2016

Master's Project: Describing Forest Structure in Southern Blue Ridge Cove Forests: A LiDAR-Based Analysis

Jamie Ervin
University of Vermont

Describing Forest Structure in Old Growth Southern Blue Ridge Cove Forests

A LiDAR-Based Analysis

Project by:
Jamie Ervin
Ecological Planning Program
University of Vermont
October 2016

Prepared for:
Mountaintrue
29 North Market Street, Suite 610
Asheville, NC 28801
(828)258-8737

The UNIVERSITY of VERMONT
RUBENSTEIN SCHOOL
OF ENVIRONMENT AND NATURAL RESOURCES

mountaintrue.
Protecting the Places We Share.

Table of Contents

List of Maps

List of Figures

List of Tables

Executive Summary

The Southern Blue Ridge Mountains contain some of the largest contiguous patches of old growth forests in the eastern United States. Known for both their beauty and biodiversity, these forests provide rare insight into what the Appalachian landscape may have looked like prior to Euro-American settlement. Mountaintrue – a North Carolina-based nonprofit conservation organization – supports ecological restoration as a guiding principle in forest management. For nearly three decades, Mountaintrue (formerly the WNC Alliance) has been a prominent local voice for protecting old growth forests on public lands.

Cove forests occupy mid-elevation slopes in concave, mesic coves throughout the Blue Ridge region. Due in part to their sheltered topography and low fire return interval, major natural disturbances are less common in cove forests than nearly any other Appalachian forest community. As a result, these forests tend to reach a late-successional old growth stage if left unlogged. Many of the region's best-known, most charismatic old growth forest stands lie in cove forests at sites like Joyce Kilmer Memorial Forest and Great Smoky Mountains National Park.

Light Detection and Ranging (LiDAR) data has become a popular tool for analyzing forest structure over large landscapes. Publicly-available, high-quality LiDAR data now exists for much of the Southern Blue Ridge. Mountaintrue has begun incorporating LiDAR analysis into their forest advocacy work, and is interested in exploring new methods for using the data to study old growth forests.

This project report is a compilation of research collected for Mountaintrue throughout the summer and fall of 2016. LiDAR-derived canopy height models are applied to existing old growth cove forest stands in Pisgah and Nantahala National Forests, and Great Smoky Mountains National Park. A unique method for studying these forests' horizontal structure is developed using grey-level co-occurrence texture statistics. Products of the project include:

- A collection of charts comparing horizontal forest structure in areas with different anthropogenic disturbance histories in Great Smoky Mountains National Park
- An analysis of statistical differences between texture values in old growth and second growth cove forests
- A collection of regression models which identify occurrences of old growth cove forests throughout the study area with ~60% accuracy based on LiDAR canopy height
- A new simple method for studying horizontal canopy heterogeneity using LiDAR canopy height models and texture-based neighborhood statistics

These results will aid Mountaintrue's work to identify old growth forests remotely, and might support future ecological restoration and forest management work in cove forests. The report concludes with a series of recommendations which may improve upon the methods used in this study.

List of Common Abbreviations:

CHM: Canopy Height Model

Ecozones: Ecological Zone Models

GLCM: Grey-Level Co-Occurrence Matrix

GSMNP: Great Smoky Mountains National Park

LiDAR: Light Detection and Ranging

TN: Tennessee

WNC: Western North Carolina

Section 1: Background

A. Study Area

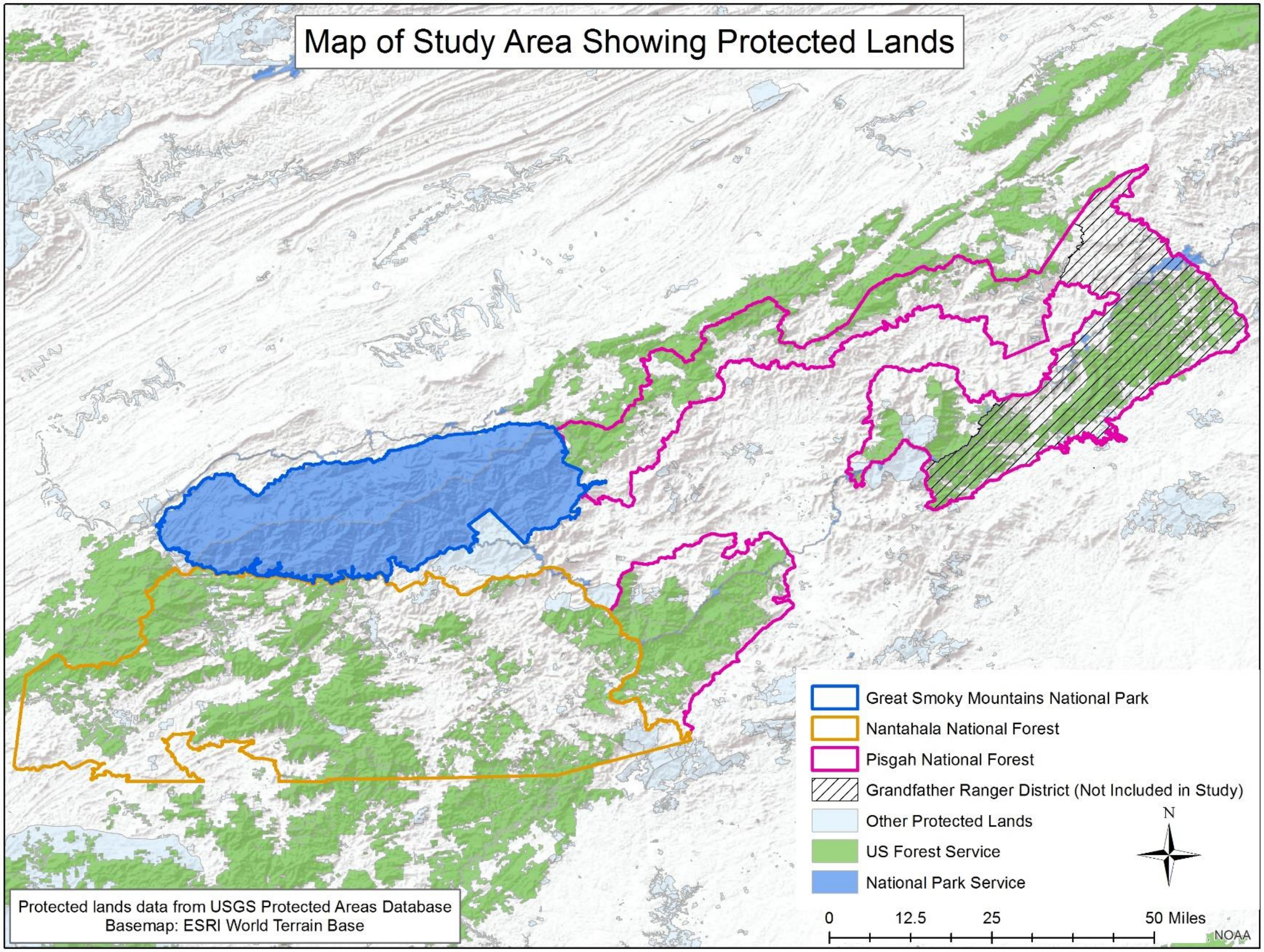
This analysis focuses on three large areas of public land – Great Smoky Mountains National Park, Pisgah National Forest, and Nantahala National Forest – within the Blue Ridge region of the Southern Appalachian Mountains. Pisgah and Nantahala National Forests both lie entirely in Western North Carolina, while Great Smoky Mountains National Park straddles the border of North Carolina and Tennessee. The study area contains a collection of major mountain ranges and geographic features including the Black Mountains, Unaka Mountains, Unicoi Mountains, Great Balsam Mountains, and the Great Smoky Mountains.

The Southern Appalachian landscape is characterized by high spatial variability in habitat conditions. Elevations across the study area range from around 1,300 feet above sea level along the French Broad River near Hot Springs, NC to 6,683 feet at the summit of Mount Mitchell, the highest peak in the Appalachian Mountains. The region receives the highest annual rainfall in the continental US east of the Cascades (SAMAB 1996). As a result, the Southern Appalachians are considered to be one of the most biodiverse temperate regions on the planet (SAMAB 1996). It is estimated that approximately 4,000 plant species occur within the region, including nearly 400 rare species, and over 250 endemics (The Nature Conservancy 2000). The region contains the highest diversity of salamander species anywhere on the planet (Petranka 1998). The study area is shown is Map 1.

B. Project Sponsor

Mountaintrue is a regional environmental nonprofit, whose mission is "championing resilient forests, clean waters and healthy communities in Western North Carolina." The organization formed in early 2015 from the merging of the Western North Carolina Alliance (Buncombe County), the Environmental and Conservation Organization (Henderson County), and the Jackson-Macon Conservation Alliance (Macon County). Mountaintrue achieves its mission through grassroots advocacy aimed at strengthening community engagement and influencing policy.

Mountaintrue's is a leading advocate for the sustainable management of public forests throughout Western North Carolina. One of the organization's (then the WNC Alliance) earliest campaigns focused on halting clearcutting practices in Pisgah and Nantahala National Forests. This early (late 1980s) "Cut the Clearcutting" campaign cemented Mountaintrue's status as a stakeholder in the management of WNC's public lands (Newfont 2012). The organization has since remained an active participant in the public lands planning process by providing input on timber sales, controlled burns, trail construction, and other management activities. Mountaintrue is currently providing input on the revision of the Land and Resource Management Plan for Pisgah and Nantahala National Forests, which will guide the management of the forests for the next 20 years. Throughout their public lands work, Mountaintrue has given specific attention to protecting old growth forests.



Map 1: Map of Study Area

C. Defining Old Growth and Primary Forests
The terms *old growth forest* and *primary forest* are used throughout this report to represent similar, but different aspects of forest disturbance history:

Primary Forest refers to forests that have never been logged or impacted by other anthropogenic land uses that alter the structure of the forest canopy. A forest's designation as primary reflects its *naturalness* rather than its *age*. This implies that a young forest, recently disturbed by a stand-replacing natural event like a fire or landslide is considered a primary forest while a mature stand recovering from a logging operation 200 years ago is not.

Old Growth Forest refers to forests that have reached the latest stages of stand development without experiencing major stand-replacing disturbances. Unlike primary forests, which are defined purely by the absence of major anthropogenic disturbances throughout their history, old growth forests are distinguished by their *age* and *structure*. This definition implies that a forest stand can regrow into an old growth state after being logged. The age at which old growth develops varies by forest community type, site conditions, and natural disturbance regime. Old growth forests are often distinguished from earlier stages of forest succession by characteristics like large trees for species and site, wide variation in tree size and spacing, presence of standing dead snags, multiple canopy layers, and the presence of canopy gaps (US Forest Service 1997).

The distinction between old growth and primary forests is important throughout this report because certain sections of the project focus on different aspects of forest disturbance history. For example, Section 5-A, which examines canopy height texture statistics in different areas of GSMNP, compares *primary* forests to second growth forests that experienced varying levels of human disturbance. By contrast, Section 5-E uses regression modeling to attempt to classify *old growth* cove forests based on their structure as represented in LiDAR canopy height models. These distinctions are noted throughout the report, and the classification "old growth/primary forest" is used in sections of the project where both primary and old growth forests were used for data analysis. This inconsistency results from a lack of true old growth field surveys for the entire study area. Specifically, the best available data for GSMNP identifies primary forests, whereas the best available data for Pisgah and Nantahala National Forests incorporates late-successional forest structure into its old growth definition and ranking system (see Section 2-B for a description of these sources).

D. Primary Forests in the Southern Appalachians
The Southern Appalachian Mountains contain some of the largest contiguous patches of primary forests in the eastern United States (Davis 2003). Of these, historical records and field surveys indicate that around 115,000 acres lie in Great Smoky Mountains National Park, while another ~90,000 acres lie scattered throughout Pisgah and Nantahala National Forests in Western North Carolina (Pyle 1988; Messick 2000). Southern Appalachian forests experienced extensive land clearing throughout the late 19th and 20th centuries. Many of the region's unlogged sites lie in the most rugged, remote parts of the region, where topography or underlying geology would have made access difficult for logging (Messick 2000). These remaining examples of primary forests serve as an important reference for what the structure, composition, and biodiversity of Southern Appalachian Forests may have been prior to Euro-American settlement. Map 2 shows known primary forests throughout the study area.

No forests in the Southern Appalachians have escaped human impact entirely. Exotic pests and pathogens like the chestnut blight (*Cryphonectria parasitica*) and the hemlock woolly adelgid (*Adelges tsugae*), both introduced to the region by humans, have dramatically altered the composition of Southern Appalachian forests. Similarly, fire suppression has indirectly altered the structure of many forest communities in the region (Flatley et al. 2015).

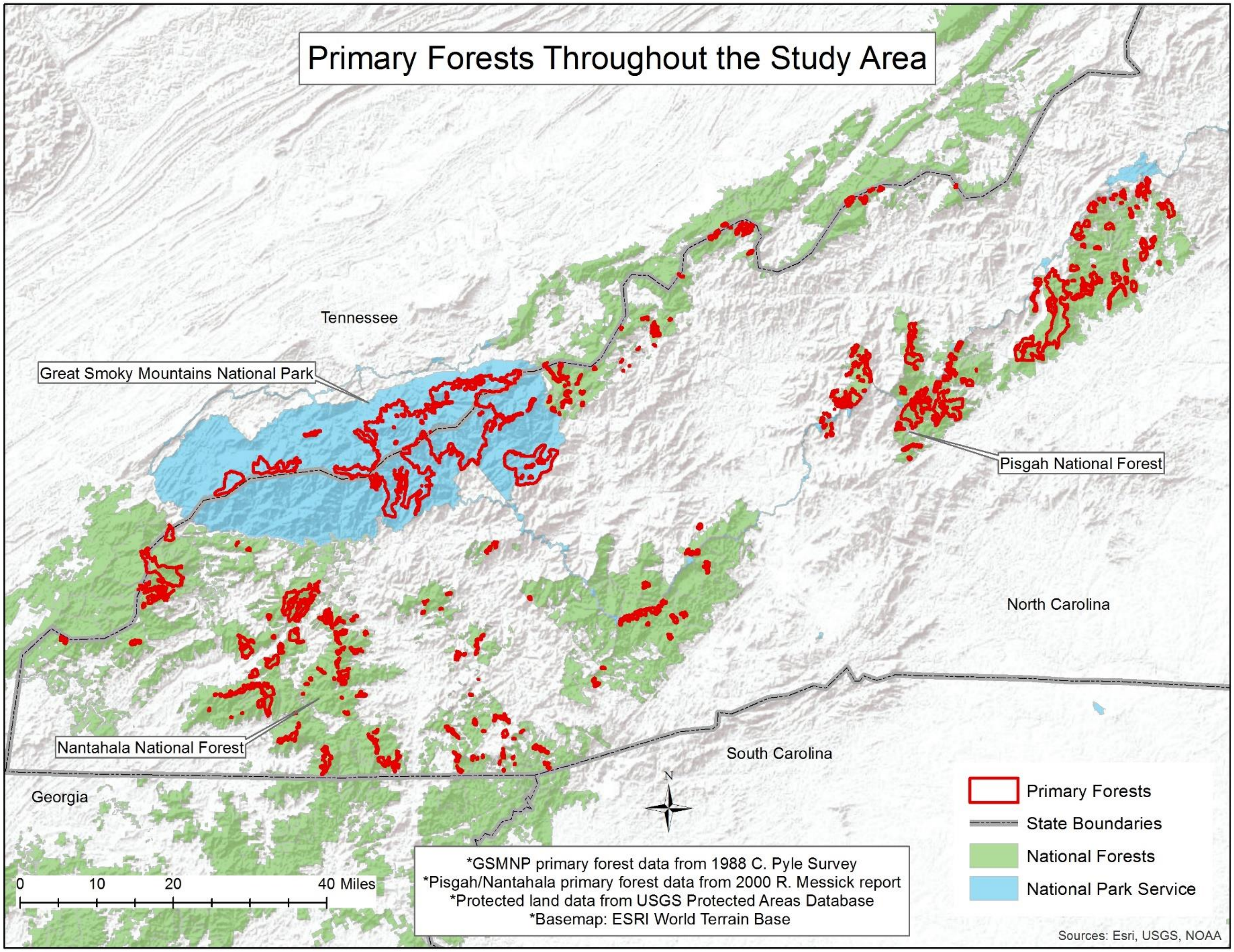
E. Southern Appalachian Cove Forests

Cove forests occupy mid-elevation (2,000-4,500 feet) slopes within concave mesic "coves" throughout the Southern Appalachian region. On an elevational gradient, cove forests represent the transition zone between higher-elevation northern hardwood forests and lower-elevation oak hickory forests. The transition between cove forests and these other communities tends to be gradual, and cove forests commonly form a mosaic with more pyrogenic oak-hickory forests outside of the most sheltered areas (Landfire 2007).

Cove forests are distinguished by a wide variety of mesophytic canopy species, with common components including tulip poplar (*Liriodendron tulipifera*), eastern hemlock (*Tsuga canadensis*), American basswood (*Tilia Americana var. heterophylla*), sugar maple (*Acer saccharum*), yellow buckeye (*Aesculus flava*), red oak (*Quercus rubra*), and white oak (*Quercus alba*). Cove forest communities are expressed in two distinct associations – rich and acidic – depending on their substrate. Rich coves are characterized by fertile soils and a diverse herbaceous layer, while acidic coves contain a dense evergreen shrub layer composed mainly of acid-tolerant species in the Heath family including rosebay rhododendron (*Rhododendron maximum*) and dog hobble (*Leucothoe fontanesiana*) (Schafale 2012).

Dense rhododendron in an acidic cove forest

Herbaceous layer in a rich cove forest



Map 2: Primary Forests in Study Area

E. Forest Structure and Biodiversity in Old Growth Cove Forests
The US Forest Service describes old growth cove forests as having uneven, or all-aged structures, and irregular diameter distributions (US Forest Service 1997). Extensive research has been done comparing long-term forest stand dynamics in old growth and second growth cove forests. Increment cores taken in Joyce Kilmer Memorial Forest revealed all-aged structures, with some individuals exceeding 400 years of age in 1976 (Lorimer 1976). Though fires do affect cove forests (Flatley et al. 2015), wind-throw and individual tree mortality commonly shape the landscape in old growth coves, creating a late-successional mosaic of canopy gaps and mature trees representing a spectrum of age classes (Runkle et al. 1987; Lorimer 1980). The death of individual trees creates gaps throughout the forest canopy, which play an important role in these forests' structure and composition over the long term. Canopy gaps allow for the regeneration of shade-intolerant species like tulip poplar (*Liriodendron tulipifera*) to regenerate with enough regularity that they remain a components of late-successional forests (Busing 1995). A 1982 field study found that canopy gaps covered approximately 9.5% of the area in a series of old growth cove forests, and that new gaps were formed at a rate of approximately 1% of the landscape per year (Runkle 182). In the Great Smoky Mountains, second growth cove forests have been shown to have smaller, more numerous canopy gaps, while their old growth counterparts house fewer, larger canopy gaps with more diversity in light conditions (Clebsch and Busing 1989).

Old growth cove forests hold ecological value as sites of high biodiversity. Herbaceous layers in primary cove hardwood forests have been found to host up to 14 plant species per square meter in the spring (Duffy and Meier 1992). Higher levels of species richness and abundance have been recorded in the herbaceous layers of old growth forests, even when compared with sites logged 100-150 years ago (Wyatt 2009). Some structural features which are common to many old growth forest communities, like standing dead snags and multiple canopy layers, are present in cove forests and also influence biodiversity. For example, old growth hemlock forests in Pennsylvania have been found to host higher levels of neotropical migratory bird territories, with some species like blackburnian warblers (*Dendroica fusca*), black-throated green warblers (*D. virens*), magnolia warblers (*D. magnolia*), solitary vireos (*Vireo solitarius*), and Swainson's thrush (*Catharus ustulatus*) present in dramatically higher numbers compared with second growth forests (Haney and Schaadt 1996). Many carnivores, including black bears (*Ursus americanus*), gray fox (*Urocyon cinereoargenteus*), and bobcat (*Lynx rufus*) use common old growth features like root masses of downed trees and fallen hollow logs for shelter (Pelton 1996).

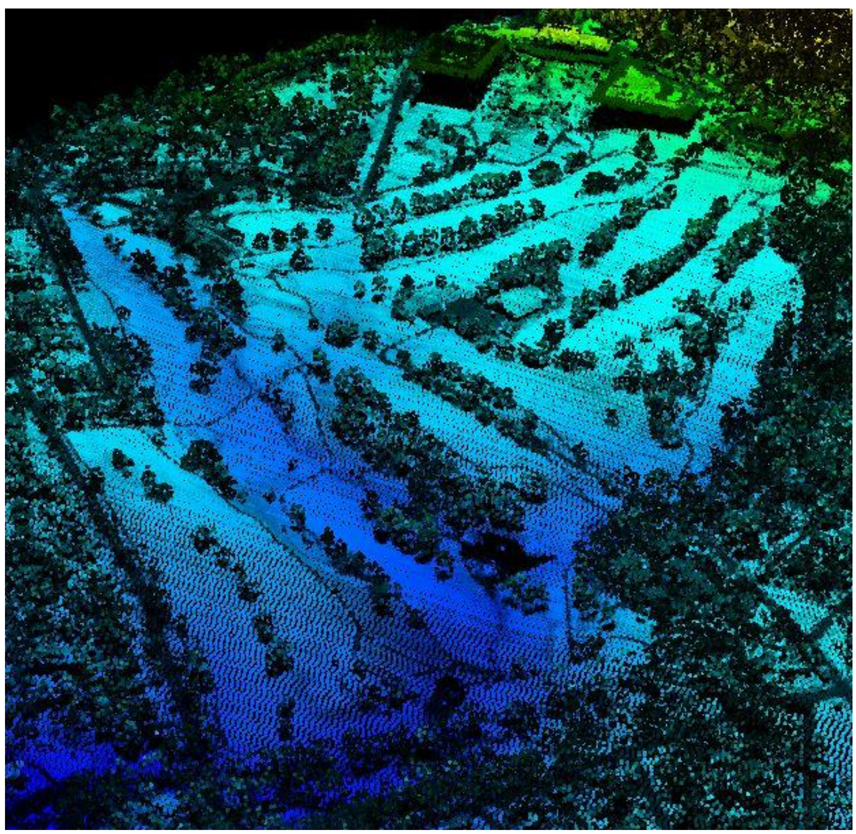
F. Why Focus on Cove Forests?
This report focuses specifically on cove forests because of their long-term disturbance regime and their prevalence throughout the study area. Due in part to their sheltered topography and mesic condition, major natural disturbances like fire are rare in cove forests. This absence of regular major natural disturbances makes cove forests less structurally variable throughout their natural range of variability when compared with more disturbance-prone communities like pine-oak heath or chestnut oak forests (Landfire 2007).

The largest concentration of old growth cove forests lies in GSMNP. As of this writing, no comprehensive inventory of old growth forests exists for the park, but a historic inventory of the park's primary forests is easily available online through the National Park Service. Although primary forests do not always have old growth structure, a time-consuming field inventory of old growth forests in GSMNP was outside of the scope of this project. The existing primary forest boundary data (see Section 2-B) represents a best guess as to where true old growth forests are likely to occur for a purely-remote study. Considering the relative infrequence of major natural disturbances in cove forests, they are more likely to exist in an old growth stage of forest succession than other communities. By this premise, studying forest structure in known primary cove forests provides an opportunity to gain insight into these forests' condition in an old growth successional state.

G. LiDAR Background

Light Detection and Ranging (LiDAR) data provides the basic toolkit used in this analysis. LiDAR is a form of active remote sensing, which portrays information via a three-dimensional "point cloud" representing the physical location of objects on the earth's surface. The data is acquired by an airborne laser scanner, which transmits a series of laser pulses and records the amount of time elapsed between the pulse transmission and its reception as it is reflected back to the scanner. These recorded values are tied to the scanner's global positioning system (GPS), and so each point is recorded with a specific horizontal and vertical position. LiDAR scanners can transmit nearly 400,000 laser pulses per second, and each pulse can record up to five *returns*, or points where the laser pulse impacted an object and was reflected back to the scanner. These multiple returns are then classified using various algorithms that estimate their correspondence to actual features on the landscape (e.g. ground and vegetation) (https://lta.cr.usgs.gov/lidar_digitalelevation).

Classified LiDAR point clouds can be processed into a variety of raster datasets representing particular aspects of the earth's surface. Common products include digital elevation models, models of canopy cover, models of understory vegetation, and models of forest canopy height. These latter products, canopy height models (CHMs) serve as the basis for the analysis in this report. CHMs are created by subtracting the ground returns (elevation of the earth's surface) from the first returns (top of the forest canopy), to generate a raster file with cell values representing the height of the forest canopy in a specified area (Lim et al. 2003).



LiDAR point cloud showing Grove Park Inn golf course, Asheville, NC (Notice higher point densities in trees and natural vegetation)

H. LiDAR in Forest Analysis

Due to its ability to portray relatively accurate information over large areas, LiDAR has become a popular tool for analyzing the physical structure of forest landscapes (Lim et al. 2003). Forest researchers frequently incorporate LiDAR analysis into their work in order to supplement or replace more time-consuming field-based measurement techniques. Researchers have demonstrated LiDAR's ability to estimate common forestry measurements including tree density, aboveground biomass (Lefsky 2005), basal area (Lefsky 2002), and fuel loads (Skowronski et al. 2007).

More recently, LiDAR datasets have been used to analyze more complex aspects of forest structure, including some applications to old growth forests. In 2009, Falkowski et al. used LiDAR to predict forest successional stage in the Northern Rockies with 95% accuracy using a series of 6 successional classes defined through field work. Between 2010 and 2011, a team from the University of Washington developed a series of LiDAR metrics to describe forest succession, patch structures, and structural complexity in a mix of primary and second growth conifer forests in the Cedar Run Watershed in the western Cascades. These metrics, most significantly 95th percentile height, rumple (crown surface roughness), and canopy density showed strong correlations with traditional field measurements of forest structure throughout multiple stages of forest succession including old growth (Kane et al. 2010a; 2010b; 2011).

LiDAR has also been used to analyze forest structure in the Southern Appalachian region. A 2012 study found that LiDAR metrics (canopy height, canopy cover, and stem density) were more accurate than Landsat imagery at classifying forest structure using supervised classification and object-oriented segmentation in Monongahela National Forest (Norman 2012). In 2013, Josh Kelly of Mountaintrue completed an assessment of ecological departure throughout Pisgah National Forest using LiDAR-derived canopy height, canopy cover, and shrub density (Kelly 2013). Most recently, Kumar et al. used various LiDAR products to classify vegetation canopy structure across Great Smoky Mountains National Park (Kumar et al. 2016).

The methods used in this analysis share some similarity with those of Dupuy et al., who incorporated texture analysis into a LiDAR-based study of forest structure on the island of Mayotte in the western Indian Ocean. Their analysis used a LiDAR-derived grey-level co-occurrence variance statistic to perform object-based classification of the island's forests into six vegetation classes. This study found a specific threshold in co-occurrence variance that separated the island's primary forests from second growth forests (Dupuy et al. 2013).

Section II: Orientation to Data

A. LiDAR Canopy Height Models

The LiDAR CHMs used in this project originated from two different sources. All data in North Carolina are derived from North Carolina's Phase III LiDAR acquisition, which covered the western corner of the state. EarthData International conducted the flights in March and April 2005, which would have been leaf-off conditions for most of the acquisition. The NC Phase III data is projected in NC State Plane in Zone 3200, NAD83. All data in Tennessee are derived from a 2011 acquisition by the Center for Remote Sensing and Mapping Science at the University of Georgia. These data are specific to the Tennessee portion of the Great Smoky Mountains National Park. The TN data were collected in April 2010 (~25% of the Park) and February-April 2011 (~75%), in leaf-off conditions. These data are projected in Universal Transverse Mercator, Zone 17, NAD83.

These two sets of raw LiDAR data were then processed in Fusion© Software into three distinct sets of CHMs covering three separate areas. The CHMs for Nantahala and Pisgah National Forests were processed by Josh Kelly, Mountaintrue's Public Lands Field Biologist. CHMs for the NC side of GSMNP were processed by Ed Schwartzman, a contracting biologist with The Nature Conservancy. CHMs for the TN side of GSMNP were processed by Jess Riddle, also a contracting biologist with The Nature Conservancy.

Due to the different coordinate systems used by the TN and NC datasets, these collections of CHMs differ in their cell size and measurement units. To match the UTM projection, the Tennessee data use a 3-meter by 3-meter cell size, and units are in 1-meter height classes. The NC Smokies CHMs use 10-foot by 10-foot cells, with units in 1-foot classes. These 10-foot cells were created in order to mimic the 3-meter TN cell size as closely as possible (3 meters = 9.84 feet). The NC Nantahala/Pisgah CHMs were created using a 9-foot cell size with units in 1-foot classes. These models cover all of Nantahala and Pisgah National Forests with the exception of the Grandfather Ranger District of Pisgah National Forest (see Map 1 for location) which was excluded from the NC Phase III LiDAR acquisition.

Nearly all of the CHMs were processed with a median filter in order to fill in "nodata" values within the model. "Nodata" replace erroneous high or low height measurements (less than 0 or higher than a specified max height). The median filter assigns the median height value in a specified neighborhood around a "nodata" value in order to smooth out the data for analysis. In some instances, a median filter was not used for the Nantahala and Pisgah data. This inconsistency in processing technique makes the CHMs less reliable for comparison across the different study areas (see Section 5-H for discussion). The following scripts are representative of those used for the TN and NC models:

For Tennessee:

..\canopymodel /outlier:-1,65 /median:3 /ground:path.dtm outfile 3 M M 1 17 2 2 tiles
..\DTM2ASCII canopydtm

For North Carolina:

\canopymodel /outlier:-3,213 /median:3 /ground:path.dtm outfile 10 F F 2 0 2 2 tiles

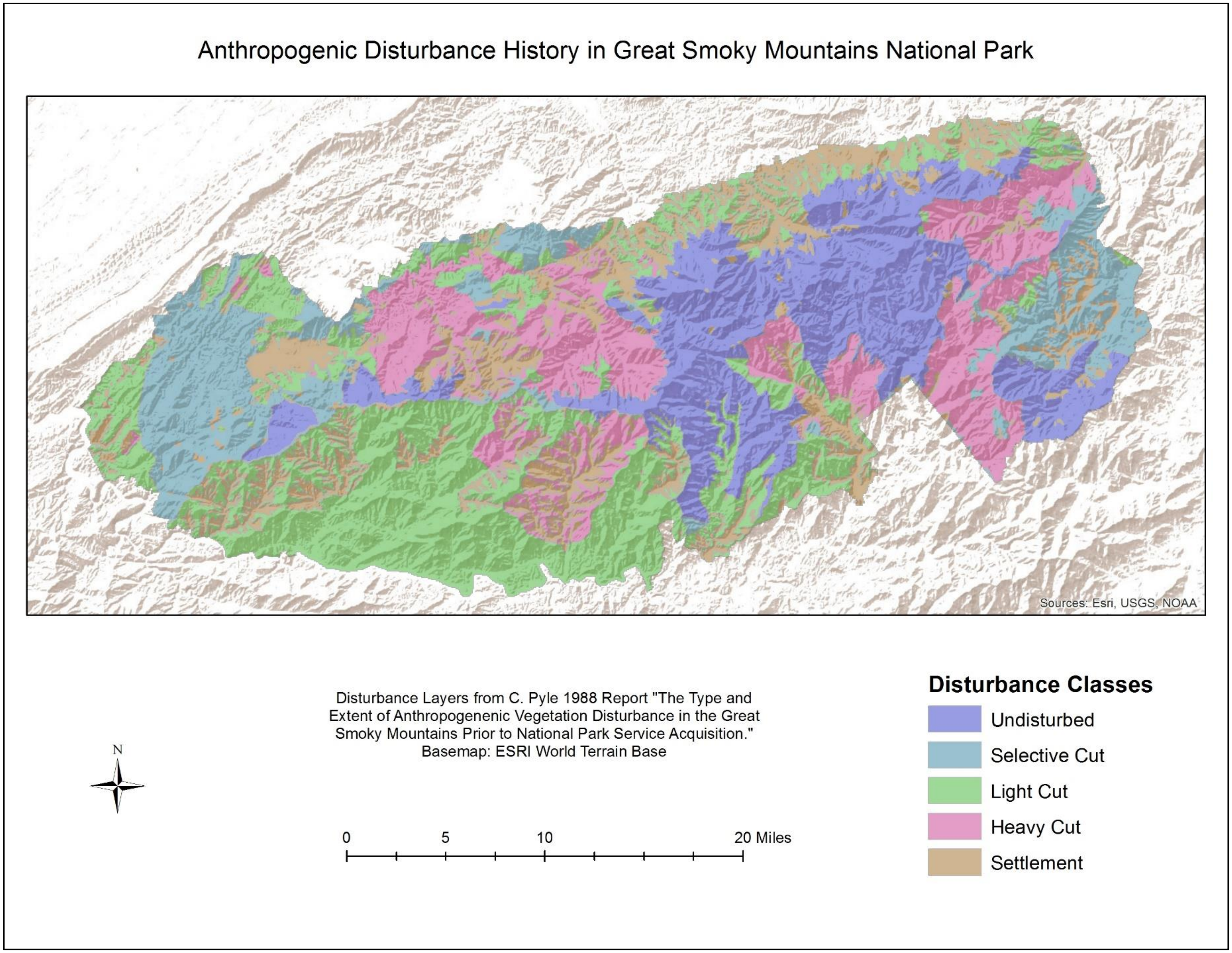
B. Forest Disturbance Surveys

Two different datasets were used for delineating the boundaries of old growth, primary, and second growth forests.

For **GSMNP**: I used a shapefile created from Charlotte Pyle's 1988 report "The Type and Extent of Anthropogenic Vegetation Disturbance in the Great Smoky Mountains before National Park Service Acquisition" (Pyle 1988). This report used archival records, and historic maps and photos to create a park-wide map of the GSMNP based on its anthropogenic disturbance history. These disturbance layers are shown in Map 3. The Pyle disturbance data divides the park into five classes based on the type and severity of the disturbance (ordered here from least-disturbed to most-disturbed):

- *Undisturbed*: These are primary forests, listed as "high in virgin forest attributes" on the original Pyle survey. They are delineated based on a lack of written records concerning anthropogenic land use and by the absence of any mapped land use records.
- *Selective Cut*: These areas are distinguished by patches of old trees which were probably never logged, intermingled with areas of light logging and other disturbance. These are shown on the original Pyle maps as "big trees with diffuse disturbance."
- *Light Cut*: These areas include much of the area mapped as "diffuse disturbance" on the original Pyle maps. They were mostly subject to small, early style logging operations, but include pockets of heavily disturbed forests. These areas may also have been affected by livestock grazing or light fires.
- *Heavy Cut*: These areas were heavily impacted by corporate logging prior to the Park's dedication in 1932. Areas designated as "heavy cut" experienced mechanized logging with little selectivity regarding species or timber quality. These forests were often heavily logged at the scale of entire watersheds, and many experienced intensive fires following logging.
- *Settlement*: These areas constitute settlements that existed prior to the designation of the park. Much of the area mapped as "settlement" comprises abandoned agricultural land.

For **Pisgah and Nantahala National Forests**: I delineated the boundaries of old growth forests using data from the 2000 unpublished report "Old-Growth Forest Communities in the Nantahala-Pisgah National Forest" compiled by Rob Messick and on file at the Mountaintrue office. This report is the result of over five hundred field outings by fifty different people. It identified 77,418 acres of old growth throughout the National Forests using a consistent ranking system which ranks sites A-C with regards to the level of human disturbance visible at the site. This analysis only uses sites ranked as Class A, where no significant signs of human disturbance were recorded in the canopy or understory. See Appendix I for the full classification system used in these surveys.



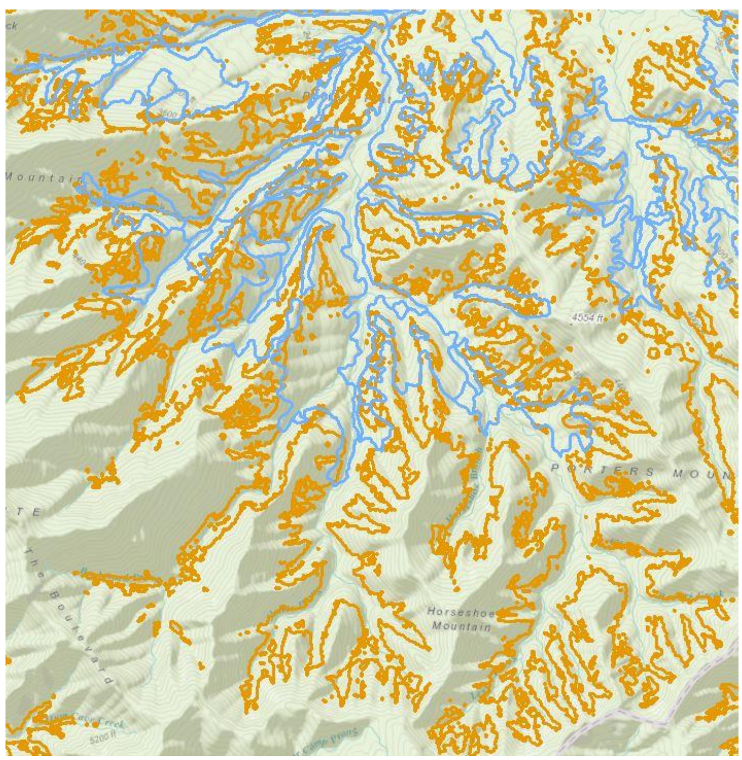
Map 3: Anthropogenic Disturbance History in Great Smoky Mountains National Park

C. Ecological Zone Models

Ecological zone models, or "ecozones," developed by Steve Simon of the US Forest Service were used to delineate the boundaries of cove forests throughout the entire study area. Ecological zones are defined as "units of land that can support a specific plant community or plant community group based upon environmental factors such as temperature, moisture, fertility, and solar radiation that control vegetation and distribution" (Simon 2011). In this manner, ecozones represent *potential* vegetation rather *actual* vegetation. These are the most accurate ecosystem maps available for the entire Southern Blue Ridge Ecoregion, and their land classification system that can be cross-walked to the LANDFIRE Biophysical Settings currently used by the US Forest Service. Most importantly for this analysis, the ecological zones were used to locate the random points for the regression-based models (Section 5-E).

D. Great Smoky Mountains Vegetation Mapping

Vector-based digital vegetation data from the University of Georgia's Center for Remote Sensing and Mapping Science aided my analysis of the Great Smoky Mountains National Park. These data were created for the entire park using a combination of analog photointerpretation, digital softcopy photogrammetry, and GIS/GPS-assisted field data (Madden et al. 2004). The data consist of nearly 50,000 polygons representing overstory vegetation communities classified at the finest division of the National Vegetation Classification System protocol. Unlike the ecological zone models, the GSMNP vegetation maps identify ecosystems based on *actual* vegetation. These data layers represent a narrow definition of cove forests, and tend to classify a significantly smaller area as cove forest than the ecozone data:



Graphic shows differences between cove forests as mapped by ecological zones (orange) and GSMNP overstory vegetation (blue)

Section III: Texture Analysis

A. Grey-Level Co-Occurrence Texture Analysis
This study uses image texture metrics generated from grey-level co-occurrence matrix (GLCM) calculations to describe patterns in the LiDAR CHMs. *Texture* can be defined as the spatial arrangement of pixel intensities within an image (Jain et al. 1995). Texture metrics quantify differences in contrast, the size of an area where changes occur, and the directionality of pixel relationships (Hall-Beyer 2007). Image analysts commonly use texture filters to pick out hard-to-describe differences that aid image classification. In the case of the CHMs, texture analyses are able to capture complex differences between forest communities that cannot be understood through simpler first-order descriptive statistics (Dupuy et al. 2013).

A GLCM calculates the conditional joint probabilities of all pair wise combinations of grey-levels in a spatial window of interest given parameters of interpixel distance and orientation (Clausi et al 2002). In more practical terms, a GLCM uses a moving window to calculate the probability that one pixel value will occur beside another. A GLCM filter assigns a value to the cell in the center of the window corresponding to a calculation that was applied to the GLCM. Consider the following example image:

0	1	0	2
0	1	3	3
2	3	3	3
1	2	1	2

The numbers above represent pixel values (or grey-level intensities) as they might appear in any image or raster file. The concentration of adjacent "3" values on the right side of the image represents an area of homogenous texture. A GLCM first requires that the user specify an angular relationship between pixels for the matrix to examine. This is accomplished by specifying a *reference* pixel and a *neighbor* pixel. These are commonly expressed in an X,Y format, which considers that any pixel will have eight neighboring pixels:

-1,-1	0,-1	1,-1
-1,0	Reference Pixel	1,0
-1,1	0,1	1,1

The X,Y (ex. 1,1 1,0 etc.) values above state the pixel orientation where the X direction equates to east and the Y direction equates to south. This is the orientation system used by ENVI, the software used to compute GLCM layers in this analysis. In this case, a 1,0 relationship would specify that the matrix consider a neighbor pixel to be one pixel to the right of the reference pixel: 1 pixel in the X direction, and 0 in the Y direction.

For any image, the GLCM will consider all possible grey level combinations. The example image above contains values 0-3, so the relationships considered by the GLCM for that image would be as follows:

Neighbor Pixel Value ----------------> **Reference Pixel Value:**	**0**	**1**	**2**	**3**
0	0,0	0,1	0,2	0,3
1	1,0	1,1	1,2	1,3
2	2,0	2,1	2,2	2,3
3	3,0	3,1	3,2	3,3

The GLCM is structured such that identical co-occurring values are spaced along the diagonal, and values with greater degrees of contrast are located farther from the diagonal. For example, for the 1,0 relationship, if the value of 0 occurs directly right of another 0 three different times, then the cell in the upper left corner of the matrix would be given a value of three. Similarly, if a value of 3 occurs to the right of the value 0 two different times, then the upper right hand corner would receive a value of two. The following is a matrix for the 1,0 relationship derived from the example image:

0	2	1	0
1	0	2	1
0	1	0	1
0	0	0	3

In order for a GLCM to be useful for image analysis, it must be converted from raw co-occurrence values into probabilities. For this to be possible, it is necessary to calculate a second matrix based on the opposite orientation from the original image. For the example above, (a 1,0 relationship) this means calculating a second matrix for the -1,0 relationship:

0	1	0	0
2	0	1	0
1	2	0	0
0	1	1	3

These two matrices are then added together to form a symmetrical matrix:

0	3	1	0
3	0	3	1
1	3	0	1
0	1	1	6

This symmetrical matrix is then normalized by dividing each value by the sum of all of the values. The equation for the normalized matrix is:

$$P_{i,j} = \frac{V_{i,j}}{\sum_{i,j=0}^{N-1} V_{i,j}}$$

i: row number
j: column number

The full normalized GLCM for the example image is as follows:

0	0.125	0.042	0
0.125	0	0.125	0.042
0.042	0.125	0	0.042
0	0.042	0.042	0.25

This normalized GLCM represents the probability that any pixel value will occur beside any other. For example, in the matrix above, a value of three has a 25% probability of occurring beside another value of three. Pixels along the diagonal of the normalized GLCM represent the probability pixel values will occur directly beside themselves. In the example image, three is the only value that occurs next to itself, so the diagonal is empty except for the cell representing the 3,3 relationship.

Information about the texture of an image can be obtained by applying different formulas (detailed below) to the GLCM probabilities. These formulas work by applying varying weights to certain values within the matrix. The values are then added up and entered into the cell in the center of the matrix. The following texture metrics were used in this report:

For all equations, i=row number, and j=column number

Entropy: $$-\sum_i \sum_j p(i,j) \log(p(i,j))$$

- Entropy calculates the randomness of grey levels within an image. Areas with high entropy values are characterized by unpredictable, highly variable patterns of grey level intensities.

Contrast: $$\sum_{n=0}^{N_g - 1} n^2 \left\{ \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j) \right\} \quad |i-j| = n$$

- Contrast places exponentially higher weights on values as they get farther from the diagonal of the matrix. Areas with high contrast are characterized by the presence of high values directly beside low values.

Dissimilarity: $$\sum_{n=1}^{N_g - 1} n \left\{ \sum_{i=1}^{N_g} \sum_{j=1}^{N_g} p(i,j)^2 \right\} \quad |i-j| = n$$

- Dissimilarity is similar to contrast, except that weights increase linearly away from the diagonal instead of exponentially.

Homogeneity: $$\sum_i \sum_j \frac{1}{1+(i-j)^2} p(i,j)$$

- Homogeneity places exponentially *decreasing* weights on values as they get farther from the diagonal. Areas with high homogeneity are characterized by large amounts of similar pixel values clustered in the same area.

<u>Second Moment:</u> $$\sum_i \sum_j \{p(i,j)\}^2$$

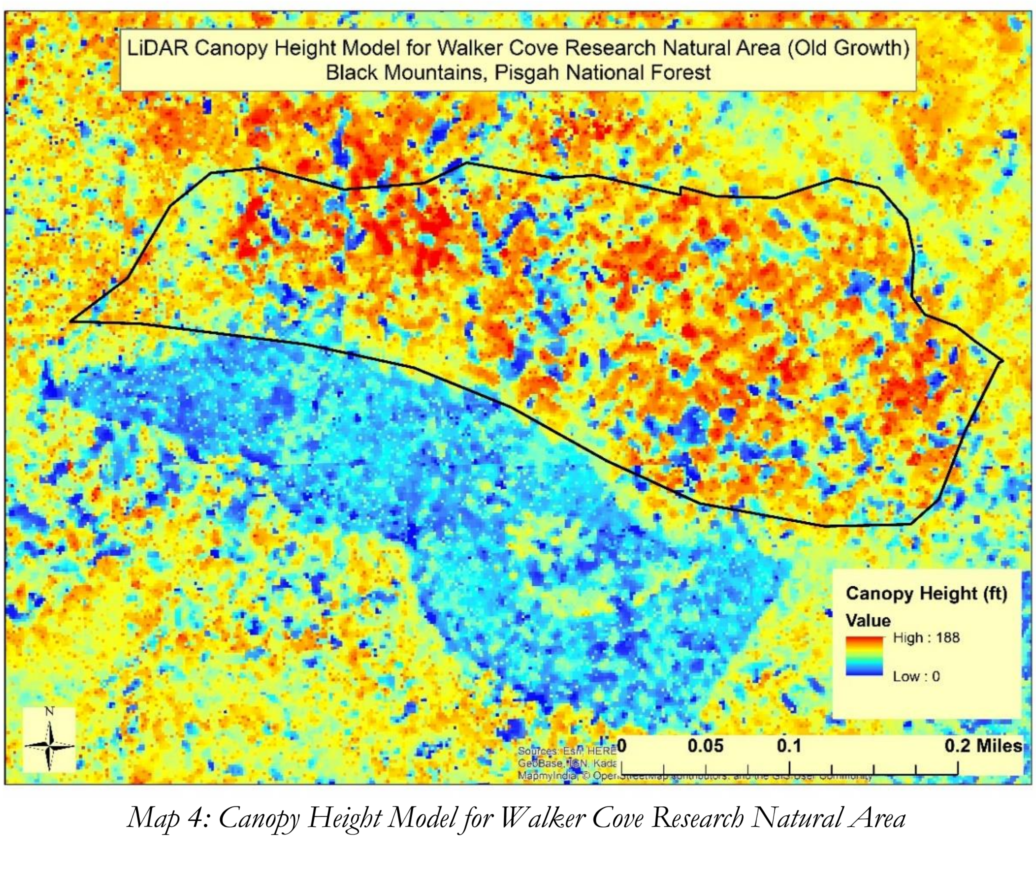
- Second moment uses the GLCM probabilities as weights for themselves. Areas with high second moment are characterized by orderly patterns (e.g. a checkerboard).

<u>Variance:</u> $$\sum_i \sum_j (i-u)^2\, p(i,j)$$

- Variance measures the dispersion of combinations of reference and neighbor pixels around the mean. Areas with high variance are characterized by widely-varying combinations of pixel pairs.

B. Connecting Texture Analysis to Forest Structure

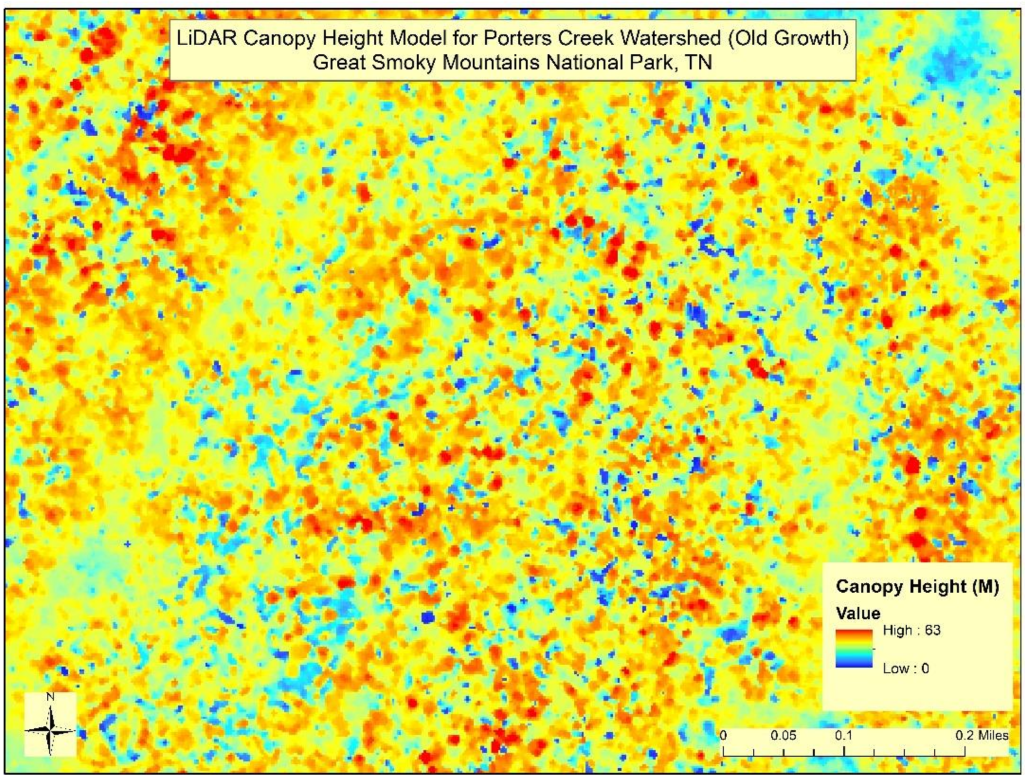
The idea for this project originated from a realization that differences between old growth and second growth cove forests are visible in the LiDAR CHMs. Aspects of old growth cove forests include an uneven-aged structure, several large canopy trees per hectare including some individuals over 200 years of age, and a mosaic of varying-sized canopy gaps (Runkle 1996). By contrast, forest structure in second growth forests tends to reflect that forests' management history. This example is demonstrated by the following map of an old growth cove forest (outlined in black) – validated by the Messick surveys – directly adjacent to a recovering clear-cut (concentration of low values represented by blue pixels):



Map 4: Canopy Height Model for Walker Cove Research Natural Area

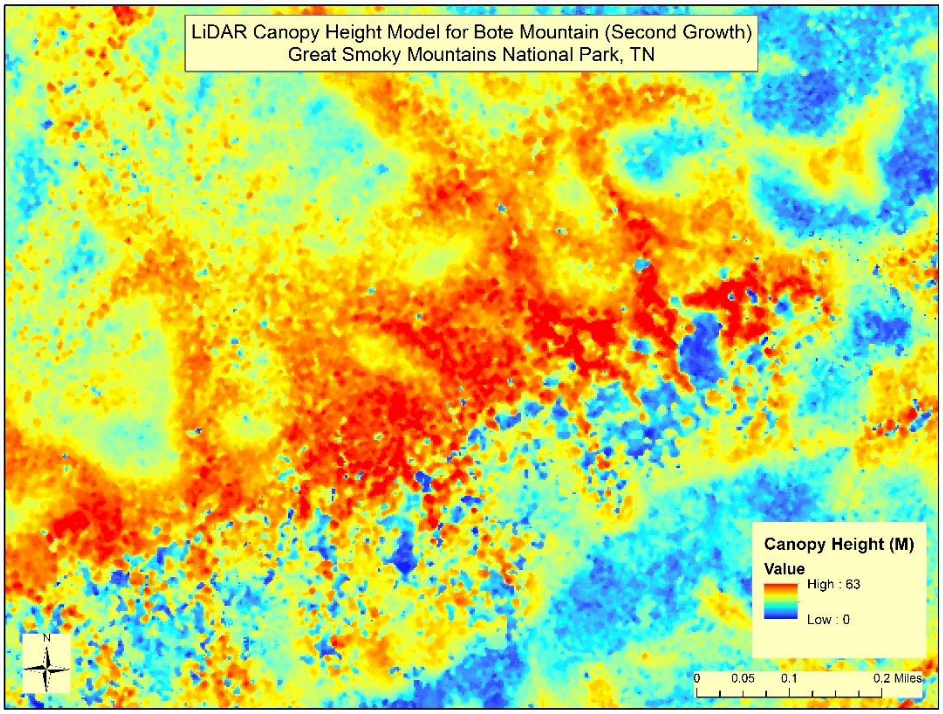
In Map 4 above, a mosaic of canopy gaps and tall trees is visible in the contrasting red and blue values throughout the old growth stand. The recovering stand to the south stands out as having height values which are generally similar to each other, reflecting the stand's even-aged structure. Considering that the CHMs are a raster grid of pixel intensity values, the differences visible between these two stands lie in both the intensity values themselves and the spatial arrangement of those values. Texture metrics, which are designed to capture aspects of pixel intensity arrangement, can provide insight into aspects of horizontal forest structure that are visible in the CHMs. For example, in terms of its texture, the second growth stand in Map 4 could be described as "homogenous," while the old growth stand is "dissimilar," "random," and full of "contrast": all terms that texture metrics are designed to capture.

Consider the following series of images. The first (Map 5) shows an old growth cove forest in the Porters Creek Watershed of GSMNP. In this image, an extensive mosaic of large trees and varying-sized canopy gaps typical of old growth cove forests is represented by sharp variations in pixel intensity:



Map 5: Canopy Height Model for Porters Creek Cove Forests

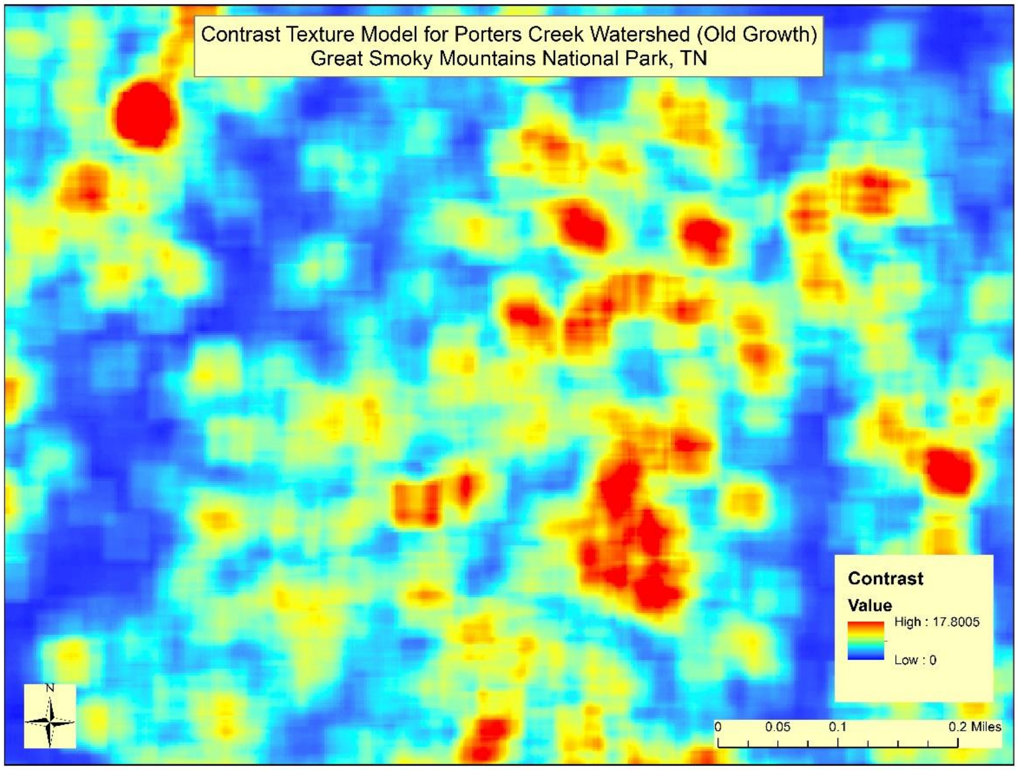
The next image (Map 6) shows a second growth cove forest just north of Bote Mountain in GSMNP TN. In the center of the image, a very high even-aged stand – likely a *Liriodendron tulipifera* stand recovering from industrial logging over 100 years ago – is shown in a large cluster of red (high) pixels. Just south of the even-aged cove stand, a montane oak hickory forest is kept open by wild grape (*Vitis spp.*) species, creating a more variable canopy structure:



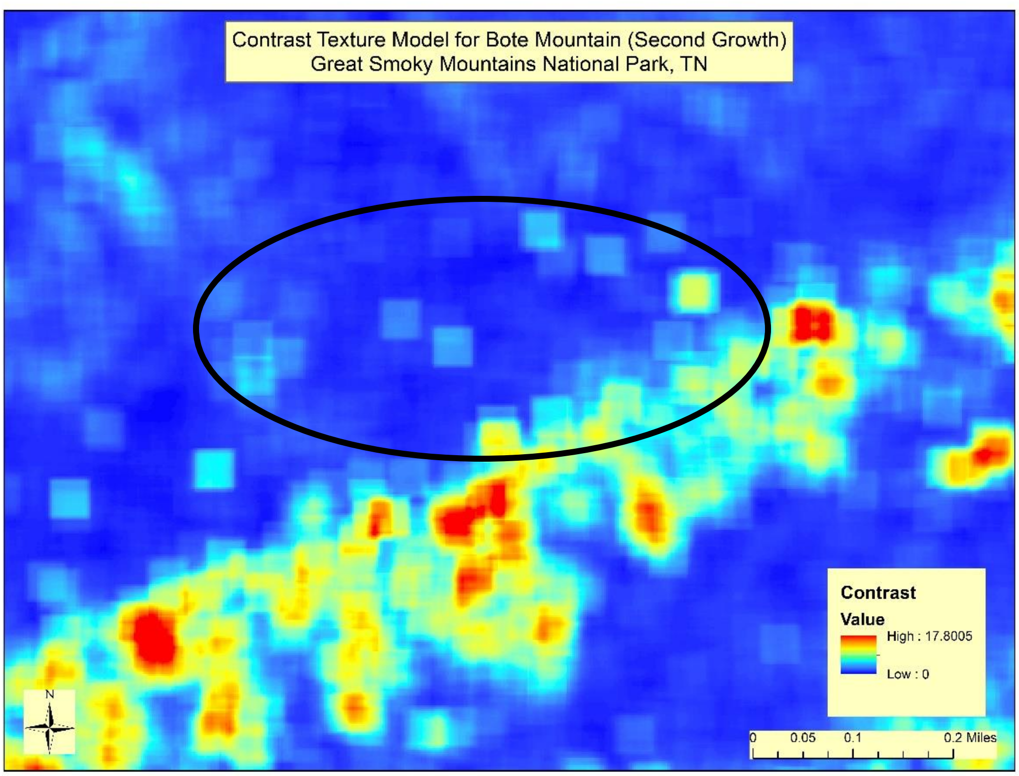
Map 6: Canopy Height Model for Bote Mountain Cove Forests

The following series of images show how the CHMs for forest stands in Maps 5 and 6 are affected by various texture measures. The old growth stand (Porters Creek) shows high levels of entropy, contrast, dissimilarity, and variance, and low levels of homogeneity and second moment. By comparison, the tall even-aged second growth stand (Bote Mountain) shows low levels of entropy, contrast, dissimilarity, and variance, and high levels of homogeneity and second moment. Texture values in the grape openings in the hickory forest tend to correspond to the texture values in the old growth forest.

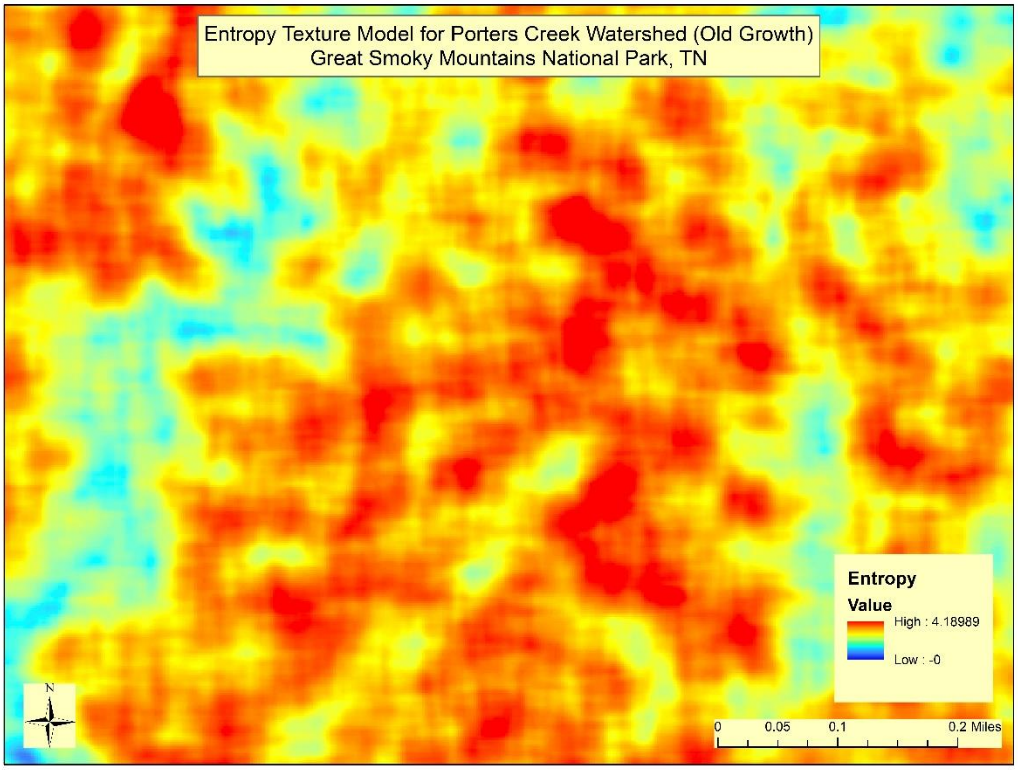
The general location of the even-aged second growth stand is marked with a black circle in all of the second growth images:



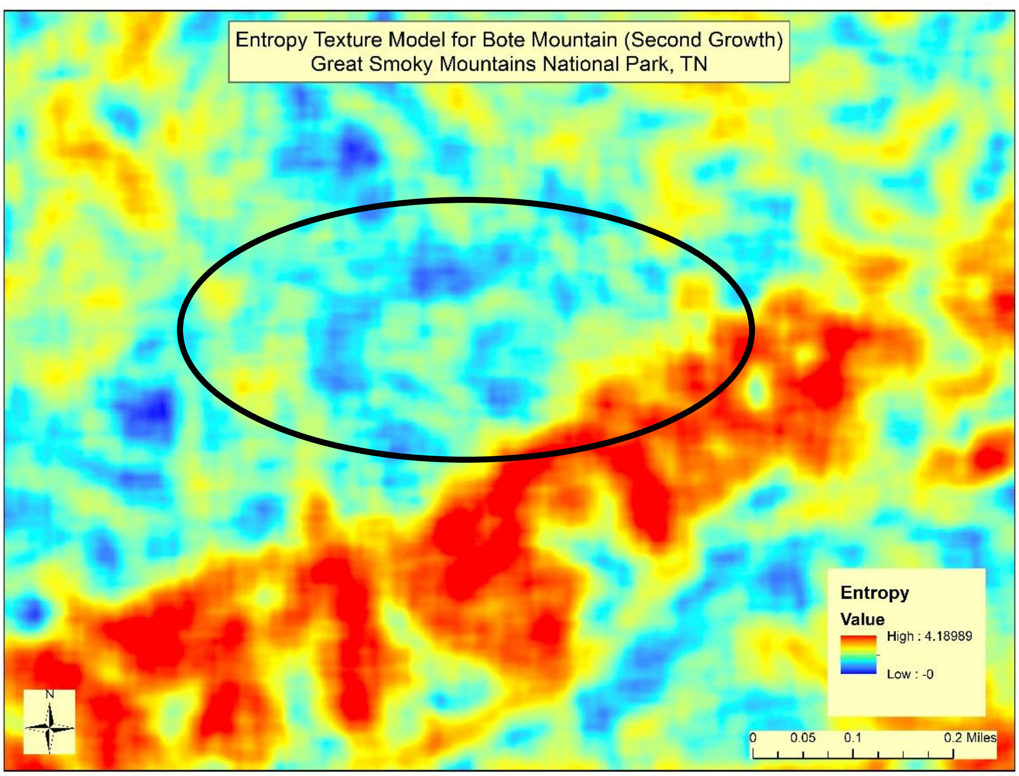
Map 7: GLCM Contrast in an Old Growth Cove Forest



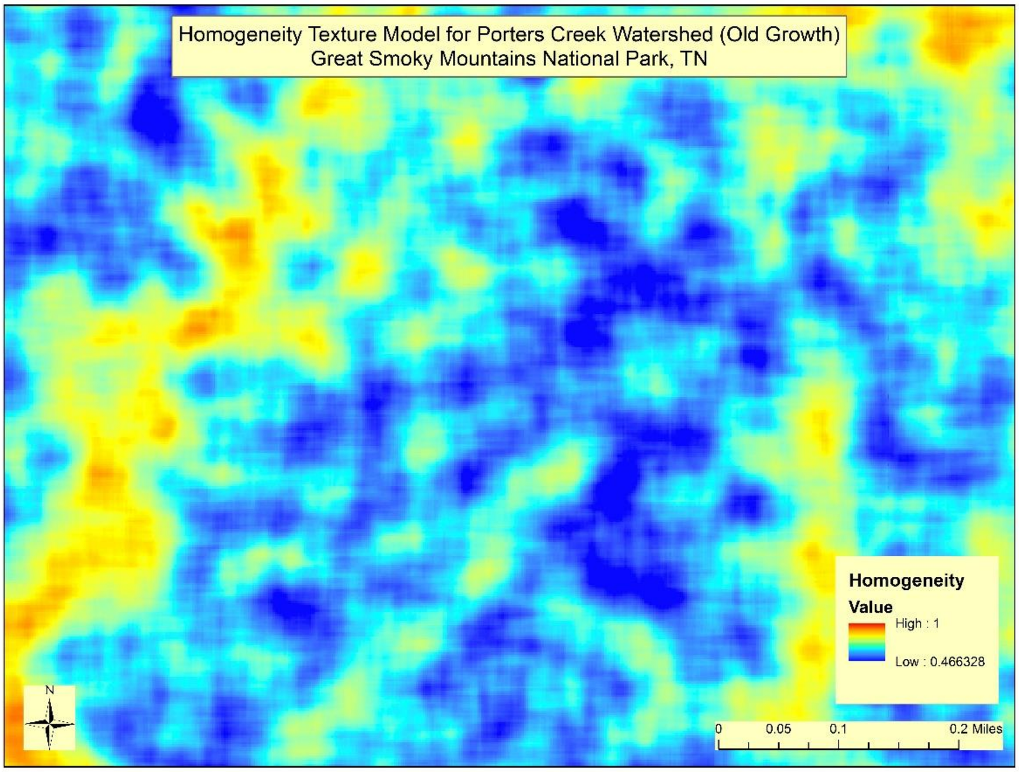
Map 8: GLCM Contrast in a Second Growth Cove Forest



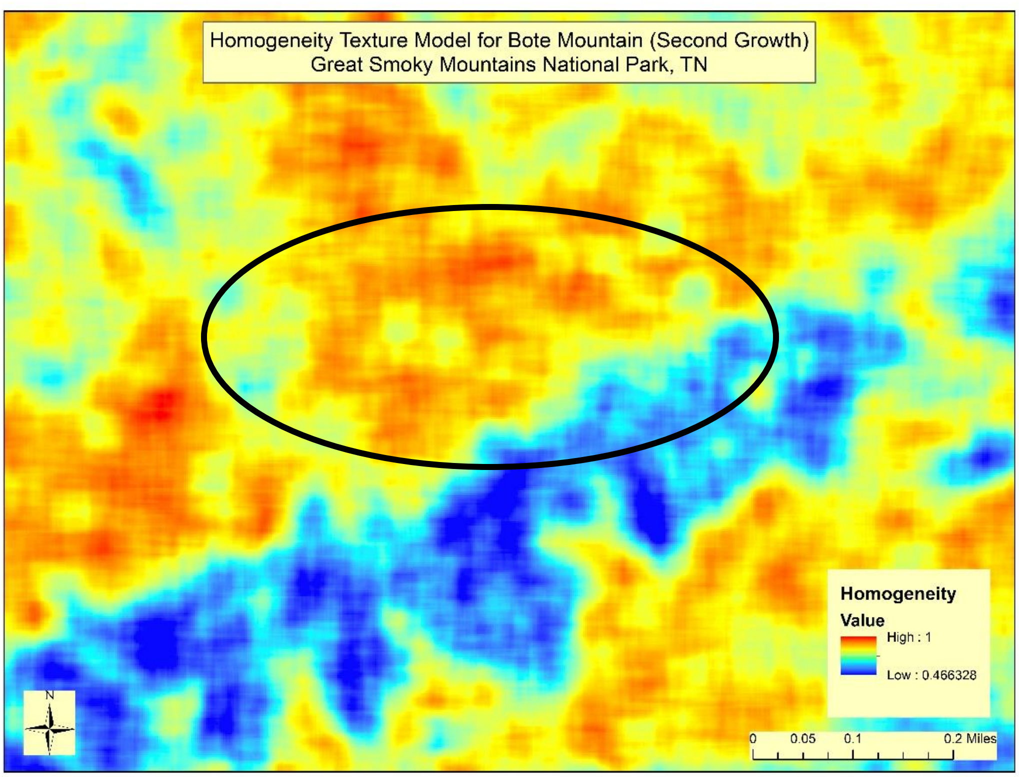
Map 9: GLCM Entropy in an Old Growth Cove Forest



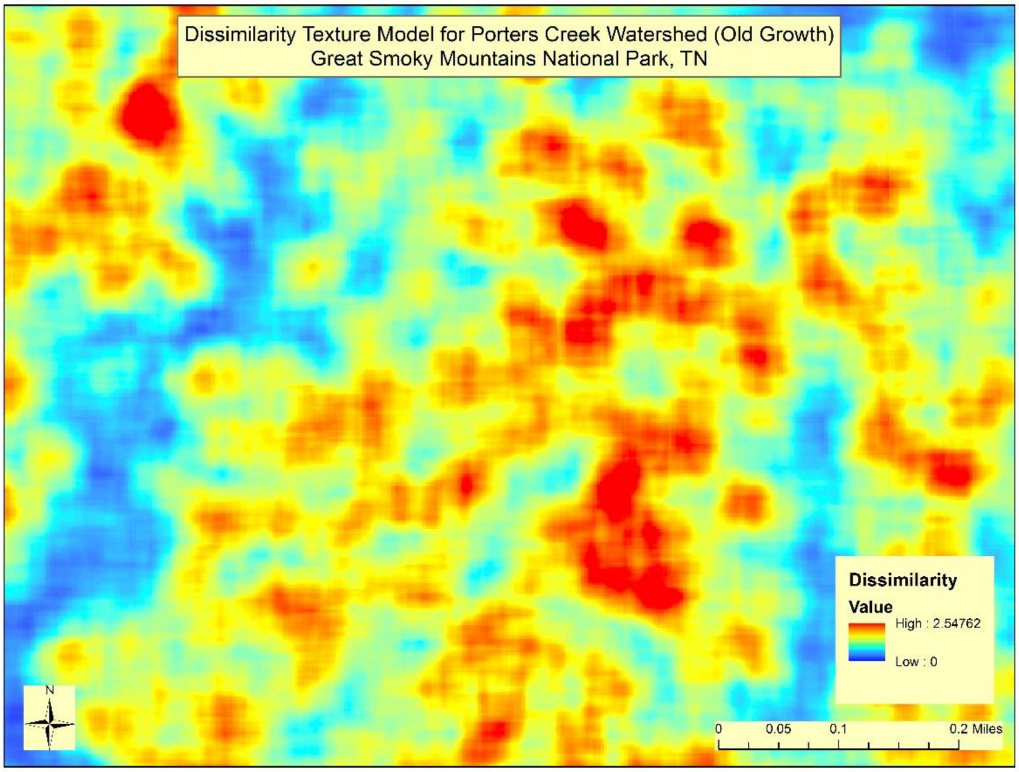
Map 10: GLCM Entropy in a Second Growth Cove Forest



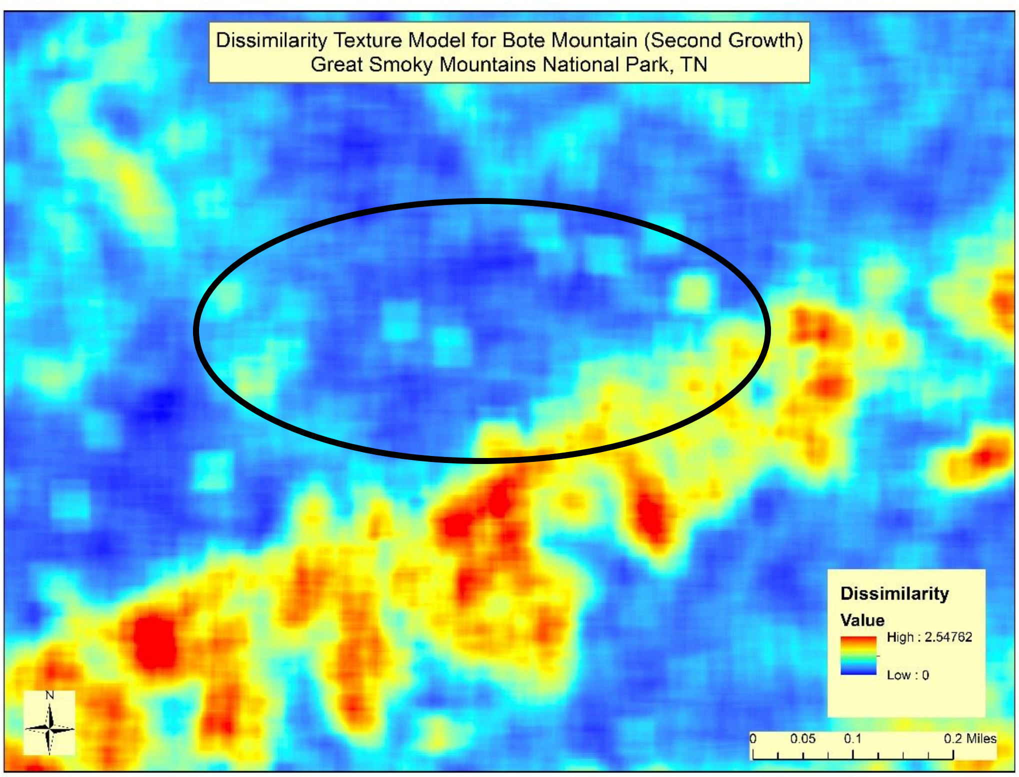
Map 11: GLCM Homogeneity in an Old Growth Cove Forest



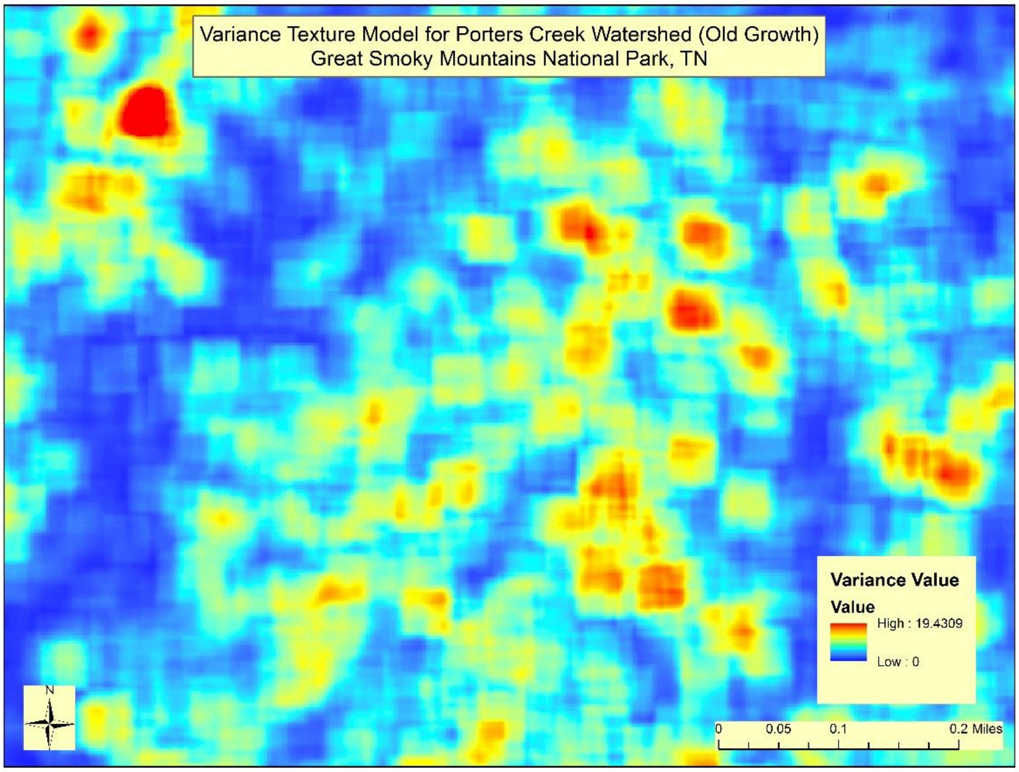
Map 12: GLCM Homogeneity for a Second Growth Forest



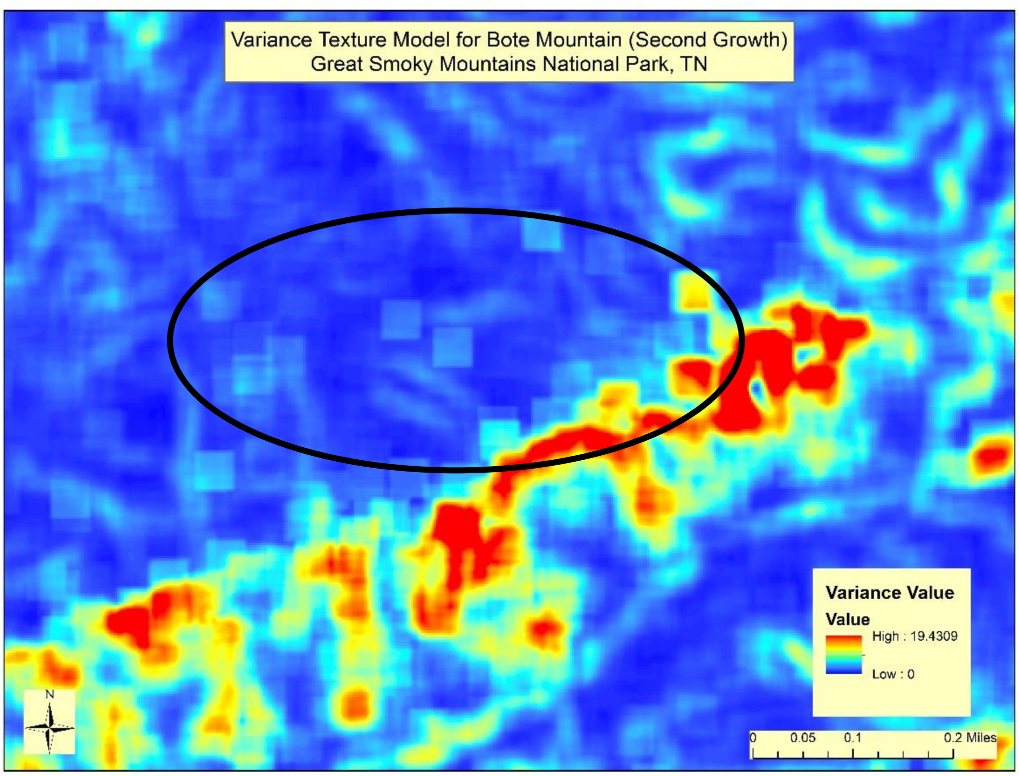
Map 13: GCLM Dissimilarity in an Old Growth Cove Forest



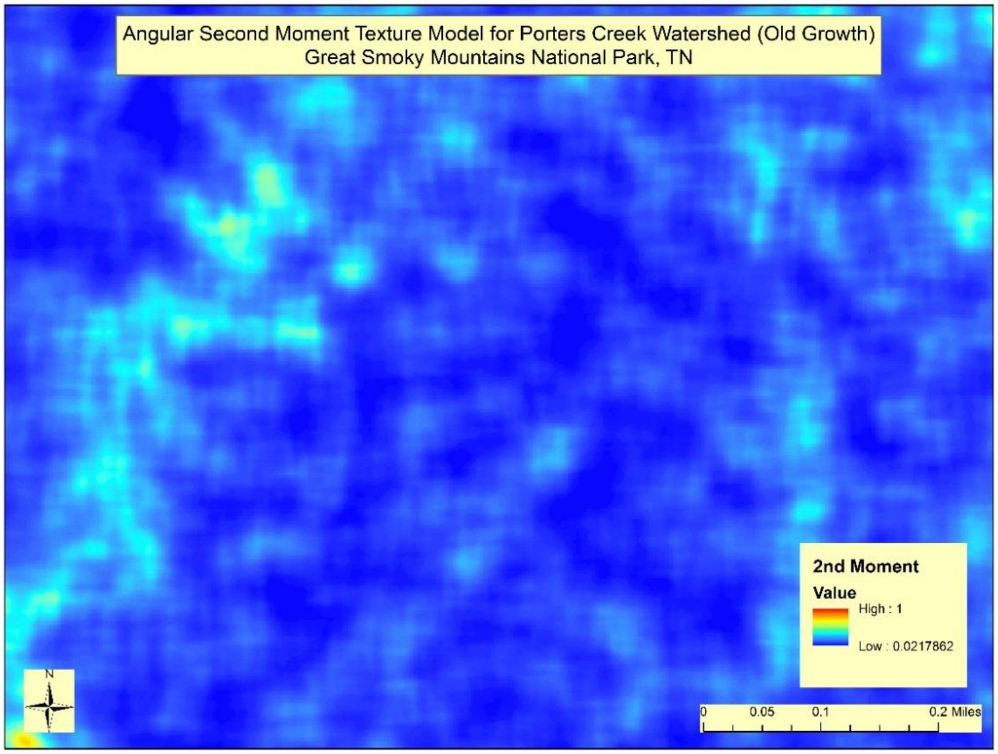
Map 14: GLCM Dissimilarity in a Second Growth Forest



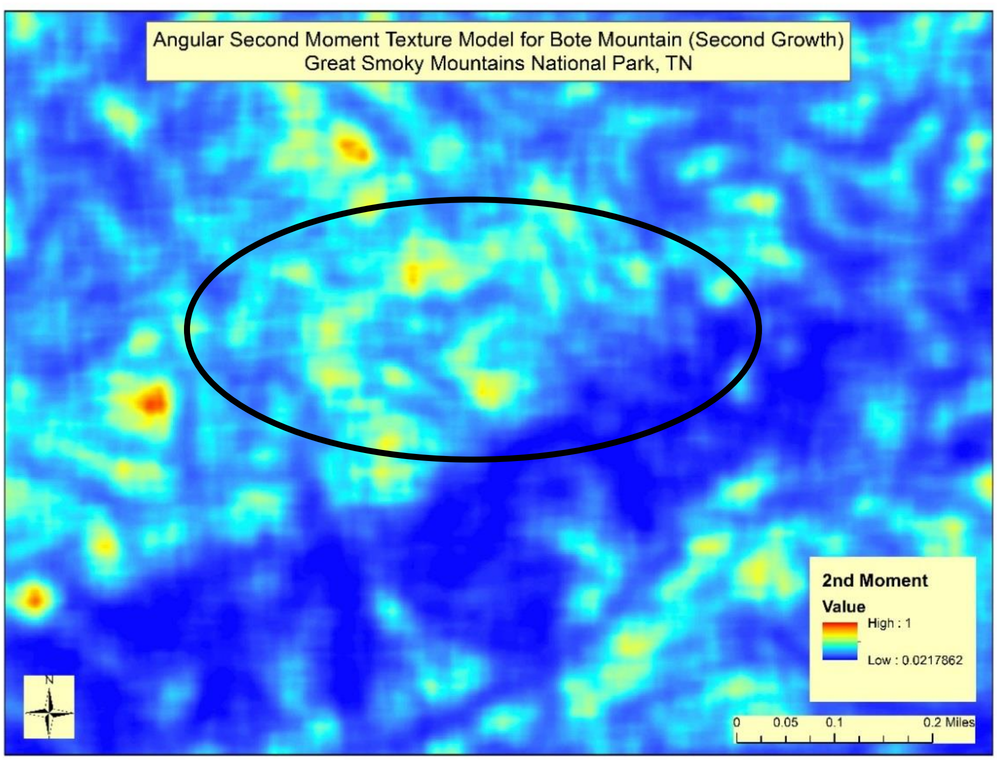
Map 15: GLCM Variance for an Old Growth Cove Forest



Map 16: GLCM Variance for a Second Growth Forest



Map 17: GLCM Second Moment in an Old Growth Cove Forest



Map 18: GLCM Second Moment in a Second Growth Cove Forest

Section IV: Methods

A. Texture Analysis
Grey level co-occurrence matrix texture layers were created in ENVI 5.3 (32-bit) software using a 21x21 cell moving window. In order to create the layers, each canopy height model was saved in ArcGIS GRID format and imported directly into ENVI. The grey-level quantization was set at 64. The GLCM layers used in the final model used the average of all four pixel orientations. The following section explains why these parameters were chosen:

<u>Window Size:</u>
A 21x21 window size was chosen by creating a series of GLCM Correlation layers with increasing window sizes. An online texture tutorial, available through the University of Calgary recommends choosing a window size by charting the GLCM Correlation values across increasing window sizes. The point at which the values level off or begin to decline can be interpreted as the approximate size of definable objects within an image (Hall-Beyer 2007).

In order to apply this concept to the LiDAR CHMs, and to reduce processing time, these correlation layers were created for a 1,000 hectare extracted area within an old growth cove forest in the Porters Creek watershed on the Tennessee side of GSMNP (Map 5). This watershed includes a relatively large area of recognizable old growth texture, and Mountaintrue biologist Josh Kelly recommended the site as representative of old growth forest structure based on his field experience in the area. For this initial analysis, the zonal mean of the correlation values for varying-sized square plots with a common center were compared at window sizes ranging from 3x3 to 71x71. The results (Figure 1) show the correlation values leveling off between the 21x11 and 41x41 window sizes:

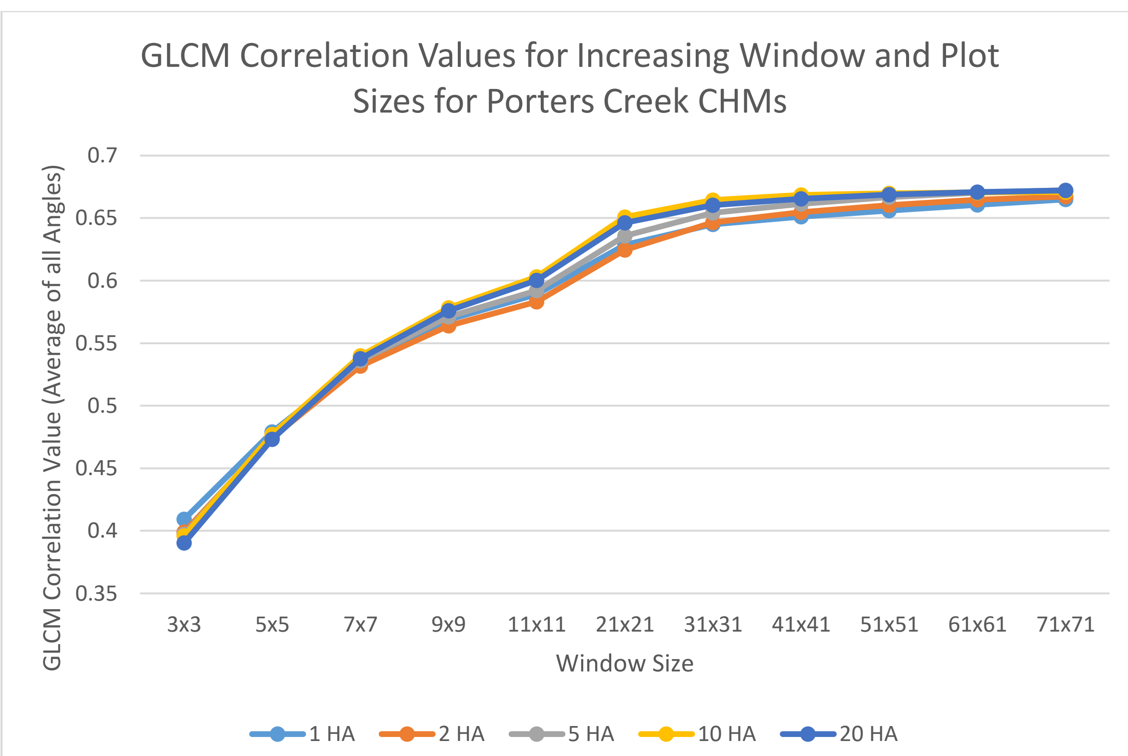


Figure 1: GLCM Correlation Values with Fixed Plots at Porters Creek, GSMNP, TN

This window size analysis was then repeated using slightly different methods. For this second analysis, polygons in the cove ecogroup from the GSMNP vegetation data were used as object "zones" instead of the square plots. This was accomplished by calculating the mean of each cove forest polygon using the ArcGIS Zonal Statistics as a Table tool. The reason for incorporating the polygons is that unlike the square plots, they roughly correspond to actual ecosystem boundaries that may provide more insight into the correct window size. These results are shown in Figure 2:

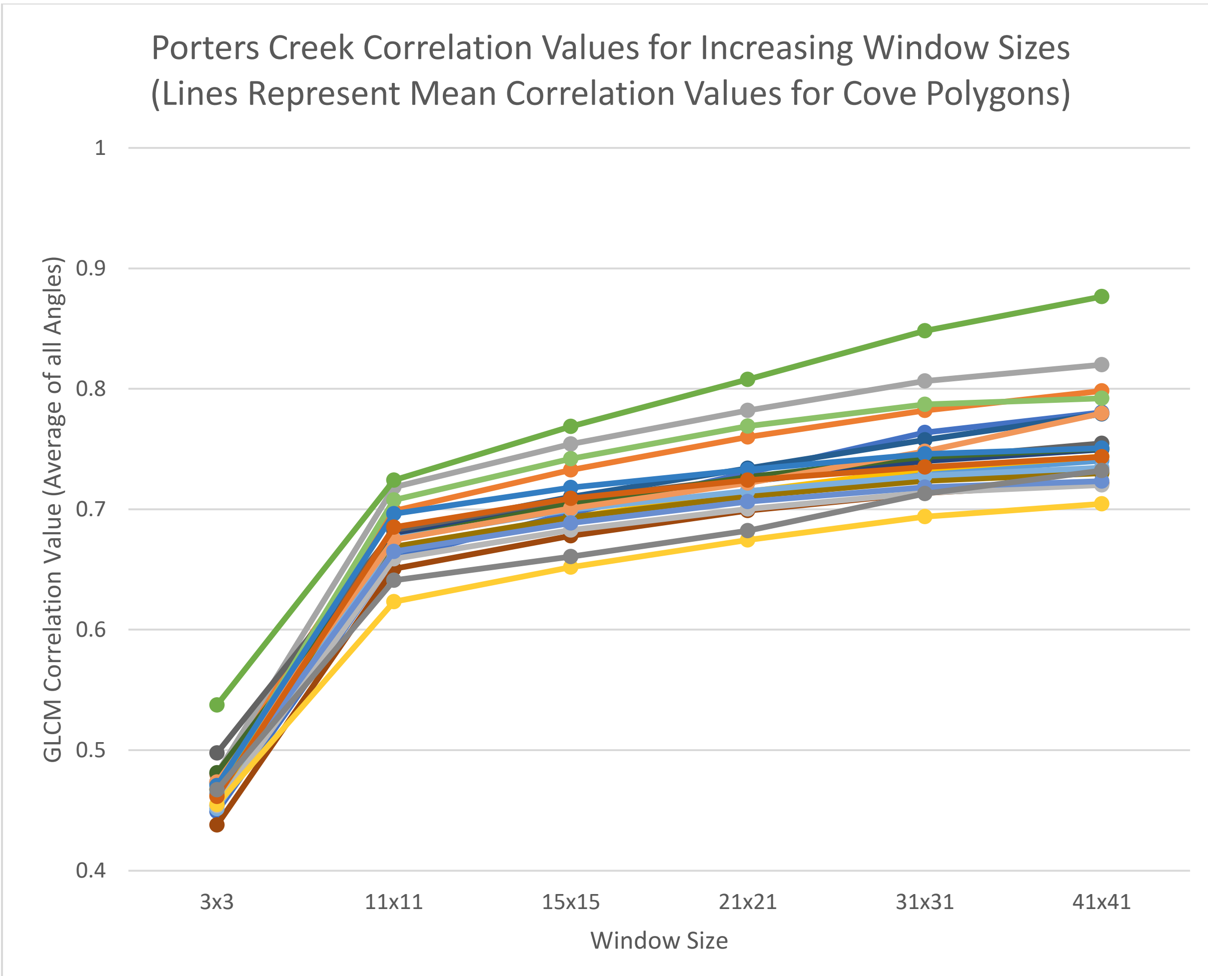


Figure 2: Mean Correlation Values for Cove Ecogroup Polygons at Porters Creek, GSMNP TN

Figure 2 also shows the correlation values leveling off gradually between an 11x11 and 31x31 window size. A third correlation window size analysis was then completed in order to visualize the correlation values in a homogenous second growth cove forest. Here, a series of GLCM Correlation layers were created for a 1,000-hectare plot surrounding an even-aged second growth stand at Bote Mountain, also on the Tennessee side of GSMNP (shown in Map 6). The mean correlation value for each cove ecogroup polygon was computed using the Zonal Statistics as a Table tool in ArcGIS. These results are shown in Figure 3, which shows the GLCM correlation values leveling off somewhere around a 21x21 or 31x31-pixel window size:

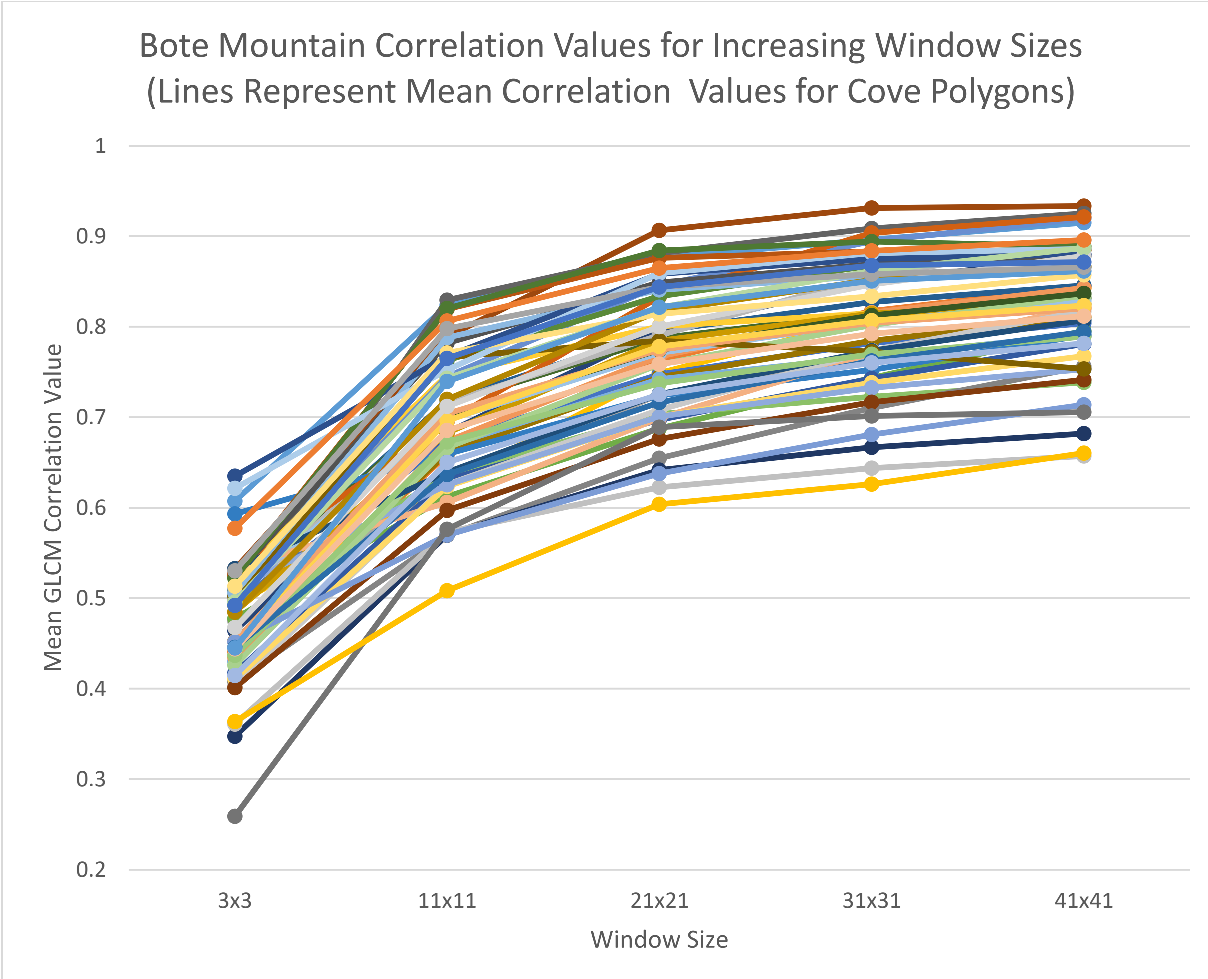
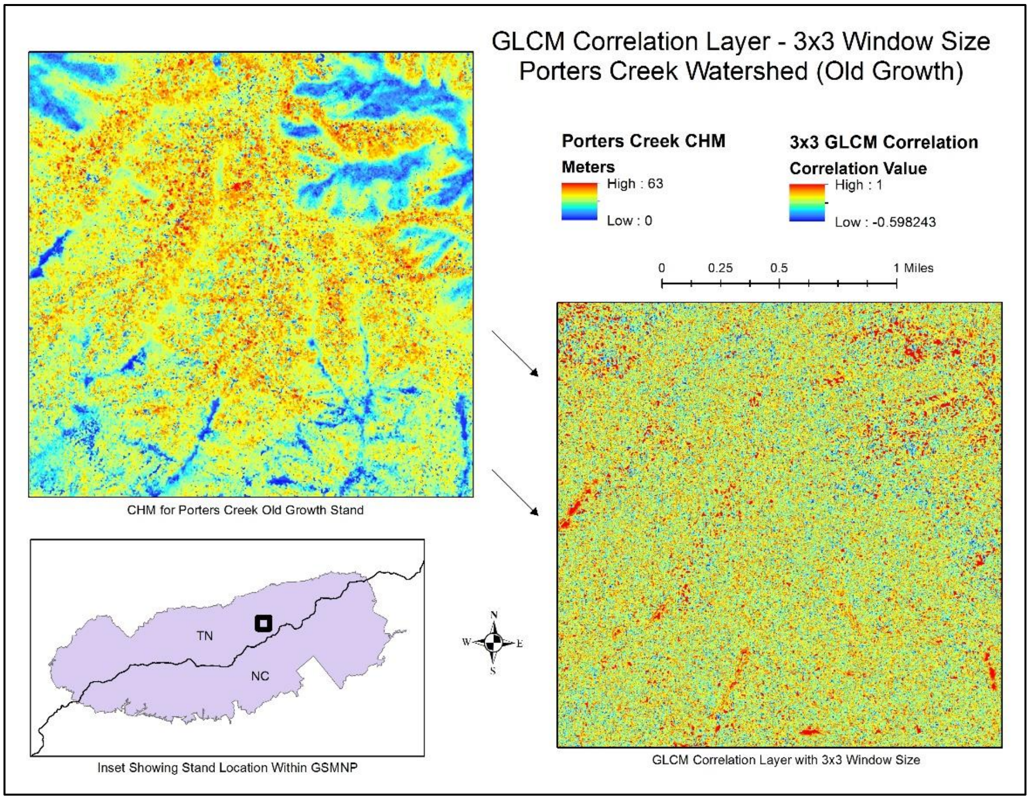


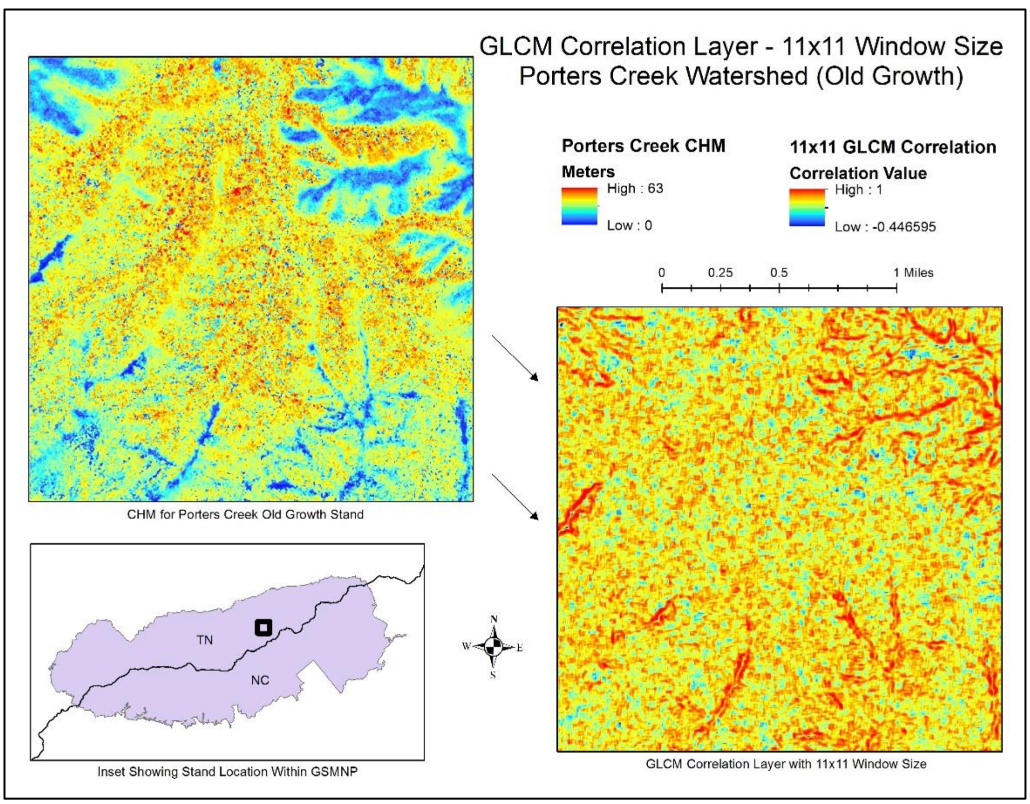
Figure 3: Mean Correlation Values for Cove Ecogroup Polygons at Bote Mountain, GSMNP TN

Figures 1-3 show the GLCM correlation values leveling off gradually at window sizes between 11x11 and 41x41 pixels with no specific point marking the spot where the curve flattens out. The raster layers created for each window size also provide some visual insight into which window size best captures the forest community boundaries represented in the CHMs. Maps 19-28 on the following pages show the GLCM correlation layers created for Porters Creek (old growth) and Bote Mountain (second growth) at increasing window sizes between 3x3 and 41x41. Both of the original CHMs (Porters Creek and Bote Mountain) show a mix of cove forests and other ecosystems. Certain ecosystem boundaries can be discerned around the concentrations of low height values symbolized by blue pixels. These ecosystem edges tend to produce high GLCM correlation values which are symbolized with red pixels in the following maps. Considering that the goal of this correlation analysis is to understand which window best accounts for the objects within an image, the visibility of ecosystem edges is important for understanding whether or not the correlation statistic is capturing landscape features of interest to this project (e.g. forest stands, canopy gaps, horizontal forest structure).

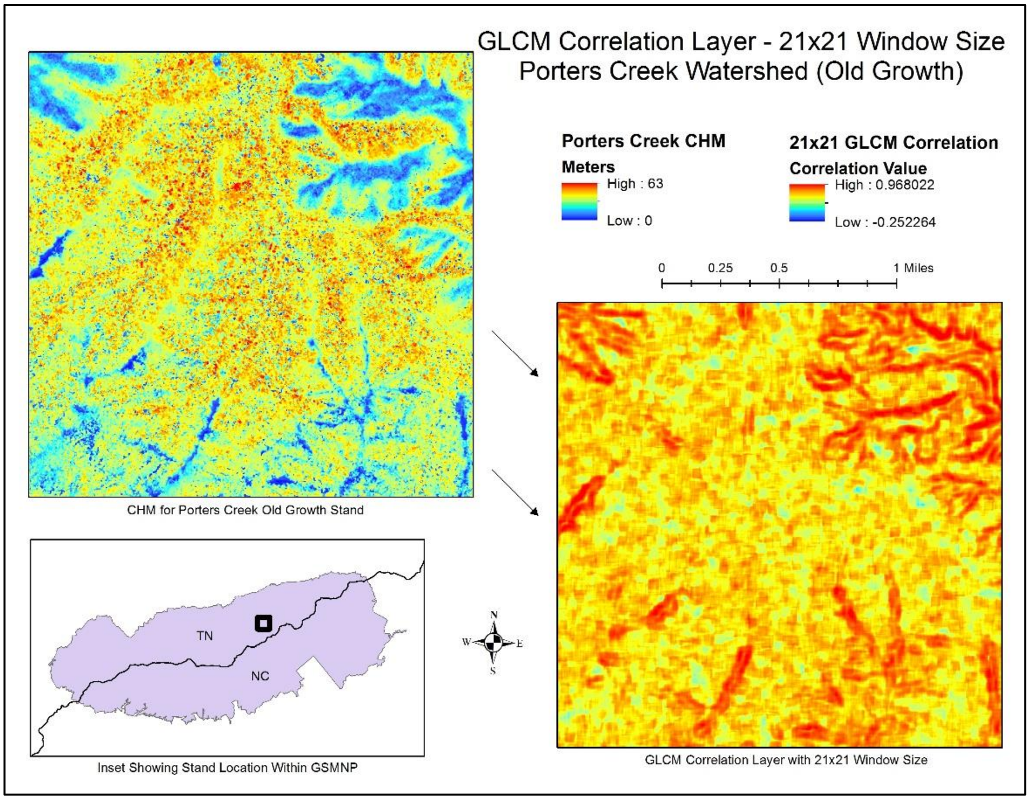
GLCM correlation formula used by ENVI: $\frac{\sum_i \sum_j (ij)p(i,j) - \mu_x \mu_y}{\sigma_x \sigma_y}$



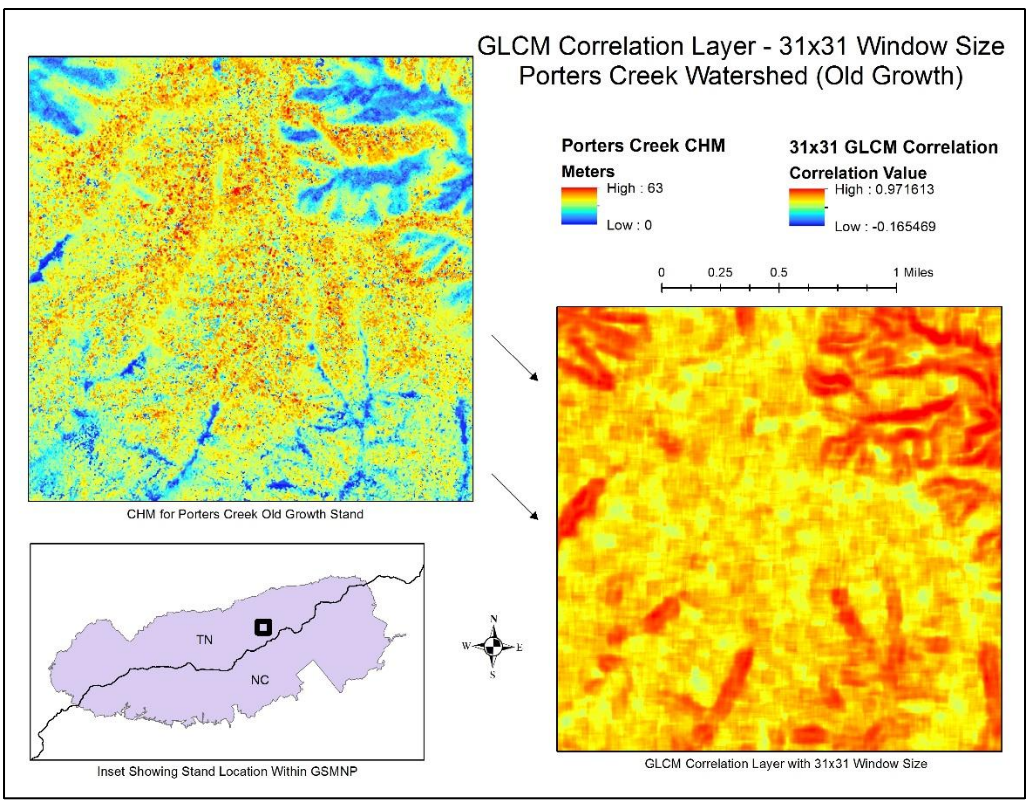
Map 19: 3x3 GLCM Correlation Visual for Porters Creek Old Growth Stand



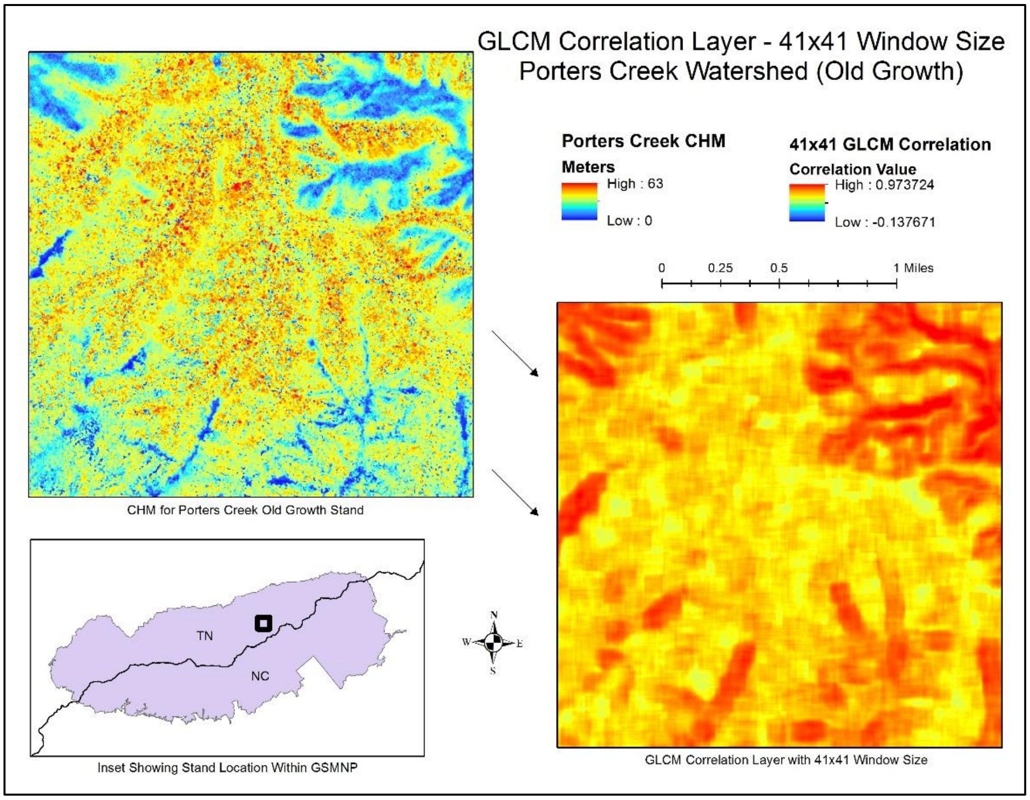
Map 20: 11x11 GLCM Correlation Visual for Porters Creek Old Growth Stand



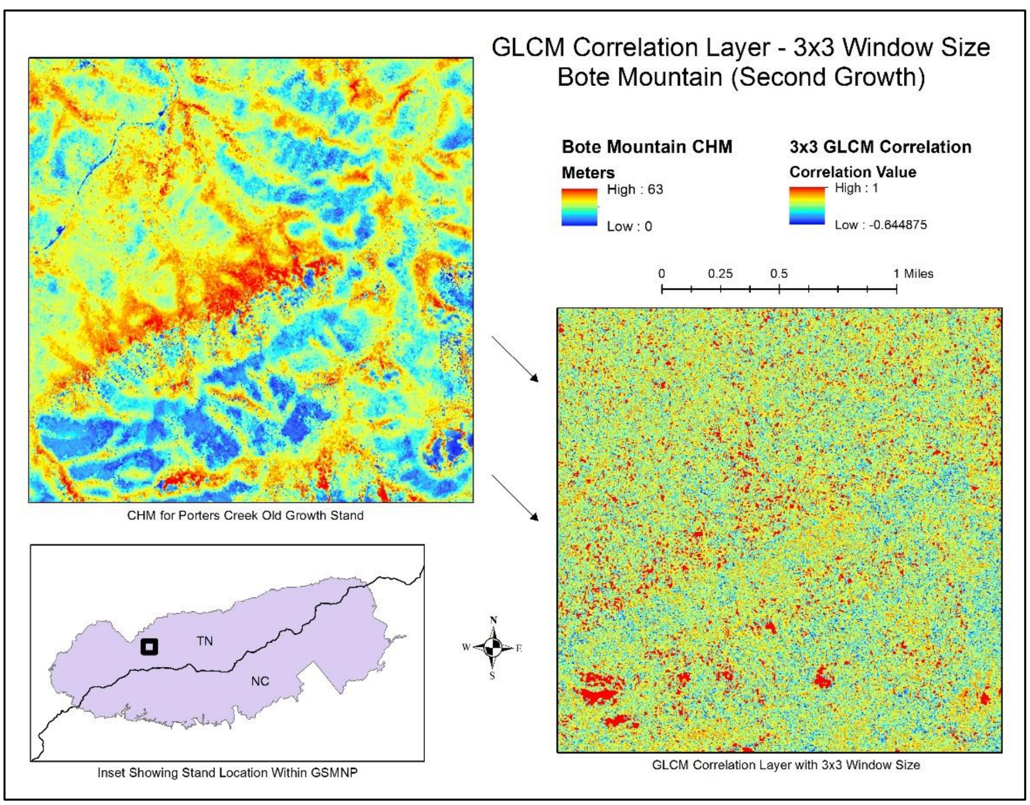
Map 21: 21x21 GLCM Correlation Visual for Porters Creek Old Growth Stand



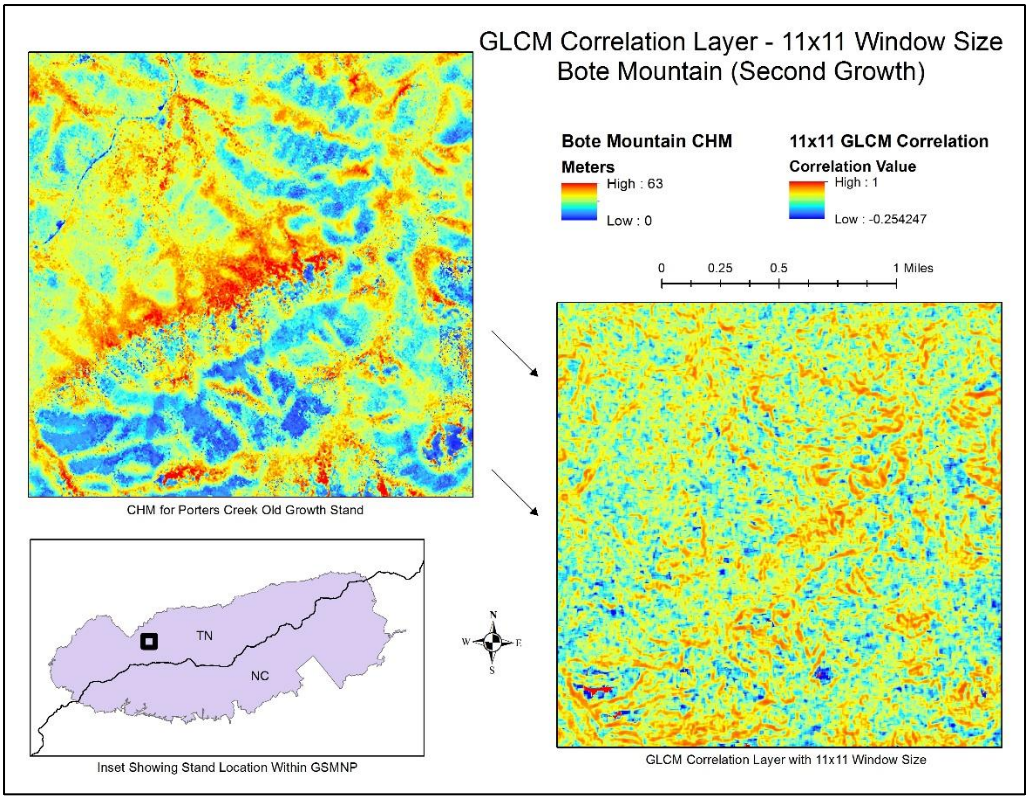
Map 22: 31x31 GLCM Correlation Visual for Porters Creek Old Growth Stand



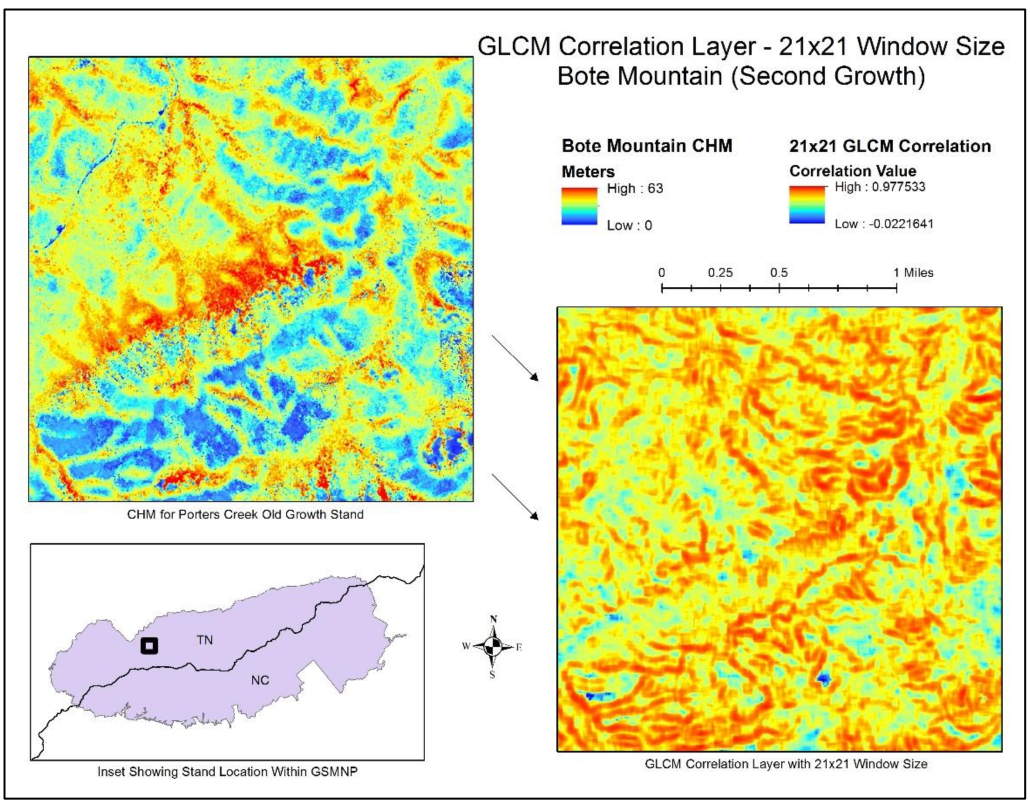
Map 23: 41x41 GLCM Correlation Visual for Porters Creek Old Growth Stand



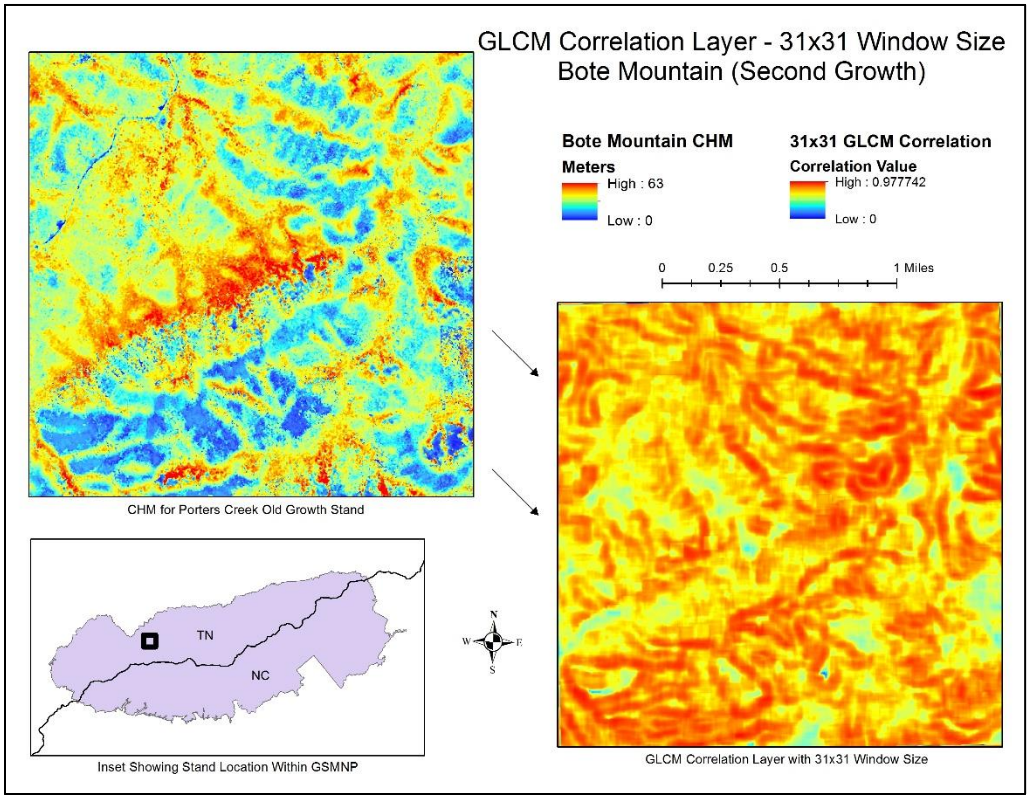
Map 24: 3x3 GLCM Correlation Visual for Bote Mountain Second Growth Stand



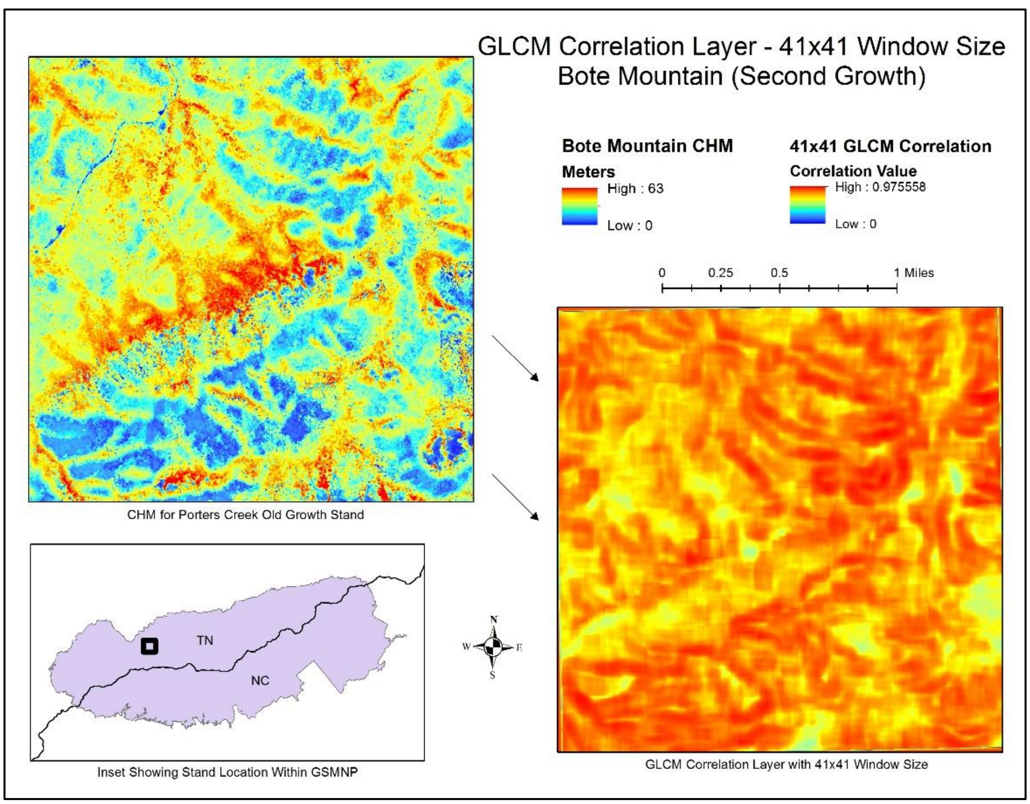
Map 25: 11x11 GLCM Correlation Visual for Bote Mountain Second Growth Stand



Map 26: 21x21 GLCM Correlation Visual for Bote Mountain Second Growth Stand



Map 27: 31x31 GLCM Correlation Visual for Bote Mountain Second Growth Stand



Map 28: 41x41 GLCM Correlation Visual for Bote Mountain Second Growth Stand

Maps 19-28 generally show higher correlation values at larger window sizes, with ecosystem edges appearing faintly visible at an 11x11 window and visible, but distorted at 41x41. A 21x21 pixel moving window was chosen for creating the final texture layers because it lies on the transitioning curve in Figures 1-3, and because it appears to pick out most ecosystem edges in Maps 21 and 26. Finally, 21x21 is preferable to larger window sizes because larger windows will include more overlap between different ecosystems. For example, if a pixel lies within a cove forest, but near a boundary with another ecosystem, the moving window used to calculate that pixel's value would include a large number of pixel relationships outside of the cove forest. Given this, at 3,969 square meters, a 21x21 window is large enough to capture old growth features like canopy gaps and tall trees, but small enough to limit overlap with other ecosystems.

Grey Level Quantization:

The CHMs used in this analysis are all in the form of 8-bit unsigned integer ArcGIS raster files. The size of the matrix being considered for a given pixel will correspond to the number of possible values allowed by a raster's bit depth. The formula for this range of possible values is 2^N, where N=bit depth. This gives an 8-bit raster 256 (0-255) possible values. Since the CHM values only ranged from 0-64 when in meters, and 0-190 when in feet, each texture matrix would include a large number of zero entries. A high proportion of 0's within a GLCM leads to "unstable calculations" (Hall-Beyer M. Personal Communication) which can alter the behavior of the calculations, causing inaccurate texture layers.

It is unlikely that any raster file will have exactly 256 values. Since the 8-bit format is common for raster data, the ENVI GLCM tool includes an option to specify the grey level quantization (number of grey levels allowed by the matrix). For this analysis, the grey level quantization was set at 64 for all texture layers in order to perfectly match the 64 height values in the Tennessee CHMs. The NC CHMs, which are in units of feet, were rescaled from ~190 values down to 64 to match the Tennessee data. ENVI uses an undisclosed rescaling algorithm to perform this data reduction.

Pixel Orientation

The GLCM is a function of the "angular relationship and the distance between two neighboring pixels" (ENVI). GLCM calculations require that a user specify the orientation to be considered between a reference pixel and a neighbor pixel. This analysis uses GLCM calculations to describe a variety of structural features over a landscape scale, and does not consider any specific directional relationship between CHM pixels (i.e. a canopy gap or tall tree can occur at any direction from a neighbor pixel). To account for this, texture layers were generated for the 0,1; 1,1; 1,0; and 1,-1 relationships. Because of the symmetrical nature of the GLCM, this averaged orientation accounts for every possible orientation with a 1-pixel displacement. The average of all four of these orientations was then calculated using the Raster Calculator tool in ArcGIS. These averaged layers were used as the texture inputs for all analyses in this report.

B. Texture Visualization in GSMNP

The following steps were completed in order to visualize and compare the differences in GLCM texture metrics between anthropogenic disturbance history classes defined by the 1988 Pyle surveys:

The GSMNP vegetation polygons were used to visualize trends in the texture metrics because they provide a more conservative estimate of cove ecosystem boundaries than the ecological zone models. All polygons in the Montane Cove Forest Ecogroup (including acid, typic, rich, and red oak cove forests) were exported to a separate feature class. These cove forest polygons were then clipped to each disturbance class in the Pyle disturbance history shapefile (undisturbed, selective cut, light cut, heavy cut, and settlement). The ArcGIS Buffer tool was used to delete an interior buffer equal to half of the window size considered by the texture calculations: for data in TN (3M cell size), a 30-M interior buffer was used, and for data in NC (10 Ft cell size) a 100-ft interior buffer was used.

The ArcGIS Extract by Mask tool was then used to extract each texture layer to the filtered cove forest polygons within each anthropogenic disturbance category. Beyond the GLCM layers, canopy height and mean height were also included. The mean height layer was created in ArcGIS using the focal statistics tool and a 21x21 square window. The ArcGIS Reclassify tool was used to convert each layer to categorical signed-integer raster files in order to visualize their distribution in the attribute table. The attribute tables from these reclassified rasters were then copied to Microsoft excel spreadsheets in order to calculate the values in Figures 4-19. These reclassified layers held extremely different numbers of pixel observations. In order to create meaningful charts showing trends within each texture metric, the data were converted into separate tables showing the percentage that each numeric class occupies within its respective anthropogenic disturbance history class (Figures 4-19).

C. Random Point Generation

The statistical comparisons and regression modeling completed in this analysis used values collected from the groups of random points scattered throughout the study area. These points were generated using the ArcGIS Create Random Points tool within certain constraining multipart feature classes. The following parameters were used in each collection of points:

GSMNP TN (1500 points):

- All points were spaced a minimum of 30M apart (to eliminate overlapping windows)
- 500 points were generated in primary cove forests. These points were confined to cove forests within the "undisturbed" section of the Park designated by the Pyle surveys. Cove forests included both the ecozones and the GSMNP cove ecogroup polygons, which were merged and dissolved into a single constraining feature class.
- 500 points were generated in second growth cove forest. These points were confined to cove forests within the "heavy cut" and "settlement" (most intensely disturbed) sections of the Park designated by the Pyle surveys. Cove forests included both the ecozones and the GSMNP cove ecogroup polygons, which were merged and dissolved into a single constraining feature class.
- 500 points were generated outside of cove forests. To create the constraining feature class for this layer, I used the ArcGIS Erase tool to erase all cove forests (ecozones and GSMNP cove vegetation polygons) from the park. These cove forests were buffered by 30M prior to running the Erase tool to ensure that no window overlap occurred between the cove and non-cove sampling areas.

GSMNP NC (1500 points):

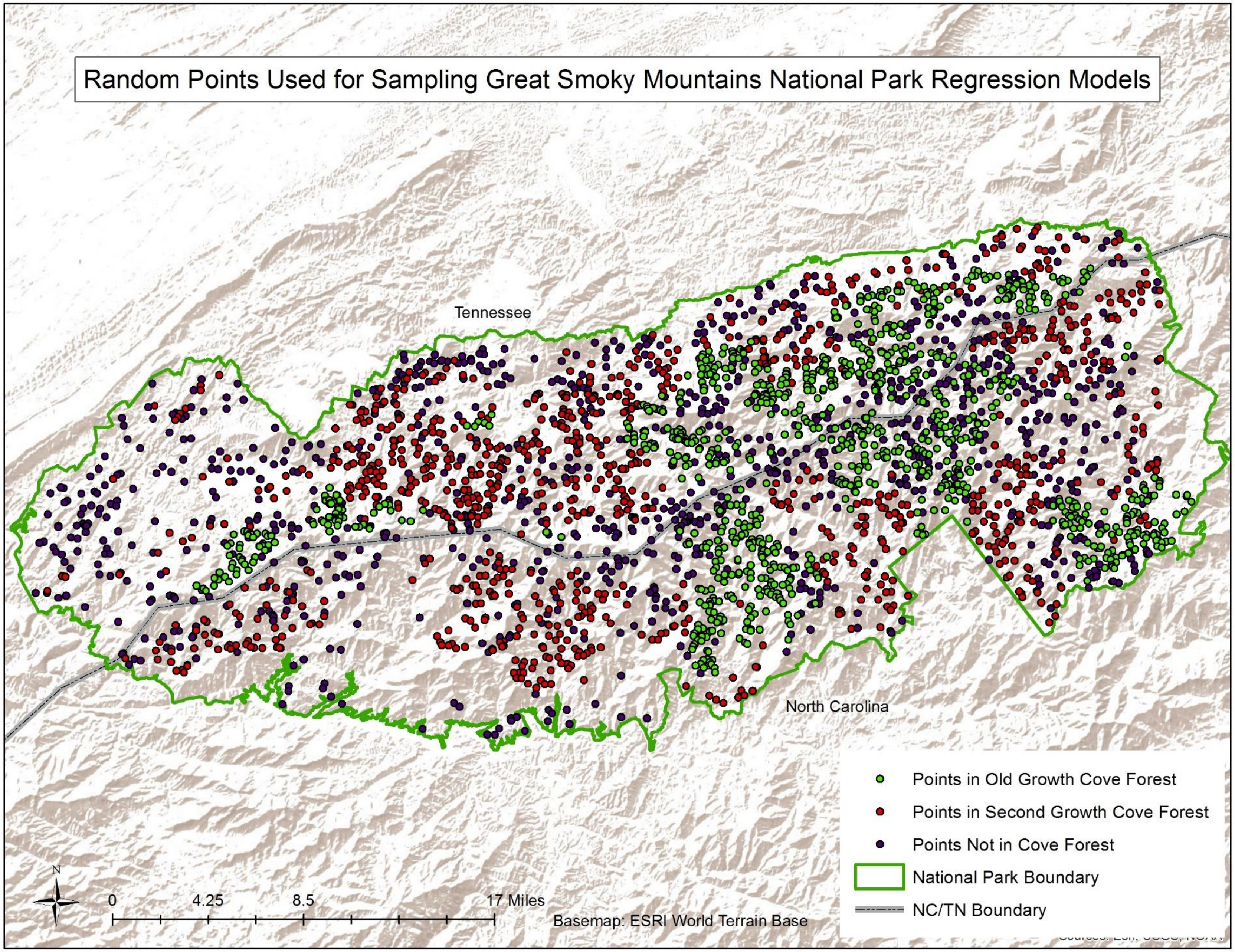
- All points were spaced a minimum of 100ft apart (to eliminate overlapping windows)
- 500 points were generated in primary cove forests. These points were confined to cove forests within the "undisturbed" section of the Park designated by the Pyle surveys. Cove forests included both the ecozones and the GSMNP cove ecogroup polygons, which were merged and dissolved into a single constraining feature class.
- 500 points were generated in second growth cove forest. These points were confined to cove forests within the "heavy cut" and "settlement" (most intensely disturbed) sections of the Park designated by the Pyle surveys. Cove forests included both the ecozones and the GSMNP cove ecogroup polygons, which were merged and dissolved into a single constraining feature class.
- 500 points were generated outside of cove forests. To create the constraining feature class for this layer, I used the ArcGIS Erase tool to erase all cove forests (ecozones and GSMNP cove vegetation polygons) from the park. These cove forests were buffered by 30M prior to running the Erase tool to ensure that no window overlap occurred between the cove and non-cove sampling areas.

Pisgah and Nantahala National Forests (2100 points):

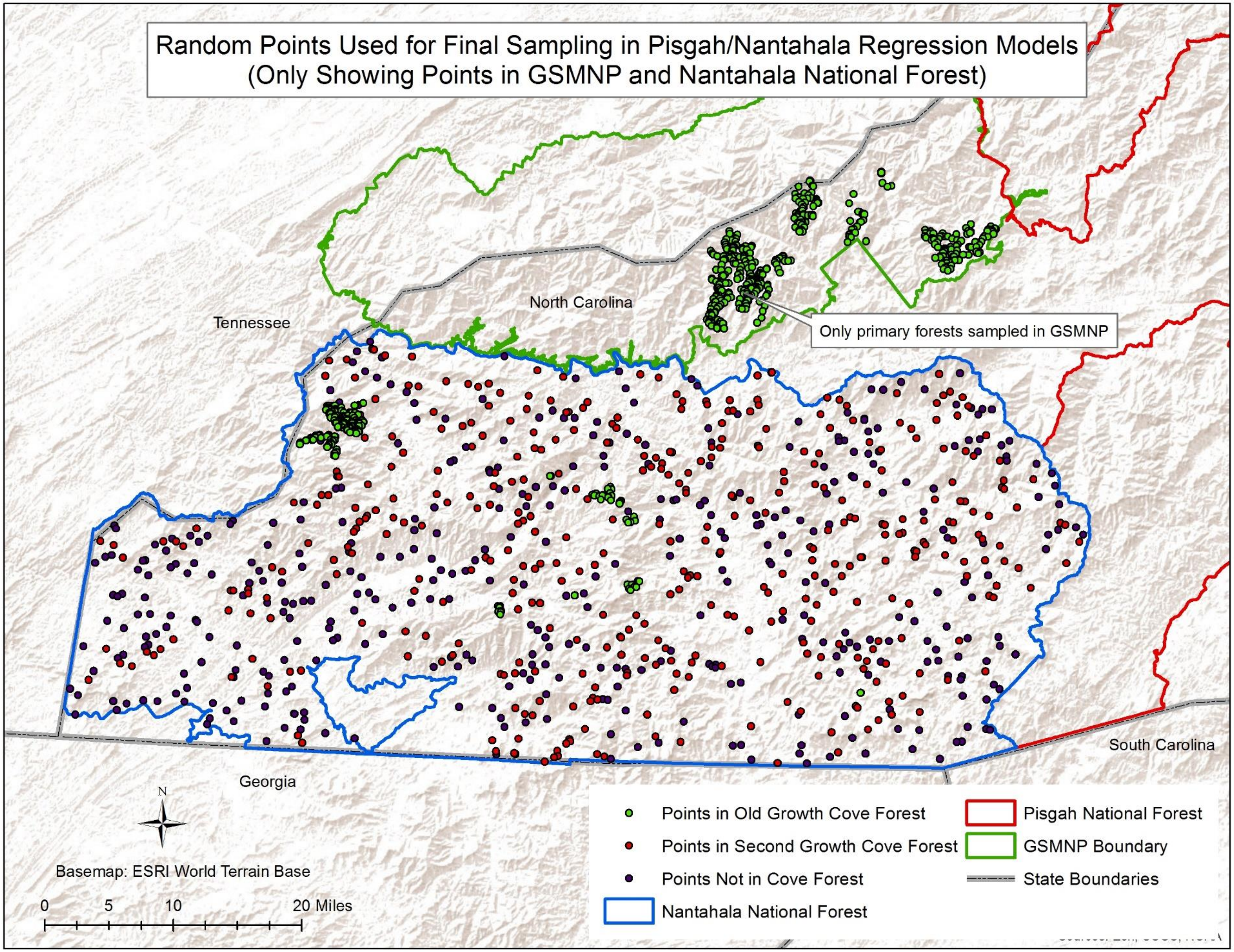
- All points in the National Forests were spaced a minimum of 90ft apart (to eliminate overlapping windows)
- All points in GSMNP were spaced a minimum of 100ft apart (to eliminate overlapping windows)
- 400 points were generated in primary cove forests on the NC side of GSMNP. These points were confined to cove forests within the "undisturbed" section of the Park designated by the Pyle surveys. Cove forests included both the ecozones and the GSMNP cove ecogroup polygons, which were merged and dissolved into a single constraining feature class.
- 200 points were generated in old growth cove forests in Nantahala National Forest. These were confined to a collection of Class A old growth sites shown on the National Forest old growth surveys (Messick et al 2000). In some locations, these surveys were edited by Mountaintrue Public Lands Field Biologist Josh Kelly based on field experience. A copy of these edited old growth sites can be seen in Appendices 2-3 Cove forests included only the ecozone polygons, which were clipped to an area below 4200 feet to eliminate overlap with Spruce Fir and Northern Hardwood ecozones.
- 100 points were generated in old growth cove forests in the Appalachian and Pisgah Ranger Districts of Pisgah National Forest. These were confined to a collection of Class A old growth sites shown on the National Forest old growth surveys (Messick et al. 200). In some locations, these surveys were edited by Mountaintrue Public Lands Field Biologist Josh Kelly based on field experience. A copy of these old growth sites can be seen in Appendices 2-3. Cove forests included only the ecozone polygons, which were clipped to an area below 4200 feet to eliminate potential overlap with Spruce Fir and Northern Hardwood ecozones.
- 400 points were generated in second growth cove forests in Nantahala National Forest. These points were constrained to cove forest ecozones, which were clipped to an area below 4200 feet to eliminate potential overlap with Spruce Fir and Northern Hardwood ecozones. Old growth sites from the National Forest old growth surveys, as well as all Significant Natural Heritage Areas (NC Natural Heritage Program) were excluded from sampling for second growth sites.

- 300 points were generated in second growth cove forests in Pisgah National Forest. These points were constrained to cove forest ecozones, which were clipped to an area below 4200 feet to eliminate potential overlap with Spruce Fir and Northern Hardwood ecozones. Old growth sites from the National Forest old growth surveys, as well as all Significant Natural Heritage Areas (NC Natural Heritage Program) were excluded from sampling for second growth sites.
- 400 points were generated outside of cove forests in Nantahala National Forest. These points were constrained to the National Forest boundary. In order to exclude cove forests, the ArcGIS Erase tool was used to erase all cove forest ecozones, all old growth sites, and all Significant Natural Heritage areas.
- 300 points were generated outside of cove forests in Nantahala National Forest. These points were constrained to the National Forest boundary. In order to exclude cove forests, the ArcGIS Erase tool was used to erase all cove forest ecozones, all old growth sites, and all Significant Natural Heritage areas.

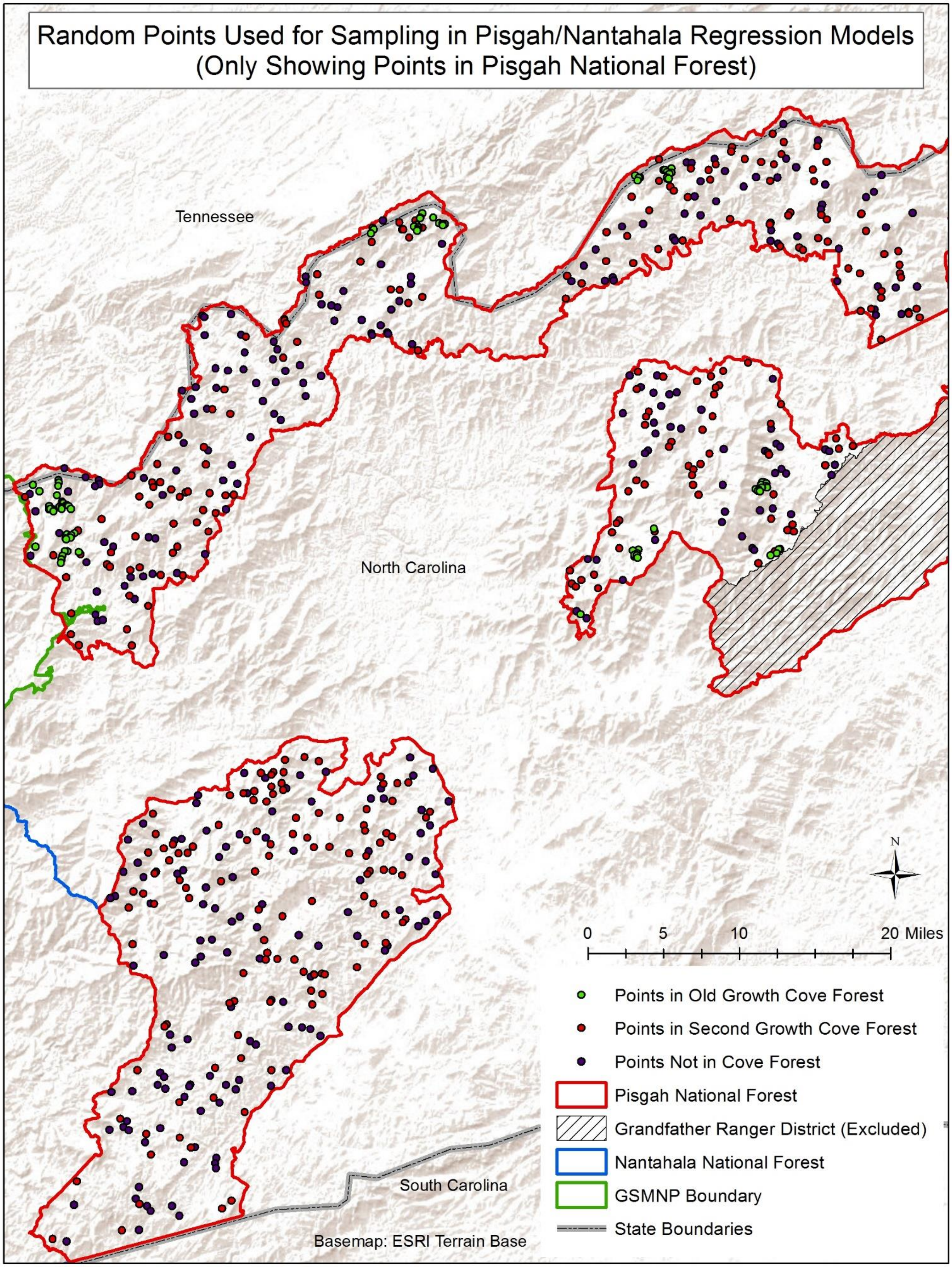
After this collection of random points was generated, the ArcGIS Extract Multi-Values to Points tool was used to extract all raster values into the attribute tables for the random point files. These tables were then saved as spreadsheets in Microsoft Excel. Maps of the random points are shown in Maps 29-31.



Map 29: Sampling Points for GSMNP Models



Map 30: Random Points for Pisgah/Nantahala Model – Nantahala National Forest and GSMNP



Map 31: Random Points for Pisgah/Nantahala Model – Pisgah National Forest

D. Statistical Comparisons

In order to test for statistical differences between the different forest disturbance classes (old growth cove forest, second growth cove forest, non-cove-forest land cover), a series of nonparametric Kruskal-Wallis tests were conducted comparing the values extracted from the random points. This nonparametric option was used in lieu of a regular ANOVA analysis because none of the metrics being compared had normally distributed data or approximately equal variance across all three forest community/disturbance classes. In total, three separate tests were conducted:

1. A comparison between old growth/primary cove forests, second growth cove forests, and non-cove-forest land cover for random points on the TN side of GSMNP
2. A comparison between old growth/primary cove forests, second growth cove forests, and non-cove-forest land cover for random points on the NC side of GSMNP
3. A comparison between old growth/primary cove forests, second growth cove forests, and non-cove-forest land cover for random points in Pisgah/Nantahala National Forests, including 400 primary forest points from the NC side of GSMNP.

Kruskal-Wallis tests were performed in JMP Pro 12 software. Each series of tests compared Entropy, Contrast, Homogeneity, Dissimilarity, Second Moment, Variance, and Mean Height variables between old growth/primary cove forests, second growth cove forests, and non-cove-forest land cover. The results of these analyses are shown in Tables 1-3.

E. Regression Modeling

The three regression models created in this analysis were generated using values derived from the three sets of random points described in Section 4C. These models were created in the JMP Pro 12 Fit Model platform. Due to the differences in the texture values generated by the different measurement units (meters vs. feet), the analysis uses three separate models corresponding to the three sets of random points (GSMNP NC, GSMNP TN, and Nantahala/Pisgah National Forests).

Entropy, Contrast, Dissimilarity, Homogeneity, Variance, Second Moment, and Mean Height were input as the X variables for each model, and the Y-variable response class was set to the disturbance category (old growth cove, second growth cove, non-cove). The "Effect Likelihood Ratio Tests" feature in the Fit Model platform was then used to determine which variables were significant components of the models. Any metrics with a Prob>ChiSq value greater than 0.05 were removed from the model one at a time in descending order from the highest Prob>ChiSq value, until all components of the model had an effect likelihood Prob>ChiSq value less than 0.05. Interestingly, this resulted in each model containing a slightly different combination of texture metrics:

- The GSMNP TN model (in meters) included Entropy, Homogeneity, Dissimilarity, Second Moment, Variance, and Mean Height
- The GSMNP NC model (in feet) included Entropy, Homogeneity, Dissimilarity, Contrast, Variance, and Mean Height
- The Pisgah/Nantahala Model (in feet) included Entropy, Homogeneity, Dissimilarity, Contrast, and Mean Height

The probability formulas for each model were saved in the JMP spreadsheet. These formulas were then entered into the ArcGIS Raster Calculator feature to create a series of inputs for the final model layers. The probability formulas for each model are listed below:

GSMNP TN:

- **Lin(SecondGrowth):** 35.0581112176678 + -6.50974893634909 * :Entropy + -20.6371647309377 * :Homogeneity + -5.23948350438013 * :Dissimilarity + -16.0568617764548 * :Second Moment + 1.30168154366883 * :Variance + 0.00407880556733335 * :Mean Height
- **Lin(NotCove):** 22.9810128104994 + -4.78364779049221 * :Entropy + -3.52335564134567 * :Homogeneity + -3.03492959069492 * :Dissimilarity + -15.439122951911 * :Second Moment + 0.5544847582633 * :Variance + -0.243856545064267 * :Mean Height
- **Prob(SecondGrowth):** 1 / (1 + Exp(-:Name("Lin[SecondGrowth]")) + Exp(:Name("Lin[Not Cove]") - :Name("Lin[SecondGrowth]")))
- **Prob(NotCove):** 1 / (1 + Exp(:Name("Lin[SecondGrowth]") - :Name("Lin[Not Cove]")) + Exp(-:Name("Lin[Not Cove]")))
- **Prob(OldGrowth):** 1 / (1 + Exp(:Name("Lin[SecondGrowth]")) + Exp(:Name("Lin[Not Cove]")))

GSMNP NC:

- **Lin(SecondGrowth):** 19.3899570065015 + 0.19256473393693 * :Contrast + -15.9305478921172 * :Homogeneity + -1.8257164976616 * :Entropy + -4.38692657761933 * :Dissimilarity + 0.0299098844576163 * :Variance + 0.0339506049962291 * :Mean Height
- **Lin(NotCove):** 13.2316484342774 + 0.172969541388208 * :Contrast + -1.24095045596307 * :Homogeneity + -1.34242569041997 * :Entropy + -1.3930098080148 * :Dissimilarity + -0.0574719330450976 * :Variance + -0.0719309417616239 * :Mean Height
- **Prob(Not Cove):** 1 / (1 + Exp(:Name("Lin[SecondGrowth]") - :Name("Lin[Not Cove]")) + Exp(-:Name("Lin[Not Cove]")))
- **Prob(SecondGrowth):** 1 / (1 + Exp(-:Name("Lin[SecondGrowth]")) + Exp(:Name("Lin[Not Cove]") - :Name("Lin[SecondGrowth]")))
- **Prob(OldGrowth):** 1 / (1 + Exp(:Name("Lin[SecondGrowth]")) + Exp(:Name("Lin[Not Cove]")))

Pisgah/Nantahala:

- **Lin(Not Cove):** 17.8219877711164 + 0.0988233261362635 * :Contrast + -2.65474683232448 * :Dissimilarity + -0.797521312716264 * :Entropy + -17.5786420662294 * :Homogeneity + -0.0256958423242377 * :Mean Height
- **Lin(Old Growth):** (-31.7556774079) + 0.192326116379111 * :Contrast + -5.27237252492664 * :Dissimilarity + 6.68936325817401 * :Entropy + 11.8287564887692 * :Homogeneity + 0.0804664298012745 * :Mean Height
- **Prob(Not Cove):** 1 / (1 + Exp(-:Name("Lin[Not Cove]")) + Exp(:Name("Lin[Old Growth]") - :Name("Lin[Not Cove]")))
- **Prob(Old Growth):** 1 / (1 + Exp(:Name("Lin[Not Cove]") - :Name("Lin[Old Growth]")) + Exp(-:Name("Lin[Old Growth]")))
- **Prob(SecondGrowth):** 1 / (1 + Exp(:Name("Lin[Not Cove]")) + Exp(:Name("Lin[Old Growth]")))

The probability layers for each model were then combined in the Raster Calculator using the following formula:

Con(("Prob_OldGrowth" > "Prob_SecondGrowth") & ("Prob_OldGrowth " > "Prob_NotCove"),1,(Con(("Prob_SecondGrowth " > "Prob_OldGrowth ") & ("Prob_ SecondGrowth" > "Prob_NotCove"),2,(Con(("Prob_NotCove" > "Prob_OldGrowth ") & ("Prob_Not_Cove" > "Prob_SecondGrowth "),3,0)))))

This formula assigns a value to each cell that corresponds to whichever disturbance class received the highest probability in the regression model. The output of the formula is a final model with the following values:

0: Unassigned Values
1: Highest Probability is Old Growth Cove Forest
2: Highest Probability is Second Growth Cove Forest
3: Highest Probability is Not Cove Forest

The "Confusion Matrix" feature in the JMP platform was used to test each model's ability to predict the random point inputs. These error matrices are shown in Tables 5-7.

Section V: Results and Discussion

My results indicate that GLCM texture metrics can be useful tools for analyzing horizontal forest structure in LiDAR canopy height models. The primary and old growth cove forests sampled generally showed higher proportions of entropy, contrast, variance, and dissimilarity, and showed lower proportions of homogeneity and second moment than second growth cove forests. Combined with a simple mean height statistic, these texture measures proved to be approximately 60% accurate at classifying areas as old growth cove forest, second growth cove forest, and non-cove-forest land cover (see Section 5-F). However, my results also show that the GLCM's usefulness as a landscape-wide forest analysis technique requires that LiDAR data be collected and processed with consistent methodology and data sampling intensity.

A. Comparing Trends in Horizontal Forest Structure Across GSMNP
The following series of charts show trends in various LiDAR-derived metrics for the different disturbance history classes outlined by the 1988 Pyle surveys. All data were derived via direct pixel extractions of each texture layer, with all layers representing cove forests delineated by the GSMNP vegetation mapping vector data (Madden et al. 2004). All cove forest polygons were buffered -30m in Tennessee and -100ft in North Carolina in order to eliminate overlap with non-cove forest communities within the moving window. All data are divided into the five anthropogenic disturbance classes used in the 1988 C. Pyle surveys. For a description and map of each disturbance class, see Section 2-B and Map 3. The area contained within each anthropogenic disturbance class varies greatly:

For Tennessee:
- Undisturbed: 1502423 cells = 1352 HA
- Selective Cut: 727212 cells = 654 HA
- Light Cut: 422817 cells = 380 HA
- Heavy Cut: 1026981 cells = 924 HA
- Settlement: 967944 cells = 871 HA

For North Carolina:
- Undisturbed: 1363263 cells = 3129 Acres
- Selective Cut: 587621 cells = 1348 Acres
- Light Cut: 2448885 cells = 5621 Acres
- Heavy Cut: 1885176 cells = 4327 Acres
- Settlement: 1270393 cells = 2916 Acres

To account for the variability in the area covered by each class, Figures 4-19 show each texture metric in terms of the percent of the total area covered within each anthropogenic disturbance class.

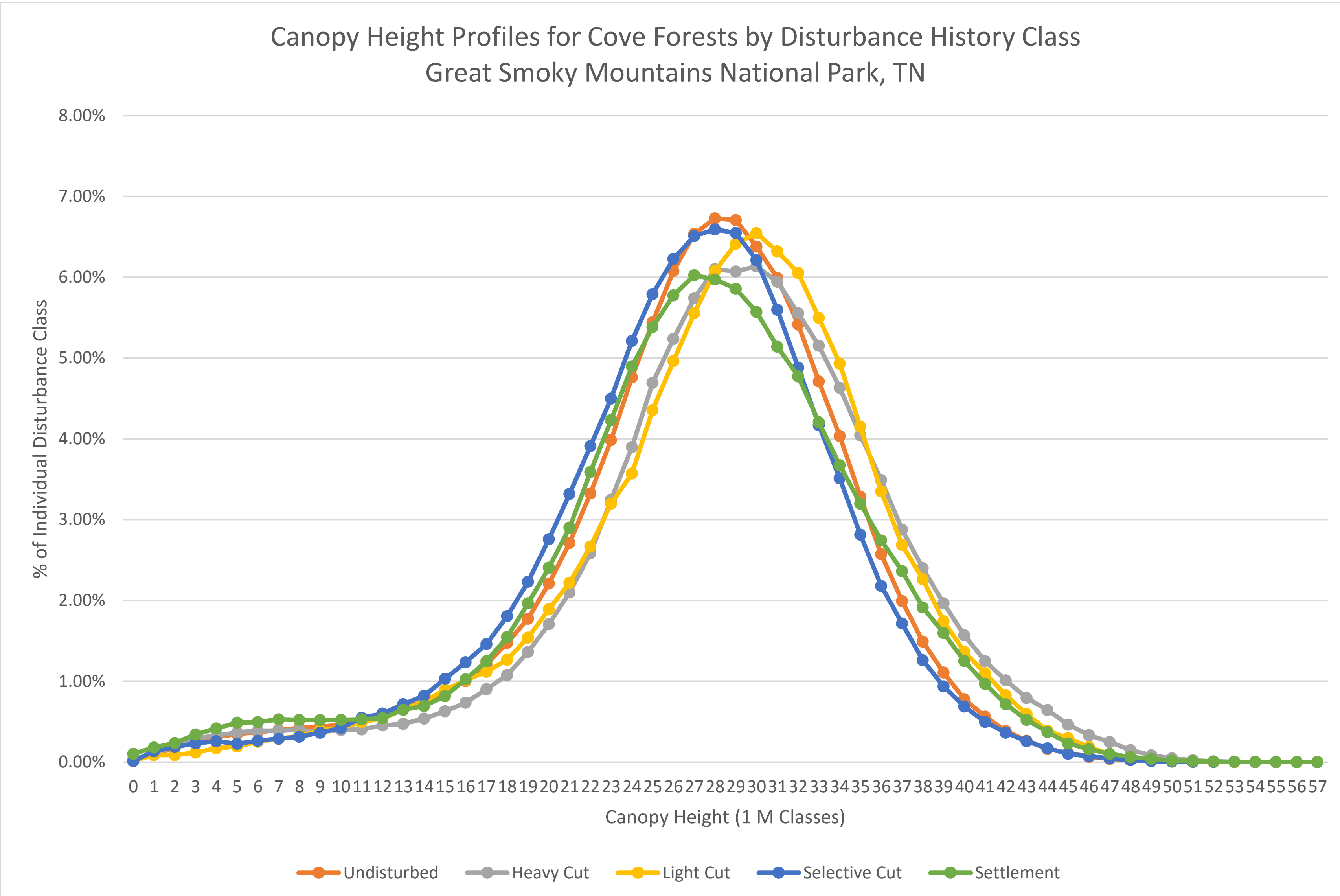


Figure 4: Canopy Height Profiles in GSMNP TN

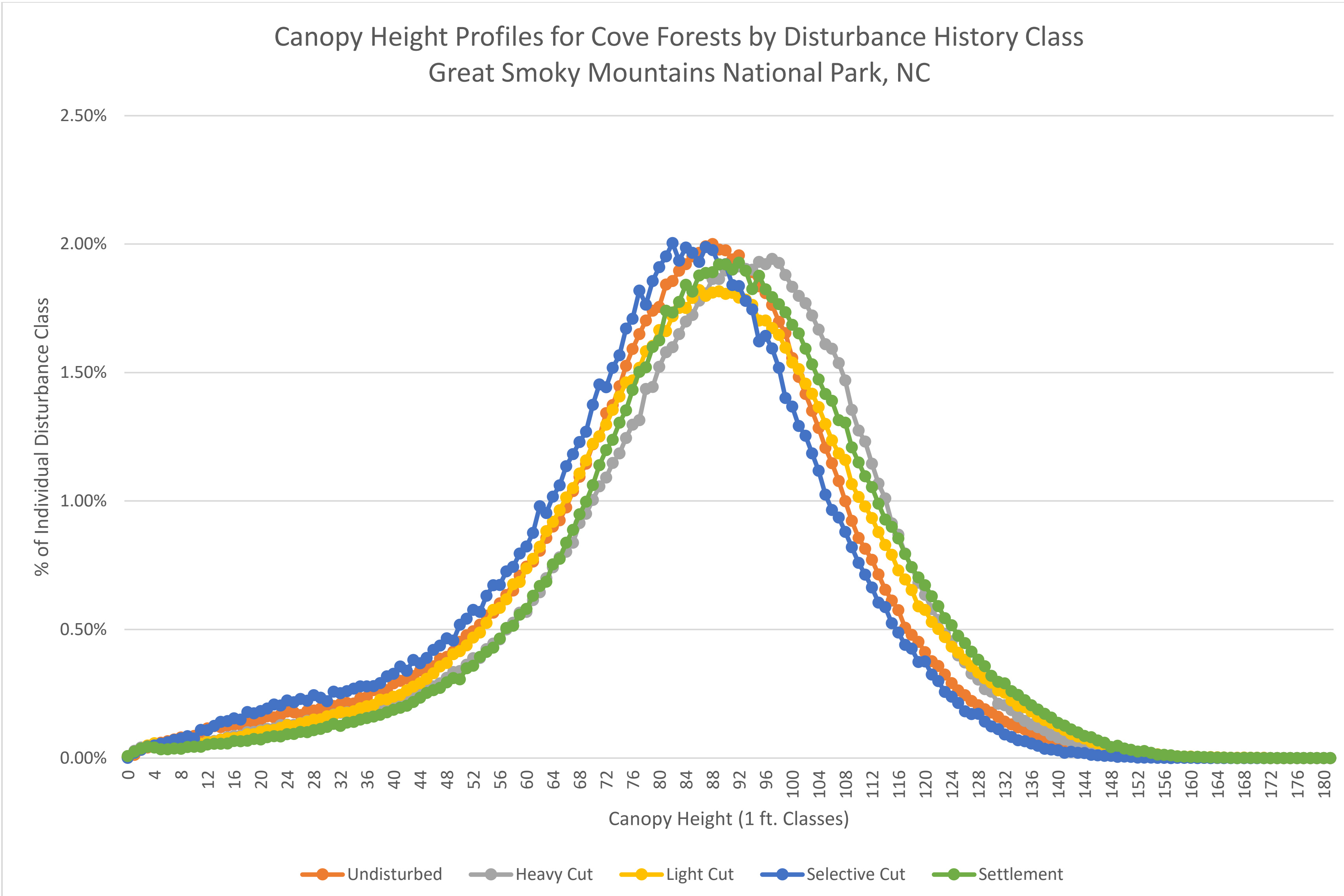


Figure 5: Canopy Height Profiles in GSMNP NC

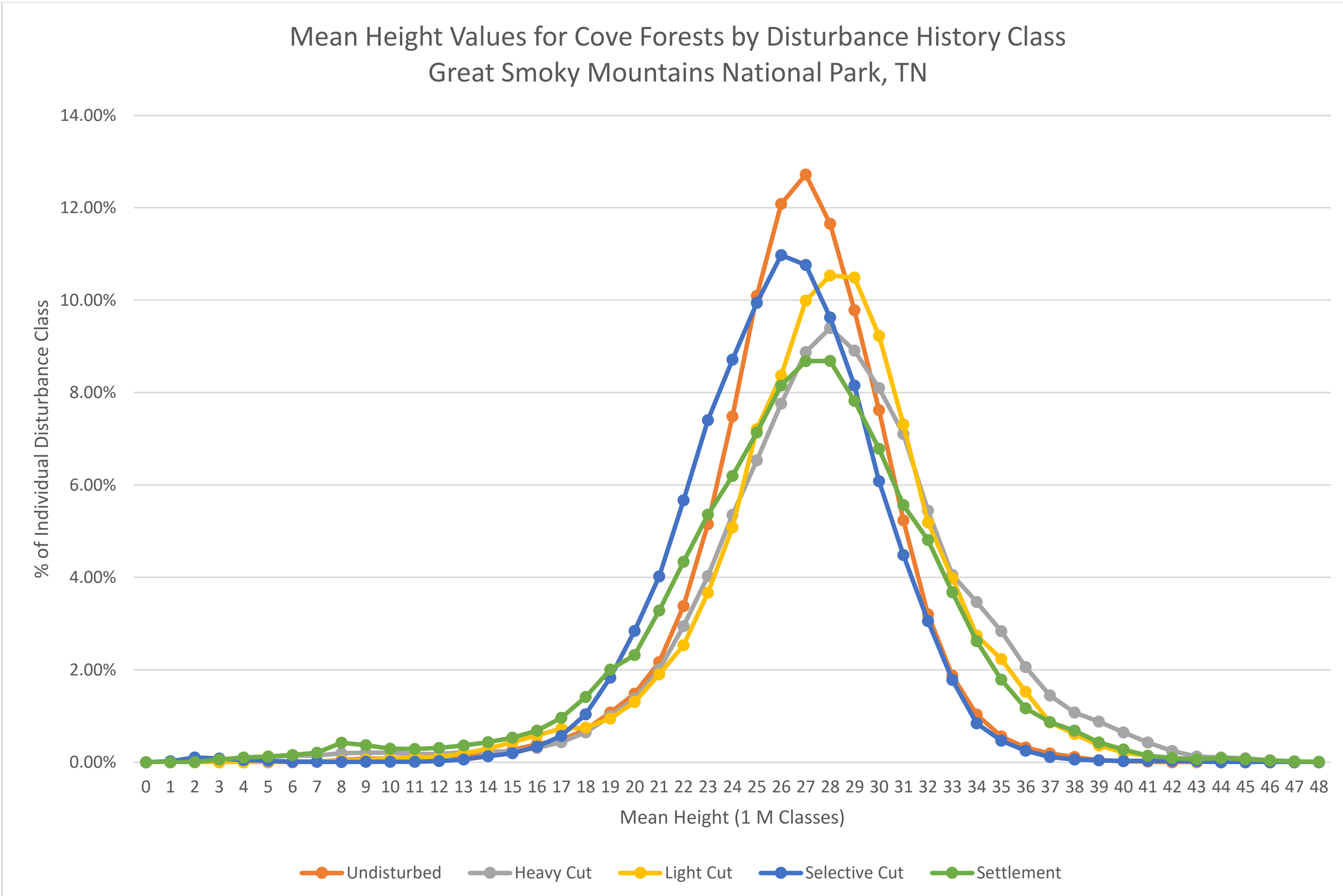


Figure 6: Height Profiles in GSMNP TN

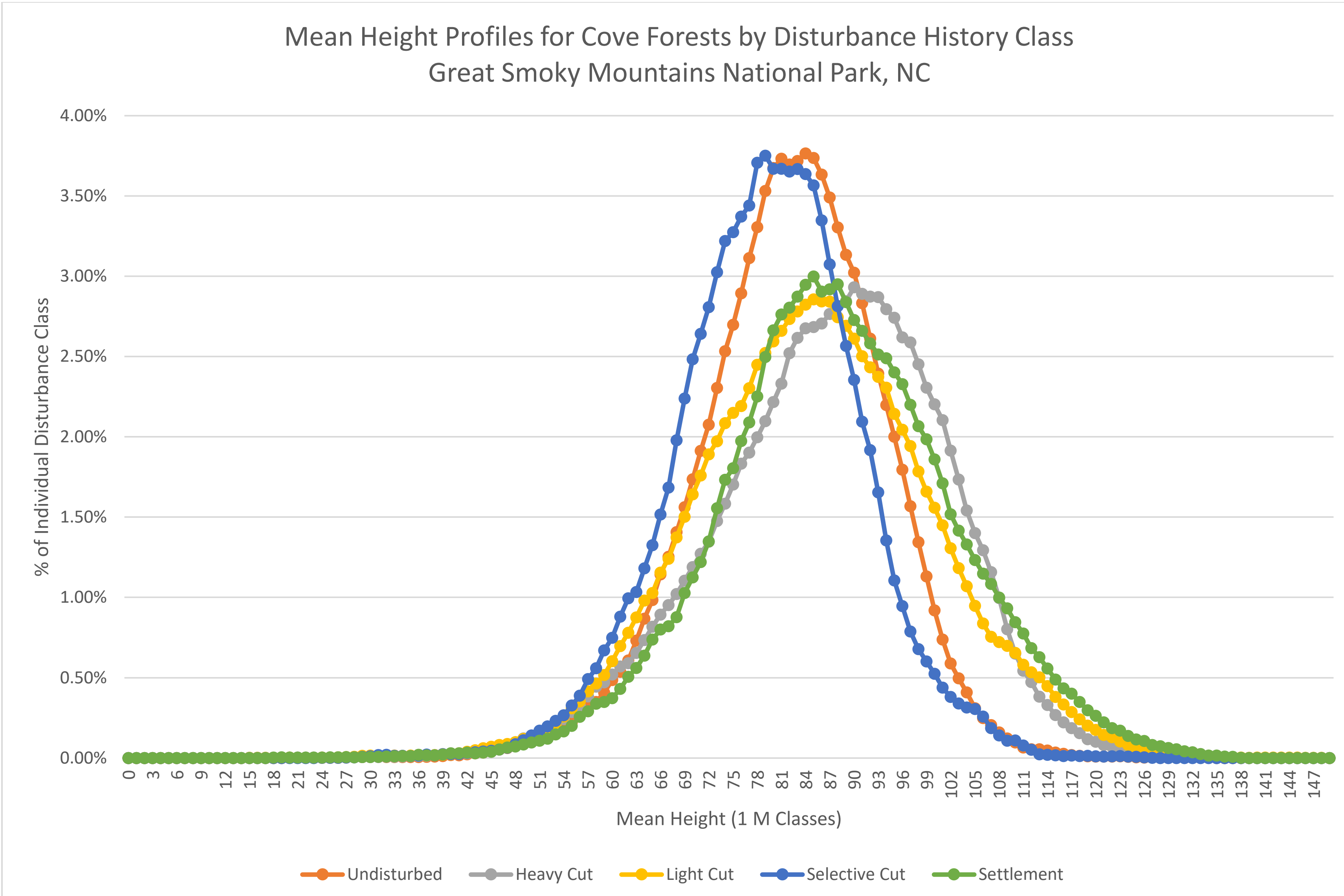


Figure 7: Height Profiles in GSMNP NC

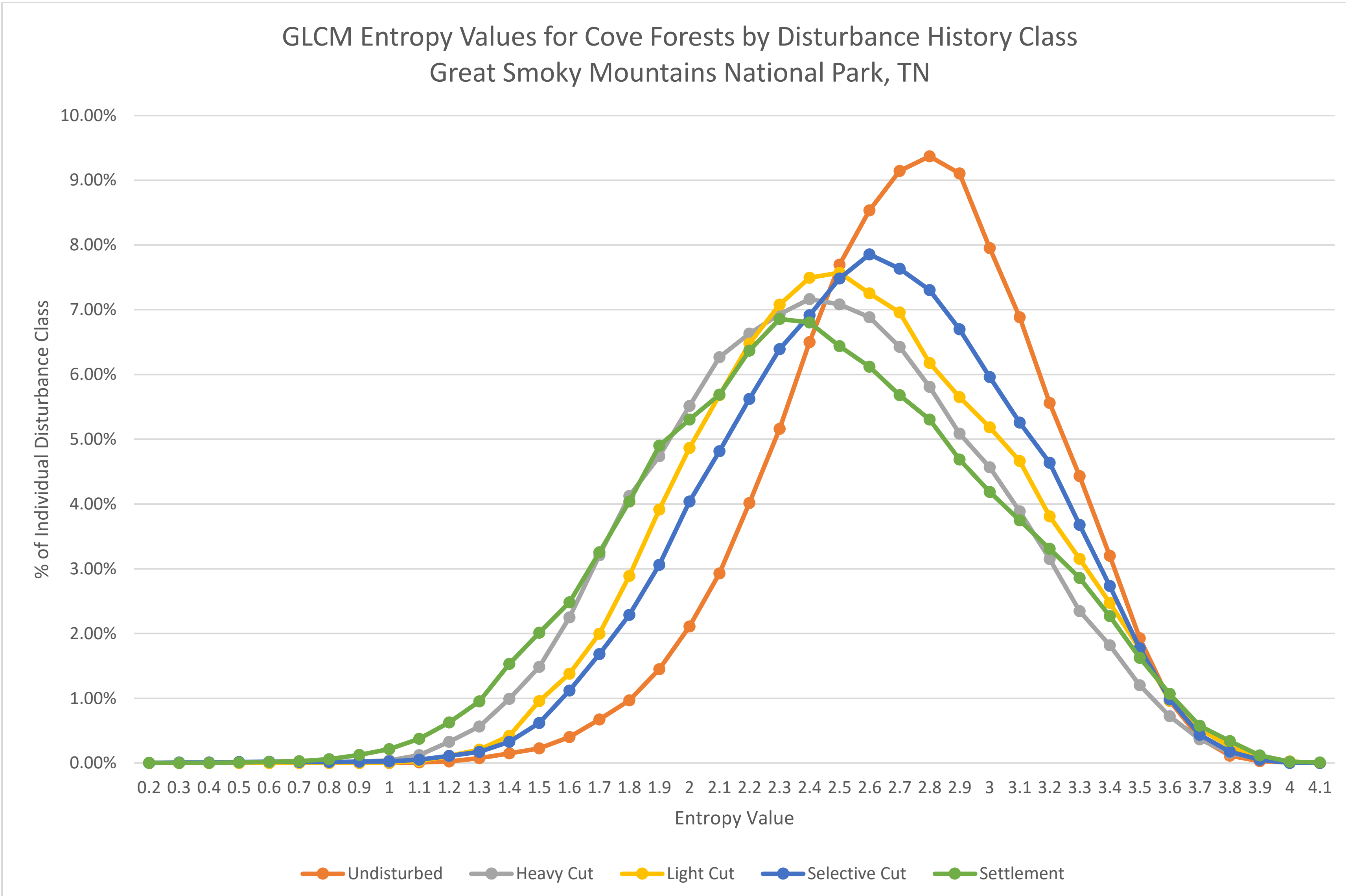


Figure 8: Entropy Profiles in GSMNP TN

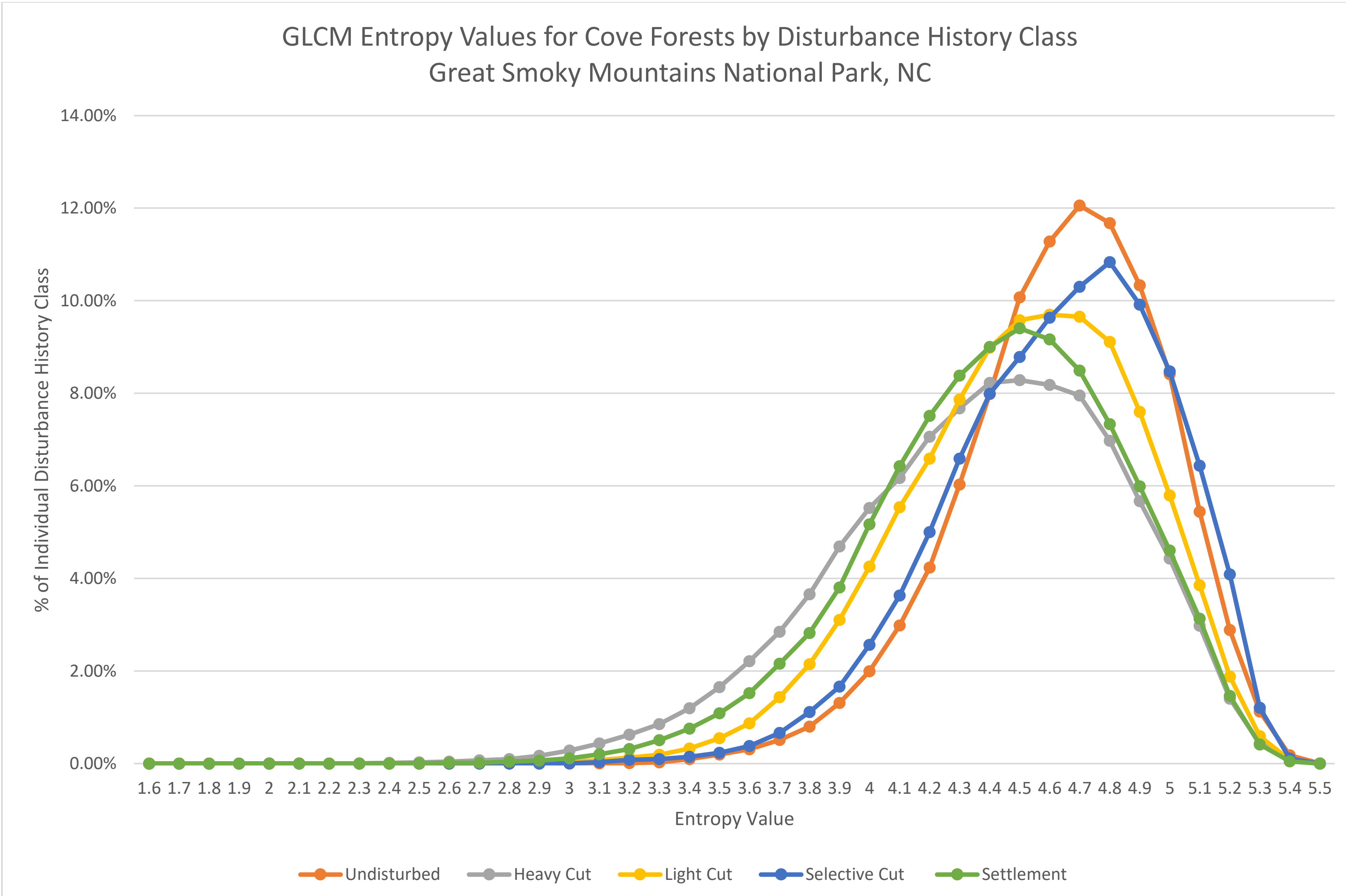


Figure 9: Entropy Profiles in GSMNP NC

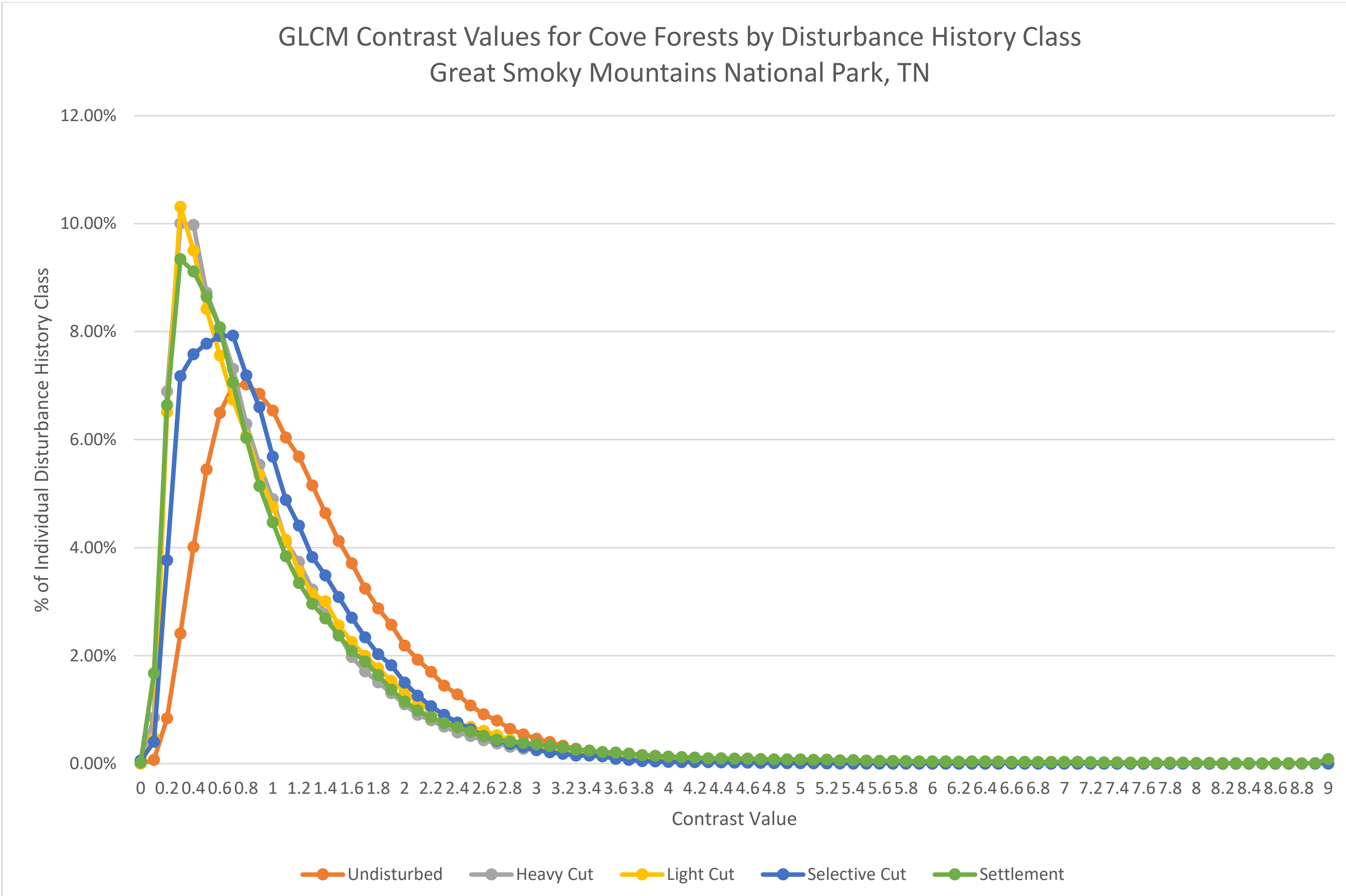


Figure 10: Contrast Profiles for GSMNP TN

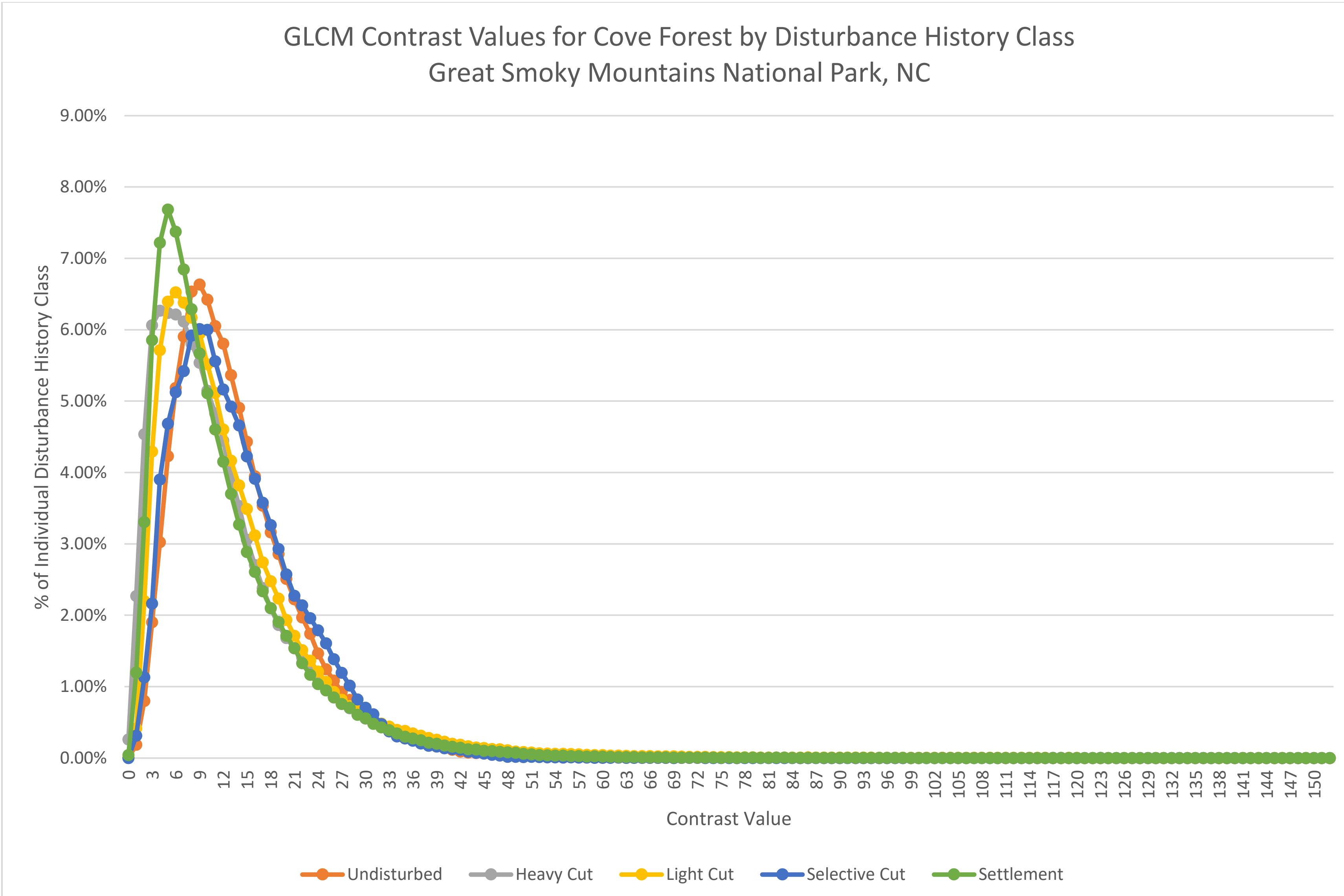


Figure 11: Contrast Profiles for GSMNP NC

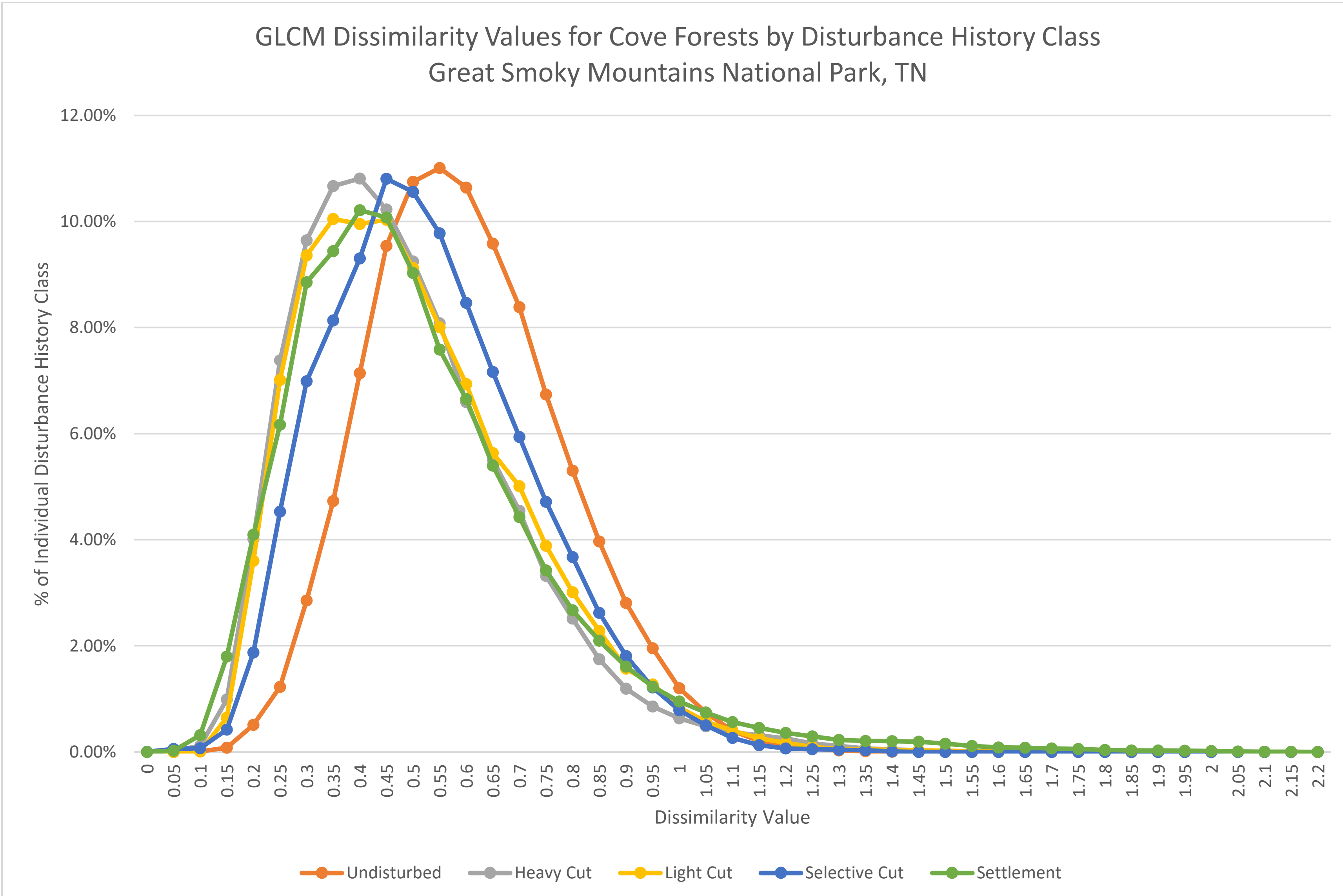


Figure 12: Dissimilarity Profiles for GSMNP TN

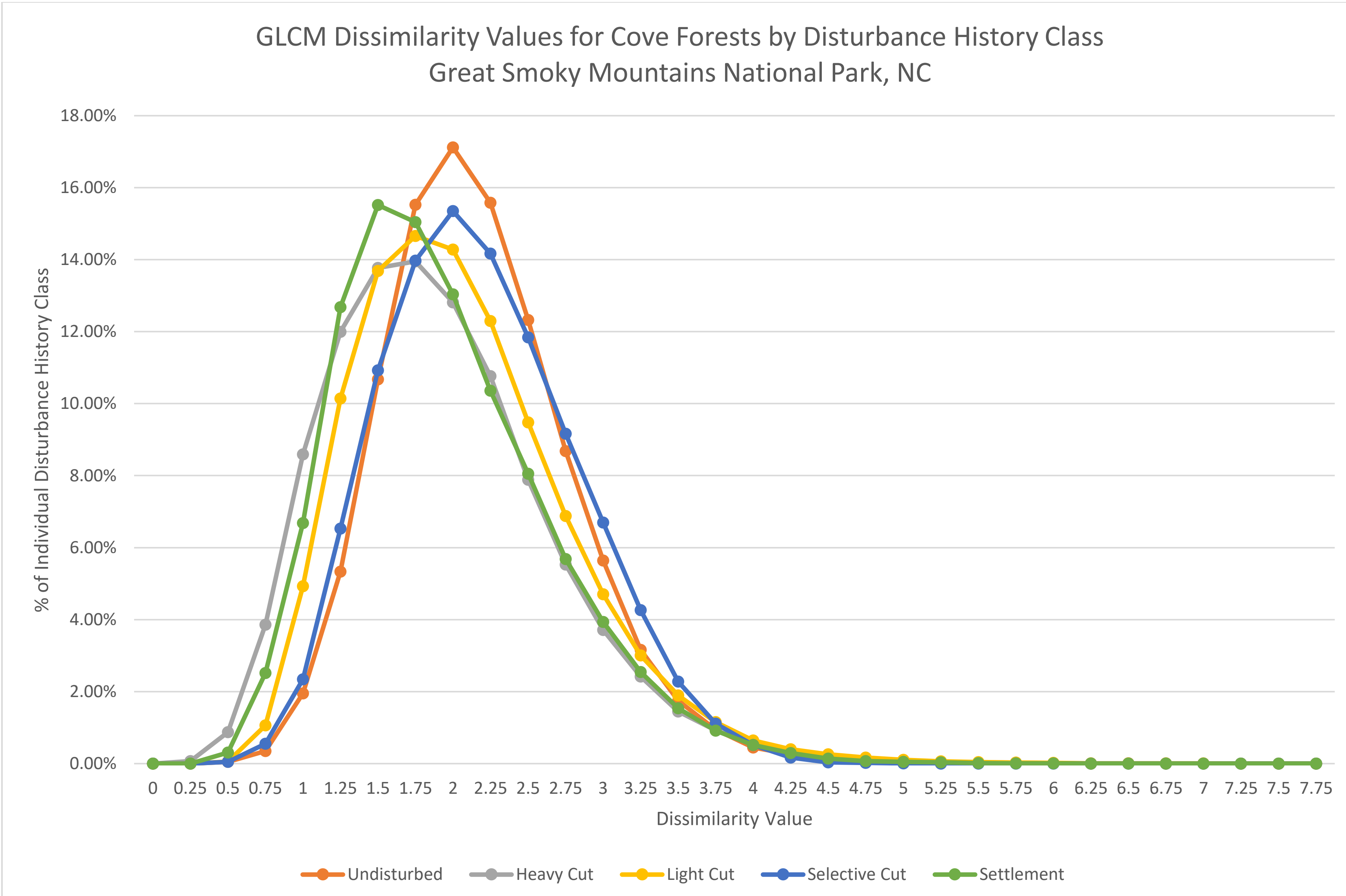


Figure 13: Dissimilarity Profiles for GSMNP NC

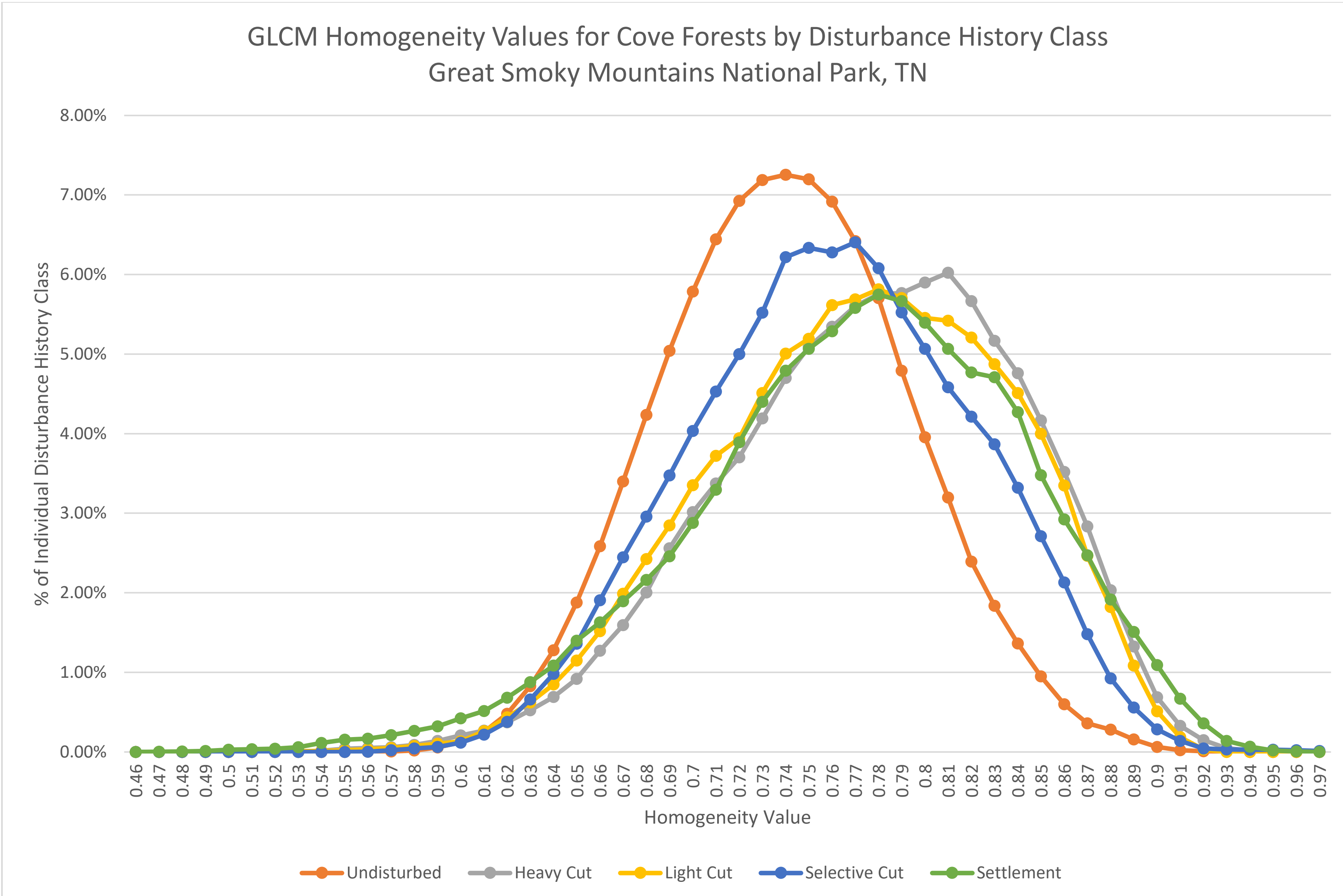


Figure 14: Homogeneity Profiles for GSMNP TN

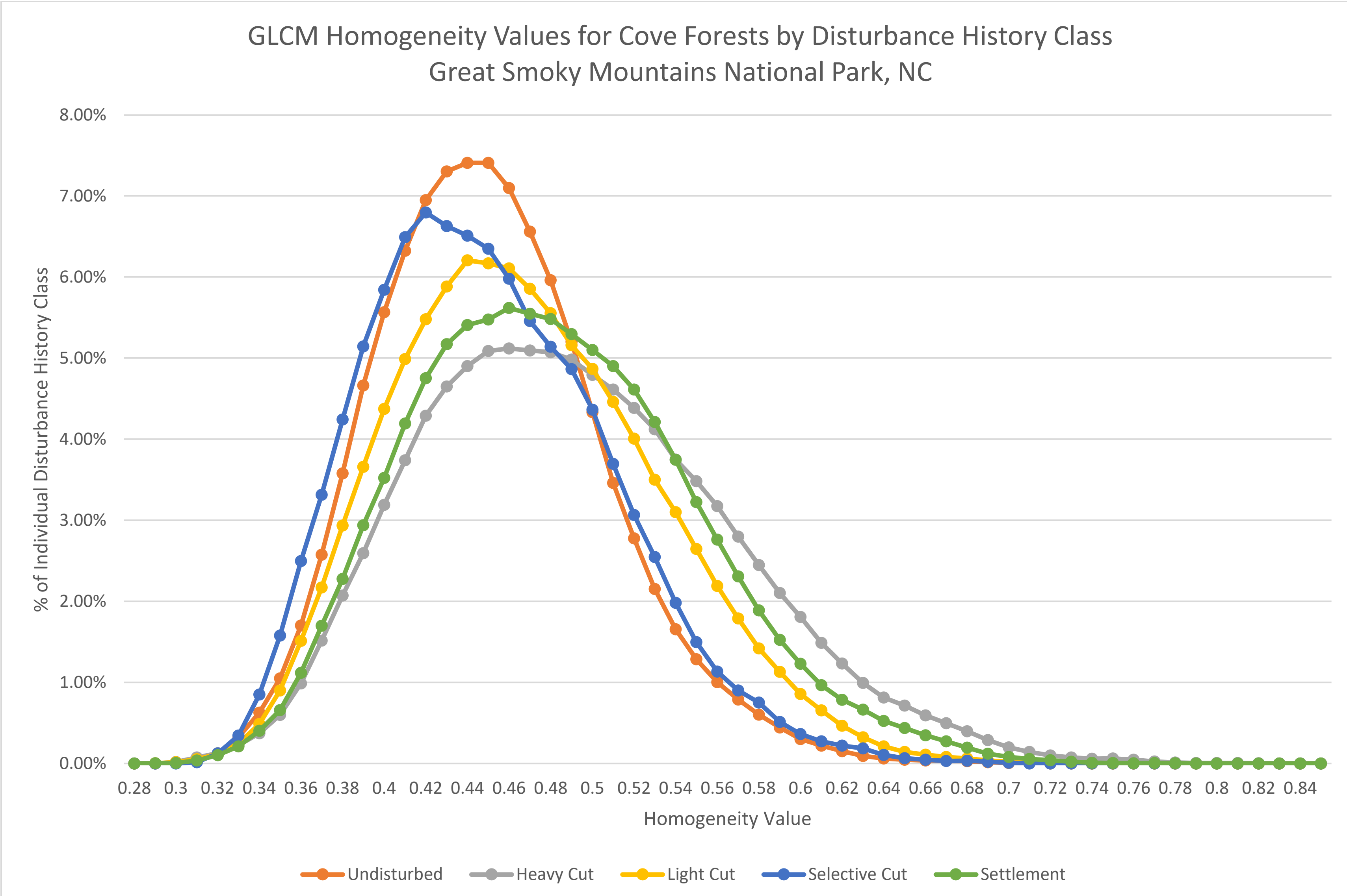


Figure 15: Homogeneity Profiles for GSMNP NC

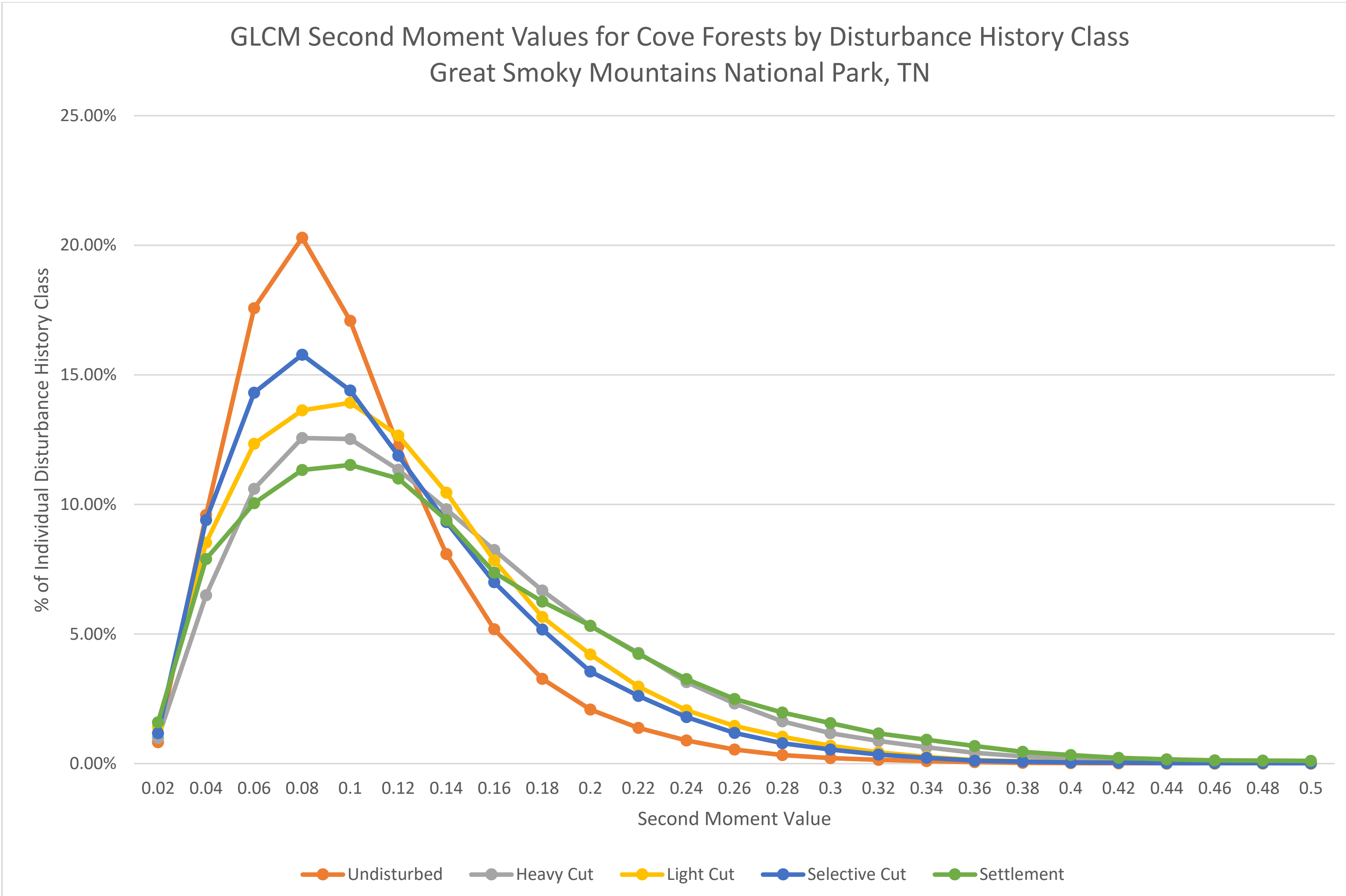


Figure 16: Second Moment Profiles for GSMNP TN

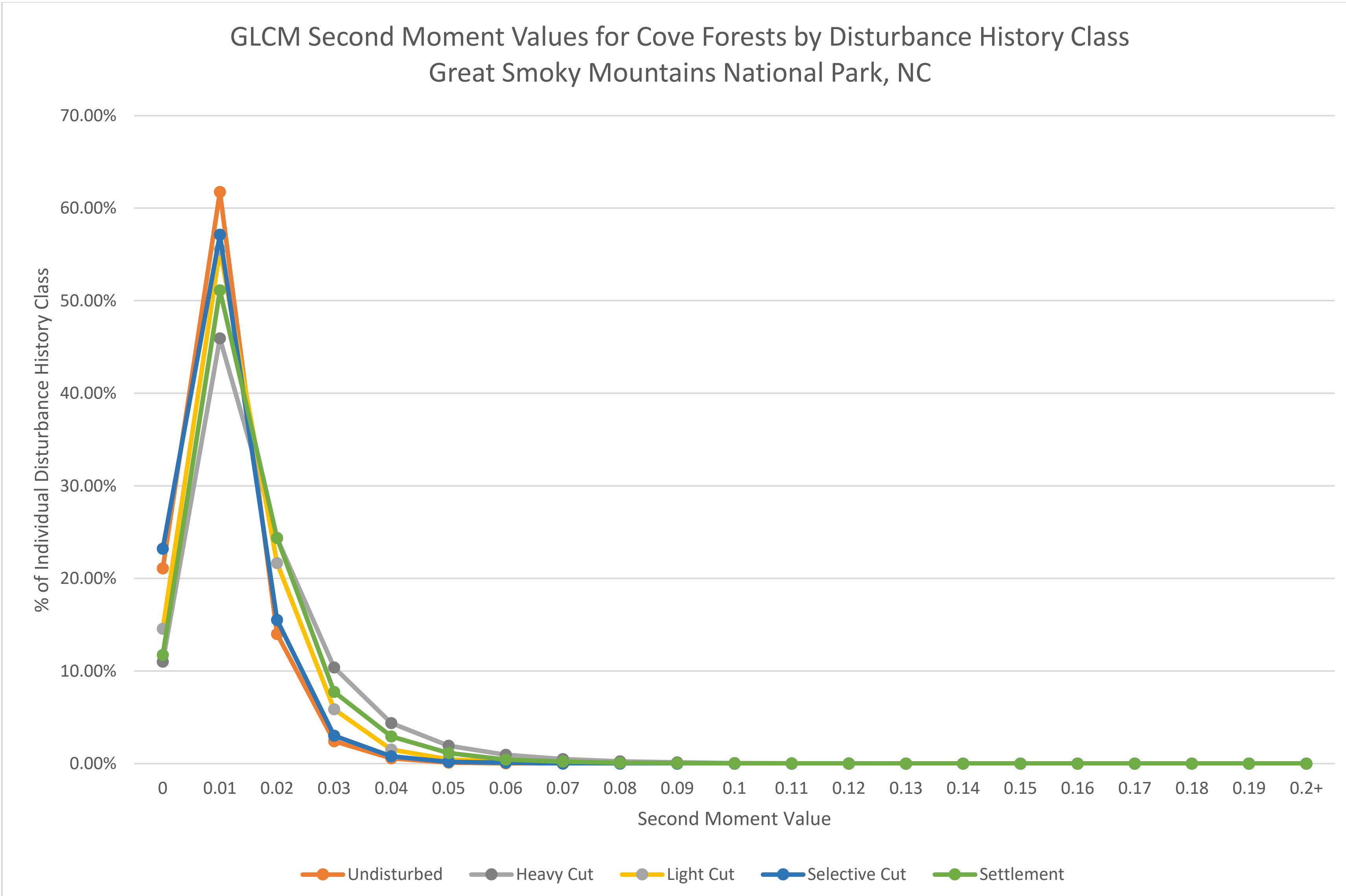


Figure 17: Second Moment Profiles for GSMNP NC

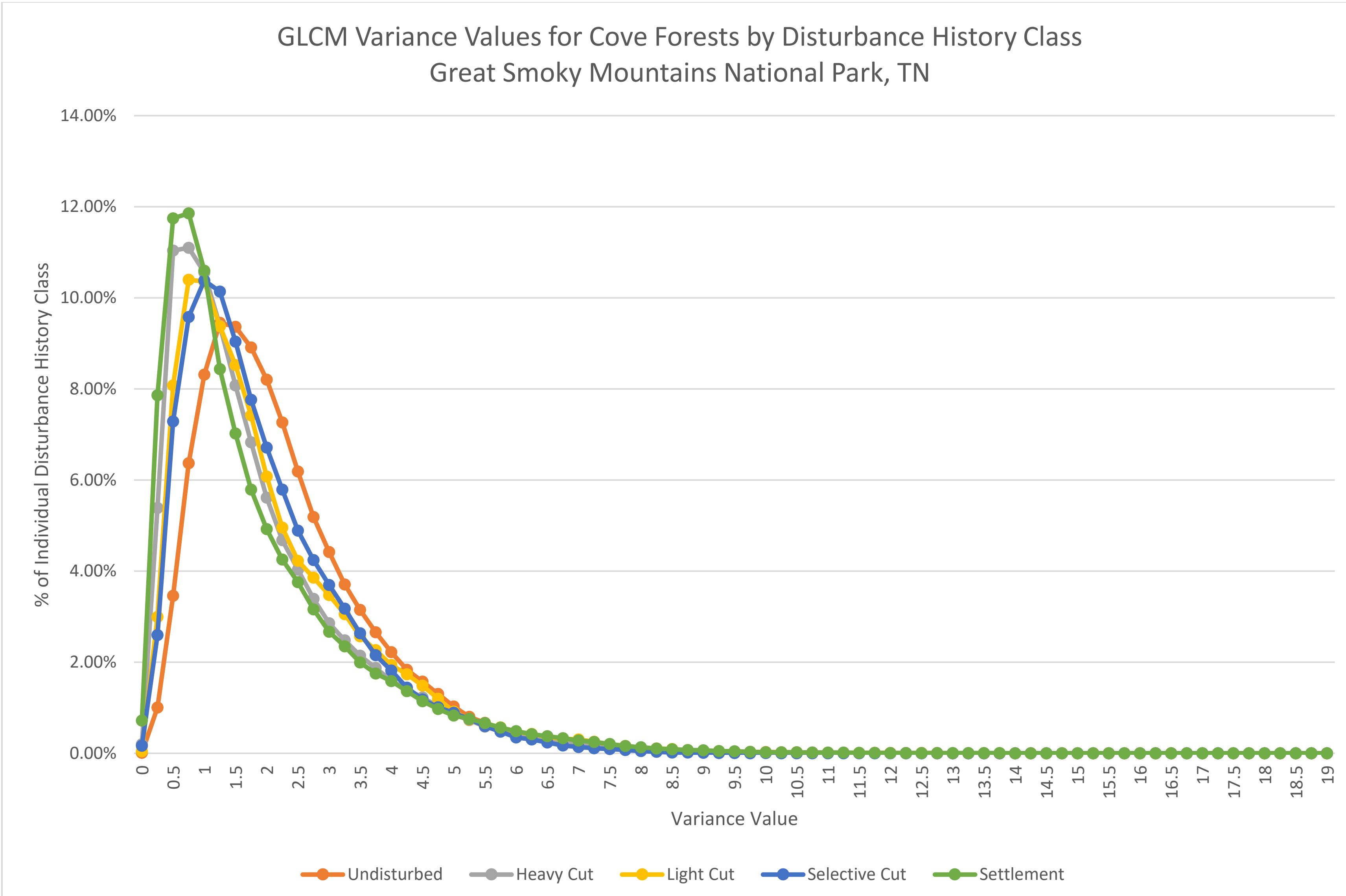


Figure 18: Variance Profile for GSMNP TN

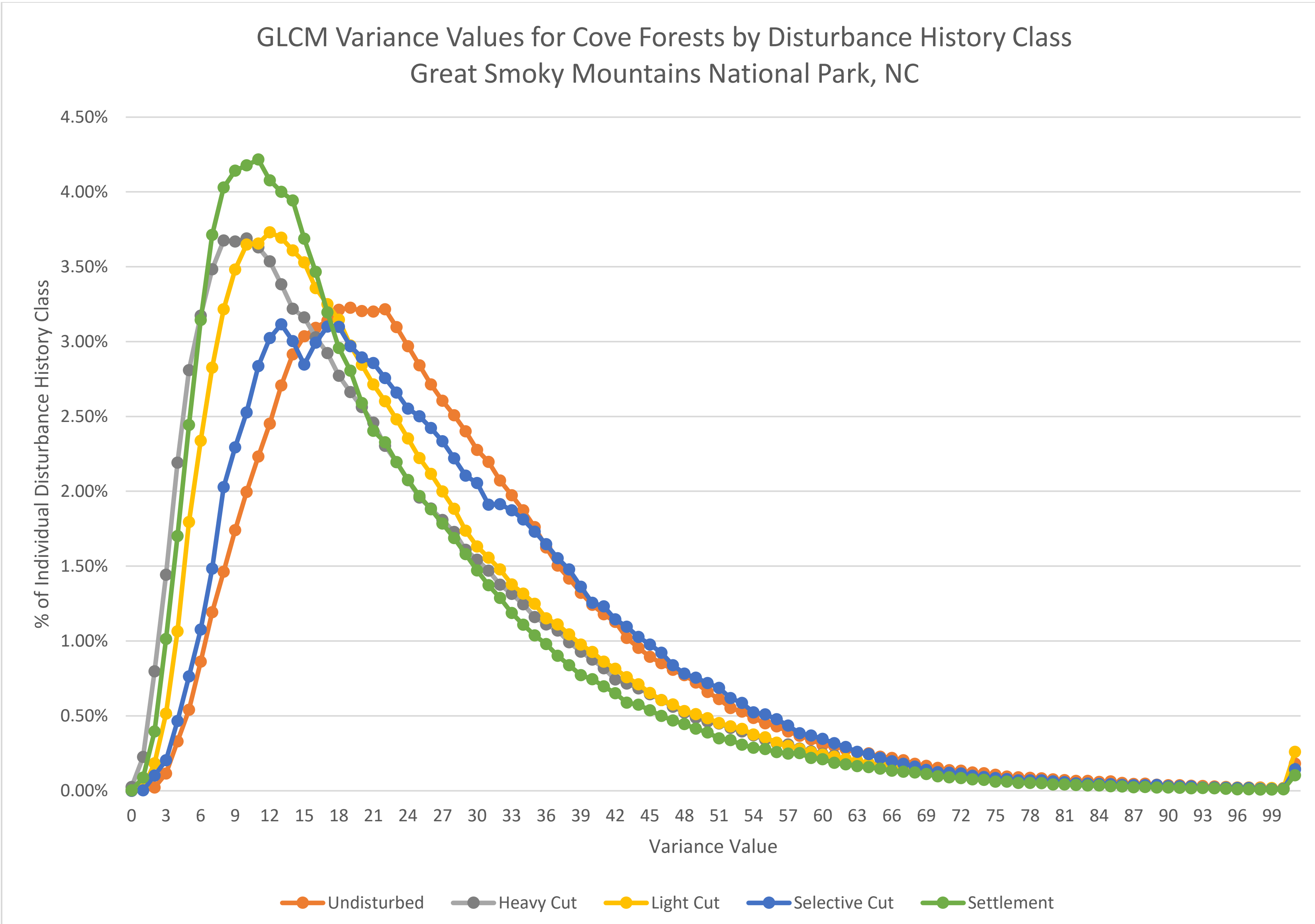


Figure 19: Variance Profile for GSMNP NC

B. Discussion of GSMNP Texture Profiles:

In Figures 4-19, the primary "undisturbed" cove forests stand out as having proportionally higher levels of entropy, contrast, dissimilarity, and variance, and proportionally lower levels of homogeneity and second moment than any of the second growth forest classes. The undisturbed forests tend to be most similar to those labeled "selective cut," which are the second least disturbed, and contain patches of primary forests intermixed with recovering second growth. Undisturbed forests were least similar to forests classified as "heavy cut" and "settlement," which experienced the most intensive logging and other land clearing prior to National Park Service Acquisition (Pyle 1988).

Despite the clear visibility of trends among these forest disturbance classes, each class varies widely throughout its dispersal. For example, the undisturbed cove forests in TN include a much higher proportion of entropy values above 2.5 than all other anthropogenic disturbance classes. However, they overlap considerably with the other classes below 2.5 entropy. There is no cutoff point at which any texture measure is clearly representing an old growth or second growth forest. However, these charts do imply that the texture metrics reflect structural patterns that are more common in primary forests than second growth forests, at least throughout GSMNP.

C. Statistical Analysis:

Values derived from the random points described in Section 4-C were used to test for statistical differences between texture values in old growth/primary cove forests, second growth cove forests, and non-cove-forest land cover. The term old growth/primary forest is used here because all points in GSMNP were located within primary forests while all points in Pisgah and Nantahala National Forests were located using the Messick surveys, which take old growth forest structure into account. Contrast, entropy, dissimilarity, homogeneity, mean height, second moment, and GLCM variance were compared for the three sets of points (GSMNP TN, GSMNP NC, and Nantahala/Pisgah National Forests).

Because the data were generally not normally distributed, nonparametric Kruskal-Wallis tests were used for all three sets of random points. Tables 1-3 and Figures 20-27 summarize the results of the statistical comparisons. For all tables, the A-C values in the first three rows correspond to the order in which the forest/disturbance classifications (old growth/primary, second growth, non-cove forest) differed significantly:

- A: Significantly higher than B and C
- B: Significantly lower than A and higher than C
- C: Significantly lower than A and B

Figures 20-27 are included to help to visualize the statistical differences between the three forest classes. These figures show the mean values for each texture metric along with 95% confidence intervals.

	Contrast	Homogeneity	Entropy	Dissimilarity	Second Moment	Variance	Mean Height
Old Growth/ Primary Cove Forest	A	C	A	A	C	A	A
Second Growth Cove Forest	B	B	B	B	B	A	A
Non-Cove-Forest Land Cover	C	A	C	C	A	B	B
DF Between	2	2	2	2	2	2	2
DF Within	1497	1497	1497	1497	1497	1497	1497
ChiSquare	294.46	251.41	211.71	272.1	150.98	175.53	417.63
P value	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
r2	0.085	0.159	0.142	0.138	0.093	0.076	0.276

Table 1: Kruskal-Wallis Results for GSMNP Tennessee Points

	Contrast	Homogeneity	Entropy	Dissimilarity	Second Moment	Variance	Mean Height
Old Growth/ Primary Cove Forest	A	C	A	A	C	A	B
Second Growth Cove Forest	B	B	B	B	B	B	A
Non-Cove-Forest Land Cover	C	A	C	C	A	C	C
DF Between	2	2	2	2	2	2	2
DF Within	1497	1497	1497	1497	1497	1497	1497
ChiSquare	250.823	250.715	264.495	253.717	241.458	249.917	464.018
P value	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
r2	0.126	0.164	0.167	0.161	0.045	0.132	0.305

Table 2: Kruskal-Wallis Results for GSMNP NC Model

	Contrast	Homogeneity	Entropy	Dissimilarity	Second Moment	Variance	Mean Height
Old Growth/ Primary Cove Forest	B	B	A	B	B	B	A
Second Growth Cove Forest	A	B	B	A	A	A	B
Non-Cove-Forest Land Cover	B	A	C	B	A	C	C
DF Between	2	2	2	2	2	2	2
DF Within	2,097	2,097	2,097	2,097	2,097	2,097	2,097
ChiSquare	84.138	41.73	68.288	70.478	74.978	88.172	708.404
P value	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001	0.0001
r2	0.044	0.027	0.047	0.035	0.038	0.034	0.305

Table 3: Kruskal-Wallis Results for Pisgah/Nantahala Model

D. Discussion of Statistical Comparisons:
Entropy was significantly highest and Second Moment consistently lowest in old growth/primary forests across all three comparisons. In GSMNP, where the LiDAR data was processed more consistently, old growth/primary forests showed significantly higher levels of contrast, entropy, dissimilarity, and variance, and significantly lower levels of homogeneity and second moment than second growth cove forests or non-cove forest communities. This distinction between the GSMNP data and the Pisgah/Nantahala data is important for three reasons:

1. Texture calculations generated inconsistent values across Pisgah/Nantahala National Forests because some CHMs were processed with a median height filter while others were not
2. The raw LiDAR data density varies more throughout Pisgah/Nantahala than it does in GSMNP (especially TN)
3. Canopy height in the areas sampled as "second growth cove forest" and "non-cove-forest land cover" is more variable in Pisgah/Nantahala than in GSMNP

These three factors may be altering whether the dissimilarity, variance, contrast and homogeneity values stand out in old growth forests. The absence of a median filter in particular causes generally higher height values which affect the GLCM's ability to reflect variations in canopy height. The LiDAR data density may also affect the texture metrics because higher height values are more commonly detected in areas of denser data collection. GLCM measures tend to capture trends rather than specific values. Any of these trends may be affected by inconsistent data. The GSMNP comparisons support the theory that the canopy in old growth cove forests is more variable, random, and contrasting, and less homogeneous than in second growth cove forests. However, low r^2 values (~0.04-0.30) indicate that these differences do not account for much of the variability throughout the datasets.

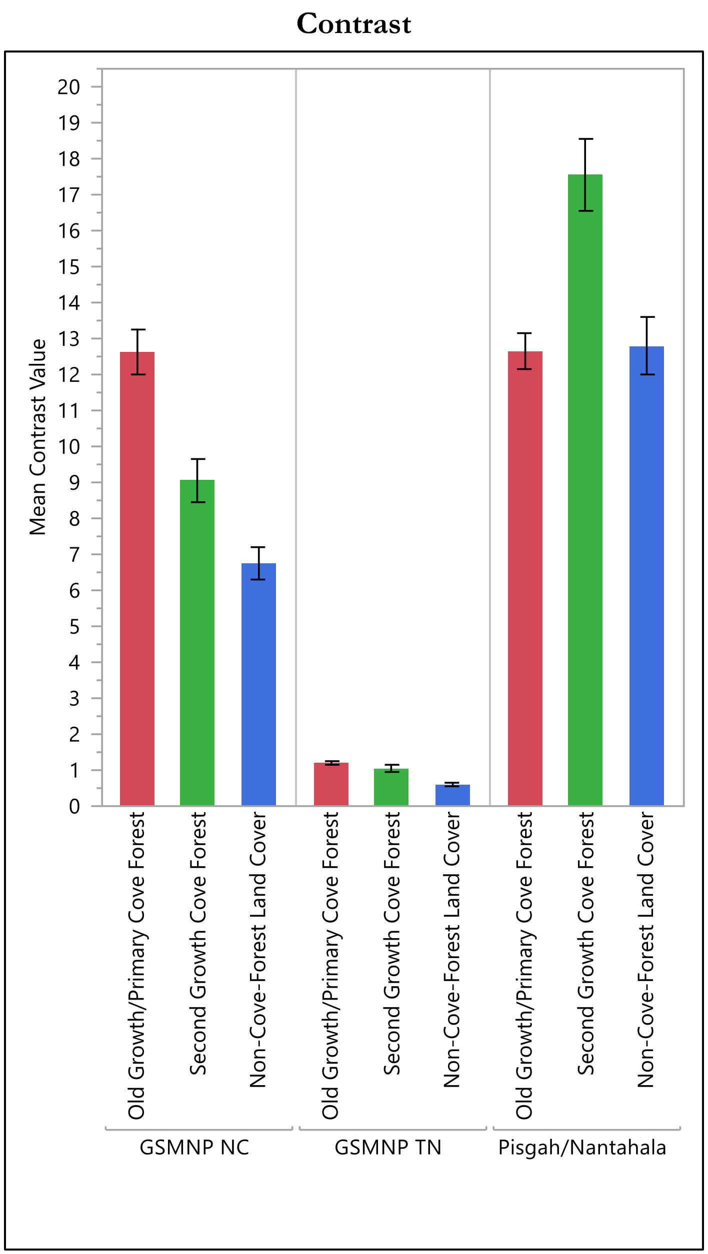


Figure 20: Means Comparisons for Contrast Values

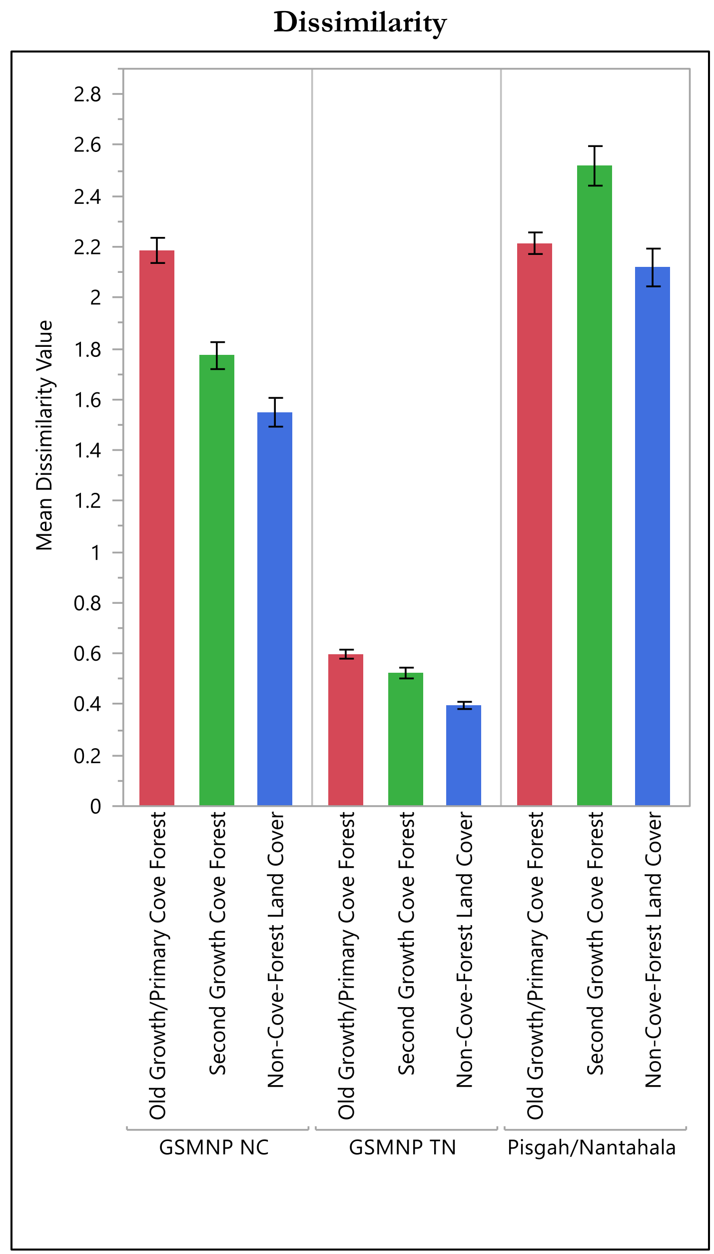


Figure 21: Means Comparisons for Dissimilarity Values

*Each error bar is constructed using a 95% confidence interval of the mean

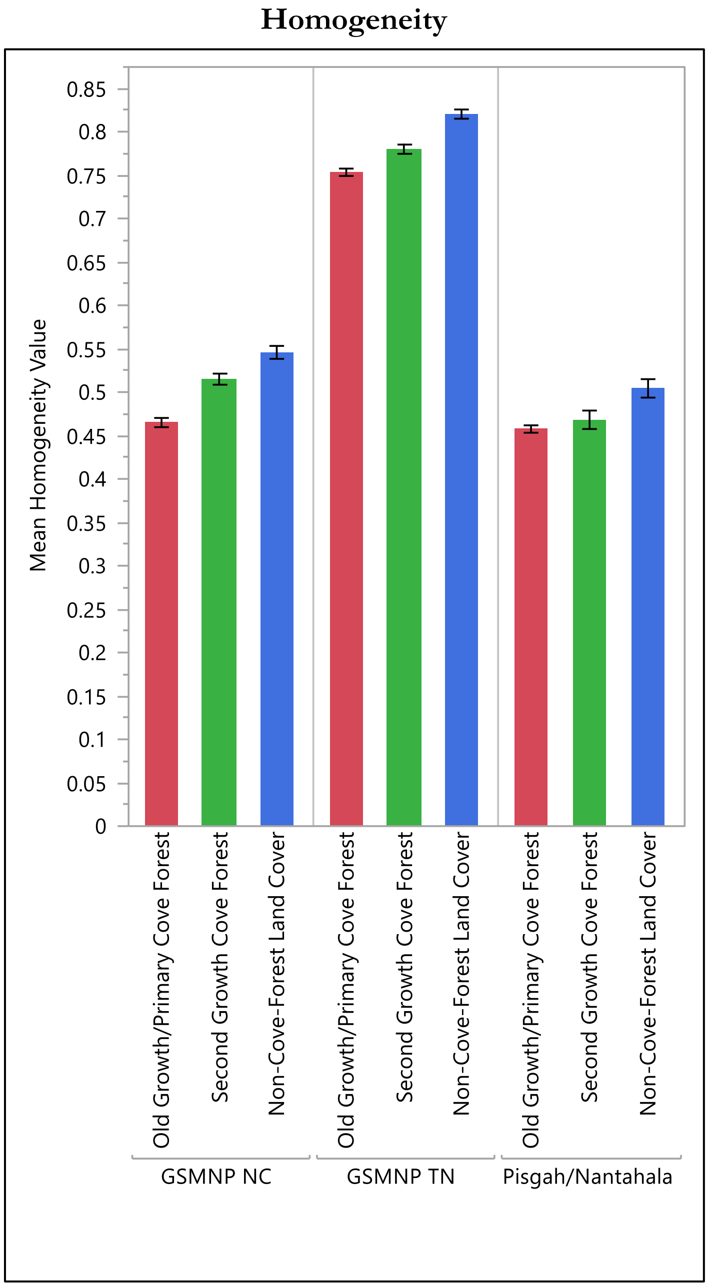


Figure 22: Means Comparisons for Homogeneity Values

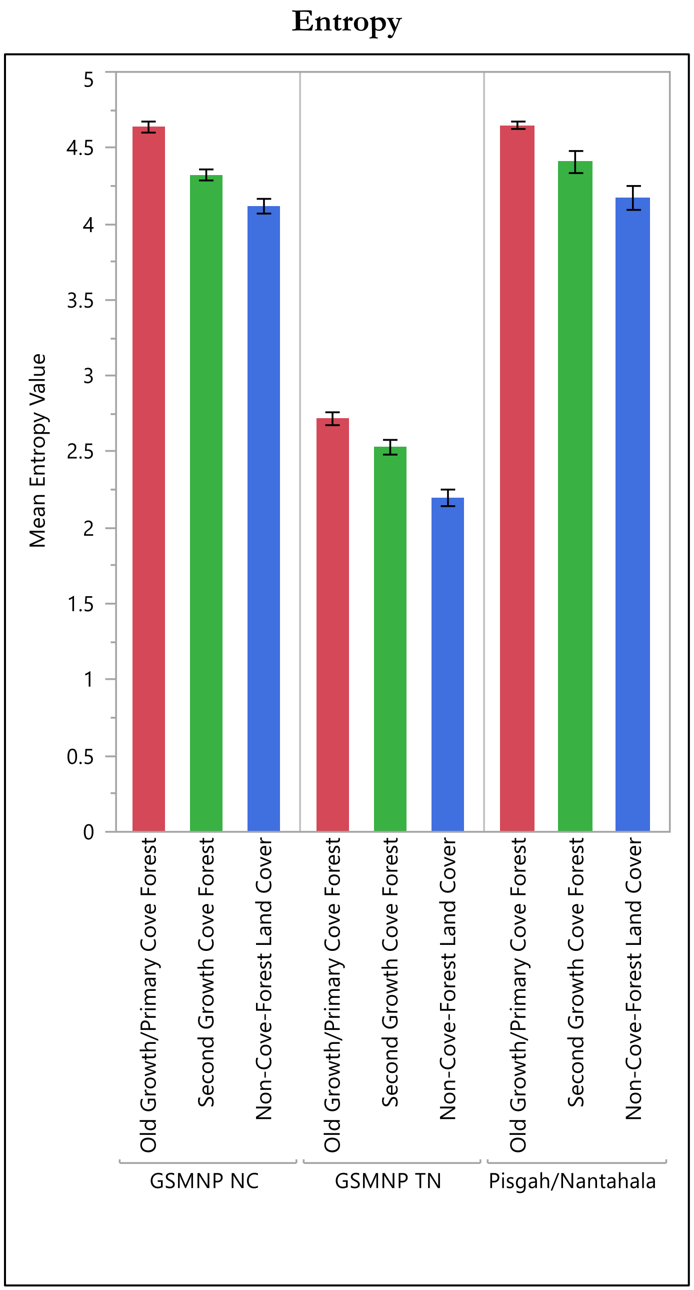


Figure 23: Means Comparisons for Entropy Values

*Each error bar is constructed using a 95% confidence interval of the mean

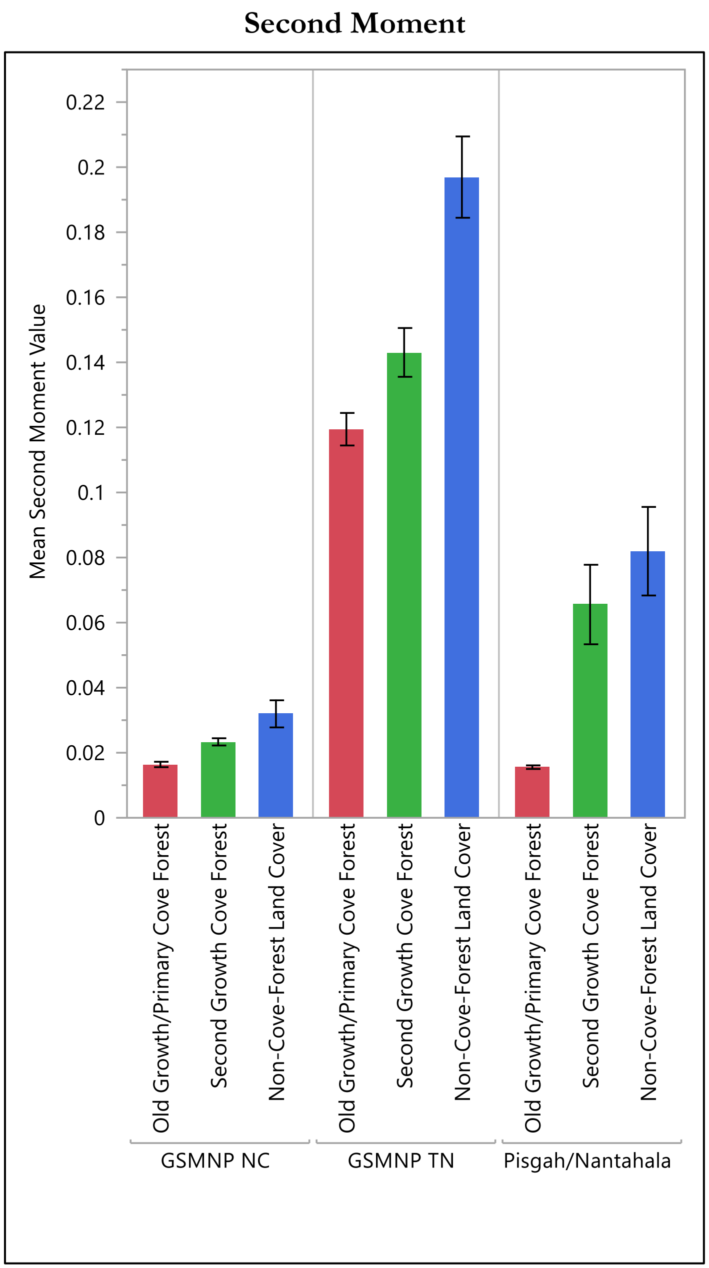


Figure 24: Means Comparisons for Second Moment Values

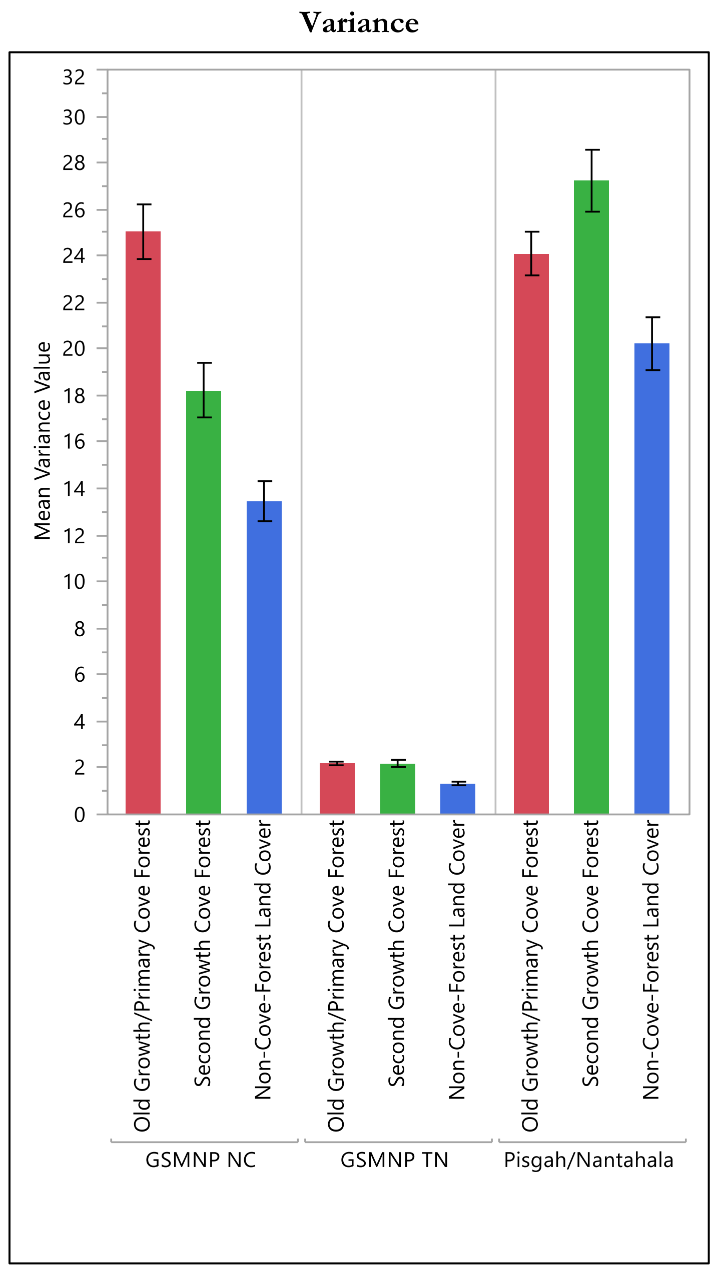


Figure 25: Means Comparisons for Variance Values

*Each error bar is constructed using a 95% confidence interval of the mean

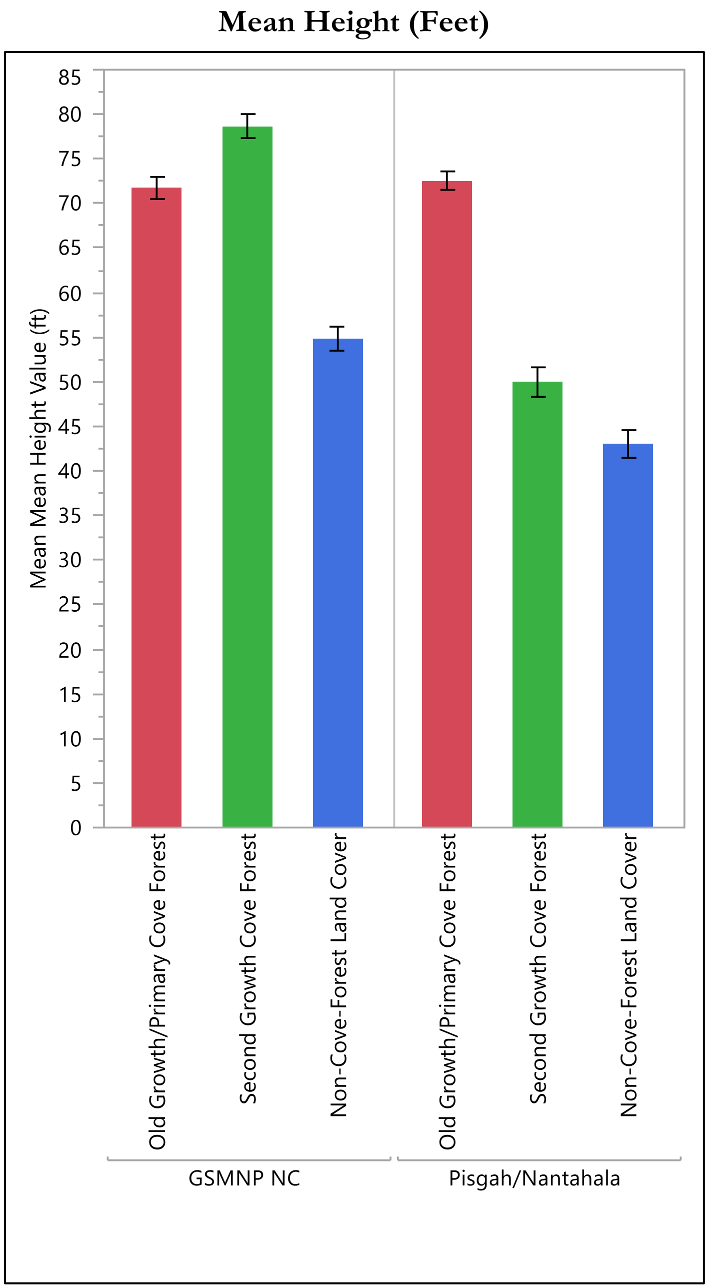


Figure 26: Means Comparisons for Mean Height Values (Feet)

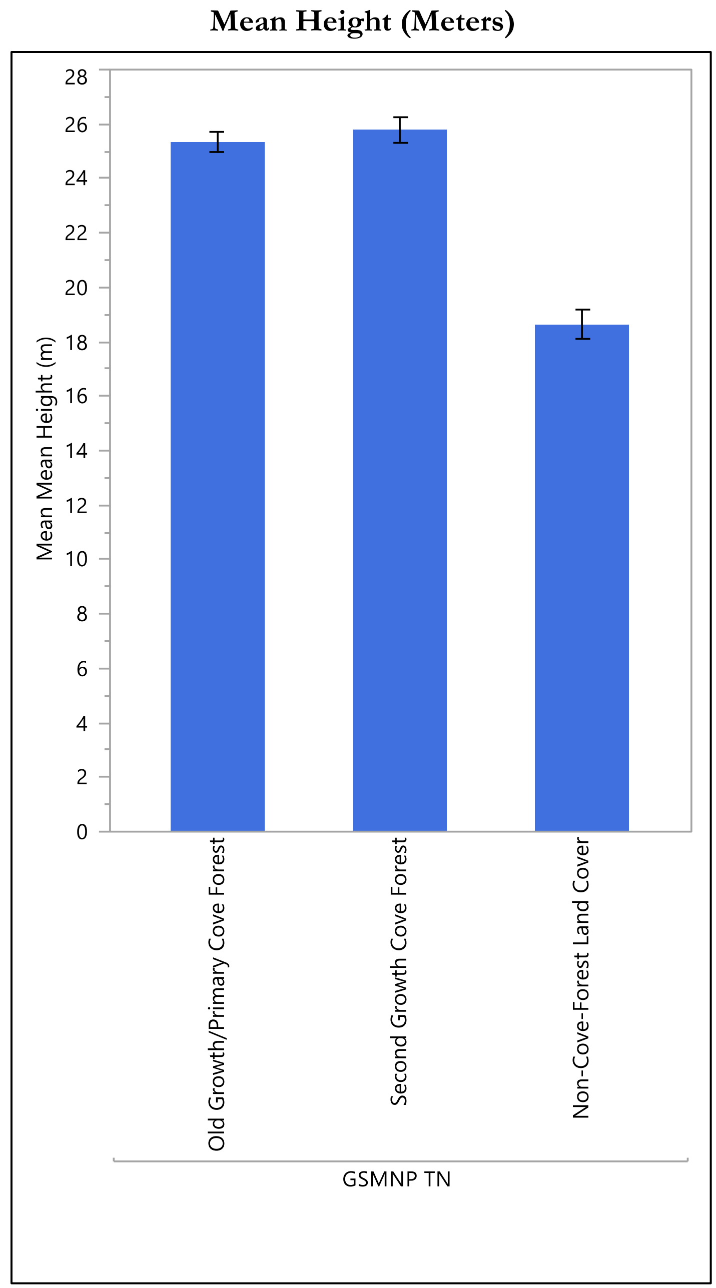


Figure 27: Means Comparisons for Mean Height Values (Meters)

*Each error bar is constructed using a 95% confidence interval of the mean

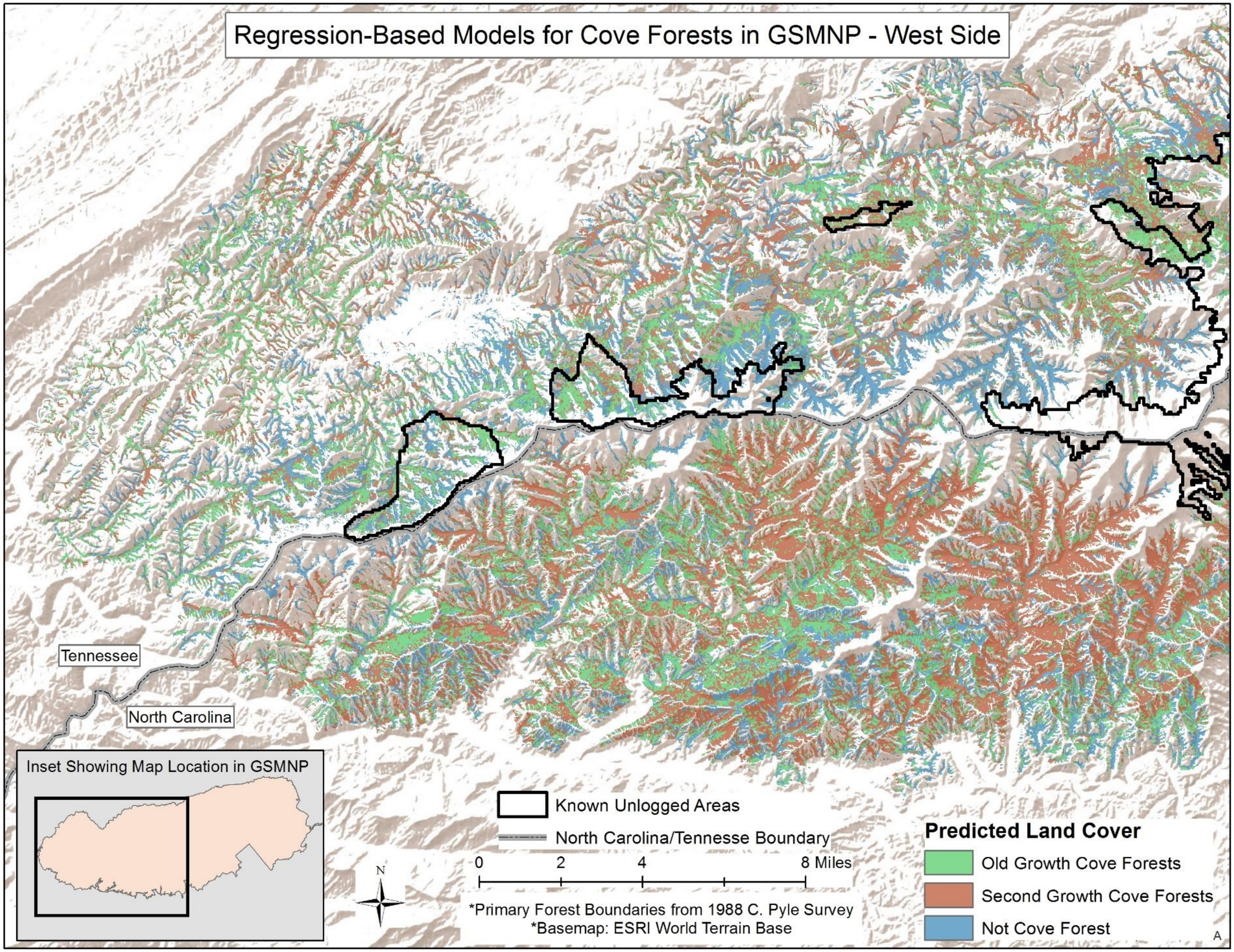
E. Regression Modeling

The final component of this project uses three multiple nominal logistic regression models to predict occurrences of old growth cove forests throughout the study area. These models classify land cover based on the three land cover classes used for the Kruskall Wallis tests (old growth cove forest, second growth cove forest, non-cove-forest land cover). The model parameters were created using texture and mean height values extracted to the 5,100 points which were also used for the statistical comparisons (Section 5-C). Section 4-E contains a more in-depth description of the methods used for creating the regression models including their equations. The regression model results are summarized in Table 4:

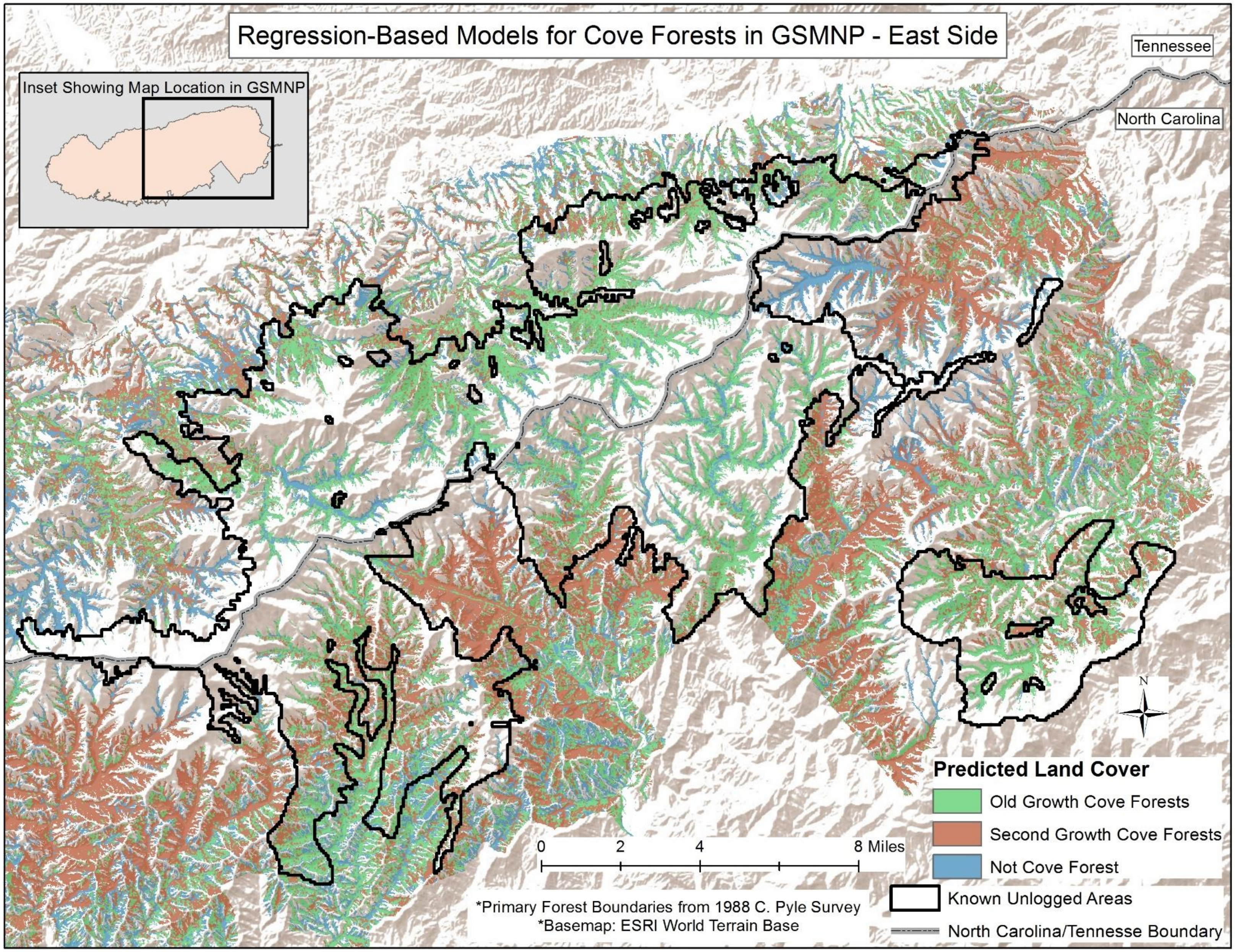
Model	ChiSquare	Prob>ChiSquare	Generalized Rsquare	Entropy Rsquare	RMSE	Misclass Rate
GSMNP TN	762.76	<0.0001	0.4484	0.2314	0.5523	0.3787
GSMNP NC	804.45	<0.0001	0.4670	0.2441	0.5457	0.3880
Pisgah/ Nantahala	1229.71	<0.0001	0.4986	0.2665	0.5389	0.3919

Table 4: Regression Model Results

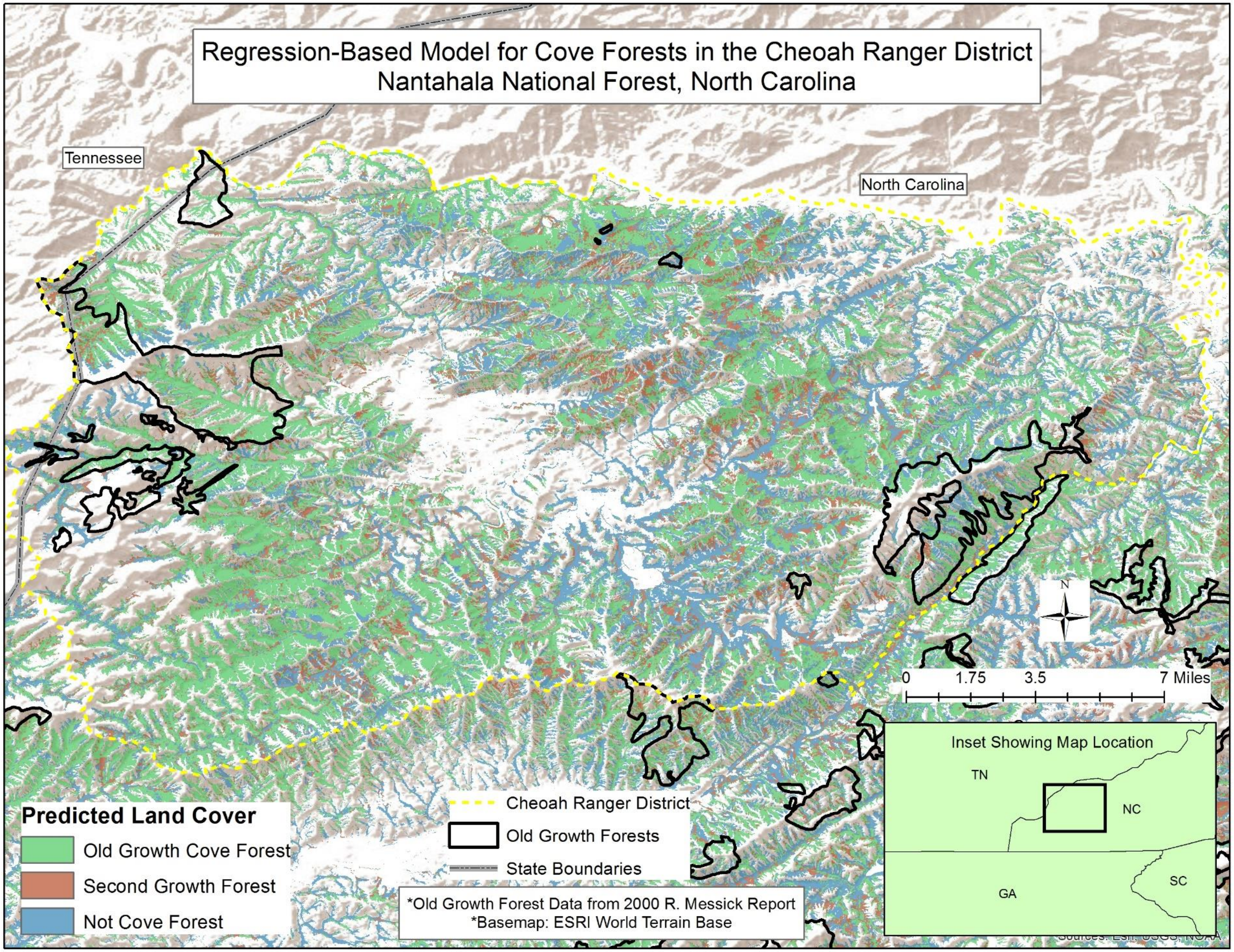
This process generated three raster layers with cell values corresponding to whichever land cover class (old growth cove forest, second growth cove forest, and non-cove-forest) received the highest probability from its respective model. Maps 31-37 show the raster layers within areas mapped as cove forest by both the ecozone and GSMNP vegetation mapping data.



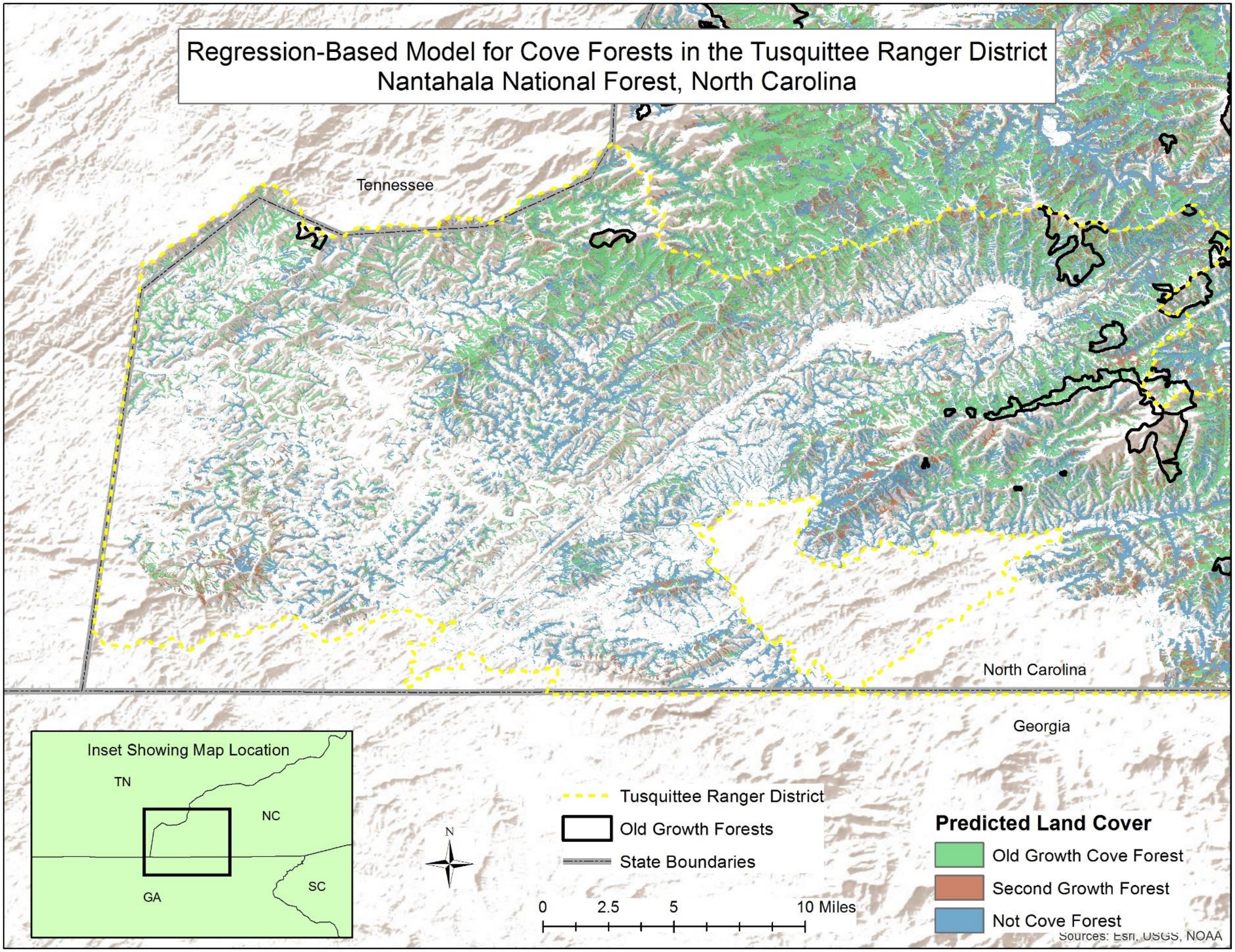
Map 31: Regression-Based Models for Cove Forests in GSMNP – West Side



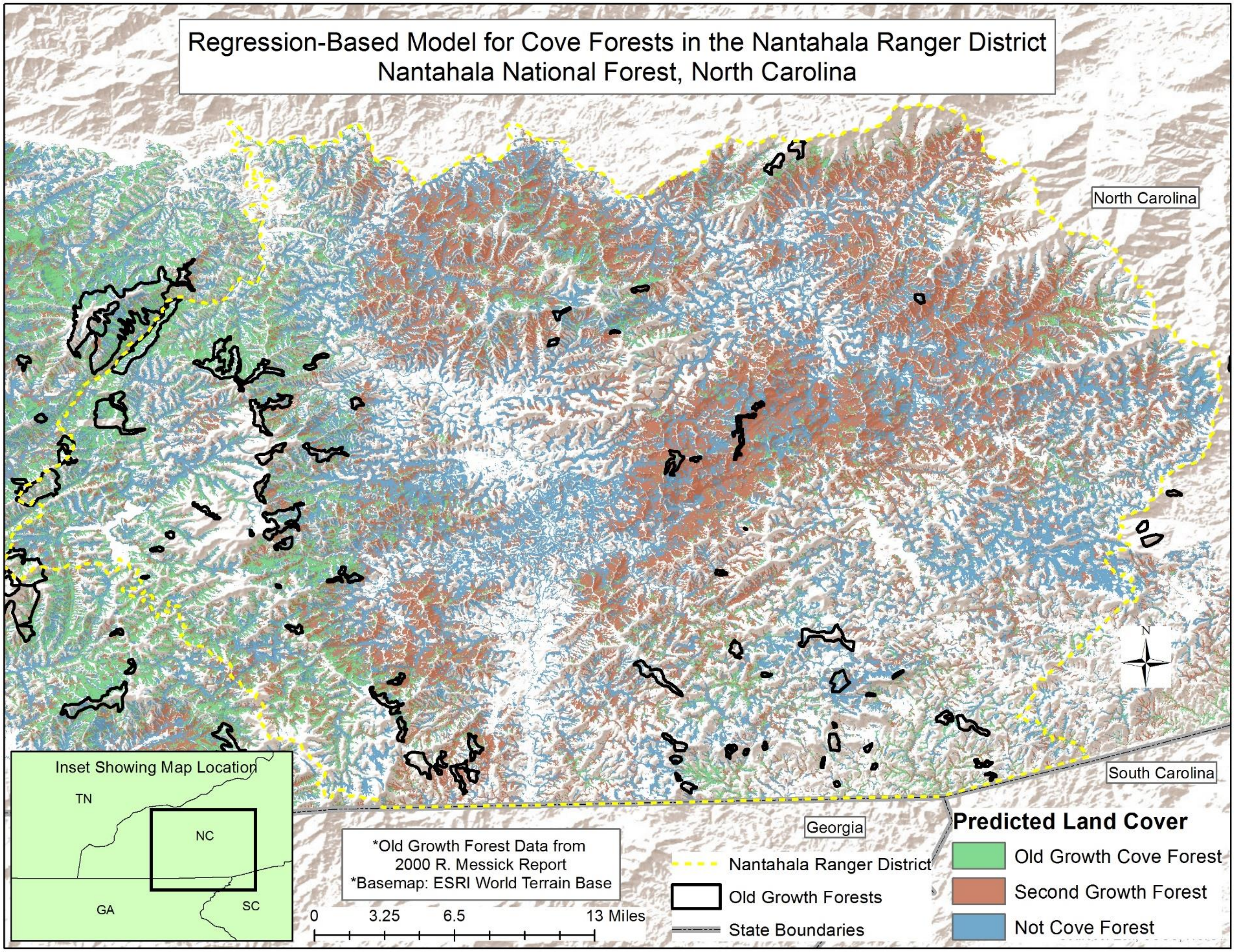
Map 32: Regression-Based Models for Cove Forests in GSMNP – East Side



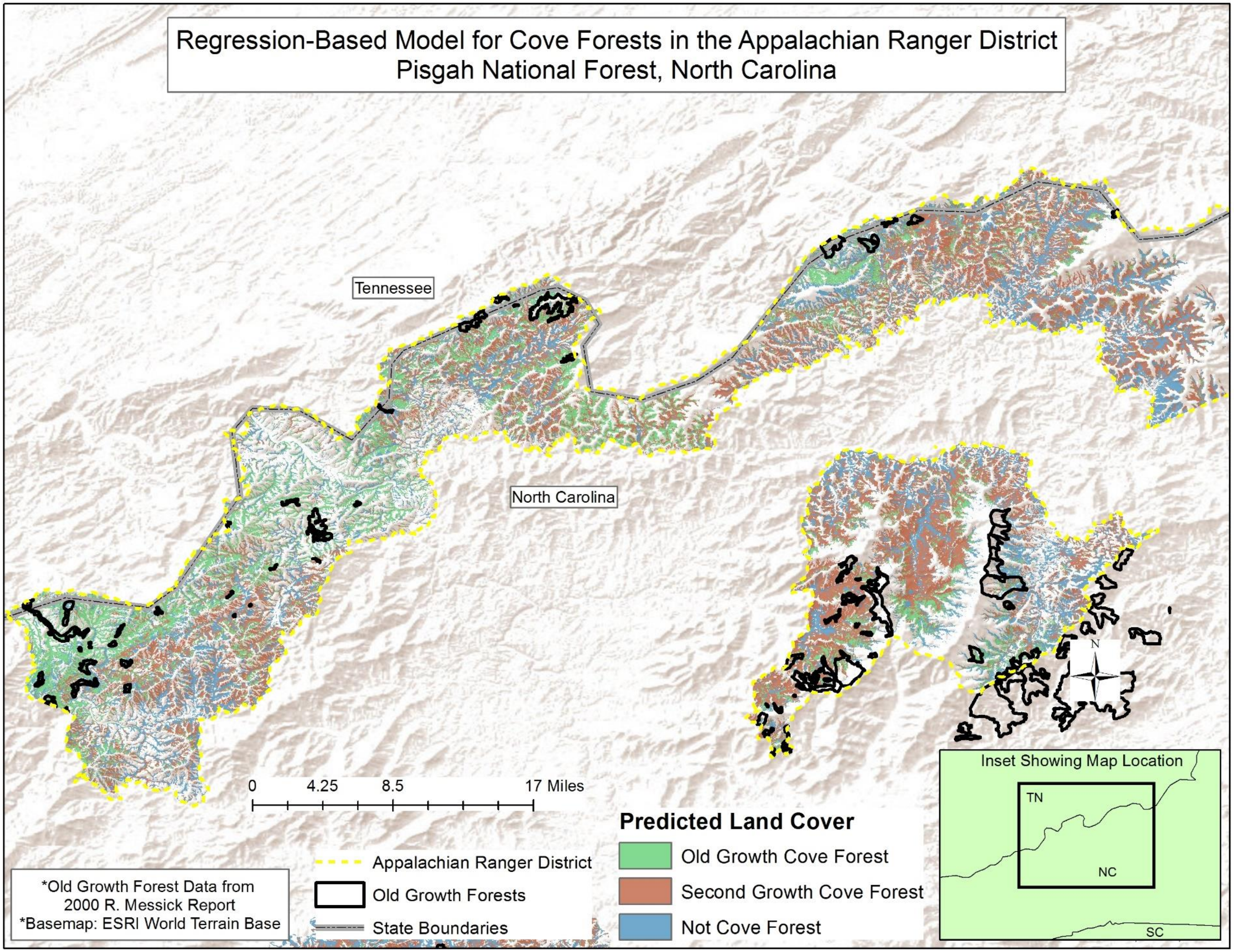
Map 33: Regression-Based Model for Cheoah Ranger District, Nantahala National Forest



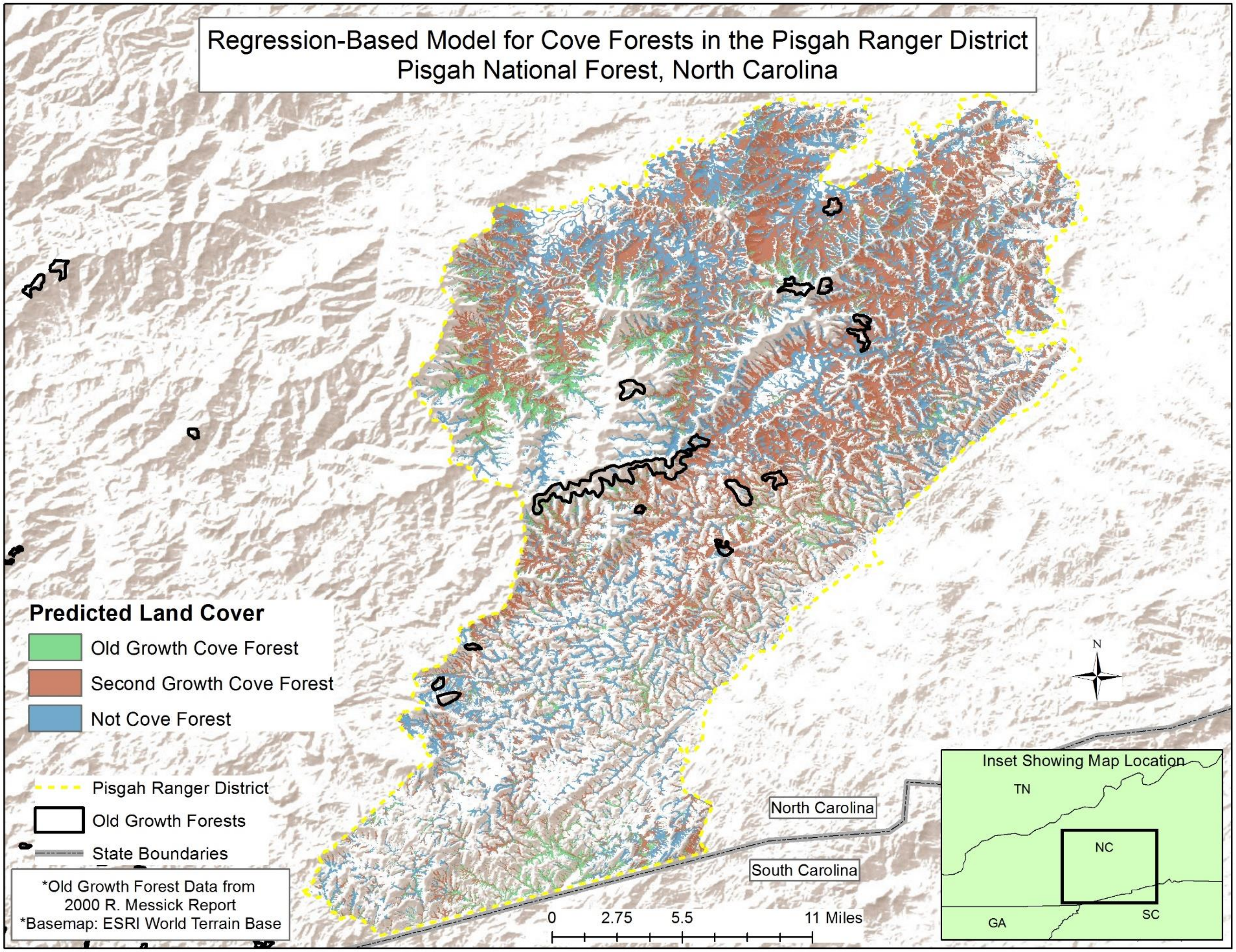
Map 34: Regression-Based Model for Tusquittee Ranger District, Nantahala National Forest



Map 35: Regression-Based Model for Nantahala Ranger District, Nantahala National Forest



Map 36: Regression-Based Model for Appalachian Ranger District, Pisgah National Forest



Map 37: Regression-Based Model for Pisgah Ranger District, Pisgah National Forest

F. Accuracy Assessment for Regression Models

The following confusion matrices represent accuracy assessments for each model's ability to predict the point values that were used as inputs for that model. These matrices are taken directly from the JMP Fit Model output.

GSMNP Tennessee Model:

The Tennessee model was able to distinguish between old growth cove forests, second growth cove forests, and non-cove-forest land cover with approximately 62% overall accuracy. Within each class, the model predicted old growth cove forests with 61% accuracy, non-cove forests with 73% accuracy, and second growth cove forests with 52% accuracy. The model tended to over-predict old growth cove forests and under-predict second growth cove forests: Old growth cove forest predictions were correct 58% of the time, non-cove predictions were correct 69% of the time, and second growth predictions were correct 58% of the time.

Confusion Matrix

Training

Actual	Predicted		
Response Class	Old Growth/Primary Cove Forest	Second Growth Cove Forest	Non-Cove-Forest Land Cover
Old Growth/Primary Cove Forest	307	120	73
Second Growth Cove Forest	154	259	87
Non-Cove-Forest Land Cover	66	68	366

Table 5: Confusion Matrix for GSMNP TN Model

GSMNP NC Model:

The North Carolina model was able to distinguish between old growth cove forests, second growth cove-forests, and non-cove-forest land cover with approximately 61% overall accuracy. Within each class, the model predicted old growth cove forests with 57% accuracy, non-cove forests with 66% accuracy, and second growth cove forests with 60% accuracy. Old growth cove forest predictions were correct 56% of the time, non-cove predictions were correct 67% of the time, and second growth predictions were correct 60% of the time.

Confusion Matrix

Training

Actual	Predicted		
Response Class	Old Growth/Primary Cove Forest	Second Growth Cove Forest	Non-Cove-Forest Land Cover
Old Growth/Primary Cove Forest	286	132	82
Second Growth Cove Forest	120	299	81
Non-Cove-Forest Land Cover	100	67	333

Table 6: Confusion Matrix for GSMNP NC Model

Pisgah/Nantahala Model:
The Pisgah/Nantahala Model was able to distinguish between old growth cove forests, second growth cove forests, and non-cove-forest land cover with approximately 61% overall accuracy. This model tended to over-predict old growth cove forests and under predict second growth cove forests. Within each class, the model predicted old growth cove forests with 82% accuracy, non-cove forests with 58% accuracy, and second growth cove forests with 42% accuracy. Old growth cove forest predictions were correct 80% of the time, non-cove predictions were correct 56% of the time, and second growth predictions were correct 53% of the time.

Confusion Matrix

Training

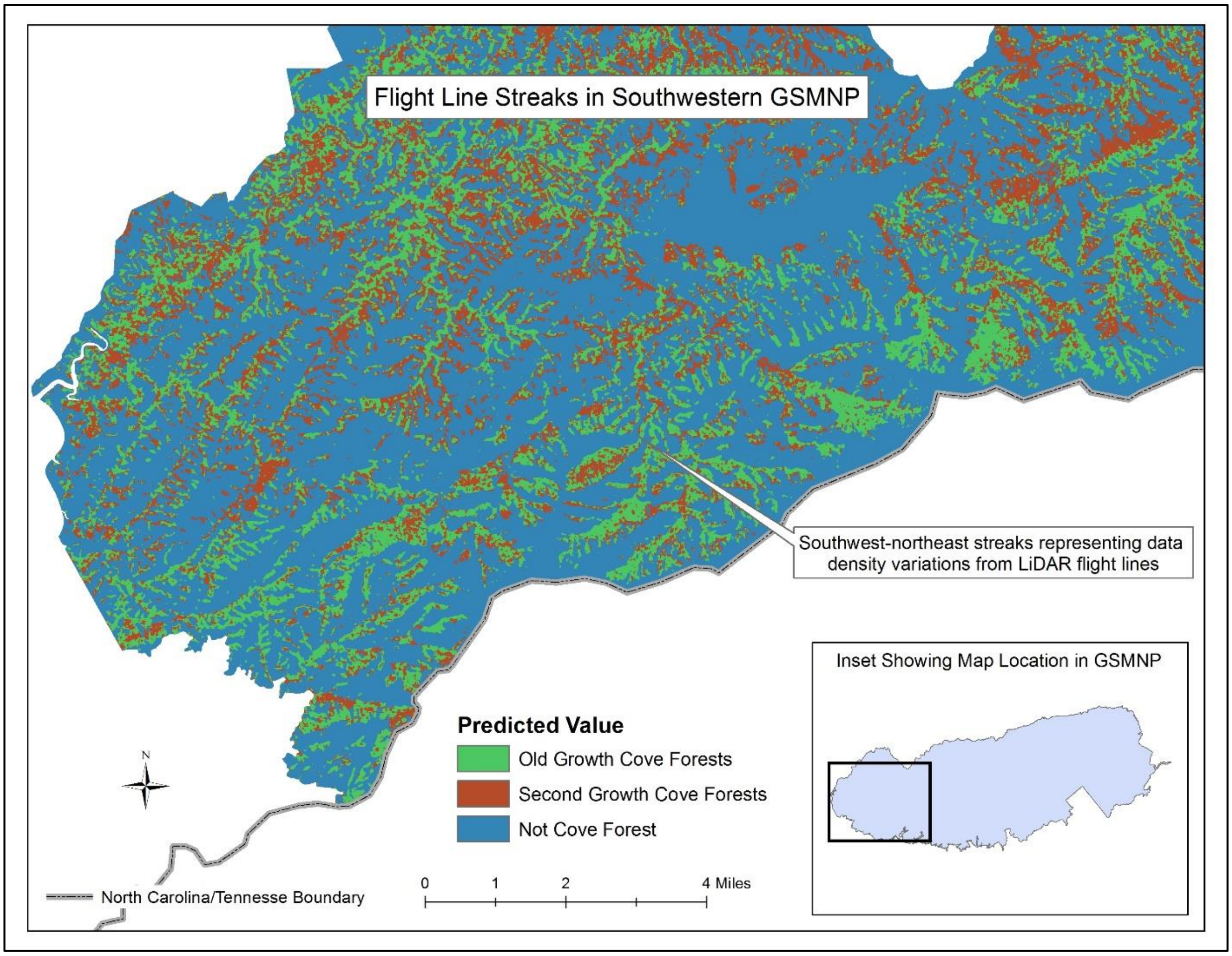
Actual Response Class	Predicted: Old Growth/Primary Cove Forest	Predicted: Second Growth Cove Forest	Predicted: Non-Cove-Forest Land Cover
Old Growth/Primary Cove Forest	575	60	65
Second Growth Cove Forest	152	295	253
Non-Cove-Forest Land Cover	91	202	407

Table 7: Confusion Matrix for Pisgah/Nantahala Model

H. Discussion for Regression Models
These models and resulting maps vary greatly in their ability to classify cove forests by their horizontal structure. Although the confusion matrices give each model approximately 60% overall accuracy with reasonable accuracy levels among the individual class predictions, close inspection of the resulting raster layers reveals a series of classification errors.

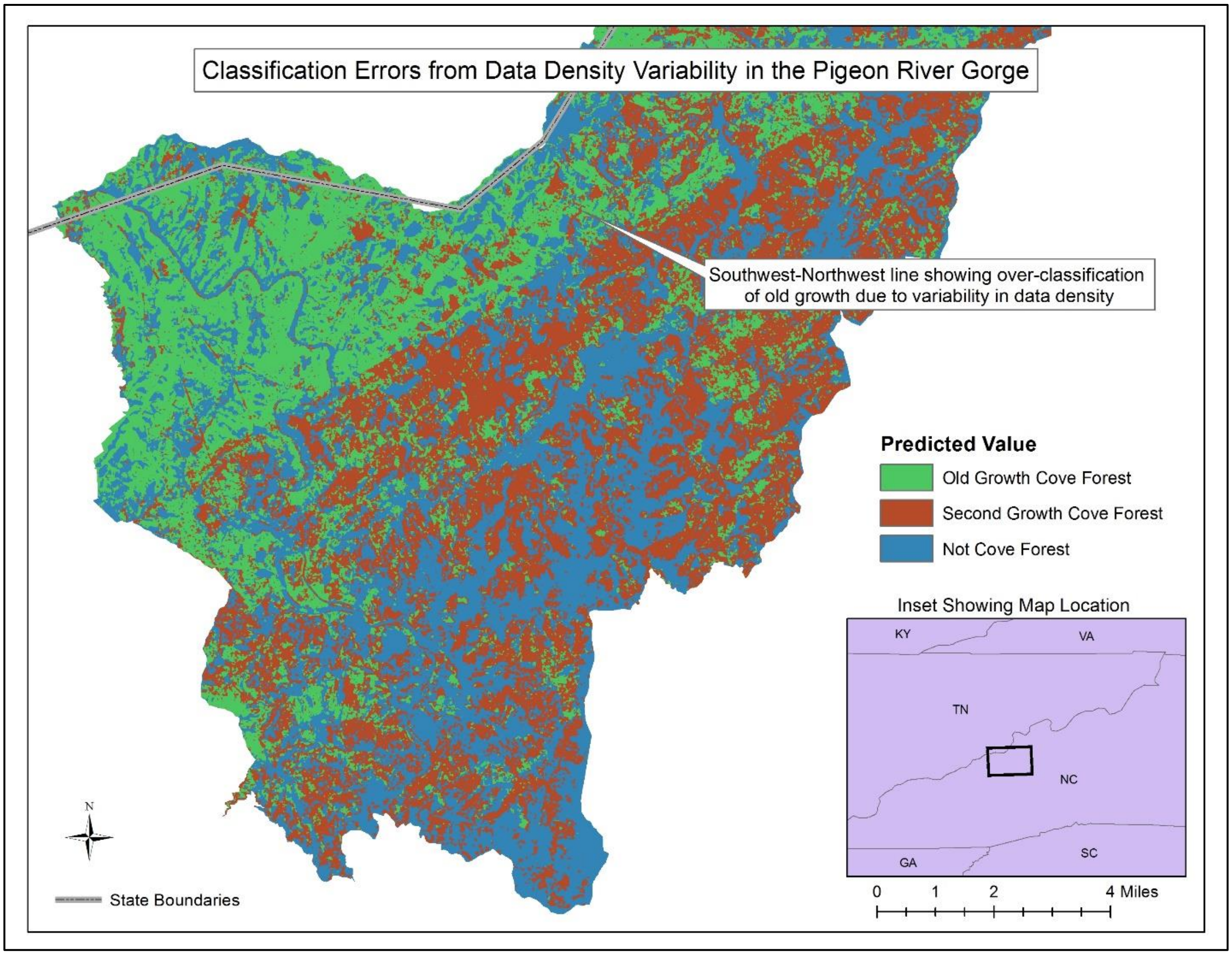
The raster layers visible in Maps 31-37 provide important insights into the accuracy of the three regression models. Upon first glance, the two GSMNP models appear to classify old growth forests fairly accurately. Areas known to be old growth forests, such as the Upper Ravens Fork Watershed include a high concentration of pixels classified correctly as old growth cove forest. Some areas like the Big Creek watershed which experienced intensive land clearing are correctly classified as second growth cove forest. Upon closer inspection, however, the maps begin to appear less accurate. In the higher elevation extents of GSMNP above 4500 feet (generally close to the NC/TN state line), extensive areas of old growth northern hardwood and transitional spruce fir forests are classified as old growth cove forests. These high-elevation cove forest classifications may be reflective of the broad definition given to cove forests by the ecological zone models.

Maps 38-40 on the following pages showcase several classification errors visible in the regression model raster layers:



Map 38: Flight Line Streaks in GSMNP TN Model

Map 38 shows a series of small classification errors on the western side of GSMNP. Here, a series of straight lines running southwest to northeast show a continuous streak of small patches classified as old growth cove forest. These lines reflect inconsistencies in the raw LiDAR point data used to create the CHMs. The SW/NE patterns most likely represent the flight lines of the planes that acquired the LiDAR data. Straight lines like these appear in small patches all over the raster layers. Overlaps in the data acquisition, and/or lighter-sampled areas between the flight lines may have caused sharp variations in both the pulse density (number of first return pulses/the area of the acquisition tile) and nominal pulse spacing (average spacing of LiDAR pulses) of the data.



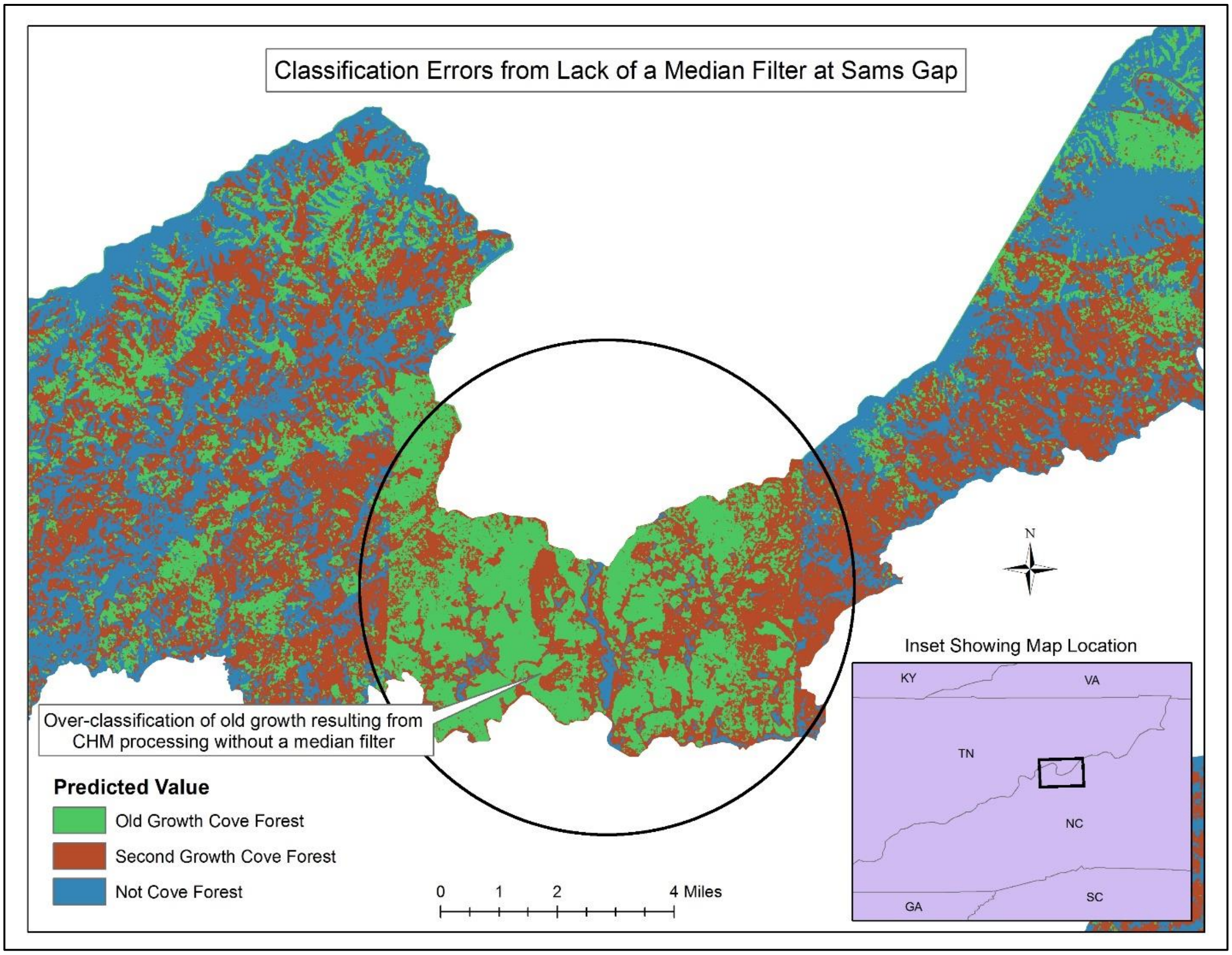
Map 39: Classification Errors in the Pigeon River Gorge

Map 39 shows an odd-shaped large area classified as old growth cove forest within the Pigeon River Gorge on the far-western end of the Appalachian Ranger District of Pisgah National Forest. This concentration of old growth is bounded on its southeastern side by a straight SW-NE line which separates it from another area which is mostly classified as second-growth cove forest. These classification differences likely resulted from inconsistencies in the data density with which the LiDAR was collected. To investigate this, I visited the USGS Earth Explorer website to check the metadata for the two adjacent areas visible in Map 39:

- The area classified as old growth was collected with a nominal pulse spacing of 7.87077 pulses/square meter, and a pulse density of 0.0161423 first returns/square meter
- The area classified as second growth was collected with a nominal pulse spacing of 5.2253 pulses/square meter, and a pulse density of 0.036625 first returns/square meter

In other words, denser raw LiDAR data caused the model to designate an inaccurately large area as old growth forest. Inconsistencies like these are apparent throughout all models and affect their usefulness and accuracy.

Inconsistent methodology for processing the LiDAR data into CHMs also affected the regression models' ability to classify old growth cove forests accurately. In most cases, a median filter was used to fill in "no data" cells throughout the CHMs. These "no data" values result from erroneous high or low points that are filtered out of the LiDAR during processing. In some instances, however, a median filter was not used, resulting in inaccurate old growth classifications. The most noticeable example of this error is visible in Map 40. In this image, inconsistencies in the data are visible in the form of a north-south trending line throughout the center of the National Forest which divides high classifications of second growth and old growth cove forests (inside the black circle).



Map 40: Classification Errors Surrounding Sams Gap

Section VI: Recommendations

This analysis represents a first attempt to explore the possibility of identifying old growth forests remotely using LiDAR. The results suggest that texture analysis of canopy height provides valuable insight into the horizontal structure of old growth cove forests. However, the methods used here are imperfect. Due to inconsistencies in the LiDAR acquisition and processing, and other data limitations, the regression models created here should not be interpreted as an accurate classification of old growth cove forests in the Southern Appalachians.

Despite its inaccuracies, the methodology used here could be improved upon to create a more accurate tool. GLCM texture analysis certainly captures some of the spatial characteristics of forest structure in old growth cove forests. The application of GLCM measures – particularly entropy – to the LiDAR CHMs captures aspects of spatial heterogeneity that descriptive statistics do not.

I make the following suggestions for any future efforts to repeat this study:

1. *Use raw LiDAR data with as consistent data density as possible*
    - Inconsistent data density in the LiDAR may have been the most limiting factor in this analysis. Better, more consistent LiDAR data could enable a much more accurate model. Alternatively, it may make sense to research LiDAR processing techniques which might account for the varying pulse densities and nominal pulse spacing inherent in the data.

2. Eliminate non-natural land cover from the sampling design
- Some of the random points used in this analysis lay in developed areas within the Nantahala National Forest Administrative Boundary. I recommend limiting random points throughout the study area to the actual national forest boundaries rather than administrative boundaries so that developed areas are not considered by the model. Alternatively, developed areas could be considered and sampled as their own land cover class.

3. Incorporate multiple forest communities into the analysis
- The study design used here probably influenced the erroneous classifications in the regression model raster layers. Treating old growth cove forest, second growth cove forest, and non-cove forest land cover as discrete classes may have been part of the problem. In particular, lumping all land cover types besides cove forest into a single "non-cove" category may have thrown off the model parameters because everything from mature oak-hickory forests to grassy balds to parking lots was included in the non-cove category.
- Multiple forest communities could be included in a similar model to create a more detailed and useful product. This might be accomplished by simplifying the ecological zone models into a smaller number of classes and then sampling throughout those classes. The 11 forest ecosystems evaluated in Josh Kelly's eCAP analysis of the Pisgah/Nantahala National Forests (Kelly 2013) could be useful for this purpose.

4. *Conduct field work to validate representative old growth forest points*
    - Field-truthing points in old growth forests would eliminate some of the uncertainties that may have resulted from extracting point values from primary forests. Using field-validated

areas of old growth forest would ensure that any model inputs that are intended to represent old growth forests represent true old growth structure.

5. *Incorporate additional ecological inputs beyond canopy height*
    - Canopy height is only one factor driving ecosystem structure in old growth forests. Incorporating additional LiDAR-derived metrics like understory density, canopy cover, and digital elevation models may produce a more accurate result. Non-LiDAR metrics like precipitation might also be useful for this purpose.

6. *Use canopy height models with a consistent cell size:*
    - The different cell sizes present in CHMs throughout the study area posed a major obstacle. Ideally, one set of texture layers could be generated for the entire Southern Blue Ridge Ecoregion. This would make it possible to incorporate data from a wider variety of old growth forests into a single, consistent final model specific to a large ecological landscape.

7. *Experiment with larger CHM cell sizes*
    - The small cell sizes used in this study (3M, 9ft, 10ft) represent a highly detailed model of the forest canopy, but may have added unnecessary variability to the texture layers. A larger cell size (perhaps 20-ft x 20-ft) might limit the effect that raw data inconsistencies have on the CHMs and texture metrics. A larger cell size might also miss some of the small canopy gaps throughout the second growth forests while detecting the larger canopy gaps in the old growth forests. This could make it easier to detect differences in the old growth and second growth forests in the CHMs. Using a larger cell size would certainly speed up processing time!

I hope that these model layers, methods, and suggestions will give Mountaintrue new ideas for expanding their LiDAR-based conservation work. I wish them luck in their efforts to protect and restore Southern Appalachian old growth!

Acknowledgements

I would like to thank the following people for their guidance and support throughout my research:

My Committee: Deane Wang, Scott Merrill, and Jen Pontius
For steering me in the right direction, for their constructive guidance and criticism, and for their good humor.

Josh Kelly
For his creativity in envisioning this research, and for offering collaborative input throughout every step of the project.

Dr. Mryka Hall-Beyer, University of Calgary
For orienting me to the finer details of GLCM texture analysis, and for fielding some very long email questions.

Dr. Marguerite Madden, University of Georgia Center for Geospatial Research
For her valuable input and research suggestions.

Kaya Lindquist
For putting up with me throughout the project and keeping my spirits high.

The Mountaintrue staff
For your excellent company throughout my project. It was an honor to be a part of such an effective organization.

My Acer R3 series laptop
For somehow processing so many texture layers against all odds!

References

Barden, L. 1983. Size, Age and Growth Rate in Trees in Canopy Gaps of a Cove Hardwood Forest in the Southern Appalachians. Castanea 48 (1): 19-23.

Busing, R. 1989. A Half Century of Change in a Great Smoky Mountains Cove Forest. Bulletin of the Torrey Botanical Club 116 (3): 283-288.

Busing, R. 1995. Disturbance and the Population Dynamics of Liriodendron tulipifera: Simulations with a Spatial Model of Forest Succession. Journal of Ecology 83 (1): 45-53.

Busing, R. 1998. Structure and Dynamics of Cove Forests in the Great Smoky Mountains. Castanea 63(3): 361-371.

Busing, R. 1993. Three Decades of Change at Albright Grove, Tennessee. Castanea 58(4): 231-242.

Clebsch, E. and R. Busing. 1989. Secondary Succession, Gap Dynamics, and Community Structure in a Southern Appalachian Cove Forest. Ecology 73 (3): 728-735.

Flatley, W., C. Lafon, H. Grissino-Mayer, and L. Laforest. 2015. Changing fire regimes and old growth forest succession along a topographic gradient in the Great Smoky Mountains. Forest Ecology and Management 350 (2015): 96-106.

Guyon, Lyle J.; Rolfe, Gary L.; Edgington, John M.; Mendoza, Guillermo A. 2003. A comparative analysis of the diversity of woody vegetation in old-growth and secondary southern Appalachian cove forests. In: Van Sambeek, J. W.; Dawson, Jeffery O.; Ponder Jr., Felix; Loewenstein, Edward F.; Fralish, James S., eds. Proceedings of the 13th Central Hardwood Forest Conference; Gen. Tech. Rep. NC-234. St. Paul, MN: U.S. Department of Agriculture, Forest Service, North Central Research Station: 75-87

Bonnet, S., R. Gaulton, F. Lehaire, and P. Lejeune. 2015. Canopy gap mapping from airborne laser scanning: an assessment of the positional and geometrical accuracy. Remote Sensing 7 (9): 11267-11294.

Clausi, D. 2002. An analysis of co-occurrence texture statistics as a function of grey scale quantization. Canadian Journal of Remote Sensing 28(1): 45-62.

Comer, P., D. Faber-Langendoen, R. Evans, S. Gawler, C. Josse, G. Kittel, S. Menard, M. Pyne, M. Reid, K. Schulz, K. Snow, and J. Teague. 2003. Ecological Systems of the United States: A Working Classification of US Terrestrial Systems. Natureserve, Arlington, VA.

Davis, M. 2003. Old Growth in the East: A Survey (Revised Edition). Appalachia-Science in the Public Interest, Mt. Vernon, KY.

Dupuy S., G. Laine, J. Tassin, and J.M. Sarrailh. 2013. Characterizations of the horizontal structure of the tropical forest canopy using object-based LiDAR and multispectral image analysis. International Journal of Applied Earth Observation and Geoinformation 25 (2013): 76-86.

Franklin, J., and R. Van Pelt. 2004. Spatial aspects of structural complexity. Journal of Forestry 102 (3): 22-28.

Frelich, L., M. Cornett, and M. White. 2005. Controls and reference conditions in forestry: the role of old growth and retrospective studies. Journal of Forestry 103(7): 339-344.

Hall-Beyer, M. 2007. Grey Level Co-Occurrence Matrix Tutorial. Accessed: http://www.fp.ucalgary.ca/mhallbey/tutorial.htm

Haney. J.C., and C. Schaadt. 1996. Functional roles of old growth in promoting forest bird diversity. In *Eastern Old Growth Forests*, ed. M.M. Davis, 76-88. Washington, DC: Island Press.

Jane, R., R. Kasturi, and B. Schunck. 1995. Machine Vision. Published by McGraw-Hill, Inc.

Kane, V., R. McGaughey, J. Bakker, R. Gersonde, J. Lutz, and J. Franklin. 2010a. Comparisons between field and LiDAR-based measures of stand structural complexity. Canadian Journal of Forest Research 40 (4): 761-773.

Kane, V., R. McGaughey, J. Bakker, R. Gersonde, J. Lutz, and J. Franklin. 2010b. Examining conifer canopy structural complexity across forest ages and elevations with LiDAR data. Canadian Journal of Forest Research 40: 774-787.

Kane, V., R. McGaughey, J. Bakker, R. Gersonde, J. Lutz, and J. Franklin. 2011. Patch dynamics and the development of structural and spatial heterogeneity in Pacific Northwest forests. Canadian Journal of Forest Research 41: 2276-2291.

Keane, R., P. Hessburg, P. Landres, F. Swanson. 2009. The use of historical range of variability (HRV) in landscape management. Forest Ecology and Management 258: 1025-1037.

Kelly, J. 2009. Using LiDAR to analyze vegetation structure and ecological departure in the Upper Warwoman Watershed, Georgia. Western North Carolina Alliance.

Kelly, J. 2013. An assessment of the ecosystems of Nantahala-Pisgah National Forest & surrounding lands: a synthesis of the eCAP Methodology and LiDAR Vegetation Analysis. Report for Western North Carolina Alliance.

Kelly, J., J. Hushaw, P. Jost, W. Blozan, H. Irwin, and J. Riddle. 2010. Using LiDAR to locate exceptionally tall trees in Western North Carolina. Bulletin of the Eastern Native Tree Society 5 (1&2): 16-21.

Landfire. Biophysical Settings Model Descriptions. URL: http://www.landfire.gov/NationalProductDescriptions20.php - Accessed 16 May 2016.

Landres, P., P. Morgan, and F. Swanson. 1999. Overview of the use of natural range of variability concepts in managing ecological systems. Ecological Applications. Vol 9 (4): 1179:1188.

Lefsky, M., D. Harding, W.B. Cohen, G. Parker, and H.H. Shugart. 2002. Surface LiDAR remote sensing of basal area and biomass in the deciduous forests of Eastern Maryland, USA. Remote Sensing of the Environment 67 (2002): 83-98.

Lefsky, M., A. Hudak, W. Cohen, and S.A. Acker. 2005. Patterns of coviariance between forest stand and canopy structure in the Pacific Northwest. Remote Sensing of Environment 95 (2005): 517-531.

Lim, K., P. Treitz, M. Wulder, B. St-Onge, and M. Flood. 2003. LiDAR remote sensing of forest structure. Progress in Physical Geography 27 (1): 88-106.

Lorimer, C. 1980. Age structure and disturbance history of a Southern Appalachian virgin forest. Ecology 61(5): 1169-1184.

Madden, M., R. Welch, T. Jordan, P. Jackson, R. Seavey, and J. Seavey. 2004. Digital Vegetation Maps for the Great Smoky Mountains National Park Final Report. Center for Remote Sensing and Mapping Science. University of Georgia. Department of Geography.

Messick, R. 2000. Old-Growth Forest Communities in the Nantahala-Pisgah National Forest. Unpublished Report.

Newfont, K. 2012. Blue Ridge Commons: Environmental Activism and Forest History in Western North Carolina. University of Georgia Press: Athens, Georgia.

Norman, D. 2012. The detection of forest structures in the Monongahela National Forest using LiDAR. Master's thesis, Marshall University, Huntington, West Virginia.

Pelton, M. 1996. The importance of old growth to carnivores in eastern deciduous forests. In *Eastern Old Growth Forests*, ed. M.B. Davis, 161-177. Washington, DC: Island Press.

Petranka, J. 2010. Salamanders of the United States and Canada. Washington, DC: Smithsonian Institution Press.

Pyle, C. 1988. The type and extent of anthropogenic vegetation disturbance in the Great Smoky Mountains before National Park Service Acquisition. Castanea 53 (3): 183-196.

Runkle, J. 1982. Patterns of disturbance in some old growth mesic forests of eastern North America. Ecology 63(5): 1533-1546.

Runkle, J. 1996. Central Mesophytic forests. In *Eastern Old Growth Forests*, ed. M.B. Davis, 161-177. Washington, DC: Island Press.

Southern Appalachian Man and Biosphere Cooperative. 1996. Southern Appalachian Assessment: Terrestrial Technical Report.

Schafale, M. 2012. Guide to the Natural Communities of North Carolina; Fourth Approximation. North Carolina Department of Natural Resources, Natural Heritage Program.

Simon, S. 2011. Ecological Zones in the Southern Appalachians; Third Approximation. Report to Nantahala-Pisgah National Forest.

Skowronski, N., K. Clark, R. Nelson, J. Hom, and M. Patterson. 2007. Remotely sensed measurements of forest structure and fuel loads in the Pinelands of New Jersey. Remote Sensing of Environment 108 (2007): 123-129.

Spies, T. 2004. Ecological concepts and diversity of old growth forests. Journal of Forestry 102 (3): 14-20.

The Nature Conservancy and Southern Appalachian Forest Coalition. 2000. Southern Blue Ridge Ecoregional Conservation Plan: Summary and Implementation Document. The Nature Conservancy, Durham, North Carolina.

United States Department of Agriculture. 1997. Guidance for conserving and restoring old growth forest communities on national forests in the Southern Region. Report of the Region 8 Old Growth Team. For. Rep. R8-FR 62. U.S. Department of Agriculture, Forest Service, Southern Region, Atlanta, GA.

Van der Sanden, J. J. and D. H. Hoekman. 2005. Review of relationships between grey-tone co-occurrence, semivariance and autocorrelation based image texture analysis approaches. Canadian Journal of Remote Sensing 38 (3): 207-213.

Appendix 1: Old Growth Classification from 2000 Messick Report:

OLD-GROWTH CLASSIFICATION SYSTEM

FOR OLD-GROWTH FOREST COMMUNITIES IN THE NANTAHALA-PISGAH NATIONAL FOREST

Class A is hard core old-growth.
Class B is soft core old-growth.
Class B+ contains both hard and soft core old-growth.

Class A old-growth forests where no significant signs of human disturbance to the forest canopy or understory could be determined. Canopies are dominated by older trees generally over 150 years of age. One hundred and fifty years is considered an appropriate coarse filter for old-growth candidacy as this corresponds to a period when logging was limited to areas near early settlement sites.

Note: Oak communities, where American chestnut was once most prolific, have often changed dramatically due to the effects of American chestnut blight. For this reason they often do not receive class A status. The effects of the blight are often less significant in Acidic Cove forest. Mesic communities such as Northern Hardwood and Rich Cove (Mixed Mesophytic) often show less effects of the blight, since American chestnut was not as prolific there.

Class B+ old-growth forests that have both class A and class B characteristics. Sites in this class tend to be large, with numerous forest communities, making it difficult to categorize the whole site. Uncut forests with canopy trees at or above 150 years may be present in these sites, yet the effects of disturbances such as blowdowns, American chestnut blight, or fire may be present in other forest communities within the site.

Class B old-growth forests exhibiting one of two different conditions:

1) the canopy is dominated by old-growth trees, yet signs of past human disturbance to the forest canopy or understory were found (generally a half century ago or longer). These stands have often been heavily impacted by American chestnut blight. Culling may also have occurred.

2) no sign of past human disturbance could be confirmed, yet the forest canopy is dominated by younger forest. These stands can range from 100 to 150 years in age and were possibly affected by natural disturbances.

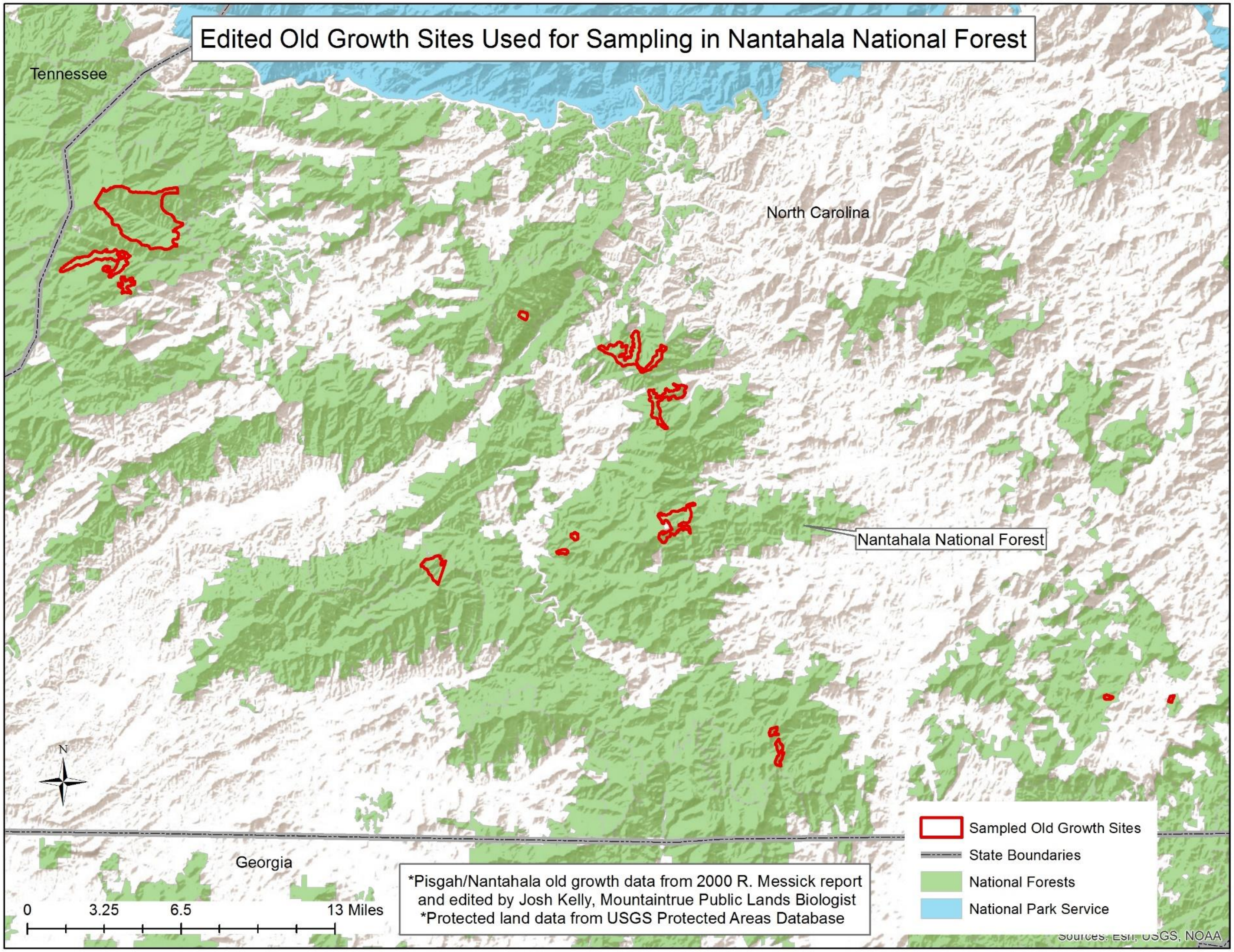
Class C forests with obvious signs of past human disturbance, yet containing appreciable old trees in the canopy or higher tree diversity than surrounding forests. **Forests in this class are suitable for old-growth recovery.** Some sites in this class are small. Others form buffers for class A, B+, or B old-growth. Forests in this class usually did not have extensive field work done in them due to time constraints.

Candidate Sites are considered worthy of a site visit due to a nomination, steep topography, or lack of access. These sites can show up as large un-inventoried stands in US Forest Service CISC data.

Classes A, B+, and B are considered existing old-growth forest. Class C and candidate sites are not.
Class B old-growth is distinct from second generation forest.

- 6 -

Appendix II: Edited Old Growth Sites for Nantahala Sampling



Appendix III: Edited Old Growth Sites for Pisgah Sampling