1-1-2012

The Detection of Forest Structures in the Monongahela National Forest Using LiDAR

Dominique Norman
norman45@marshall.edu

The Detection of Forest Structures in the Monongahela National Forest Using LiDAR

Thesis submitted to
the Graduate College of
Marshall University

In partial fulfillment of the
requirements for the degree of
Master of Science

Geography

by

Dominique Norman

Approved by

Dr. Anita Walz, Committee Chairperson
Dr. Godwin Djietror
Dr. Min Kook Kim

Marshall University

May 2012

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF TABLES

ABSTRACT

The mapping of structural elements of a forest is important for forestry management to provide a baseline for old and new-growth trees while providing height strata for a stand. These activities are important for the overall monitoring process which aids in the understanding of anthropogenic and natural disturbances. Height information recorded for each discrete point is key for the creation of canopy height, canopy surface, and canopy cover models. The aim of this study is to assess if LiDAR can be used to determine forest structures. Small footprint, leaf-off LiDAR data were obtained for the Monongahela National Forest, West Virginia. This dataset was compared to Landsat imagery acquired for the same area. Each dataset endured supervised classifications and object oriented segmentation with random forest classifications. These approaches took into account derived variables such as, percentages of canopy height, canopy cover, stem density, and normalized difference vegetation index, which were converted from the original datasets. Evaluation of the study depicted that the classification of the Landsat data produced results ranging between 31.3 and 50.2%, whilst the LiDAR dataset produced accuracies ranging from 54.7 to 80.1%. The results of this study increase the potential of LiDAR to be used regularly as a forestry management technique and warrant future research.

CHAPTER 1

Introduction

The identification of structural elements of a forest is linked to the identification and studies of wildlife habitats, carbon sequestration, and impacts of climate change. Forests are one of the world's most important renewable resources. They help stimulate the economy and absorb an estimated 0.68 ± 0.34 billion tons of carbon per year in northern forests (Myneni et al., 2001). Above all, forests are a crucial part of biodiversity. Maintaining biodiversity increases biological and social benefits that we all can take advantage of. Nationwide, forests are in decline due to deforestation, climate change, and an increase of pollution. Old growth trees have experienced dramatic declines in the Appalachians, especially in the West Virginia region. Remote sensing procedures can be deployed to capture and study these declines.

The techniques and technologies utilized for forestry has shifted over the last century due to advances in technology. These advances include technologies such as the global positioning system (GPS), aerial, and satellite photography (Leckie, 1990) Traditional forestry methods involved time consuming fieldwork and inventory processes, but the advancements in remote sensing and geographic information systems (GIS) have propelled measurement, inventory, and sampling, which are all components of forestry management. Forest management is the division of forestry concerned with the overall administrative, economic, legal, and regulations of the forest (Waring & Schlesinger, 1985). Remote sensing technologies, such as light detection and ranging

(LiDAR) gives governmental, private, and non-profit organizations the ability to answer questions about forestry management for restoration, conservation, and policy creation.

Although the integration of LiDAR and forestry is still developing, studies thus far depict a great relationship that creates detailed information about forest structure and composition. In past years, LiDAR has been engrossed in terrain mapping and atmospheric research (Wandinger, 2005). Current LiDAR research focuses on the use of LiDAR for 3-D modeling, watershed, flood, and coastal mapping. Within forestry, the focus has been on identifying individual tree species and producing stand maps (Reutebuch et al., 2005). Less emphasis has been given to the use of LiDAR as a functioning tool for forestry management. Therefore, the purpose of this master's thesis is to use LiDAR data to identify and characterize vertical structures of a mixed forest for the purpose of forestry management.

The conceptual and methodological procedures developed for this research will aid in the analysis of the forest structure by providing a baseline for comparison and measurement of forest structure shifts and declines. In addition, the comparison of LiDAR and Landsat remote sensing technologies will be examined. This examination will determine which sensor provides the best overall results. In addition, the comparison of classification techniques will be conducted to determine which algorithm is more robust. The methods presented in this research will contribute to knowledge about forest structures, remote sensing, and classification techniques.

CHAPTER 2

Background and Literature Review

2.1 Decline of Montane Forests in West Virginia

Since the industrial revolution, the economic value of trees has been increasingly important, but the ecological value is often ignored until it is too late. The decline can be seen in the case of the montane forest in West Virginia. Due to many anthropogenic and natural disturbances, montane red spruce and douglas fir are now considered endangered. During the 19th century, the distribution of these trees in the Appalachians was very extensive, but now they are highly restricted (Noss et al., 1995; Adams et al., 2009). There are several old growth red spruce stands in the Appalachians that were “over looked” by loggers. West Virginia's Shavers Mountain and Gaudineer Knob (Rollins, 2005; Adams et al., 2009) have beautiful old growth virgin stands that are now federally protected. Montane forests provide an important habitat for many endangered animals. These consist of the cheat mountain salamander, the sow whet owl, and the recently de-listed northern flying squirrel (Byers et al., 2010; Loeb et al., 2000). Also, red spruce and douglas fir’s shallow root system is a vital source for the stabilization of watersheds which protects soils from erosion (Rollins, 2005). These watersheds are very important to West Virginia's ecosystems and economy. They provide fresh and clean water for wildlife, drinking, and recreation. In addition, these forests form a vital ecotone which provides a transition zone between northern hardwoods and spruce-fir forests (Battles & Fahey, 2000). If the anthropogenic and natural disturbances continue on the same trend, then forest tree species with limited ranges are expected to become extinct as their access to suitable habitat becomes more limited (Potter et al., 2009).

Wood has been an invaluable resource to mankind. It is only relatively recently that this resource has become abused. With the impulse of technology and land use conversion certain habitats have been destroyed. Within America, European settlers were the first to make the expansion westward. Since the 1700s, European settlers have logged many old growth conifer stands for economic and technological growth. Red spruce, one of the dominant trees at the time, along with balsam firs, hemlocks and eastern white pines were some of the first species to experience a decline (Nowacki et al., 2009). Since this initial disturbance, the spruce-fir forests have been in a detrimental state.

In addition to logging, climate change has been affecting forests all over the world from the beginning of time, but the recent influx of anthropogenic induced climate change is negatively affecting forests everywhere. It is fair to say that logging has affected spruce-fir forests more, but several studies have shown that climate change and various forms of atmospheric pollutants have influenced a lack of growth increase in spruce-fir (Adams et al., 2009; Dale et al., 2001; Johnson et al., 1985). This recent evidence depicts that global climate change is being effected by a large anthropogenic component (Houghton et al., 1996). This component will create an increase in timing, frequency, and extent of natural disturbances that will force forests to face rapid alternations (Dale et al., 2001). Montane forests are already in a critical state, and with the increase of disturbances they could face extinction.

2.2 Overview of LiDAR

LiDAR, is an optical and active remote sensing technique that measures scattered light from target items (NOAA 2008). These data can be used to measure distance, speed, rotation, and chemical composition of clouds. Traditionally these data is acquired through airborne mechanism, but there is also space-based LiDAR that is becoming increasingly popular. In addition to how LiDAR is acquired, there are now two forms of LiDAR. They are discrete and full waveform LiDAR. Although full waveform LiDAR data are thought to provide more information, because it is relatively new and denser, the amount of research compared to discrete data has been minimal. LiDAR data can be applied to all environments including urban and forestry. More and more LiDAR based research has been based on the application of LiDAR in a forestry application, but mostly within the context of measuring tree crowns and vegetation biomass (Lim et al., 2003; Dubayah & Drake, 2000).

LiDAR is considered to be a "breakthrough technology for forestry applications" (Dubayah & Drake, 2000). Many studies show that LiDAR is an appropriate dataset for the analysis of vegetative structures. Alone, the LiDAR point clouds aid in the visualization of a forestry structure, but with the creation of several algorithms, these point clouds can depict much more. LiDAR transforms a traditional 2-D representation to a 3-D representation of a forest. This visualization gives a unique insight into the structure of the forest. LiDAR datasets can be transformed into canopy height models, canopy surface models, canopy cover models, digital terrain models, intensity images, and many more derived products. The creation of these derived products aid in the analysis of silvicultural studies, such as vegetation distribution, habitat mapping, and species identification.

The classification or identification of tree species has been a recent trend in LiDAR research.   There are numerous studies describing the accuracies of LiDAR data for separating and classifying deciduous and coniferous tree species. These studies use statistical methods such as regression analysis and discriminant analysis to determine tree species classifications using variables derived from LiDAR. The results of these studies show that LiDAR data can accurately classify tree species as well as tree structure or landcover (Song et al., 2002; Boyd & Hill 2007; Reitberger et al., 2008; Kim et al., 2008). Kim et al. (2008) reported that LiDAR has the ability to distinguish between species within the broadleaved deciduous and conifers. This study reports classification accuracies which range from 88.8% to 98.2%.

The intensity data provided for both discrete and full waveform LiDAR data can provide much needed information about an object's reflectance properties. Intensity is related to reflectance properties of vegetation, absence of foliage, type of foliage, as well as canopy openness (Kim et al., 2008). The observed intensity properties are similar to the near infrared reflectance properties for vegetation. Essentially, the healthier vegetation that is present the "brighter" the intensities are. Other spectral reflectance studies have reported this direct relationship (Ahokas et al., 2006 & Kim et al., 2008). Conversely, Song et al., (2002) reported that intensity values do not conform to theoretical reflectance properties of materials. The reflectance values follow a path of relative magnitudes of reflectance that allows separability. Because intensity values do not adhere to true reflectance properties, the spectral signatures from separate datasets cannot be directly compared. In addition, intensity values are noisy. The amount of noise available is a product of gaussian, impulse, and speckle noise (Xudong et al., 2005). With an increase in noise there is a decrease in separability. Song et al., (2002); Yan & Shaker, (2006); Xudong et al., (2005) agree that using a filtering method to decrease the noise within the recorded intensities will allow more separability, thus allowing for a more detailed classification.

Intensity values derived from LiDAR have not been studied fully so more research needs to be conducted to fully understand their suitability in forestry applications. Lim et al., (2003) revealed that intensity values from different sensors could not be compared because of the lack of calibration amongst LiDAR sensors. In addition the intensity values resulting from LiDAR cannot be directly compared to intensity values from other remote sensing technologies. Similarly, to fully understand intensity values, there is a need to create more algorithms for these data. These algorithms could lead to more accurate measurements of forest parameters and vegetation classifications as well as having a more automated process for tree delineation and classification.

CHAPTER 3

Conceptual Model

This thesis research proposes a remote forest classification (RFC) process for the investigation of the structural elements of a forest. The guiding principles for the RFC process were established by several studies which have identified individual tree species and tree structures using LiDAR products (Kim et al., 2008; Sullivan, 2008). These principles require the model to be driven by an understanding of remote sensing, forestry management, scientific and statistical methods. The model for this study will be built on a foundation of existing successful LiDAR remote sensing principles. In addition, the model will be adaptive so it can evolve with continuing research in remote sensing as well as forestry inventory management. To be successful, the model must be able to address current natural resource questions which include: 1) What are the proportions of old growth and new growth trees in our national forests? 2) Are the threatened and endangered tree species declining locally and nationally? Each of these issues requires addressing the fundamental question: What are the tree structures in a particular study area? The RFC process will enhance an overall understanding of the canopy and sub-layers, determine the stand quality and status of the resources, and discern between old and new trees. The proposed framework formally links natural resources and remote sensing in an iterative process which aids in the inventory process of forest management. This linkage will provide the information needed to better understand the theory behind the methodology for tree species classification for remote sensing technologies. The conceptual model proposed is illustrated in Figure 1.

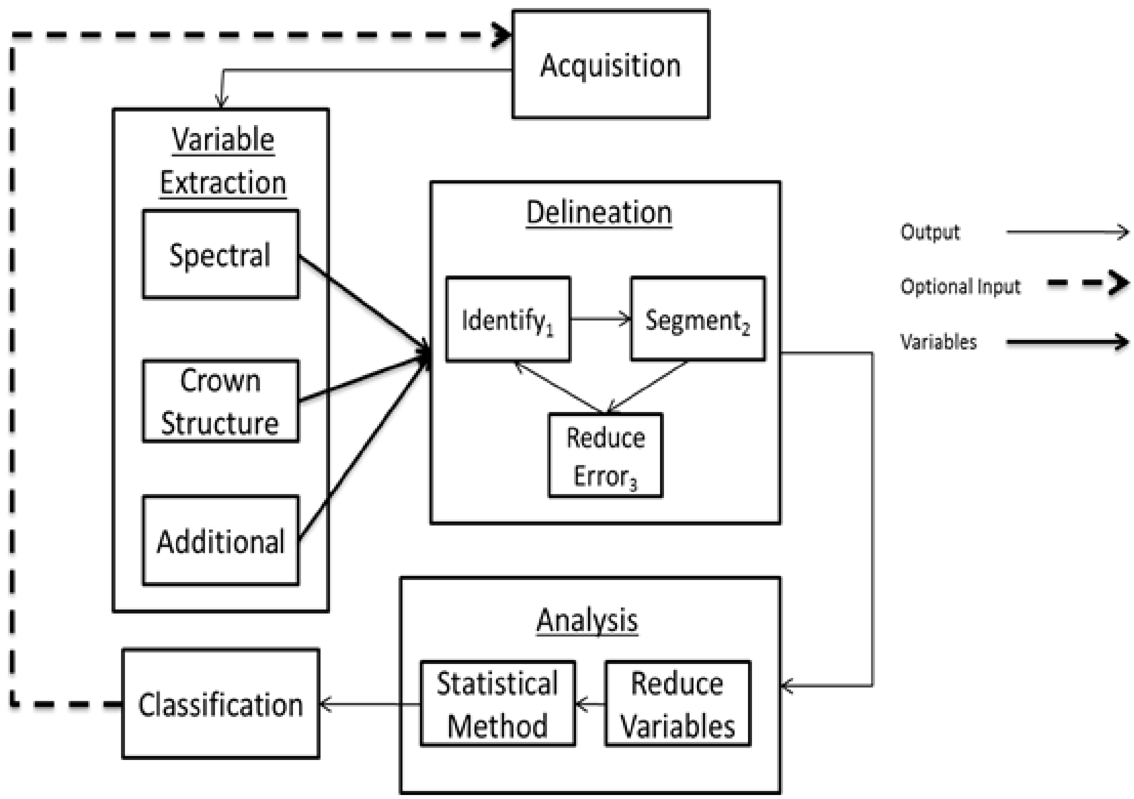


Figure 1: Remote Forest Classification Process (RFC)

In this study, the goal of the RFC process is to classify delineated tree stands using remotely sensed data. It consists of five steps, namely, acquisition, variable extraction, tree isolation, analysis, and classification. The first segment, acquisition, represents the procurement of the remotely sensed data needed to start the RFC process. These data is restricted to remotely sensed data because of the spectral information that is captured.

The second segment represents the types of variables that can be derived from remotely sensed data for the isolated trees. Spectral variables refer to any reflectance data that can be used to distinguish between vegetation or structure types. Crown structure variables refer to measurements of the crown. These measurements could include crown width, crown length, crown height, stem density, and percentage of canopy cover. An additional variables section allows for any information that can be derived from other sensors. Which includes information gained from sensors such as synthetic aperture radar or hyperspectral data.

The third segment represents the tree stand delineation process. To decrease error from other types of vegetation, it is important to isolate trees, tree canopies, or tree stands. This segment depicts an iterative three step process for stand delineation. First, the scale at which the landscape is delineated must be identified. These scales can range from individual trees, tree stands, or larger landscape sections. Second, the identified scale must be delineated or segmented. This process can be automated or digitized. And last, the delineated segments must be inspected for reduction of error, which can be achieved by a visual inspection or statistically. For smaller or more detailed segments this process can be iterated until satisfactory stand delineation is achieved.

The fourth segment represents the analysis of the variables. First, these variables must be reduced to maximize the amount of variance between variables. Second, a method must be applied which analyzes the variables and groups them statistically for classification. The fifth and final segment is the classification process. The results of the statistical methods will allow for the classification of the groupings. The classifications are analyzed to determine if the results are satisfactory. If not, additional data can be acquired or the delineation process can be repeated for different tree segments.

The most significant aspect of the conceptual model is the universal nature at which any remote sensing technology can adapt the model for the purpose of tree stand classification. This universality is indicated in the simplicity of the model. The acquisition, tree isolation, variables, analysis, and classification segments are coordinated in a fashion that provides a framework for structural classification for trees without being specific to one remotely sensed technique or technologies.

The universal and iterative nature proposed in this model will provide several benefits. The first benefit is the optimization of spatial resolution. This optimization is accomplished by the iterative process for tree isolation by identifying the optimum geographic scale for species classification for a specific sensor. A second benefit is the ability to compare temporal resolutions. This comparison is accomplished through a time series analysis that can detect variability in the composition of the forest and the identification of any extreme changes. Last, the universal nature of the RFC process allows for the measurement of variables at the same locations using different remote sensing techniques, which allows for the comparison of imagery from various remote sensing sensors.

Remote sensing has become more advanced and there is a need for sophisticated techniques to analyze the increasingly detailed data being made available. But the nature of forest planning and management necessary to address current and future questions about natural resources will depend on institutional policies and logistics. The overarching benefit of this conceptual model is the availability of a framework by which a certain region can be assessed for its forestry composition and be applied to specific natural resource issues or questions.

CHAPTER 4

Methodology

4.1 Introduction

LiDAR has many advantages over other aerial and satellite remote sensing technologies. LiDAR has the ability to capture structures in three dimensions while capturing dense data which result in high resolution imagery. The height information obtained during a LiDAR acquisition contains important information which correlates to tree heights. These heights, along with coordinate information, can provide canopy heights, canopy densities, and percentages of canopy cover. All of these variables have been used in previous studies to classify forest structures and species. Brennan and Webster (2006) reported the effect of intensity, canopy height, surface height, and density of LiDAR returns on land cover features. This study recommends the use of LiDAR height derived information to distinguish between land cover classifications. This study also reports that dense coniferous forest stands were harder to classify because of the inability of LiDAR to penetrate the canopy. These results directly relate to LiDAR and forest structures. Lo and Chen (2008) developed a workflow which utilizes canopy heights derived from LiDAR to analyze vertical profiles of trees to delineate individual tree crowns. This study found that individual tree crowns were harder to distinguish due to the nature of a dense forest.

To better understand the utility of LiDAR, it must be compared with another remote sensing dataset. Landsat is an easily available moderate resolution dataset which provides spectral information ranging from 0.45 – 12.5 μm (Campbell, 2002). This spectral information can be used in a variety of vegetation indices that estimates vegetation cover and density. The most popular vegetation index is the Normalized Difference Vegetation Index (NDVI). Freitas et al., (2005) analyzed the relationship between NDVI, moisture vegetation index (MVI) and the structure of an Atlantic Rainforest. This study reported that NDVI is a good indicator of biomass in deciduous and dry forests, whereas MVI is a better indicator for rainforests.

In addition, the supervised classification and random forest classification methods must be compared to determine which is more robust for the classification of tree structures. Classification is the decision-making process that is used to understand large quantities of data (Ayhan and Kansu, 2010). The supervised classification technique requires defined training areas to determine the characteristics of each class. Lee et al., (2005) reported that the use of supervised classification yielded better results than unsupervised classification of Interferometric synthetic aperture radar (InSAR) data. In contrast, the random forest classification is a machine learning, rule based, classification which has many decision trees. Pal and Mather (2003) assessed the efficiency of decision tree algorithms for land cover classification. They concluded that the use of decision trees ultimately were computationally faster and had the ability to handle data of any scale with no statistical assumptions. In addition they found that once high-dimensional data were introduced that decision tree classifiers no longer exceeded maximum likelihood classifiers in accuracy. To better utilize the random forest algorithm, this study employed the object oriented segmentation process which groups features into homogenous objects or segments. According to Geneletti and Gorte (2003), the object oriented segmentation process enhanced the accuracies of land-cover classifications when compared to pixel based classifiers.

The Monongahela National Forest is a suitable study area to analyze the utility of using remote sensing to determine tree structures. This study site has few man-made structures which aid in each remote sensing acquisition to capture pure vegetation values. The presence of these structures can skew digital numbers by scattering the energy source used for each remote sensing technique (Campbell, 2002). This section presents a workflow which transforms variables which represent heights and densities related to the canopy and sub-layers of the forest.

4.2 Site Description

The Monongahela National Forest lies within the Alleghany mountains valley and ridge system and is described as lying within the strategic heart of the Appalachians (Mueller, 2003). This forest is more than 1.7 million acres and is comprised mostly of mixed, deciduous and conifer, forest types. The elevation ranges from 1,000 to 4,863 feet and is home to Canaan Valley, the highest valley east of the Mississippi River. This forest is owned by the United States Forest Service (USFS) and provides an ample amount of outdoor recreation during all seasons. For this study, the Monongahela National Forest was subset to a region spanning 4,568 acres. This region is comprised of mature and young stands which depict most of the tree species composition present throughout the Monongahela National Forest. Figure 2 depicts the location of the study area.

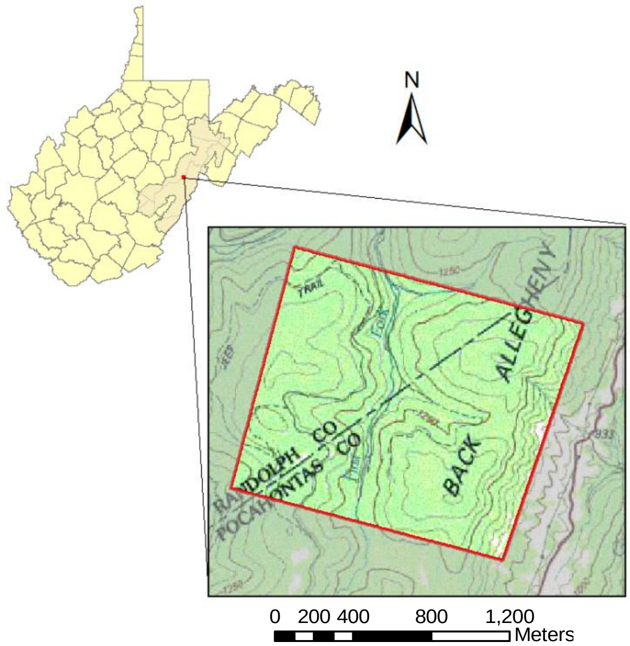


Figure 2: Study area in Randolph and Pocahontas county of West Virginia

4.3 Data Acquisition

The LiDAR data used for this study were collected in November of 2007 by Canaan Valley Institute using a small footprint, high density scanning system. The data were captured during the winter and is leaf off. This system acquires discrete multi-return LiDAR data. The Natural Resource Analysis Center (NRAC) pre-processed the data by removing outliers and characterizing LiDAR point data based on ground-truthed data points. Table 1 outlines the specifications for the LiDAR dataset. The LiDAR dataset was compared to the Landsat dataset to determine the full utility of using LiDAR to classify forest structures. Figure 3 depicts an overview and side view of the LiDAR imagery. Landsat thematic mapper was chosen because it is a readily available dataset which provides full coverage for the study area.

Table 1: Specifications for LiDAR Acquisition	
Acquisition Date	November 2007
Laser Scanner	ALTM 3100
Flying Height	1000 feet
Laser Pulse Density	1 meter
Max returns per pulse	5 returns
Angle of Incidence	18 Degrees

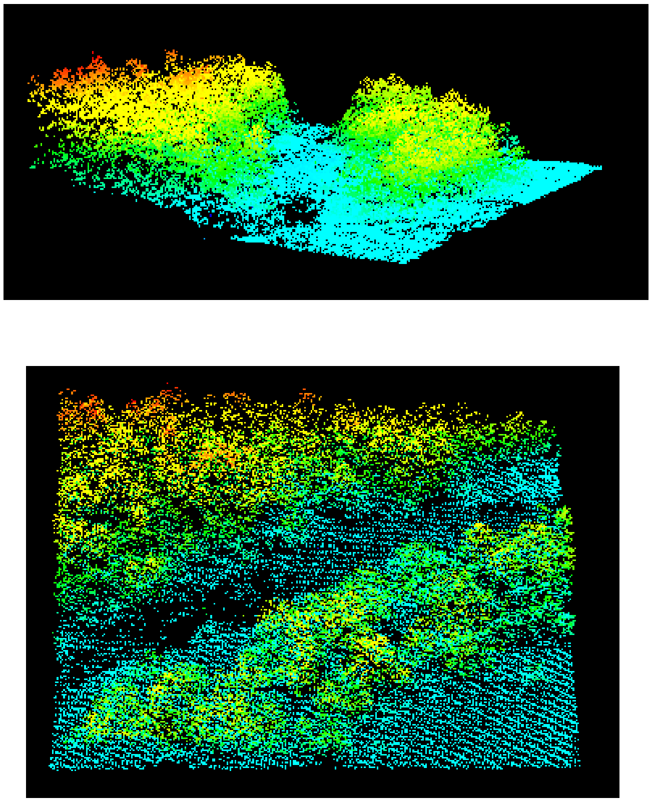


Figure 3: Top; side view of study area in LiDAR. Bottom; Overview of study area in LiDAR. Imagery depicts the LiDAR point clouds for the study area.

The Landsat data used for this study were captured from the United States Geological Survey (USGS). This scene was captured on March 3rd, 2008. This image was captured during late winter/ early spring, which indicates that the trees are in the beginning stages of sprouting new leaves. The date for this scene was chosen based on its closeness in date to the LiDAR acquisition while having zero percentage of cloud cover present. The LiDAR and Landsat acquisition does differ, but it does not make a significant difference. Stereńczak (2010) reported that seasonal changes do have an influence on LiDAR values, but the influence measured was significantly small. Figure 4 depicts the Landsat imagery for the study area. This study is only focusing on height and densities while disregarding differences in vegetation types. The presence of clouds in imagery creates scattering which can affect the data collected by skewing the digital numbers (Campbell, 2002). This Landsat image was radiometrically enhanced based on the first level of standard terrain correction (Level 1T). This type of correction rectifies an image by removing random radiometric noise (Meyer et al., 1993). In addition, the level 1T corrects for geometric errors. Table 2 outlines the specifications for the Landsat dataset.

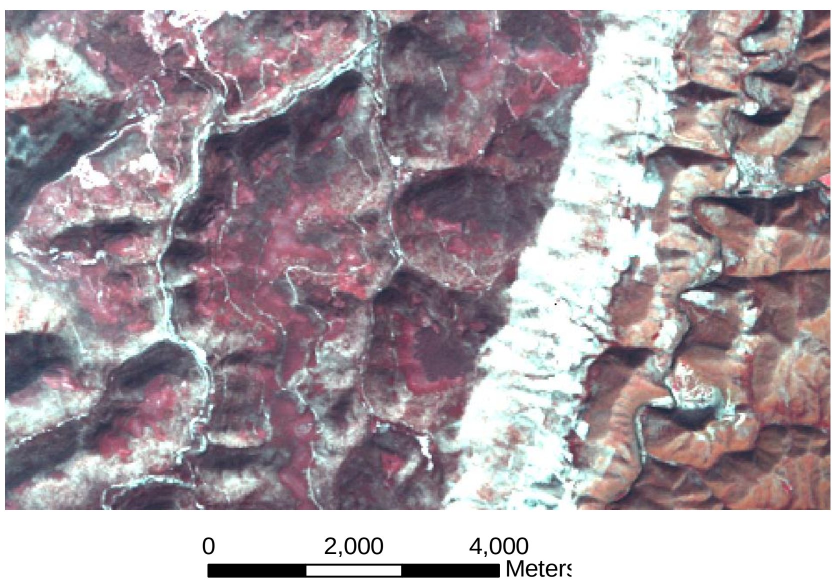


Figure 4: Landsat image of study area. Image acquired from USGS Earth Resources Observation and Science (EROS) center (http://eros.usgs.gov/).).

Table 2: Specifications for Landsat Acquisition	
Acquisition Date	March 3rd 2008
Sensor	Landsat Thematic Mapper 5
Resolution	30 meters
Cloud Cover	0%
Correction	Standard Terrain Correction (level 1T)

4.4 Pre-Processing

The processing procedures of LiDAR and Landsat data used in this research were derived from literature and previous studies depicting the uses of remote sensing techniques in forestry management and inventory (Brandtberg et al., 2003, Tiede et al., 2007, Koukoulas & Blackburn, 2005; Kwak et al., 2007). Canopy height, percent canopy cover, stem density, and the normalized vegetation difference index are commonly used variables for vegetation and forestry analysis (Kwak et al., 2007, Kim et al., 2008, Koukoulas & Blackburn, 2005; Pascual et al., 2010). Algorithms, such as canopy height models and canopy cover models were used to transform the LiDAR and Landsat data into products used for the segmentation and classification process. A methodology was developed that compares LiDAR and Landsat imagery for the classification of forest structures. This process gains insight into the effects of resolution and sensor type on identifying structural elements of a forest.

Figure 5 depicts the steps of analysis for the comparing the results of LiDAR and Landsat data for structural classification. The flow chart describes the steps, processes, and software used to create the LiDAR and Landsat products for segmentation and classification. The data pre-processing and analysis are divided into four main steps. Table 3 outlines the variables derived from the LiDAR and Landsat imagery.

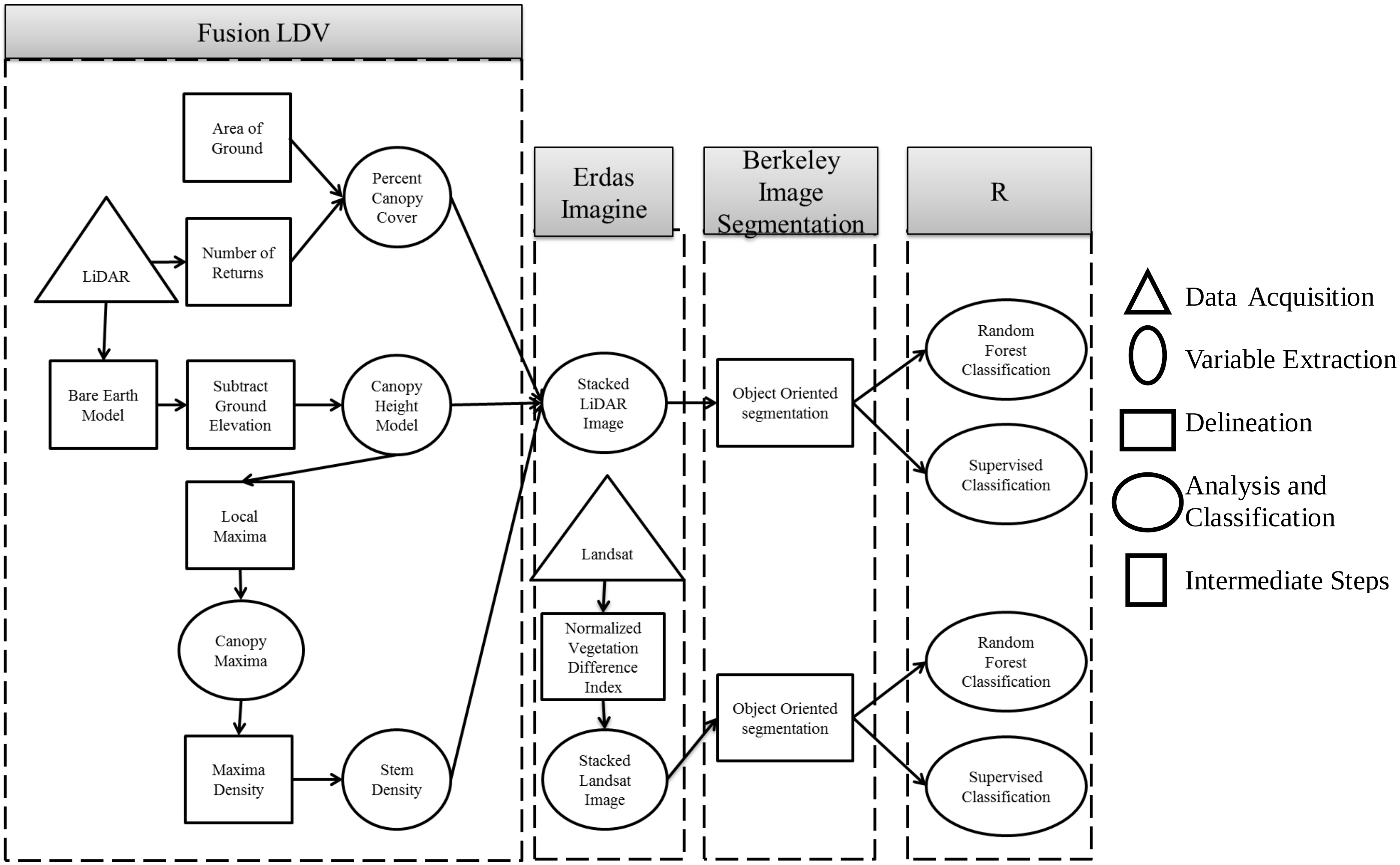


Figure 5: Flow Chart of Pre-processing process

Table 3: Summary of Variables	
Canopy Height	Height of tree from base to canopy.
Canopy Cover	Portion of ground that is covered by trees in a specified area.
Stem Density	Number of plants in a specified area.
Normalized Difference Vegetation Index	Measure of “greenness” in vegetation

4.5 Variable Extraction

The first stage of this study was to complete the pre-processing steps that resulted in a stacked LiDAR and Landsat image. These resulting images are represented in Figure 6. The Fusion LDV 2.90 was used for the pre-processing steps for the LiDAR dataset. This software is a visualization system that transforms and analyzes LiDAR data while providing a viewer that displays LiDAR in its native form. This step resulted in the percent canopy cover, canopy height model, and stem density raster datasets. To obtain the resulting imagery raw LiDAR data were used in the first three steps of the data processing. The output from FUSION LDV was viewed and stacked within Erdas Imagine to create one image. Erdas Imagine 2010 is software used for advance remote sensing analysis. The resulting stacked image was then segmented based on object oriented rules within the Berkeley Image Segmentation software. This software uses a “region merging technique to obtain a complete spatial partition of the input image pixels” (Clinton et al., 2010). Once the image was segmented, each section was classified using random forest and supervised classification techniques.

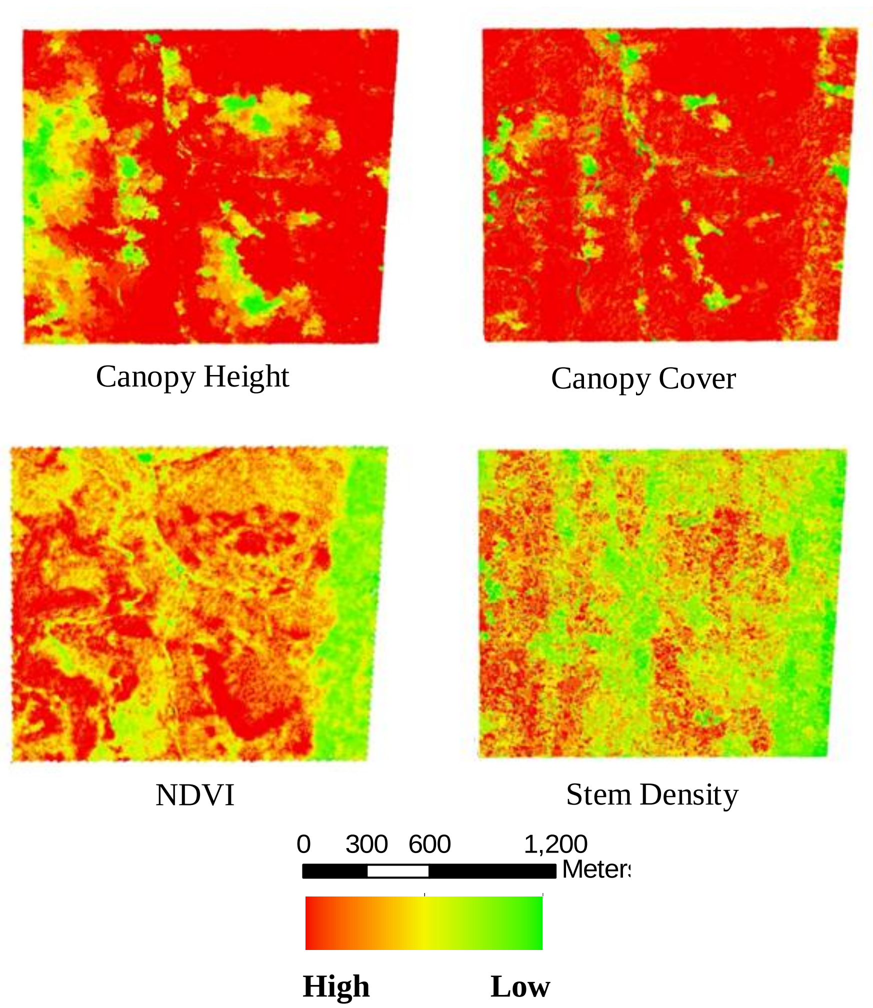


Figure 6: pre-processing images for the Landsat and LiDAR data. Each image is rotated 15°.

In addition the bulk of the Landsat data pre-processing was conducted within Erdas Imagine software package. Similar to the LiDAR data, the resulting stacked Landsat image was segmented within the Berkeley Image Segmentation software package then classified using random forest and supervised classification.

First, a canopy height image was created to determine tree heights. Canopy height is a 3D representation of a forest canopy with regard to height and shape for a resolution of 1 meter. It was obtained by subtracting the ground elevations from the first returns within the LiDAR data (Reutebuch et al., 2005). It was created from raw LiDAR data and a bare earth model. The bare earth model allows the CHM to be used for the comparison of tree heights without the effect of elevation (McGaughey, 2010).

Second, the stem density image is created to determine the percentage of vegetation present per 1 meter pixel, which is created by obtaining the amount of local maxima points in each cell. Once this count is obtained the stem density algorithm assigns a percentage which is the relative density between the forestry stands. A local maximum refers to a maximum point or height within a certain neighborhood. Essentially, the local maxima, which are generated from the CHM, will represent tree tops. These tree tops are created by an algorithm similar to the work of Popescu et al. (2002) and Popescu and Wynn (2004), which uses variable window sizes based on the CHM to detect local maximas (McGaughey, 2010). The window size used changed based on forest stand and maturity level being analyzed.

Third, canopy cover is created to determine the percentage or density of the tree canopy present for each pixel. This model uses the raw LiDAR data and the ground model to estimate these percentages using the Fusion LDV software. The first-returns over a height break of 3 meters were parameters for the canopy cover model algorithm. The height break refers to the height at which all vegetation above the specified height is included in the analysis. This height break was chosen to exclude any small brush which could potentially skew the resulting image. In addition the pixel size for the canopy cover model is 15 meters. According to McGaughey (2010), the pixel for the canopy cover algorithm must be wider than individual tree crowns. Smaller pixel values will skew the output of the canopy cover model by placing more emphasis on areas with an absence of tree crowns and an area with an abundance of tree crowns.

Last, the NDVI image estimates the percentage of vegetation cover from the reflective bands of the original Landsat TM image. This index allows for the detection of vegetation throughout the pre-processing process. It uses the near infrared and red bands of Landsat TM to find the difference in brightness values for vegetation. The brightness values represent the abundance of vegetation present of a scale of -1 to 1 (Tucker, 1979). Pascual et al. (2010) reported that there is a correlation between LiDAR heights and NDVI. This correlation allows for comparison between the Landsat and LiDAR imagery for the forest structure classification.

Initially, the principal component analysis (PCA) was to be applied to the stacked LiDAR and Landsat imagery. This process transforms a set of images into a new set that have less correlation between components. The first component from the principal component analysis can be used to guarantee the highest amount of separability or variance between the image components (Ricotta et al., 1999). But, according to Mutlu & Popescu (2006), PCA decreases the accuracy of remotely sensed data for classification. Therefore, the data reduction for this study was conducted by limiting the amount of Landsat bands and Landsat imagery being used in the analysis. Rather than employing the PCA transformation, data reduction proceeded by limiting the amount of data applied to the stacked Landsat and LiDAR imagery.

4.6 Analysis

4.6.1 Object Orientated Segmentation

Once the pre-processing steps were completed, the object based image analysis was employed. The object oriented segmentation process, developed by the Berkeley Environmental Technology, partitions an image into a set of objects that provide a visual representation in the form of a vector. *BerkeleyImageSeg* utilizes the region merging technique for its segmentation algorithm (Clinton, 2010). Region merging is the process of eliminating false boundaries and regions by combining neighboring objects of the same characteristics defined by decision rules. This technique utilizes hypothesis testing for the probability that each region or object will statistically have similar distributions of intensity values (Harris et al., 1998). Each image is segmented based on a set of predefined criteria which includes threshold, shape, and compactness. Threshold refers to the iteration of the merging cycles. Essentially, the higher the threshold value the larger the objects will be due to the increased amount of merging. Shape refers to “a value that describes the improvement of the shape with regard to smoothness and compactness of an object’s shape” (Benz et al., 2004, p. 246). Compactness refers to the texture of the image and its ability to maintain smooth edges for the segmentation process.

Three parameter combinations were examined for threshold, shape and compactness: {10, 05, 05}, {30,05,05}, {50,05,05}, respectively. The shape and compactness values remained the same for each segmentation process so the comparison of each combination would be solely on the size of each object. The parameters for the segmentation process were chosen based on a visual inspection of size and scale. Each segment resulted in objects that represent tree crowns, cleared areas, and sparse vegetation. Ultimately, the {30, 05, 05} segmentation gave the best representation of the area from a visual point of view. Representations of the segments are displayed in Figure 7 and 8.

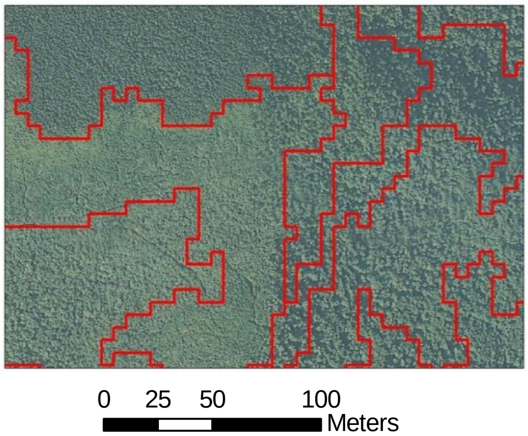


Figure 7: Object oriented segmentation of Landsat image overlaid on NAIP imagery. Image zoomed in to show size and texture.

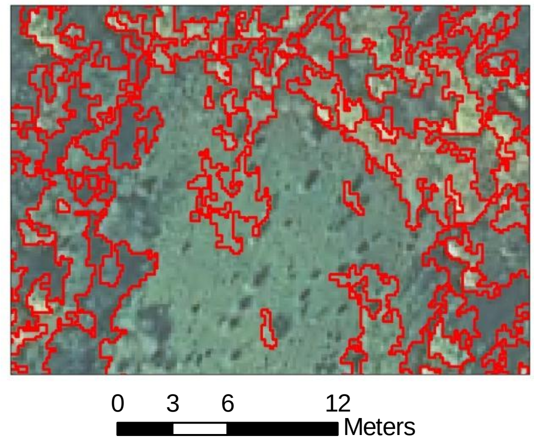


Figure 8: Object oriented segmentation of LiDAR image overlaid on NAIP imagery. Image zoomed in to show size and texture.

4.6.2 Random Forest Analysis

Succeeding the object oriented segmentation process, a random forest analysis was applied to the LiDAR and Landsat stacked and segmented images. A random forest analysis is a classification and regression tree (CART) analysis which grows many classification trees versus one. Random forest “reputably delivers considerable robustness to noise, outliers, and overfitting when compared to a single tree classifier” (Williams, 2009). This algorithm is a non-parametric decision tree learning method. Decision trees are created by rules that are based on the variables in the training dataset. These rules are selected to obtain the best split amongst values to differentiate between observed classifications based on the dependent variables. The random forest process stops when no further splits can be detected. Each decision tree has a series of child and terminal nodes which are used to reflect the recursive nature of the classification process. Each terminal node contains one class which is defined by the rules created.

To obtain the random forest classification a set of parameters were defined to analyze the dataset. The number of tress, number of variables, and sample size are parameters used in the analysis. According to Williams (2009), five hundred decision trees are an efficient amount and still avoids over-fitting. The sample size was determined by a random sampling of one third of the dataset.

4.6.3 Supervised Classification

To compare the accuracy of the random forest algorithm a supervised classification was employed. This classification method extracts quantitative information from remotely sensed images (Richards, 1993; Strahler, 1980). This process uses *priori* knowledge to generate representative parameters or areas of interest for each class. Creating these areas of interest is referred to as *training*. The maximum likelihood algorithm was used to statistically analyze the LiDAR and Landsat data in order to provide a correct classification. The maximum likelihood algorithm is a usual method for classification in remote sensing. This method assumes that each class in each band is normally distributed, which allows for each pixel to be assigned a probability that it belongs to a class. Each pixel is assigned to the class with the highest probability (Richards, 1993). The training for each class was obtained by locating spectral signatures within segments which were digitized based on priori knowledge. This knowledge is a form of empirical knowledge which is defined by observation (Barthelemy, 1985). This method minimizes a potential problem associated with speckle amongst classes.

4.6.4 Accuracy Assessment

An accuracy assessment compares derived classifications with ground-truth or reference data. This method is accomplished by evaluating how well the classifications represent "the real world." The accuracy assessment allows a classified map to be used for more than a reference image. The accuracy of the random forest analysis and supervised classification were evaluated with error matrices. An error matrix is a common procedure for the accuracy assessment for imagery classification. It depicts the usefulness of a classification by assessing the user's, producer's, and overall accuracy. The user's accuracy refers to the error of commission. This error investigates the usefulness of the classified map from the perspective of the user. This error represents the amount of pixels that are dedicated to an incorrect class. In addition, the producer's accuracy refers to the error of omission. This error investigates the usefulness of the classified map from the perspective of the map maker. This represents the number of pixels that are labeled correctly on the map. Also, the overall accuracy represents the average between user's and producer's accuracy. These accuracies are presented in an error matrix which clearly displays the classes which correctly and incorrectly identified pixels (Rossiter, 2004).

To perform the accuracy assessment analysis a systematic sampling method was employed. With the systematic sampling method there is a possibility of missing sites due to the evenly spaced pattern. However, this method avoids the issue of low sample concentration when compared to random sampling methods. For this technique, a uniform grid of 1850 sample points with a spacing of 100 meters was created. Traditional approaches to sample size suggest mathematical solutions which yield enormous sample sizes and require additional effort. Congalton & Green (1999) suggest collecting a minimum of 50 samples for each class. This study exceeded the minimum sample size to ensure that each class was represented homogeneously. The amount of points per classification varies due to the difference in each reference image. The resulting grid was overlaid on the classified Landsat and LiDAR

4.7 Reference Images

A set of reference images were created by utilizing criteria proposed by studies such as Photoscience (2011), Sullivan (2008), and Roy et al. (1996). This method was implemented due to a lack of primary data which would provide a source of ground-truthing. The use of reference imagery is a form of secondary data which is not 100% accurate. This disadvantage has the ability to decrease the overall accuracies which are presented in this study. If any errors are present in reference imagery then any correctly classified pixels may be incorrectly assigned. In addition, the use of reference imagery could introduce a conservative estimate of the classification accuracy. This error is attributed to cell size and what a cell truly represents on the ground. These cells only represent one class whereas it may represent many classes at a smaller scale (Verbyla and Hammond, 1995).

4.7.1 Reference Image 1

The first reference image was created utilizing the criteria presented by PhotoScience (2011) and applied to the 2007 LiDAR dataset for the Monongahela National Forest established plot information for the study area. Photoscience (2011) provided a baseline for the composition and canopy characteristics for the Monongahela National Forest. The parameters for the variables given were applied to the LiDAR imagery to create a guideline for the accuracy assessment. A series of conditional statements was used to create reference image 1. Appendix 1 includes detailed information regarding the habitat variables. The *West Virginia Northern Flying Squirrel (WVNFS) Vegetative Habitat* project analyzed 5 million acres for forest cover type, size class, canopy cover, and crown condition. Each plot was analyzed using 2003 leaf off aerial photographs.

To define the habitat types forest cover, size class, canopy cover, and crown condition was measured. Forest cover represents spruce, northern hardwoods-conifer, conifer, and other vegetative types. Size classes refer to seedlings/saplings, pole-timber, saw-timber, and mature trees. Canopy cover describes the percentage of the canopy covering the minimum mapping unit of 5 acres. The values for canopy cover are grouped into intervals that are less than 10%, 10-29.9%, 30-59.9%, 60-79.9%, and 80-100%. In addition, the tree condition for the minimum mapping unit is described by less than 15% of tree mortality, between 15% and 20% tree mortality, and over 50% tree mortality (PhotoScience, 2011). For this study, canopy cover, canopy height, and size classes were used and translated into the existing variables. The variables that were measured for each plot were used in the classification criteria for this thesis research.

The parameters from the study above were combined with the naming convention presented by Sullivan (2008) and the canopy heights which are depicted in McGaughey (2010). The parameters are as follows:

Mature/ Old Growth:

Vegetation density >25 percent
Canopy height >70 feet
Canopy density >50 percent

Young/Understory 1:

Vegetation density <25 percent
Canopy height >20 and <70 feet
Canopy density<50 percent

Young/Understory 2:

Vegetation density <25 percent
Canopy height >1and <20 feet
Canopy density<50 percent

Thinned/Clearcut:

Vegetation density <25 percent
Canopy height <4 feet
Canopy density<25 percent

4.7.2 Reference Image 2

The second reference image was created by utilizing criteria proposed by Roy et al., (1996). This criterion was applied to Landsat thematic mapper imagery captured in November of 2003. The *Stratification of density in dry deciduous forest using satellite remote sensing digital data* analyzed a central part of India, which is a hot and dry climate. This climate is different from the temperate climate of West Virginia. Because of the differences in climate, the spectral reflectances will differ. However, the imagery used for the Roy et al., (1996) was leaf off and both forests have the same structural elements. Thus, the same criteria can be used. They utilized Landsat thematic mapper data from January 1990. They presented a model which utilized many vegetation indices which are: advanced vegetation index (AVI), bare soil index (BI), canopy shadow index (SI), middle infrared, and normalized vegetation index (NDVI). These indices were used to classify landcover based on a stratification of density. Only the SI and AVI were utilized for forest density. To have continuity between both reference images, the same naming convention was applied. The following criteria for a rule-based approach for forest density classifications are as follows:

If SI is between 177-192 and
    AVI is between 0-22, then non vegetated
    AVI is between 23-31, then fallow
    AVI is between 32-40, then grass
    AVI is between 41-69, then crop **Thinned/Clearcut**

If SI is between 193-203 and
    AVI is between 21-53, then grassland
    AVI is between 59-93, then crop

If SI is between 204-214 and
    AVI is between 19-37, then scrub **Young/Understory 2**
    AVI is between 38-55, then forest 10-20%

    AVI is between 56-66, then forest 20-40%
    AVI is between 67-135, then forest 40-60% **Young/Understory 1**

If SI is between 215-224 and
    AVI is between 17-56, then forest 60-80% **Thinned/Clearcut**
    AVI is between 56-128, then forest > 80%

CHAPTER 5

Results and Discussion

This study resulted in a set of classified maps which can be compared to determine the best process possible for characterizing the structure of a forest. Four different classification approaches were implemented. The results are presented in Figure 9. A maximum likelihood and a random forest classification were applied to Landsat and LiDAR imagery. Each resulting image produced the same classification scheme but, differences in resolutions and sensor types resulted in different classification accuracies. In addition, the results for both LiDAR datasets were resampled to better compare it with the Landsat imagery. This process revealed that the random forest classification process is more robust when compared to the supervised classification. In addition, these classification processes have revealed a notable difference between the classification accuracies of the Landsat and LiDAR imagery.

5.1 Classification Results of Supervised Classification

Initially, a visual inspection of the supervised classification results for both Landsat and LiDAR data were conducted. The resulting LiDAR image depicts more detail in the clearcut/thinned class by deciphering roads in the study area. Also, this image displays more “speckle” within each class. Speckle is a type of noise which degrades the quality of the data. In contrast, the resulting Landsat imagery depicts less detail but displays fewer speckle within each class. Both images have the same general placement of young 1 and young 2 classes but the Landsat image embellished the quantity available. Overall, the differences in the Landsat and LiDAR supervised classifications can be attributed to the difference in cell size and the nature of the supervised classification being a pixel by pixel classifier.

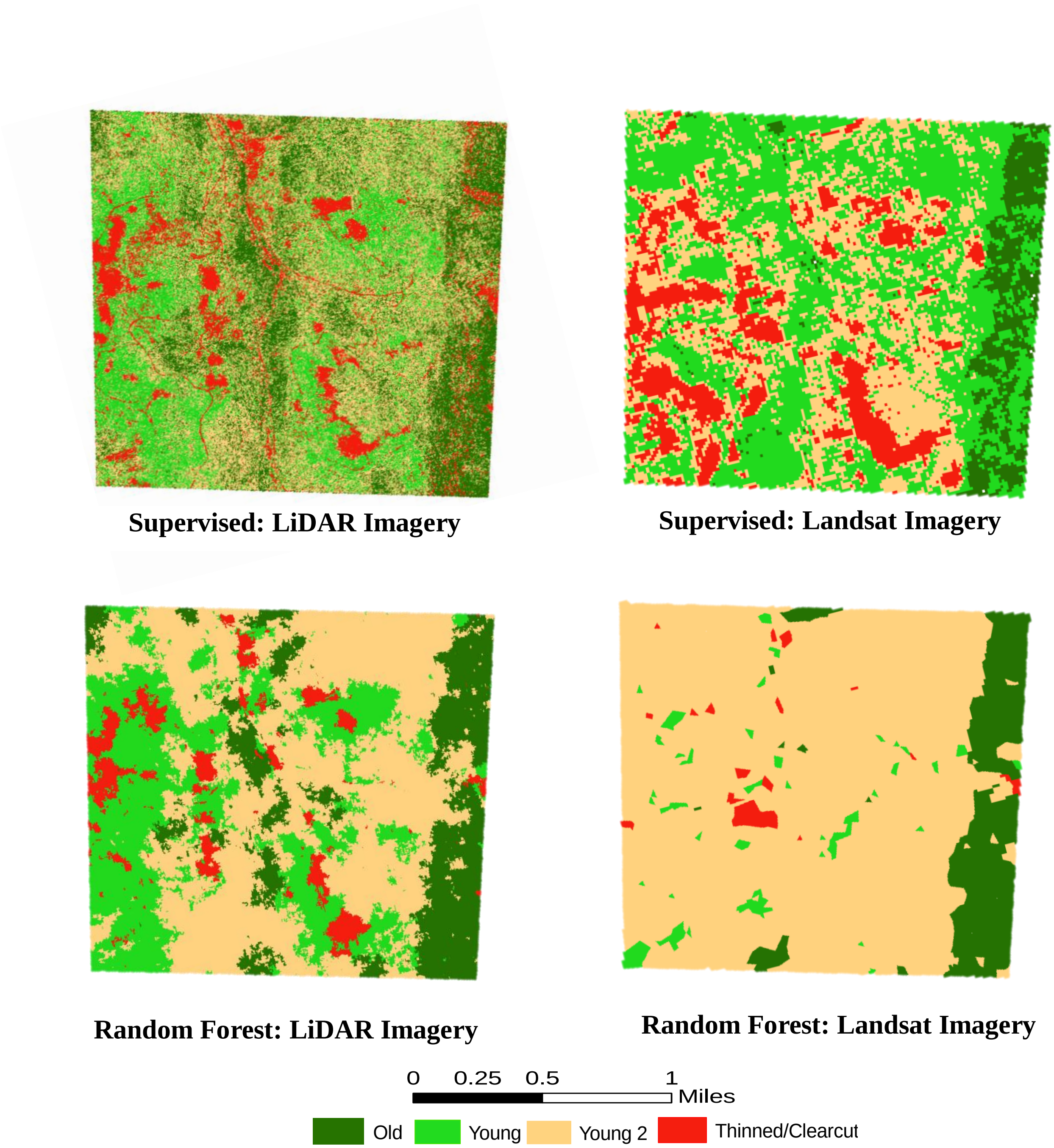


Figure 9: Supervised and random forest classification of LiDAR and Landsat data. Each image is rotated 15°.

5.2 Classification Results of Random Forest Classification

Similarly, the resulting images of the random forest classification were visually inspected. These images depict less detail when compared to the supervised classification method. The decrease in detail is attributed to the object oriented segmentation. This segmentation process used on the LiDAR and Landsat imagery did not distinguish smaller features, such as roads. In addition, the object oriented segmentation process resulted in different size segments for each type of data. This error can be attributed to cell size but, the nature of how the data are collected and portrayed is the biggest contributing factor. The Landsat imagery was collected via a NASA Satellite using the electromagnetic spectrum. This imagery gives a varied digital number for any surface. In contrast, the LiDAR imagery was collected via an airborne sensor, which provides the LiDAR data with very dense information which has the ability to capture heights. For this thesis research, the object oriented segmentation process produced better segments for the LiDAR imagery because of its ability to provide data for smaller areas on earth, as well as its ability to capture heights which directly correlates to structures involving height.

Naturally, the random forest classification process resulted in a more aesthetically pleasing image for the LiDAR dataset opposed to the Landsat. The success of the random forest process depends on the success of the object oriented segmentation. Because, the LiDAR imagery produced smaller amounts a more diverse set of classifications could be produced, thus providing higher classification results.

5.3 Resampling of LiDAR

To better compare both datasets, the LiDAR imagery was resampled to 30 meters to match the Landsat imagery, which allows for a direct comparison. The image was resampled with a bi-linear method. This resampling method uses a weight average of a group of pixels to create a new image. According to Suwendi and Allebach (2008), the bilinear method results in a sharper image without the blocky appearance that is a result of the nearest neighbor resampling method. The resulting resampled images produced more generalized results. Figure 10 depicts the resampling of each image.

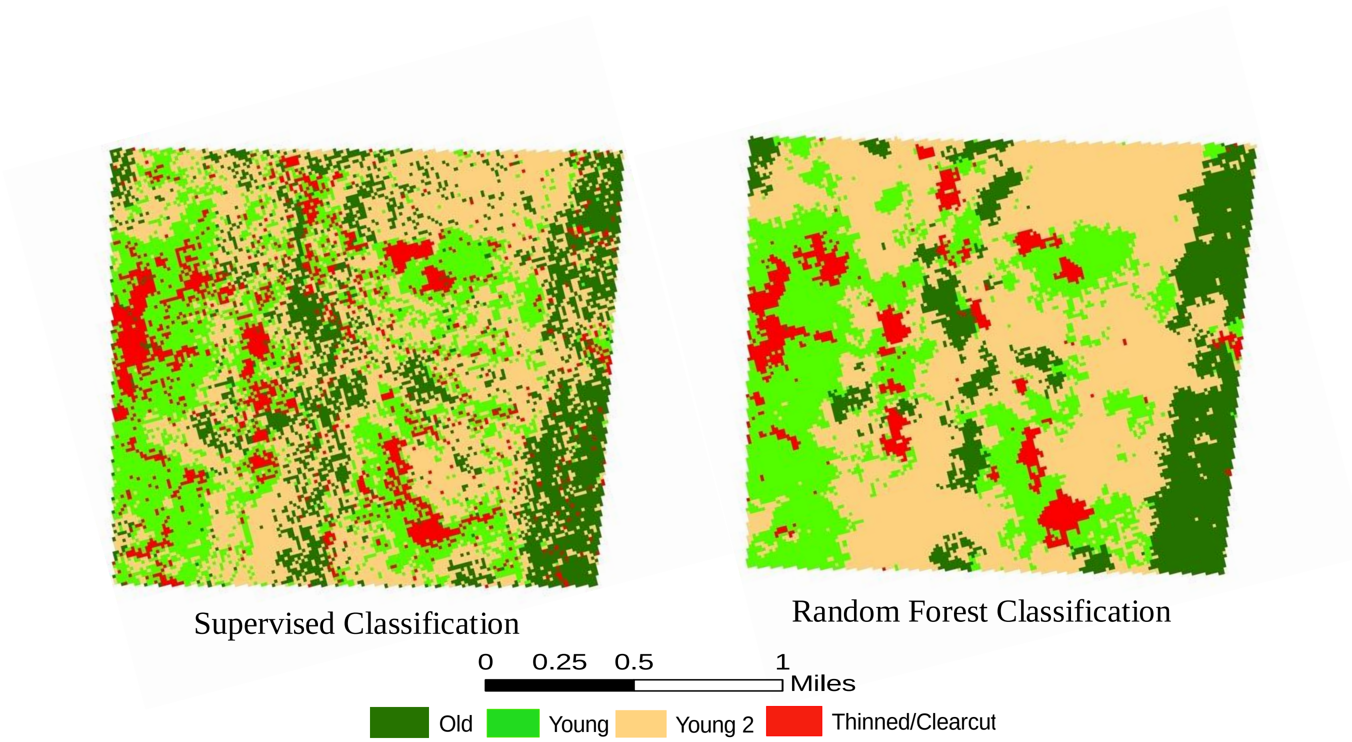


Figure 10: Images of resulting resampling of LiDAR data to 30 meters. Each image is rotated 15°

5.4 Accuracy Assessment

The classification accuracy for forest structures was analyzed using an error matrix for each resulting image. First, an accuracy assessment was performed on all of the imagery using both reference images. This process defines which reference image best represents the study area. The accuracy results, presented in Table 4, revealed that the random forest classification with the object oriented segmentation better distinguished between the forest cover classes. Furthermore, the LiDAR imagery presented higher accuracies for both reference images and classification techniques. “Current techniques used for forest stand delineation are variable across landowners and are expected to produce accuracies of about 80% to 90%” (Sullivan, 2008). The accuracies for reference image 1 range from 31.3% to 80.1% whilst the accuracies for reference image 2 range from 26.8% to 32.0%. In addition, the overall accuracies for the LiDAR imagery range from 26.8% to 80.1%, whereas the overall accuracies for the Landsat imagery range from 30.0% to 50.2%.

The resampled LiDAR image provided better results for the supervised imagery, whilst the percentages decreased for the random forest image. These results prove that resampling an image does favor the pixel based classifier versus the object oriented segmentation classifier. The bi-linear resampling method combines pixels into groups. The object oriented segmentation already grouped the similar pixels, so the resampling method grouped pixels that were drastically different, thus resulting in lower accuracies.

This series of accuracies is intended to provide a baseline for the comparison of imagery for classification, reference imagery, and classification techniques. Differences in each classification method were expected due to the differences in each remote sensing datasets. The differences in classification accuracies for each reference image were highly variable. In fact, there is a 48.1% difference between the highest accuracies of each reference image.

Table 4: Results of confusion matrix which represents the percentage of points in the accuracy assessment grid that was accurately and in-accurately classified during each method.

Reference Image 1								
	LiDAR				Landsat			
	Supervised		Random		Supervised		Random	
	Producer	User	Producer	User	Producer	User	Producer	User
Oldgrowth	60.5	51.7	69.8	83.7	24.0	60.1	54.7	72
Young 1	53.7	55.1	80.2	79.2	26.5	16.6	3.8	44.2
Young 2	51.8	58.3	87.4	78.9	33.7	42.4	90.5	46.5
Clear-cut/Thinned	57.4	49.4	71.8	82.4	46.1	34.0	5.6	42.9
Overall	**54.7**		**80.1**		**31.3**		**50.2**	

Reference Image 2								
	LiDAR				Landsat			
	Supervised		Random		Supervised		Random	
	Producer	User	Producer	User	Producer	User	Producer	User
Oldgrowth	19.9	60.0	41.0	85.7	62.9	59.1	87.2	59.2
Young 1	0.0	0.0	0.0	0.0	0.0	0.0	0.0	0.0
Young 2	2.6	30.9	1.3	21.8	0.0	0.0	3.7	0.1
Clear-cut/Thinned	87.4	13.1	99.1	7.2	100.0	18.4	1.6	92.6
Overall	**26.8**		**32.0**		**30.0**		**30.5**	

LiDAR Resampled								
	Reference Image 1				Reference Image 2			
	Supervised		Random		Supervised		Random	
	Producer	User	Producer	User	Producer	User	Producer	User
Oldgrowth	71.6	70.7	56.0	64.0	72.1	28.0	84.8	39.6
Young 1	71.8	70.0	61.8	60.6	0.0	0.0	0.0	0.0
Young 2	75.4	76.9	75.2	63.8	26.4	1.9	18.9	1.1
Clear-cut/Thinned	58.8	59.6	33.5	64.3	12.6	95.7	6.8	98.2
Overall	**69.3**		**59.9**		**39.6**		**31.2**	

5.4 Discussion

The methods outlined in this thesis research are intended to be applied to aerial or airborne remote sensing imagery to produce products which depict vertical structures of a woody vegetated area. The final maps will be used to determine which classification method and which remote sensing dataset are better suited for identifying structural elements of a forest. In addition, the resulting images are intended to be used for planning and management purposes.

The user's accuracy impacts the overall quality of the product from the end user's perspective. Overall, most of the user's accuracies were below the industry standard of 80%. However, all of the supervised classification accuracies were below this percentage. In addition, all of the classification accuracies for reference image 2 were below the industry standard. Similar to Chen (1999), the speckled imagery that resulted from this classification method is a contributing factor to the low user's accuracies. This is due to the nature of the object oriented segmentation. The segmentation allows for the classification of a group of pixels versus individual pixels.

In addition, the young 1 and young 2 classes were the hardest classes to decipher amongst both classification methods. According to Sullivan (2008), the confusion between the "young 1" and "young 2" classifications are attributed to presence of gaps each study area as well as the similarity between the two classes. Both of these classes represent the mid-story of the forest but symbolize different heights. This is a contributing factor to the low accuracies for these classes. Conversely, if the two classes are combined into one "young" class the accuracies dramatically increase. For reference image 2, the accuracies of the new "young" class did not increase. Conversely, the accuracies for reference image 1 yielded accuracies ranging between 76.1 to 94.2%.

Overall, the user's accuracies were higher for the random forest classification of the LiDAR imagery. In addition, LiDAR provides information which is directly attributed to height, which is a key element in classifying structural elements. This element directly attributes the higher accuracy of LiDAR when compared to Landsat. Furthermore, the percentages of old growth detected in each image ranges from 2.4 to 17.2%. The LiDAR random forest image had the highest percentage. With that being said, the LiDAR Imagery with the random forest classification and the object oriented segmentation yields higher results and provides a better process for forestry management. In addition, reference image 1 yields higher results for both Landsat and LiDAR imagery.

The resolutions vary for each dataset and contribute to the differences in sizes for each class within the classification. The overarching disparity is the availability of true height data in each dataset. LiDAR data does capture true heights which correlate to the resulting imagery. In contrast, the Landsat dataset provides density information which is based off of the reflectance of the vegetation, but this cannot provide true heights. The essence of classifying structural elements of the forest is combining variables which depict density and height.

CHAPTER 6

Conclusion

The research presented in this study contributes to the overall knowledge of LiDAR and forestry. The primary objectives of this thesis research were to determine if LiDAR data can classify structural elements of a forest when compared to Landsat imagery, as well as determining which classification method was more robust. The utility of the LiDAR data collected show the ability to extract variables which directly correlate to tree structures. The result of the supervised and the random forest classification shows that there is a significant difference between the use of LiDAR and Landsat data in forestry. The clustering of the young 1 and young 2 classes increased the overall accuracies of the datasets. When using the object oriented segmentation, the clusters of the Landsat imagery were not formed well during this process. It should be noted that the object oriented segmentation does fare well for the red and near-infrared bands used for the Landsat imagery. For the purpose of classifying forest structures from remote sensing imagery, using variables derived from LiDAR showed the best classification accuracies. In addition, the random forest classification with the object oriented segmentation proved to be the best classification method.

The outcome of this study warrants future research. The conceptual model and methodology presented allows for this study to be repeated for multiple remote sensing images. In addition, the study can be repeated for different forest types. This model consisted of 4 main stages: conversion, segmentation, reduction, and the classification of the data. Also, the fusion of variables from LiDAR and Landsat can also be explored. This would allow information about height and density to be coupled with information from the electromagnetic spectrum. In addition the use of the intensity values captured during the LiDAR acquisition could provide additional information which could further aid in the classification of forest structures.

Although this study has reached its objectives, there were unavoidable limitations. The reference imagery being used for the accuracy assessment was created in an unconventional method. Both reference images used in this study were created with the current LiDAR dataset and a different Landsat dataset using parameters derived by another study. Initially, a set of ground truthed points were to be provided by the USFS and the WVDNR. Unfortunately, these points were not acquired due to the lack of information from both agencies. To combat this limitation, future research will create a new reference image based on field data. The importance of new reference imagery is eminent, because it allows for the repeatability of the proposed methodology with ground truthed data.

The additional limitation refers to the temporal and spatial difference between the LiDAR and Landsat data. First, the data were collected at significantly different resolutions. These spatial difference contributes to how detailed the imagery will be. Secondly, the images were collected during different seasons. There is a four-month difference, which resulted in the LiDAR imagery being collected in mid-winter and the Landsat imagery in early spring. The differences in temporal resolutions could contribute to the difference in density of vegetation due to the time in which deciduous trees regaining their leaves.

Although there have been many research studies in the realm of LiDAR and its uses in forestry, little research focuses on classification of the structure of these forests. As forests become more threatened by climatological factors, there will be a need for speedy analysis of these forests using remote sensing technologies. The monitoring of these montane forests is important for the preservation and conservation of beautiful old-growth trees.

.

Bibliography

Adams, H., Stephenson, S., Rollins, A., & Adams, M. (2009). The Isolated Red Spruce Communities of Virginia and West Virginia. Ecology and Management of High Elevation Forests in the Central and Southern Appalachian Mountains (pp. 1-12). Slatyfork, West Virginia: United States Department of Agriculture.

Ahokas, E., Kaasalainen, S., Hyypa, J., & Suomalainen, J. (2006). Calibration of the Optech ALTM 3100 laser scanner intensity data using brightness targets. *ISPRS Journal of Photogrammetry and Remote Sensing*.

Ayhan, E., & Kansu, O. (2010). Analysis of image classification methods for remote sensing. *Experimental Techniques, 36*, 18-25.

Battles, J., & Fahey, T. (2000). Gap Dynamics Following Forest Decline: A Case Study of Red Spruce Forests. Ecological Applications, 10 (3), 760-774.

Barthelemy, W. (1985). *The structure of empirical knowledge* . Cambridge: Harvard University Press.

Benz, U., Hofmann, P., Willhauck, G., Lingenfelder, I., & Heynen , M. (2004). Multi-resolution, object-oriented fuzzy analysis of remote sensing data for GIS ready information. *ISPRS Journal of Phogrammetry & Remote Sensing, 58*, 239-258.

Boyd, D., & Hill, R. (2007). Validation of airborne LiDAR intensity values from a forested landscape using hymap data: Preliminary Analysis. *ISPRS Journal of Photogrammetry & Remote Sensing*.

Brandtberg, T., Warner, T., Landenberger, R., & McGraw, J. (2003). Detection and analysis of individual leaf-off tree crowns in small footprint, high sampling density LiDAR data from the eastern deciduous forest in North America. *Remote Sensing of Environment, 85,* 290-303.

Brennan, R., & Webster, T. (2006). Object-oriented land cover classification of LiDAR-derived surfaces. *Canadian Journal of Remote Sensing, 32*(2), 162-172.

Byers, E., Vanderhorst, J., & Streets, B. (2010). Classification and Conservation Assessment of Upland Red Spruce communities in West Virginia. Elkins, West Virginia: West Virginia Division of Natural Resources, Wildlife Resources Section.

Campbell, J. (2002). Introduciton to Remote Sensing (3rd Edition). New York, NY: The Guilford Press.

Chen, C.-h. (1999). *Information processing or remote sensing.* River Edge, New Jersey: World Scientific Publishing Co. Pte. Ltd. .

Clinton, N. (2010). BerkeleyImageSeg User's Guide. *CiteSeerX*.

Clinton, N., Holt, A., Scarborough, J., Yan, L., & Gong, P. (2010). Accuracy assessment measures for object-based segmentation goodness. *Photogrammetric Engineering & Remote Sensing, 76*(3), 289-299.

Congalton, R. (1991). A review of assessing the accuracy of classifications of remotely sensed data. *Remote sensing Environment, 37*, 35-46.

Congalton, R., & Green, K. (1999). *Assessing the accuracy of remotely sensed data: principles and practices.* Boca Raton, FL: Lewis Publishers.

Dale, V., Joyce, L., McNulty, S., Neilson, R., Ayres, M., Flannigan, M., et al. (2001). Climate Change and Forest Disturbances. BioScience3, 51(9), 732-734.

Dubayah , R., & Drake , J. (2000). LiDAR remote sensing for forestry. *Journal of Forestry, 98*(6), 44-46.

Freitas, S., Mello, M.., Marcia, C., & Cruz, C. (2005). Relationships between forest structure and vegetation indices in Atlantic Rainforest. *Forest Ecology and Management, 218*, 353-362.

Geneletti, D., & Gorte, B. (2003). A method for object-oriented land cover classification combining Landsat TM data and aerial photographs. *International Journal of Remote Sensing, 24*(6), 1273-1286.

Harris, K., Efstratiadsi, S., Maglaveras, N., & Katsaggelos, A. (1998). Hybrid image segmentation using watersheds and fast region merging. *IEEE Transactions on Image Processing, 7*(12), 1684-1699.

Houghton, J., Meira-Filho, L., Callander, B., Harris, N., Kattenberg, A., & Maskell, K. (1996). Climate Change 1995: The Science of Climate Change. New York: Cambridge University Press.

Johnson, A., Cook, E., & Siccama, T. (1985). Climate and Red Spruce Growth and Decline in the Northern Appalachians. Applied Biology, 85, 5369-5373.

Kim, S., Mcgaughey, R., & Anderson, H. (2008). Individual tree species identification using LIDAR intensity data. *ASPRS.*

Koukoulas , S., & Blackburn, G. (2005). Mapping individual tree location, height, and species in broadleaved deciduous forest using airborne LIDAR and multi-spectral remotely sensed data. *International Journal of Remote Sensing, 26*(3), 431-455.

Kwak, D., Lee, W., Lee, J., & Gong, P. (2007). Detection of individual trees and estimation of tree height using LIDAR data. *Journal of Forest Research, 3*, 425-434.

Leckie, D. (1990). Advances in remote sensing technologies for forest surveys and management. *Canadian Journal of Forest Research, 20*(4), 464-483.

Lee, J., Grunes, M., Ainsworth, T., Hainsek, L., Mette, T., & Papathanassiou, K. (2005). Forest Classification Based on L-Band Polarimetric and Interferometric SAR Data . *Proceedings of the 2nd International Workshop POLINSAR (ESA SP-586)*, ISBN: 92-9092-892-2.

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

Lo, C., & Chen, L. (2008). Canopy extraction using airborne laser scanning data in forestry areas. *International Society for Photogrammetry and Remote Sensing*.

Loeb, S., Tainter, F., & Cazares, E. (2000). Habitat associations of hypogenus fungi in the southern Appalachians: Implications for the endangered northern flying squirrel (Glaucomys sabrinus coloratus). American Midland Naturalist, 144, 286-296.

Mahesh, P., & Mather, P. (2003). An assessment of the effectiveness of decision tree methods for land cover classification. *Remote Sensing of Environment, 86*(4), 554-565.

McGaughey, R. (2010). FUSION/LDV: Software for LiDAR Data Analysis and Visualization.

Meyer, P., Itten, K., Kellenberger, T., Sandmeier, S., & Sandmeier, R. (1993). Radiometric corrections of topographically induced effects on Landsat TM data in an alpine environment. *Journal of Photogrammetry and Remote Sensing*, *48*(4), 17-28.

Mueller, R. F. (2003). Centeral Appalachian Wilderness in Perspective: The Monongahela National Forest. *Wild Earth 2.3*, 56-60.

Mutlu, M., & Popescu, S. (2006). Assessing forest fuel models using LiDAR remote sensing. *ASPRS 2006 Annual Conference.* Reno, Nevada: ASPRS.

Myneni, R., Dong, J., Tucker, C., Kaufmann, R., Kauppi , P., Liski , J., et al. (2001). A large carbon sink in the woody biomass of Northern forests. *Proceedings o the National Academy of Sciences, 98*(26), 14784 - 14789.

National Oceanic and Atmospheric Administration (NOAA) Coastal Services Center. (2008). “LiDAR 101: An Introduction LiDAR Technology, Data, and Applications.” Charleston, SC: NOAA Coastal Services Center.

Noss, R., LaRoe, E., & Scott, J. (1995). Endangerd ecosystems of the United States: A preliminary assessment of loss and degradation. Washington, DC: National Biological Service, Biological Report 28.

Nowacki, G., Carr, R., & Dyck, M. (2009). The Current Status of Red Spruce in the Eastern United States: Distribution, Population Trends, and Environmental Drivers. Ecology and Management of High Elevation Forests in the Central and Southern Appalachian Mountains (pp. 140-162). Slatyfork, West Virginia: United States Department of Agriculture.

Pal, M., & Mather, P. (2003). An assessment of the effectiveness of decision tree methods for land cover classification. *Remote Sensing of Environment, 86*(4), 554-565.

Pascual, c., Garcia-Abril, A., Martin-Fernandez, S., & Cohen, W. (2010). Relationship between LiDAR-derived forest canopy height and landsat images. *International Jornal of Remote Sensing, 31*(5), 1261-1280.

PhotoScience. (2011). *West Virginia Northern Flying Squirrel (WVNFS) Vegetative Habitat.* Elkins, West Virginia: United States Forest Service.

Popescu , S., Wynne, R., & Nelson, R. (2002). Estimating plot-level tree heights with LiDAR: local filtering with a canopy-height based variable window size. *Computers and Electronics in Agriculture, 37*(1-3), 71-95.

Popescu, S., & Wynne, R. (2004). Seeing the trees in the forest: Using LiDAR and multispectral data fusion with local filtering and variable window size for estimating tree height. *Photogrammetric Engineering & Remote Sensing, 70*(5), 589-604.

Potter, K., Hargrove, W., & Koch, F. (2009). Predicting Climate Change Extirpation Risk for Central and Southern Appalachian Forest Tree Species. Ecology and Management of High Elevation Forests in the Central and Southern Appalachian

Mountains (pp. 179-189). Slatyfork, West Virginia: Untied States Department of Agriculture.

Reitberger, J., Krzystek P., & Stilla , U. (2008). Analysis of full waveform LiDAR data for tree species classification. *International Society for Photogrammetry and Remote Sensing*, Commission III.

Reutebuch, S., Andersen, H., & McGaughey, R. (2005). Light Detection and Ranging (LIDAR): An Emerging gtool for Multiple Resource Inventory. *Journal of forestry*, 286-292.

Richards, J. (1993). *Remote sensing digital image analysis: an introduction.* Springer; 2nd edition.

Ricotta, C., Avena, G., & Palma, A. (1999). Mapping and monitoring net primary productivity with AVHRR NDVI time-series: statistical equivalence of cumulative vegetation indices. *ISPRS Journal of Photogrammetry and Remote Sensing, 54*(5-6), 325-331.

Rollins, A. (2005). Analysis of Red Spruce(picea rubens) Regeneration in Pocahontas, Randolph, and Tucker counties, West Virginia. Morgantown, West Virginia : Unpublished Thesis, Department of Forestry, West Virginia University.

Rossiter, D. (2004). *Technical Note: Statistical Methods for Accuracy Assessment of Classified Thematic Maps.* Department of Earth Systems Analysis International Institute for Geo-information Science & Earth Observation (ITC), Enschede (NL).

Roy, P., Sharma, K., & Jain, A. (1996). Stratification of density in dry deciduous forest using satellite remote sensing digital data: An approach based on spectral indices. *Journal of biosciences, 21*(5), 723-734.

Song, J., Han, S., Yu, K., & Kim, Y. (2002). Assessing the possibility of land cover classification using LiDAR intensity data. *Internatiobnal Society for Photogrammetry and Remote Sensing Commission III PCV02 Symposium*.

Sterenczak, K. (2010). DTM generated from LiDAR data under foreest conditions. *International Society for Photogrammetry and Remote Sensing*.

Strahler, A. (1980). The use of prior probabilities in maximum likelihood classification of remotely sensed data. *Remote Sensing of Environment, 10*, 135-163.

Sullivan, A. (2008). LIDAR Based Delineation in Forest Stands. Master of Science *Thesis: University of Washington* .

Sumendi, A. & Allebach J. (2008). Nearest-neighbor and bilinear resampling factor estimation to detect blockiness or blurriness of an image. *Journal of Electronic Imaging*, 17(02), 023005 (16 pages).

Tiede, D., Lang, S., & Maier, B. (2007). Transferability of a tree crown delineation approach using region-speciic segmentation. *Simposio Brasileiro de Sensoriamento Remoto*, 1883- 1890.

Tucker, C. (1979). Red and photographic infrared linear combinations for monitoring vegetation. *Remote Sensing of the Environment, 8*, 127-150.

Verbyla , D., & Hammond, T. (1995). Conservative bias in classification accuracy assessment due to pixel-by-pixel comparison of classified images with reference grids. *International Journal of Remote Sensing, 16*, 581-587.

Wandinger, U. (2005). *Introduction to LiDAR* (Vol. 102). SpringerLink.

Waring, R., & Schlesinger, W. (1985). Forest Ecosystems: Concepts and Management. London: *Academic Press*.

Williams, G. (2009). Rattle: A data mining GUI for R. *The R Journal, 1*(2), 45-55.

Xudong, L., Xuedong, Z., & Youchuan, W. (2005). A kind of filtering algorithms for LiDAR intensity image based on flatness terrain. *ISPRS Journal of Photogrammetry and Remote Sensing, XXXVI-2*(W25).

Yan, W., & Shaker, A. (2006). Radiometric calibration of airborne LiDAR intensity data for land cover classification. *ISPRS Journal of Photogrammetry and Remote Sensing, commission I, WG I/2* .

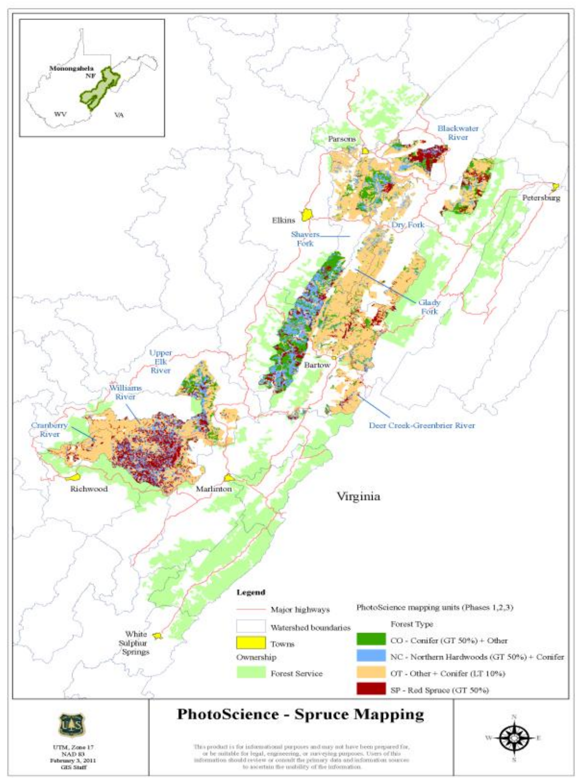
Appendix

MONONGAHELA NATIONAL FOREST

WEST VIRGINIA NORTHERN FLYING SQUIRREL (WVNFS) VEGETATIVE HABITAT

February 5, 2011

The PhotoScience Red Spruce mapping project involved mapping Forest Service lands falling within eight 5th level hydrologic units that fall on or intersect with the Monongahela National Forest (MNF) for four forest types as correlates with West Virginia Northern Flying Squirrel habitat. These forest types are spruce (SP), northern hardwoods-conifer (NC), conifer-other (CO), and other (OT). The study area encompasses approximately 780,000 total acres of which 500,000 acres fall within Forest Service ownership. The project primarily maps those 500,000 acres of high-elevation Forest Service lands (greater than 2,500 feet) within the Forest boundary. Due to funding constraints the 280,000 acres of non-Forest Service land were not consistently mapped.



OBJECTIVES

This project's primary objective was to map the identified forest cover types and through the use of physiognomic modifiers to further refine the habitat classification. The four forest types include spruce (SP), northern hardwoods-conifer (NC), conifer-other (CO), and other (OT). Physiognomic modifiers included size class, coverage density, and crown condition for all forest types with the exception of the “other” category. This work facilitates the identification of West Virginia Northern Flying Squirrel habitat across the Forest.

STANDARDS AND SPECIFICATIONS

The minimum standards for delineating the vegetative polygons are:

- The polygons should be generally homogeneous in species, size class, canopy cover, and crown condition.
- The minimum mapping unit for non-riparian polygons is five acres.
- A two- acre minimum mapping unit will be applied to polygons with riparian vegetation along streams or wet areas.

Polygons are delineated based on identifiable characteristics as seen on digital ortho-photographs, satellite imagery, or aerial photography. Polygon boundaries are drawn around homogenous vegetation conditions and obvious changes in the delineation criteria. Delineation criteria can be subtle, but most often are readily seen on a stereo view of the photo.

DELINEATION CRITERIA

Black and white aerial photographic prints, at the 1:15840 scale and dated from the spring of 2001, were originally used as stereo pairs to determine polygon characteristics for the initial phase of the three-phase project. Subsequent work incorporated the use of spring, 2003 leaf-off aerial photography shot at the 1:4800 scale and converted to raster imagery. Delineation criteria are based on a map unit code (i.e., forest cover type) and physiognomic modifiers.

The **Forest Cover Type** represents one of four vegetative types: spruce, northern hardwoods-conifer, conifer-other, and other. The spruce type (SP) is made up of red spruce that comprises a plurality of stocking (greater than 50%) with potential minor components of hemlock, norway spruce, balsam fir, and hardwood species. The northern hardwoods-conifer type (NC) of sugar maple,beech, and birch which exist singularly or in combination and comprise the plurality of the stocking (greater than 50%) with potential minor subcomponents of conifer (greater than 10%). The conifer-other type (CO) is composed of white pine, hemlock, red pine which exist singularly or in combination and comprise a plurality of stocking (greater than 50%) with a subcomponent of spruce and/or hardwood species. The other type (OT) is a stand which is predominantly hardwood, but may contain a minor components of conifer species (less than 10%).

The map unit code represents the Forest Cover Type for a particular polygon. The map unit code is made up of two alpha characters.

Physiognomic Modifiers

The first physiognomic modifier code represents the **Size Class**. The size classes for trees are seedlings/saplings, poletimber, sawtimber, and mature based on diameter of the stem at breast height (DBH).

Record the sizes of tree using the tree size classes in the Table below. The diameters associated with each size class are interpreted from the height and crown structure, unless measured in the field. Do not include seedlings unless they are the dominant vegetation. This modifier is a one-character alpha code and follows the forest type.

Code	Size Class (dbh)	Description
E	0 – 4.99”	Seedlings/saplings
P	5 – 8.99”	Poletimber
S	9 – 21.99”	Sawtimber
M	22” or greater	Mature

The second physiognomic code represents **Canopy Cover**. The canopy cover describes the plurality of canopy cover based on the dominant species. The modifier is a single numeric code from one (1) through five (5). All vegetation map unit codes receive this modifier.

Code	Crown Condition
A	Signs of declining crowns absent, less than 15% of the component
D	Dead crowns comprise 15 – 50% of the component
T	Tree mortality, over >50% of the component

The third physiognomic modifer represents the **Crown Condition** of the polygon. The polygon must contain greater than 15% crown damage to qualify for a damaged crown condition of D (containing over 15% and less than 50% dead crowns) or T (containing over 50% tree mortality). This modifier is a two-character alpha code.

Code	Canopy Cover
1	Less than 10%, sparse vegetation
2	10 – 29.9%, canopy cover
3	30 – 59.9%, canopy cover
4	60 – 79.9%, canopy cover
5	80 – 100%, canopy cover

ACCURACY

The vegetation maps was based on the National Map Accuracy Standards for positional accuracy with a minimum class accuracy goal of 80 percent.

ACCURACY ASSESSMENT

A partial ground-based accuracy assessment was completed to validate and verify the classification. Sample points were randomly selected from the classified type, located using GPS technology, evaluated for typing criteria, photographed for illustrative purposes, and digitally mapped in GIS.

TO BE PROVIDED BY MNF

- Project validation sites (100+)
- Red Spruce vegetation layer
- Forest Visitor Map (digital and hardcopy)
- Forest stands layer with CDS forest type attribute
- Digital SAMB 2003 leaf-off, spring true color 2-foot digital orthophotography
- NAIP summer 2007 leaf-on true color 1-meter digital orthophorography
- Forest ownership
- Road + trail layers
- Stream network (from NHD)
- Digital elevation models (30-meter)
- Digital raster graphics (DRG) or 1:24000 scale topographic quadrangles.

CONTRACTOR DELIVERABLES

The following deliverables were required:

- Digital classified forest type layer in ArcGIS geodatabase/coverage format with polygon modifiers.
- Modifier data entered, delivered, and provided as attribues in the ArcGIS geodatabase/coverage
- FGDC-compliant metadata for coverage
- Maps in both hard-copy and digital format.
- All field data will be provided in a digital database management system (DBMS).
- Photographs of accuracy assessment locations.
- Map accuracy verification (including error matrix).