Topographic Analysis and Predictive Modeling using Geographic Information Systems

Steven Hall
*Clemson University*, hall.stevent@gmail.com

Follow this and additional works at: [https://tigerprints.clemson.edu/all_dissertations](https://tigerprints.clemson.edu/all_dissertations)

Part of the *Agriculture Commons*

Recommended Citation
[https://tigerprints.clemson.edu/all_dissertations/322](https://tigerprints.clemson.edu/all_dissertations/322)

This Dissertation is brought to you for free and open access by the Dissertations at TigerPrints. It has been accepted for inclusion in All Dissertations by an authorized administrator of TigerPrints. For more information, please contact kokeefe@clemson.edu.
TOPOGRAPHIC ANALYSIS AND PREDICTIVE MODELING USING GEOGRAPHIC INFORMATION SYSTEMS

A Dissertation
Presented to
the Graduate School of
Clemson University

In Partial Fulfillment
of the Requirements for the Degree
Doctor of Philosophy
Forest Resources

by
Steven Thomas Hall
December 2008

Accepted by:
Dr. Chris J. Post, Committee Chair
Dr. Elena A. Mikhailova
Dr. Thomas A. Waldrop
Dr. Patrick D. Gerard
ABSTRACT

This dissertation describes three GIS models developed to better model topographic features and the occurrence of mountain laurel (*Kalmia latifolia*) in the southern Appalachian Mountains. The first study presented “A LiDAR based GIS model to calculate Terrain Shape Index on a landscape scale”, attempts to develop a GIS based model to calculate the Terrain Shape Index (TSI). TSI is typically collected in the field using a series of elevation measurements to determine the average elevation change within the study plot. In this study, a GIS model is developed and TSI values compared to those collected using conventional methods. The second study, “A GIS model for determining landform type and slope position”, uses a progressive scanning method developed within a GIS to identify ridges and subsequently landform type. The results from this study are compared to landform classifications made visually by a group of volunteers. The third study presented, “A predictive GIS model for determining the probability of mountain laurel occurrence in the southern Appalachian Mountains”, attempts to develop a statistical model to better predict the presence or absence of mountain laurel on the landscape. Mountain laurel, often associated with decreased hardwood regeneration and its role as a vertical fuel, is important in both stand and fire management. In this study, results from a comprehensive, long term field study are used predict the occurrence of mountain laurel across the landscape. The GIS models described herein were designed to be efficient, user friendly and accurate in their results as well as easily transferable between parties and locations.
DEDICATION

This manuscript is dedicated to my daughter, Audrey. The beauty of her smile and the sound of her laughter make even the hardest of days easy.
ACKNOWLEDGMENTS

This dissertation is as much my own as it is my adviser’s, Dr. Chris Post. His friendship and support over the preceding years helped to make this research possible, and without it this document would not exist. I thank him for not only academic guidance, but also professional and personal.

Secondly, I would be remiss for not mentioning the rest of my committee. Their guidance on my research has been unparalleled. I would especially like to thank Dr. Pat Gerard for his statistical expertise and Dr. Tom Waldrop for providing years of data to serve as a confirmation of my work. Funding for Dr. Waldrop’s research was provided by the USDA\USDI Joint Fire Science Program and SRS-415 Center for Forest Disturbance Science.

Of course, this research would not have been possible at all had the LiDAR data not been provided, for free, from the North Carolina Floodplain Mapping Program. Their generosity is recognized on every page of this dissertation.
# TABLE OF CONTENTS

<table>
<thead>
<tr>
<th>CHAPTER</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td>I.</td>
<td>INTRODUCTION .......................... 1</td>
</tr>
<tr>
<td></td>
<td>Literature Cited ...................... 6</td>
</tr>
<tr>
<td>II.</td>
<td>A LIDAR-BASED GIS MODEL TO CALCULATE TERRAIN SHAPE INDEX ON THE LANDSCAPE SCALE ................. 9</td>
</tr>
<tr>
<td></td>
<td>Abstract ..................................... 9</td>
</tr>
<tr>
<td></td>
<td>Introduction ................................ 10</td>
</tr>
<tr>
<td></td>
<td>Methodology ................................ 12</td>
</tr>
<tr>
<td></td>
<td>Results ..................................... 15</td>
</tr>
<tr>
<td></td>
<td>Discussion .................................. 18</td>
</tr>
<tr>
<td></td>
<td>Conclusion ................................ 21</td>
</tr>
<tr>
<td></td>
<td>Literature Cited ...................... 22</td>
</tr>
<tr>
<td>III.</td>
<td>A GIS MODEL FOR DETERMINING LANDFORM TYPE AND SLOPE POSITION ................................... 25</td>
</tr>
<tr>
<td></td>
<td>Abstract ..................................... 25</td>
</tr>
<tr>
<td></td>
<td>Introduction ................................ 26</td>
</tr>
<tr>
<td></td>
<td>Methodology ................................ 28</td>
</tr>
<tr>
<td></td>
<td>Results ..................................... 33</td>
</tr>
<tr>
<td></td>
<td>Discussion .................................. 36</td>
</tr>
<tr>
<td></td>
<td>Conclusion ................................ 39</td>
</tr>
<tr>
<td></td>
<td>Literature Cited ...................... 41</td>
</tr>
</tbody>
</table>
# LIST OF TABLES

<table>
<thead>
<tr>
<th>Table</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Hypothetical error in GPS derived TSI values</td>
<td>20</td>
</tr>
<tr>
<td>3.1</td>
<td>Overall agreement of derived slope position values</td>
<td>35</td>
</tr>
<tr>
<td>4.1</td>
<td>Classification of LiDAR vegetation data into function classes</td>
<td>51</td>
</tr>
<tr>
<td>4.2</td>
<td>Average Values Used in Statistical Models</td>
<td>53</td>
</tr>
<tr>
<td>4.3</td>
<td>Cross validation table</td>
<td>54</td>
</tr>
</tbody>
</table>
## LIST OF FIGURES

<table>
<thead>
<tr>
<th>Figure</th>
<th>Description</th>
<th>Page</th>
</tr>
</thead>
<tbody>
<tr>
<td>2.1</td>
<td>Map of North Carolina study site</td>
<td>14</td>
</tr>
<tr>
<td>2.2</td>
<td>Basic GIS model used to calculate TSI</td>
<td>15</td>
</tr>
<tr>
<td>2.3</td>
<td>Scatter plot comparing LiDAR and field derived TSI values</td>
<td>17</td>
</tr>
<tr>
<td>2.4</td>
<td>Map of LiDAR derived TSI values</td>
<td>18</td>
</tr>
<tr>
<td>3.1</td>
<td>Map of North Carolina study site</td>
<td>29</td>
</tr>
<tr>
<td>3.2</td>
<td>Three-dimensional representation of elevation model with identified ridges</td>
<td>34</td>
</tr>
<tr>
<td>3.3</td>
<td>Three-dimensional representation of slope position</td>
<td>35</td>
</tr>
<tr>
<td>4.1</td>
<td>Map of North Carolina study site</td>
<td>49</td>
</tr>
<tr>
<td>4.2</td>
<td>Profile view of LiDAR data</td>
<td>52</td>
</tr>
<tr>
<td>4.3</td>
<td>Map of predicted mountain laurel presence\absence</td>
<td>56</td>
</tr>
<tr>
<td>4.4</td>
<td>Probability map of mountain laurel presence</td>
<td>57</td>
</tr>
<tr>
<td>4.5</td>
<td>Map of vertical structure</td>
<td>58</td>
</tr>
<tr>
<td>4.6</td>
<td>Comparison between vertical structure and probability maps</td>
<td>61</td>
</tr>
</tbody>
</table>
CHAPTER I
INTRODUCTION

As technology advances, so does the ability to predict and model real world phenomenon. Typical modeling attempts to accurately demonstrate and quantify the interaction of multiple environmental variables in order to not only understand the role of such parameters in nature, but also to predict their behavior under varying conditions.

One of the more common methods of modeling and predicting spatial data is through the use of a Geographic Information System (GIS). A GIS is a “computer based system to aid in the collection, maintenance, storage, analysis, output, and distribution of spatial data and information” (Bolstad 2003). Described in lay terms, a GIS is comprised of specialized software designed with the explicit purpose of handling spatial data. Spatial data is simply data – such as soil temperature – that includes a spatial reference (e.g. latitudinal and longitudinal coordinates). The ability to associate a dataset with its spatial location allows for more than standard statistics to be performed; it allows researchers to incorporate geographic locations into their analysis and also understand what role, if any, the spatial location itself fills.

The use of GIS is most often coupled with environmental research. For example, GIS has been used in a multitude of environmental research studies, such as habitat identification (Creque et al. 2005; Eikaas et al. 2005; Pettorelli et al. 2005; Store and Jokimaki 2003), habitat restoration (Gillenwater et al. 2006; Greer 1994; Richardson and Gatti 1999), soil science (de Paz et al. 2006; Liu et al. 2004; Lufafa et al. 2003), water quality (Baker et al. 2001; Vivoni and Richards 2005; Wang and Yin 1997) and fire
ecology (Keane et al. 2001; Ruiz-Gallardo et al. 2004; Theodore 2004) to name just a few. Other fields taking advantage of the many capabilities of a GIS include city and regional planning (Jat et al. 2008), public health (Lu 2005) and even sociology (Brown 2003). In truth, GIS is not limited by the field in which it is used, but only by the availability of spatial information.

A large portion of GIS analysis is performed using ArcGIS, a software package developed by Environmental Systems Research Institute (ESRI 2006). In it, users can create anything from simple, explanatory maps to complex, analysis-driven models. In ModelBuilder®, users can connect various analysis tools in the form of a workflow diagram to aid in model creation and processing. ModelBuilder® keeps track of all analysis tools, their interdependencies and gives users a way to create large, complicated models with little effort. Additionally, ModelBuilder® allows models to be quickly and easily edited, and analysis tools can be run separately or in unison (Ormsby et al. 2004). ModelBuilder® is used extensively in each of the chapters presented in this paper.

Each of the succeeding chapters represents a different approach to modeling environmental variables using a GIS. All spatial analysis was completed using ArcMap 9.2 (ESRI 2006) and statistical analysis using SAS 9.13 (SAS Institute 2008) and Maple 11 (Maplesoft 2008). Since each project was undertaken individually with the explicit purpose of publication, some information contained within the chapters may overlap.

The overall theme behind the studies was to develop methodologies that better represent the physical world using a GIS. In each case, information typically derived from rigorous field work was generated using a variety of spatial analysis tools. In the
first study, a GIS model was developed to mimic the procedure used by McNab (1989) to determine the Terrain Shape Index. The Terrain Shape Index, or TSI, has been used in numerous studies, but is often related to site quality and tree growth (McNab 1989). Hutto et al. (1999) used TSI data collected from the field in a GIS model to predict the landscape ecosystem classification unit (LEC). Essentially, TSI measures the microtopography of a site, classifying it as either convex or concave.

To best simulate the methodology used in the field to calculate TSI, a raster based analysis was developed using a custom kernel, or neighborhood. Results for the study were compared to field derived data at 245 plots located in western North Carolina. These sites were generated randomly as part of a forest fuels study (Waldrop et al. 2007). All of the confirmation data used for each of the studies contained herein were obtained from Waldrop et al. (2007).

The second study discussed here is the ability to determine slope position – ridge, side slope or cove – in an automated fashion using a GIS model. Slope position has been cited as an important factor in many studies, but is often cited in fire ecology studies, particularly those concerned with fuel loading (Waldrop et al. 2007). A universal method for determining slope position was first developed by McNab (1993). In his methodology, McNab used the degree of inclination to surrounding landforms to determine on what type of landform a study site resided.

To best estimate slope position in a GIS model, a progressive scanning technique was developed and employed. This allowed for the slope position to be determined across an entire region, instead of at only individual study sites. Similar to the TSI study, the
study sites from Waldrop et al. (2007) were used for comparison. However, instead of using the predefined slope position used in that study, a group of volunteers were asked to manually classify all 250 field sites. Doing so helped to eliminate user bias as well as create a distribution of the most common slope classifications.

The third and final study conducted was the development of a predictive statistical model, estimating the probability for mountain laurel (*Kalmia latifolia*) presence. Mountain laurel has been shown to increase the likelihood of crown fires (Vose et al. 1999; Waldrop and Brose 1999; Waldrop et al. 2007), and thus it serves an important role in wildfire management and planning. In order to predict the presence of mountain laurel, various types of spatial data – including a LiDAR based canopy model – were used to create a discriminant function based on presence and absence data collected by Waldrop et al. (2007). Using the discriminant function, it was possible to calculate a probability raster that denotes the likelihood of mountain laurel presence in the southern Appalachian Mountains included in this study. Results of this study may be found in Chapter IV.

The difference between these studies and other GIS models that have attempted to produce similar output is the use of LiDAR – Light Detection and Ranging – as the foundation for the elevation model. LiDAR, while an advanced remote sensing technique, is very simple in nature. A laser, mounted to the underside of an aircraft, travels over an area of concern firing thousands of laser pulses to the ground per second. As the laser pulses are reflected from the earth’s surface, the elevation of the reflected pulses can be determined allowing for the creation of highly accurate elevation models (Dubayah and Drake 2000; Lefsky et al. 2002).
The inclusion of LiDAR data in each of these studies offers several advantages over conventional methods. Because LiDAR is collected at such a high density (in some cases, a data point every 1.5 m) and often has a vertical accuracy of less than 25 cm, it may help to minimize the error found in ordinary elevation models, which, at best, have a vertical accuracy of 1 m and horizontal resolution of 30 m (USGS 1998). More accurate elevation models are available from the federal government, but coverage of the continental United States is not yet complete (Anderson et al. 2006).

At first glance, each of the three studies mentioned here may seem only loosely associated with one another. However, when considered at a landscape scale and the potential savings in time, money and resources – the efficacy of the GIS models presented here is clear. Not only is there potential to eliminate certain aspects of data collection, but most importantly, there is great potential for advances in fire ecology and wildfire management should reliable predictions of mountain laurel presence be possible.
LITERATURE CITED


CHAPTER II

A LIDAR-BASED GIS MODEL TO CALCULATE TERRAIN SHAPE INDEX ON THE LANDSCAPE SCALE

ABSTRACT

Terrain Shape Index is a common measure for determining the shape of the micro landform. Previous studies have shown that terrain shape is significantly correlated with site quality and tree growth. In this study, a GIS model was developed to accurately calculate the Terrain Shape Index using a LiDAR derived elevation model. Index values from the GIS model were compared to Terrain Shape Index values calculated on 245 field sites located in western North Carolina. Results indicate that the GIS based Terrain Shape Index calculations are similar to field collected values. A T-test showed no significant difference (p=0.26) between the means of the two datasets, and when plotted there was no significant difference between the intercept and slope (p=.055), implying agreement among the datasets. The $R^2$ for a linear model, however, is .42. A low explanation of variance between the models may at least be partially explained by error associated in both LiDAR and field collected data, particularly if the plots are located in dense vegetation and elevation data is collected with a handheld GPS device. In total, the GIS model correctly predicted nearly 3 out of 4 study sites as being either convex or concave when compared to field collected data.
INTRODUCTION

Since first being introduced to mainstream ecology, the Terrain Shape Index (McNab 1989) has been used as the standard method to quantify the effects of minor landform on vegetation growth. McNab’s (1989) objective was to develop a methodology to express landform as a continuous, objectively measured variable. He did this by calculating the mean relative difference in elevation between the center of the study plot and eight surrounding points at 45° intervals. When divided by the plot radius, the result quantifies the shape of the minor landform; negative indices indicate that the study site is on a convex landform and positive indices indicate that the study site is on a concave landform. McNab concluded that the Terrain Shape Index (TSI) was significantly related to the total tree height of yellow-poplar trees in his study. Since then, many studies have used his TSI method to help quantify vegetation patterns and growth (Abella and Covington 2006; Humphries et al. 2008; Thompson et al. 2006).

In this study, a Geographic Information System (GIS) model was created to calculate TSI using Light Detection and Ranging (LiDAR) data. Results were compared to field calculated TSI values at 245 study sites. Comparisons were made to determine whether or not computer based TSI models would be an accurate substitute for field generated data.

A GIS is a “computer based system to aid in the collection, maintenance, storage, analysis, output, and distribution of spatial data and information” (Bolstad 2003). A GIS is a means of using the spatial information (such as latitude and longitude) of a feature to perform analysis on it. Analyses can range from simple to the complex; in this case, a
GIS is used to calculate the TSI values for a given region. All GIS analyses in this study were performed using ArcMap 9.2 (ESRI 2006).

Since its inception, GIS analysis has been used in many ecological studies. Some common applications of GIS include habitat restoration (Massey et al. 2008; Mollot and Bilby 2008), soil mapping (Lewis et al. 2005; Tischler et al. 2007) and hydrology (Frankenberger et al. 1999; Zheng and Baetz 1999) as well as several diverse landform mapping studies (Blaszczyński 1997; Gallant et al. 2005; Jenson and Domingue 1988). This study differs from previous attempts to calculate TSI using a GIS (Humphries et al. 2008) in that it attempts to calculate TSI identical to McNab’s (1989) method using a LiDAR surface model and an automated mathematical approach.

Despite being an advanced form of remote sensing, the underlying concept behind LiDAR is simple in nature. Instruments mounted on the underside of an aircraft emit laser pulses at a specified target. The return signal from these pulses (i.e. the reflected light) from the target’s surface are collected by a receiver on the LiDAR instrument (Dubayah and Drake 2000; Lefsky et al. 2002). By measuring the time elapsed between the initial laser pulse and the reception of the return signal, along with the distance traveled, the surface elevation can accurately be calculated (Dubayah and Drake 2000; Lefsky et al. 2002). Furthermore, laser pulses often travel through vegetation (including some dense vegetation), giving researchers elevation data of both the ground surface and forest canopy. When combined with the number of elevation points retrieved per second, this allows for exceptionally accurate surface models to be produced (Lefsky et al. 2002; Means et al. 1999). For the purposes of this study, a LiDAR derived digital elevation
model was used as the foundation for TSI calculation. The objective of this study was to develop a computer based model to accurately calculate the TSI for a region and be comparable to values derived from field observations.

METHODOLOGY

LiDAR data for this study was acquired from April to December 2003 by EarthData International of North Carolina for the Floodplain Mapping Program, part of the North Carolina Division of Emergency Management. The horizontal datum is NAD83 North Carolina State Plane Feet and the vertical datum NAVD88 US Survey Feet.

The area of concern for this project is a 33.7 Km² (13 mi²) section of North Carolina, covering a portion Macon county, NC (Figure 2.1). The landscape is considered mountainous, as the Appalachian Mountains are the dominant topographical feature for this region.

The initial Digital Elevation Model (DEM) was created using the LiDAR points identified as bare earth. These points were interpolated in ArcMap 9.2 (ESRI 2006) using Kriging, a commonly used geostatistical method for creating surface models (Lloyd and Atkinson 2006). Optimal Kriging parameters were automatically determined in ArcMap using the Geostatistical Analyst extension. The cell size for the DEM is 5 feet, which is comparable to the average posting between the LiDAR data points and matches the units used for both the horizontal and vertical datum.

To best duplicate McNab’s (1989) procedure, it was necessary to develop a raster based model that queried cells at 45° intervals, 18 m (59 ft) from the plot center. In this
study, individual plot locations were not used as plot centers; instead, each cell in the
DEM was considered a plot center and a predictive TSI raster was created for the entire
study area.

To best create the required interval, a new, single point shapefile was created and
placed in an empty map document. From this point the Euclidean direction was
calculated and extracted to fit the 18 m boundary required. Map Algebra was used to
identify the eight boundary measurements (1 every 45°) and the result of which was
converted into an ASCII text file. The ASCII file was then used as the input kernel file
for a focal statistics irregular neighborhood. A second ASCII text file was created that
represented only the center point of the 18m plot.

The sum of the boundary measurements were calculated, and from that, the
elevation from the center cell, multiplied by eight, was subtracted. The product of this
was divided by eight and multiplied by the plot radius as seen in the following equation:

\[
\frac{(\sum Z_{1-8}) - (Z_0 \times 8)}{8 \times \text{Radius}},
\]

ultimately producing the final TSI values for the raster.

Although the method of calculation is slightly different than McNab’s (1989)
method described earlier, identical TSI results may be obtained regardless of whether or
not McNab’s approach or the one described here is used. This approach was used over
the conventional approach because it limits computer processing time and minimizes the
size of the final model (Figure 2.2).
Figure 2.1: Map of North Carolina study sites, all of which were contained within the Nantahala National Forest.
To best confirm the results, a simple linear regression was used to predict field based TSI measurements using LiDAR data. The slope and intercepts of each dataset were compared, testing whether the slope is equal to 1 and the intercept equal to zero, implying agreement among the datasets. Additionally, the correlation coefficient was calculated as well as frequency tables to determine how often the two datasets agreed according to landform shape (concavity and convexity). A paired T-test was also conducted to determine whether or not a statistical difference between the means of the two datasets existed. All statistical analysis was performed using SAS 9.13 (SAS Institute 2008).

RESULTS

Basic statistical analysis shows that the two datasets are comparable. A paired T-test reveals that there is no statistical difference ($p=0.2615$) between the mean observations from the field collected TSI values and those generated by ArcMap.
Additionally, both the field TSI and the LiDAR derived TSI values are normally distributed (p=0.4461 and p=0.1879, respectively). The Pearson Correlation Coefficient for LiDAR and field derived TSI values were significant at .65 (p < 0.0001), and frequency tables reveal that 73% – nearly 3 out of 4 study sites – match one another in terms of convexity and concavity. A scatter plot (Figure 2.3) shows the linear relationship between the TSI values. Comparisons of the slope (0.984) and intercept (-0.007) of both datasets reveal that they are in agreement (p=0.5521) with one another. It should be noted, however that the R^2 for the linear model is 0.4285. Further discussion as to why this may be is found later in this chapter.

The final output from the GIS analysis is a raster layer containing calculated TSI values (Figure 2.4). In it, high TSI values (concave landforms) are bright white in color. Low TSI values (convex landforms) are dark in color.
Figure 2.3: Scatter plot comparing LiDAR and field derived TSI values
DISCUSSION

Results indicate that the datasets do not statistically differ from one another, and that 3 out of 4 study sites were correctly identified as either convex or concave. However, it should be noted that a regression model using the LiDAR derived values accounts for
only 42% of the variability seen in the field calculated values. There are several potential explanations for this apparent lack of predictive ability.

One rationalization for poorly explained variability in this model is the inherent error associated with LiDAR data. Given that LiDAR depends on light reflectance to accurately determine surface height, areas with little to no vegetative cover may have a higher density of surface points (and thus result in a more accurate surface model) than comparable regions with a dense canopy (Barber and Shortridge 2005; Cowen et al. 2000; Jensen et al. 1987) because not all light pulses will reach the earth’s surface. Cowen et al. (2000) estimates that when the canopy is 80-90 percent closed, only 10-40 percent of the light pulses reach the ground and are reflected back to the LiDAR sensor. Additionally, LiDAR collected for this study are horizontally and vertically accurate within 1-25 cm (.01 -.2 5m or .03-.8 2ft), thus there is potential for some of the elevation values to be inflated, albeit by a negligible amount.

A second potential cause for disagreement is the accuracy of field collected data. Perhaps the most common and likely form of error found in the field observations is the error associated with setting up the field plot itself. In cases where vegetation is dense and the topography highly variable, it can be very difficult to accurately determine the plot boundaries and thus take plot measurements in the proper location. Furthermore, all of the elevation values for the field collected TSI dataset were obtained using a Trimble Geo Explorer 3 Global Positioning System (GPS) as a C/A code receiver. Typically, a C/A code receiver has an accuracy of 3-30m; post-differential correction may improve the accuracy of GPS points to within a few meters (Bolstad 2003). Still, even with post-
differential correction, changes of even a few meters in horizontal or vertical position can potentially have a large influence on the final TSI.

For example, a typical field plot as described in this paper requires eight measures of elevation around the plot center in order to calculate the TSI. If, hypothetically, four of the eight measurements experienced even minor amounts of error (in this example, a difference of 3 m – well within the range of expected GPS error) – whether due to satellite obstruction, poor signal or any other number of causes – the final TSI can be considerably different (Table 2.1). It is reasonable, then, to consider the possibility that even a small amount of error within GPS derived elevations could mean the difference between classifying a study site as convex or concave.

Table 2.1: This table outlines a hypothetical situation, showing the difference in final TSI values should only half of the field collected points have a negative 3m error associated with it.

<table>
<thead>
<tr>
<th>Point 1</th>
<th>Point 2</th>
<th>Point 3</th>
<th>Point 4</th>
<th>Point 5</th>
<th>Point 6</th>
<th>Point 7</th>
<th>Point 8</th>
<th>Avg. Elev. Difference</th>
<th>TSI</th>
</tr>
</thead>
<tbody>
<tr>
<td>GPS Elev.</td>
<td>1</td>
<td>0</td>
<td>-1</td>
<td>2</td>
<td>2</td>
<td>3</td>
<td>4</td>
<td>-2</td>
<td>1.125</td>
</tr>
<tr>
<td>True Elev.</td>
<td>-2</td>
<td>0</td>
<td>-4</td>
<td>2</td>
<td>-1</td>
<td>3</td>
<td>1</td>
<td>-2</td>
<td>-.375</td>
</tr>
</tbody>
</table>

Ultimately, it is doubtful that the low level of explained variance can be solely attributed to LiDAR, GPS or field error. Instead, it is likely to be a combination of all factors. However, when considering the inherent error associated with both forms of measurement, the LiDAR based TSI calculations may serve as an adequate supplement or perhaps even replacement of field based TSI measurements. The computer based model allows the user to calculate TSI over a large landscape (instead of on a point-specific basis) and is simple to construct in addition to being time efficient.
CONCLUSION

In this study a computer based model was created using ArcMap 9.2 to calculate the Terrain Shape Index for a 33.7 Km$^2$ (13 mi$^2$) area. Results were compared to 245 study sites where field based TSI measurements were taken. A 73% accuracy rate of identifying a site as either convex or concave was achieved. Most of the minor inconsistencies between the two datasets are likely due to the inherent error present in both LiDAR and GPS derived measurements, as well as field measurement error associated with dense vegetation or other obstructions. With this in mind, results indicate that the LiDAR based model produces an accurate representation of minor landform shape on a landscape. This methodology offers a simple and quick alternative to field based measurements as well as producing results for an entire area of concern instead of being limited to individual study sites.


CHAPTER III
A GIS MODEL FOR DETERMINING LANDFORM TYPE AND SLOPE POSITION

ABSTRACT

Landform identification (ridges, side slope, cove) is often a part of field based ecology studies. In this study, a GIS model was developed to progressively scan a digital elevation model (DEM) and determine the slope position. Once the ridges were identified, a stream layer was used as to represent the bottom of the slope. With both the ridges and the slopes identified, the data were reclassified into three equally divided groups – ridges, side slopes and coves. To confirm the results generated in the GIS model, 245 randomly generated study sites were classified according to their slope position on a digital elevation model by a group of 16 volunteers. Each volunteer worked independently when classifying the study sites, the results of which were compiled and compared to those generated in the GIS model. Only study sites which had more than nine (56%) volunteers classify the sites equally were used. Sites not fitting this criterion were not believed to have a distinct classification, and therefore would produce misleading results. Two-hundred fifteen sites met the criterion. When compared to the results generated by the GIS model, correct identification of slope position was consistently equivalent to the standards placed upon the volunteers. When there was 60% agreement among volunteers, the GIS model correctly identified 60% of slope positions. This trend is visible throughout the study, including when there is 100% agreement of site classification among volunteers.
INTRODUCTION

The characteristic shape and form of a landmass – often referred to as the landform – has long been regarded as an important factor affecting the distribution of forest tree species. Some of the earliest studies to recognize the importance of landform on vegetation include Auten’s (1945) yellow-poplar study, where he describes the affect of landform on soil moisture and poplar distribution. As McNab (1993) explains, Auten (1945) went so far as to classify landform into coves, slopes and ridges, but little more.

Since then, the significant influence of landform on environmental attributes has become well known. Numerous studies show the effects of landform on variables such as soil moisture (Hanna et al. 1982; Helvey et al. 1972; Lookingbill and Urban 2004; McNab 1993; Yeakley et al. 1998) and vegetation distribution (Abella 2003; McNab 1993; Miller and Franklin 2002).

Despite the number of studies that incorporate landform as a variable, there did not exist a standard method for determining landform shape until McNab (1993) introduced his Landform Index (LFI) methodology. The calculation of McNab’s index is straightforward and relatively simple to compute. Based on the observers’ physical position on the landscape, one may use a clinometer to measure the gradient from the plot center to the horizon in eight directions, averaging all of the gradients to determine the index. A concave landform (such as a valley) will have a positive index and a convex landform (such as a ridge) a negative index.
Although the LFI index is an effective and accurate method of identifying topographic landform, it has the potential to be time and labor intensive should a large number of study plots need to be surveyed. The objective of this study was to develop a customizable Geographic Information System (GIS) procedure that identifies landform type (ridge, slope and cove) as well as the approximate slope position based on a 0-100 scale (100 being the bottom of a slope, 0 being the ridge).

A GIS is a “computer based system to aid in the collection, maintenance, storage, analysis, output, and distribution of spatial data and information” (Bolstad 2003). A GIS is a means of using the spatial coordinates (such as latitude and longitude) of a feature to perform analysis on it and its related attributes. Analyses can range from a straightforward proximity analysis to a complex, model-driven procedure.

Over its relatively short lifespan, GIS analysis has been used in many ecological studies. Some of the more common applications of GIS in ecology include habitat restoration (Massey et al. 2008; Mollot and Bilby 2008), soil mapping (Lewis et al. 2005; Tischler et al. 2007) and hydrology (Frankenberger et al. 1999; Zheng and Baetz 1999) as well as several diverse landform mapping studies (Blaszczynski 1997; Gallant et al. 2005; Jenson and Domingue 1988). This study differs from previous landform characterizations in that it attempts to identify landform type using a LiDAR (Light Detection and Ranging) surface model as well as an automated mathematical approach in conjunction with known hydrologic pathways.

LiDAR has become a popular and cost effective method of retrieving three dimensional surface data. Although a technologically advanced technique, the underlying
The concept behind LiDAR is simple in nature. Instruments – usually mounted on the underside of an aircraft – emit laser pulses at a specified target. These pulses are reflected (often referred to as the return signal) off the target’s surface and collected by a receiver on the LiDAR instrument (Dubayah and Drake 2000; Lefsky et al. 2002). By determining the time elapsed between the initial laser pulse and the reception of the return signal, along with the distance traveled, one can accurately calculate elevation of a given LiDAR point (Dubayah and Drake 2000; Lefsky et al. 2002). Additionally, laser pulses will often travel through vegetation (including some dense vegetation), giving researchers elevation data of both the ground surface and forest canopy; when combined with the number of elevation points retrieved per second, this allows for exceptionally accurate surface models to be produced (Lefsky et al. 2002; Means et al. 1999). In this study a LiDAR derived digital elevation model is used as the foundation for landform identification. The objective of this study was to identify topographical ridges and calculate, on a 0-100 scale (0 being the ridge) the approximate slope position.

**METHODOLOGY**

LiDAR data for this study was collected and finalized between April and December 2003 by EarthData International of North Carolina for the Floodplain Mapping Program, part of the North Carolina Division of Emergency Management. The horizontal datum is NAD83 North Carolina State Plane Feet and the vertical datum NAVD88 US Survey Feet. The area of concern for this project is a 33.7 Km² (13 mi²) section of North Carolina, including parts of both Clay and Macon counties (Figure 3.1). Topography for
Figure 3.1: Study sites were located in the Southwestern corner of Macon County, NC, adjacent to Clay County and were part of the Nantahala National Forest.
this region is considered mountainous, as the southern Appalachian Mountains are the dominant topographical feature.

ArcGIS 9.2 (ESRI 2006) was used to create the digital elevation model (DEM) as well as the corresponding landform identification model. The cell size for the DEM was 5 ft (1.5 m) to best match the resolution of the LiDAR data. Elevation data was interpolated using Kriging, a common and often recommended method for creating accurate and high resolution DEMs (Lloyd and Atkinson 2006).

In order to best identify ridgelines on a DEM, focal statistics were used to progressively scan the elevation model both horizontally and vertically. Focal statistics, one of the analysis tools available in ArcGIS, is an overlapping form of analysis that allows the user to calculate simple statistics – such the mean, standard deviation, minimum and maximum – for a specified neighborhood in a raster data layer. The shape and size of the neighborhood is customizable by the user. For the purposes of this study, a custom kernel, or neighborhood, was created.

The kernel developed for this study was linear in nature. For both the horizontal and vertical scans, the kernel was 500 cells long (with respect to orientation) and one cell wide. Two kernels were developed each for the horizontal and vertical scans for a total of four kernel files. Each of the kernels collected the maximum elevation value for different sections of the neighborhood. For example, Horizontal Kernel 1 (HK1) collected the maximum elevation value for the left third of its 500 cell neighborhood. HK2 collected elevation values for the right third. The vertical kernels performed similar operations.
A kernel length of 500 cells was chosen because it is the best compromise between correct ridge identification and processing speed. A shorter kernel size could be used if topographic features were closer together. For example, should one be interested in not only the main ridges but all convex areas along a slope, a shorter kernel would be appropriate. Conversely, a longer kernel would add to processing time but would identify only the most major of ridges.

With data layers describing the topographic trends, it was then possible to combine the information and determine where ridgelines existed. Ideally, ridgelines exist where a cell’s elevation is greater than the surrounding cells in at least three directions to help ensure saddles (a ridgeline that connects two higher ridges) are identified. Essentially, the maximum elevation value in the elevation model must be greater than the maximum elevation value found in at least three out of four combinations of the other kernels (HK1, HK2, VK1 and VK2).

A slope and elevation parameter were added into the model to allow a user to more strictly define where ridgelines should be identified. For example, should a user wish to only identify ridges with a slope less than 20%, the slope parameter can constrain the results to slopes less than 20%. Similarly, the elevation parameter allows a user to adjust the sensitivity of ridge identification. The larger the elevation parameter is, the more likely the model is to identify a cell as a ridge.

With the ridges identified, it was possible to gauge the slope position, on a zero to one-hundred scale (zero being the ridge) using a stream layer as the reference for the
bottom of the slope. The stream layer was acquired from the National Hydrography Dataset (USGS 2008a).

By calculating the Euclidean distance (straight line distance) from the identified ridges and the streams, and performing simple arithmetic, one can approximate the slope position by dividing the distance from the nearest ridge by the sum of the stream distance and ridge distance:

\[
\frac{\text{RidgeDist}}{\text{StreamDist} + \text{RidgeDist}}.
\]

This allows the model to create a 0-100 scale of the slope position, with 0 being the ridge and 100 the bottom of the slope. Reclassifying the slope position into three broad categories makes it possible to label high (ridge), medium (side slope) and low slope positions. A ModelBuilder® model was created to streamline the workflow for this project (Appendix A).

To best confirm the results of this study, the slope position from 245 randomly generated study sites were used for comparison. More information on the study sites may be found in Waldrop et al. (2007), as the sites were used in an extensive forest fuels study in the southern Appalachian Mountains.

Classification of study sites as existing on a ridge, medium or low slope was performed independently by 16 volunteers (all familiar with using a GIS), each classifying all 245 study plots. General instructions on how to classify the points were provided to all volunteers, including a sample classification for all three classes. Volunteers used a 1/3 Arc Second elevation model from the USGS National Elevation
Dataset (USGS 2008b) for their classifications. It should be noted that the elevation model used for the volunteer classification was independent of the one created for this study to eliminate any misclassification due to elevation model error. The sites with the highest levels of agreement among volunteer classifications were used to confirm the results of this study. Volunteer classifications where more than half (9 out of 16 volunteers) agreed on the slope position were used. Multiple volunteers were used to classify the data such that a consistent, independently generated classification could be achieved with minimal user bias. This approach, when compared to field based observations, offers a more objective perspective than would be possible with field based observations. Field based observations, particularly for the large number of study plots used here, would be difficult to coordinate using the same number of volunteers. Furthermore, field based observations of this type can often be obscured by vegetation, further limiting the accuracy of such measurements.

RESULTS

Output from the GIS model designed for this study consisted of three raster (grid) layers. The first and arguably most important output from the model is the ridge identification raster. This raster displays where ridges (in red) occur on the landscape as defined by the GIS model (Figure 3.2). Visually, the majority of ridges and ridge saddles are identified.
The second useful output from the GIS model is the slope position, on a 0-100 scale with 0 being the top of the ridge. Using hydrography data as a base of calculation, the percentage distance between known streams and ridges were calculated. Slope position can then be reclassified into three categories: High, Medium and Low slope position with respect to ridges, side slopes and drainages (Figure 3.3).

The final slope position raster was used as a comparison against the classifications made by the group of volunteers. Study sites with less than half of the volunteers in agreement were not included; any results derived from including them would have been misleading knowing that a consistent classification was not available. Much of the disagreement found in the sites not used occurred in areas where a site could be classified as either low or medium, or medium and high depending on the volunteer’s interpretation of the elevation model.

Figure 3.2: Three-dimensional representation of the elevation model with the identified ridges in red.
Of the 245 study sites, 215 had nine or more (i.e. 56%) volunteer classifications in agreement. A complete breakdown of how the user classified points compared to the GIS generated classifications is available in Table 3.1.

Figure 3.3: Three-dimensional representation of the elevation model with the slope position (high, medium, low) as an overlay.

Table 3.1: Overall agreement between volunteer derived classifications and the GIS derived classification. As user agreement rises, so does agreement with the GIS model.

<table>
<thead>
<tr>
<th>Level of Agreement</th>
<th>Volunteer Agreement</th>
<th># Plots</th>
<th>Model Derived Plots in Agreement</th>
<th>Percent Agreement</th>
</tr>
</thead>
<tbody>
<tr>
<td>&gt; 50%</td>
<td>9 of 16</td>
<td>215</td>
<td>137</td>
<td>63.72%</td>
</tr>
<tr>
<td>80%</td>
<td>13 of 16</td>
<td>84</td>
<td>69</td>
<td>82.14%</td>
</tr>
<tr>
<td>90%</td>
<td>14 of 16</td>
<td>55</td>
<td>50</td>
<td>90.91%</td>
</tr>
<tr>
<td>95%</td>
<td>15 of 16</td>
<td>26</td>
<td>25</td>
<td>96.15%</td>
</tr>
<tr>
<td>100%</td>
<td>16 of 16</td>
<td>5</td>
<td>5</td>
<td>100.00%</td>
</tr>
</tbody>
</table>
DISCUSSION

Typically, when the slope position is determined in the field, it is done through visual estimation, which may or may not be obscured by either vegetation or a topographical feature. In this study, computer based evaluations of slope positions were used to determine the accuracy of the GIS model created. Results of this study indicate that the GIS model described herein may be an effective means of determining slope position to approximately the same level of accuracy achieved when computer based visual classification is employed. As more stringent requirements are placed on visually classified study sites, a higher degree of correct identification can be seen in the GIS model.

A number of factors play a role in the effectiveness of both visual identification and computerized classification. Perhaps the most significant challenge to address when identifying topographic landforms is the schema by which the landforms are defined. During the course of having the volunteers classify all 245 study sites, they were given only minor instructions on how to classify the study sites. The purpose was to let their interpretations of what a high, medium and low slope should be dictate how they classified the study sites. Too much instruction would have skewed their results to match that of the GIS model and therefore potentially invalidating some or all of this study. Thus, the user classifications compared to the results from the GIS model are completely independent of one another; the volunteers had not seen the results of this study before performing the classification and were unaware of how a high, medium and low slope was defined within the context of the GIS model.
The use of the GIS model described in this paper offers several advantages over conventional slope position identification. First, and likely the most appealing aspect of this study, is the ability to identify landforms across an entire region instead of just at individual study sites, as would normally be the case. To confirm the results of this study, 245 study sites were used; in reality, however, any number of confirmation points – 500, 1,000, 10,000 or more – could have been generated and compared against the results because landform was identified throughout the region. Regardless of the number of confirmation sites, it takes the same amount of time to identify the landform type using the GIS model, whereas field identification of the points would take an exponentially longer amount of time as the number of plots increases. This opens the potential to not only consider the landform type of a particular study site within an analysis, but also the landform surrounding the study site.

A second point of consideration is the accuracy of field based position estimates. Often times field measurements are visually estimated, and unless the observer can reasonably see both the top and bottom of the slope, it can be difficult to accurately estimate one’s position, particularly if the observer is surrounded by dense vegetation. The GIS model described in this paper negates the effect of dense vegetation on position estimates.

A third advantage of this GIS model when compared to conventional methods is the ability to customize it to one’s analysis. Because the model is heavily dependent on the accuracy of hydrographic data, it allows a significant amount of flexibility when determining the final results. For example – stream networks are relatively simple to
delineate and divide into classes (ephemeral, intermittent etc…). Different classes of hydrographic data could potentially be used to alter the level of detail and specificity found in the final slope position classification.

There are, however, two items to consider that may produce unwanted effects within the GIS model. First is the reliability of the elevation model that is being used as the basis for ridge identification. In this study, a LiDAR DEM was created and used as the foundation for all the analysis. Great effort was put forth to ensure that the original LiDAR data consisted of entirely bare earth points with no remnants of vegetation. However, it is highly unlikely that all traces of vegetation were removed from the initial bare earth data set, meaning that some vegetation artifacts may have been incorporated in the DEM used in this project. While this alone likely had a little to no effect on the final results, it is possible that in a similarly designed project using a poorly classified DEM, vegetation artifacts could have a negative influence on the validity of the analysis.

A second item to consider is the precision of computer based modeling. While it can be challenging to visually see (both on a computer screen and in the field) every undulation of topography, it is not necessarily a difficult task for a computer. It was found during the course of this study that the GIS model described here has the ability to identify not only topographic ridges, but also all convex regions naturally found in the landscape. It was necessary to design the model to limit the output to only main ridges, and this was accomplished by using the large kernel neighborhoods described earlier. Smaller kernels identify the same ridges as the large kernels, but also identify many of the small regions not of concern to this study. It is important to consider the type of
topography being analyzed with the model. The main difference in criteria when deciding to use a large versus small kernel is the distance found between the ridges themselves. If it is expected to be a large gap between ridgelines, a large kernel will help to identify only the major ridges and eliminate the smaller, negligible areas that may otherwise be identified as a ridge. However, in areas where ridgelines are close together, a smaller kernel size would be needed to identify ridges that are close to one another, such as one only 100 to 200 cells in length. In truth, it is unlikely that any one kernel length will accurately identify all ridges in a given region. A comprise may need to be made to optimize ridge identification and computer processing time.

It is also possible that a similar model could be developed using a non-LiDAR based elevation dataset so long as the resolution would support such an intensive analysis. Although it was not considered during the course of this study, it seems only logical that as DEM resolution decreases, the accuracy of the slope position predictions would follow suit.

**CONCLUSION**

The objective of this study was to develop a GIS based model to identify topographic ridges, and using that information, calculate the approximate slope position on a 0-100 scale (0 being the ridge) and then break the slope position into three broad categories – high, medium and low slope.

To best accomplish this, a progressive scanning technique in ArcMap 9.2 was developed using Focal Statistics and kernels creating a neighborhood specific to this
study. Ridges were identified using a combination of methods, but most heavily relied on the elevation of the surrounding topography.

Sixteen volunteers independently classified 245 study sites as residing on either a high, medium or low slope. Once their classifications were compiled, those study sites that had a high level of agreement among classifications (at least 9 out of 16 volunteers had to agree on a site’s classification) were used to confirm the results of the study. The results from the GIS model followed an almost identical trend to that of the volunteer classifications: the higher user agreement led to higher agreement with the GIS derived values.

In total, the authors of this study feel as though the GIS model described here identifies slope position and landform type as well as, if not slightly better, than what can be accomplished through visual identification. The GIS model has several advantages over conventional landform identification in that it requires no field work, can approximate landform type over entire regions (as opposed to single points) and is highly customizable. Furthermore, it eliminates the risk of error associated determining slope position in the field in an area with dense vegetation. Conversely, there are certain considerations to be made about the GIS model. Primarily, the accuracy of the elevation model as well as the model sensitivity to ridge identification has potential to influence the final results given with this GIS model.


CHAPTER IV

A PREDICTIVE GIS MODEL FOR DETERMINING THE PROBABILITY OF MOUNTAIN LAUREL OCCURRENCE IN THE SOUTHERN APPALACHIAN MOUNTAINS

ABSTRACT

The presence of mountain laurel (*Kalmia latifolia*) is often associated with diminished seedling establishment in hardwood forests and its function as a vertical fuel in forest fires. In this study, a GIS model was developed to predict the presence or absence of mountain laurel over a 33.7 Km² study area using 245 study sites with field observations for the basis of the analysis and confirmation of the results. Within the GIS, numerous raster layers were created to supplement the field collected data. Of these layers, four were found to be statistically significant in predicting the presence of mountain laurel using stepwise discriminant analysis. These included the canopy structure, aspect, canopy height and elevation, all of which were created using LiDAR derived models. Cross validation showed that the statistical model correctly predicted 81% of mountain laurel occurrence among the 245 study sites. A linear discriminant function was applied in the GIS to create both a binary presence/absence map and a posterior probability map. Results of the binomial and probability maps agree with those found in the statistical analysis. Accurate prediction of mountain laurel presence may significantly influence hardwood regeneration management as well as forest fire mitigation.
INTRODUCTION

Mountain laurel (*Kalmia latifolia*) extends in range from New England to the Gulf coast (Kurmes 1967). Other reports indicate that mountain laurel extends as far north as Ontario and New Brunswick (Kurmes 1967; Munns 1938; Preston 1961). One region in which mountain laurel has received considerable attention is the southern Appalachian Mountains. Numerous studies have attempted to quantify the importance of mountain laurel in both vegetation dynamics and its response to fire.

One of the earlier studies to consider the affects of mountain laurel is Day and Monk (1974). In their study, focusing on the Coweeta Basin in Franklin, NC, it was found that 58.5% of all stems ≥ 2.5cm dbh were either mountain laurel or rhododendron (*Rhododendron maximum*), and both were considered to be the most important understory species. They also noted that certain patterns emerged for both understory species, such as rhododendron primarily occurring low on NE-facing slopes and mountain laurel appearing largely on or near ridges (Day and Monk 1974).

Other studies agree that the presence of mountain laurel significantly affects the community structure found in hardwood forests (Mcgee and Smith 1967; Monk et al. 1985). Several studies, however, report that mountain laurel presence does not affect survivorship or establishment of hardwood species. Waterman et al. (1995) concluded that the removal of mountain laurel from study plots was not a significant factor in pitch pine communities. Their results confirmed those found by Kittredge and Ashton (1990), where it was determined that the density of mountain laurel had no affect on hardwood regeneration.
Waterman et al. (1995) did note, however, that although the density of mountain laurel did not have a measurable affect on hardwood regeneration, mountain laurel litter may. It has been noted numerous times in the literature that other laurel species, perhaps through allelopathy, limit the germination and growth of competing vegetation (Damman 1971; Mallik 1987; Zhu and Mallik 1994). Were mountain laurel litter to play an important role in hardwood regeneration, it would only add to the importance of fire and prescribed burning in the ecosystem.

The use of low intensity prescribed fires has been experiencing a surge in popularity as an understanding of the relationship between fire and oak regeneration improves (Brose et al. 2001). The objective behind using low intensity prescribed fires is to remove thin barked mid-story shrubs as well as consume the litter layer and is believed to encourage grasses, herbaceous vegetation and oak regeneration (Brose et al. 2001).

Several studies have investigated the affect of prescribed burning on the density of mountain laurel thickets and the subsequent diversity of regeneration. Ducey et al. (1996) determined that after low intensity fires in a mountain laurel dominated landscape, most species responded with an increase in density and frequency. They did note, however, that the diversity of species on moderately burned sites was lower than expected, perhaps due to the rapid regeneration of mountain laurel compared to other species. Similar results were recorded by Elliott et al. (1999).

In a study comparing fire intensity levels necessary for table mountain pine (Pinus pungens) stand replacement, the density of mountain laurel was a significant factor in determining different fire intensity levels (Waldrop and Brose 1999). Furthermore, in
areas with mountain laurel present, fire intensity is related to the mountain laurel density because the shrub serves as a vertical fuel. Waldrop and Brose (1999) found that low intensity fires had 30% or less mountain laurel cover, while medium and high intensity fires had 41 and 86% cover, respectively.

One of the central themes that associate both hardwood regeneration and forest fire intensity is the presence of mountain laurel. In this study, a statistical model was created that accurately predicts the presence or absence of mountain laurel based on a number of environmental factors. To accomplish this task, high resolution remote sensing data – Light Detection and Ranging – were used to develop an elevation model, canopy model, and vertical structure model within a Geographic Information System (GIS). This data was then exported for use in the statistical model.

A GIS is a “computer based system to aid in the collection, maintenance, storage, analysis, output, and distribution of spatial data and information” (Bolstad 2003). A GIS is a method of combining ordinary data with their spatial location (i.e. latitudinal and longitudinal coordinates). GIS analyses can range from simple explanatory maps to an intricate, model-driven product. In this study, ArcGIS 9.2 (ESRI 2006) is used exclusively for all GIS analysis.

Light Detection and Ranging (LiDAR) is an advanced form of remote sensing. Originally, desired output derived from LiDAR data was simply high resolution elevation models. With advances in technology, not only are highly accurate elevation models a possibility, but also canopy models (Lefsky et al. 1999; Lovell et al. 2003; Lovell et al.
LiDAR data are collected by a laser mounted to the underside of a plane. Thousands of light pulses are emitted from the laser and as the pulses are reflected from the earth’s surface, the return pulses are received by the LiDAR transmitter. Knowing the altitude and location of the plane and the time it took for the light pulse to travel from the plane and return to it, it is possible to calculate the elevation of the reflective surface with high accuracy (Dubayah and Drake 2000; Lefsky et al. 2002). One of the characteristics of LiDAR that makes it so attractive to ecological studies is the fact that the laser pulses emitted by the plane are reflected off all surfaces which they strike – vegetation, bare earth, all man-made structures and others. Because of this diversity of data, it is possible to create detailed canopy models for a region of concern (Lefsky et al. 1999; Lovell et al. 2003; Lovell et al. 2005).

The objective of this study was to develop a GIS model to accurately predict the occurrence of mountain laurel in the southern Appalachian Mountains.

**METHODOLOGY**

The project area is a 33.7 Km$^2$ (13 mi$^2$) section of North Carolina, including portions of Clay and Macon counties (Figure 4.1). As a whole, the topography for this section of North Carolina is considered mountainous. The Appalachian Mountains are the dominant topographical feature for this region.
Figure 4.1: Study sites were located in the Southwestern corner of Macon County, NC, adjacent to Clay County and part of the Nantahala National Forest.
Two-hundred forty five study plots were randomly located throughout the study region, and at each plot a multitude of topographical data was gathered including Terrain Shape Index (McNab 1989), Landform Index (McNab 1993) and a complete vegetation survey, paying particular attention to ericaceous shrubs such as mountain laurel and rhododendron. Used as part of a forest fuels study, complete information on the design and designation of study plots may be found in Waldrop et al. (2007).

LiDAR data for this study was acquired between April and December 2003 by EarthData International of North Carolina for the Floodplain Mapping Program, part of the North Carolina Division of Emergency Management. The horizontal datum is NAD83 North Carolina State Plane Feet and the vertical datum NAVD88 US Survey Feet.

All GIS analysis was completed in ArcGIS 9.2 (ESRI 2006). Kriging, a form of geostatistical interpolation, was used to generate the surface model and canopy model incorporated in this project. Kriging is the recommended method for interpolating LiDAR data (Lloyd and Atkinson 2006). All parameters necessary to perform Kriging were automatically calculated in ArcGIS.

LiDAR data points were initially classified by EarthData International prior to delivery to the North Carolina Floodplain Mapping Program. Bare earth points, as classified by EarthData International, were the source of points for the elevation model. Canopy points used for interpolation were all of the remaining points not classified as bare earth. Because the study area is located in a very mountainous region, no man-made structures appeared in the LiDAR.
An initial canopy model was created from Kriging using the vegetation height values found in the LiDAR data. These values would be the height of the tree (or other vegetation) plus that of the elevation itself. Once the first canopy model was created, it was subtracted from the surface model to reveal the actual canopy heights without the influence of the underlying topography.

In order to create a map layer representing the structure of the forest, the vegetation points were separated into five classes (Table 4.1). The classes were designated in order to best capture the height ranges in which vegetation existed.

<table>
<thead>
<tr>
<th>Class</th>
<th>Height Range</th>
<th>Vegetation Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>0-10ft (0-3m)</td>
<td>Understory</td>
</tr>
<tr>
<td>2</td>
<td>10-30ft (3-9m)</td>
<td>Low-Midstory</td>
</tr>
<tr>
<td>3</td>
<td>30-50ft (9m-15m)</td>
<td>Midstory</td>
</tr>
<tr>
<td>4</td>
<td>50-80ft (15-25m)</td>
<td>Mid-Upper Canopy</td>
</tr>
<tr>
<td>5</td>
<td>Greater than 80ft (25m)</td>
<td>Upper Canopy</td>
</tr>
</tbody>
</table>

With the vegetation points separated into their respective classes, each was converted to a raster, reclassified into a single value, and “layered” on top of one another. Rasters for this portion of the project were 25 ft (8 m) in resolution. This resolution was chosen to ensure that some overlap between cells occurred, but not in all cases. Using a smaller cell size would have resulted in almost no overlap, and therefore no structural information. Too large of a cell size would have resulted in the majority of cells overlapping and thus creating what would appear to be uniform vertical structure throughout the study area. The resolution chosen was thought to be a good comprise between the two extremes.
Summing the vegetation classes present at each of the study plots allowed for the creation of 15 different structural classes among the study plots. For example, in Figure 4.2, Line a, vegetation data exists in height classes 1 and 3 as they are described in Table 4.1. This would be given a structural class of 4. A full description of the structural classes may be found in Appendix B. Other data generated to help predict the presence of mountain laurel included flow direction, flow accumulation, aspect and slope. Because aspect is directional, it was broken into two categorical groups dividing the aspect into Northeast\west (315° – 135°) and Southeast\west (135°- 315°) segments. Terrain Shape Index, slope position, maximum canopy heights, distance from the nearest stream and elevation were also included in the statistical model.

![Figure 4.2: Profile view of LiDAR data. Black dots represent individual LiDAR points broken into 5 vegetation height classes. Blue vertical lines intersect some of the points. This intersection made it possible to break the height classes into structural classes based on where LiDAR points were present in the vertical column. For example, line a has vegetation points in classes 1 and 3, line b has vegetation present in class 1, 2, 4 and 5 and line c has vegetation present in class 1, 3 and 5. In total, there were 15 different combinations of vertical structure present at the study site.](image)

SAS 9.13 (SAS Institute 2008) was used for initial statistical analysis. Stepwise discriminant analysis (PROC STEPDISC) was used to model the percent mountain laurel coverage to the environmental variables. With significant variables identified, discriminant analysis (PROC DISCRIM) was performed with cross validation to assess
the accuracy of the model. Maple 11 (Maplesoft 2008) was then used to determine the linear discriminant function, and, using this function, a probability and binary presence\absence map was calculated in ArcGIS.

**RESULTS**

Of the 245 study sites included in this study, 62 (25%) had mountain laurel presence. Four environmental variables were identified as significant parameters in the identification of mountain laurel presence using stepwise discriminant analysis. These variables include vertical structure, aspect, canopy height and elevation. The remaining variables were not significant to the model. Average values at each study plot for the four variables included in the statistical model may be found in Table 4.2.

**Table 4.2: Average values for the four variables included in the statistical model on sites with and without mountain laurel present.**

<table>
<thead>
<tr>
<th>Variable</th>
<th>Average Value (ML Present)</th>
<th>Average Value (ML Not Present)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Structure</td>
<td>9</td>
<td>11</td>
</tr>
<tr>
<td>Aspect</td>
<td>.71</td>
<td>0.40</td>
</tr>
<tr>
<td>Canopy Height</td>
<td>31.1</td>
<td>53.8</td>
</tr>
<tr>
<td>Elevation</td>
<td>3817</td>
<td>3992</td>
</tr>
</tbody>
</table>

Cross validation performed during discriminant analysis achieved an overall accuracy of 77%. Individually, 138 (75%) of 183 sites with no mountain laurel present were correctly classified, and 45 (25%) of 183 were incorrectly classified as having mountain laurel present when there was not. Conversely, 50 (81%) of 62 sites with mountain laurel present were correctly classified using this model, and 12 (19%) were
incorrectly classified as not having mountain laurel present when in fact it was. These values are summarized in Table 4.3.

Table 4.3: Cross validation table generated in SAS 9.13

<table>
<thead>
<tr>
<th>Field Observations</th>
<th>ML Not Present</th>
<th>ML Present</th>
<th>Total</th>
</tr>
</thead>
<tbody>
<tr>
<td>ML Not Present</td>
<td>138 (75%)</td>
<td>45 (25%)</td>
<td>183</td>
</tr>
<tr>
<td>ML Present</td>
<td>12 (19%)</td>
<td>50 (81%)</td>
<td>62</td>
</tr>
<tr>
<td>Total</td>
<td>150</td>
<td>95</td>
<td>245</td>
</tr>
</tbody>
</table>

With the discriminant model in place, a linear discriminant function was developed using Maple 11 (Maplesoft 2008), resulting in both a probability map and binary presence\absence map. The linear discriminant function is:

\[
P(ML\text{presence}) = \frac{e^{-\frac{1}{2}d_i}}{e^{-\frac{1}{2}d_i} + e^{-\frac{1}{2}d_0}}
\]

where

\[
d_i = (x - \mu)^T \Sigma^{-1} (x - \mu).
\]

And where \(x\) = the environmental variable (such as aspect), \(\mu\) = the average value for the environmental variable and \(\Sigma^{-1}\) the inverse covariance matrix (Johnson 1998).

Results from the discriminant function applied in the GIS model are very similar to those obtained with cross validation in SAS. Of the 62 sites with mountain laurel present, 51 (82%) were correctly classified in the binary presence\absence map (Figure 4.3) and 51 (82%) were correctly classified according to the probability function (where greater than 50% indicates presence), seen in Figure 4.4. Conversely, 46 (25%) sites, which did not have mountain laurel present, were incorrectly classified in the binary
presence\absence model. Lastly, 46 sites (25%) were incorrectly classified as having mountain laurel present using the probability model when in reality mountain laurel is not present at the site. The overall accuracy of the GIS based model was 77%.

Of the eleven study sites where mountain laurel was present but not accurately predicted (according to the presence\absence probability maps), all but two were located less than 10 ft (3 m) from at least one cell identified as possibly having mountain laurel. The average distance between the misclassified study sites and cells identified as mountain laurel was 9 ft (3 m), with a maximum distance of 50 ft (15 m).

**DISCUSSION**

Of the four variables identified as significant in the statistical model, the most unique is canopy structure. Numerically, the lower the structure classification, the simpler the forest structure is at that cell. For example, a structure class of one implies that vegetation is only present from 0-10 ft (0-3 m) above the earth’s surface. As the structure classification increases, so does the complexity of the forest structure. The highest structure class implies that there is vegetation present throughout each of the five vegetation classes described in Table 4.1 (Figure 4.5).

Of the fifteen structure classes present, the average structure class in which mountain laurel was present was class nine, and the average class in which mountain laurel was not present was class eleven. Class nine was typically comprised of vegetation relatively evenly dispersed among the height classes, but on average only two to three height classes were represented at sites with mountain laurel present. Class
Figure 4.3: Map of mountain laurel presence\absence as predicted using discriminant analysis.
eleven was comprised of vegetation that also appears to be evenly dispersed throughout
the five height classes, however sites without mountain laurel present on average had
vegetation in three to four of the height classes. A complete description of structural
classes may be found in Appendix A.
Numerous theories may be posed as to why mountain laurel preferred structural class nine. An important observation to make is that the average canopy height on sites where mountain laurel was present was 31ft. This seems to support, at least to some degree, the tendency for mountain laurel to be present on sites with only two to three
Figure 4.5: Map of vertical structure complexity (bottom right). Darker colors represent areas with higher amounts of vertical structure. Visual similarities between the vertical structure and aerial imagery (top left) are evident.
height classes present. Alternatively, the average canopy height on plots without mountain laurel present was 53 ft, also supporting the tendency for sites without mountain laurel to have more height classes present. It should be noted, however, that as canopy density increases, LiDAR penetration decreases (Barber and Shortridge 2005; Cowen et al. 2000; Jensen et al. 1987). Cowen et al. (2000) estimated when the canopy is 80-90 percent closed, only 10-40 percent of the light pulses reach the ground and are reflected back to the LiDAR sensor. Some caution should be exhibited when interpreting the potential relationship between vertical structure and mountain laurel presence, as the structure classes presented here are wholly dependent upon the presence of LiDAR data in each of the classes.

Other results show differences in the average values among plots with and without mountain laurel. Considering the aspect was converted to a binary variable (Northeast\west aspects equaling zero and Southeast\west aspects, where mountain laurel is expected to be found, equaling one), on average, plots with mountain laurel occurred on southeast\westerly aspects 71% of the time whereas plots without mountain laurel occurred on southeast\westerly aspects only 40% of the time. This is in agreement with numerous other studies that cite mountain laurel’s preference for southerly aspects (Day and Monk 1974; Waldrop et al. 2007). Additionally, mountain laurel preferred sites, on average, 175 ft (53 m) lower in elevation than sites without mountain laurel.

Other interesting results from this study are how well the statistical model, based on the field plots, translated into a prediction raster in ArcGIS. Results from the cross validation performed in SAS were nearly identical to the results attained when the
discriminant function was applied to the entire landscape. While it is logical to expect the study sites to have similar results in the GIS model as they did in SAS, the visual patterns displayed in Figures 4.3 and 4.4 were unexpected.

The resemblance between the structural classes and the predicted presence of mountain laurel, however, cannot go unnoticed (Figure 4.6). It is possible that the true relationship between the vertical structure classes and mountain laurel presence may not be fully realized in this study. Based on these results, it would be prudent to further investigate the potential relationship between the two variables, and perhaps even improve upon the predictions described here.

It is also necessary to consider the scale of this study. When comparing the GIS model results to that of field observations, it is crucial that more than just the single spatial location of a study site be considered. In this study, 12 plots were known to have mountain laurel present but were not identified as such in the statistical model. However, it should be taken into consideration that 10 of the 12 incorrectly identified sites were all within 10 ft (3 m) of an area likely to have mountain laurel present. One might consider then, that although not all of the study sites matched the statistical results perfectly, that, when applied in a real world setting, it may be possible to attain results better than described here, should the strictest of interpretations not be taken. It should also be noted that the study sites were located in the field using a GPS (Global
Figure 4.6: A comparison between vertical structure and mountain laurel probability.
Positioning System) device which typically has an error rate of 3-30 m; post-differential correction may improve this to within a few meters (Bolstad 2003), however the average distance between misclassified points and regions with a high probability of mountain laurel presence is still within the acceptable range of GPS error.

The ability to predict mountain laurel presence is one that would be of great service to not only forest managers with the goal of stand regeneration, but also to fire ecologists attempting to better understand wildfire dynamics. One potential extension of research similar to this would be to predict the actual percent cover of mountain laurel over a given landscape. Knowing that mountain laurel often serves as vertical fuel in forest fires, it may be possible to predict, at least to some degree, where a low intensity fire may transition to a medium or high intensity fire based upon predicted density. Waldrop and Brose (1999) state that low intensity fires had 30% or less percent cover of mountain laurel, while medium and higher intensity fires had 40 to 80% mountain laurel coverage, respectively. Not only would statistical models that predict the percent cover of mountain laurel be useful during a forest fire, but also during early stages of prevention or mitigation.

CONCLUSION

In this study, a LiDAR based GIS model was used to statistically predict the presence or absence of mountain laurel in the southern Appalachian Mountains. Data from 245 study plots were collected, and data ranged from basic topographical measures (such as TSI and LFI) to intensive vegetation surveys.
Once all of the data was compiled, multiple GIS based map layers were created using high resolution LiDAR data. Data included an elevation model, canopy model and a model of vertical forest structure. A sample of all plot and GIS data was used as the input for discriminant analysis, and four environmental variables were identified as statistically significant to the presence of mountain laurel: elevation, canopy height, aspect and vertical structure. A linear discriminant function was produced, and GIS based presence\absence and probability models were generated.

It was found that there was over 80% agreement on sites known to have mountain laurel present and those that were predicted both in SAS and in the GIS model. One of the more interesting results of this study is the potential association between mountain laurel presence and vertical forest structure given the stark visual similarities between the two.

The potential uses for accurate predictions of mountain laurel presence are broad. The two most obvious applications of such results deal with hardwood regeneration and wildfire mitigation and prevention. In both cases, the removal or management of mountain laurel in high risk areas may aid hardwood regeneration and perhaps lower the risk of mountain laurel being used as vertical fuel, transitioning low intensity fires to more severe, high intensity fires.
LITERATURE CITED


Munns, E.N. (1938). The distribution of important forest trees in the United States. In U.S.D.o. Agriculture (Ed.) (p. 176)


CHAPTER V

CONCLUSION

The underlying theme behind the studies presented here was the continued and increasing ability to more accurately describe and model the topographic landscape of the southern Appalachian Mountains. In the preceding pages, it was shown how GIS modeling may be used to calculate the Terrain Shape Index (TSI), slope position, and probability of mountain laurel presence across a landscape.

In each of the studies, an attempt was made to develop not only accurate GIS models, but also ones that are easily reproduced and transferable among users and locations. In Chapter II, a GIS based model was developed to calculate TSI. TSI, first introduced by McNab (1989), quantified the degree of convexity or concavity for a given study plot. Numerous studies cite using TSI as an indicator of site quality and as a measure for projected vegetative growth (Abella and Covington 2006; Humphries et al. 2008; Thompson et al. 2006). Limitations do exist, however. Field based TSI observations are limited to a fixed number of sites simply because the amount of time it takes to collect field data is directly related to how many sites are included in the study. Additionally, only information regarding the study site is available; a classification of the surrounding landform is not included unless other study sites are located nearby.

Given these limitations, it is clear where the GIS model presented in Chapter II may have certain advantages over conventional measurement practices. In that study, a straightforward GIS model is described that calculated TSI not only for individual study sites, but also for the entire landscape. Using LiDAR (Light Detection and Ranging) data,
high resolution elevation models may be generated and TSI values calculated. Following the procedure laid forth, nearly 3 out of every 4 study plots were correctly classified as either convex or concave when compared to field based observations. It is possible that any remaining disagreement between the field based observations and GIS based calculations may be a result of measurement error in both the field based TSI values and those derived from within the GIS model.

Nonetheless, the ability to calculate TSI values in a GIS offers many advantages. Primarily, it provides a simple and efficient method of generating values across an entire landscape. Furthermore, the model described here is simple and uses functions available to anyone with basic GIS knowledge and access to the ArcGIS Spatial Analyst extension.

The second study described in this paper details the methodology behind using a GIS model to determine the slope position (ridge, side slope and cove) at a given location. Most conventional methods used when determining slope position in the field involve either simple visual estimation, or use of McNab’s (1993) Landform Index methodology. In this study, a progressive scanning technique was employed that, simply put, looked for local maximums in elevation in order to identify ridges. Combined with the location of known streams and rivers, it became possible to measure the distance between the ridges and the bottom of the slope, thus calculating the percent slope position across the landscape.

The results of this study were nearly equivalent to those obtained when slope position is classified manually using a digital elevation model. The potential uses of this GIS model are multifold; not only appear to accurately classify slope position across the
landscape, but it also provides an estimated slope position on a scale of 0-100, with 0 being the ridge top. Conventional field based methods visually estimate the scalar slope position, which at times may be obstructed by vegetation or other topographic features making the true slope position difficult to gauge. The GIS model presented here removes the affect of vegetation and other sources of error.

The third and final study presented here moves away from typical topographical modeling and focuses on the development of a GIS model to predict the presence or absence of mountain laurel, an ericaceous shrub known to serve as a vertical fuel in forest fires (Waldrop et al. 2007) and also potentially limit hardwood regeneration in areas where mountain laurel growth is dense (Damman 1971).

The objective of this study was to accurately predict the presence or absence of mountain laurel in the southern Appalachian Mountains. To do this, a number of raster based GIS layers were created, including flow direction, accumulation, slope, aspect, a canopy height profile and a layer describing the vertical structure of the region. This study made significant use of LiDAR based data, and it was found that several of the LiDAR derived models (canopy height, vertical structure, aspect and elevation) were significant to the presence or absence of mountain laurel. With this information available and combined with data collected in the field, a nonlinear discriminant function was developed and applied in the GIS model. It determined that just over 80% of the known locations of mountain laurel were accurately predicted using the discriminant analysis.

Significant results as those seen in this study have the potential to greatly alter future management practices for both hardwood regeneration and fire
prevention/mitigation. Data on the probability of mountain laurel occurrence, as developed in this study, may provide managers concerned with mountain laurel presence with the information they need to focus their efforts on key geographical areas likely to harbor mountain laurel. Whether it is to remove leaf litter (thought to inhibit seeding establishment) or the mountain laurel entirely, there is potential for significant amounts of time, capital and effort to be saved by focusing on areas with a high likelihood of mountain laurel occurrence.

As a whole, this paper presents several advances in topographic and predictive modeling in a GIS. It has been shown that a GIS model, properly configured, can produce TSI values similar to those collected using conventional field based methods, identify slope positions to the same accuracy that can be achieved by manually classifying a location on an elevation model, and also predict the presence or absence of mountain laurel across a landscape. Perhaps the most alluring attribute of each of the models is that they can each be readily transferred and used by other parties. No complicated programming was used; in fact, only readily available analysis tools in ArcGIS were employed throughout each study. Lastly, each has the potential to be used in real world applications, potentially saving time, money and effort for those involved.
LITERATURE CITED


APPENDICES
Appendix A

GIS Model for the Determination of Slope Position and Ridge Identification

Figure A-1: ModelBuilder Model of the processes involved in ridge identification and slope position calculation.
Appendix B

Height Class Combinations Used to Define Structure Classes

<table>
<thead>
<tr>
<th>Height Class 1</th>
<th>Height Class 2</th>
<th>Height Class 3</th>
<th>Height Class 4</th>
<th>Height Class 5</th>
<th>Structure Class</th>
</tr>
</thead>
<tbody>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>0</td>
</tr>
<tr>
<td>X</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>1</td>
</tr>
<tr>
<td>-</td>
<td>X</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>2</td>
</tr>
<tr>
<td>X</td>
<td>X</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>3</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>X</td>
<td>-</td>
<td>-</td>
<td>3</td>
</tr>
<tr>
<td>X</td>
<td>-</td>
<td>X</td>
<td>-</td>
<td>-</td>
<td>4</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
<td>X</td>
<td>-</td>
<td>4</td>
</tr>
<tr>
<td>-</td>
<td>X</td>
<td>X</td>
<td>-</td>
<td>-</td>
<td>5</td>
</tr>
<tr>
<td>X</td>
<td>-</td>
<td>-</td>
<td>X</td>
<td>-</td>
<td>5</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>X</td>
<td>5</td>
</tr>
<tr>
<td>X</td>
<td>X</td>
<td>X</td>
<td>-</td>
<td>-</td>
<td>6</td>
</tr>
<tr>
<td>-</td>
<td>X</td>
<td>-</td>
<td>X</td>
<td>-</td>
<td>6</td>
</tr>
<tr>
<td>X</td>
<td>-</td>
<td>-</td>
<td>-</td>
<td>X</td>
<td>6</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>X</td>
<td>X</td>
<td>-</td>
<td>7</td>
</tr>
<tr>
<td>X</td>
<td>X</td>
<td>-</td>
<td>X</td>
<td>-</td>
<td>7</td>
</tr>
<tr>
<td>-</td>
<td>X</td>
<td>-</td>
<td>-</td>
<td>X</td>
<td>7</td>
</tr>
<tr>
<td>X</td>
<td>-</td>
<td>X</td>
<td>X</td>
<td>-</td>
<td>8</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>X</td>
<td>-</td>
<td>X</td>
<td>8</td>
</tr>
<tr>
<td>X</td>
<td>X</td>
<td>-</td>
<td>-</td>
<td>X</td>
<td>8</td>
</tr>
<tr>
<td>-</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>-</td>
<td>9</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>-</td>
<td>X</td>
<td>X</td>
<td>9</td>
</tr>
<tr>
<td>X</td>
<td>-</td>
<td>X</td>
<td>-</td>
<td>X</td>
<td>9</td>
</tr>
<tr>
<td>X</td>
<td>-</td>
<td>-</td>
<td>X</td>
<td>X</td>
<td>10</td>
</tr>
<tr>
<td>X</td>
<td>X</td>
<td>X</td>
<td>-</td>
<td>X</td>
<td>10</td>
</tr>
<tr>
<td>-</td>
<td>X</td>
<td>X</td>
<td>-</td>
<td>X</td>
<td>10</td>
</tr>
<tr>
<td>-</td>
<td>X</td>
<td>-</td>
<td>X</td>
<td>X</td>
<td>11</td>
</tr>
<tr>
<td>X</td>
<td>X</td>
<td>X</td>
<td>-</td>
<td>X</td>
<td>11</td>
</tr>
<tr>
<td>-</td>
<td>-</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>12</td>
</tr>
<tr>
<td>X</td>
<td>X</td>
<td>-</td>
<td>X</td>
<td>X</td>
<td>12</td>
</tr>
<tr>
<td>X</td>
<td>-</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>13</td>
</tr>
<tr>
<td>-</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>14</td>
</tr>
<tr>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>X</td>
<td>15</td>
</tr>
</tbody>
</table>

Table B-1: Structure classes were defined by summing the presence of vegetation within each height class. For example, structure class 3 implies vegetation was present in height classes 1 and 2 (1 + 2 = 3) but no others.