2017

Habitat Modeling and Vegetation Mapping of Mammoth Cave National Park Using LiDAR data and Multispectral Imagery

Hongli Yang

HABITAT MODELING AND VEGETATION MAPPING OF MAMMOTH CAVE NATIONAL PARK USING LIDAR DATA AND MULTISPECTRAL IMAGERY

A Thesis

Presented to

The Faculty of the Department of Geosciences

Murray State University

Murray, Kentucky

In Partial Fulfillment

Of the Requirements for the Degree

Of Master of Science

By

Hongli Yang

2017

VEGETATION MAPPING AND HABITAT MODELING OF MAMMOTH CAVE NATIONAL PARK USING MULTISPECTRAL IMAGERY AND LIDAR DATA

DATE APPROVED: _____

_____

Thesis Advisor

_____

Member, Thesis Committee

_____

Member, Thesis Committee

_____

Member, Thesis Committee

_____

Collegiate Graduate Coordinator

_____

Dean of the College

_____

University Graduate Studies Coordinator

_____

Provost

ACKNOWLEDGEMENT

First of all, I would like to express my heartfelt gratitude to my thesis committee members, Professors Robin Q. Zhang, Bassil El Masri, Kate S. He and Jane Benson. I am really grateful for their valuable suggestions and advice for successful completion of this research.

I would like to render a very special thanks to my thesis Advisor, Dr. Robin. Q. Zhang for her relentless support and directions in terms of academics and future career throughout the duration of my study at Murray State University. Thank you so much for valuable comments, contributions, suggestions, patience, enthusiasm, and immense knowledge. Her guidance helped me in all the time of research and writing of this thesis. I could not have imagined having a better advisor and mentor for my master study.

My sincere thanks also go to staff from Mammoth Cave National Park, Rick Olson, Rick Toomey, Lillian Scoggins, Shannon Trimboli, and Bobby Carson. Thank you for providing the data and host us during our field trips to the Park. Special thanks go to Rick Olson for his extensive knowledge about the Park and being the expert and guide in the field. His assistance and patience played an important role in my work.

Deep appreciation goes to Dr. George Kipphut, previous Chair of the Department of Geosciences and Dr. Haluk Cetin, Graduate Coordinate, for their constant encouragement and assistance for me to accomplish my goal. Also, I would like to express sincere appreciation to all my friends and faculty members in the Department of Geosciences.

Last but not the least, I would like to thank my family: my parents, for their unconditional love and support throughout my life.

ABSTRACT

An up-to-date and detailed vegetation map provides critical information for habitat management. In addition, a habitat model is necessary for Park's Fire Management, for classification of fuel types, and for delineation of fire management units. Several attempts to map the vegetation at the Mammoth Cave National Park were conducted in 1934, 1975, 1997 and 2011. This essential goal of this study was to produce a new vegetation habitat model and update the vegetation map for the Park. Landsat-8 Operational Land Imager (OLI) imagery, LiDAR and bedrock dataset were used for habitat model configuration and vegetation mapping at Mammoth Cave National Park.

Vegetation habitat types were determined by a combination of slope, aspect and bedrock. The habitat model indicated that Acid and Calcareous were the two dominant habitats within the park, accounting for 46.24% and 49.74% of the total park area respectively. Among the ten habitat types, Acid Mesic and Calcareous Sub-Mesic occupied the largest areas, which accounted for 29.26% and 21.03% respectively. The habitat was observed and described at 29 ground reference sites due to limited accessibility. The habitat types of 22 sites (76%) predicted by the model were consistent with field observations. The discrepancy between model result and field observation at three sites was likely due to previous human disturbance. And the model needs further improvement to accurately predict Acid Xeric habitat locations.

Principal Component Analysis (PCA), enhanced vegetation index (EVI), and unsupervised classification, were applied to map the vegetation types. Five classes were mapped: barren land/ man-made structure, evergreen, deciduous, mixed forests and water. In the resultant map, deciduous trees accounted for the largest area in the park and most of the evergreen and

mixed trees were found in the southern part of the park. The classification results were evaluated by 398 deciduous, 76 evergreen and 65 mixed field plots data. The overall accuracy of PCA technique and EVI Index was 85%, 7% higher than using PCA technique alone and 13% higher than NLCD 2011.

The influence of historic disturbance which occurred before the establishment of the Park can still be seen today. Approximately 70% of the evergreen forests, dominated by eastern red cedar (Juniperus virginiana) are found in previously cropland and pasture fields. They are the first successional forest in the area. While 40% of coniferous trees are currently in Xeric or Sub-Xeric habitat types with favorable conditions to support coniferous species, the remaining 60% will likely to be replaced by deciduous trees in the future.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

CHAPTER ONE

DIMENSIONS OF VEGETATION HABITAT MODELING AND VEGETATION MAPPING

1.0 INTRODUCTION

Vegetation has considerable impacts on almost all land surface energy exchange processes, acting as an interface between land and atmosphere. It affects local and regional climate (Arora, 2002; Douville et al., 2000) and hydrologic balance of the land surface (Eugster et al., 2000). The dynamics of vegetation are of primary importance in global terrestrial ecosystem change (Suzuki et al., 2007 and Kelly et al., 2011). Vegetation not only forms essential habitats for plant and animal species but is also a prerequisite for ecosystem function. Vegetation provides many ecosystem services, principally through the protection of the land surface, the amelioration or modification of the local climate, the maintenance of critical ecosystem processes, and the conservation of biodiversity (Hölzel et al., 2012). Vegetation types represent different stages in vegetation restoration and succession and are closely related to soil properties, water runoff, soil erosion, as well as ecological stability (Jiao et al., 2008; Nagase and Dunnett, 2012; Qiu et al., 2010; Wang et al., 2011). Therefore, it is important to accurately distinguish vegetation types in ecological studies.

Aspect and slope are alternatives for the spatial and temporal distribution of factors such as solar radiation, moisture and temperature that affect species composition and productivity (Stage and Salas, 2007). Differences in insolation period and intensity change with aspect, thereby forming a range of microclimates in multifaceted landscapes (Holland and Steyn, 1975). In general, aspect can have important influences on climate as well as the distribution of

vegetation types. In the northern hemisphere, the north side of slopes often have more shaded area than the south side, which receives less solar radiation. South-facing slopes tend to be more xeric (dry) due to high levels of evapotranspiration than a north-facing slope. For example, Pinder et al. (1997) have found that shrub communities were more abundant on south-facing slopes in Larsen Volcanic National Park in northern California. Also, Desta et al. (2004) found that deciduous forest on the north and east aspects were 27–50% more productive than the west and southwest aspects in an Appalachian watershed. This suggests that variations in aspect may have a great influence on the floristic and life-form composition of the vegetation (Armesto and Martinez, 1978). The steepness of a slope can also affect the growth of plants because it affects the amount of solar radiation received. In addition, the gradient of slope influences the availability of water to the vegetation. The steeper the slope, the more likely that rain will run off rather than infiltrate. Therefore, steep slope tends to hold less water and the soil will be more xeric. Airborne Light Detection and Ranging (LiDAR) offers detailed and precise elevation measurements that can be used to calculate aspect and slope measurements of much higher resolution and accuracy than traditional measurements based on topographic maps.

Traditionally, vegetation maps were completed by using topographic maps, panchromatic aerial photography and field sampling. These methods, however, are ineffective to acquire vegetation coverage for a large area because they are time-consuming and often too costly. Remote sensing techniques provide a more practical, rapid and economic means to study vegetation habitats, especially over large areas. Landsat data provide a wide area coverage and medium spatial resolution to monitor vegetation changes. The frequent revisit makes it possible for the satellites to collect data for the study of landscape dynamics and monitoring of forest habitats in time and space.

1.1 Purpose of the study

There have been several attempts to map the vegetation and habitats at the Mammoth Cave National Park over the last decades with limited success due to difficulty of access to some parts of the park and lack of high temporal and spatial resolution remotely sensed data. In 2010, the Park acquired high resolution LiDAR data for detailed mapping. The Park is now in the process of updating their Fire Management Plan that calls for an updated habitat map to facilitate the designation of fuel types. The specific objectives of this research are: (1) to develop a predictive habitat model based on a limited number of physical attributes (bedrock, slope, and aspect), (2) to produce an up-to-date vegetation map, (3) to evaluate the accuracy of the habitat model and vegetation map, and (4) to analyze the distribution of habitat and vegetation with reference to each other and the land use history of the Park. Data collected by Landsat-8 Operational Land Imager (OLI) is used to categorize vegetation types according to their reflectance characteristics. Digital Elevation Model (DEM) derived from LiDAR data is used to calculate and categorize slope and aspect throughout the Park. The resultant vegetation and habitat maps are valuable for fire management and wildlife habitat and biodiversity conservation analysis.

1.2 Organization of the research

The thesis is organized into six chapters. The first chapter introduces the theme of research, including purpose of the study. Chapter Two contains literature review pertaining to this research. The study area of this research, including its physiography, climate, and geology, are discussed in Chapter Three. The following chapter provides a detailed description of the

methodology including data sources and types, habitat modeling, vegetation classification, and accuracy assessment. The results are presented in Chapter Five. Finally, Chapter Six presents the conclusion of the study and provides some recommendations.

CHAPTER TWO

VEGETATION HABITAT MODELING AND VEGETATION MAPPING – AN OVERVIEW

2.0 Introduction

It is an important task to classify and map vegetation for managing natural resources as vegetation provides a base for all living beings and plays an essential role in affecting global climate change, such as influencing terrestrial CO_2 (Xiao et al., 2004). Vegetation mapping also provides valuable information for understanding the natural and man-made environments through quantifying vegetation cover from local to global scales at a given time period or over a continuous period (Xie et al., 2008). It is critical to obtain current states of vegetation cover in order to initiate vegetation protection and restoration programs (Egbert et al., 2002; He et al., 2005).

Habitat type mapping is not only applicable in forestry purposes but also valuable in conservation (Räsänen et al., 2014). In forestry, habitat type maps and other thematic maps are used for strategic analysis in forest management planning (Tomppo et al., 2008). In conservation perspective, habitat type maps can be used in mapping biodiversity patterns (e.g. Kerr and Ostrovsky, 2003; Turner et al., 2003). Habitat types can also provide information for classification of fuel types. Fuel characterization is key to wildfire prevention as forest fuel is one of the primary factors affecting wildfire risk and behavior (Marino et al., 2016).

2.1 Vegetation habitat and vegetation type mapping using remote sensing and LiDAR

Traditional methods for vegetation analysis, such as field surveys, map interpretation and ancillary data analysis are often ineffective because they are time consuming, expensive, and

often provide date lagged information (Xie et al., 2008). Remote sensing data brings another perspective to vegetation studies, because it provides possibilities of detecting the patterns at different spatial scales, which may not be feasible through field observations. Therefore, this technique helps in the characterization of ecosystems at various spatial extents. Apart from this, remote sensing data archives have great potential for facilitating systematic temporal analysis at various scales from recent past to several decades back (Xie et al., 2008). Lefsky et al. (2002) stated that analogue aerial photography has been the oldest, most frequently used and best understood form of remote sensing. High-resolution historic aerial photographs dating back to 1930s are effective in the mapping of small ecosystems, fine-scale landscape features and successional pathways in some cases (Green and Hartley, 2000; Morgan et al., 2010). Automated digital image analysis techniques provide a time-saving solution and eliminate the influence of the interpreter's subjectivity in vegetation delineation. The optimal approach depends primarily on the definition of the output products (e.g., the type of the maps) and is influenced by spatial resolution and inter-pixel variance (Wulder et al., 2004).

The application of remote sensing techniques in forest mapping is possible because of the high reflectance values from forested areas in the near-infrared, moderate reflectance in the middle-infrared and low reflectance in the red spectral regions compared with non-forested areas (Xiao and McPherson, 2005). Vegetation index (VI), defined as the arithmetic combination of the near-infrared and red bands related to the spectral characteristics of vegetation, has been widely used for phenologic monitoring and biophysical derivation of radiometric and structural vegetation parameters (Huete and Justice, 1999). Normalized Difference Vegetation Index (NDVI) is the most commonly used vegetation index, which can cancel out a large proportion of the noise caused by changing sun angles, topography, clouds, or shadow, and atmosphere (Huete

and Justice, 1999). Enhanced Vegetation Index (EVI) has been considered as modified NDVI with improved biomass regions and improved vegetation monitoring capability (Huete and Justice, 1999). Compared to NDVI, EVI is more responsive to canopy variation, canopy type and plant physiognomy.

To extract land cover information from remotely sensed imagery, image classification methods are usually applied. Cluster analysis is commonly used for unsupervised classification (Choi et al., 2004; Mukherjee and Lal 2014; Abdi and Williams 2010; Lu et al., 2014). The goal of cluster analysis is to assign observations to groups ("clusters") where observations within each group are similar to one another with respect to variables or attributes of interest, and the groups themselves stand apart from one another (Beauchaine and Beauchaine, 2002). Although these clusters are not always equivalent to actual classes of land cover, this method can be used without having prior knowledge of the ground cover in the study site (Nie et al.2001). Traditional unsupervised classification algorithms, such as k-means (Duda et al. 2001) and the Iterative Self-organizing Data Analysis Techniques Algorithm (ISODATA) (Ball and Hall 1965), use iterative calculations to find an optimum set of decision boundaries for clustering. The ISODATA is a more sophisticated version of k-means, which allows classes to be split and merged (Zhong et al. 2011).

Principal Component Analysis (PCA) is a mathematical procedure that uses orthogonal transformation to convert a set of observations of possibly correlated variables into a set of values of linearly uncorrelated variables called principal components (Abdi and Williams, 2010; Jolliffe, 2005). PCA transforms a number of detected variables into a smaller number of principal components (PCs) (Guo et al. 2015). This transformation is defined in such a way that the first PC has the largest possible variance, and each succeeding component in turn has the

highest variance possible under the constraint that it be orthogonal to (i.e., uncorrelated with) the preceding components (Mari et al. 2012). PCA is known to be effective to reduce the dimensionality in Landsat imagery (Kwarteng and Chavez, 1989) because several Landsat bands are highly correlated.

Light Detection and Ranging (LiDAR) systems uses artificial laser light to measure distances to the earth. LiDAR offers detailed and accurate elevation measurements that can provide high spatial resolution digital elevation models (DEMs), and slope and aspect (Gould et al., 2013). Small-footprint, discrete-return system have the ability to penetrate surface vegetation and yield multiple returns from canopy and underlying terrain (Gould et al., 2013). The dense collection of elevation data makes LiDAR an attractive data source for the production of high-resolution digital elevation models (DEMs) used in many GIS application (Gould et al., 2013). Research has been conducted to explore the effect of DEM resolution on terrain representations, such as slope and aspect mapping (Chow and Hodgson. 2009). The study of Chang and Tsai (1991) showed that both the accuracy of slope and aspect decreased with coarser DEM resolution.

2.2 Vegetation habitat and vegetation mapping in Mammoth Cave National Park

The first vegetation field mapping of the Park was completed by Ivan Ellsworth (1934), and included 47,348 acres within the minimum proposed boundary. The resultant map contained seventeen categories of four major forest types plus two categories of cleared areas (Ellsworth 1936). Much of these disturbed areas, now in the middle of succession, can still be recognized today. Though aerial photographs were available at the time, they were not used in Ellsworth study for vegetation mapping which was accomplished by field teams armed with topographic

maps. The next vegetation map was produced by Franklyn Hambly (1966) as part of the Park's Resource Management Plan. The map featured five forest cover types, and park biota were assigned to one or more of the cover types. No details of mapping methods were provided. The third vegetation map of the park was produced by Adolf Faller and Marion Jackson (1975). They used 1966 panchromatic aerial photography (1:20,000) to establish individual mapping units. Seven major vegetation types were established with 18 "subunits characteristic of successional types on particular landforms" (Faller and Jackson, 1975). It was a thorough study, but the map sections had boundary discrepancies, which prevented assembly into one coherent map. The fourth vegetation map of the park was prepared by Kemuel Badger (Badger et al., 1997), and employed a stratified random field sampling approach within a variety of habitat types. Eight community types were delineated with GIS-based modeling, but significant discrepancies between field data and community types on the map persisted (Olson et al. 2000). The latest vegetation map was produced in 2011 by Olson et al. (2013). They used 2008 Landfire map to designated 24 sub-categories and grouped them into four vegetation categories. Barrens and Prairie Plantation categories were added as superimposed polygons, and the same approach was taken for both fire and storm-linked forest canopy gaps (Olson et al., 2013). Accuracy assessment data points were selected randomly and the cumulative average accuracy for this map was 66%, which was below the acceptable 80%. Meanwhile, it was identified then that another vegetation mapping effort based on Landsat data would be necessary to compare to the mapping result of 1997 (Olson et al., 2013).

CHAPTER THREE

STUDY AREA AND DATA AQUISISTION

3.1 Location

Mammoth Cave National Park (Figure 1) is located at latitude 37.2° north and longitude 86.1° west in the south-central Kentucky. Mammoth Cave National Park has the world's largest network of natural caves and underground passageways, which are characteristics of limestone formations. Established in 1941, Mammoth Cave National Park is also a World Heritage Site. The park and its underground network of more than 560 surveyed-km of passageways are home to a varied flora and fauna, including a number of endangered species. The park's 52,830 acres (21,380 ha) are located primarily in Edmonson County, Kentucky, with small areas extending eastward into Hart County and Barren County.

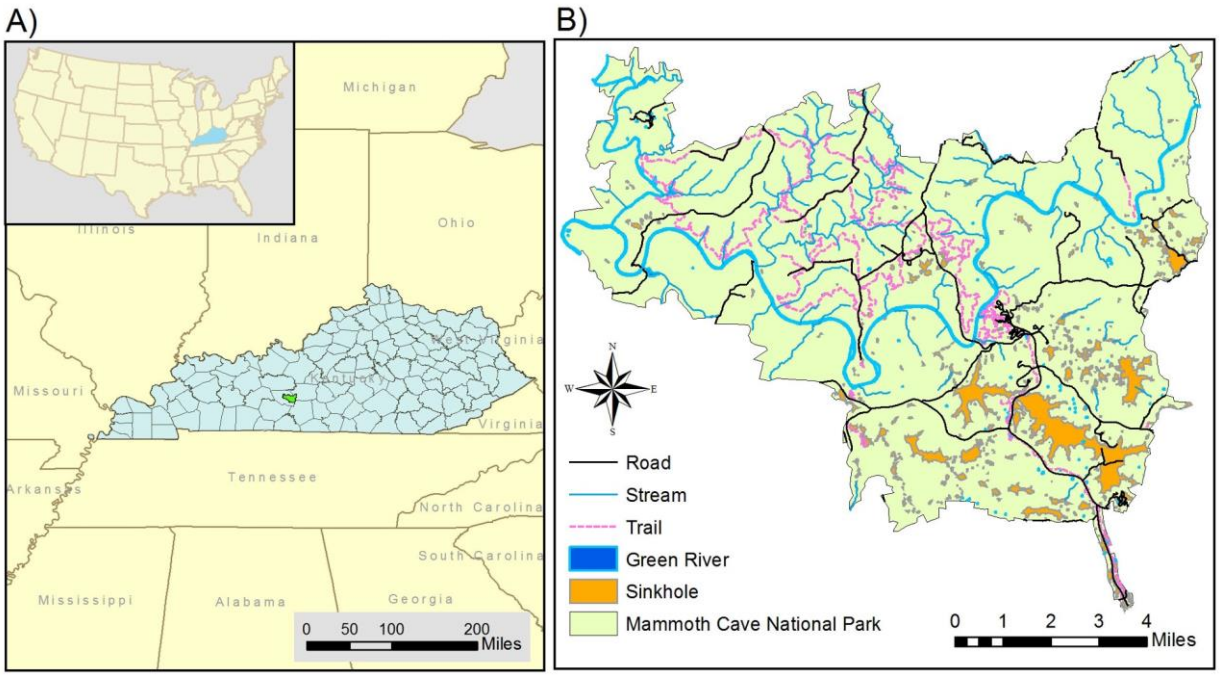

Fig.1: A) The geographic location of the Mammoth Cave National Park

B) The geographic features of Mammoth Cave National Park

3.2 Physiography

Mammoth Cave National Park lies in the South-Central Kentucky karst, which is a crossroad of carbonate bedrock stretching north to Indiana, east to the Cumberland Plateau, south to Georgia and west to the Ozarks. The park is bisected east to west by the Green River, which defines the hydrologic base-level and divides the region into two distinct physiographic areas. North of the river is an alternating series of limestone and insoluble rocks are exposed with the main limestone strata accessible only near the river and in the bottom of a few deeply incised valleys (National Park Service, 2015).

3.3 Climate

Kentucky has a moderate climate, characterized by warm, yet moist conditions. Summers can average in the mid-90s (F) (32.22°C), while winters average in the low 40s (F) (9.44°C). Much of the park's average annual 52 inches (132cm) of precipitation falls in the spring. Storms occur year-round, though most occur March-September. Year-round, the cave temperature in interior passages fluctuates from around 54º (F) (12.2°C) to 60º (F) (15.5°C). Winter temperatures, however, can be below freezing at the cave entrances.

3.4 Geology

The Mammoth Cave National Park is part of South-Central Kentucky Karst, which is characterized by subterranean drainage to springs on major rivers. From the southeast to the northwest portion of the landscape, there is a gradient of decreasing maturity in karst development, which corresponds to the regional dip of the bedrock.

For a given climate, bedrock largely determines soil types, and whether surface or subsurface (karst) drainage prevail. Due to the tendency for subsurface drainage to develop in

calcareous bedrock such as limestone, these sites will be more xeric than areas underlain by sandstone or shale. The magnitude of this general difference appears to be minimized on the steepest slopes due to rapid surface drainage.

3.5 Data Sources

To effectively distinguish evergreen and deciduous trees, one leaf-off and one leaf-on Landsat-8 OLI image were obtained from U.S Geological Survey (USGS) Earth Explorer website. The images were acquired on January 2nd and June 10th, 2016. Landsat 8 images consist of nine spectral bands and all seven reflective bands were used. The spatial resolution of the images is 30 meters. Figure 2 shows January 2nd imagery in Band 6, 5, 4 as RGB. Figure 3 shows June 10th imagery in Band 6, 5, 4 as RGB.

Airborne Light Detection and Ranging (LiDAR) dataset was provided by the Mammoth Cave National Park. The LiDAR dataset was acquired between Oct 13th and 17th, 2010, which was used to produce a digital elevation model (DEM) of 1 m pixel resolution (Figure 4).

The Park provided the bedrock dataset. The bedrock categories are Alluvium, Calcareous, Calcareous caprock, and sandstone/shale. “Alluvium” referrs to river lain sediment. “Calcareous” refers to carbonate bedrock, which results in more alkaline soil. Non-carbonate sandstone and shale bedrock conditions would result in more acidic soil. Figure 5 shows the cross section of the cave and its relationship to the surroundings and geology. Massive cross-bedded sandstone and shale beds cross on top of older limestone and as erosion continues, these sandstone and shale form protective caprock for the cave system in the Mammoth Cave region. Subsurface drainage on limestone tends to be more xeric than on sandstone as most of the water flows into the sinkhole on limestone while more water is retained in soil on sandstone.

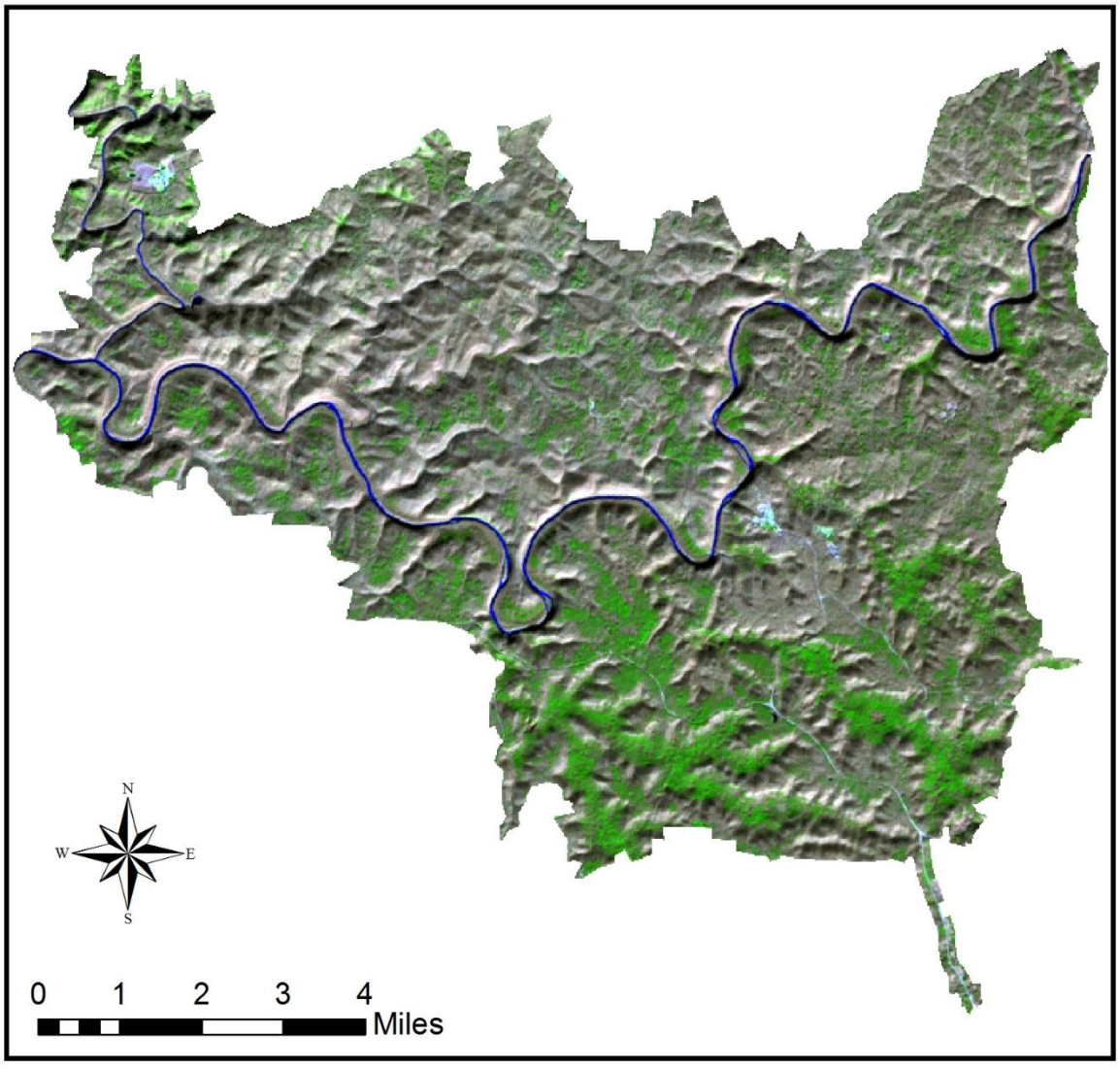

Fig.2: False color composite of Landsat-8 OLI image of Mammoth National Park, January 2nd, 2016 (Band 6, 5, 4 as RGB)

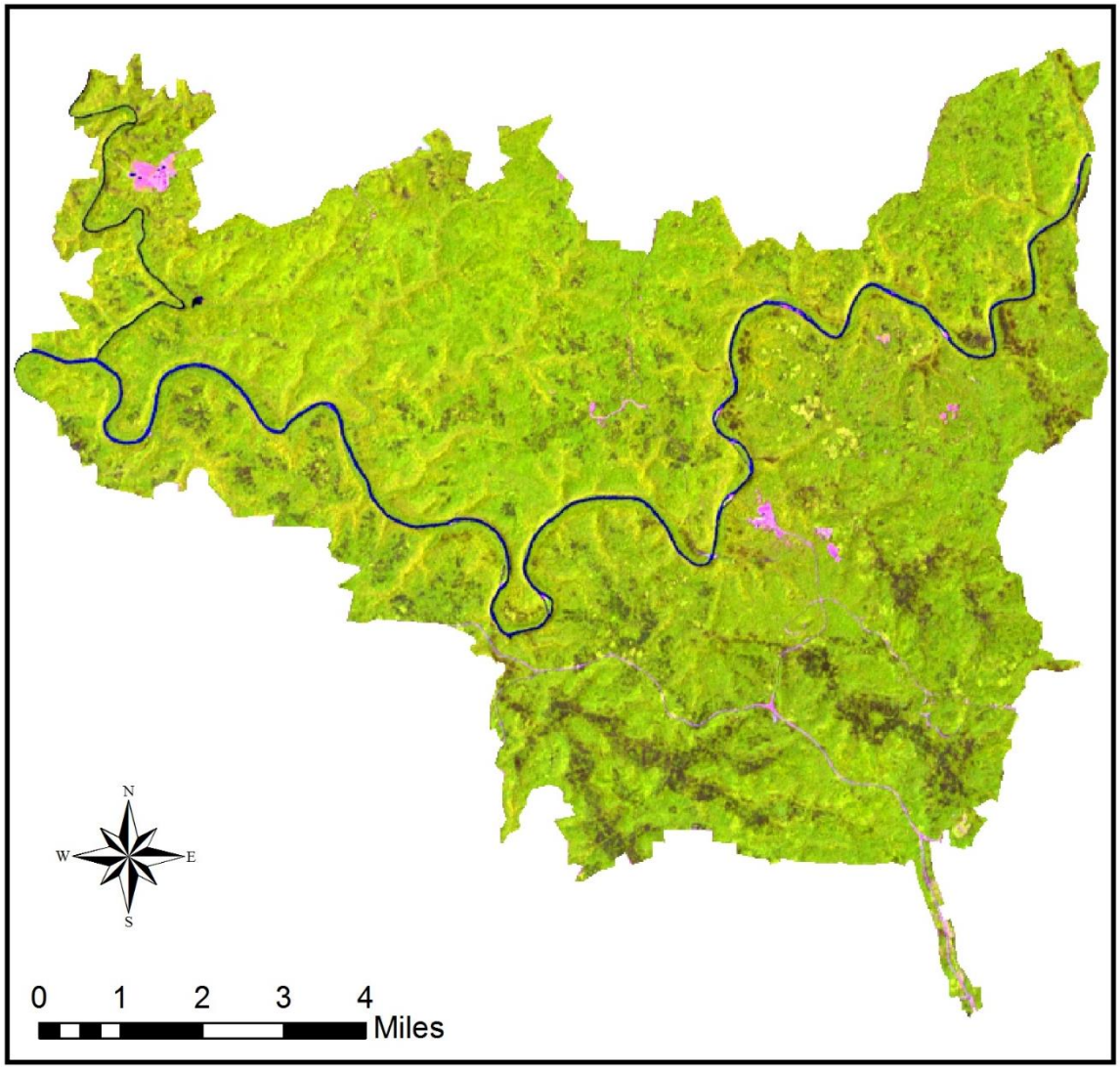

Fig.3: False color composite of Landsat-8 OLI image of Mammoth National Park, June 10th, 2016 (Band 6, 5, 4 as RGB)

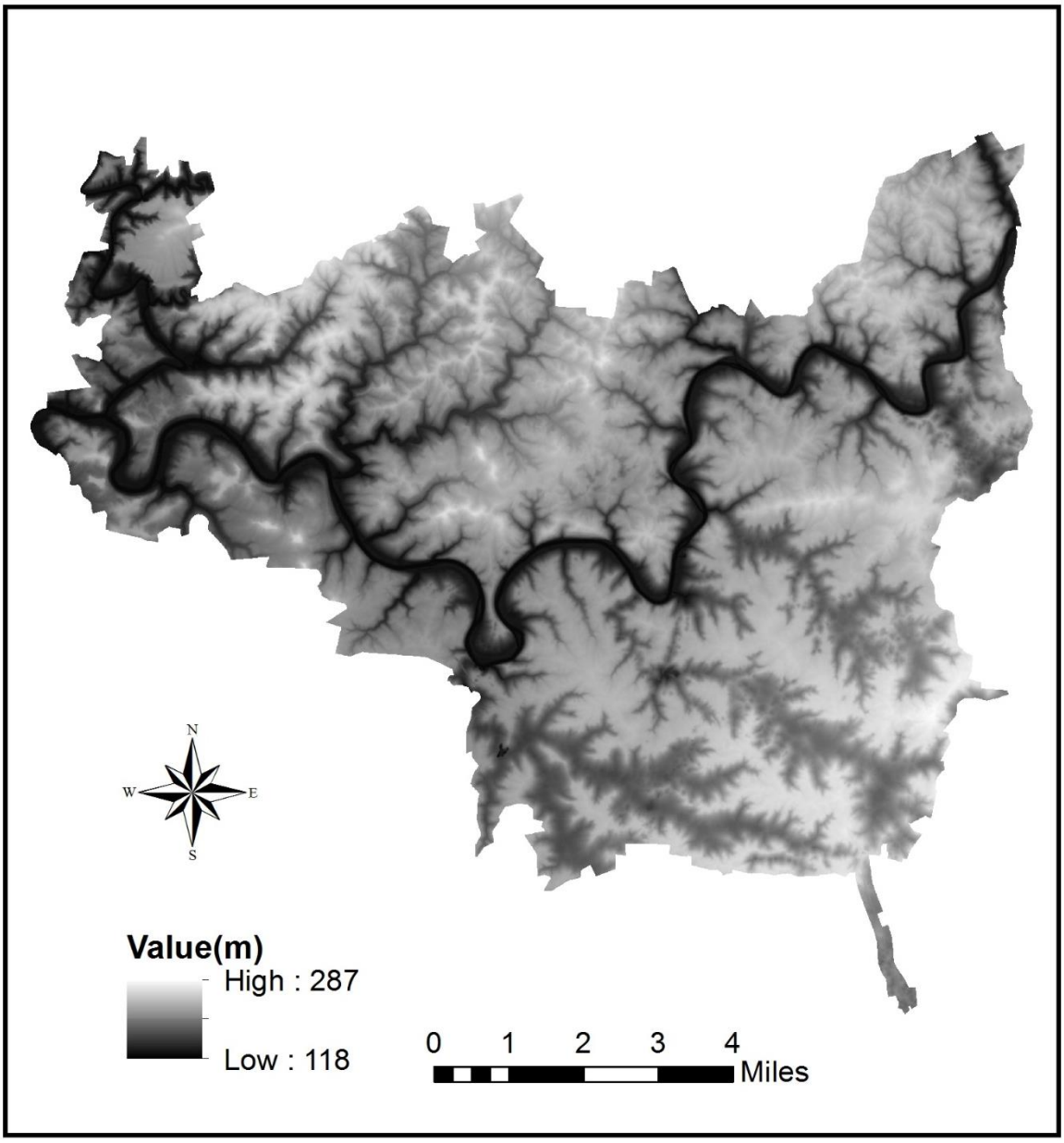

Fig.4: Derived Digital Elevation Model of LiDAR Dataset

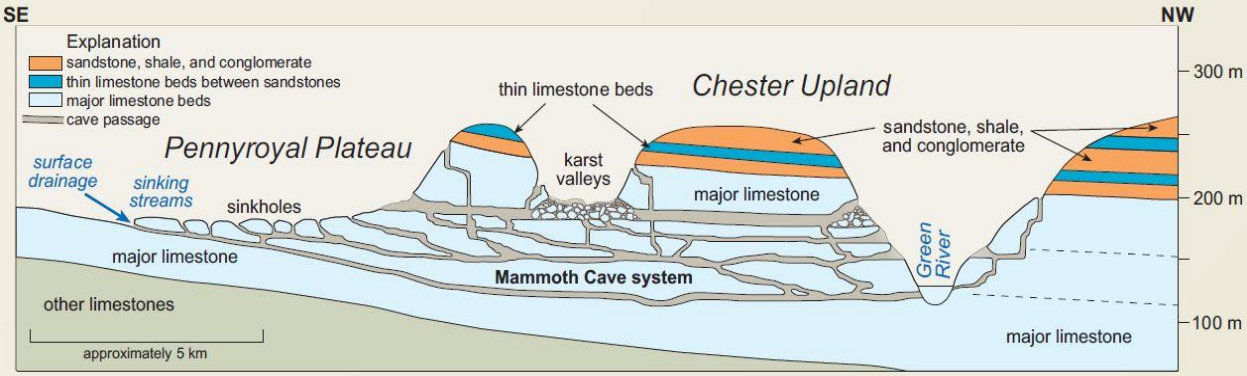

Fig.5. Cross section through the Mammoth Cave area showing the relationship of the cave to the surrounding land surface and geology (Crawford et al., 2013).

In addition, the land use map from 1936 (before the official establishment of the Park) was obtained to provide historic human disturbance data. Natural color digital orthophotographs were acquired the same time of LiDAR data acquisition. These 1 m spatial resolution photographs were used as reference during the class labeling of the unsupervised classification process. The Park also provides field plots data for accuracy assessment of the vegetation mapping results.

CHAPTER FOUR

METHODOLOGY

4.1 Vegetation Habitat Model

4.1.1 Model Configuration

Discussions with the Park staff and field trips to the Park provided the knowledge and scheme for the habitat model based on three physical attributes: bedrock type, slope (Figure 6), and aspect (Figure 7). Bedrock (Figure 8) is aggregated into four broad types: limestone (main cave), sandstone/shale, limestone caprock (exposed thin limestone beds overlaying sandstone/shale) and alluvium (flood plain deposits). Habitat developed over limestone bedrock is drier (more xeric) than that of same topographic condition in sandstone areas because water infiltrates limestone better while poorly drained soil over sandstone is more moist. The thin layer of limestone caprock with insoluble sandstone underneath creates more moist conditions. Slope is classified into three categories: flat (below 5°), moderate (between 5° and 23°), and steep (23° and above). Steeper slopes create more xeric conditions due to faster runoff thus less infiltration. The 360 degrees of aspect are grouped into 16 wedges of 22.5 degrees each. South and Southwest facing slopes create more xeric condition due to longer hours of sun exposure. For different degrees of slopes, the corresponding aspect ranges determined how xeric or mesic the habitat type is. Supra-mesic conditions were on the moist end of the mesic, but not saturated. Mesic conditions were moderately moist. Sub-mesic conditions were less moist compared to mesic. Sub-xeric conditions were intermediate between xeric and mesic. After several rounds of try and error working with Park Ecologist, Rick Olson, the final scheme for habitat prediction based on these three physical attribute is decided for the full array of habitat types (Table 1).

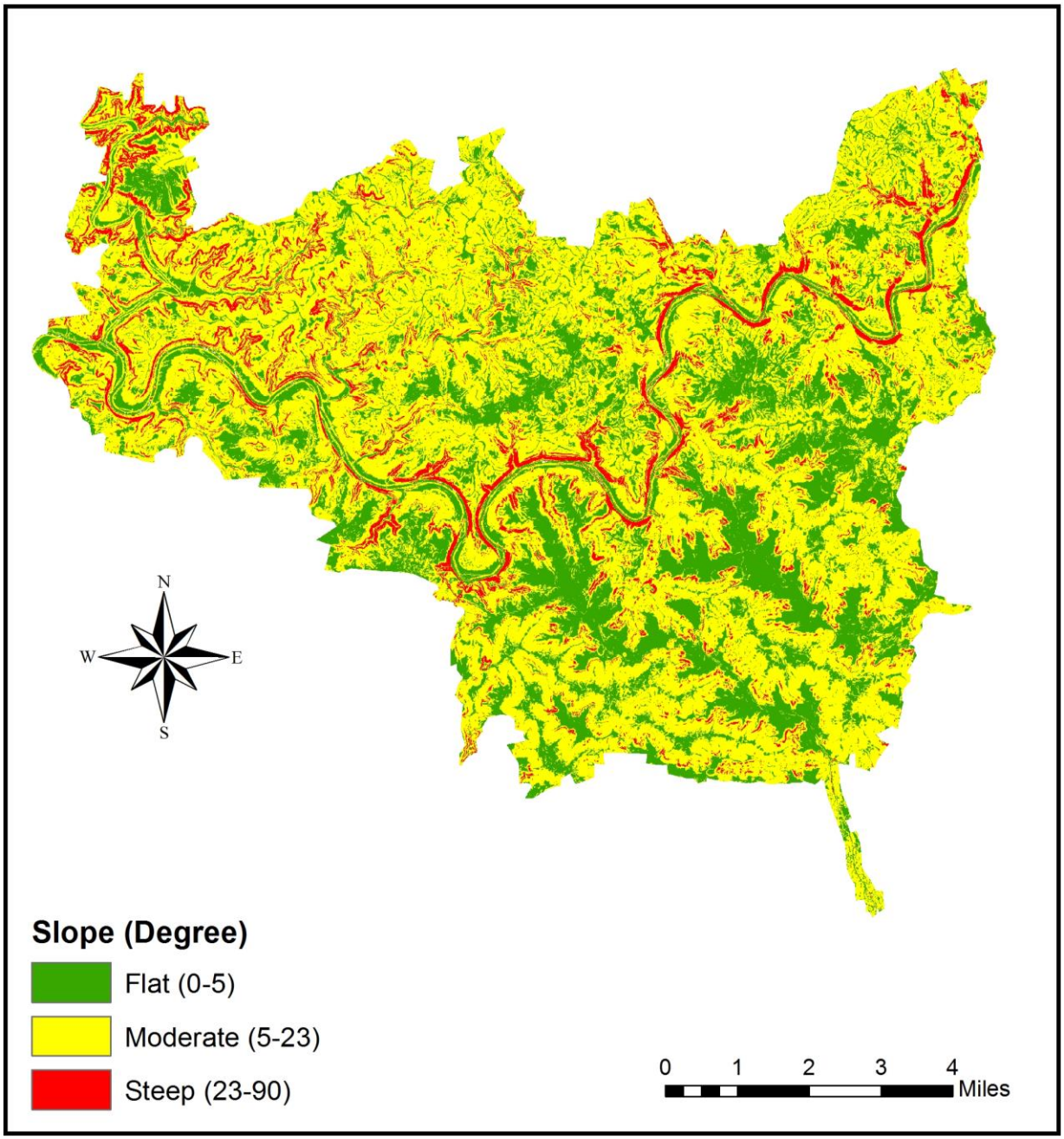

Fig.6: Slope Reclassification of Mammoth Cave National Park

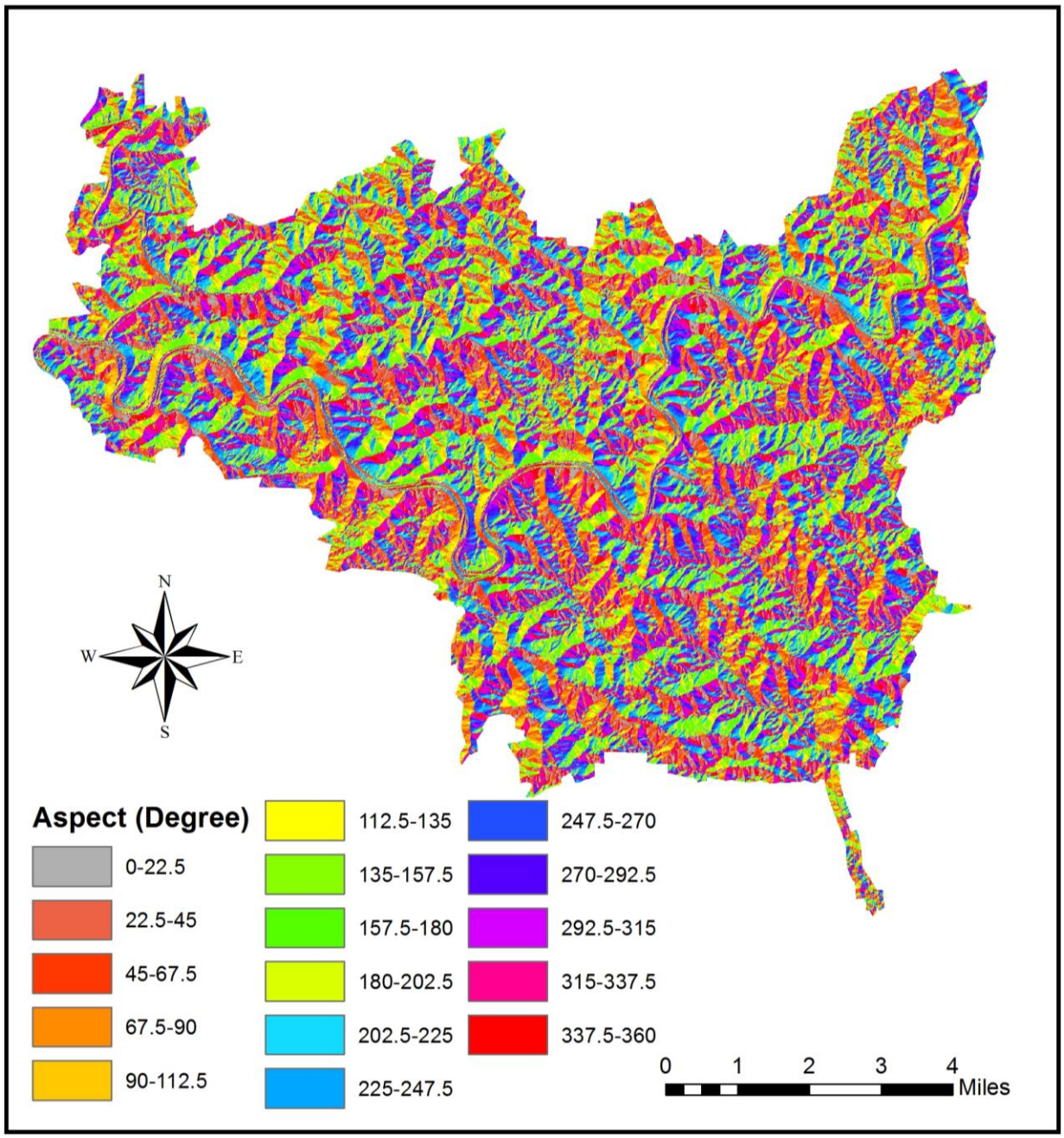

Fig.7: Aspect Reclassification Result

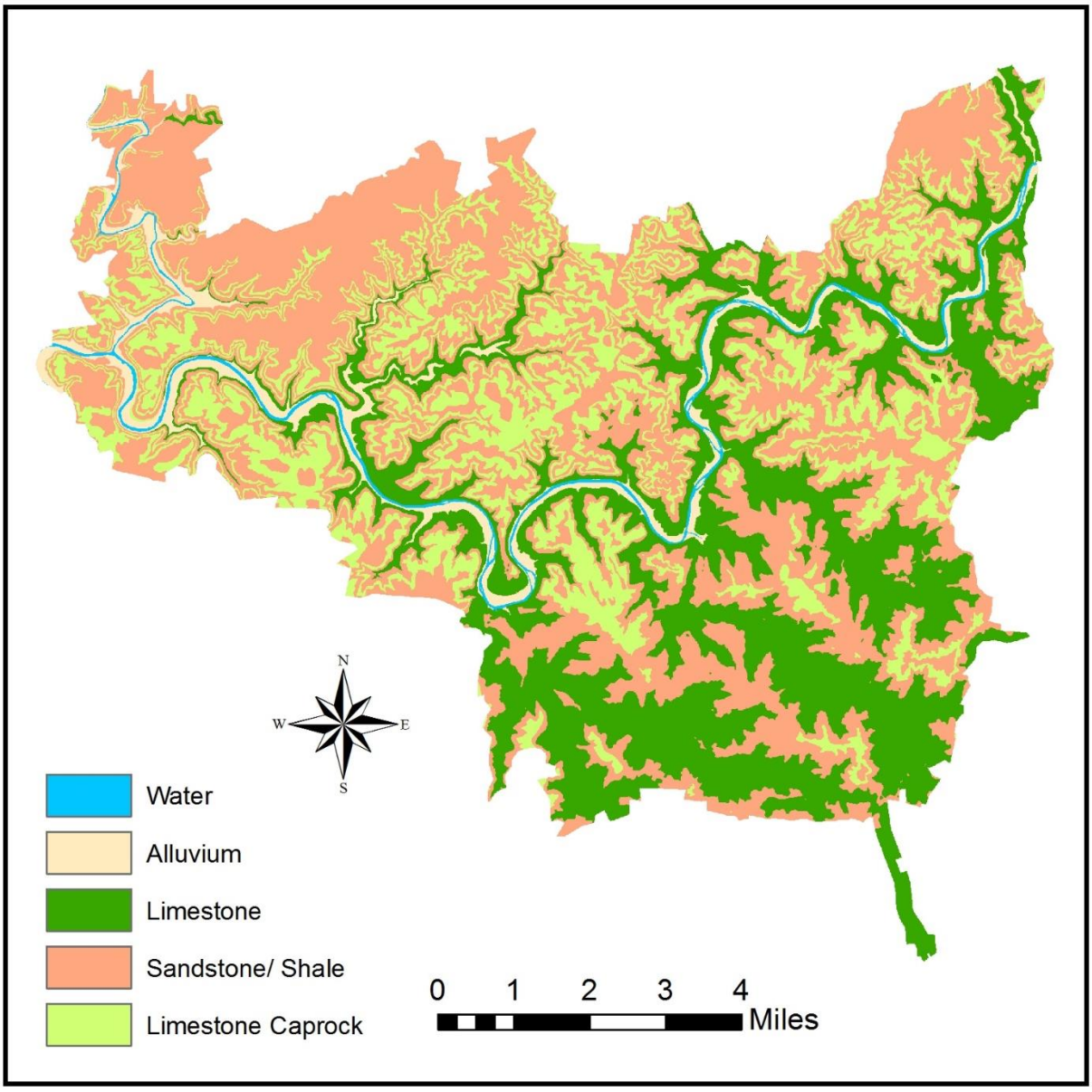

Fig.8: Bedrock Reclassification Result

Figures 9 and 10 show how the combination of moderate and steep slopes with limestone or sandstone bedrock are assigned to habitat classes based on aspect.

Table 1. Habitat physical attribute classification

1. Calcareous Xeric	Compass bearings
a. Southeast to West	180-247
2. Calcareous Sub-Xeric	
a. Flat	–
b. Moderate Southeast	112-180
c. Moderate Southwest	247-292
d. Steep Southeast	90-180
e. Steep Southwest	247-315
3. Calcareous Mesic	
a. Moderate Northwest to Southeast	292-112
b. Steep Northeast	045-090
4. Calcareous Supra-Mesic	
a. Steep Northwest to Northeast	315-045
5. Acid Xeric	
a. Steep Southeast to Southwest	157-247
6. Acid Sub-Xeric	
a. Moderate Southeast to West	135-270
b. Steep Southeast	135-157
c. Steep Southwest	247-270
7. Acid Mesic	
a. Flat (+ Hydro-Mesic vernal Ponds)	–
b. Moderate West to Northwest	270-315
c. Moderate Northeast to Southeast	045-135
d. Steep West to Northwest	270-315
e. Steep Northeast to Southeast	045-135
8. Acid Supra-Mesic	
a. Moderate Northwest to Northeast	337-022
b. Steep Northwest to Northeast	315-045
9. Floodplain Alluvium	–

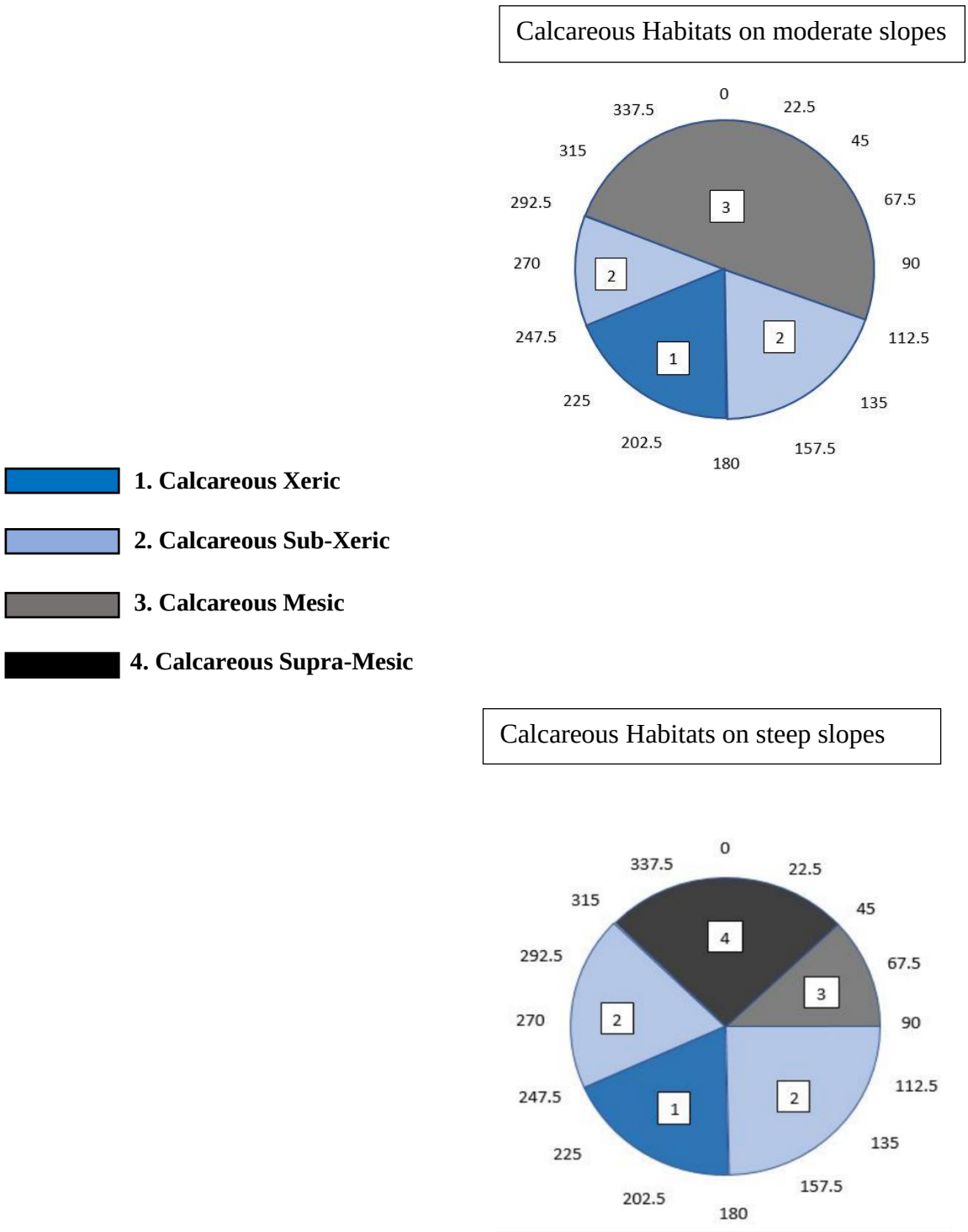

Fig.9: Calcareous habitat model with regard to slope and aspect.

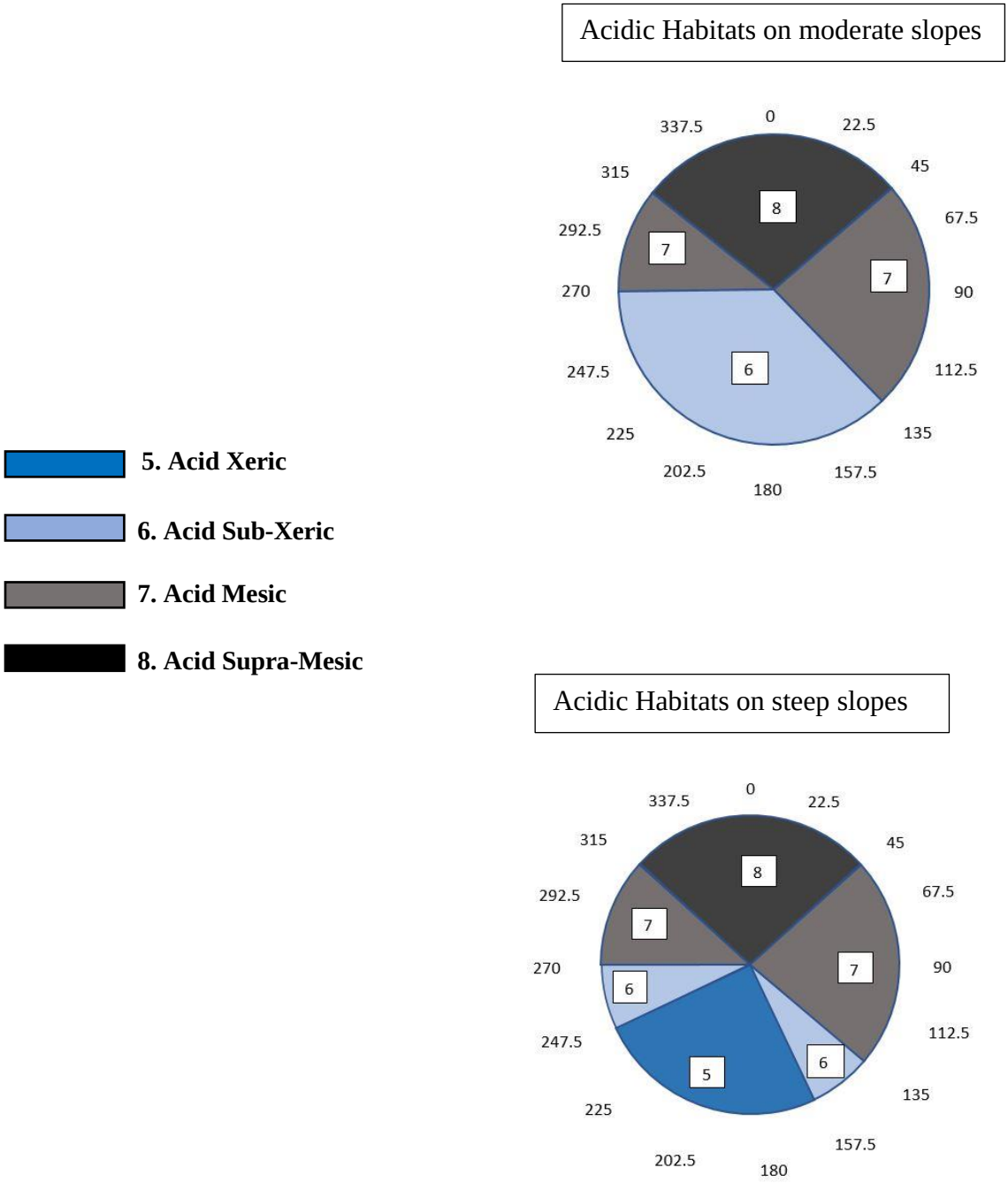

Fig.10: Acidic habitat model with regard to slope and aspect

4.1.2 Model application

Aspect and Slope were calculated from LiDAR-derived DEM data and reclassified. The bedrock dataset was simplified into four broad categories. To make the data processing manageable, it was determined that 5 m by 5 m pixels be an appropriate resolution. GIS overlay function was applied to determine habitat types by integrating aspect, slope and bedrock. Given the bedrock types, certain slope combined with certain aspect determines the habitat type (Table 1). For example, habitat located over limestone bedrock on moderate slope with an aspect of 135° (southeast facing) is determined to be Calcareous Sub-Xeric type.

4.1.3 Accuracy assessment of the vegetation habitat model

A total of 600 random points were generated using stratified random sampling strategy. Among all 600 points, around 20 points for each category that have relatively easy access to roads were selected as possible candidates for field verification. The field reference data were collected between April 28th and 30th, 2017. We were able to visit a total of 30 points during the three days with the Park Ecologists, Mr. Olson. The aspect and slope were measured onsite. Bedrock, vegetation species, and habitat type were determined and recorded as well (Appendix I).

4.2 Vegetation Type Mapping

4.2.1 Image Preprocessing

Radiation from Earth's surface undergoes significant interaction with the atmosphere before it reaches the satellite sensor (Hadjimitsis et al., 2010). The atmosphere can always influence the radiation from the ground to the sensor. Therefore, it is essential to consider the effects caused by the atmosphere by applying an efficient method during pre-processing of

digital data. Top of Atmospheric (TOA) reflectance and sun angle correction were executed for the Landsat 8 images (USGS, 2016).

PCA is known to be effective to reduce the dimensionality in Landsat imagery (Kwarteng and Chavez, 1989). It transforms the original image into a set of uncorrelated variables that represents most of the information present in the image. The two Landsat-8 OLI images from January and June were combined into one dataset, and PCA was performed on the combined two-date dataset. The first four components that accounted for 99% of the variance were selected for image classification later.

4.2.2 Enhanced Vegetation Index Calculation

Vegetation indices have been derived from Landsat to capture characteristics of vegetation. The Enhanced Vegetation Index (EVI) was developed to optimize the vegetation signal with improved sensitivity in high biomass regions and improved vegetation monitoring through a de-coupling of the canopy background signal and a reduction in aerosol influences (Huete et al. 2002):

$$EVI = G * ((\rho\text{NIR} - \rho\text{red})) / (\rho\text{NIR} + \text{C1} * \rho\text{red} - \text{C2} * \rho\text{blue} + \text{L}) \quad (1)$$

where ρNIR, ρred, and $\rho blue$ are the atmospherically corrected reflectance for the NIR, red, and blue spectral bands; L is the canopy background adjustment factor that addresses nonlinear, differential NIR and red radiant transfer through a canopy; and C_1 and C_2 are the coefficients of the aerosol resistance term, which uses the blue band to correct for aerosol influences in the red band. In the MODIS EVI algorithm, the coefficients of $L = 1$, $C_1 = 6$, $C_2 = 7.5$, and $G = 2.5$ are adopted (Huete et al.2002; Huete et al. 1997, which are also applicable for Landsat-8 OLI products (USGS, 2017; Masek et al. 2006). For this research, EVI was calculated for the January and June images respectively for integration into the classification process.

4.2.3 Image Classification

Unsupervised classification is an effective method of partitioning remote sensing image data in multispectral feature space and extracting land cover information (Loveland et al., 1999; Huang, 2002). It requires minimum knowledge of the study area, which mainly uses some clustering algorithm to classify image data (Richards, 1993). In this research, the ISODATA algorithm was applied twice to produce two vegetation maps of the Park based on two different imagery inputs: 1) the imagery composed of the four principal components, and 2) the imagery composed the four principal components and the two EVI images. The two resultant maps were then compared for their accuracy.

For classification of each image, 250 initial clusters were generated by the ISODATA algorithm. They were then examined, labeled, and aggregated into five classes: Barren Land/ Man-made Structure, Evergreen, Deciduous, Mixed forest, and Water. The orthophotographs were used as reference for the labeling of deciduous, evergreen and mixed forest. Applying the same definition as the National Land Cover Dataset (NLCD, Homer et al., 2012), if evergreen or deciduous trees occupied 25-75% in one Landsat pixel based on the reference orthophotos, the pixel would be "mixed".

4.2.4 Accuracy assessment of classification results

Accuracy assessment quantitatively determines how effectively pixels were grouped into their corresponding classes in the investigated area. Each classification entails assessment of the consistency and reliability of the results. The accuracy assessment includes three steps: using field plot data as testing samples, preparing a confusion matrix resulting from field plots and classified clusters, and calculating the accuracy measurement.

The field plots data contains 750 deciduous points, 47 evergreen points and, 65 mixed points. Instead of applying all the deciduous points for accuracy assessment, I randomly selected 398 out of 750 points were randomly selected for accuracy assessment (the remaining 352 points were used in the cluster labeling process during classification). All 47 evergreen plots and 65 mixed plots were applied for accuracy assessment.

The error matrix is the most common method of reporting classification assessment. An accuracy assessment report depicts the user's accuracy and producer's accuracy. The producer's accuracy indicates how well a certain study area can be classified (omission error). In contrast, the user's accuracy is a measure of commission error.

Accuracy assessment report also includes the Kappa statistics as well (Equation 2). The Kappa Coefficient of Agreement measures the performance of the classification compares to the possibility to assign pixel randomly into classes. The results of Kappa closer to 1 means that the classification is considered much better than chance agreement, and closer to zero means that the classification is no better than chance (Congalton, 1991).

$$\hat{k} = \frac{N\sum_{i=1}^{r} x_{ii} - \sum_{i=1}^{r} (x_{i+} \cdot x_{+i})}{N^2 - \sum_{i=1}^{r} (x_{i+} \cdot x_{+i})} \qquad \text{(Equation: 2)}$$

Where: r = the number of the row in the matrix,

X_{ii} = sum of diagonal,

X_{i+} = the total of observation in row I,

X_{+i} = the total observations in column I, respectively,

N = the total number of the samples

CHAPTER FIVE
RESULTS AND DISCUSSION

5.1 Results of vegetation habitat modeling

The LiDAR-Derived DEM shows that the highest elevation is 287 m and the lowest 107 m (Figure 4). Lower elevation values are near streams and the Green River while higher elevation areas are in northwest and southeast portions of the Park.

Figures 6 and 7 show the resultant map of slope and aspect reclassification. Most areas at lower elevation (except water areas) had moderate to steep slopes, indicating that gullies or depressions existed in those regions. Therefore, these shaded regions tended to be more mesic (moisture-laden) than sun-exposed areas. In contrast, areas at higher elevation with steep slopes tend to divert water away fast, which caused more xeric (dry) conditions. The amount of solar radiation on the landscape changes during a day and seasonally, according to the aspect that the slope is facing. Typically, south-facing slopes are exposed to much more sunlight compared to north-facing slopes in the northern hemisphere. Thus, south-facing areas would be more xeric while north-facing slopes regions more mesic.

The bedrock reclassification map (Figure 8) shows the distribution of Calcareous, Acid (Sandstone/Shale) and Calcareous Caprock bedrock types. Lower elevation areas south of the Green River are mostly main cave Calcareous bedrock. Acidic bedrock, on top of the main cave limestone, spreads throughout the park. Limestone caprock is found on higher elevation overlaying insoluble sandstone.

Figure 11 is the result of habitat modeling considering bedrock, aspect, and slope. The two major habitat types were largely determined by bedrock (Table 2). The variation within the

two main habitat types formed under different slope and aspect conditions. Most Calcareous Xeric and Calcareous Sub-Xeric habitat areas are in the southeastern part while Calcareous Sub-Mesic habitat spreads across the Park. The Acid Xeric and Acid Sub-Xeric habitat mostly were in the northwestern part while moderate elevation regions in the southeast formed Acid Mesic habitat.

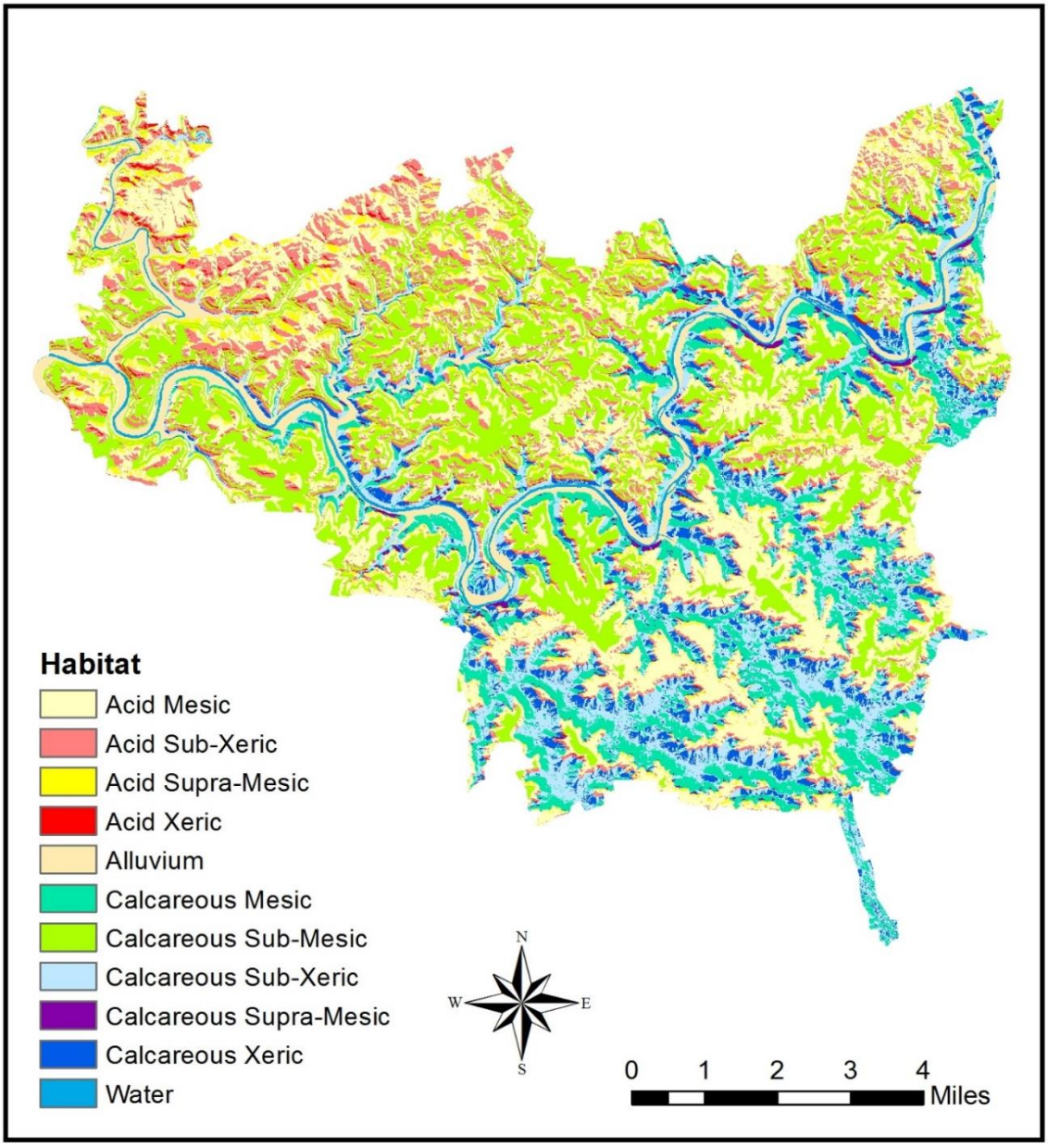

Fig.11: Habitat modeling result based on geology, slope, and aspect

Calcareous Xeric and Calcareous Supra-Mesic habitats are displayed in Figures 12 and 13. The two habitats occur on relatively steep slopes, with Xeric habitat facing southeast and Supra-Mesic habitat facing north.

Fig.12. Calcareous Xeric Habitat

Fig.13. Calcareous Supra-Mesic Habitat

Acid Sub-Xeric (Figure 14), Calcareous Sub-Xeric (Figure 15), Acid Mesic (Figure 16) habitats are displayed as examples to compare the differences between each other. Mesic habitats are dominated by deciduous trees with abundant understory growth while the density of vegetation in Sub-Xeric habitats is less dense with more evergreen trees.

Fig.14. Acid Sub-Xeric Habitat

Fig.15. Calcareous Sub-Xeric Habitat

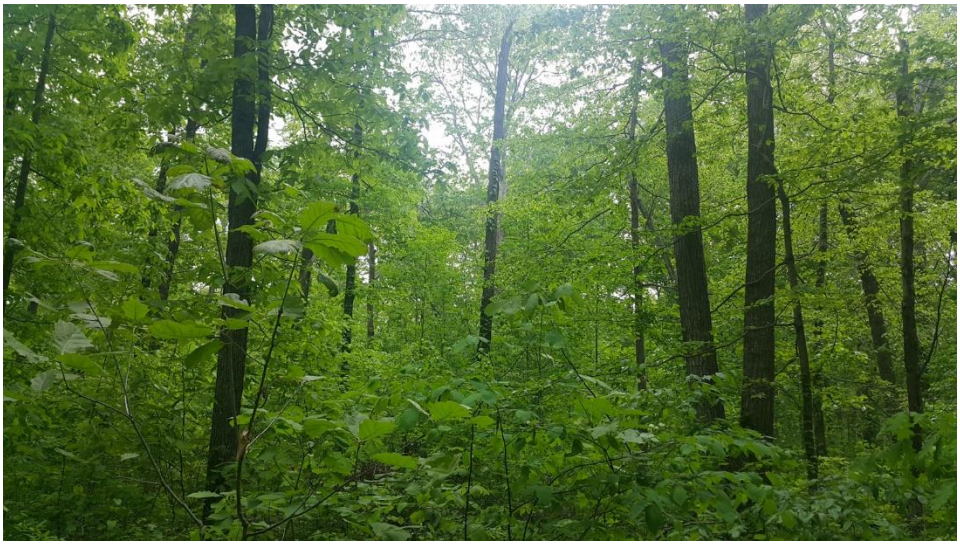

Fig.16. Acid Mesic Habitat

Table 2. Habitat type area in hectares and the corresponding percentages by the Park in 2017

Habitat Type	Acreage	Percentage of the Park (%)
Calcareous Supra-Mesic	**440.41**	**0.87**
Calcareous Mesic	**5,667.45**	**11.18**
Calcareous Sub-Mesic	10,662.63	21.03
Calcareous Sub-Xeric	5,995.60	11.83
Calcareous Xeric	2,448.75	4.83
Alluvium	**2,036.59**	**4.02**
Acid Supra-Mesic	**2,180.86**	**4.30**
Acid Mesic	14,832.60	29.26
Acid Sub-Xeric	5,962.35	11.76
Acid Xeric	466.28	0.92

Acid and Calcareous are the two dominant habitats within the park, which accounted for 46.24% and 49.74% of the total park area respectively (Table 2). Acid Mesic had the largest areas of a single category followed by Calcareous Sub-Mesic, which accounted for 29.26% and 21.03% respectively. Calcareous Supra-Mesic and Acid Xeric each accounted for less than 1%. Calcareous Mesic, Calcareous Sub-Xeric and Acid Sub-Xeric habitat types occupy about the same area in the Park (11-12%). The remaining three habitat types are less than 5% each.

Habitat types in Table 2 follow the classification system of the Kentucky State Nature Preserves Commission (Evan 1991). Table 3 shows the typical species of each habitat type in deciduous, mixed and coniferous forests. Various types of oak trees dominant the deciduous forests in more xeric conditions, such as Chestnut oak, post oak, Chinkapin oak, and blackjack oak. White oak, black oak and pignut hickory are commonly found in mesic upland deciduous forest. In the mesic valleys and floodplain, sugar maple, beech, box elder, and sycamore trees are more common. In mixed forests, the dominant deciduous species are red maple, tulip popular, dogwood, and sweetgum. Eastern red cedar and Virginia pine are the two major coniferous species found throughout the park.

Table 3. Typical vegetation species for each habitat of the Park (Olson and Noble, 2005).

Vegetation	Habitat type	Typical Species	Fire frequency
1. Sub-Xeric deciduous forest/ savanna	Acid Sub-Xeric Calcareous Sub-Xeric	Chestnut oak, post oak Chinkapin oak, blackjack oak	Frequent
2. Mesic upland deciduous	Acid Mesic Calcareous Sub-Xeric (thin beds)	White oak Pignut hickory, black oak	Frequent
3. Mesic hollow/ floodplain deciduous forest	Calcareous Mesic Acid Mesic Alluvium	Sugar maple Beech Box elder, sycamore	Rare
4/5. Mixed deciduous/ coniferous Mixed coniferous/ deciduous forest	Acid Mesic Calcareous Sub-Xeric, Alluvium	Red maple, tulip popular Dogwood, sweetgum, cedar/pine	Infrequent
6. Coniferous forest	Acid Xeric to Mesic Calcareous Xeric to Sub-Xeric	Virginia pine Eastern red cedar	Infrequent
7. Prairie/ open area	Calcareous Sub-Xeric Acid Mesic	Native Grasses and Forbs Mown Grass	

Olson and Noble (2005) noted that fire was rare in mesic hollow/floodplain deciduous and infrequent in coniferous forest and mixed forests, while fire was frequent in all other types of deciduous forest (Table 3). Based on that information, habitat types in regular typeface in Table 2 are capable of carrying fire during the spring and fall fire seasons (Olson and Noble. 2005). These habitat types account for approximately four-fifths of the Park. Habitat types in bold in Table 2, which account for approximately one-fifth of the park, do not support fire independent or fire-tolerant plant communities (Olson and Noble, 2005). Figure 17 shows the location of fire-sensitive and fire-tolerant habitat types. In general, the fire-sensitive habitat types are surrounded by fire-tolerant habitat types, a good sign for management.

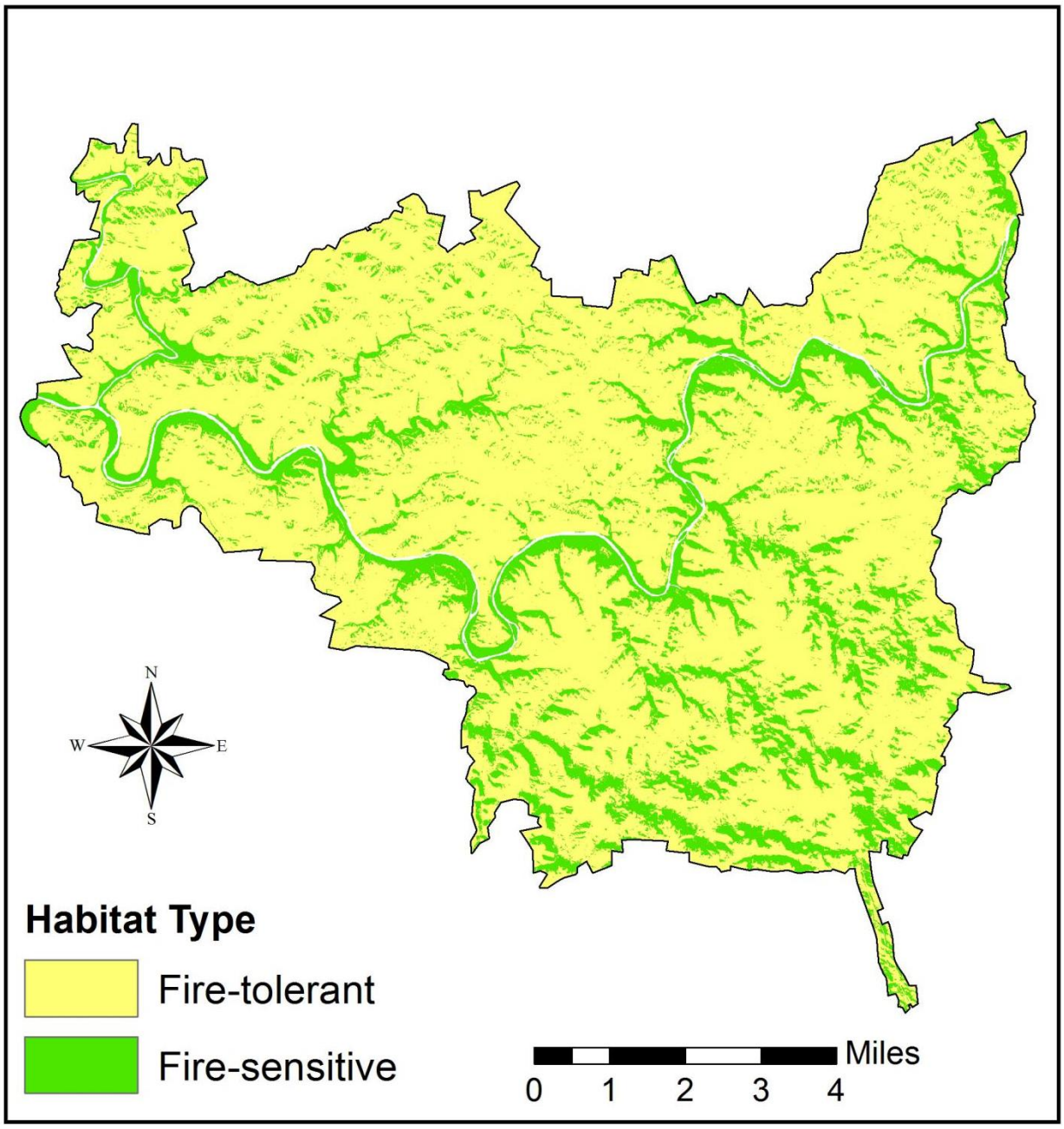

Fig.17. The distribution of fire-sensitive and fire-tolerant habitat types

5.2 Accuracy of habitat modeling

Figure 18 shows the location of the 30 sites we were able to visit in the field, to evaluate the accuracy of the habitat model. One site was discarded due to its location at a road cut. For the remaining 29 sites, the habitat types as predicted by the model for 22 sites were consistent with

field observations (76%). The model results for five out of the nine acidic and calcareous habitat categories matched the field observations completely.

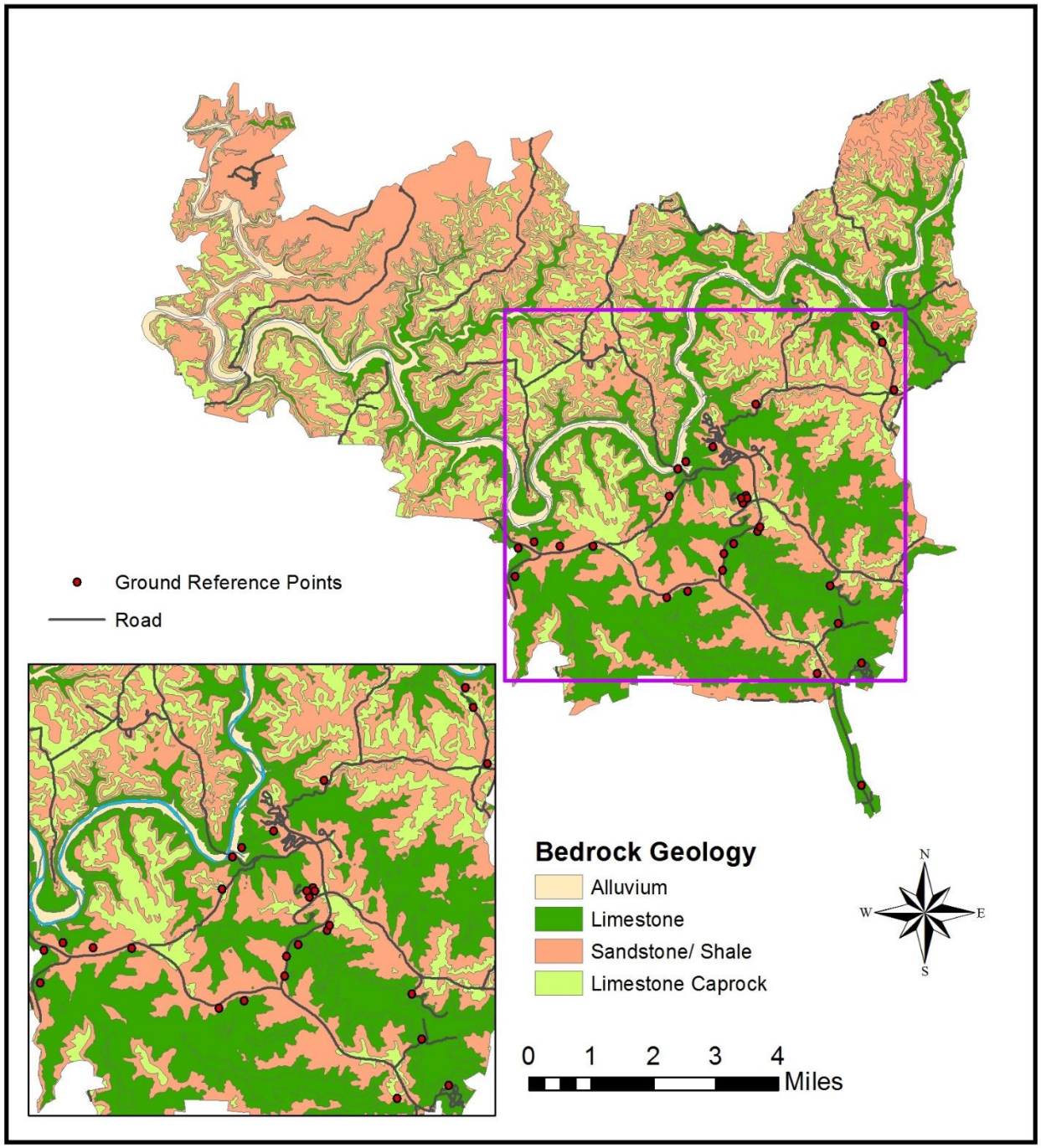

Fig.18. The distribution of sites visited for ground truthing

Table 4 shows the seven (7) sites where inconsistency between model result and ground truth was identified and the reasons for inconsistency. At three of the locations, previous human disturbance was the most likely reason for the inconsistency (sites ID #437, #434 and #122). The model was clearly problematic for Acid Xeric habitat. Four sites where the model predicted Acid Xeric, all turned out to be Sub-Xeric. Acidic Xeric habitat accounted for a very small proportion of the Park. It would be found on top of steep sandstone cliffs usually marked by a pine tree stand. These locations have proven difficult to identify due to their small size and the 5 m resolution of the model.

Table 4. Habitat Modeling and Ground Truth comparison of 7 observed sites

<table>
<tr><th>Site Name</th><th>Habitat modeling result</th><th>Ground truth</th><th>Reason</th></tr>
<tr><td>ID#437</td><td>Acid Supra-Mesic</td><td>Calcareous Sub-Mesic or Acid Mesic</td><td>It is an old crop field where habitat types are influenced by previous human disturbance.</td></tr>
<tr><td>ID#521</td><td rowspan="4">Acid Xeric</td><td rowspan="4">Acid Sub-Xeric</td><td rowspan="4">Acid Xeric habitats are at specific locations with steep sandstone cliff, with a pine stand on top. It is difficult to locate.</td></tr>
<tr><td>ID#550</td></tr>
<tr><td>ID#555</td></tr>
<tr><td>ID#557</td></tr>
<tr><td>ID#434</td><td>Calcareous Sub-Xeric</td><td>Calcareous Mesic</td><td>It is in a transition zone from Calcareous Mesic and Calcareous Sub-Xeric, with previous disturbance likely fire.</td></tr>
<tr><td>ID#122</td><td>Calcareous Xeric</td><td>Calcareous Sub-Mesic to Sub-Xeric</td><td>It is an old crop field. Previous human disturbance likely contributed to current site condition.</td></tr>
</table>

5.3 Results of vegetation type mapping

The classification results include five classes: barren land/ man-made structure, evergreen (eastern red cedar, ferns), deciduous (sugar maple, pignut hickory, tulip popular), mixed forest, and water. Figure 19 displays the results of classification using principal components alone and Figure 20 shows the result of using both principal components and EVI images. Deciduous forest, which accounted for the largest area, was found throughout the Park. Most evergreen forest was located in the southern part of the park. Mixed forest was in close proximity to evergreen and it had the smallest amount of area. Comparing the two results, the inclusion of EVI index helped distinguish mixed and evergreen forest better. Some pixels of mixed forest were misclassified as evergreen forest in Figure 19; mixed forests are better defined in Figure 20. To put it in perspective, the classification results were compared to National Land Cover Database (NLCD) 2011 (Homer et al., 2012), the only other land cover map readily available. The same field plots data were used to evaluate the accuracy of all three land cover products.

Tables 5-7 show the error matrices for the two classifications and NLCD. The overall accuracy for PCA technique and EVI index was 85%, which was relatively higher compared to the PCA technique (78%) and NLCD 2011 (72%). Among the different types of forest, deciduous was the most accurately mapped with between 87% and 97% in both producer's and user's accuracy in this study. Incorporating the EVI in classification boosted the accuracy of both evergreen and mixed forests. The accuracy for mixed forest was lower than any other class, which had the highest percent correct with the inclusion of EVI index. Using the PCA alone, the evergreen and deciduous forests also had lower accuracies (59% for the producer's accuracy and 43% for the user's accuracy).

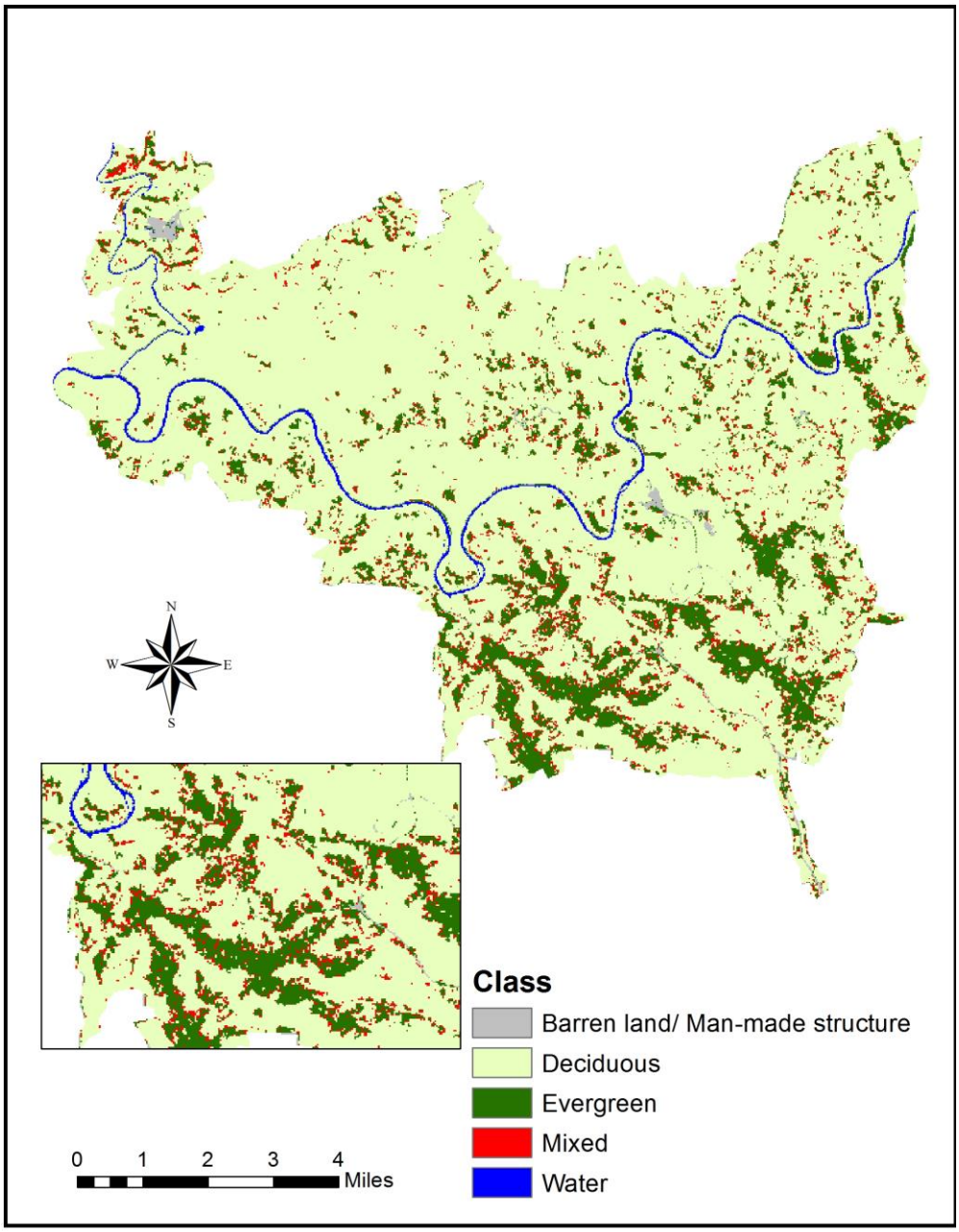

Fig.19: The classification result of PCA technique of Mammoth Cave National Park

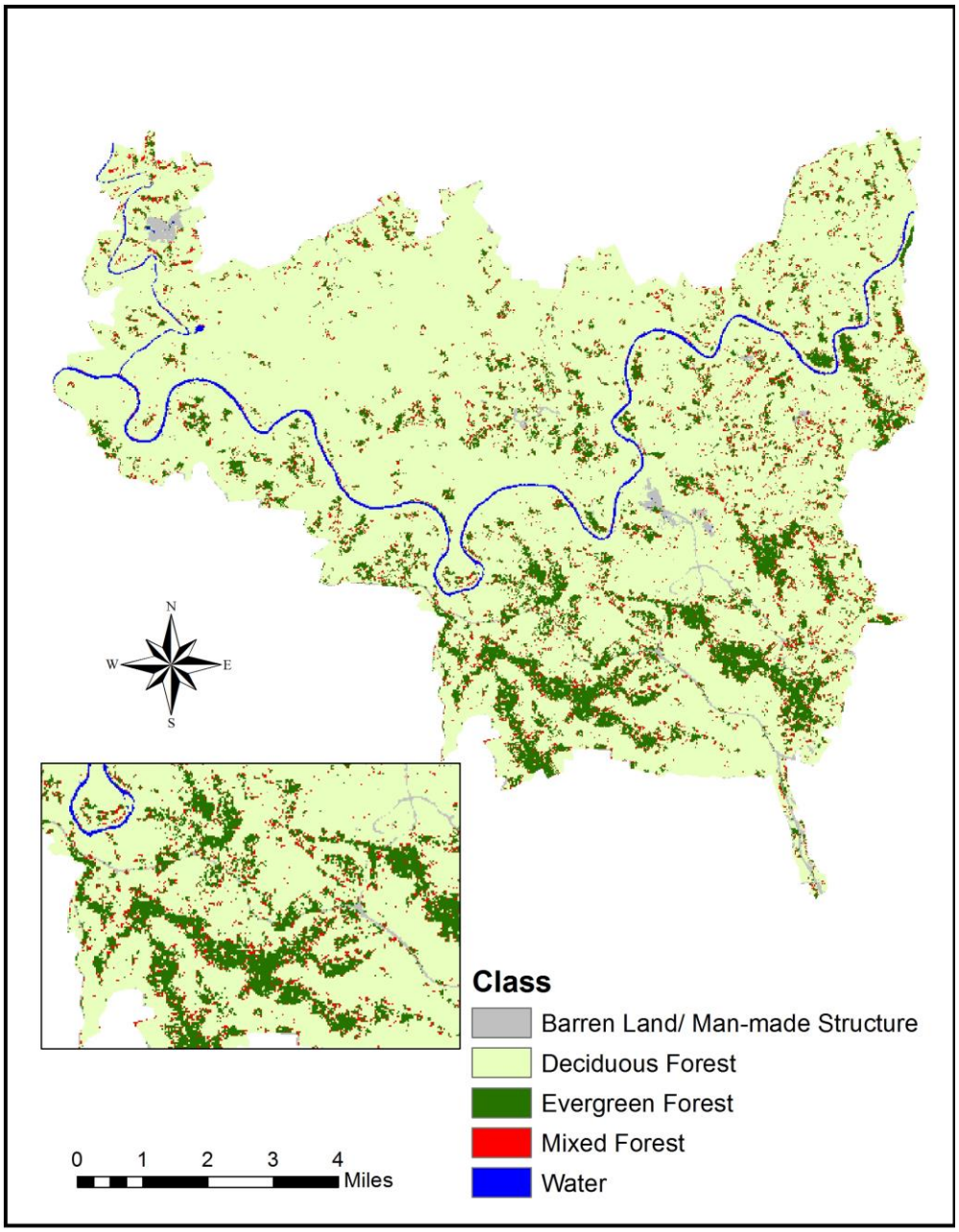

Fig.20: The classification result of PCA technique and EVI Index of Mammoth Cave National Park

Table 5. Error Matrix for the classification of PCA technique

		Reference Data					
Classified Data		Barren Land/ Man-made Structure	Evergreen	Deciduous	Mixed	Water	Total
	Barren Land/ Man-made Structure		1	1			2
	Evergreen		42	12	24		78
	Deciduous		25	368	29		422
	Mixed		8	14	12		34
	Water			2			2
	Total		76	397	65		538

Table 6. Error Matrix for the classification of PCA technique and EVI Index

		Reference Data					
Classified Data		Barren Land/ Man-made Structure	Evergreen	Deciduous	Mixed	Water	Total
	Barren Land/ Man-made Structure		1	1			2
	Evergreen		48	7	17		72
	Deciduous		22	385	26		433
	Mixed		5	3	22		30
	Water			1			1
	Total		76	397	65		538

Table 7. Error Matrix for the classification of National Land Cover Database

		Reference Data					
Classified Data		Woody Wetlands	Evergreen	Deciduous	Mixed	Water	Total
	Woody Wetlands			18			18
	Evergreen		60	42	41		143
	Deciduous		15	320	15		350
	Mixed		1	13	9		23
	Water			4			4
	Total		76	397	65		538

Table 8. Accuracy report of the classification results for vegetation types

	Accuracy of PCA technique		Accuracy of PCA technique and EVI index		Accuracy of NLCD 2011	
Class Name	Producer's Accuracy	User's Accuracy	Producer's Accuracy	User's Accuracy	Producer's Accuracy	User's Accuracy
Evergreen	55%	54%	63%	66%	79%	42%
Deciduous	92%	87%	97%	89%	80%	91%
Mixed	18%	35%	34%	73%	14%	39%
Overall classification Accuracy	78%		85%		72%	
Overall kappa statistics	0.45		0.61		0.42	

The comparison of NLCD between classified images using different techniques demonstrated that PCA technique and EVI index improved the overall classification accuracy to some extent. In addition, mixed forests' low classification accuracy in NLCD also demonstrated the classification limitation for classifying mixed forests. The inclusion of EVI index improved the accuracy of mixed forest to some extent.

5.4 The impact of historic disturbance on vegetation

Since the evergreen, particularly eastern red cedar, the dominate evergreen trees in the Park, is considered the first successional forest after restocking on previously cropland or pasture – "old field", it would be worthwhile to cross examine the current evergreen coverage and historic disturbance (Figure 21). The overlay result shows that most of evergreen forests (70%) at present time are located in old fields, where ecological succession occurred after pre-park pasture and row crop use. The old fields are generally found in three habitat types (Olson and Noble, 2005): 1) on relatively level uplands with interbedded sandstone and limestone, 2) in sub-xeric limestone habitats found in karst valley, and 3) on floodplain alluvium deposit.

Considering future succession stages, the eastern red cedar dominated evergreen forest will only be supported in xeric and sub-xeric habitats. We can predict that most cedar trees currently found in mesic habitat will be replaced by deciduous trees as succession continues and the forest matures. Figure 22 shows the overlay result of the combination of Xeric and Sub-Xeric habitat types with evergreen forests growing in previously disturbed areas. The evergreen forests currently growing in Xeric or Sub-Xeric habitat types amounted to about 40% of the total evergreen forests and cedar trees in these area will likely to survive. The remaining 60% of evergreen forests neither grow on Xeric nor Sub-Xeric habitat will likely be succeeded by deciduous forests in the future. How further succession of the coniferous forests by deciduous forests will affect the fire regime will depend on the habitat characteristics of the location (Table 3).

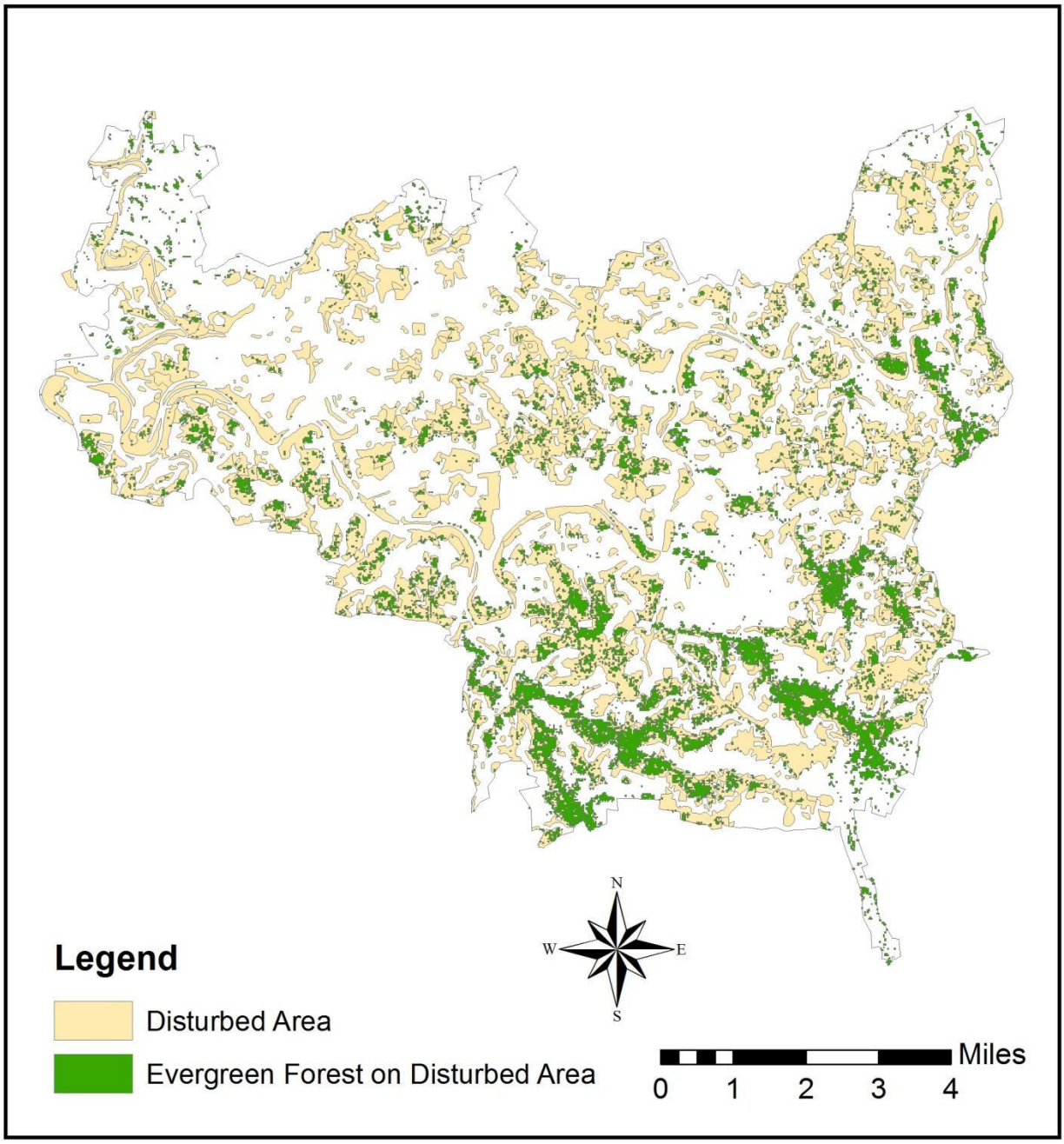

Fig.21: The result of evergreen forests on disturbed areas in the Park

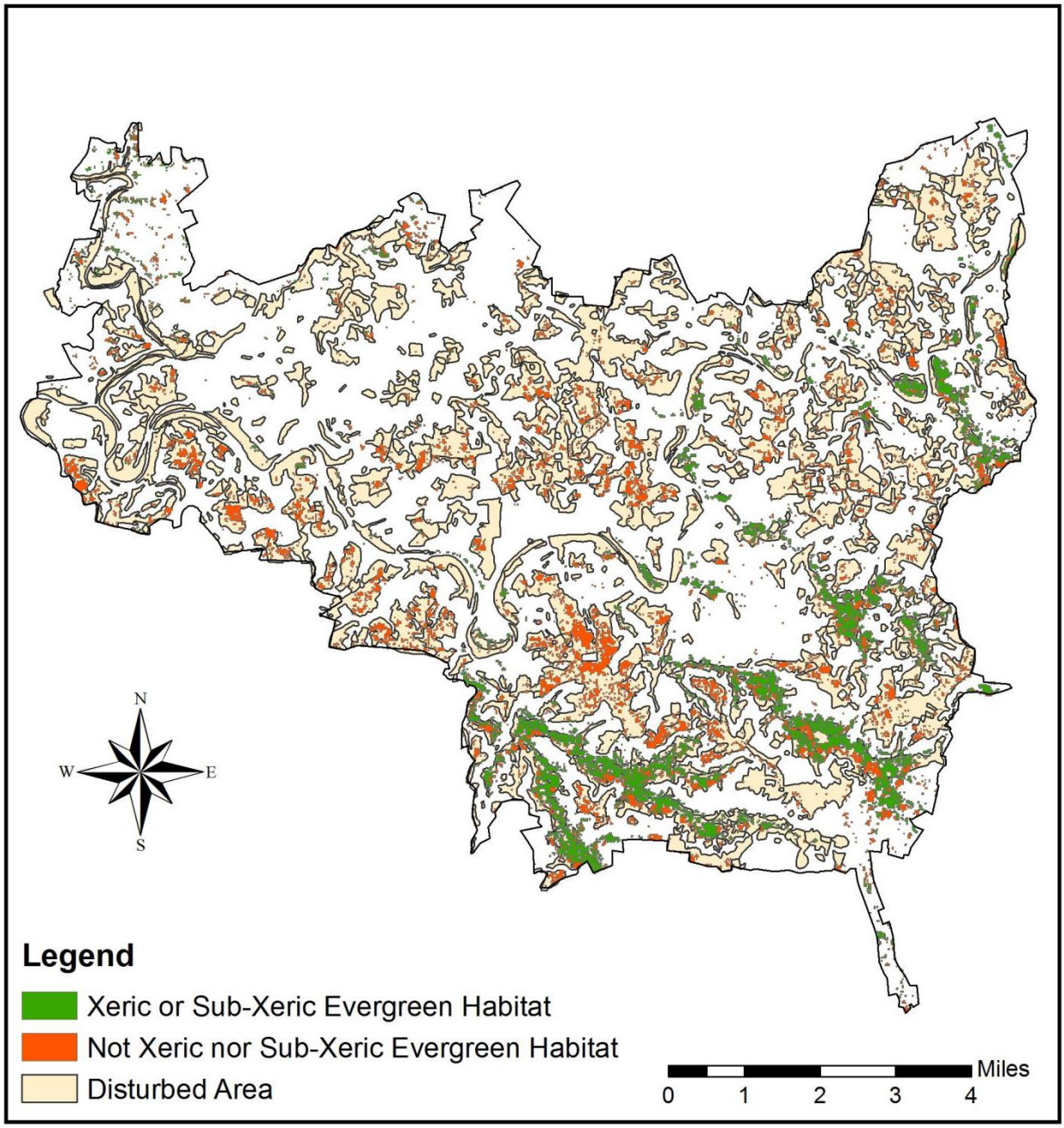

Fig.22: The location of Xeric or Sub-Xeric Evergreen habitat on Disturbed areas

CHAPTER SIX
CONCLUSIONS AND RECOMMENDATIONS

This study demonstrated the procedures and potential of integrating remote sensing and GIS techniques to model habitat types and map vegetation types of Mammoth Cave National Park. This research has followed several main steps, which included: (1) the acquisition of remotely sensed data, LiDAR data, bedrock, and other ancillary data, (2) image preprocessing, such as atmospheric correction, PCA transformation, and EVI calculation, (3) habitat type modeling using the combination of aspect, slope and bedrock, (4) unsupervised classification using ISODATA algorithm for two different datasets: PCA, and PCA and EVI Index, (5) interpretation of the results and accuracy assessment.

Vegetation habitat modeling attempts to predict the geographic distribution of plant cover types from mapped environmental variables. The research demonstrated that a simple model using three environmental factors – slope, aspect and bedrock – was able to determine to a large degree the different habitat types in Mammoth Cave National Park. The variation of aspect and slope affect the amount of solar radiation and water available to vegetation, which influences the contrasting habitat types formed in the long term. Bedrock, one of the most influential factors in the study area, primarily controls the soil types and drainage conditions that support the various habitat types. Field verification of the results at 29 sites selected from the pool of 600 random locations showed that the model correctly predicted the habitat types of 22 sites (76%). Inconsistencies found at three sites were due to human disturbance before the establishment of the Park. The only category that would need further work is Acid Xeric type where the model missed the mark due to specific condition where this habitat could be found and the resolution of the model.

The habitat model result indicates that Acid and Calcareous are the two dominant habitat categories within the Park, which accounted for 46.24% and 49.74% of the total area respectively. Acid Mesic and Calcareous Sub-Mesic were the two most popular habitats, representing 29.26% and 21.03% of the area respectively. Fire-sensitive habitat types are surrounded by fire-tolerant types which acts as a buffer when fire events happen.

Leaf-on and leaf-off Satellite images from two different dates were used to map the vegetation types. PCA technique was used to transform the two-date image composite to four principal components that accounted for 99% of the variance, which was used as input to image classification. EVI Index was incorporated to improve classification accuracy. Accuracies of classification results were evaluated by reference data from 538 field plots. The overall accuracy of PCA and EVI Index was 85%, better than using PCA alone or the NLCD 2011. The accuracies of mixed and evergreen categories were significantly improved with the inclusion of EVI Index compared to using PCA technique alone. Comparing the accuracy report of NLCD 2011, the accuracy of mixed category was significantly improved in both producer's and user's accuracy. Deciduous forest had higher accuracies among all forest types. With the inclusion of EVI Index, the producer's accuracy of deciduous forest was 97%.

Forest restocking on previous cropland or pasture has resulted in ecological succession. Evergreen forests made of mostly eastern red cedar was the first forest to establish. At present time, about 70% of evergreen forest is found in previously disturbed area (historic cropland and pasture). Later succession of vegetation would depend on the habitat types. Overlay of the habitat model result and vegetation type showed that about 40% of the evergreen forests are located in Xeric or Sub-Xeric habitats. They are likely to prevail while the remaining 60% of evergreen forests will likely be succeed by deciduous forests in the future.

There are some limitations of habitat modeling as well as vegetation mapping in this study. The result of habitat modeling was restricted by the precision of the geology map, the criteria that distinct different habitat types, and the resolution of the model. In the Park, Acid Xeric habitat type occurs at specific locations with steep sandstone cliff, usually indicated by a pine stand on top, which the model missed. In addition, the transition zone between different habitat types is hard to identify due to previous disturbance.

The following factors have limited the results of vegetation mapping: 1) spatial resolution of Landsat-8 OLI imagery has made the distinction of mixed forests from evergreen forests difficult; 2) many areas in the shadow as a result of topography and sun angle were hard to classify, regardless of sun angle correction; 3) even with the inclusion of EVI Index, spectral confusion remain among mixed, deciduous, and evergreen forest pixels; and 4) the limitation of field reference data due to area accessibility.

Vegetation habitat types provide baseline data set for the development of successional plant community classification for Mammoth Cave National Park (Cooper et al., 1991). Habitat also provides a natural plant stratification within the Park area (Cooper et al., 1991). Furthermore, it acts as a means of predicting both site quality and response following disturbance (Cooper et al., 1991). The fire-vulnerable habitat types account for one-fifth of the total Park area and the rest are fire-resistant habitat types. The results of habitat modeling and vegetation mapping from this study provide up-to-date information for fire management planning and resource management at Mammoth Cave National Park.

Reference Cited

Abdi, H. and Williams, L. J., 2010. Principal component analysis. WIREs Comp Stat, 2: 433–459. doi:10.1002/wics.101.

Armesto, J., and Martínez, J., 1978. Relations Between Vegetation Structure and Slope Aspect in the Mediterranean region of Chile. *The Journal of Ecology* 66(3): 881-889. doi:10.2307/2259301

Arora, V., 2002. Modeling vegetation as a dynamic component in soil-vegetation-atmosphere transfer schemes and hydrological models. *Rev. Geophys*., 40, 1-25.

Badger, K., Taylor, J., Jones, B., and Shell, M. 1997. Mammoth Cave National Park Forest Vegetation Study. Cooperative Agreement No. CA-5530-3-9001, Subagreement No. CA-5530-3-9003. Ball State University, Muncie, Indiana, 1-7.

Ball, G. H. and Hall, D. J., 1965. *ISODATA, a novel method of data analysis and pattern classification: technical report*. Menlo Park, Calif, Stanford Research Institute.

Beauchaine, T. P. and Beauchaine III, R. J., 2002. A comparison of maximum covariance and K-means cluster analysis in classifying cases into known taxon groups. *Psychological methods*, *7*(2), p.245.

Chang, K. and Tsai, B., 1991. The effect of DEM resolution on slope and aspect mapping. *Cartography and Geographic Information Systems*, 18: 69–77.

Choi, Y. H., Kim, H. K., Hazekamp, A., Erkelens, C., Lefeber, A. W., Verpoorte, R., 2004. Metabolomic differentiation of Cannabis sativa cultivars using 1H NMR spectroscopy and principal component analysis. *Journal of Nature Products*, 67: 953–957.

Chow, T. E. and Hodgson, M. E., 2009. Effects of lidar post-spacing and DEM resolution to mean slope estimation. *International Journal of Geographical Information Science*, *23*(10): 1277-1295.

Congalton, R. G., 1991. A Review of Assessing the Accuracy of Classifications of Remotely Sensed Data. *Remote Sensing of Environment*, 37:35-46.

Cooper, S. V., Neiman, K. E. and Roberts, D. W., 1991. Forest habitat types of northern Idaho: a second approximation. Rev. April 1991. *General technical report INT (USA). no. 236.*

Crawford, M. M., Olson, R. A., Toomey R. S., Scoggins. L. J., 2013. Geology of Mammoth Cave National Park, Kentucky. Kentucky Geological Survey, Mammoth Cave National Park, and Mammoth Cave National International Center for Science and Learning. Electronic document, http://kgs.uky.edu/kgsweb/olops/pub/kgs/mcs186_12.pdf. Last accessed June 20, 2017.

Desta, F., Colbert, J. J., Rentch, J. S. and Gottschalk, K. W., 2004. Aspect induced differences in vegetation, soil, and microclimatic characteristics of an Appalachian watershed. Castanea, 69(2): 92-108.

Douville, H., Planton, S., Royer, J. F., Stephenson, D. B., Tyteca, S., Kergoat, L., Lafont, S., and Betts, R. A., 2000. Importance of vegetation feedbacks in doubled-CO2 climate experiments. *Journal of Geophysical Research,* 105(D11): 14841–14861. doi:10.1029/1999JD901086.

Duda, R. O., Hart, P. E. and Stork, D. G., 2001. *Pattern Classification*, 2nd ed. (New York: John Wiley & Sons).

Egbert, S. L., Park, S., Price, K.P., Lee, R. Y., Wu, J. and Nellis, M. D. 2002. Using conservation reserve program maps derived from satellite imagery to characterize landscape structure. *Comput Electron Agric* 37(1):141–56.

Ellsworth, I. 1934. Forest cover type map survey of proposed Mammoth Cave National Park. United States Department of the Interior, Office of National Parks, Buildings and Reservations, p.2.

Ellsworth, I. 1936. Forest cover type map of proposed Mammoth Cave National Park. United States Department of the Interior, Office of National Parks, Buildings and Reservations.

Eugster, W., Rouse, W. R., Pielke, S. R. A., Mcfadden, J. P., Baldocchi, D. D., Kittel, T. G. F., Chapin, F. S., Liston, G. E., Vidale, P. L., and Vaganov, E., 2000. Land-atmosphere energy exchange in Arctic tundra and boreal forest: Available data and feedbacks to climate. *Glob. Chang. Biol.*, 6, 84–115.

Evans, M. 1991. Kentucky Ecological Communities. Draft. Kentucky State Nature Preserves Commission, Frankfort, KY.

Faller, A. and M. Jackson., 1975. The Plant Ecology of Mammoth Cave National Park, Kentucky. Indiana State University, Terre Haute, Indiana, 41, 126.

Gould, S. B., Glenn, N. F., Sankey, T. T. and McNamara, J. P., 2013. Influence of a dense, low-height shrub species on the accuracy of a LiDAR-derived DEM. *Photogrammetric Engineering & Remote Sensing*, *79*(5): 421-431.

Green, D. and Hartley, S., 2000. Integrating Photointerpretation and GIS for Vegetation Mapping: Some Issues of Error. In *Vegetation Mapping: from Patch to Planet*, R. Alexander and A. C. Millington (ed.), Wiley, 2000: 103-134.

Guo, J., Chen, Q., Wang, C., Qiu, H., Liu, B., Jiang, Z., and Zhang, W., 2015. Comparison of two exploratory data analysis methods for classification of Phyllanthus chemical fingerprint: unsupervised vs. supervised pattern recognition technologies, *Analytical & Bioanalytical Chemistry*, 407, 5: 1389-1401.

Hadjimitsis, D. G., Papadavid, G., Agapiou, A., Themistocleous, K., Hadjimitsis, M. G., Retalis, A., Michaelides, S., Chrysoulakis, N., Toulios, L., and Clayton, C. R. I., 2010. Atmospheric correction for satellite remotely sensed data intended for agricultural applications: impact on vegetation indices. *Natural Hazards and Earth System Sciences*, 10(1): 89-95.

Hambly, F. 1966. Mammoth Cave National Park Resources Management Plan, p.35.

He, C., Zhang, Q., Li, Y., Li, X. and Shi, P., 2005. Zoning grassland protection area using remote sensing and cellular automata modeling—a case study in Xilingol steppe grassland in northern China. *Journal of Arid Environments*, *63*(4): 814-826.

Holland, P.G. and Steyn, D. G., 1975. Vegetational responses to latitudinal variations in slope angle and aspect. *Journal of Biogeography*. 2:179-183.

Homer, C. H., Fry, J. A., and Barnes C. A., 2012. The National Land Cover Database, U.S. Geological Survey Fact Sheet 2012-3020, p.4.

Hölzel, N., Buisson, E., and Dutoit, T., 2012. Species introduction—a major topic in vegetation restoration. *Applied Vegetation Science*, 15: 161–165.

Huang, K., 2002. A Synergistic Automatic Clustering Technique for Multispectral Image Analysis, *Photogrammetric Engineering & Remote Sensing*, 68(1): 33-40.

Huete, A., Liu, H., Batchily, K., and Van Leeuwen, W., 1997. A comparison of vegetation indices over a global set of TM images for EOS-MODIS. *Remote Sensing of Environment, 59*: 440–451.

Huete, A. R. and Justice, C., 1999. MODIS vegetation index (MOD13) algorithm theoretical basis document. Ver. 3.

Huete, A., Didan, K., Miura, T., Rodriguez, E., Gao, X., and Ferreira, L., 2002. Overview of the radiometric and biophysical performance of the MODIS vegetation indices. *Remote Sensing of Environment, 83*: 195–213.

Jiao, J., Tzanopoulos, J., Xofis, P., and Mitchley, J., 2008. Factors affecting distribution of vegetation types on abandoned cropland in the hilly-gullied Loess Plateau Region of China. *Pedosphere*, 18: 24–33.

Jolliffe I, 2005. Principal component analysis. Encyclopedia of Statistics in Behavioral Science.

Kelly, M., Tuxen, K. A., and Stralberg, D., 2011. Mapping changes to vegetation pattern in a restoring wetland: Finding pattern metrics that are consistent across spatial scale and time. *Ecol. Indic*., 11: 263–273.

Kerr, J. T. and Ostrovsky M., 2003. From space to species: ecological applications for remote sensing. *Trends in Ecology & Evolution*, 18(6): 299-305.

Kwarteng, P. S. and Chavez, A.Y., 1989. Extracting spectral contrast in Landsat Thematic Mapper image data using selective principal component analysis. *Photogrammetric Engineering & Remote Sensing, 55*: 339-348.

Lefsky, M. A., Cohen, W. B., Parker, G. G., and Harding, D. J., 2002. Lidar remote sensing for ecosystem studies. *BioScience*. 52, 1: 19-30.

Loveland, T. R., Zhiliang, Z., Ohlen, D. O., Brown, J. F., Reed, B. C and Yang L., 1999, "An Analysis of the IGBP Global Land Cover Characterization Process," *Photogrammetric Engineering & Remote Sensing*, 65(9): 1021-1032.

Lu, Y., Gao, B., Chen, P., Charles, D., and Yu, L. L., 2014. Characterisation of organic and conventional sweet basil leaves using chromatographic and flow-injection mass spectrometric (FIMS) fingerprints combined with principal component analysis. *Food Chemistry*, 154: 262–268.

Mari, A., Montoro, P., Pizza, C., and Piacente, S., 2012. Liquid chromatography tandem mass spectrometry determination of chemical markers and principal component analysis of Vitex agnus-castus L. fruits (Verbenaceae) and derived food supplements. *J Pharm Biomed Anal*,70:224–230.

Marino, E., Ranz, P., Tomé, J. L., Noriega, M. Á., Esteban, J. and Madrigal, J., 2016. Generation of high-resolution fuel model maps from discrete airborne laser scanner and Landsat-8 OLI: A low-cost and highly updated methodology for large areas. *Remote Sensing of Environment*, 187: 267-280.

Masek, J. G., Vermote, E. F., Saleous, N., Wolfe, R., Hall, F. G., Huemmrich, F., Gao, F., Kutler, J., and Lim, T. K., 2006. A Landsat surface reflectance data set for North America, 1990-100, *IEEE Geoscience and Remote Sensing Letters*, 3: 68-72.

Morgan, J. L., Gergel, S. E. and Coops, N. C., 2010, Aerial Photography: A Rapidly Evolving Tool for Ecological Management. *In BioScience*, 60, 1: 47-59.

Mukherjee, A., and Lal, R., 2014. Comparison of Soil Quality Index Using Three Methods. *PLoS ONE* 9(8): e105981. doi:10.1371/journal.pone.0105981.

Nie, Y., Kafatos, M., and Wood, K., 2001. Estimating soil-type pattern from supervised and unsupervised classification: Case study in Cuprite, Nevada. Project Report, pp. 28. George Mason University.

Nagase, A., and Dunnett, N., 2012. Amount of water runoff from different vegetation types on extensive green roofs: effects of plant species, diversity and plant structure. *Landscape and Urban Planning*, 104: 356–363.

National Park Service. 2015. Geologic Formations – Mammoth Cave. Electronic document, https://www.nps.gov/maca/learn/nature/geologicformations.htm. Last accessed June 19, 2017.

Olson, R., Franz, M., and Ghitter, G., 2000. A vegetation map of Mammoth Cave National Park using satellite remote sensing data. In Proceedings of the 8th Mammoth Cave Science Conference, Division of Science and Resource Management, Mammoth Cave National Park, Mammoth Cave, Kentucky, USA., 61–68.

Olson, R. and Noble, C., 2005. The Geological Foundation for Prescribed Fire in Mammoth Cave National Park. *Geodiversity & Geoconservation*, *22*(3): 22-28.

Olson, R., Scoggins, L., Toomey, R., and Burton, J., 2013. 2011 Vegetation Map for Mammoth Cave National Park. Mammoth Cave Research Symposia. p. 4. Electronic document,

http://digitalcommons.wku.edu/mc_reserch_symp/10th_Research_Symposium_2013/Research_Posters/4, accessed June 18th, 2017.

Pinder, J. E., Kroh, G. C., White, J. D. and May, A.B., 1997. The relationships between vegetation type and topography in Lassen Volcanic National Park. *Plant ecology*, *131*(1): 17-29.

Qiu, L., Zhang, X., Cheng, J., and Yin, X., 2010. Effects of black locust (Robinia pseudoacacia) on soil properties in the loessial gully region of the Loess Plateau, China. *Plant and Soil*, 332: 207–217.

Räsänen, A., Kuitunen, M., Tomppo, E., and Lensu, A., 2014. Coupling high-resolution satellite imagery with ALS-based canopy height model and digital elevation model in object-based boreal forest habitat type classification. *ISPRS Journal of Photogrammetry and Remote Sensing*, 94: 169-182.

Richards, J. A., 1993. Remote sensing digital image analysis: an introduction (second edition), p 231.

Stage, A. R. and Salas, C., 2007. Interactions of elevation, aspect, and slope in models of forest species composition and productivity. *Forest Science*, *53*(4): 486-492.

Suzuki, R., Masuda, K., and Dye, D. G, 2007. Interannual covariability between actual evapotranspiration and PAL and GIMMS NDVIs of northern Asia. *Remote Sensing of Environment*, 106: 387–398.

Tomppo, E., Haakana, M., Katila, M., and Peräsaari, J., 2008. Multi-source national forest inventory – methods and applications. Managing Forest Ecosystems 18. *Springer*. 374 p. ISBN 978-1-4020-8712-7.

Turner, W., Spector, S., Gardiner, N., Fladeland, M., Sterling, E., and Steiniger, M., 2003. Remote sensing for biodiversity science and conservation. *Trends in Ecology & Evolution*,18(6): 306-314.

USGS, 2016. Landsat 8 (L8) data users handbook. Electronic document, https://landsat.usgs.gov/sites/default/files/documents/Landsat8DataUsersHandbook.pdf, accessed May 15th, 2017.

USGS, 2017. Landsat Surface Reflectance-Derived Spectral Indices. Corrected typographical error in EVI formula. Electronic document, https://landsat.usgs.gov/sites/default/files/documents/si_product_guide.pdf, accessed May 23rd, 2017.

Wang, L., Wei, S., Horton, R., and Shao, M. A., 2011. Effects of vegetation and slope aspect on water budget in the hill and gully region of the Loess Plateau of China. *Catena*, 87: 90–100.

Wulder, M., Hall, R. J., Coops, N. C. and Franklin, S. E., 2004. High Spatial Resolution Remotely Sensed Data for Ecosystem Characterization. *In BioScience*, 54, 6: 511-521.

Xiao, X. M., Zhang, Q. and Braswell, B., 2004. Modeling gross primary production of temperate deciduous broadleaf forest using satellite images and climate data. *Remote Sensing of Environment*, 91: 256–70.

Xiao, Q. and McPherson, E. G., 2005. Tree health mapping with multispectral remote sensing data at UC Davis, California. *Urban Ecosystems*, 8(3): 349-361.

Xie, Y., Sha, Z. and Yu, M., 2008. Remote Sensing Imagery in Vegetation Mapping: A Review. *In Journal of Plant Ecology,* 1, 1: 9-23.

Zhong, Y., Zhang, L. and Gong, W., 2011. Unsupervised remote sensing image classification using an artificial immune network. *International journal of remote sensing*, *32*(19): 5461-5483.

Appendix I

Field notes for 30 selected sites for the accuracy assessment of habitat modeling result

Sample#	ID#	Slope	Aspect	Bedrock	Habitat	Species
1	352	7°	285°	Sandstone	Acid Mesic	Post oak, sugar maple, pignut hickory, white ash, maple trees
2	557	20°	245°	Sandstone	Acid Sub-Xeric	White oak, pignut hickory, thinner trees over story and under tulip popular and black gum
3	463	25°	43°	Limestone	Calcareous Supra-Mesic	Green ash, turnip popular dominant, sugar maple, middle canopy pawpaw, ferns, spice bush, maidenhair fern
4	157	0°	0°	Sandstone	Acid Mesic	Chestnut oak, white ash, red oak, no ferns
5	152	3°	327°	Sandstone	Acid Mesic	East red cedar, sugar maple, blackberry, southern red oak, white oak, old field, cedar succession mixed, Christmas fern
6	295	20°	190°	Sandstone and limestone boundary	Acid Sub-xeric	Southern red oak, black oak, sugar maple, white ash, Christmas fern
7	185	16°	45°	Limestone	Calcareous Mesic	Black oak, cedar maple, southern red oak, Christmas fern.
8	434	19°	100°	Limestone	Calcareous Mesic	Large and medium tulip popular, red oak, American chestnut, muscle beech, American columbo.
9	228	27°	68°	Sandstone	Acid Mesic	White oak(several), dogwood, black gum (understory), pignut hickory, red sweet gum
10	521	24°	231°	Sandstone	Acid Sub-Xeric	White oak, pignut hickory, sugar maple.
11	353	4°	215°	Limestone	Calcareous Sub-Mesic	Post oak, black oak, pignut hickory, ostrya southern fern
12	258	20°	351°	Sandstone	Acid Supra-Mesic	Big white oak dominate, black oak, young tulip popular, sugar maple, beech, spice bush, fern. Young tulip popular (many of them are under story), sassafras
13	95	3°	154°	Sandstone	Calcareous Sub-Mesic	Upland swamp, southern red oak, black gum, black oak, winged elm, tulip popular, red maple, Christmas fern
14	383	32°	354°	Limestone covered with sandstone	Calcareous Mesic	Beech, tulip popular, red oak, sugar maple, fern, spicebush

15	156	4°	189°	Limestone	Calcareous Sub-Mesic	Old field, cedar, beech, black cherry, black oak, sugar maple, black gum, sourwood, dogwood, ferns
16	22	28°	154°	Sandstone	Acid Sub-Xeric	White oak, red maple, scarlet oak, black oak, sugar maple, black gum, sourwood, dogwood, ferns
17	437	6°	4°	Limestone	Calcareous Sub-Mesic	Old field, Virginia pine, red cedar, dogwood, sparkle berry, southern red oak
18	296	25°	45°	Limestone	Calcareous Supra-Mesic	Tulip popular, white oak, shagged-bark hickory, sugar maple, (possible) shell-bark hickory, pawpaw, maidenhair fern, mayapple, wild yam, Solomon seal, bed strawl
19	429	13°	195°	Limestone	Calcareous Xeric	Eastern red cedar (not successional), not very tall, chinkapin oak, shagged-bark hickory, blue ash, erosion mark on Calcareous rock, rillenkarren, sugar maple should not be here
20	100	17°	229°	Sandstone and limestone boundary	Acid Sub-xeric	Black oak, white oak, dogwood, cedar, scarlet oak, pignut hickory, red bud seeding, Solomon seal, grape vine
21	550	26°	172°	Sandstone	Acid Sub-Xeric	White oak, pignut hickory, tulip popular, black walnut, black gum, Christmas fern, beech(small) trillium, Indian cup.
22	16	27°	334°	Limestone	Calcareous Supra-Mesic	Horse chestnut, green and white ash, sugar maple, red oak, spice bush, dense spice bush shrubs, wild ginger, wingstem, several kinds of ferns
23	178	10°	237°	Limestone	Calcareous Xeric	Shagged-bark hickory, white ash, Chinkapin oak, blue ash, cedar (eastern red), (chestnut oak sapping), sugar maple (small), rock cress
24	456	18°	169°	Limestone	Calcareous Sub-Xeric	Chinkapin oak, cedar, red bud, shagged-bark hickory, sugar maple, unknown shrub, boundary on Calcareous Xeric
25	233	18°	179°	Sandstone	Acid Sub-Xeric	Ferns, cedar, scarlet oak, black gum, beech(lack of fire), southern red oak, pignut hickory, California buckthorn
26	361	7°	127°	Limestone	Calcareous Sub-Xeric	Cedar, sugar maple, chinkapin oak, dead white ash trees (6), red bud.

						One live white ash, blueberry bush, American columbo
27	555	24°	248°	Sandstone	Acid Sub-Xeric	White oak, tulip popular, sour wood, sugar maple, little under story, not steep enough, not xeric. Xeric would be on steep cliff pine stand
28	448	14°	56°	Limestone	Calcareous Mesic	Tulip popular over story with pignut hickory, dogwood, muscle beech, southern red bud, Christmas fern, American columbo, native iris, mayapple
29	402	Artificial cliff by road cut, no good				
30	122	11°	180°	Limestone	Calcareous Sub-Mesic	Eastern red cedar, white oak, chinkapin oak, old field, ash, mayapple, Christmas fern, spice bush.