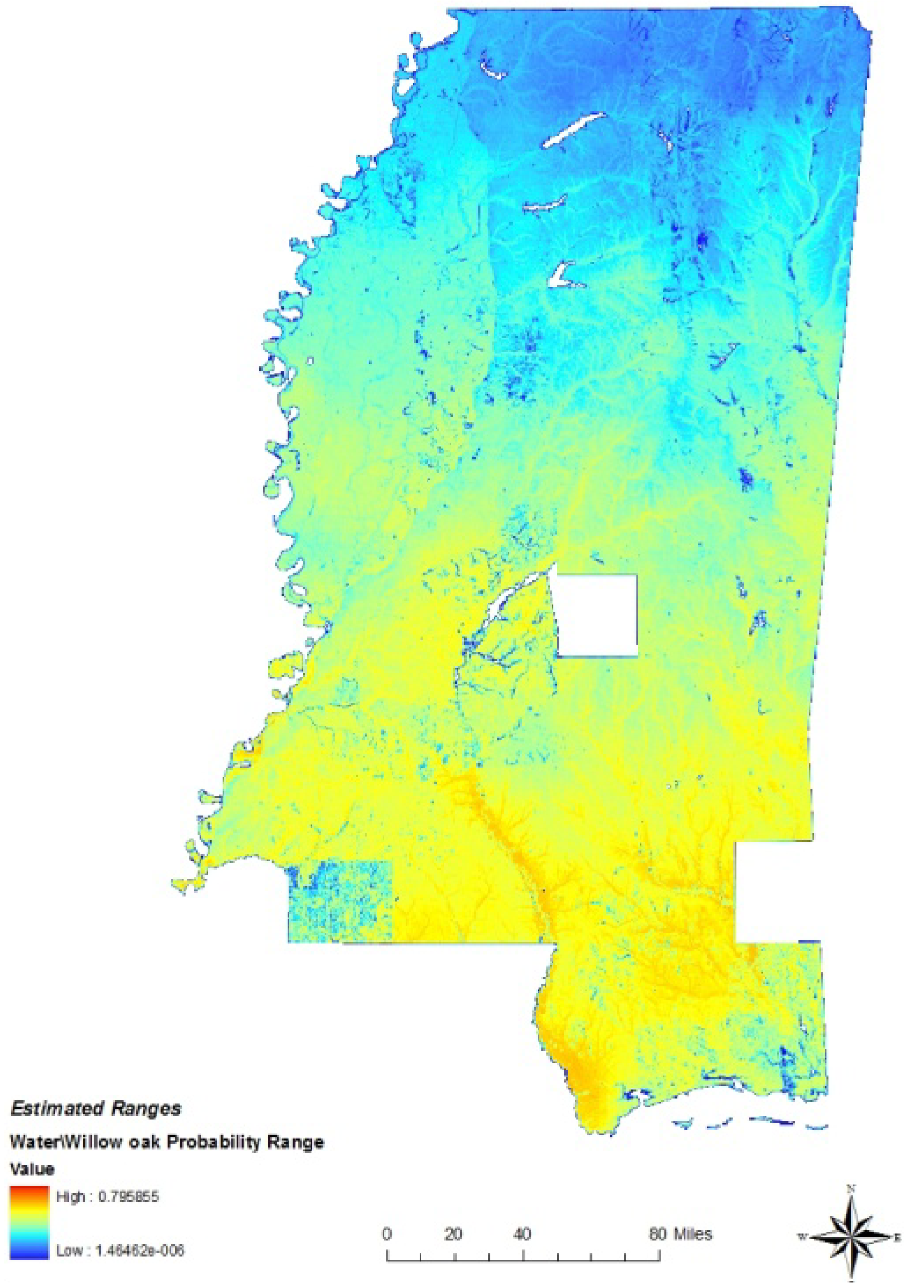

Figure 19 Probability map showing the predicted range of water/willow oak (*Q. nigra/phellos*) derived from the stepwise logistic regression parameters of the significant variables-precipitation, temperature, percent slope and flow accumulation.

Shumard oak

Precipitation, temperature, and percent slope were the only three variables found to be significant in the occurrence of Shumard oak (P< 0.05; Table 10, Figure 20). The test regression iterations revealed that only one variable, temperature, was not significant over 50 % of the iterations (Appendix-Table 16). Precipitation and percent slope were only significant 53% and 56% of the times respectively (Appendix -Table 16). The chance of the occurrence of Shumard oak will decrease with every unit of increase of precipitation by a factor of 0.74543 (Table 10). An increase of temperature will increase the probability of Shumard oak occurrence the most out of all the significant variables with an increase by 1.37732.

Table 10 Stepwise logistic regression coefficients and fit statistics for significant variables correlated with occurrence of Shumard oak in Mississippi

Variable	Coefficients	Standard Error	Odds Ratio	Lower Confidence Level	Upper Confidence Level	P-value
Intercept	-12.35415	5.65138				
Climate variables						
precipitation (in.)	-0.29379	0.06745	0.74543	0.65312	0.85078	0.00001
mean annual temperature (F)	0.32014	0.09183	1.37732	1.15046	1.64891	0.00049
Topographic variable						
percent slope	0.12083	0.02402	1.12843	1.07654	1.18283	0.00000

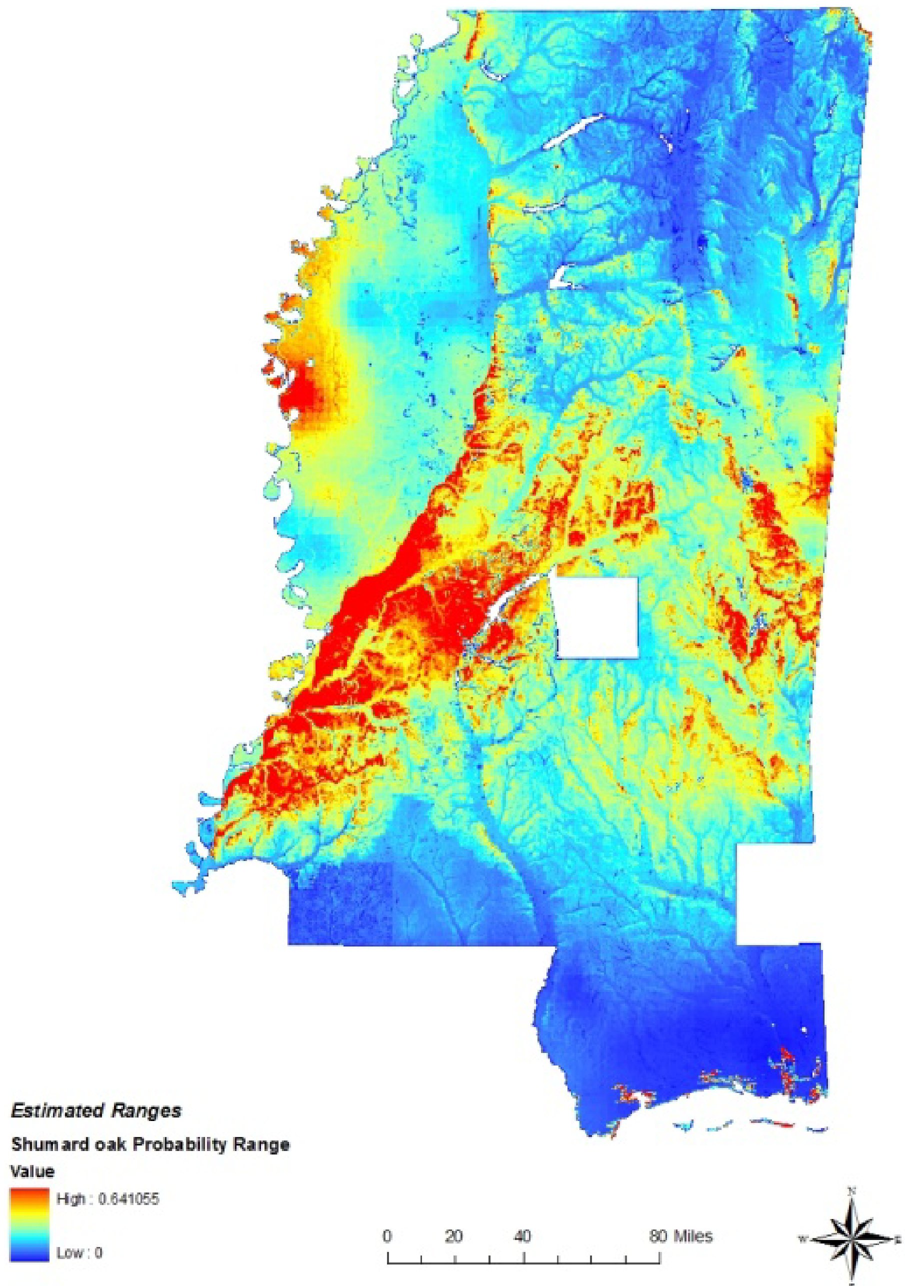

Figure 20 Probability map showing the predicted range of Shumard oak (*Q. shumardii*) derived from the stepwise logistic regression parameters of the significant variables: precipitation, temperature and percent slope.

Post oak

Elevation, percent slope, flow accumulation, and texture class of H1 and H2 were significant in the occurrence of post oak (P< 0.05; Table 11, Figure 21. The test regression iterations revealed that only variables, precipitation and H2 texture class, were not significant over 50 % of the iterations. Flow accumulation and textures of H1 are negatively related to the occurrence of post oak. Whereas, soils with textures of silt and clay in H2 are the most significant determinates on the occurrence of post oak with an increase of probability by factors of 1.85488 and 1.90103.

Table 11 Stepwise logistic regression coefficients and fit statistics for significant variables correlated with occurrence of post oak in Mississippi.

Variable	Coefficients	Standard Error	Odds Ratio	Lower Confidence Level	Upper Confidence Level	P-value
Intercept	-3.40820	0.24004				0.00000
Topographic variables						
elevation	0.00172	0.00044	1.00172	1.00085	1.00259	0.00011
percent slope	0.03032	0.01372	1.03079	1.00344	1.05888	0.02708
flow accumulation	-0.38893	0.09691	0.67778	0.56053	0.81956	0.00006
Soil variables						
H1 silty*	-0.35060	0.11207	0.70426	0.56538	0.87726	0.00176
H1 clay*	-1.13735	0.36569	0.32067	0.15660	0.65664	0.00187
H2 silty	0.61782	0.23440	1.85488	1.17164	2.93654	0.00839
H2 clay	0.64240	0.24181	1.90103	1.18349	3.05362	0.00789

*H=horizon silty= soil having a particle size between 0.002-0.05 clay=soil having a particle size <0.002

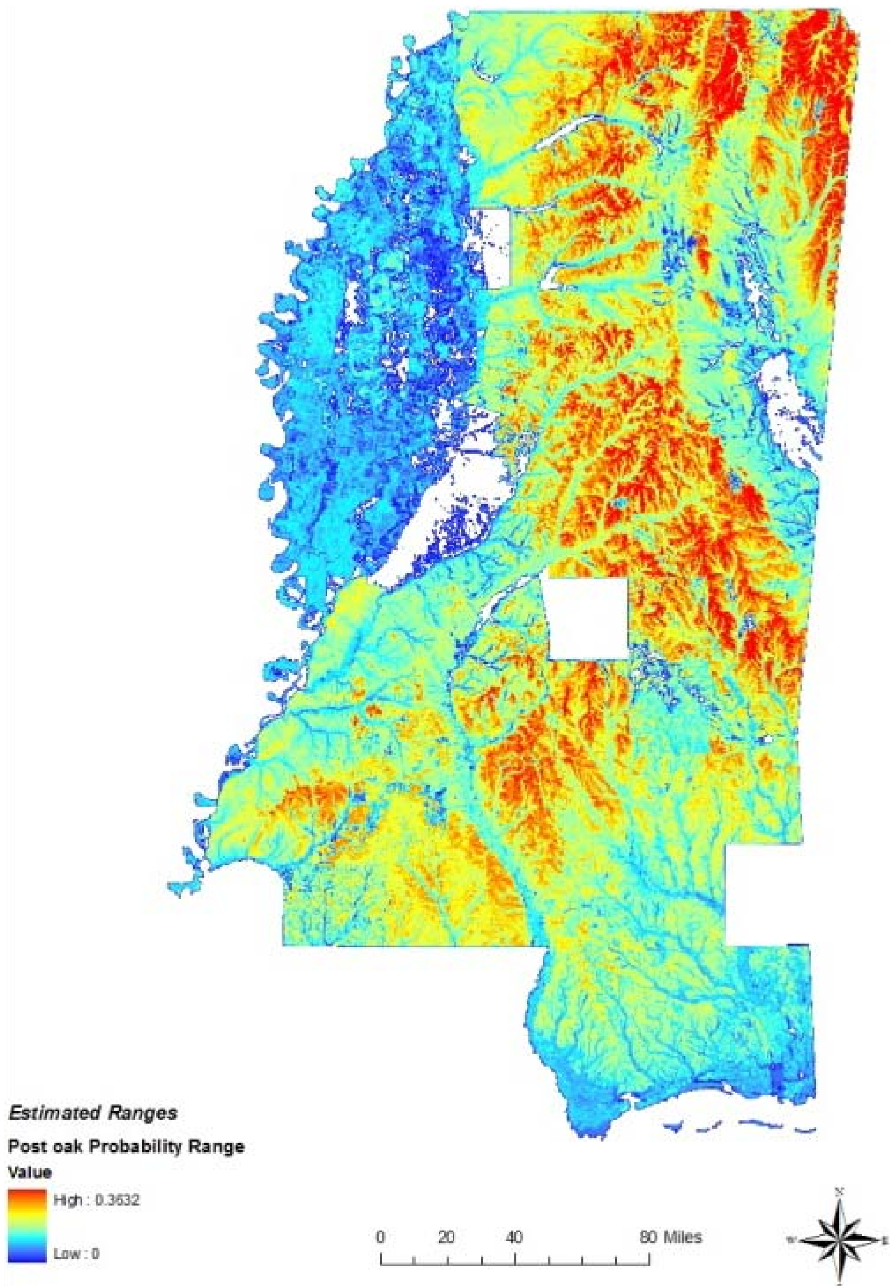

Figure 21 Probability map showing the predicted range of post oak (Q. stellata) derived from the stepwise logistic regression parameters of the significant variables: elevation, percent slope, flow accumulation, H1 and H2 texture classes.

Model Validity

The accuracy of the probability maps yielded from the logistic coefficients (Figures 14-21) can be seen in Table 12 where the total numbers of plots known to have the specific oaks were used to validate the models. The mean and maximum columns are the validation plots probability values with known species occurrence derived from the probability maps. All species had the majority of presence points falling on relatively low probability portions of the maps with water/willow oak having the highest mean accuracy of 38%. Nuttall oak had the highest maximum accuracy with at least one point accurately being predicted to occur on a site in which occurrence had been visually verified. Only four species were able to correlate with the probability of occurrence better than 50 % with maximums of 0.60 and greater.

Table 12 The mean and maximum prediction probability of occurrence for validation plots for *Quercus* spp.

Oak Species	Count	Mean (%)	Maximum (%)
blackjack	66	0.27	0.65
cherrybark	129	0.12	0.33
Nuttall	30	0.08	1.00
post	129	0.10	0.21
Shumard	18	0.01	0.04
southern red	272	0.21	0.46
water/willow	651	0.38	0.72
white	270	0.22	0.60

CHAPTER IV

DISCUSSION

Soil Database

The SSURGO database data caused many problems when it came to the mapping of soils across Mississippi. There are several unavoidable flaws in these data that incorporated inaccuracies into the mapping of expected probability of suitable abiotic site characteristics for southern oaks across Mississippi. The SSURGO database is derived from soil surveys collected at a county level. According to Hodgkins, *et al.* (1979), the accuracy of interpretation derived from these soil surveys will be "highly unreliable" for several reasons. SSURGO fails as a unified statewide system because of "intensive revisions and refinements in the Cooperative Soil Survey system" (Hodgkins, *et al.*, 1979). These revisions affect changes in descriptions of soil types, making it possible that adjacent counties soil characteristics if collected years apart will not transition smoothly over the arbitrary county lines. This is easily seen in the smooth edges and rectangular shapes in the soil data around Holmes, Tallahatchie, and Yazoo counties. The variation in collection techniques is compounded when some of the soil surveys that make up the state of Mississippi were published over 40 years ago. Although the SSURGO database is not ideal to use at a statewide level, it is the most comprehensive database publically available.

Significant Variables

Trees in the genus *Quercus* can occupy number of niches due to these species abilities to survive and grow over a range of environmental conditions (Johnson, *et al.*, 2002). The niche of an individual species is narrower due to the ecological gradients the species prefer. The ecological amplitude or range of habitat conditions that an oak species can tolerate often exhibits a bell-shaped curve form (Johnson, *et al*, 2002). This distribution became evident in the results of this study; often times it was not a specific texture or slope that was preferred by the oaks but a range of conditions, making it difficult to pinpoint oak site preference by a significant variable.

White oak

According to the *Silvics of North America Volume 2 Hardwoods* (Rodgers, 1990), the preferred site conditions influencing the occurrence of white oak are wide ranging and variable. White oak is typically "found on sandy plains, gravelly ridges, rich uplands, coves and well-drained loamy soils" (Rodgers, 1990). Topography and aspect play a role on occurrence, white oak grows mainly on "upland aspects and slopes" except for "extremely dry, shallow-soil ridges; poorly drained flats; and wet bottom land" (Rodgers,1990). White oak also occurs under a wide variety of temperatures and precipitation. White oak can be found in areas with mean annual temperatures ranging from 45°to 70°F and annual mean precipitations of 30 to 80 inches (Rodgers, 1990). The results of the stepwise logistic regression were explanatory of published site preference descriptions when viewed as a whole. The texture classes identified in the model support the notion that white oaks prefer soils that have good internal drainage with the fact that all three-horizon texture classes were significant. The sandy texture class was the most significant texture class in the first and third horizons. Although it was negatively correlated, sandy soils were the least negatively correlated. Clay in H2 is beneficial in holding the water and is the most significant soil texture for that horizon. Since sandy soil is the preferred texture in H1, this will eliminate sites in which ponding may occur since water moves freely into the second horizon when the first is comprised of high proportions of sand. The logistic regression showed that flow accumulation was significant, meaning that the more runoff that a site has entering it from up slope, the higher the chance of white oak occurring, to a certain extent. Since elevation and percent slope are also positively correlated to the occurrence of white oak, white oak would not be expected to inhabit swampy bottoms that would accumulate the majority of runoff. Supported for this if found by Rodgers (1990) statement, that white oak likes coves and well-drained soils. Little's (1971) range map and the probability range derived from the stepwise logistic regression coefficients are very similar to one another (Figure 22). The probability map stresses a higher probability on the ridges and higher elevations than the bottoms and with small probability of occurrence in the northwest portion of Mississippi in the region of the Mississippi and Yazoo River floodplain, here after known as the Mississippi Delta, and along the Gulf coast.

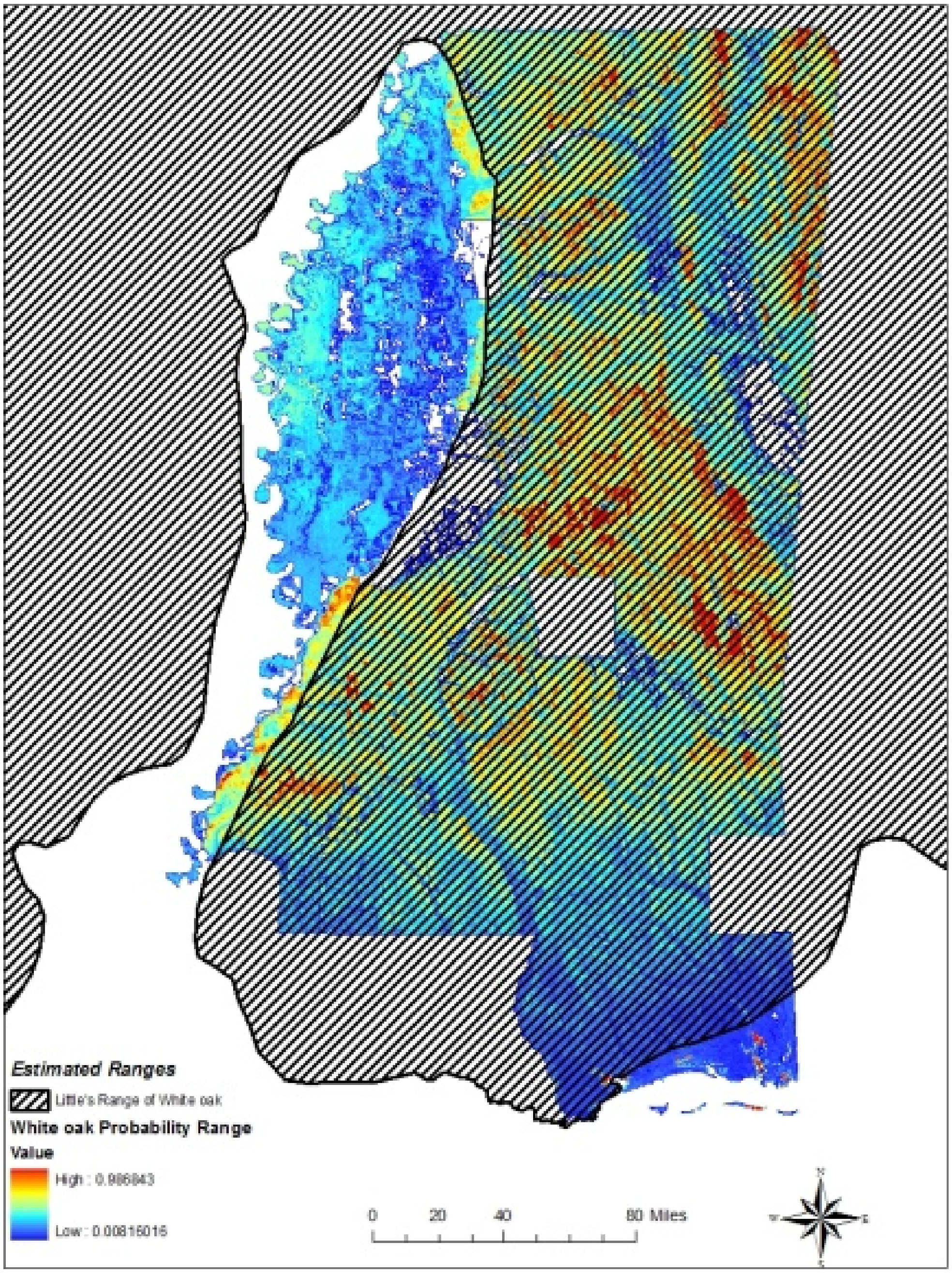

Figure 22 Probability map derived from the significant variables logistic regression parameters showing the predicted range of white oak (*Q. alba*) in comparison to the Elbert Little's (Little, 1971) range map of white oak (*Q. alba*).

Southern red oak

According to Belanger (1990) in Silvics of North America Volume 2 Hardwoods, southern red oak is "one of the common upland oaks" that prefers to grow on dry, sandy clay soils. Southern red oaks grow on dry ridge tops and upper slopes and are commonly found on slopes with aspects that are south and west facing. The stepwise logistic regression results support these findings (Table 5). The logistic regression indicated a predicted increase of occurrence of southern red oak with an increase in percent slope, showing that steeper the slope, higher the probability of the occurrence of southern red oak which is exactly what Belanger inferred. The decrease of probability of occurrence with the increase of flow accumulation of a site also supports that southern red oak prefers dry sites. This preference is also seen in the fact that well-drained sites had the least negative effect on the expected occurrence of southern red oak and that sand is the most influential texture of H1 and H3. Although textures in these horizons are negatively correlated with occurrence, sandy soil was the least negatively correlated. The significance of clay in H2 combined with the significance of sand in H1, supports that southern red oak prefers dry sites that are "dry, sandy, clay soils"(Belanger, 1990).

The other site variable that had a negative impact on the occurrence of southern red oak was precipitation. This is reasonable since, according to Belanger (1990), the preferred average annual precipitation for this species is between 40 to 50 inches. According to the Prism data raster, the average annual precipitation for Mississippi is between 51 and 68 inches a year, slightly more than what is preferred (Figure 7). It does, however, occur readily here, and the cause of this disagreement could be. The apparent role that precipitation plays on southern red oak distribution is so minute and variable that it overall has little influence on its occurrence in Mississippi. When compared, Little's

(1971) range map is not similar to the probability range derived from the stepwise logistic regression coefficients (Figure 23). When Little's (1971) range map was derived, the distributions of southern red oak and cherrybark oak were given by the same range map because taxonomically, they were considered as varieties of the same species. Cherrybark oak is generally found on wetter sites and is more likely to be found in areas like the Mississippi Delta than southern red oak. This is one example to support the need of more updated range maps. The probability map for southern red oak more logically stresses a higher probability on the ridges and higher elevations than the bottoms and with small probability of occurrence in the Delta and along the Mississippi coast.

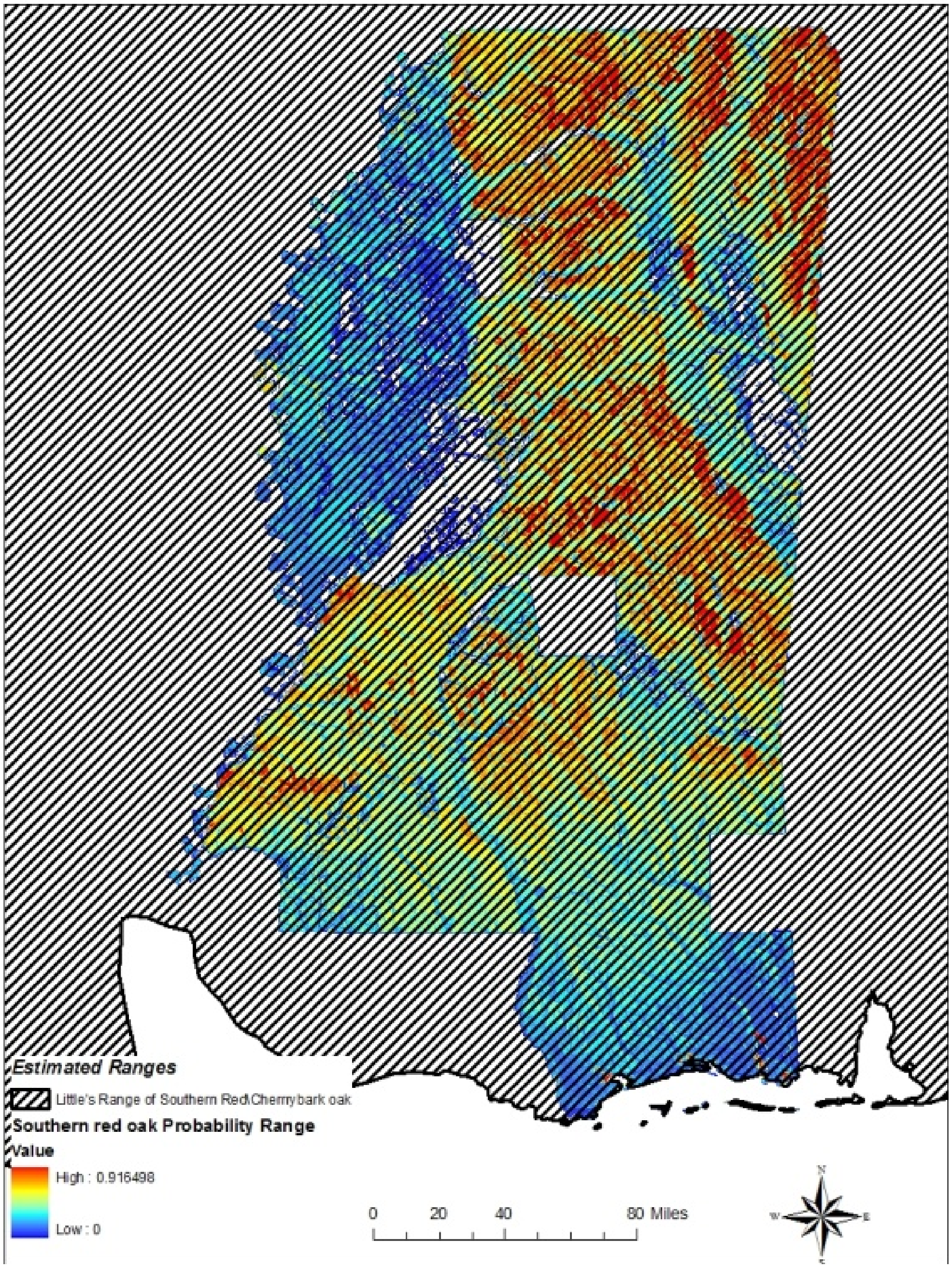

Figure 23 Probability map derived from the significant variables logistic regression parameters showing the predicted range of southern red oak (*Q. falcata*) in comparison to the Elbert Little's (1971) range map of southern red\cherrybark oak.

Blackjack oak

The results for blackjack oak differ strongly from site ecology reported in the literature. Carey (1992) states that blackjack oak is a semi-xeric species that can be found on xeric sandy deposits and on dry upper slopes and ridges. This species of oak is reported to occur on southerly and westerly aspects that are dry and nutrient poor (Carey, 1992). The logistic regression results of elevation, precipitation, percent slope, and sandy texture in H2 having a positive significance on the occurrence of blackjack oak and generally supports what Carey (1992) reported. The deviation occurs with the high significance of clay in H3. Clay in H3 is the most significant variable in the presence of blackjack oak according to the regression results. This is contrary to the actual distribution of blackjack oak. When compared, the contradiction between the probability map of blackjack oaks from regression model and Little's (1971) range map is quite evident (Figure 24). A cause of this differing could be because the SSURGO texture data for the Mississippi Delta indicates high proportions of clay in all the horizon layers.

There can be other possible causes for the apparent erroneousness of the analysis. First and foremost, it could be a result of misidentification of the species by field crews. Blackjack oak leaves might be mistaken for the shade leaves of southern red oak (*Quercus falcata*) or post oak (*Quercus stellata*). This could be coupled with the fact that only 181 cases of occurrence were used to derive the logistic estimates. Also of these 181 cases, 102 occur on clay textures. The test regression iterations revealed that the H3 texture class was significant only for 6% of the iterations (Appendix-Table 16). The significance of clay seen in the full model could have been due to the overestimation of absences.

Soils data, mentioned previously as being inconsistent, are likely the third source of error for this species. The high clay content of the soils for Holmes, Tallahatchie and Yazoo counties and the Blackbelt Prairie region in east-central Mississippi are a cause for concern because of symmetrical shapes of these features and dissimilarity to the surrounding county mapunits. These areas are easily seen in the probability map (Figure 24). The combination of these flaws came together to yield the probability map seen in comparison to Little's (1971) map with readily apparent overall differences for blackjack oak distributions (Figure 24).

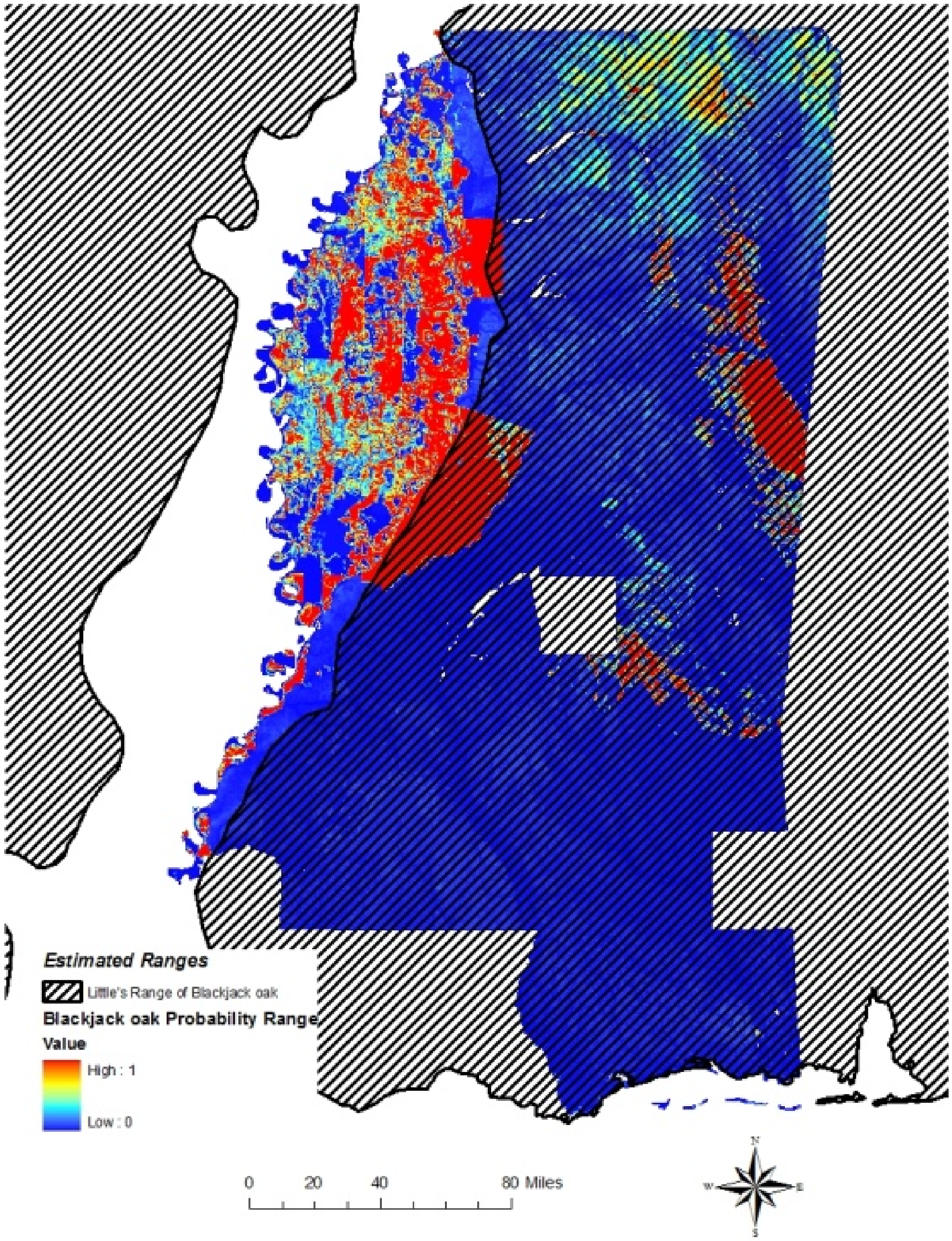

Figure 24 Probability map derived from the significant variables logistic regression parameters showing the predicted range of blackjack oak (*Q. marilandica*) in comparison to the Elbert Little's (1971) range map of blackjack oak (*Q. marilandica*).

Nuttall oak

According to *Silvics of North America Volume 2 Hardwoods* (Filer, 1990), Nuttall oak occurs on clay flats that are poorly drained and in low bottoms of the Gulf Coastal Plain. This species of oak grows "best on alluvial bottomlands of the Mississippi River and its tributaries" since it prefers heavy, poorly drained alluvial clay soils (Filer, 1990). These soils are found in "the first bottoms of the Mississippi Delta region" (Filer, 1990). Nuttall oak is also said to be commonly found on clay ridges (Filer, 1990). The results from the stepwise logistic regression strongly support the preference of clayey soils. Clay was the most significant factor in the presence of Nuttall oak over any other variable. This preference for clay could be correlated to why elevation is negatively influential to the presence of Nuttall oak. Clay tends to occur more in the H1 at lower elevations due to particle size. The clay particles are so small that they tend to stay suspended in water longer than courser sediments and thus settle in low areas with little water movement. Nuttall oak preference for soils with high clay content was shown consistently to be significant through the iterations (Appendix Table 16) indicating that it is consistently an important explanatory variable across the state. Filer (1990) states that Nuttall oaks are not found in permanent swamps.

Precipitation proved to be negatively correlated with the presence of Nuttall oak. According to Filer (1990), Nuttall oaks grow in regions of the country that are characteristic of annual precipitation of 50 to 65 inches and mean temperature ranging between 45° F and 80° F. Mississippi's annual precipitation falls completely within this with a range between 51 to 68 inches.

When comparing the probability map derived from regression analysis with Elbert Little's (1971) range map (Figure 25), it is easily seen that they vary slightly. The probability map's red areas are areas that, as compared to Figures 9-11, are largely in the overall clay texture class. The main variation between the two ranges occurs in the area of the Tombigbee River in the Blackbelt Prairie region in northeastern Mississippi. This area is high in clay according to the SSURGO data and at a low relative elevation from the landscape perspective since it is the flood plain of this main river and its tributaries. These areas and the areas east of the Mississippi Delta region are somewhat misrepresented also due to the anomaly of the SSURGO data delineating these counties with such high clay content. Little's (1971) range map is also a more generalized version of the range, whereas the probability map is more site/area specific.

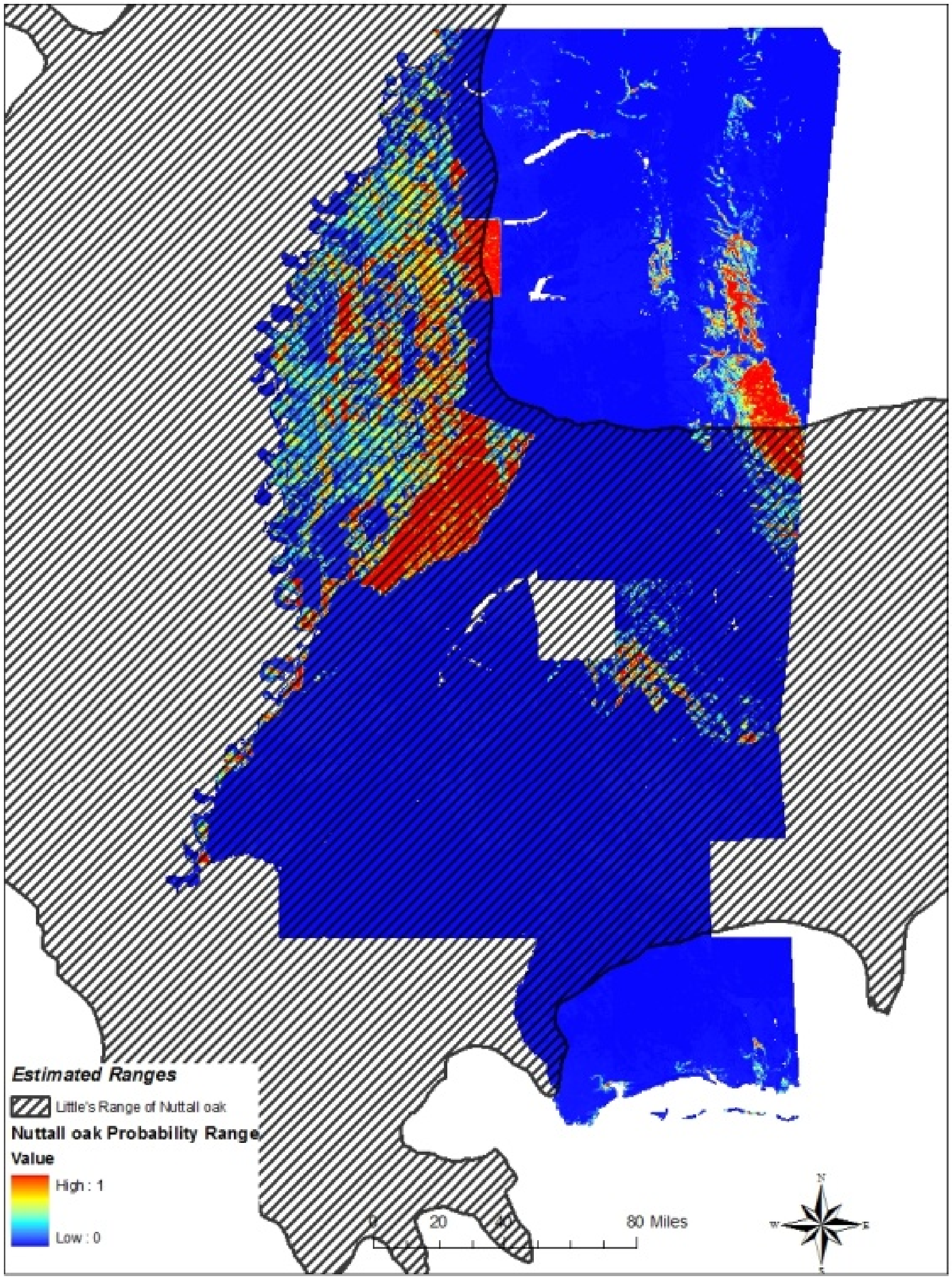

Figure 25 Probability map derived from the significant variables logistic regression parameters showing the predicted range of Nuttall oak (*Q. nuttallii*) in comparison to the Elbert Little's (Little, 1971) range map of Nuttall oak (*Q. nuttallii*).

Cherrybark oak

According to *Silvics of North America Volume 2 Hardwoods* (Krinard, 1990), cherrybark oak is a lowland tree that is generally not found on excessively wet or swampy soils. The results of the stepwise logistic regression indicated elevation as one of the significant factors that explains the occurrence of cherrybark oak. Elevation as an estimator was negatively correlated with the occurrence of this oak, meaning that it prefers lower sites and that increases in elevation, decrease the probability of presence. The majority of significant variables negatively affected the expected presence of cherrybark oak. Temperature and silt in H1 were the only two factors that were shown to be positively correlated with the presence of cherrybark oak. This was very similar to what was written by Krinard (1990). The two most decisive things that Krinard (1990) could say about cherrybark oak was that it grows in areas that have hot summers and mild winters and is "best on loamy sites". Krinard (1990) goes on to say that, cherrybark oak can be found on "first bottom ridges, well-drained terraces, and colluvial sites". Drainage is key to the sites in which cherrybark occurs according to Krinard, since it prefers sites that are loamy but with the adequate drainage, cherrybark oak can be found on clayey soils. It is rare to find it in the lower Mississippi Delta, although, cherrybark oak is found in "areas of loessial soil" and "the rolling hills of the lower Piedmont and certain uplands of the upper Coastal Plain."(Krinard, 1990) The stepwise logistic regression results show that drainage negatively affected the predicted occurrence of cherrybark oak. Drainage classes of well drained and poorly drained were the two that had the least negative effect on the occurrence of cherrybark oak.

As previously stated, Little's (1971) range map for cherrybark oak was the same as the range map of southern red oak so there is little ability to compare the two since for this study, the species were considered separate. The probability map for cherrybark oak showed no one area to be highly probable for the occurrence of cherrybark oak (Figure 26). The only area of high occurrence appears to be a small pocket near the coast. This is believed to be an abnormality that may have been caused by the errors in input data and not true of the population. Cherrybark oak occurs readily all across Mississippi, making it difficult to pinpoint significant occurrence variables. Thus, the ambiguity of site preferences made this specific model apparently ineffective in predicting occurrence. The low probability in the loess hills adjacent to the Mississippi Delta and in the region known as the Blackbelt Prairie caused by the SSURGO delineation of clay in these areas resulted in a display similar to areas with no data. Elevated probabilities were evident in areas where one might expect greater occurrence, riverine system and in the Mississippi Delta (indicated as lighter blue tones in Figure 26).

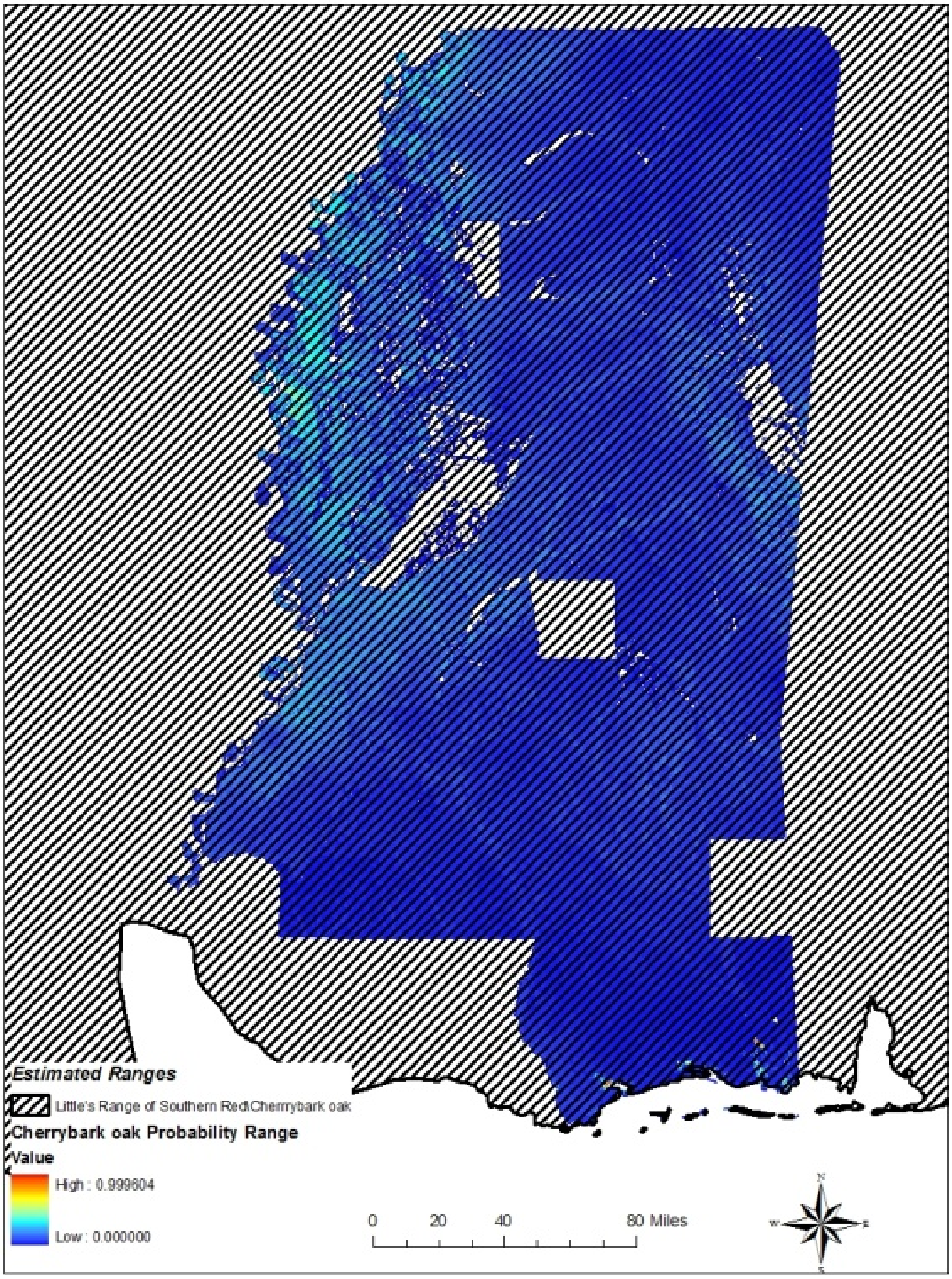

Figure 26 Probability map derived from the significant variables logistic regression parameters showing the predicted range of cherrybark oak (*Q. pagoda*) in comparison to the Elbert Little's (Little, 1971) range map of southern red\cherrybark oak.

Water/Willow oak

Water and willow oak both grow on a variety of similar sites. Water oak (*Quercus nigra*) is found on sites varying "from wet bottomlands to well drained uplands" (Vozzo, 1990). Water oak reaches its optimal quality on the well drained "silty clay or loamy soils on high flats or ridges of alluvial stream bottoms" (Vozzo, 1990). Willow oak (*Quercus phellos*) is found on ridges, flats and sloughs (Schlaegel, 1990). Willow oak is found on a "variety of alluvial soils" (Schlaegel, 1990).

It is due to this ambiguity in site between the two species, that the probability map and Little's (1971) range map (Figure 27) shows them occurring all over Mississippi. It was hard to pinpoint exact sites because of the ranges of sites that water/willow oak occurs on. The results from the logistic regression and the probability map indicate the probability of occurrence is mainly correlated with temperature and precipitation. According to *Silvics of North America Volume 2 Hardwoods*, the preferred annual precipitation for these two species range from 40 to 60 inches and temperatures between 50-70° F (Schlaegel [1990] and Vozzo [1990]). The probability map shows that the probability of occurrence decreases as it moves north (Figure 27). This could be a result of the tendency of water/willow oak to retain leaves longer into the dormancy season, which may result in greater susceptibility to damage from extreme cold. The odd square shape in the bottom left corner of southwest Mississippi (Amite County) is due to lack of soil data texture for this area.

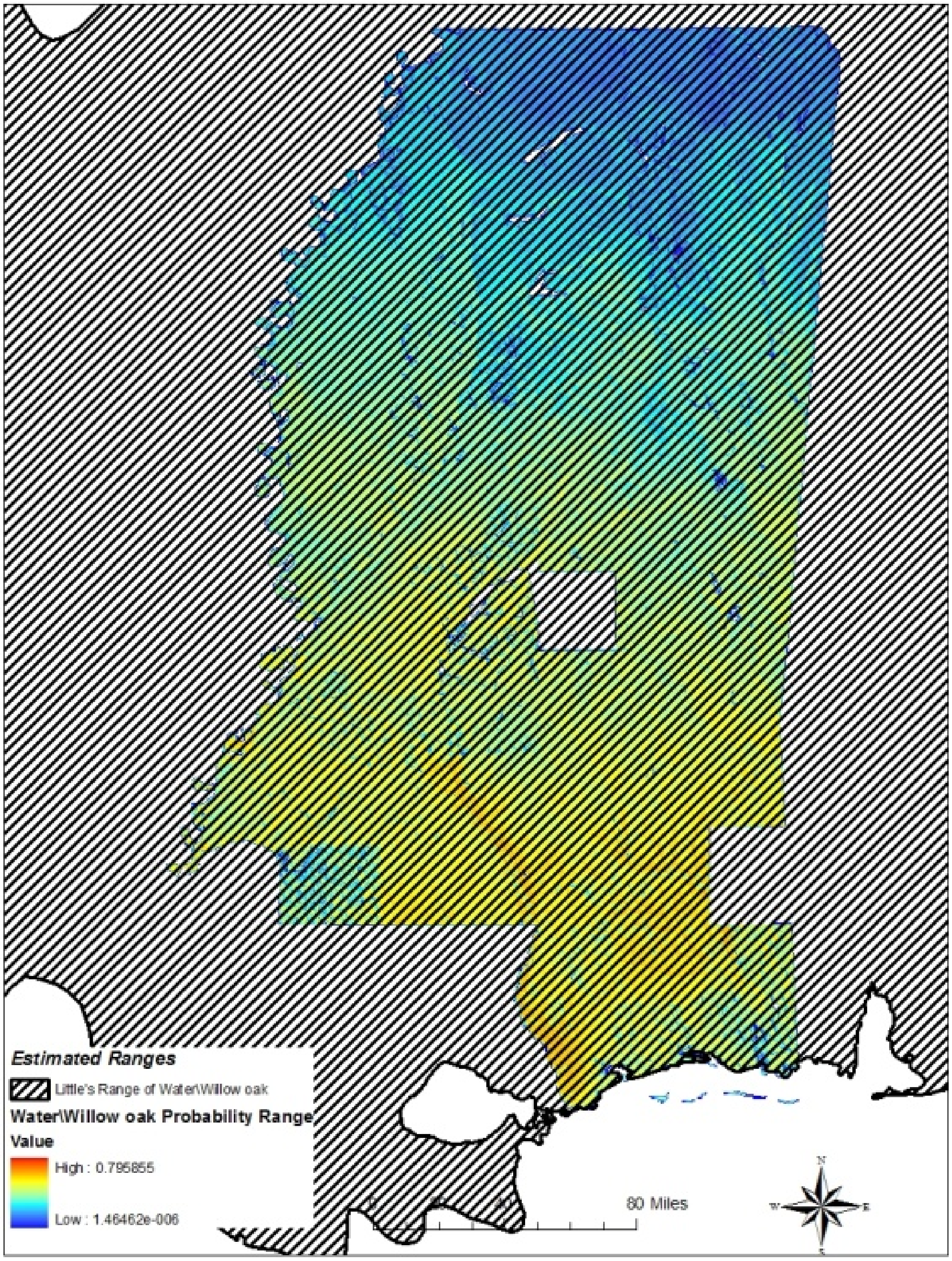

Figure 27 Probability map derived from the significant variables logistic regression parameters showing the predicted range of water/willow (*Q. nigra/phellos*) oak in comparison to the Elbert Little's (Little, 1971) range map of water/willow oak (*Q. nigra/phellos*).

Shumard oak

According to the *Silvics of North America Volume 2 Hardwoods* (Edwards, 1990), Shumard oak is a red oak that prefers "moist, well-drained loamy, soils found on terraces, colluvial sites, and adjacent bluffs associated with large and small streams". The logistic regression results agreed with this assessment with percent of slope being significant. As you increase the steepness of the slope, the likelihood of Shumard oak occurring, increases.

The results also showed that temperature and precipitation were important in predicting the occurrence of Shumard oak. Precipitation is negatively correlated to the presence of this oak. Edwards (1990) states that, Shumard oak grows in areas in which the average annual precipitation range from 45 to 55 inches. This is a little low for what we found was average for Mississippi. Average precipitation for Mississippi is 51 to 68 inches. The likelihood of Shumard occurrences decreases with an increase of precipitation. Temperature was only significant in 37% of the test regression iterations (Appendix-Table 16) showing that this variable may have been determined to be significant due to a unique physical geographic region that is not consistent throughout the state or by the overestimation of absences in the logistic regression.

The probability map compared with Little's (1971) range map are very similar to one another (Figure 28). The probability range differs from the current range map in that it shows that the range could be cut off much further from the coast than is shown by Little (1971). A high probability pocket is indicated on the western side of the Mississippi Delta. This is a result in the negative correlation of precipitation on the occurrence of Shumard oak. This area was classified as having an annual average precipitation around 50 inches, which is ideal for Shumard. Another possible cause for this high probability pocket, could be due to Shumard preference for soils with high pH and the planting of this species on the Mississippi Delta's high pH (7.8-8.0) soils (Kennedy and Krinard, 1985). Another possible concern is in the high probability of Shumard indicated in the central Mississippi hilly region. This could have been caused by the misidentification of scarlet oak (*Quercus coccinea*) as Shumard oak. This potential misidentification may also have caused elevation to be more significant in the overall prediction probability.

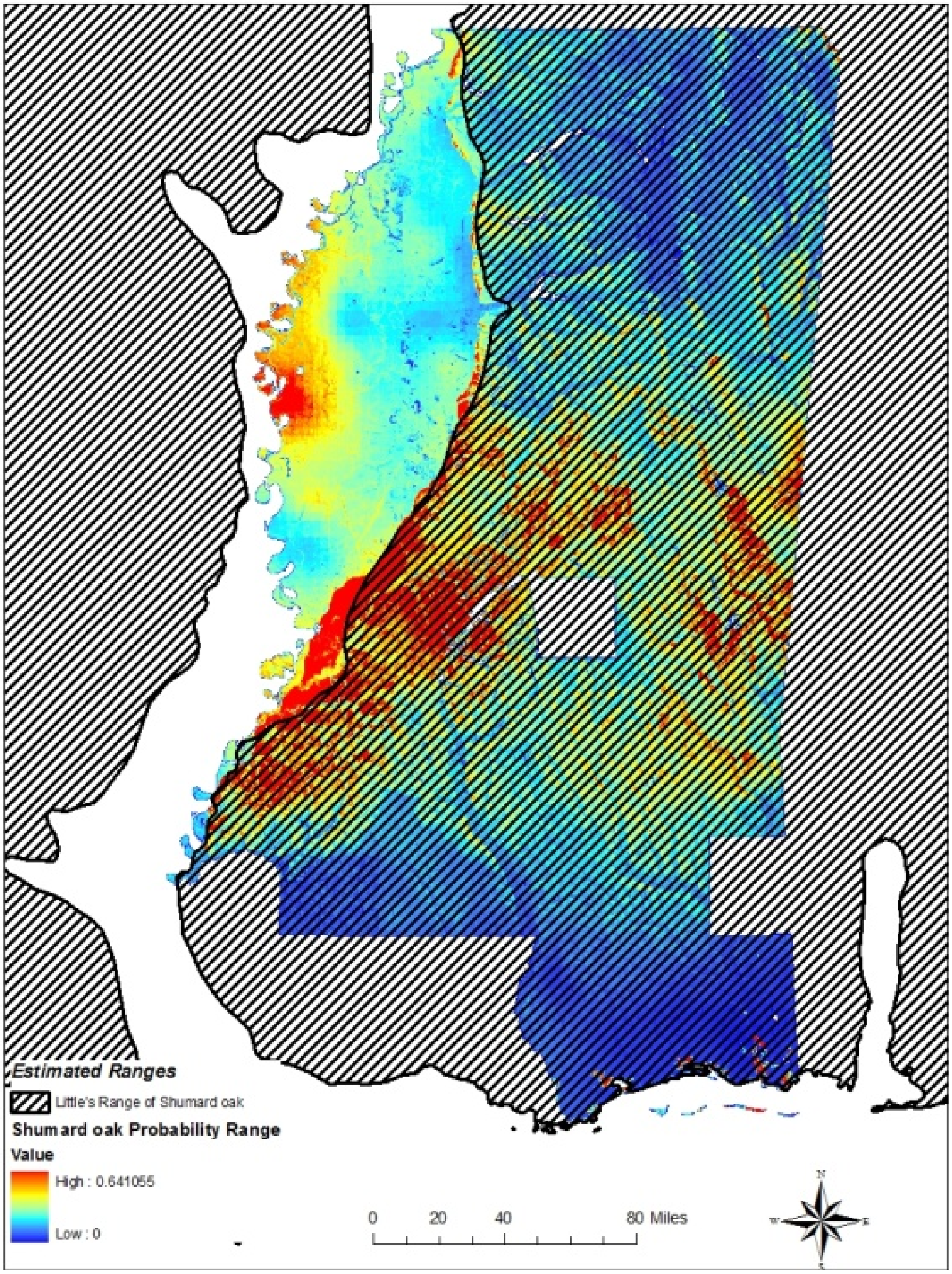

Figure 28 Probability map derived from the significant variables logistic regression parameters showing the predicted range of Shumard oak (*Q. shumardii*) in comparison to the Elbert Little's (Little, 1971) range map of Shumard oak (*Q. shumardii)*.

Post oak

According to the *Silvics of North America Volume 2 Hardwoods*, post oak typically grows on dry sites such as rocky outcrops, ridges, and upper southerly or westerly slopes (Stransky, 1990). These sites usually are poor in nutrients and well drained with course sandy textures (Stransky, 1990). The results from the analysis of the physical characteristics of the sites generally concur with these descriptions. Sand in H1 was most significant for the presence of the occurrence of post oak with clay being so in H2. Sand in H1 aids in internal drainage of the surface while clay in the second horizon often influences increased water and nutrients retention.

From the results, it is also evident that post oak prefers sites that are high in elevation with steep slopes since both variables are positively correlated with the occurrence of post oak with occurrence probability increasing as elevation and slope increase. Lower flow accumulation values are significant to the presence of post oak and correlated with the increase in elevation and are indicative of well drained sites that Stransky (1990) states that post oak prefers.

Figure 29 shows the probability map for post oak in comparison to Elbert's Little's (1971) range map and both are similar. The only deviation that the probability map makes from Little's (1971) is in the Mississippi Delta area. Little's (1971) encompasses a slight part of the Delta while the probability map cuts off the range at the Loess bluffs. This may have been due to some over simplification of the range by Little due to the mapping capabilities at the period or perhaps by inclusion of delta post oak (*Quercus similis*) in the identification of this species.

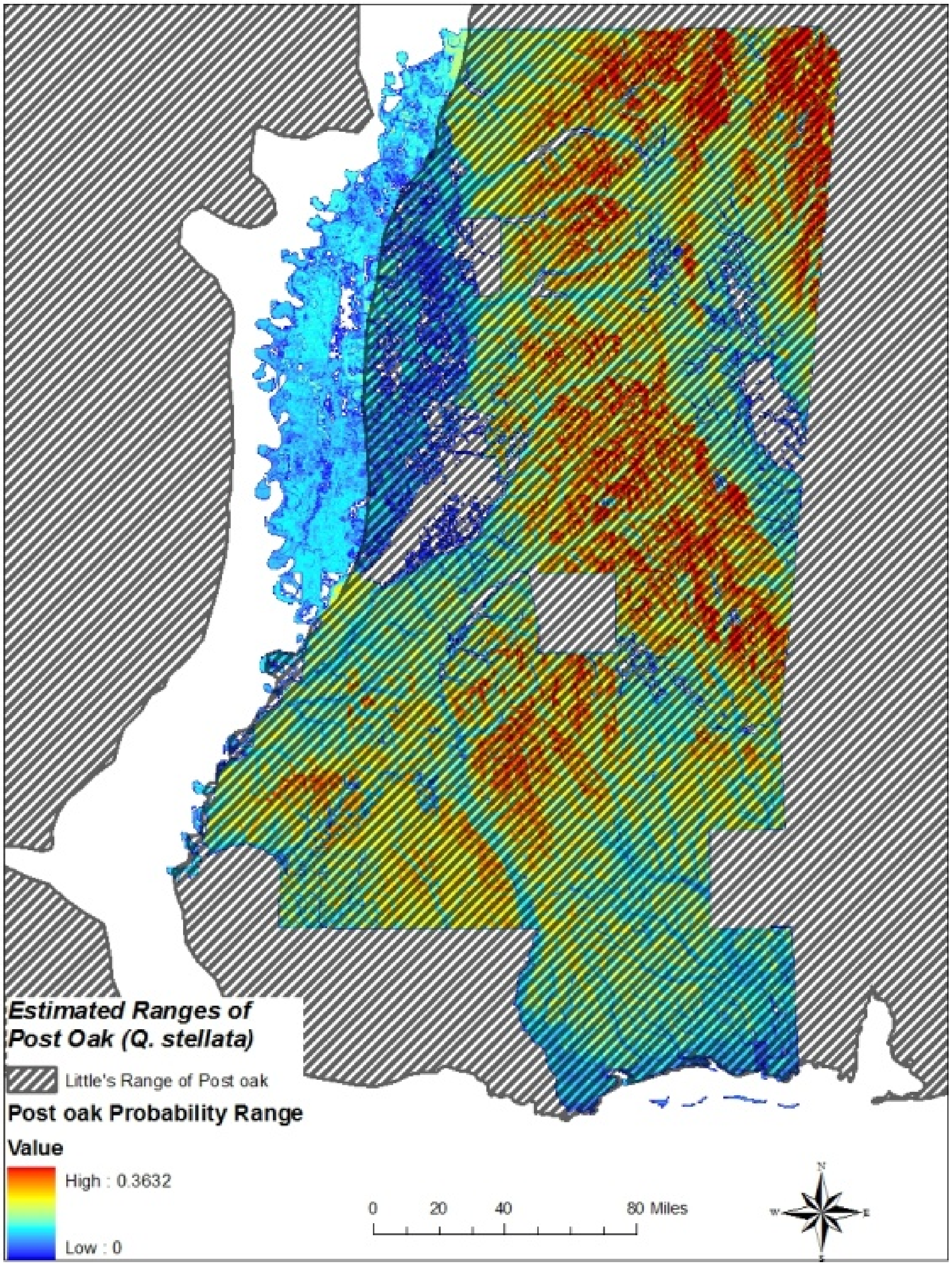

Figure 29 Probability map derived from the significant variables logistic regression parameters showing the predicted range of post oak (*Q. stellata*) in comparison to the Elbert Little's (1971) range map of post oak (*Q. stellata*).

Model Validity

The accuracies of the probability maps as compared to the validation plot set were all low. The lack of confirmed validation sites for species falling in areas of high probability is a cause of some concern. The inability of some of the models to predict sites with high probability is disconcerting due to the fact that all species mapped are known to occur in these areas and would be expected to show higher probability in the prediction of occurrence. Several species probabilities only predicted sites with a highest probability of less than 50%. The causes of these inaccuracies and limitations are likely a direct result of differences between input variable in both spatial resolution and geographic precision. Management regime and species interactions were not taken into consideration for this study due to unavailability in spatial format, but likely also played a role in species composition at individual sites.

The main validation and modeling issues come from the resolution of the environmental data largely at the landscape scale being compared to the species composition plot data. The field plot observations came from areas of 1/5th, 1/10, or 1/20th of an acre. The soils, terrain, and climate data were at much courser resolution. The DEM used to derive the elevation, slope, aspect and flow accumulation was at a 10-meter resolution although it was derived from topographic data including 7.5 minute quadrangle elevations sampled from a 30-by-30-meter grid. This 10-meter resolution was also the resampling resolution for all other raster datasets. The PRISM data started out at resolution of 4km and were resampled to 10-meter resolution. The climate data at a course resolution of 4km were inadequate to describe local, topographically-driven variation in these explanatory variables on a 1/5th acre plot. The best spatial precision for explanatory variables was to infer that the value was valid for 10 –meter from plot center although the actual accuracy of some of these high-resolution data have been subject to also question earlier.

The second cause of data error and lack of validity of our models comes from the plot species data. The MIFI data were collected by numerous people with assumed but not confirmed varying ability in tree identification. It is possible that some of the species were misidentified. This is believed to be the case with some Shumard oak observations. Shumard oak and scarlet oak are morphologically similar in many respects and since scarlet oak occurs at sites higher in topographic position, this could cause the site probability maps to show Shumard at a slightly higher elevation than where it actually occurs. A second example of this comes from a species of oak that this study had hoped to investigate but after further investigation of plot data was unable to model. This study had planned to map the probability range of turkey oak (*Quercus laevis*), but was unable to because some of the MIFI data showed the species occurring in northern Mississippi. this was deemed to be an identification error since this red oak is found near the coast in dry sandy soils (Harlow and Harrar, 1969). It is extremely unlikely that the species was widely transplanted north of its range due to the overall lack of commercial value. It is more likely that southern red oak (somewhat similar) was misidentified as turkey oak in the northern tier of counties. So instead of going to every site and confirming these misinterpretations, the species was eliminated from the study.

The third possibility for low validity in the models is the SSURGO data. The SSURGO data were a generalization of soil characteristics many of which were not confirmed at every site. The SSURGO data were collected by soil scientists that visited sites taking only a few soil cores from the area and interpolating their findings across broad areas based on similarities in topography to the sample sites. Three sets of 10 observational transects were done for each map unit. The confidence level in the preciseness of the mapunits is about 3 acres (NRCS, 2009); because of limitation, it would be difficult to determine accurate site ecology for the species. Many of the soil types were classified as complexes, associations or undifferentiated groups. Complexes and associations are map units that are made up of "two or more dissimilar components that occur in a regularly repeating pattern" (NRCS, 2009). Undifferentiated groups are map units that have components that do not occur together but are similar in use and management. These generalizations made it difficult to get accurate soil data to associate with each oak species since the main component or most commonly occurring component was used for this study. It is possible that a species of oak would be keying in on a more subtle secondary soil characteristics that was not taken into account due to limitations of the soils data in this study. The impreciseness of the SSURGO made it hard to derive precise preferred soils for the species.

Inconsistency in county soil delineations and classifications resulted in too many combinations of soil variables that in some cases could not be reconciled at the desired level of detail. Therefore, components had to be combined and generalized in order to achieve statistical convergence. Black oak is another species of oak that this study had planned to map but was unable to do so due to inaccuracies in the soil data. The black oak model never achieved statistical convergence due to too much variation in the data even after the elimination of horizon 4. Based on these problems, it was felt that this species should be eliminated rather than eliminate an important variable from consideration in all other species models.

CHAPTER V

CONCLUSIONS

Development of probability maps derived from stepwise logistic regression is the first step needed to up-date tree species range maps for commercial and ecologically important tree species. Little's (1971) maps were made in the most accurate and up-to-date way possible during that time period, but with the technical tools currently available, it is believed that more accurate and detailed maps can be produced to depict current tree species distribution. The probability maps and techniques discussed in this study are a positive step to achieve this goal. Such was the case for post oak and white oak, which modeled closely to what is reported about the site preferences of these species. In several cases, the probability maps were, in overall good visual agreement with Little's (1971) range maps. The ranges deviated slightly from one another with the exception of blackjack oak, which was probably a result of problems reported for the soils data. The data used to derive these new ranges have errors and inaccuracies but it are the most complete data that are available. It is felt that anomalies, such as shown in Appendix-Table 16, where percent slope within the bottomland oak group varied from 0 to 100%, could be indicative of problems associated with the unequal size and uniformity of the datasets. However, to completely verify this would entail a reconstruction of the data set and plots. Unfortunately this could not be covered during the scope of this study.

It is believed that with the data that are currently available, these probability maps are the most accurate that can be developed. In order to get more accurate ranges of southern oaks and perhaps other important species, soil characteristics should also be documented from field sites on which species composition is recorded in the future.

REFERENCES

Abrams, M.D. 2003. Where has all the white oak gone? Bioscience 53:927-939.

Baker, J.B. and W. M. Broadfoot. 1977. Site evaluation for eight important southern hardwoods. USDA For. Serv. Gen. Tech. Rep. SO-14.: 31 p.

Baker, J.B. and W. M. Broadfoot. 1979. A Practical field method of site evaluation for commercially important southern hardwoods. USDA For. Serv. Gen. Tech. Rep. SO-26.: 51 p.

Belanger, R.P. 1990. Southern red oak. In: R.M. Burns and B.H. Honkala, eds. Silvics of North America: Volume 2, Hardwoods. Agriculture Handbook 654. USDA For. Serv.: pp 640-644.

Carey, J. H. 1992. Quercus marilandica. In: Fire Effects Information System, [Online]. USDA For. Serv., Rocky Mountain Research Station, Fire Sciences Laboratory (Producer). Available: http://www.fs.fed.us/database/feis/: Accessed Jan. 25, 2009.

Cassens, D.L. 2007. Hardwood lumber and veneer series: White oak. Purdue University. www.ces.purdue.edu/extmedia/FNR/FNR-292-W.pdf: Accessed 2009

Daly, C., G. Taylor, and W. Gibson. 1997. The PRISM Approach to Mapping Precipitation and Temperature, 10th Conf. on Applied Climatology, Reno, NV, Amer. Meteor. Soc.: pp 10-12.

Daly, C., G. Taylor, and W. Gibson, 2002a. 103-Year High-Resolution Precipitation Climate Data Set for the Conterminous United States. Corvallis, Oregon. The PRISM Climate Group, Oregon

Daly, C., W. Gibson, G. Taylor, G. Johnson, and P. Pasteris. 2002b. A knowledge-based approach to the statistical mapping of climate. Climate Research 22:99–113.

Edwards, M.B. 1990. Shumard Oak. In: R.M. Burns and B.H. Honkala, eds. Silvics of North America: Volume 2, Hardwoods. Agriculture Handbook 654. USDA For. Serv.: 734-738.

Felicisimo, A.M., E. Frances, J.M. Fernandez, A. Gonzalez-Diez and J. Varas. 2002. Modeling the potential distribution of forests with a GIS. Photogrammetric Engineering and Remote Sensing 68: 455-461.

Filer, T.H. Jr. 1990. Nuttall Oak. In: R.M. Burns and B.H. Honkala, eds. Silvics of North America: Volume 2, Hardwoods. Agriculture Handbook 654. Washington, D.C.: USDA For. Serv.: pp 704-708.

Hall, B, G., D.R. Motzkin, M. Foster, Syfert, and J.Burk. 2002. Three hundred years of forest and land-use change in Massachusetts, USA. Journal of Biogeography 29: pp 1319-1335.

Harlow, W. M. and E. S. Harrar. 1969. Textbook of Dendrology, Covering the Important Forest Trees of the United States and Canada. New York, McGraw-Hill.: 329-330

He, H.S., D.C. Dey, X. Fan, M.B. Hooten, J. M. Kabrick, C. K. Wikle, and Z. Fan. 2006. Mapping pre-European settlement vegetation at fine resolutions using a hierarchical Bayesian model and GIS. Plant Ecology 191: pp 85-94.

Hodgkins, E. J., M. S. Golden, and W. F. Miller: 1979, Forest Habitat Regions and Types on a Photomorphic-Physiographic Basis: A Guide to Forest Site Classification in Alabama-Mississippi, Southern Coop. Series 210, Alabama Agriculture Experiment Station, Auburn University, Auburn, AL, U.S.A., 63 .

Iverson, L. and A. Prasad. 2001. Using FIA data to assess current and potential future tree species importance values in the eastern United States. Proceedings of the Third Annual Forest Inventory and Analysis Symposium. 2001 October 17-18; Traverse City, MI., Delaware, OH., USDA For. Serv.: 171-172.

Iverson, L and A. Prasad. 1998. Predicting abundance of 80 tree species following climate change in the eastern United States. Ecol. Monogr. 68. 465–485.

Johnson, P.S., S.R. Shifley, and R. Rogers, 2002. Ecology and Silviculture of Oaks. CABI Publishing, Wallingford. 8-53

Kennedy, H.E., JR., and R.M. Krinard. 1985. Shumard oaks successfully planted on high pH soils. US For. Serv. Res. Note SO-321. 3 p.

King, G., and L. Zeng. 2001. "Logistic Regression in Rare Events Data." Political Analysis 9 (2): 137-63

Krinard, R.M. 1990. Cherrybark Oak. In: R.M. Burns and B.H. Honkala, eds. Silvics of North America: Volume 2, Hardwoods. Agriculture Handbook 654.USDA For. Serv.: 644-649.

Kutner, M.H, C.J. Nachtsheim, and J. Neter. 2004 Applied Linear Regression Methods (4th ed.). Chicago, IL: McGraw-Hill/Irwin: 604-608

Lister, A., R. Riemann, and M. Hoppus. 2000. Use of regression and geostatistical techniques to predict tree species distributions at regional scales. 4th International Conference on Integrating GIS and Environmental Modeling (GIS/EM4): Problems, Prospects and Research Needs. 2000 September 2-8; Banff, Alberta, Canada: University of Colorado: 10 p.

Little, E.L. 1971. Atlas of United States Trees: volume 1. conifers and important hardwoods. USDA For. Serv. Misc. Publ. 1146 .1-7.

Merwade, Venkatesh.2008. Downloading SSURGO Soil Data from Internet. http://web.ics.purdue.edu/~vmerwade/education/download_ssurgo.pdf

NRCS. 2009. Soil Survey Staff, Natural Resources Conservation Service, United States Department of Agriculture. Soil Survey Geographic (SSURGO) Database for Mississippi. Available online at http://soildatamart.nrcs.usda.gov. Accessed 2009-2010.

Ohmann, J.L. and M.J. Gregory. 2002. Predictive mapping of forest composition and structure with direct gradient analysis and nearest-neighbor imputation in coastal Oregon, USA. Canadian Journal of Forest Research 32:725-741.

Rodgers, R.. 1990. White Oak. In: R.M. Burns and B.H. Honkala, eds. Silvics of North America: Volume 2, Hardwoods. Agriculture Handbook 654. USDA For. Serv.: 605-613.

Schlaegel, B.E., 1990. Willow oak. In: R.M. Burns and B.H. Honkala (Coordinators), Silvics of North America: 2. Hardwoods. Agricultural Handbook 654. USDA For. Serv.: 715-720.

Shostak, D. I., E. F. Loewenstein, and M. R. Dubois. 2004. Topographic position and site index: an oak regeneration relationship. USDA For. Serv. Gen. Tech. Rep. SRS-71.: 243-248.

Smith, W.B., P.D. Miles, J.S. Vissage, and S.A. Pugh, 2004. Forest Resources of the United States, 2002. USDA For. Serv. Gen. Tech. Rep. NC-241

Stransky, J.J. 1990. Post Oak. In: R.M. Burns and B.H. Honkala, eds. Silvics of North America: Volume 2, Hardwoods. Agriculture Handbook 654. USDA For. Serv.: 738-743.

Tucker, W. 2010. Mississippi Institute of Forest Inventory (MIFI). http://www.mifi.ms.gov/. Accessed 2009-2010.

Vozzo, J.A. 1990. Water oak. In: Burns, R.M. and Honkala, B.H. (Eds) Silvics of North America. USDA For. Serv. Agriculture Handbook 654. 701–705

APPENDIX A

RECLASSIFIACTIONS AND ITERATION TABLES

Table 13 SSURGO soil texture classes and the reclassification and regrouping of texture classes for oak sites suitability models analysis.

Textures	SSURGO Abbreviation of textures	Reclass texture group	Numerical classification
stratified loamy sand to sand	SR LS S	Sandy	1
stratified sand to loamy sand	SR S LS		
sand	S		
loamy very fine sand	LVFS		
loamy sand	LS		
loamy fine sand	LFS		
loamy course sand	LCOS		
course sandy loam	COSL		
course sand	COS		
stratified to fine sandy loam to weathered bedrock	SR FSL WB		
stratified fine sandy loam to loamy fine sand	SR FSL LFS		
stratified sandy loam to fine sandy loam	SR SL FSL		
stratified sand to fine sandy loam	SR S FSL		
stratified loamy sand to fine sandy loam	SR LS FSL		
stratified loamy fine sand to very fine sandy loam	SR LFS VFSL		
stratified loamy fine sand to fine sandy loam	SR LFS FSL		
stratified coarse sand to sandy loam	SR COS SL		
silty loam	SIL		
fine sandy loam	FSL		
fine sand	FS		
very fine sandy loam	VFSL		
stratified loamy sand to sandy clay loam	SR LS SCL		
mucky sand	MKS		

Table 13 cont.

stratified to loam to weathered bedrock	SR L WB	Silty	2
silt	SI		
loam, variable	L VAR		
loam	L		
stratified very fine sandy loam to silt loam to silty clay loam	SR VFSL SIL SICL		
stratified very fine sandy loam to silty clay loam	SR VFSL SICL		
stratified sandy loam to silty clay loam	SR SL SICL		
stratified silt loam to silty clay loam	SR SIL SICL		
stratified sandy loam to silty clay loam	SR SL SICL		
silty clay loam, variable	SICL VAR		
silty clay loam	SICL		
stratified muck to loam	SR MUCK L		
mucky silt loam	MKSIL		
silty clay loam, silty clay, clay	SICL SIC C	Clay	3
sandy clay loam, variable	SCL VAR		
sandy clay loam	SCL		
sandy clay	SC		
clay loam, variable	CL VAR		
clay loam	CL		
stratified very fine sandy loam to silty clay	SR VFSL SIC		
stratified sand to clay	SR S C		
stratified fine sandy loam to clay	SR FSL C		
stratified sandy loam to clay	SR SL C		
silty clay, clay	SIC C		
silty clay	SIC		
clay, variable	C VAR		
clay	C		
mucky clay	MKC		
muck	MUCK		

Table 13 cont.

very gravelly sandy loam	GRVSL		
very gravelly sandy clay loam	GRV SCL		
gravelly sandy clay loam	GRV SCL		
gravelly fine sandy loam	GRV FSL		
angular cobbly fine sandy loam	CBA FSL		
weathered bedrock, fine sandy loam	WB FSL	Rock	4
weathered bedrock	WB		
unweathered bedrock	UWB		
stratified weathered bedrock to loam	SR WB L		
stratified weathered bedrock to fine sandy loam	SR WB FSL		
sr to weathered bedrock	SRWBFSL		

Table 14 SSURGO drainage classes and the reclassification and regrouping of drainage classes for oak sites suitability models analysis.

SSURGO drainage classes	Numerical reclassification groupings	Generalized groupings
excessively drained somewhat excessively drained	1	excessively drained
moderately well drained	2	moderately well drained
well drained	3	well drained
somewhat poorly drained poorly drained	4	poorly drained
very poorly drained	5	excessively poorly drained
other water	0	eliminated from study

Table 15 Aspect output and the reclassification of aspect for oak site suitability models analysis.

Aspect	Rasterized groupings	Reclassification grouping	Aspect generalized	Rasterized groupings
Flat	1	0	Eliminated from study	
North	022.5	1	North	022.5
Northeast	22.567.5			
East	67.5112.5	2	East	22.5157.5
Southeast	112.5157.5			
South	157.5202.5	3	South	157.5202.5
Southwest	202.5247.5			
West	247.5292.5	4	West	202.5337.5
Northwest	292.5337.5			
North	337.5360	1	North	337.5360

Table 16 The percentage of times in the test regression iterations that each variable proved to be significant for each species out of 1000 stepwise logistic regression.

Oak species	temperature (F)	elevation (ft)	percent slope	precipitation (in)	flow accumulation	H1 texture class	H2 texture class	H3 texture class	drainage class
	_	_	_	_	%	_	_	_	_
white	68	86	100	100	47	100	82	60	16
southern red	5	97	100	84	27	80	41	35	58
Nuttall	20	9	10	31	1	99	0	8	1
blackjack	99	99	96	85	9	11	82	6	4
cherrybark	37	99	0	100	23	99	7	5	38
willow\water	100	0	100	81	100	4	1	3	23
Shumard	33	2	56	53	3	4	1	7	7
post	2	99	30	5	99	54	11	11	22