10-11-2016

Using LiDAR to Determine Early Successional and Shrubland Structure in Eastern Connecticut

Jacob P. Conshick

University of Connecticut - Storrs, jacob.conshick@uconn.edu

Recommended Citation
http://digitalcommons.uconn.edu/gs_theses/1006

This work is brought to you for free and open access by the University of Connecticut Graduate School at DigitalCommons@UConn. It has been accepted for inclusion in Master's Theses by an authorized administrator of DigitalCommons@UConn. For more information, please contact digitalcommons@uconn.edu.
Using LiDAR to Determine Early Successional and Shrubland Structure in Eastern Connecticut

Jacob Conshick

B.S., University of Connecticut, 2013

A Thesis
Submitted in Partial Fulfillment of the
Requirements for the Degree of
Master of Science
At the
University of Connecticut
2016
Masters of Science Thesis
Using LiDAR to Determine Early Successional and Shrubland Structure in Eastern Connecticut

Presented by

Jacob Conshick, B.S.

Major Advisor

Dr. Daniel Civco

Associate Advisor

Dr. Chadwick Rittenhouse

Associate Advisor

Dr. Howard Kilpatrick

Associate Advisor

Dr. Thomas Worthley

University of Connecticut

2016
ACKNOWLEDGMENTS

I would like to thank all the people who are part of my life who have made the completion of this thesis and my MS Degree possible. I would like to thank my advisor Dr. Daniel Civco, for helping me and providing the guidance necessary for completing my thesis. It was a pleasure and a privilege to learn from someone with so much experience in remote sensing, along with support he provided, and his technical expertise which he used to help me grow. Dr. Civco had the confidence in me allowing me to work independently and falter at time, but would always be there to provide the necessary help when needed. For allowing me to work through my problems but never leaving my side, I thank him; it truly was an honor to work with him.

I would also like to thank all of my committee members for all of their support and for all that they have contributed in helping me write my thesis: Dr. Chadwick Rittenhouse for helping me understand classification of land cover in a unique way and for helping me set up a plan of action; Dr. Howard Kilpatrick for providing me with the opportunity to work on this conservation project with the Department of Energy and Environmental Protection (DEEP) and for providing me information on various topics of interest that helped me understand my research even better; and Dr. Thomas Worthley for his help in the field, whereby giving me a better understanding of early successional make-up. I found the field work very valuable in helping me understand what I was classifying in the imagery.

I would also like to thank my family – my parents, grandmother, aunts, uncles, and cousins for their love and support. Along with my advisors and committee members I could not have finished my thesis without them. Without the unwavering confidence and support for me, I could never have reached this point. Thank you all very much.
I would like to give a special thanks to some special friends in my life that have lent me their support not only through my Master’s but in years past and in the years to come. Thank You; Thomas Scanlon, Marco Balassone, Caleb Rose, Sean Cochran, Coast Guard Ensign Jacquelyn Kubicko, Coast Guard 1/c Angela Johnson, and Coast Guard 1/c Phong Vu. Thank you all for everything that you have done for me over the years. Your support is one of the many reasons for my success.

I would like finally to thank the professors and staff of the Natural Resources and the Environment Department with whom I have worked. Over the years you have all provided me with both professional and personal support, and for that I am grateful. It was a pleasure to work with all of you. The past several years have been some of the best of my life. Thank you again for the opportunity to work and learn from you, you have allowed me to grow not only as a scholar but as a person.
# Table of Contents

Acknowledgements .................................................................................................................. iv

Table of Contents ...................................................................................................................... vi

List of Figures ........................................................................................................................ x

Chapter 1: Introduction .............................................................................................................. 1
  1.1 The Problem .................................................................................................................... 1
  1.2 Objectives ....................................................................................................................... 2

Chapter 2: Literature Review .................................................................................................. 3
  2.1 Early Successional Structure ......................................................................................... 3
  2.2 Measuring Early Successional Structure ....................................................................... 5
  2.3 Terrain Ruggedness Index ............................................................................................. 7

Chapter 3: Data and Methodology .......................................................................................... 10
  3.1 Study Area ..................................................................................................................... 12
  3.2 Software Used ............................................................................................................... 12
  3.3 Data ............................................................................................................................... 13
  3.4 Processing Raw LiDAR Files and Initial Land Cover ..................................................... 14
  3.5 Creating Terrain Roughness Index ................................................................................ 17
  3.6 Creating Land Cover Mosaic ........................................................................................ 19
  3.7 Creating the Structural Land Cover Raster .................................................................... 21
3.8 Testing the Accuracy ......................................................................................24

Chapter 4: Results ..........................................................................................28

4.1 Final Structure Land Cover Raster ..............................................................28

4.2 Trial Accuracy Results .............................................................................31

4.2.1 Trial Windows Standard Calculation Results ....................................34

4.2.2 Trial Windows Area Adjusted Calculation Results .........................36

4.3 Final Accuracy Assessment (3-Classes) ..................................................38

4.4 Final Accuracy Assessment (4-Classes) ..................................................39

Chapter 5: Discussion ......................................................................................40

5.1 Relation to Other Studies ........................................................................40

5.2 Management Purposes ............................................................................41

5.3 Constraints and Limitations ....................................................................47

5.4 Further Research ......................................................................................48

References ........................................................................................................49

Appendix A: The Study Area ........................................................................52

Appendix B: LiDAR Coverage Area .................................................................53

Appendix C: Aerial Coverage Area .................................................................54

Appendix D: Aerial Zoom ................................................................................55

Appendix E: Compiled Maximum Height Digital Surface Model ...............56
Appendix F: Terrain Roughness Raster.................................................................57

Appendix G: 3x3 Window Accuracy Error Matrix .................................................58

Appendix H: 3x3 Window Accuracy Error Matrix (Continued) ............................59

Appendix I: 5x5 Window Accuracy Error Matrix..................................................60

Appendix J: 5x5 Window Accuracy Error Matrix (Continued) ............................61

Appendix K: 9x9 Window Accuracy Error Matrix ..................................................62

Appendix L: 9x9 Window Accuracy Error Matrix (Continued) ............................63

Appendix M: 15x15 Window Accuracy Error Matrix.............................................64

Appendix N: 15x15 Window Accuracy Error Matrix (Continued) ..........................65

Appendix O: Terrain Ruggedness Script .............................................................66

Appendix P: Early Successional Structure Script .................................................70

Appendix Q: Python Module – Land Cover Boxplot List .....................................78
List of Figures

Figure 1: Early Successional Structure from Bear Hill Wildlife Management Area Bozrah, CT ...4

Figure 2: Coordinate System for The Terrain Roughness Index .............................................9

Figure 3: Aerial Image of Early Successional Structure .........................................................11

Figure 4: Study Area of The 12 Towns ................................................................................12

Figure 5: 2012 Aerial Image Example of Columbia CT ................................................................13

Figure 6 Digital Surface Model Example of Columbia, CT .........................................................13

Figure 7: Land Cover Example of Columbia, CT .................................................................14

Figure 8: Lidar Point Cloud .......................................................................................................16

Figure 9: Model Creating a Mosaic from The Lidar Height Files .............................................17

Figure 10: Compiled Height Raster of Non-Ground Return Lidar ............................................18

Figure 11: Land Cover Map Used to Identify Urban Development .........................................20

Figure 12: Land Cover with The Urban Development Changed to No Data ..........................21

Figure 13: Terrain Roughness Raster with The Urban Development Removed ......................22

Figure 14: Box Plot Depicting the Land Cover Thresholds ..................................................23

Figure 15: Output from The Clump Analysis (Top) Showing The Various Groups and The Eliminate (Bottom) Showing The Removal of the Salt and Pepper Effect ........................................24

Figure 16: Map Depicting Structure Classes Based On A 3 Class ........................................29

Figure 17: Map Depicting the Structure Classes Based On A 4 Class Assessment ..................30
Figure 18: Graphical Representation of the Overall Standard Calculation Trend ..........................32

Figure 19: Graphical Representation of the Overall Area-Adjusted Calculation Trend  .............32

Figure 20: Site in Lebanon, CT Proposed for Field Analytic Derived from The Tri Land Cover Raster ..................................................................................................................................................................42

Figure 21: Another Site Located in Colchester, CT for Site Analytics ......................................43

Figure 22: Backside of Babcock Pond Wildlife Management Area Depicting Remoteness from Urban Development ..................................................................................................................................................................45

Figure 23: Site in East Hampton, CT What Would Benefit from Clear Cutting to Revert the Land Back to Early Stage Successional Structure .............................................................................................................46
Chapter 1: Introduction

1.1 The Problem

Early successional structure is described as assemblages of woody shrub-like plants which grow in close proximity to each other creating a dense structure mass. Wildlife species such as New England cottontail and woodcock use early successional structure for shelter, food, and protection. Early successional structure in the state of Connecticut has been declining in recent years, due to a decrease in disturbances that cause and promote the growth of early successional structure. In the 1800’s there was an increase in agricultural production throughout Connecticut’s landscape leaving about 25 percent of Connecticut’s forests intact by 1825 (Hochholzer, 2010). Deforestation in Connecticut was the disturbance that allowed early successional structure to expand.

It was not until the decline of agriculture, due to the opening of the Erie Canal and the Industrial Age, that allowed for the conversion of farmland back to forests (Ward, Worthley, Smallidge, & Bennett, 2013). It was during this period between the landscape not being truly agriculture nor forest that one could see the development of early successional structure (Smith, 2007). To help promote early structural growth, in the late 1800s to the early 1900s a demand for charcoal increased causing the new growth forests to be cut down again. It was through the demand for charcoal that allowed for the growth period of the early structural growth to be extended.

Species dependent on this type of structure, such as the New England cottontail (Sylvilagus transitionalis) and American woodcock (Scolopax minor) moved into Connecticut. These species populations have started to decline over recent years primarily due to a decline in the early succession structure. According to the Natural Resource Conservation Service, the
Northern bobwhite quail population has declined over the past 20 years by 65 percent mainly due to a loss of habitat (Smith, 2007).

The structure of early successional habitat can be quantified by metric derived from remote sensing data. Terrain roughness, as calculated with LiDAR data has the potential for efficiently identifying early successional structure. The terrain roughness algorithm derives the textural properties of the landscape cover which can be used to characterize early successional structure. Terrain roughness can help in the identification of potential earl successional structure thereby, decreasing the amount of time in the field by finding plausible locations via the remotely sensed data.

There are three different management categories that can be used to create, enhance, or manage early successional structure: (1) Forest Harvests, (2) Site Preparation, and (3) Improvement Practices (Smith, 2007). By creating a map showing the location of early successional structure and how it is distributed across the Connecticut landscape, managers can create plans that would best suit structure creation for a particular location.

1.2 Objectives

1. Create protocols for locating early successional and shrubland structure.

2. Assist wildlife managers in determining early successional and shrubland structure locations.

3. Assess the accuracy of identifying early successional and shrubland structure from remote sensing data.
Chapter 2: Literature Review

2.1 Early Successional Structure

Land cover classes such as Forest (both coniferous and deciduous), water (lakes, streams, rivers, etc.), and agricultural land can all be easily and clearly defined. “Early successional structure”, on the other hand, has several interpretations, based on vegetation types, animals, and disturbance (both manmade and natural). Each has a slightly different way of viewing and defining early successional structure.

Early successional structure can fall into an ecological category called secondary succession. Secondary succession occurs on areas that contain the remnants of plant communities after a disturbance event such as clear cutting a forest, fires, hurricane, and tornadoes (Bolen & Robinson, 2003). To add further to the definition of early successional composition, Ricklefs reiterates what Bolen and Robinson assert but adds another dimension in stating that the size of the disturbance determines the growth of the smaller plants. Small gaps in the canopy can be easily filled whereas the larger gaps take longer to fill in and allow for the secondary plant growth to happen (Ricklefs, 2008).

The concept of secondary succession proposed by Ricklefs describes the condition that is suitable for early stage plant life. Plant vegetation that makes up early successional structure need larger disturbances where they have ample access to the sun. Researchers have expanded this textbook definition to encompass different requirements between types of disturbances and stand types. King and Scott (2014) makes a distinction between what people term as early successional structure and a young forest structure, stating that early successional structures are areas of land that have pioneer (shade-intolerant) plants dominating (King & Schlossberg.
The key term is “shade intolerant”, since early successional structure happens where gaps in the canopy have occurred, and there is ample direct sunlight, forcing the plants to be tolerant of “no shade” situations.

An early successional structure ceases to exist when it turns into a young forest structure which can be a thin line of transition. A young forest structure is considered to be stands recovering from disturbances through the recruitment of canopy species from advanced regeneration (King & Schlossberg, 2014).

The variables that make up the early successional structure includes height of vegetation (the vertical profile), horizontal patchiness, density, diameter of stem, and finally the proportion of woody to herbaceous plants (King & Schlossberg, 2014). Understanding the successional structure of the stand helps in the proper identification of the land cover type of the stand. Figure 1 shows an example of early successional structure located in the Bear Hill Wildlife Management Area in Bozrah, CT.

Figure 1 Early successional structure from Bear Hill Wildlife Management Area Bozrah, CT
2.2 Measuring Early Successional Structure

Scientists have mapped early successional and shrubland structure using a variety of remote sensing techniques. A study “Characterizing Forest Succession with LiDAR Data: An Evaluation for the inland Northwest, USA” mapping forest succession classes using LiDAR in the Northwest United States. The study aimed to determine if LiDAR would be useful for predicting forest processes such as long-term carbon sequestration or creating accurate forest classifications to achieve forest management goals effectively (Falkowski, Evans, Martinuzzi, Gessler, & Hudak, 2009). The study looked at the textural properties that can be obtained through the use of LiDAR, for example a forest that is undergoing stand initiation has reflective LiDAR pulses at or near the ground level with a few pulses off seedlings and saplings (Falkowski, Evans, Martinuzzi, Gessler, & Hudak, 2009).

The LiDAR for their study was flown in the summer of 2003 for the entire study area. A digital surface model (DSM) was created by subtracting the “bare earth” elevation from the height layer. Using the above ground surface model, 34 predictive variables that are used in identifying forest structure were created. Using these predictive variables, a Random Forest algorithm was used to classify the forest successional structure (Falkowski, Evans, Martinuzzi, Gessler, & Hudak, 2009). The study determined five advantages over traditional forest classification: (1) Using the bootstrap approach can achieve higher classifications while addressing over-fitting problems at the same time, (2) It can create robust predictions based on bootstrap replicates, (3) It is nonparametric and unaffected by distributional assumptions, (4) The GENI static can integrate non-linear variable interactions, (5) Has a reliable internal estimate of classification (Falkowski, Evans, Martinuzzi, Gessler, & Hudak, 2009).
The study created two forest successional maps, the first had seven classes and the second had 6 classes with mature multistory (MMS) and closed stem exclusion (CSE) combined. They chose to combine the MMS and CSE due to the structural similarity between the two classes creating a problem in detecting and identify understory in a multi-story or closed canopy forest (Falkowski, Evans, Martinuzzi, Gessler, & Hudak, 2009). The study showed that the Random Forest algorithm had an overall accuracy of 90.1% for the seven-class successional classification. Confusion did occur between the MMS and CSE classes, with error rates at 27.0% and 36.0% respectfully. The overall accuracy for the six-class successional classification (which has the MMS and CSE aggregated together) was 95.5% with a further reduction in class error (Falkowski, Evans, Martinuzzi, Gessler, & Hudak, 2009). The study determined that using the Random Forest algorithm was an effective way to accurately classify forests which can be used to achieve multiple forest management goals.

Another way that early successional structure can be classified is through remote sensing techniques using Landsat digital image data. Rittenhouse (2014) estimated the amount of early successional structure in Connecticut. The goal of the study was to map the disturbance / regeneration of the forests and map the afforestation located in Connecticut. The imagery used in the study was from 30-m resolution Landsat Thematic Mapper (TM) and Enhanced Thematic Mapper Plus (ETM+). Using remote sensing techniques, such as satellite imagery, has the potential to identify mature forests and forest understory (which have similar characteristics to early successional structure) (Rittenhouse, 2014).

The Vegetation Change Tracking (VCT) algorithm, designed by Huang et. al. (2010), was used in 22 Landsat footprints to map forested regions. The down side to using the VCT is the low capability to distinguish between afforestation and forest regeneration (Rittenhouse,
To assist with classification, the VCT is used in conjunction with a Support Vector Machine (SVM) algorithm. The SVMs primary function is to classify the areas of afforestation that might not have been picked up by the VCT.

The overall accuracy using the Landsat imagery with the VCT and SVM was 76.0%. According to the literature the overall accuracy of 76.0% is comparable to the 2001 NLCD Region 10 at 78.0% (Rittenhouse, 2014). The afforestation class obtained a user’s accuracy range between 19.0% and 87.0%, which has an average higher than the 37.0% accuracy obtained from 2006 NLCD change for region 10 (Rittenhouse, 2014).

2.3 Terrain Ruggedness Index

Early successional structure exhibits particular textural properties that might be detectable with the appropriate remote sensing data. Nearest neighbor imputation methods are one way that textural properties of early successional structure can be viewed. Imputation methods have grown in favor for their ability to relate multiple attributes of interest to satellite data. One study looked at various imputative methods to predict timber volume in a stand (Hudak, Crookston, Evans, Hall, & Falkowski, 2007).

Nearest neighbor imputation is a form of nonparametric regression. The objective of using a nonparametric regression is to predict response variable irregularly across the landscape from predictor variables determined continuously through the landscape and finally partitioned into continuous pixels (Hudak, Crookston, Evans, Hall, & Falkowski, 2007). To test the various imputation methods, 165 field plots were geolocated, and the basal area and tree diameter were determined for each species of tree in the plot. The predictor variables (n= 60) were derived from airborne LiDAR during the summer. A digital surface model, or DSM, assumed to depict vegetation only, was created by removing ground surface elevation. To eliminate high
correlation between variables the random forests (RF) model (which can bootstrap data) was used to prune down the variables used in the analysis (Hudak, Crookston, Evans, Hall, & Falkowski, 2007).

Using the refined variables eight imputation methods were modeled to determine their effectiveness; (1) Euclidean Distance (EUC), (2) RAW, (3) Mahalanobis Distance (MAL), (4) Independent Component Analysis (ICA), (5) Most Similar Neighbor (MSN), (6) Most Similar Neighbor 2 (MSN2), (7) Gradient Nearest Neighbor (GNN), and (8) Random Forest (RF). The MSN and MSN2 were the performed poorly. The EUC, MAL, and RAW were in the middle of the spectrum, leaving the RF, GNN, and ICA the best methods of imputative methods to use. It should be noted that while RF rated highly, it also had the lowest distribution of RMSE value (Hudak, Crookston, Evans, Hall, & Falkowski, 2007). The study showed that imputation methods using LiDAR to determine forested structure can be useful and accurate to use.

The primary focus of this thesis revolved around the work performed by Riley and colleagues (1999) who created a terrain ruggedness index that can quantify terrain heterogeneity. Since early successional structure occupies a specific height stratum within an ecosystems niche, determining the magnitude of this structure compared other cover types can isolate early successional locations. The terrain ruggedness method proposed by these scientists calculates the sum of change in elevation using a grid cell system (Riley, DeGloria, & Elliot, 1999). The center value is calculated in relation to the eight surrounding neighbors (Figure 2), using the terrain ruggedness equation (Equation 1).
Equation 1 Terrain roughness equation

\[
TRI = \sqrt{(O - A)^2 + (O - B)^2 + (O - C)^2 + (O - D)^2 + (O - E)^2 + (O - F)^2 + (O - G)^2 + (O - H)^2}
\]

The output raster was classified into seven equal intervals (which range from the lowest value (level) to the maximum value (extremely rough) derived from the equation) giving each class a specific name (Riley, DeGloria, & Elliot, 1999).
Chapter 3: Data and Methodology

The purpose of the study is to 1. Determine the location of early successional and shrubland structure in a portion of eastern Connecticut, 2. Assist wildlife managers in locating early successional structure, and 3. Assess the accuracy of the early successional and shrubland structure from remotely sensed data. A data fusion between orthoimagery and Light Detection and Ranging (LiDAR) data is used to determine the early successional vegetation within the defined study area. Using the orthoimagery, general land cover types are defined using polygons (these are the controls for creating the thresholds for identifying similar cover types through the rest of the study area). The land cover thresholds for the cover types are determined and applied to all other pixels within the study area. Due to a “salt-and-pepper” effect caused by the high resolution data, a “clump and eliminate” tool was run to generalize patches comprised by only one or two pixels. The product raster from the clump and eliminate process is then assessed for its accuracy using 1-ft. orthoimagery flown in 2012. The resulting map depicting the locations of early successional structure can be used to help ecologists determine where the best locations would be to preserve and restore early successional structure for those animals that require these conditions. Figure 3 depicts early successional structure in natural color (a) and color infrared (b) aerial photographs. Early successional structure is characterized by low woody vegetation, high neighbor density, and a high diversity of plants.
Figure 3. Aerial image of early successional structure. (a) Natural color. (b) Color infrared.
3.1 Study Area

The study area consisted of twelve towns in the central part of Eastern Connecticut. The twelve towns are: Bozrah, Colchester, Columbia, East Haddam, East Hampton, Franklin, Haddam, Hebron, Lebanon, Marlborough, Montville, and Salem, covering approximately 954,340,650 m² (282476.3 acres) (See Appendix A and Figure 4). The study area defines the spatial extent of: (1) the LiDAR tiles needed for the terrain roughness analysis (See Appendix B), (2) orthoimages needed for land cover analysis (See Appendix C and Appendix D), and (3) the constraining area for the creation of the land cover training polygons.

3.2 Software Used

ESRI ArcMap 10.2 was the primary program used in the analysis. Python scripting was also used with various modules and extensions. ERDAS Imagine was used to create the ‘clump’ and ‘eliminate’ raster as well as to test the accuracy.

Figure 4 Study area of the 12 towns
3.3 Data

Data used in this research consisted of leaf-off orthoimagery (CT DEEP, 2012) Figure 5, a height raster of all non-bare ground LiDAR (CT DEEP, 2010) returns Figure 6, and land cover derived from the LiDAR and orthoimagery Figure 7, the latter two data sets derived by Parent et al. (2015).

Figure 5 2012 Aerial Image example of Columbia CT

Figure 6 Digital Surface Model example of Columbia, CT
3.4 Processing Raw LiDAR Files and Initial Land Cover

Processing the raw LiDAR files is the first step in the analysis. The LiDAR files are in the LAS standard format, consisting of millions of points for each 1 km by 1 km tile. To use the information in the point cloud (Figure 8), the data are converted to a raster grid format. Using a script named “LiDAR_rawData_Processing” the point cloud is converted into 1m by 1m pixel grid cells (Parent, Volin, & Civco, 2015). The script creates a 1m by 1m area cell and calculates the height at that particular point which is used as the raster cell value. This method is performed for the entire study area creating a digital surface model (DSM) (See Appendix E).

A preliminary land cover assessment can be performed on the compiled raster. The preliminary land cover analysis is used to determine the location of urban development and impervious surfaces. Urban development is considered to be manmade structures such as buildings, and other flat impervious surfaces, such as parking lots and roadways. These classes are used as a mask of pixels to be excluded in the creation of the terrain roughness index. Using a python script, variables such as the slope and infrared intensity are used to generate a land cover
raster. Introducing aerial imagery into the script allows for another validation of the land cover (Parent, Volin, & Civco, 2015).
Figure 8 (a) LiDAR Legend, (b) LiDAR point cloud for the area portrayed in Figure, (c) 3-D LiDAR view of the point cloud, (d) Profile view of the LiDAR point cloud.
3.5 Creating the Terrain Roughness Index

The terrain roughness index raster, derived from the LiDAR surface height data, is used to help differentiate land cover classes by calculating the magnitude of difference for a single pixel and its surrounding 8 neighbors. A pixel with a greater neighborhood difference in vegetative height obtains a higher magnitude. Python was used in the creation of the terrain roughness index. This step had three major components (1) creating the training polygons for the analysis, (2) developing the statistics for the terrain roughness raster, and (3) creating the terrain roughness raster.

ArcMap Model Builder was used to create a single mosaicked surface height raster of all rasters generated from the LiDAR point cloud processing script (Figure 9). A geodatabase is used to collect the single raster in a raster catalog. All of the individual rasters were uploaded into the raster catalog. Finally, the raster catalogs were converted to a raster dataset ready to be analyzed (Figure 10).

---

**Creating a Mosaicked Raster from a Raster Catalog**

![Diagram](image_url)

*Figure 9. Model used to create the mosaic from the LiDAR height files.*
The Terrain Roughness Index was calculated from the LiDAR mosaic of the DSM. A simplified version of Riley’s equation (Equation 1) is used to make the processing easier (Cooley, 2014):

**Equation 2 Simplified terrain roughness algorithm.**

\[
TRI = \sqrt{|Focal \ Statistics \ Maximum \ (3x3 \ Grid)^2 - Focal \ Statistics \ Minimum \ (3x3 \ Grid)^2|}
\]

Using python, the mosaicked DSM raster can be processed into the terrain roughness raster. The focal statistics from the nearest neighbor toolset (located in arcpy spatial analyst)
calculates a maximum value raster and minimum value raster for each cell using a 3 by 3 neighborhood grid.

### 3.6 Creating a Land Cover Mosaic

The preliminary land cover raster is created from the land cover script described in section 3.4. The purpose of this script is to identify urban development, which will be used to eliminate those cells in the terrain roughness raster.

As with the terrain roughness raster, the land cover raster is processed into 1 km by 1 km tiles. When the land cover raster is registered to the terrain roughness raster it has to be mosaicked into one large raster dataset. This is done by performing the same model function displayed in Figure 9. A raster catalog is created containing all the land cover rasters pertaining to the study area. Then the raster catalog is converted to a raster dataset (Figure 11).
Figure 11 Land cover map used to identify urban development. Missing data occurred due to errors in processing the LiDAR point cloud.
3.7 Creating the Structural Land Cover Raster

To create the structural land cover raster, training polygons were drawn for each of the land cover classes (Table 1). These polygons take up a small portion of the study area and are used as samples to generate the threshold values for land cover delineation. The training polygons were randomly placed in areas representative of the various land cover types.

To ensure the urban development was not included in the calculation, those pixels were removed from further analysis. The land cover raster generated in section 3.4 was reclassified so that all undeveloped areas were given a value of 1 and all impervious surfaces and urban development given a NoData value (Figure 12). The reclassified land cover raster is multiplied by the terrain roughness index raster, preserving the TRI values for land cover types other than urban and assessing the NoData value to remove urbanization. The result is the terrain roughness index raster with no urban development (Figure 13).

<table>
<thead>
<tr>
<th>No.</th>
<th>Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Bare Earth</td>
</tr>
<tr>
<td>2</td>
<td>Coniferous Forest</td>
</tr>
<tr>
<td>3</td>
<td>Deciduous Forest</td>
</tr>
<tr>
<td>4</td>
<td>Early Successional</td>
</tr>
<tr>
<td>5</td>
<td>Grassland/Pasture</td>
</tr>
<tr>
<td>6</td>
<td>Water</td>
</tr>
</tbody>
</table>

Figure 12. Land cover with the urban development changed to NoData
Figure 13. Terrain roughness raster with the urban development removed

With the urban development removed the land cover thresholds could be determined. Using python, 50 points were placed randomly inside each of the training polygons. To ensure that points were not duplicated and that there was a reasonable distance between points each point was given a buffer distance of two meters from any neighbor.

Using “Zonal Statistics as Table” found in spatial analyst the magnitude at that location was determined for each of the land cover types defined by the training polygons (five reports in total). Using python, the magnitude value was obtained from the reports for each land cover type and placed into respective blank lists. These land cover lists become the dataset values used in creating a box plot which shows the thresholds for each land cover type defined by the training
polygons (Figure 14).

Using “matplotlib” the first and third quartile (upper and lower limits) for each class were determined. These values were used in reclassifying the terrain roughness index raster remap range threshold values. If a pixel fell within one of these ranges, it would be classified the corresponding land cover type, the result was a five class raster. The two forest classes were combined into one forest class creating four final classes (Table 2).

Table 2. Final land cover classes for the four class structure map

<table>
<thead>
<tr>
<th>Class Number</th>
<th>Abbreviated Class Name</th>
<th>Class Name</th>
</tr>
</thead>
<tbody>
<tr>
<td>1</td>
<td>Flat</td>
<td>Water, Bare Earth, Grass</td>
</tr>
<tr>
<td>2</td>
<td>ES Early</td>
<td>Early Stage Successional Structure</td>
</tr>
<tr>
<td>3</td>
<td>LS Late</td>
<td>Late Stage Successional Structure</td>
</tr>
<tr>
<td>4</td>
<td>Forest</td>
<td>Forest (Deciduous or Coniferous)</td>
</tr>
</tbody>
</table>

Figure 14. Box plot depicting the land cover thresholds
3.8 Testing the Accuracy

Accuracy assessment was performed in ERDAS Imagine. Before the assessment was performed, the image was spatially-generalized to reduce the “salt-and-pepper” effect. Using ERDAS Imagine, the structural raster was “clumped”, which groups contiguous groups of pixels in one thematic class (CITE ERDAS). Next, the clump raster was processed with the “Eliminate” tool, which removes pixel clumps smaller than the specified threshold (in this case the threshold was three), replacing the value of the pixel with that of the nearby larger clumps (CITE ERDAS) (Figure 15).

The refined raster was used to check the validity of the structural raster. The appropriate window size needed to validate the structural raster was chosen from four individual trials. Each window size contains a threshold value, which is the minimum number of pixels per class that should be represented in the window before a random point is created (Table 3). Each window was given 80 random points with 20 points per class to test which window gave the best accuracy results.

<table>
<thead>
<tr>
<th>Window Size (m)</th>
<th>Maximum Threshold</th>
<th>Threshold Used</th>
<th>Percent Used</th>
</tr>
</thead>
<tbody>
<tr>
<td>3x3</td>
<td>9</td>
<td>9</td>
<td>100.0%</td>
</tr>
<tr>
<td>5x5</td>
<td>25</td>
<td>25</td>
<td>100.0%</td>
</tr>
<tr>
<td>9x9</td>
<td>81</td>
<td>75</td>
<td>92.6%</td>
</tr>
<tr>
<td>15x15</td>
<td>225</td>
<td>100</td>
<td>44.4%</td>
</tr>
</tbody>
</table>

Figure 15. Output from the Clump Analysis (Top) showing the various groups and the Eliminate (Bottom) showing the removal of the salt and pepper effect.
The output from each accuracy assessment report contained the area-adjusted producer’s accuracy, user’s accuracy, errors of commission, and errors of omission for each class. These values were graphed to determine which window would be the best to use for the final accuracy assessment.

The final accuracy assessment shows the overall data accuracy and map accuracy, using a stratified random accuracy assessment, which selects the number of samples proportional to the areal extent of each class. The number of random points (samples) needed to perform the assessment is calculated from the multinomial distribution equation. The multinomial equation (Equation 3) takes into account the proportion each class has in relation to the overall area and the size of the sample being used (Congalton & Green, 2009):

**Equation 3 Multinomial Distribution Equation.**

\[ N = \frac{\chi^2 \Pi_i (1 - \Pi_i)}{b_i^2} \]

Equation 3 shows the multinomial distribution equation used to calculate the number of samples per class. \( N \) is the number of samples, \( \chi^2 \) is chi squared, \( \Pi \) is the proportion that class had to the overall area, and \( b \) is the precision. The samples per class were based on a 90.0% confidence interval with 10.0% precision, making \( \alpha = 0.1 \). To calculate \( \chi^2 \) (chi squared), the probability and degrees of freedom must be determined. The degrees of freedom for this assessment is equal to one. Equation 3 shows the equation to calculate the probability where alpha is the precision and \( k \) is the number of classes. Using

**Equation 4 Calculating the probability**

\[ 1 - \frac{\alpha}{k} \]

\[ 1 - \frac{0.1}{4} = 0.975 \]
python, the chi squared value is determined to be 3.841. The proportion for each class in relation to the overall area can be seen in (Table 4).

<table>
<thead>
<tr>
<th>Class</th>
<th>Number of Pixels/Class</th>
<th>Proportion</th>
</tr>
</thead>
<tbody>
<tr>
<td>Bare Earth, Grass, Water</td>
<td>169,276,905</td>
<td>17.7%</td>
</tr>
<tr>
<td>Early Stage Successional Structure</td>
<td>126,858,516</td>
<td>13.2%</td>
</tr>
<tr>
<td>Late Stage Successional Structure</td>
<td>144,236,630</td>
<td>15.1%</td>
</tr>
<tr>
<td>Forest (Deciduous or Coniferous)</td>
<td>513,968,599</td>
<td>53.8%</td>
</tr>
<tr>
<td><strong>Total # of Pixels</strong></td>
<td>954,340,650</td>
<td></td>
</tr>
</tbody>
</table>

Bare Earth, Grass, Water = \( \frac{3.841 \times 0.1774(1 - 0.1774)}{0.1^2} = 56 \text{ Samples} \)

Early Stage Successional Structure = \( \frac{3.841 \times 0.1329(1 - 0.1329)}{0.1^2} = 42 \text{ Samples} \)

Late Stage Successional Structure = \( \frac{3.841 \times 0.1511(1 - 0.1511)}{0.1^2} = 48 \text{ Samples} \)

Forest (Deciduous or Coniferous) = \( \frac{3.841 \times 0.5386(1 - 0.5386)}{0.1^2} = 170 \text{ Samples} \)

The number of samples is calculated for each class then summed to give the total number of 316 samples. When the assessment is completed the producer’s accuracy, user’s accuracy, errors of commission, and errors of omission are determined to give the per class and overall accuracy of the sample of points representing the dataset for both the standard calculation and area-adjusted calculation of accuracy. The standard overall accuracy is calculated by taking the sum of correctly classified points divided by the total number of sum of all sample points. The
area-adjusted accuracy is calculated by taking the sum of the correctly classified adjusted class probability. The adjusted class probability is calculated by taking the number of correctly classified for that class divided by the total number of the total number of points in that class times area weight.
Chapter 4: Results

4.1 Final Structure Land Cover Raster

The product of the terrain roughness index algorithm are two maps showing the potential location of early successional structure in the study area. The area of coverage portrayed by all for classes is approximately 954,340,650 m².

The first map (Figure 16) depicts the location of early stage successional structure and late stage successional structure as one class called Class 2. Class 1 portrays the location of bare earth, grass, and water (referred to as “Flat”). Class 3 shows the location of deciduous and coniferous forests. Of the approximate 950,000,000 m² classified by the terrain roughness algorithm 28.4% (66,989.07ac) of land are considered to be in a state of either early or late successional growth. Bare earth, grass, and water (Class 1) make up 17.7% (41,829.23ac) of the area and the forest land (Class 3) makes up 53.9% (127,004.41) of the area.

The second map (Figure 17) depicts the location of early successional structure and late stage successional structure as two distinctly different classes. Out of the approximate 200-thousand acres classified by the terrain roughness algorithm, 13.3% (31,347.42ac) is classified as early stage successional structure (Class 2). Late stage successional structure (Class 3) makes up 15.1% (144,105,438.15 m²) of the approximate 950,000,000 m² area. Bare earth, grass, and water (Class 1), and forest (Class 4) have the same area as in the previous map.
Figure 16 Map depicting structure classes based on a 3 Class Assessment. Missing data in the map is a result in unprocessed LiDAR point cloud data and the removed urban development.
Figure 17 Map depicting the structure classes based on a 4 Class Assessment. Missing data in the map is a result in unprocessed LiDAR point cloud data and the removed urban development.
**4.2 Trial Accuracy Results**

The four windows (3m x 3m, 5m x 5m, 9m x 9m, and 15m x 15m) were processed to determine which window size was most viable for performing a full accuracy assessment. The assessment showed that the 9m x 9m window had the highest result with a standard accuracy (SA) overall of 80.0% and an area-adjusted calculation (AAC) overall accuracy of 80.9% (Tables 5 and 6), ranking it the highest threshold window tested. The smallest window size (3m x 3m) had the lowest SA overall at 47.5% and an AAC overall of 45.3% (See Tables 5 and 6). The 5m x 5m and the 15m x 15m windows are in the middle on each side of the 9m x 9m window. The 5m x 5m window shows an incline of accuracy to the highest threshold value while the 15m x 15m window shows the declining trend from the highest threshold window (See Figures 18 and 19).

<table>
<thead>
<tr>
<th>Grid</th>
<th>3x3 (m)</th>
<th>5x5 (m)</th>
<th>9x9 (m)</th>
<th>15x15 (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Acc.</td>
<td>47.5%</td>
<td>68.8%</td>
<td>80.0%</td>
<td>75.0%</td>
</tr>
<tr>
<td>K-hat</td>
<td>0.30</td>
<td>0.58</td>
<td>0.73</td>
<td>0.67</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Grid</th>
<th>3x3 (m)</th>
<th>5x5 (m)</th>
<th>9x9 (m)</th>
<th>15x15 (m)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Overall Acc.</td>
<td>45.3%</td>
<td>66.0%</td>
<td>80.5%</td>
<td>75.7%</td>
</tr>
<tr>
<td>K-hat</td>
<td>0.45</td>
<td>0.66</td>
<td>0.8</td>
<td>0.76</td>
</tr>
</tbody>
</table>
For each of the land cover classes in the four window sizes, the producer’s accuracy (PA), the user’s accuracy (UA), the error of commission (EC), and the error of omission (EO) were determined.
Since the primary purpose of this research is to determine the location of early successional structure (class 2) throughout a portion of the eastern of Connecticut, the main focus of the assessment statistics focuses on class 2.
4.2.1 Trial Windows Standard Calculation Results

Four windows (3m x 3m, 5m x 5m, 9m x 9m, and 15m x 15m) were processed to determine which window size was most viable for performing a full accuracy assessment used in the standard calculation assessment (SCA). Table 5 shows the 9m x 9m window size has the highest overall accuracy at 80.0% and a k-hat of 0.73 making it the most suitable for the final accuracy assessment. The 9m x 9m window has a producer’s accuracy of 95.0% for the flat class, 83.3% accuracy for identifying early stage successional structure, 60.9% accuracy for identifying late stage successional structure, and 84.2% accuracy for identifying forested regions (both coniferous and deciduous). In comparison the 9m x 9m window size has a user’s accuracy of 95.0% accuracy for identifying the flat class, 75.0% accuracy for identifying early stage successional structure, 70.0% accuracy for identifying late stage successional structure, and 80.0% accuracy for identifying forested regions (both coniferous and deciduous). Appendix I shows the statistical break down for the 9mx 9m window accuracy assessment.

The 15m x 15m accuracy window came in second for the SCA with an overall accuracy of 75.0% and a k-hat of 0.67. This shows a 5.0% difference overall between the 15m x 15m and the 9m x 9m window sizes. Breaking down the 15m x 15m into the individual classes looking at the producers and user’s accuracy specifically for the early stage successional structure (70.6% and 60.0% respectfully) and late stage successional structure (53.57% and 75.0% respectfully). For the early stage successional structure, there is 12.7% decrease in producer’s accuracy and a 15.0% decrease in user’s accuracy compared to the 9m x 9m window. The late stage successional structure decreases in producer’s accuracy by 21.4% and increases in user accuracy by 5.0%. Appendix J shows the statistical break down for the 15m x 15m window accuracy assessment.
The worst window size for the SCA was the 3m x 3m window size. The overall accuracy for the 3m x 3m window size was 47.5% which is 32.5% lower than the 9m x 9m and a k-hat of 0.3 which is the lowest value for all four windows. Breaking down the producers and user’s accuracy for the 3m x 3m window; the producer’s accuracy for the early stage successional structure was 40.0% which is a 43.3% decrease from the 9m x 9m window size. The user’s accuracy for the early stage successional structure was also 40.0% which is a 35.0% decrease from the 9m x 9m window size. The producer’s accuracy and user’s accuracy for the late stage successional structure are the same as for the early stage successional structure. The producer’s accuracy and user’s accuracy have decreased by 20.9% and 30.0% respectfully when compared to the 9m x 9m window size.

The 5m x 5m window size fell in-between the 3m x 3m window size and the 15m x 15m window size. The overall accuracy is 68.8% which is a 11.3% decrease from the 9m x 9m window size but is a 21.3% increase from the 3m x 3m window size. The k-hat for the 5m x 5m window size is 0.58. The producer’s accuracy and user’s accuracy for the early stage successional structure for the 5m x 5m window is 81.4% and 65.0%, respectfully. Compared to the 9m x 9m window size the producer’s accuracy is has only decreased by 2.1% where the user’s accuracy has decreased by 10.0%. The producer’s accuracy and user’s accuracy for the late stage successional structure are 50.0% and 60.0% respectfully. These compare to the 9m x 9m having a 10.9% decrease for the producer’s accuracy and a 10.0% decrease for the user’s accuracy.
4.2.2 Trial Windows Area-Adjusted Calculation Results

The area adjusted accuracy assessment (AAA) takes the matrix formed in the SCA and applies a weight based on the proportion of area in each class. The total study area contains an area of 954,340,650m$^2$; with the flat class having 169,276,905m$^2$, early stage successional structure area of 126,858,516m$^2$, late stage successional structure area of 144,236,630m$^2$, and forested (both coniferous and deciduous) area of 513,968,599m$^2$. The largest area class being the forested class.

Table 5 shows the 9m x 9m window size has the highest overall accuracy at 80.8% making it the most suitable for the final accuracy assessment. The 9m x 9m window has a producer’s accuracy of 95.7% for the flat class, 80.6% accuracy for identifying early stage successional structure, 42.9% accuracy for identifying late stage successional structure, and 95.0% accuracy for identifying forested regions (both coniferous and deciduous). In comparison the 9m x 9m window size has a user’s accuracy of 95.0% accuracy for identifying the flat class, 75.0% accuracy for identifying early stage successional structure, 70.0% accuracy for identifying late stage successional structure, and 80.0% accuracy for identifying forested regions (both coniferous and deciduous). Appendix I shows the statistical breakdown for the 9mx 9m window accuracy assessment.

The 15m x 15m accuracy window came in second for the SCA with an overall accuracy of 75.7%. This shows a 5.2% difference overall between the 15m x 15m and the 9m x 9m window sizes. Breaking down the 15m x 15m into the individual classes looking at the producers and user’s accuracy specifically for the early stage successional structure (66.4% and 60.0% respectfully) and late stage successional structure (37.6% and 75.0% respectfully). For the early
stage successional structure there is a 12.7% decrease in producer’s accuracy and a 15.0% decrease in user’s accuracy compared to the 9m x 9m window. The late stage successional structure shows an increase of 1.4% for producer’s accuracy and a 5.0% decrease in the user’s accuracy. Appendix J shows the statistical break down for the 15m x 15m window accuracy assessment.

The overall accuracy for the 3m x 3m window size was 45.3% which is 35.2% lower than the 9m x 9m and lowest for all four windows. Examining the producers and user’s accuracy for the 3m x 3m window; the producer’s accuracy for the early stage successional structure was 25.0% which is a 55.6% decrease from the 9m x 9m window size. The user’s accuracy for the early stage successional structure was also 40.0% which is a 35.0% decrease from the 9m x 9m window size. The late stage successional structure producer’s accuracy 18.5% which is a 24.4% decrease and a user’s accuracy of 75.0% which is a 5.0% decrease from the 9m x 9m window size.

The 5m x 5m window size fell in-between the 3m x 3m window size and the 15m x 15m window size. The overall accuracy is 66.0% which is a 14.5% decrease from the 9m x 9m window size but is a 20.7% increase from the 3m x 3m window size. Breaking down the producer’s accuracy and user’s accuracy for the early stage successional structure for the 5m x 5m window is 77.4% and 65.0% respectfully. Compared to the 9m x 9m window size the producer’s accuracy is has only decreased by 3.3% where the user’s accuracy has decreased by 10.0%. The producer’s accuracy and user’s accuracy for the late stage successional structure are 27.3% and 60.0% respectfully. These compare to the 9m x 9m having a 13.6% decrease for the producer’s accuracy and a 10.0% decrease for the user’s accuracy.
4.3 Final Accuracy Assessment (3 Classes)

The final overall accuracy for the combined early successional structure (ES) (early stage successional structure and late stage successional structure) assessment is 91.1% for both the standard calculation and area-adjusted calculation (Table 7 and 8). The producer’s accuracy (PA) for flat (class 1) and forest (class 3) in Tables 7 and 8, respectively, are the same. The combined early successional structure (ES), is higher using the standard calculation (Table 7) than that of the area-adjusted calculation (Table 8) by 0.1%. The user’s accuracy (UA) in both Table 7 and Table 8 are the same. The errors of omission (O) are low for classes 1 (Flat) and 3 (Forest) at 3.6% and 3.2%, respectfully, for both the standard calculation and the area-adjusted calculation. Class 2 (ES) has a 20.0% standard calculation error of omission (Table 7) which is 16.4% higher than class 1 (Flat) and 16.8% higher than class 3 (Forest). The area-adjusted calculation has a 20.1% error of omission (Table 8), which is 16.5% higher than class 1 (Flat) and 16.9% higher than class 3 (Forest). The area-adjusted calculation error of omission is 0.1% higher than the stand calculation error of omission. The error of commission for both

Table 7. Final overall accuracy assessment for the standard calculation (not adjusted) 3 classes

<table>
<thead>
<tr>
<th></th>
<th>PA</th>
<th>UA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flat</td>
<td>96.4%</td>
<td>94.6%</td>
</tr>
<tr>
<td>ES</td>
<td>80.0%</td>
<td>93.3%</td>
</tr>
<tr>
<td>Forest</td>
<td>96.8%</td>
<td>88.8%</td>
</tr>
</tbody>
</table>

Overall accuracy = 91.1%
K-hat = 0.85

Table 8. Final overall accuracy assessment for the area-adjusted calculation 3 classes

<table>
<thead>
<tr>
<th></th>
<th>PA</th>
<th>UA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flat</td>
<td>96.4%</td>
<td>94.6%</td>
</tr>
<tr>
<td>ES</td>
<td>79.9%</td>
<td>93.3%</td>
</tr>
<tr>
<td>Forest</td>
<td>96.8%</td>
<td>88.8%</td>
</tr>
</tbody>
</table>

Overall accuracy = 91.1%
K-hat = 0.91
Tables 11 and 12 for each class from 5.4%, 6.7%, and 11.2% (flat (class 1) and late early successional structure (class 3) respectfully).

4.4 Final Accuracy Assessment (4 Classes)

The final accuracy assessment for the early stage early successional structure (Early ES) and late stage early successional structure (Late ES) classes had an accuracy of 85.1% for both the standard calculation and area adjusted calculation. This is a 6.0% decrease in accuracy from the final 3 class assessment. The producer’s accuracy (PA) for Flat, Early ES, and Forest were the same at 96.4%, 66.7%, and 96.8%, respectfully, for both the standard and area adjusted calculations. There is a 0.1% decrease in the producer’s accuracy in the Late ES class going from the standard calculation (Table 9) to the area-adjusted calculation (Table 10). The user’s accuracy (UA) for all four classes remained the same for both the standard calculation (Table 9) and the area adjusted calculation (Table 10). This was also true for the errors of omission (O) and errors of commission in both tables.

Table 9 Final overall accuracy assessment for the standard calculation (not adjusted) 4 classes

<table>
<thead>
<tr>
<th></th>
<th>PA</th>
<th>UA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flat</td>
<td>96.4%</td>
<td>94.6%</td>
</tr>
<tr>
<td>Early ES</td>
<td>66.7%</td>
<td>76.2%</td>
</tr>
<tr>
<td>Late ES</td>
<td>57.9%</td>
<td>68.8%</td>
</tr>
<tr>
<td>Forest</td>
<td>96.8%</td>
<td>88.8%</td>
</tr>
</tbody>
</table>

Overall accuracy = 85.1%
K-hat = 0.77

Table 10 Final overall accuracy assessment for the area-adjusted calculation 4 classes

<table>
<thead>
<tr>
<th></th>
<th>PA</th>
<th>UA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flat</td>
<td>96.4%</td>
<td>94.6%</td>
</tr>
<tr>
<td>Early ES</td>
<td>66.7%</td>
<td>76.2%</td>
</tr>
<tr>
<td>Late ES</td>
<td>57.8%</td>
<td>68.8%</td>
</tr>
<tr>
<td>Forest</td>
<td>96.8%</td>
<td>88.8%</td>
</tr>
</tbody>
</table>

Overall accuracy = 85.1%
K-hat = 0.85
Chapter 5: Discussion

5.1 Relation to Other Studies

The research documented in this thesis relates back to other research performed looking for similar structural properties in vegetation research characterizing forest succession using LiDAR in the Northwest United States had an overall accuracy of 90.1% with the mature multistory (MMS) and closed stem exclusion (CSE) broken out and a 95.5% when the MMS and CSE were combined (Falkowski, Evans, Martinuzzi, Gessler, & Hudak, 2009). In accordance with the two land cover rasters produced from the TRI, the three-class land cover raster was 0.9% better than the Northwest Study with the MSS and CSE broken out, but was 4.4% lower when the MSS and CSE were combined.

In comparison the four-class structure raster is lower than the study performed in the Northwest forests for the MSS and CSE broken out as well as aggregated together. The difference between the MSS and CSE aggregated together and the four-class system hold a difference of 10.0%. With the MSS and CSE broken into their own classes the difference is lower at 4.6%. The northwestern forest study had a higher accuracy overall potentially due to the number of variable they introduced into their equation, where this study only used one dimension. The northwestern study though offers a good comparison to this study, due to both looking at structure of forested areas. Even though the two sites are across the country early successional structure has similar characteristics across the board. Where you would find a difference is in the habitat make up such as describing the vegetative species that make up early successional structure.

The research performed in the Northwest Forests did explain that there were issues in classifying MSS and CSE which is similar to the early stage successional structure (ESSH) and
late stage successional structure (LSSH). The rasters created through this research showed the same result to create a land cover raster when the ESSH and LSSH are combination. The research was able to show stand delineation similar to the Northwest forest research, only using one set of parameters rather than the 30+ parameters used in the other research. The inclusion of more parameters to the dataset might increase the accuracy of the land cover raster.

Given the results from the study performed by Rittenhouse (2014) there is a higher accuracy of 15.1% between the three-class land cover classification and the land cover raster produced by Dr. Rittenhouse. With the early stage successional structure and late stage successional structure broken out the accuracy is 9.1% higher than Dr. Rittenhouse’s land cover raster. Seeing these results LiDAR has the high potential for classifying land cover at a generic level but when the land cover classes get specific it has a harder time in rendering one class from another. The same drop in accuracy happened using LiDAR in the research on the forests in the northwest portion of the United States.

5.2 Management Purposes

The structure maps produced from the TRI algorithm have the potential to assist wildlife managers and agencies, such as the Connecticut Department of Energy and Environmental Protection: Forestry and Wildlife Divisions, in managing wildlife species that utilize early successional structure. In Chapter 1: Introduction, examples of wildlife species were discussed as users of early successional structure either for shelter or predation. In September the New England cottontail was denied to be added to the endangered species list. As such management plans were developed outlining what can be done to recover this species. Using this technique there is a potential of helping improve the management plan in recovering the New England cottontail population.
The land cover maps created depicting the location of early successional structure has the potential to assist wildlife managers in maintaining wildlife populations. The land cover maps show wildlife managers areas of early successional structure which has a threefold capability. The first shows managers where site work can be conducted reducing search time for finding suitable areas to conduct analysis. For example, Figure 20 shows an area in Lebanon that would be suitable for performing site analytics.

Figure 20 Site in Lebanon, CT proposed for field analytic derived from the TRI land cover raster. (a) Orhtoimagery of the site and (b) the land cover representation derived from the TRI algorithm.
The example area was determined from visual inspection of the land cover raster as a suitable site. Further analysis (i.e. an on-site field assessment) of the area would determine the plausibility for an event such as live trapping. Even though an onsite verification would be needed the land cover raster directed the wildlife manager to the spot without having to randomly search high and low or the area. This opens the possibilities of areas that cannot be seen from aerial photography as well as from the road. Figure 21 located in Colchester shows the location of early successional structure that is set off the road.

Figure 21 Site located in Colchester, CT for field analytics. The site is located further off the road than the Lebanon site (Figure 20). (a) Orthoimage of the Colchester site, (b) a land cover assessment at the Colchester Site
The land cover raster was able to show this hidden location that could have otherwise gone over looked.

The second capability of land cover maps is to assist wildlife managers for site restoration. Connecticut has been losing area of early successional structure to forests and human involvement, which can push out species dependent on that cover type to survive. The type land cover maps portray the possible locations of early succession structure distributed through the study area. Managers can use these maps to find patches of early successional structure that are most effectively enlarged or improved upon. Instead of choosing a patch of forest that may or may not be suitable to establish an early successional patch, wildlife managers can turn to the land cover maps and find areas where early successional growth is already established but needs more room to grow.
For example, Figure 22 depicts an area of the back side of Babcock Pond Wildlife Management Area by the Colchester/ East Haddam line.

Figure 22 Interior of Babcock Pond Wildlife Management Area depicting remoteness from urban development but the predominate cover type being forest not the most ideal for early successional wildlife. (a) Orthoimage of Babcock Pond Wildlife Management Area, (b) Land cover for Babcock Pond Wildlife Management Area.

The area is remote, preferable to a wildlife manager wanting to release animals from a breeding program. The problem becomes looking at the land cover type, which is primarily dominated by forest. To make the area suitable for a species that thrive in early successional
structure all of the trees would have to be cut down creating a fragmented forest which is not good for those species that require a continuous forest.

On the other hand, Figure 23 shows an area of land in East Hampton with a far less area of forest growth, and more growth leaning toward the early successional stage going into an afforestation stage.

Figure 23 Site in East Hampton, CT what would benefit from clear cutting to revert the land back to early stage successional structure. The forest area that makes up the patch is less than in Babcock Pond Wildlife Management area, preventing forest segmentation. (a) Orthoimage of the East Hampton, CT site, (b) Land cover for the East Hampton, CT site.

Since the land is already in a state of successional growth it would be more prudent and wise to maintain the parcel of land in state of successional growth. The animal species that survive on that parcel have already adapted to that way of life. Therefore, it has the potential to
have less of an impact on the ecosystem then clear cutting the center of a forest. There are also economic advantages to harvesting from the area again keeping it in a state of early successional growth. The one being the cost of clear cutting is reduced; since most of the trees have been cut down it would cost as much to finish the job unlike the amount it would cost for a true forest stand to be cut down (which would involve having more machinery and a larger crew).

5.3 Constraints and Limitations

Understanding the constraints and limitations of the land cover maps allows the user to use the maps effectively and efficiently. One constraint comes from the algorithm used in relation to identifying dense tree canopy structure; an example of this would be a grove of pine trees. Since the canopy in a pine stand can get dense the magnitude of difference can represent that of early stage successional structure. The misclassification can be mitigated if large stands of mature pine groves are removed from the analysis. One way to remove the large stands of pine is to use the height of the trees and set a threshold value to the minimum height of a pine tree that should be included from removal of the dataset.

The limitations on the map accuracy come from both the pixel size of the raster and the 9m x 9m window used in the final assessment. The pixel size of the window is 1m x 1m therefore interpolating smaller pixel sizes creates the chance of introducing error into the dataset. Taking in account the assessment window size is also important, for utilizing the map. Choosing a window size smaller or greater than the optimal window used reduces the accuracy of the map. Adhering to the pixel accuracy and to the assessment window the two land cover maps can be used effectively and efficiently for management practices.
5.4 Further Research

The analysis performed within the research was confined to using the difference in magnitude base on the maximum height of the DSM. Expanding the research using other variable as inputs for the TRI such as the minimum height, mean height, or the standard deviation of height related to the DSM could refine the product produced from the initial analysis. Also applying the infrared band from the LiDAR has the potential to break the flat class into subclasses such as water, grass, and barren as well as break the forest class into coniferous and deciduous forests. There is also another potential advantage for combining the magnitude with the infrared, which is identifying evergreen shrubs such as mountain loral, which can be hard to distinguish in an aerial photograph. Using more than one variable such as the minimum height, or different return combination can assist in furthering segmentation of the early successional forest class.
References
   Francisco: Pearson Benjamin Cummings.

   Boca Raton: Taylor & Francis Group.


   Characterizing forest succession with lidar data: An evaluation for the Inland Northwest,


   neighbor imputation of species-level, plot-scale forest structure attributes from LiDAR

King, D. I., & Schlossberg, S. (2014, July 15). Synthesis of the conservation value of the early-

   mapping with airborne LiDAR and high resolution multispectral imagery in a forested
suburban landscape. 104. Retrieved from


Appendix A: The Study Area
Appendix B: LiDAR Coverage Area
Appendix C: Aerial Coverage Area

Study Area Location in Connecticut

Aerial Image of Study Area

Legend

Author: Jacob Conshick
Date: 7/21/2015
Appendix E: Compiled Maximum Height Digital Surface Model

Maximum Height Non-Ground Surface Raster

Legend
- Maximum Height
- Focus Area

Value
- High: 101.85
- Low: 0

Author: Jacob Conshick
Date: 7/31/2015
### Appendix G: 3x3 Window Accuracy Error Matrix

#### Standard error matrix

<table>
<thead>
<tr>
<th>Classification</th>
<th>Flat</th>
<th>Early ES</th>
<th>Late ES</th>
<th>Forest</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flat</td>
<td>14</td>
<td>3</td>
<td>2</td>
<td>1</td>
<td>20</td>
</tr>
<tr>
<td>Early ES</td>
<td>4</td>
<td>8</td>
<td>1</td>
<td>7</td>
<td>20</td>
</tr>
<tr>
<td>Late ES</td>
<td>2</td>
<td>6</td>
<td>8</td>
<td>4</td>
<td>20</td>
</tr>
<tr>
<td>Forest</td>
<td>0</td>
<td>3</td>
<td>9</td>
<td>8</td>
<td>20</td>
</tr>
<tr>
<td>Sum</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>20</td>
<td>80</td>
</tr>
</tbody>
</table>

#### Class distribution in classified map

<table>
<thead>
<tr>
<th>Weights</th>
<th>km²</th>
<th># Pixel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flat</td>
<td>0.1774</td>
<td>152349.21</td>
</tr>
<tr>
<td>Early ES</td>
<td>0.1329</td>
<td>114172.66</td>
</tr>
<tr>
<td>Late ES</td>
<td>0.1511</td>
<td>129812.97</td>
</tr>
<tr>
<td>Forest</td>
<td>0.5386</td>
<td>462571.74</td>
</tr>
</tbody>
</table>

#### Adjusted class probabilities

<table>
<thead>
<tr>
<th>Classification</th>
<th>Flat</th>
<th>Early ES</th>
<th>Late ES</th>
<th>Forest</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flat</td>
<td>0.1242</td>
<td>0.0266</td>
<td>0.0177</td>
<td>0.0089</td>
<td>0.1774</td>
</tr>
<tr>
<td>Early ES</td>
<td>0.0266</td>
<td>0.0532</td>
<td>0.0066</td>
<td>0.0465</td>
<td>0.1329</td>
</tr>
<tr>
<td>Late ES</td>
<td>0.0151</td>
<td>0.0453</td>
<td>0.0605</td>
<td>0.0302</td>
<td>0.1511</td>
</tr>
<tr>
<td>Forest</td>
<td>0.0000</td>
<td>0.0808</td>
<td>0.2424</td>
<td>0.2154</td>
<td>0.5386</td>
</tr>
</tbody>
</table>

#### Proportions belonging to class

<table>
<thead>
<tr>
<th></th>
<th>Flat</th>
<th>Early ES</th>
<th>Late ES</th>
<th>Forest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flat</td>
<td>0.7000</td>
<td>0.1500</td>
<td>0.1000</td>
<td>0.0500</td>
</tr>
<tr>
<td>Early ES</td>
<td>0.2000</td>
<td>0.4000</td>
<td>0.0500</td>
<td>0.3500</td>
</tr>
<tr>
<td>Late ES</td>
<td>0.1000</td>
<td>0.3000</td>
<td>0.4000</td>
<td>0.2000</td>
</tr>
<tr>
<td>Forest</td>
<td>0.0000</td>
<td>0.1500</td>
<td>0.4500</td>
<td>0.4000</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>EPS</th>
<th>EV</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>EPS</td>
<td>0.1659</td>
<td>0.2059</td>
<td>0.3272</td>
</tr>
<tr>
<td>EV</td>
<td>0.0006</td>
<td>0.0026</td>
<td>0.0043</td>
</tr>
<tr>
<td>SE</td>
<td>0.0246</td>
<td>0.0513</td>
<td>0.0653</td>
</tr>
</tbody>
</table>

*EPS = Estimated proportion of samples in class*

*EV = Estimated variance of EPS*

*SE = Standard error of EV*
Appendix G: 3x3 Window Accuracy Error Matrix (Continued)

PA = producer's accuracy
UA = User's accuracy

<table>
<thead>
<tr>
<th>Area</th>
<th>PA</th>
<th>UA</th>
<th>Area</th>
<th>PA</th>
<th>UA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flat</td>
<td>70.0%</td>
<td>70.0%</td>
<td>Flat</td>
<td>74.9%</td>
<td>70.0%</td>
</tr>
<tr>
<td>Early ES</td>
<td>40.0%</td>
<td>40.0%</td>
<td>Early ES</td>
<td>25.8%</td>
<td>40.0%</td>
</tr>
<tr>
<td>Late ES</td>
<td>40.0%</td>
<td>40.0%</td>
<td>Late ES</td>
<td>18.5%</td>
<td>40.0%</td>
</tr>
<tr>
<td>Forest</td>
<td>40.0%</td>
<td>40.0%</td>
<td>Forest</td>
<td>71.6%</td>
<td>40.0%</td>
</tr>
</tbody>
</table>

Overall accuracy = 47.5%  
K-hat = 0.30

Overall accuracy = 45.32%  
K-hat = 0.45
Appendix H: 5x5 Window Accuracy Error Matrix

**Standard error matrix**

<table>
<thead>
<tr>
<th>Classification</th>
<th>Reference</th>
<th>Flat</th>
<th>Early ES</th>
<th>Late ES</th>
<th>Forest</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flat</td>
<td>Flat</td>
<td>18</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Early ES</td>
<td>0</td>
<td>13</td>
<td>4</td>
<td>3</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Late ES</td>
<td>0</td>
<td>1</td>
<td>12</td>
<td>7</td>
<td>20</td>
</tr>
<tr>
<td></td>
<td>Forest</td>
<td>0</td>
<td>0</td>
<td>8</td>
<td>12</td>
<td>20</td>
</tr>
<tr>
<td>Sum</td>
<td></td>
<td>18</td>
<td>16</td>
<td>24</td>
<td>22</td>
<td>80</td>
</tr>
</tbody>
</table>

**Class distribution in classified map**

<table>
<thead>
<tr>
<th>Classification</th>
<th>Weights km²</th>
<th># Pixel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flat</td>
<td>0.1774</td>
<td>152349.21</td>
</tr>
<tr>
<td>Early ES</td>
<td>0.1329</td>
<td>114172.66</td>
</tr>
<tr>
<td>Late ES</td>
<td>0.1511</td>
<td>129812.97</td>
</tr>
<tr>
<td>Forest</td>
<td>0.5386</td>
<td>462571.74</td>
</tr>
<tr>
<td>Sum</td>
<td>1.00</td>
<td>858907</td>
</tr>
</tbody>
</table>

**Adjusted class probabilities**

<table>
<thead>
<tr>
<th>Classification</th>
<th>Reference</th>
<th>Flat</th>
<th>Early ES</th>
<th>Late ES</th>
<th>Forest</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flat</td>
<td>Flat</td>
<td>0.1596</td>
<td>0.0177</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.1774</td>
</tr>
<tr>
<td></td>
<td>Early ES</td>
<td>0.0000</td>
<td>0.0864</td>
<td>0.0266</td>
<td>0.0199</td>
<td>0.1329</td>
</tr>
<tr>
<td></td>
<td>Late ES</td>
<td>0.0000</td>
<td>0.0076</td>
<td>0.0907</td>
<td>0.0529</td>
<td>0.1511</td>
</tr>
<tr>
<td></td>
<td>Forest</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.2154</td>
<td>0.3231</td>
<td>0.5386</td>
</tr>
</tbody>
</table>

**Proportions belonging to class**

<table>
<thead>
<tr>
<th>Classification</th>
<th>Flat</th>
<th>Early ES</th>
<th>Late ES</th>
<th>Forest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flat</td>
<td>0.9000</td>
<td>0.1000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>Early ES</td>
<td>0.0000</td>
<td>0.6500</td>
<td>0.2000</td>
<td>0.1500</td>
</tr>
<tr>
<td>Late ES</td>
<td>0.0000</td>
<td>0.0500</td>
<td>0.6000</td>
<td>0.3500</td>
</tr>
<tr>
<td>Forest</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.4000</td>
<td>0.6000</td>
</tr>
</tbody>
</table>

| EPS | 0.1596 | 0.1117 | 0.3327 | 0.3960 |
| EV  | 0.0001 | 0.0004 | 0.0041 | 0.0041 |
| SE  | 0.0122 | 0.0204 | 0.0640 | 0.0637 |

*EPS = Estimated proportion of samples in class*

*EV = Estimated variance of EPS*

*SE = Standard error of EV*
Appendix H: 5x5 Window Accuracy Error Matrix (Continued)

PA = producer's accuracy
UA = User's accuracy

<table>
<thead>
<tr>
<th></th>
<th>Standard calculation (not adjusted)</th>
<th>Area-adjusted calculation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PA</td>
<td>UA</td>
</tr>
<tr>
<td>Flat</td>
<td>100.0%</td>
<td>90.0%</td>
</tr>
<tr>
<td>Early ES</td>
<td>81.3%</td>
<td>65.0%</td>
</tr>
<tr>
<td>Late ES</td>
<td>50.0%</td>
<td>60.0%</td>
</tr>
<tr>
<td>Forest</td>
<td>54.6%</td>
<td>60.0%</td>
</tr>
<tr>
<td>Overall accuracy</td>
<td>68.8%</td>
<td></td>
</tr>
<tr>
<td>K-hat</td>
<td>0.58</td>
<td></td>
</tr>
</tbody>
</table>
Appendix I: 9x9 Window Accuracy Error Matrix

### Standard error matrix

<table>
<thead>
<tr>
<th>Classification</th>
<th>Reference</th>
<th>Flat</th>
<th>Early ES</th>
<th>Late ES</th>
<th>Forest</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flat</td>
<td></td>
<td>19</td>
<td>1</td>
<td>0</td>
<td>0</td>
<td>20</td>
</tr>
<tr>
<td>Early ES</td>
<td></td>
<td>0</td>
<td>15</td>
<td>5</td>
<td>0</td>
<td>20</td>
</tr>
<tr>
<td>Late ES</td>
<td></td>
<td>1</td>
<td>2</td>
<td>14</td>
<td>3</td>
<td>20</td>
</tr>
<tr>
<td>Forest</td>
<td></td>
<td>0</td>
<td>0</td>
<td>4</td>
<td>16</td>
<td>20</td>
</tr>
<tr>
<td>Sum</td>
<td></td>
<td>20</td>
<td>18</td>
<td>23</td>
<td>19</td>
<td>80</td>
</tr>
</tbody>
</table>

### Class distribution in classified map

<table>
<thead>
<tr>
<th>Classification</th>
<th>Weights</th>
<th>km²</th>
<th># Pixel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flat</td>
<td>0.1774</td>
<td>152349.21</td>
<td>169276905</td>
</tr>
<tr>
<td>Early ES</td>
<td>0.1329</td>
<td>114172.66</td>
<td>126858516</td>
</tr>
<tr>
<td>Late ES</td>
<td>0.1511</td>
<td>129812.97</td>
<td>144236630</td>
</tr>
<tr>
<td>Forest</td>
<td>0.5386</td>
<td>462571.74</td>
<td>513968599</td>
</tr>
</tbody>
</table>

| Sum            | 1.00    | 858907 | 954340650 |

### Adjusted class probabilities

<table>
<thead>
<tr>
<th>Classification</th>
<th>Reference</th>
<th>Flat</th>
<th>Early ES</th>
<th>Late ES</th>
<th>Forest</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flat</td>
<td></td>
<td>0.1685</td>
<td>0.0089</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.1774</td>
</tr>
<tr>
<td>Early ES</td>
<td></td>
<td>0.0000</td>
<td>0.0997</td>
<td>0.0332</td>
<td>0.0000</td>
<td>0.1329</td>
</tr>
<tr>
<td>Late ES</td>
<td></td>
<td>0.0076</td>
<td>0.0151</td>
<td>0.1058</td>
<td>0.0227</td>
<td>0.1511</td>
</tr>
<tr>
<td>Forest</td>
<td></td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.1077</td>
<td>0.4308</td>
<td>0.5386</td>
</tr>
</tbody>
</table>

### Proportions belonging to class

<table>
<thead>
<tr>
<th>Classification</th>
<th>Flat</th>
<th>Early ES</th>
<th>Late ES</th>
<th>Forest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flat</td>
<td>0.9500</td>
<td>0.0500</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>Early ES</td>
<td>0.0000</td>
<td>0.7500</td>
<td>0.2500</td>
<td>0.0000</td>
</tr>
<tr>
<td>Late ES</td>
<td>0.0500</td>
<td>0.1000</td>
<td>0.7000</td>
<td>0.1500</td>
</tr>
<tr>
<td>Forest</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.2000</td>
<td>0.8000</td>
</tr>
</tbody>
</table>

### EPS, EV, SE

<table>
<thead>
<tr>
<th></th>
<th>EPS</th>
<th>EV</th>
<th>SE</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flat</td>
<td>0.1761</td>
<td>0.1237</td>
<td>0.0117</td>
</tr>
<tr>
<td>Early ES</td>
<td>0.2467</td>
<td>0.0029</td>
<td>0.0190</td>
</tr>
<tr>
<td>Late ES</td>
<td>0.4535</td>
<td>0.0268</td>
<td>0.0536</td>
</tr>
<tr>
<td>Forest</td>
<td>1.0000</td>
<td>0.0509</td>
<td>0.0536</td>
</tr>
</tbody>
</table>

*EPS = Estimated proportion of samples in class
EV = Estimated variance of EPS
SE = Standard error of EV
Appendix I: 9x9 Window Accuracy Error Matrix (Continued)

PA = producer’s accuracy
UA = User’s accuracy

<table>
<thead>
<tr>
<th>Area</th>
<th>PA</th>
<th>UA</th>
<th>PA</th>
<th>UA</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flat</td>
<td>95.0%</td>
<td>95.0%</td>
<td>Flat</td>
<td>95.7%</td>
</tr>
<tr>
<td>Early ES</td>
<td>83.3%</td>
<td>75.0%</td>
<td>Early ES</td>
<td>80.6%</td>
</tr>
<tr>
<td>Late ES</td>
<td>60.9%</td>
<td>70.0%</td>
<td>Late ES</td>
<td>42.9%</td>
</tr>
<tr>
<td>Forest</td>
<td>84.2%</td>
<td>80.0%</td>
<td>Forest</td>
<td>95.0%</td>
</tr>
</tbody>
</table>

**Overall accuracy** = 80.00%

K-hat = 0.73

**Overall accuracy** = 80.5%

K-hat = 0.80
### Appendix J: 15x15 Window Accuracy Error Matrix

#### Standard error matrix

<table>
<thead>
<tr>
<th>Classification</th>
<th>Flat</th>
<th>Early ES</th>
<th>Late ES</th>
<th>Forest</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flat</td>
<td>18</td>
<td>2</td>
<td>0</td>
<td>0</td>
<td>20</td>
</tr>
<tr>
<td>Early ES</td>
<td>0</td>
<td>12</td>
<td>8</td>
<td>0</td>
<td>20</td>
</tr>
<tr>
<td>Late ES</td>
<td>0</td>
<td>3</td>
<td>15</td>
<td>2</td>
<td>20</td>
</tr>
<tr>
<td>Forest</td>
<td>0</td>
<td>0</td>
<td>5</td>
<td>15</td>
<td>20</td>
</tr>
<tr>
<td>Sum</td>
<td>18</td>
<td>17</td>
<td>28</td>
<td>17</td>
<td>80</td>
</tr>
</tbody>
</table>

#### Class distribution in classified map

<table>
<thead>
<tr>
<th>Weights</th>
<th>km2</th>
<th># Pixel</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flat</td>
<td>0.1774</td>
<td>152349.21</td>
</tr>
<tr>
<td>Early ES</td>
<td>0.1329</td>
<td>114172.66</td>
</tr>
<tr>
<td>Late ES</td>
<td>0.1511</td>
<td>129812.97</td>
</tr>
<tr>
<td>Forest</td>
<td>0.5386</td>
<td>462571.74</td>
</tr>
</tbody>
</table>

| Sum  | 1.00 | 858907 | 954340650 |

#### Adjusted class probabilities

<table>
<thead>
<tr>
<th>Classification</th>
<th>Flat</th>
<th>Early ES</th>
<th>Late ES</th>
<th>Forest</th>
<th>Sum</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flat</td>
<td>0.1596</td>
<td>0.0177</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.1774</td>
</tr>
<tr>
<td>Early ES</td>
<td>0.0000</td>
<td>0.0798</td>
<td>0.0532</td>
<td>0.0000</td>
<td>0.1329</td>
</tr>
<tr>
<td>Late ES</td>
<td>0.0000</td>
<td>0.0227</td>
<td>0.1134</td>
<td>0.0151</td>
<td>0.1511</td>
</tr>
<tr>
<td>Forest</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.1346</td>
<td>0.4039</td>
<td>0.5386</td>
</tr>
</tbody>
</table>

#### Proportions belonging to class

<table>
<thead>
<tr>
<th>Classification</th>
<th>Flat</th>
<th>Early ES</th>
<th>Late ES</th>
<th>Forest</th>
</tr>
</thead>
<tbody>
<tr>
<td>Flat</td>
<td>0.9000</td>
<td>0.1000</td>
<td>0.0000</td>
<td>0.0000</td>
</tr>
<tr>
<td>Early ES</td>
<td>0.0000</td>
<td>0.6000</td>
<td>0.4000</td>
<td>0.0000</td>
</tr>
<tr>
<td>Late ES</td>
<td>0.0000</td>
<td>0.1500</td>
<td>0.7500</td>
<td>0.1000</td>
</tr>
<tr>
<td>Forest</td>
<td>0.0000</td>
<td>0.0000</td>
<td>0.2500</td>
<td>0.7500</td>
</tr>
</tbody>
</table>

 EPS = Estimated proportion of samples in class
 EV = Estimated variance of EPS
 SE = Standard error of EV
Appendix J: 15x15 Window Accuracy Error Matrix (Continued)

PA = producer's accuracy
UA = User's accuracy

<table>
<thead>
<tr>
<th></th>
<th>Standard calculation (not adjusted)</th>
<th>Area-adjusted calculation</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>PA</td>
<td>UA</td>
</tr>
<tr>
<td>Flat</td>
<td>100.0%</td>
<td>90.0%</td>
</tr>
<tr>
<td>Early ES</td>
<td>70.6%</td>
<td>60.0%</td>
</tr>
<tr>
<td>Late ES</td>
<td>53.6%</td>
<td>75.0%</td>
</tr>
<tr>
<td>Forest</td>
<td>88.2%</td>
<td>75.0%</td>
</tr>
</tbody>
</table>

Overall accuracy = 75.0%
K-hat = 0.67

Overall accuracy = 75.67%
K-hat = 0.76
Appendix K: Terrain Ruggedness Script

import arcpy
import numpy as np
from arcpy.sa import *
import os, os.path

## Tool Name: Terrain Roughness Model
## Source Name: Terrain_Roughness_v1.py
## Version: 1.1
## Author: Jacob Conshick

## Required Arguments:
## Set the arcpy workspace
## Set the output folder
## Set the LiDAR type

## Description:
## Create a Terrain Roughness Index by using the Facal Statistics of the Maximum and the Minimum. To get the 7 classes for the roughness index bring the raster into ArcMap and right click and under symbology click classified and then change the classes to 7. The 7 classes created are as followed 1- Level, 2- Nearly Level, 3- Slightly Rugged, 4- Intermediately Rugged, 5- Moderately Rugged, 6- Highly Rugged, 7- Extremely Rugged.

## This script was created using an equation for Terrain Roughness by Shuan Riley.

# ---------------------------------------
# import arcpy
# import numpy as np
# from arcpy.sa import *
# import os, os.path
arcpy.env.overwriteOutput = 1
arcpy.CheckOutExtension("spatial")

#ONLY CHANGE THESE THREE PARAMETERS BELOW!!!!
#ALL PARAMETERS SHOULD BE STRINGS!!!!

#Set the folder containing the raster you would like to run the
#Terrain Roughness on.
arcpy.env.workspace = "

#Set the folder where you want the created rasters to end up
Output_Folder = "

#Set the Type of LiDAR being Processed
#Examples = DEM, Intensity, Max_Height, Median_Height...etc
LiDAR_Type = 'Median_Height'

#Create the folder to house the Minimum Focal Statistics
Output_FC_Min = '%s\Focal_Statistics_Min'%(Output_Folder)
if not os.path.exists(Output_FC_Min):
    os.makedirs(Output_FC_Min)

#Create the folder to house the Maximum Focal Statistics
Output_FC_Max = '%s\Focal_Statistics_Max'%(Output_Folder)
if not os.path.exists(Output_FC_Max):
os.makedirs(Output_FC_Max)

#Create the folder to house the Terrain Roughness Index Rasters

Output_TRI = '%s\Terrain_Roughness_Index'%(Output_Folder)

if not os.path.exists(Output_TRI):
    os.makedirs(Output_TRI)

#---------------------------------------------------------------#

###MAIN SCRIPT###

raster_Lst = arcpy.ListRasters()

list_length = len(raster_Lst)

cnt = 1

for raster in raster_Lst:
    print "Starting Terrain Roughness Index Cell %s..."%(cnt)
    Neighbor = NbrRectangle (3, 3, "CELL")

    #Minimum Focal Statistics
    print "Focal Min...%s"%(cnt)
    FC_Min = r"%s\LiDAR_%s_Focal_Statistics_3x3min_%s.img"%(Output_FC_Min, LiDAR_Type,cnt)
    outFocalMin = FocalStatistics (raster, Neighbor, "MINIMUM")
    outFocalMin.save(FC_Min)

    #Maximum Focal Statistics
    print "Focal Max...%s"%(cnt)
    FC_Max = r"%s\LiDAR_%s_Focal_Statistics_3x3max_%s.img"%(Output_FC_Max, LiDAR_Type,cnt)
    outFocalMax = FocalStatistics (raster, Neighbor, "MAXIMUM")
    outFocalMax.save(FC_Max)

    #TRI
TRI = SquareRoot(Abs(Minus(Square(FC_Max),Square(FC_Min))))

TRI.save (r'\%s\Median_Height_TRI_Cell_%s.img'% (Output_TRI,cnt))

print "Completed Terrain Roughness Index Cell %s..."%(cnt)

cnt += 1
Appendix L: Early Successional Structure Script

## Tool Name: LandCover Validation

## Source Name: Module LandCover Validation 2.2

## Version: ArcGIS 10.2.1

## Author: Jacob Conshick

##

## This tool performs constrained aggregative clustering based on traditional k-means and spatial k-means based on a minimum spanning tree algorithm:

##

## Source:


## Efficient regionalisation techniques for socio-economic geographical units using minimum spanning trees.

## "International Journal of Geographical Information Science"

import arcpy
import glob
import os
import os.path
import traceback
import sys
from arcpy.sa import *
from operator import itemgetter
import matplotlib.pyplot as figure
import SubModuleLandcoverBoxPlotList
import SubModuleRemapTable

print 'All Land Cover Modules Imported'

##System##
arcpy.env.overwriteOutput = 1
arcpy.CheckOutExtension ('spatial')
arcpy.env.outputCoordinateSystem = arcpy.SpatialReference(26918)

##Variables##

#Set the study name of the area being looked at
StudyName = 'Thesis'

#Input the Landcover Raster generated from the Landcover Script
LandcoverRaster = r''

#Input the Terrain Roughness Raster
Raster = r''

#Input the file location that holds the training polygons for the classifying the landcover
LCShapefileWKSP = r""

#Input the folder where you want the Rasters generated
TRIOOutput = r""

wksp = LCShapefileWKSP
arcpy.env.workspace = wksp

##Create Folders##
FldRndPts = '%s\Random_Points'%(wksp)
if arcpy.Exists(FldRndPts) == False:
    print 'Creating Random Point Folder'
    os.makedirs(FldRndPts)
else:
    print "Random Point Folder Created"

FldStats = '%s\Statistics'%(wksp)
if arcpy.Exists(FldStats) == False:
    print "Creating Spatial Statistics Folder"
    os.makedirs(FldStats)
else:
    "Random Point Folder Created"
print "Spatial Statistics Folder Created"

##Script##

#Dodging the Image#
#This section of the script creates the image used in burning out the urban development from the script
#There are no input variables in this part of the script

#------------------------------#

##Removing Urban Development##
#This takes the raster created above and removes all the urban development located in the Terrain Roughness
#Ratser. This ensures that the statistics used for creating the box plots below do not contain the urban development.
InputRaster = r'\%s\%s_TRI_UrbanRemoved.img'%(TRIOutput,StudyName)
if arcpy.Exists(InputRaster) != True:
    Dodge = arcpy.sa.Times(Raster,LandcoverRaster)
    Dodge.save(InputRaster)

#------------------------------#

LCLst = []
LCMeanLst = []
LCShpLst = arcpy.ListFeatureClasses()
print LCShpLst
LenLCShpLst = len(LCShpLst)
cnt = 1
for Shp in LCShpLst:
    basename = os.path.splitext (Shp)
    print basename
    name = basename[0]
    print name
    key = name[0:4]
    print key
    LCLst.append(key)

#------------------------------------------------------------------------------------------#
#Create the Random Points for the Module
rpts = r'\%s\%s_%s_Random_Points.shp'%(FldRndPts,StudyName,name)
print rpts
if arcpy.Exists(rpts) == False:
    print "Creating Random Points %s" %cnt
    arcpy.CreateRandomPoints_management (FldRndPts, rpts, Shp, "", 50 , '2 Meters','MULTIPOINT', '')
else:
    print "Random Points %s Created" %cnt
#------------------------------------------------------------------------------------------#
#Create the Statistics from the Random Points
StatsInput = r'\%s\%s_%s_Random_Points.shp'%(FldRndPts,StudyName,name)
print StatsInput
outTable = r'\%s\%s_%s_Stats_Table.dbf'%(FldStats,StudyName,name)
if arcpy.Exists(outTable) == False:
    print "Creating Statistics \%s\" \%cnt
    ZonalStatisticsAsTable (StatsInput, 'FID', InputRaster, outTable, "DATA", 'MEAN')
else:
    print "Statistics \%s Created" \%cnt

#-----------------------------------------------------------------------------------------#

rows = arcpy.SearchCursor(outTable)
MeanLst = []
for row in rows:
    mean = row.getValue('MEAN')
    MeanLst.append(mean)
LCMeanLst.append(MeanLst)
cnt+=1
print 'Random Points Processed'
print 'Creating Figure'

##Create the Graph##

figure.boxplot(LCMeanLst, sym = 'r+', vert = True, notch = True)
labels = LCLst
left = SubModuleLandcoverBoxPlotList.ListChoice(LenLCShpLst)
figure.xticks(left,labels,size = 8, color = 'r')
figure.show()
RemapLst = []
Quants = []
for i in range(0, LenLCShpLst, 1):
    Quan = figure.boxplot(LCMeanLst)['boxes'][i].get_ydata()[1:3]
    x = i + 1
    QuanLst = Quan[1]
    Remap = [LCLst[i], Quan[1], x]
    RemapLst.append(Remap)
    Quants.append(QuanLst)

print RemapLst
Quantiles = sorted(Quantiles)
print Quantiles

#---------------------------------------------------------------------------------------------------------------------------#

##Reclassifying the Raster##

print ""
print "Creating Reclassify Raster"
Reclassify = r'%s\TRI_%s_Reclassify.img' % (TRIOutput, StudyName)
if arcpy.Exists(Reclassify) != True:
    remapTable = SubModuleRemapTable.remap (LenLCShpLst, Quantiles)
    print remapTable
    remap = arcpy.sa.RemapRange(remapTable)
    print remap
outReclassify = arcpy.sa.Reclassify (InputRaster, 'Value', remap)
outReclassify.save(Reclassify)

print ""
print ""
print 'Finished Reclassify Starting Extraction'

sys.exit()

#---------------------------------------------------------------#

##Calculate the Area##

rows = arcpy.SearchCursor(Reclassify)
valueLst = []
for row in rows:
    value = row.getValue('VALUE')
    valueLst.append(value)

print ""

print ""
print ""
print "ALL PROCESSES HAVE BEEN COMPLETED"
Appendix M: Python Module – Land Cover Boxplot List

def ListChoice(LenLCShpLst):
    if LenLCShpLst == 1:
        return [1]
    if LenLCShpLst == 2:
        return [1,2]
    if LenLCShpLst == 3:
        return [1,2,3]
    if LenLCShpLst == 4:
        return [1,2,3,4]
    if LenLCShpLst == 5:
        return [1,2,3,4,5]
    if LenLCShpLst == 6:
        return [1,2,3,4,5,6]
    if LenLCShpLst == 7:
        return [1,2,3,4,5,6,7]
    if LenLCShpLst == 8:
        return [1,2,3,4,5,6,7,8]
    if LenLCShpLst == 9:
        return [1,2,3,4,5,6,7,8,9]
    if LenLCShpLst == 10:
        return [1,2,3,4,5,6,7,8,9,10]
    if LenLCShpLst == 11:
        return [1,2,3,4,5,6,7,8,9,10,11]
    if LenLCShpLst == 12:
        return [1,2,3,4,5,6,7,8,9,10,11,12]
if LenLCShpLst == 13:
    return [1,2,3,4,5,6,7,8,9,10,11,12,13]

if LenLCShpLst == 14:
    return [1,2,3,4,5,6,7,8,9,10,11,12,13,14]

if LenLCShpLst == 15:
    return [1,2,3,4,5,6,7,8,9,10,11,12,13,14,15]