Mixed Forest Image Classification of Paper Birch:
Using AVIRIS Bandwidths Ranging from 530 to 745 nm

by

Kyle Benjamin Uhler

A Thesis Presented to the
Faculty of the USC Graduate School
University of Southern California
In Partial Fulfillment of the
Requirement for the Degree
Master of Science
(Geographic Information Sciences and Technology)

August, 2018

To my wife Meghan

Table of Contents

List of Figures

List of Tables

Acknowledgements

I wish to acknowledge the patience and guidance that Professor Dr. Laura Loyola and my thesis advisor Dr. Elisabeth Sedano gave throughout this thesis. I am grateful to all my peers who helped edit this report and facilitate discussion about this domain. Thank you to my thesis committee members Dr. Andrew Marx and Dr. Steven Fleming for their input and advice. I want to acknowledge my employer, the Minnesota Board of Water and Soil Resources, for facilitating project-based learning concerning remote sensing. Finally, I would like to thank Jeff Walsh from Peregrine Aerospatial LLC for assisting with this project's direction and for insights into commercial remote sensing applications.

List of Abbreviations

AVIRIS	Airborne Visible/ Infrared Imaging Spectrometer
CORS	Continually operating reference station
DBH	Diameter at breast height
DN	Digital numbers
EM	Electromagnetic
FSA	Farm Service Agency
GCPs	Ground control points
GIS	Geographic information systems
LMF	Laurentian mixed forest
MBS	Minnesota Biological Survey
MN DNR	Minnesota Department of Natural Resources
NAIP	National Agriculture Imagery Program
NIR	Near-infrared
NSFC	North Shore Forest Collaborative
NSH	North Shore Highlands
RGF	Raster group files
RS	Remote sensing
SGCN	Specie of greatest conservational need

Abstract

Paper birch (*Betula papyrifera*) is a dominant specie within Northern Minnesota's Laurentian mixed forest. Though these trees are common place, paper birch populations have been in decline for the past couple decades along Lake Superior. Due to the reduced replacement rate of this specie, organizations are implementing management strategies to promote healthy forests. This thesis investigates remote sensing techniques to predict paper birch locations remotely. The thesis tests individual specie level spectral signature effectiveness to classify community level data of the same informational group. The project uses hyperspectral Airborne Visible/ Infrared Imaging Spectrometer (AVIRIS) flights data for its imaging platform. Two spatial resolution scenes were classified for this project. A 4.3 m pixel resolution image was used for defining spectral signatures based on ground truth data. Then a lower resolution 16.5 m image was used to apply the produced paper birch signature as a classifier to test functionality of these methods on known paper birch communities. Pixels were used as the final classification unit. A linear unmixing soft classification was utilized to produce percent signature contained within pixels. The classification resulted in ~93% of forest plots containing some pixels with 95% similar spectral signatures to paper birch. Total classified coverage of validation forest plots was low with only ~30% covered with 75% similar signature or greater. The long-term objective for this project is to automate specie identification to monitor trees. Further research is needed to streamline classification and refine procedures, yet current findings can help forest managers and conservationists identify priority sites to both map current specie distribution and implement restoration activities.

Chapter 1 Introduction

Paper Birch (*Betula papyrifera*) populations on Minnesota's North Shore Highlands (NSH) are on the decline. Many factors influence the reduction of paper birch populations. Some contributing variables that can directly affect paper birch are deer browsing, insect infestations, soil compaction, and regional warming (NSFC 2017). Historic stands of white pine and cedars were extensively harvested in the 1800s, altering the landscape into a successional habitat. In those altered areas pioneer species like paper birch have thrived. Climate models used by the Northern Institute of Applied Climate Sciences for forest adaptability assessments predict winter temperatures will continue to increase. This increase in temperature could reduce the regional snowpack, which makes up a majority of its annual precipitation. This spells a disaster for the cold climate species of the North Shore like paper birch. Management of paper birch populations depends upon identification of those populations. This paper demonstrates methods to identify birch populations in mixed forest habitats using remotely sensed data.

This project is in response to the loss of paper birch on the landscape. This research explores informational group classifications of paper birch in the Laurentian mixed forest (LMF) habitats of the North Shore. It incorporates remote sensing techniques to better equip forest managers to monitor paper birch remotely along the North Shore and use in the management decision process. This project does not seek to model future vegetation, but to assist in monitoring of past and current paper birch stands to give managers another tool to track plant community change. The goals are to produce methods that maximize correct classification of paper birch within the mixed forest and reduce time taken to model that plant community.

This study could help future researchers establish best practices for remote detection of paper birch and plants with similar plant morphologies. Utilizing findings form this project could help future monitoring efforts. This chapter introduces paper birch and the habitat it occupies. It discusses habitat change and ongoing management of the landscape. The chapter then introduces remote sensing applications for environmental research. In conclusion, an overview of this project's study area is described.

1.1. Laurentian Mixed Forest Habitat

Northern Minnesota is dominated by the LMF. This habitat type surrounds much of the Great Lakes region. The LMF covers some 23 million acres of Minnesota and is divided into 26 subsections based on soil, climate, and vegetation (MN DNR 2006). The habitat is dominated by conifer and deciduous broadleaf forests, swamps, and bogs with a variety of niche communities. This landscape is marred by historic logging, land development, and more recently tourism. Many factors have their effect on what this environment looks like today.

The LMF is a habitat that is influenced by historical geologic underpinnings. The Laurentian name refers to the ancient continent and a precursor to North America. The bedrock of the North Shore is situated along a mid-continental rift of this ancient landscape. The 1.1 billion-year-old basalt rift is greatly exposed, defining the province. The more recent geologic events of glaciation have ground and eroded soils within the region. Due to these glacial events this area is characterized by shallow coarse aggregate and topography with great relief.

The prevalent driver of climate to the near shore environments is Lake Superior. This is a unique habitat that receives most of its annual precipitation in the form of snow. The lake acts as a climate moderator (MNDNR 2006), both a source of heat in the winter and as an air conditioner in the summer. The moderate cool annual climate gives rise to cold-weather vegetative communities.

Both the presettlement and current vegetation communities are forest dominated. These forests are fire-dependent communities that rely on episodic successional pioneer species. After disturbance events stands of conifers are replaced by species like aspen, birch, and maple. This project seeks to identify paper birch communities, which are deciduous-broadleaf species.

1.1.1. Paper Birch

Paper birch trees are easily identifiable specie on the northern landscape. They have iconic bark that is thin, white, and peeling in nature. Paper birch trees are known to grow up to 30 m in height but are usually 20 m or shorter (Marshall 1785). Trees are most commonly found with a singular main trunk but can be seen growing with multiple trunks. Multiple trunked trees will usually be defined by shorter canopy heights. Paper birch trees are shallow-rooted with few roots deeper than 60 cm below the soil surface (Safford 1990). A key factor in this research is the average crown width. The North Dakota Tree Handbook states canopies at maturity can be observed from 6 to 12 m in width. These birch crowns are comprised of simple leaves that are 2-4 in. long with narrow 1 in. stems and can be seen in Figure 1.

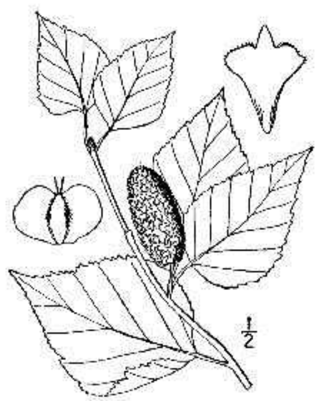

Figure 1 Paper birch leaf and seed morphologies (USDA)

Paper birch form nearly pure pioneer communities on disturbed sites and are rare in late successional forests (Uchytill 1991). In the LMF paper birch can be found in mixed established communities due to the distribution of the tiny seeds on the wind. Compared to other tree species birch are relatively short-lived species. Most birch trees will not grow past the age of 125 years old (Day 1981). Paper birch have a fast grow rate, which is useful for long-term revegetation and soil stabilization of severely disturbed sites (Uchytill 1991).

A multitude of pests can affect the health of birch trees, yet few pests are culpable for much of birch harm. A major insect pest of birch trees is the bronze birch borer. Individual trees compromised by environmental stresses are susceptible to infestation of the bronze birch bore (Conklin 1969). Over-browsing can knock back birch populations too. Although there are many browsing animals that affect birch, white-tailed deer eat substantial amounts of the trees' foliage later in the season, and heavy browsing by deer and moose populations have been shown to prevent or setback forest regeneration (Irwin 1985).

1.1.2. Habitat Change and Conservation

Minnesota's NSH is uniquely situated along many environmental gradients. The area is at the north edge of many warm-climate deciduous tree species and the southern limit of many cold-weather species. Due to this fact, slight changes to climate gradients can alter the vegetative community makeup. The changing forest can increase soil erosion, river turbidity, and sediment deposition that affects the environment (Malgorzata 2010). Increasing habitat change throughout the region has led conservation groups to develop restoration plans to combat the transformations.

The North Shore Restoration Project is one program planned to manage 39,000 acres of the Superior National Forest that is part of the NSH. It was implemented as a response to climate change. It assessed the area and found that 80% of the birch stands are of mature age and dying. This would be of no concern if the trees were being replaced by other age stands, but that is not the case. The historically forested habitat is not regenerating and is being replaced by grasses as seen in Figure 2.

Figure 2 Dead and dying paper birch trees on the North Shore (MN DNR)

The Superior National Forest assessment predicts an unsustainable future habitat due to the lack of tree stand regeneration. Many projects are collaborating on reforesting efforts. The

North Shore Forest Collaborative (NSFC) is a promising initiative to promote healthy forests. The NSFC strives to bring together a multitude of partners to restore a productive ecosystem and they promote native forests. In 2015 alone, this project planted many thousands of trees, over some 1000 acres of the North Shore. Attempts to reforest this region are underway.

The Minnesota Biological Survey (MBS) has been identifying species in greatest conservational need (SGCN) for the NSH. The MBS found 84 wildlife SGCN that utilize this region. Table 1 displays nine problems that contribute to specie decline and the percentage that the problem influences the SGCN. Habitat loss and degradation lead the list, showing the importance of managing these habitats appropriately. Monitoring this habitat can aid management in restoring populations of greatest concern. Projects like these need a baseline monitoring protocol to gauge progress of these habitat and implement management.

Table 1 SGCN study from the MBS

Problem	Percentage of SGCN in the Subsection for Which This Is a Problem
Habitat Loss in MN	73
Habitat Degradation in MN	82
Habitat Loss/Degradation Outside of MN	38
Invasive Species and Competition	26
Pollution	26
Social Tolerance/Persecution/Exploitation	25
Disease	1
Food Source Limitations	2
Other	6

This project is motivated by a desire to help the restoration and management of Minnesota's NSH and other similarly endangered forests. Furthermore, this project's scope covers a wide-range of topics within the domain of remote sensing. Finally, this project brings together areas of personal interest.

1.2. Remote Sensing

The domain of remote sensing (RS) is simplistically described as obtaining information from a distance. The science of RS has been changing rapidly since the launch of Earth-orbiting satellites in the 1970s. NASA's Landsat mission's objective was to make repeated Earth observations through time. These missions gave researchers multispectral data over large extents of the globe. The images from Landsat and other RS platforms allowed for image analysis on a large scale. As RS advanced, so did geographic information systems (GIS). The ability to pair geospatial material with RS increased the potential of geographic analysis.

1.2.1. Imagery Classification

Cover class description is an objective for researchers in the RS domain. Research and development in classification methods have improved since early aerial photography. The early classifications were accomplished through photo interpretation by trained specialists. Computer algorithms now have the power to identify objects and even classify the unique characteristics of features' makeup. Classification of imagery pixels in this study is used to identify paper birch on the landscape. Pixels contain values from multiple bands. Spectral difference between pixels can be compared and assigned to informational groupings. This thesis investigates spectral differences between pixels of known objects to define classes. Supervised classification of the landscape was used to form a final project product. Supervised classifications require inputs to identify pixels related to ground features. The spectral signature of paper birch is extracted with use of an algorithm.

1.3. Methods Overview and Study Area

Image classification of paper birch within the Laurentian mixed forest habitat represents a challenging scenario. Though paper birch represents a sizable portion of the individuals on the landscape the trees' morphological traits lead to mixed pixel outputs. Discrimination of this plant due to this fact can be problematic.

This project investigates areas of the NSH subsection of the LMF. This subsection of the LMF is adjacent to Lake Superior, only 25 miles inland at its furthest. The study area presents a mixture of swamps, bogs, lakes, conifer, and hardwood forests. Figure 3 shows the extent of the LMF in Minnesota (purple) and the 1.4-million-acre extent of the NSH subsection (orange).

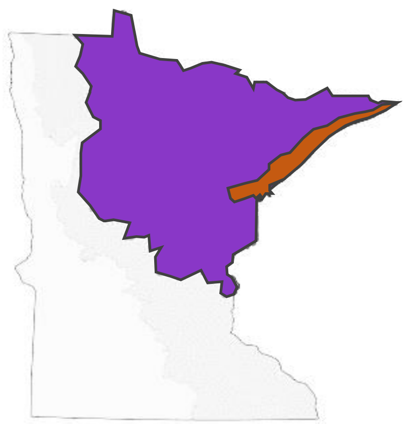

Figure 3 Laurentian mixed forest seen in purple
and the North Shore Highlands subsection in orange (MN DNR)

The overall objective of this project was to develop individual level signature data and apply those signatures to a community level scene to predict paper birch occurrence. This project tested supervised classification using training data collected in the field. The training data selection involved quality controls to ensure quality data to model the spectral class. The supervised image classification section describes utilizing a linear unmixing technique to assess percent similar pixel signature. This technique is a soft classification used in mixed pixel situations.

In this thesis, two AVIRIS flight scenes were acquired over NE Minnesota. The classification of both AVIRIS resolutions were tested for accuracy with ground truth data collected and open source forest stand inventory data. Subpixel scale was used to represent canopy locations that were based on percent spectral response. The data processing phase is ongoing throughout this project and includes clipping data into workable datasets, georeferencing, translating files into useable formats, and data exploration. Figure 4 displays a general schematic of this project.

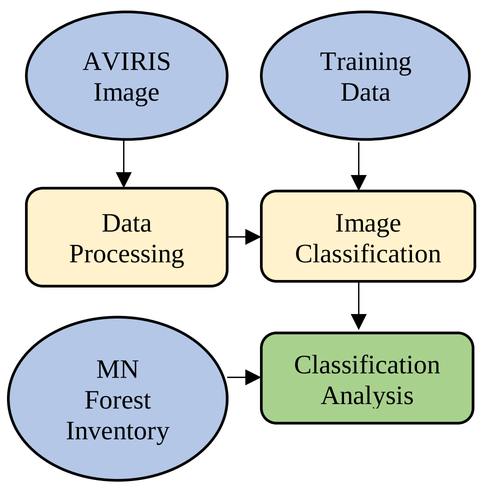

Figure 4 General methods schematic

The study area within the NSH was chosen based on public access, historic survey records, and availability of imagery. Areas between Two Harbors and Lutsen, MN exhibit those factors. AVIRIS images located on the North Shore within 20 miles of one another were used for classification. Figure 5 shows the locations of both images in relationship to one another and in Minnesota. These areas have surveyed cover classes delineated in a MN Forest Inventory layer.

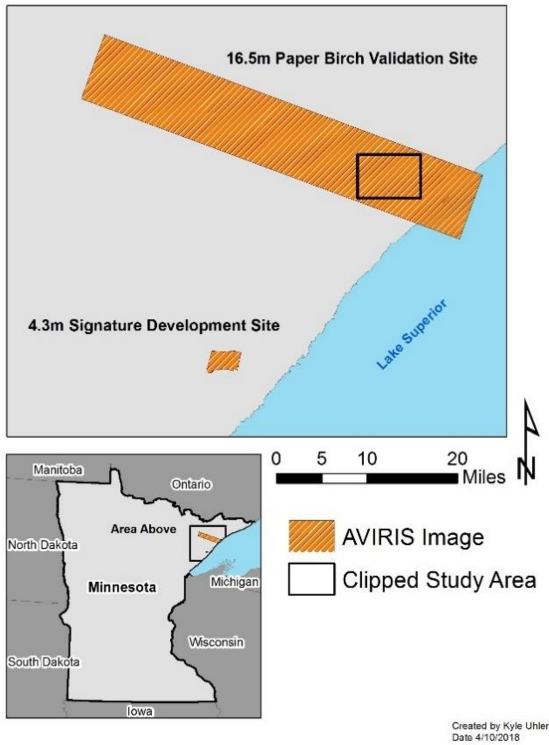

Figure 5 Study area overview

The following chapter examines background research in RS classification and specie identification using hyperspectral AVIRIS datasets. This study utilizes prior investigations in hyperspectral species classification to shape methodologies and those studies are discussed in the literature review. Many aspects of remote sensing sciences were used to complete this project and methods applied are described in Chapter 3. Then Chapter 4 presents results of both data processing and image classification. Finally, Chapter 5 discusses recommendations for further and future work as well as applications of this project.

Chapter 2 Related Work

Traditionally, remote sensing has used imagery to identify vegetative areas form non-vegetative areas of the Earth's surface. Many applications have utilized community-based observations because individual plants can be difficult to verify with low satellite resolutions. The vegetative areas imaged can be subdivided into classes of plants with similar reflections called spectral classes. It is important to understand the foundations of remote sensing prior to completing this project. The following sections describe related work in remote sensing that gives a background to species level identification. It is crucial to understand the foundational research into remote sensing systems and the electromagnetic (EM) interactions with our planet's surface. Following sections also investigate hyperspectral data uses in vegetation classification. Additionally, knowledge for this project will utilize conventional classification methodologies of data and will be discussed in this chapter.

2.1. Background

Remote sensing has many useful applications such as identifying land cover and more specifically specie identification. Land cover classifications involve classifying large extents vs. specie delineations happening on smaller scales e.g. stand level in forest ecology. Full cover classification of specie level data is extremely limited. Using RS applications over more traditional cover surveys can save investigators time in the field.

Digital imaging platforms coinciding with GIS has allowed raster pixels to be interpreted as ground features. The individual pixels contain information beyond the georeferenced location and area covered. Imaged pixels are commonly formatted as digital numbers (DN), which is the measured intensity value. DN are usually a conversion from radiance captured at the sensor. DN can be formatted as different binary digits or bits. The number of bits collected will regulate the radiometric resolution of the image. For example, an 8-bit data represents pixels defined between 0-255, whereas 16 bit allows for values range of 0-65536. Many raster images are formatted as 8-bits, because it represents decent diversity for a scene. Higher radiometric resolution allows greater detection of small variation in ground reflectance. In observing the radiation emitted by these objects classification can be obtained. This project will exploit the difference between known objects' EM relative emittance and predict occurrence.

2.1.1. Vegetation and Atmospheric Interactions with Electromagnetism

The EM spectrum is divided into regions base on historic uses within disciplines. The EM spectrum ranges from < 0.03 nm to > 30 cm. Any radiation captured with RS technology inevitably must travel through the Earth's atmosphere. This radiation doesn't move unimpeded in the atmosphere. The atmosphere can scatter, refract, and absorb incoming EM energy. Observed scattering of particles is greatest near the blue end of the spectrum. Refraction of light by atmospheric interactions will depend on humidity and thickness of atmosphere at image capture time. Absorption of energy happens by molecules like ozone and carbon dioxide within the image scene. There are multiple bandwidths in the atmosphere that absorbs energy; water vapor absorption is centered at approximately 0.94, 1.14, 1.38 and 1.88 µm, an absorption band for oxygen is at 0.76 µm, and a carbon dioxide absorption band is near 2.08 µm (Gao 2009). A researcher may avoid use of these absorption points to limit effects on datasets.

The removal of atmospheric effects is necessary when preparing RS data. Remote sensing sensors on airborne platforms can minimize atmospheric effects by lowering flight altitude. This will mitigate some atmospheric effects by reducing the amount of atmosphere images are captured through. Remotely sensed areas with sparse vegetation can omit both atmospheric and background noise when using indices that account for those effects (Giannico 2004). The atmospheric effects can increase as spatial resolution decreases; and spectral resolution increases. The energy that is not affected by atmospheric distortions are called atmospheric windows. These are the regions that allow light to return to the sensor for capture.

Many methods for atmospheric correction have been established, ranging from simple calibrations to complex models. The Gao et al. (2009) experiment reviewed multiple approaches to correct atmospheric effects for routine processing of imaging data. They found that radiative transfer model approaches are sufficient and can be used for preprocessing of hyperspectral data. Radiative transfer models are tools to represent both scattering and absorption of radiation. When multi-temporal images are used for image classification, atmospheric calibration is mandatory (Lu 2007).

Once the images are corrected for atmospheric manipulation surface vegetation light interaction can be measured. The reflectance of any plant depends on the leaf structure, pigments, and water content within those parts. Species with similar leaf structures will produce similar spectral response (Hoffer 1969). The typical spectral response of green vegetation is shown in Figure 5. The troughs represent wavelengths at which absorption by pigment, then by water occur within the plant. The human-visible spectrum occupies wavelengths between 380-720 nm. In Figure 6 one can see pigment absorption near 400 and 650 nm. The infrared area of the EM spectrum is a much larger section than visible light and is important to the identification of plants. The near-infrared (NIR) wavelengths are longer than those within the visible spectrum (720 nm -1300 nm). The air-spaces in leaves lower cross-sectional structure slows the incoming light. Larger leaves air spaces have longer wavelengths in the NIR (Abrams 1990). These variations in NIR response can be used to define plant species within a scene.

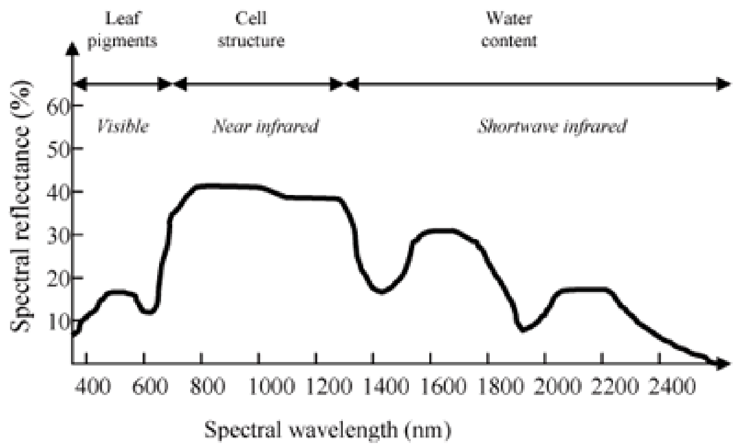

Figure 6 Common spectral reflectance of green vegetation (Gaussman 1977)

Variation in spectral reflectance from tree canopy level can be influenced by location, specie, leaf structure, leaf angle, water content, background, varied irradiance and pigment concentration (Ollinger 2011). It is difficult to reproduce canopy reflectance; attempts have been made by stacking leaves on top of each other from below as in a canopy spectral measurement (Coops 2003). Spectral differences within tree canopies is known in the literature and is significant with respects to position in the canopy (Danson 1995). Spectral variation in the visible wavelengths are due to chlorophyll concentrations (Cochrane 2000). Plant senescence in the autumn changes the overall levels of chlorophyll contain in the leaf. Gitelson et al. (1994) found that leaves with high chlorophyll concentrations had low pigment variation. Maximum deviations of pigment where observed near 550-560 and 700-710 nm. These ranges are most sensitive to pigment concentrations (Blackburn 2007). This project uses variations in these ranges to classify paper birch.

Distinguishing variables characterize a plant's functional type in the environment. The plant's functional type is defined as the vegetative structure, biochemical physiology, and phenology. These plant variables affect the spectral return of the plant. Spectral returns that have been identified at the leaf scale in studies have also been applied generally at both the canopy and landscape scales (Zarco-Tejada 2000). This study will use canopy level spectral response due to the spatial resolution of the raster images obtained.

Phenology does vary between plant species, but expression will fluctuate based on temperature, moisture, and photoperiod (Sayn-Wittgenstein 1978). Although there is season variability photosynthesis is highly reliant on temperature and photoperiod (Reed et al. 1994). Thus, single year observations can generalize based on phenological differences.

These phenotypic seasonal differences are exploited in this project. In general, many remote sensing studies focus on coarse resolution data to delineate forest communities and have not used individual trees as a classification unit (Key et al. 2001).

2.2. Image Classification

Imaging landscapes for remote sensing purposes has long interested researchers. Classification of the landscapes can produce a wide range of accuracies because of scene complexity, informational groupings, image acquisition, and classification approach. The challenges of RS can be overcome with insight about the study areas cover. When beginning analysis of an imaged scene most researchers have some prior knowledge of the locational cover. Furthermore, when using prior surveyed data for a site, the researcher can select sites with individual spectral groups or plant communities of interest. If historic surveys acquired were sampled within similar communities as the research questions, investigators can maximize accuracies of classification (Lauver 1997).

Utilizing prior knowledge about surface cover to inform a classifier is considered a supervised classification. Supervised classifications lean on the researcher's prior knowledge of the image location to reduce the errors in classification. The prior site knowledge is selected as training data to produce similar information groups. In using supervised classifications, the selection of training data is a crucial step in a project work flow and can affect overall accuracy in capturing the attended informational group. In an elevation-based vegetation classification study, Gartzia (2013) found that the proper selection of training data was the most principal factor in improving accuracies. The problems in selecting enough quality data in mixed landscapes is that it is difficult and is complicated when coarse spatial resolution datasets are used for classification. These factors lead to increased mixed pixels (Lu 2007).

The habitat preferences of paper birch are commonly associated with many northern forest species that don't necessarily create homogenous timber stands. This leads to the increased likelihood of mixed pixel outputs depending on resolution. Commonly the number of mixed pixels increases as resolution decreases (Crapper 1984). Mixed pixels act to average the brightness of a scene by subdivision. Mixed pixels are not inherently bad, if the spectral signatures averaged are of similar groups the average signature of the pixel will not be affected much. When the pixel overlaps distinctly different spectral signatures within a pixel the average can change greatly (Franklin 2002). The mixing of vegetation and soil reflectance is an example of possible distinct spectra in a mixed pixel composition. The combined signature of the distinct groups may not match any of the groups captured in the pixel area (Campbell 1987).

The spectral signature contains all materials present in the training pixels and can neglect the influence of the mixed pixels. The characteristics of each pixel spectra then represents the objects acquired in a hyperspectral image (Campbell 2011). This thesis discriminates spectral classes within mixed forest vegetation from other information classes within the imagery.

Mixed forest habitats by nature are difficult to distinguish informational groups that exists within the community. Early studies like Franklin (1994) found that out of 12 classes used on Landsat TM multi-spectral imagery of mixed deciduous forest they achieved accuracies of only 77%. That study was using the entire community types and showed the difficulty of mixed pixel identification. Leckie et al. (2005) sub-divided informational groups to increase the representation of the group and bypass intra-group variation by environmental conditions.

Establishing a robust training data sample will limit interspecies variation. Fassnacht et al. (2016) reviewed prior work and established six criteria on which training data selection should successfully be based: classes much match the research question(s); data selected is representative of study location; spatial scale should match the scale of the question; assumptions underlying the methodologies applied; observed errors should be known and discussed for impact; and samples should be spatially independent. The training data for this project will use cover type class categories established by the MN DNR to distinguish areas to select the best training data locations. In general, a good representative dataset for every class is fundamental in carrying out a supervised classification. (Lu 2007).

The classification of objects with transitional boundaries are inherently difficult to represent spatially. Fuzzy classifications or soft classifications offer a different approach to classifying objects. The soft classifications result in a degree of membership to each informational group. Unlike soft classifiers, hard classifications use well defined groups for each pixel location. In mixed pixel images traditional hard classifiers are unsuitable (Atkinson et al. 1997). Long-established classifiers such as the maximum likelihood classifier utilizes the spectral bands to resolve what class each pixel is most likely to belong to. The goal of a soft classifier is not to categorize each pixel into one group but have it as a sum of its components (Atkinson et al. 1997).

Soft classifiers are intrinsically difficult to develop accurately, but the produced pixel mixtures can lead to more accurate representation of land-cover estimations compared to hard classifiers (Ichoko et al. 2009). Linear unmixing is one such method that produces fuzzy memberships. These categories can be hardened based on input criteria. Linear unmixing functions have been found quick and useful for image classification. Linear unmixing models every pixel as a linear function of the classes input. At its simplest linear unmixing is the minimum difference between the training data spectra and the unknown spectra. This is based on all possible composition of the training data. The process of spectral unmixing involves breaking down pixel spectra into subpixel component spectra, commonly referred to as endmembers.

Studies have applied linear unmixing models to identifying rock substrates, crop cover estimates, and land cover estimates. In arriving at land cover estimates, Foody and Cox (1994) assumed pixels were pure spectra that defined endmember spectra. They found a strong correlation between actual and predicted percent cover with a 99 percent confidence. Although all endmember categories tested highly, grasses had the most accurate predictions. They noted a tendency to underestimate the percent cover of trees in some pixels because trees tend to not cover pixels homogeneously. A concern in using linear unmixing models is that in the study area there may not exist a pure class spectra pixel, depending on image resolution. Selecting appropriate training data for endmember spectra can increase accuracies (Atkinson et al. 1997).

2.1.1. AVIRIS Data Classification of Vegetation

Many early studies spotlight the use of broad spectral bandwidths in the visible and NIR for use in vegetation classification. Over time innovative technologies increased the number of bands reducing the spectral ranges. Recent work has targeted the emphasis of using narrow-band regions like the red edge to distinguish image elements (Blackburn 2007). Hyperspectral images consist of a hundred or more narrow spectral bands. Hyperspectral imaging is a specialized discipline within RS sciences. The spectral bandwidths are contiguous throughout the EM spectrum. Hyperspectral data offers a unique view of forest vegetation. This study takes advantage of a narrow band range of hyperspectral raster imagery and will apply classification methods to identify plants in the scene.

The AVIRIS platform has been utilized for countless classification studies. This is an airborne image acquisition platform. Data acquired form the AVIRIS optical sensor captures wavelengths from 400 to 2500 nanometers. This spectrum is obtained through 224 spectral bands. These data are collected on a NASA airplane. Lower altitude fights can capture greater than 3 m resolution. The higher resolution images are collected on smaller extent strips. Disadvantages of limited spatial extent is overcome by increased spectral resolution to define objects on the landscape. The pixels at sub-five meters begin to be able to isolate spectral signatures of individual tree canopies and can give a better distinction between objects. Imaging spectrometers collect images as contiguous spectral bands like AVIRIS. Complete reflectance spectrum can be derived from the wavelength region covered in every pixel (Goetz 1985). Hyperspectral imagery can be used to generate spectral signatures of vegetation species. Studies have used hyperspectral data to identify and map forest species and have achieved a certain degree of success (Ruiliang 2009). In nature, everything has a unique spectral characteristic that can be exploited to identify information about the object (Parker 1965). The assumption for this project is that spectral differences between species are greater than the variation with the target specie.

Prior work has utilized many different classifiers at a multitude of scales. The increased spectral resolutions obtained by hyperspectral datasets when remote sensing in adjacent narrow bands can produce elevated levels of autocorrelation (Blackburn 2007). Autocorrelated data is redundant in nature and is of low importance when identifying spectral variation. This needs to be addressed in methodologies that use hyperspectral data. Misclassification of hyperspectral imagery can be a result of many different variables. Problems can occur due to image timing creating shadowing. Also, areas with high relief can cause issues with classification (Mehner 2004). Individual plant species can vary on reflective signatures independently of pigment (Blackburn 2007). So, classification evolves to use both the average and diversity of each spectral group or class. The advent of hyperspectral imagery allows more stratified ranges to be analyzed.

Chapter 3 Methods

Methodologies presented in this thesis work to classify paper birch's likelihood of being represented by any given pixel within the study areas scope. Methods used for this study can be grouped into three general types of tasks: Data Acquisition and Preparation, Image Classification, and Spatial Analysis. This chapter describes these tasks.

3.1. Data Acquisition and Preparation

Geospatial data and RS data are becoming more accessible in online open source formats. Datasets obtained for this thesis project were all open-sourced or self-collected. Various resolutions and extents could be obtained for this projects requirement.

3.1.1. Data Acquisition

This project relied on AVIRIS datasets to perform all subsequent work. Hyperspectral datasets were acquired from NASA's JPL AVIRIS platform for the study site. Data was downloaded from NASA's JPL AVIRIS data portal (https://aviris.jpl.nasa.gov/alt_locator/). All files needed to be extracted with Zip7 multiple times to open the (.TAZ) compressed packaging. Images were imported into TerrSet with use of the GDAL protocol. This procedure converted ort.img to Idrisi native .rst format (RST Idrisi). The 224 band images could then be packaged as raster group files (RGF). This format allows for multi- and hyperspectral image sets to be input into a variety of protocols within TerrSet.

Two flight datasets were used for this project: f120930t01p00r11 (4.3 m resolution) and f061002t01p00r11 (16.5 m resolution). The AVIRIS data utilized were captured on 30 September 2012 and 2 October 2006 respectively. The AVIRIS platform collects data in continuous bandwidths ranging from 400 to 2500 nanometers (nm). AVIRIS uses a scanning mirror as a "whisk broom" the scene over the 224 band detectors at the sensor. This project's methods only imported bandwidths or bands between 530 and 745 nm and bands 15 through 43. These bands would cover the EM spectrum from green to near-infrared. Furthermore, it reduced function processing time. There are multitudes of scanned regions on Minnesota's North Shore. The two flights chosen were due to similar time of year capture date, time of day, and azimuth angle. The lower resolution scene also needed some level of accessibility for capturing training data.

The classification analysis utilized previously identified forest community patches. Forest stand data was needed to validate plant community type by percent coverage of classification within the low resolution 16.5 m scene. Minnesota Department of Natural Resources (MN DNR) Forest Stand Inventory was obtained through the Minnesota Geospatial Commons (https://gisdata.mn.gov/dataset/biota-dnr-forest-stand-inventory). The dataset had all the attributes needed for this project's scope. The data is updated, and field checked by individual administration areas throughout the state on an as-needed basis.

All datasets needed to be georeferenced to allow for data overlay and analysis. Imagery from the Farm Service Agency (FSA) through the National Agriculture Imagery Program (NAIP) was acquired to georeferenced the AVIRIS imagery. NAIP produces true color red, green, and blue images with a near-infrared band. NAIP images covering the entire AVIRIS images were downloaded through the USDA Geospatial Data Gateway (https://gdg.sc.egov.usda.gov/GDGHome_DirectDownLoad.aspx). Minnesota's Lake County imagery was obtained for this project.

All self-collected field data were acquired between 10 and 24 September 2017. Ground sampling was performed to establish training data locations within the high resolution 4.3 m

AVIRIS image. Ground truth data locations were selected based on proximity to trail access, Forest Stand Inventory data, and Park Manager notes on specie distribution. Field observations were noted on density of birch trees sampled and areas with homogenous stands to minimize background signal. Areas with steep sloping terrain were avoided when sampling to reduce terrain effects. Samples were distributed diversely through study images. Figure 7 shows ground truth sample locations within the AVIRIS 4.3 m image. Trimble's GeoExplorer 6000 series was used to sample tree species. Pathfinder office was employed to create a data dictionary. The data dictionary used generic point for vegetation format type. One could select for tree species of concern or create comment notes about another feature collected. The Trimble unit employed TerraSync software as a platform on which to collect GIS field data.

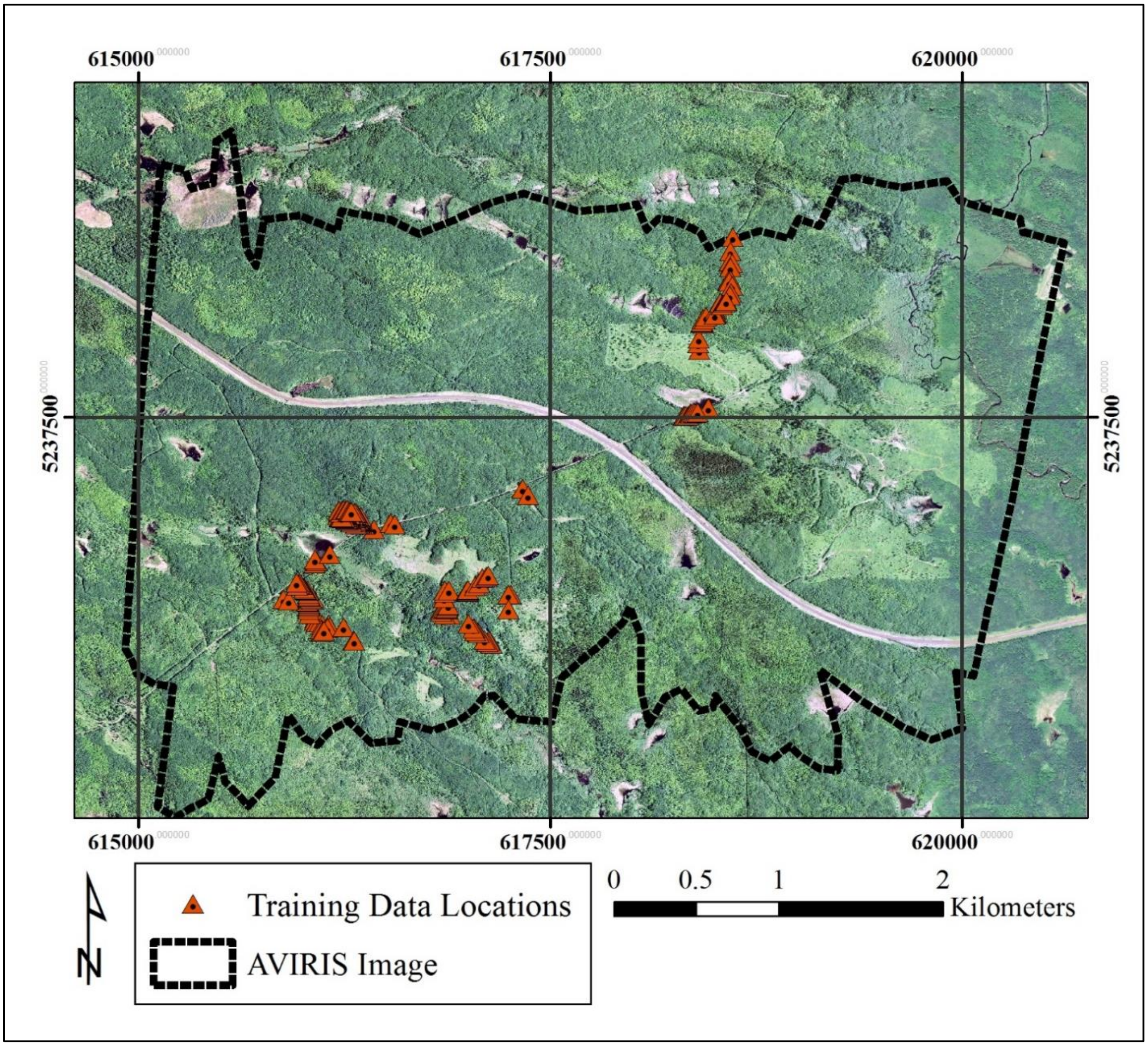

Figure 7 AVIRIS 4.3 m pixel resolution extent and ground truth collection locations

When performing field identification of tree species three reference sources were used: *Minnesota Flora: An Illustrated Guide to the Vascular Plants of Minnesota*, Steve W. Chadde, 2013; *Trees and Shrubs of Minnesota*, 1st Ed. Welby R. Smith, 2008; *Trees of Minnesota: Field Guide*, Stan Tekiela, 2002. Ground sampling utilized alive or recently dead trees with a diameter at breast height (DBH) of equal or less than 20 cm. This would ensure that trees were in the image scene during the acquisition date. Under these parameters, 57 validation tree locations were identified. Table 2 summarizes the data used in this project. Data and layers were integrated to perform classification and analysis. Table 3 show software needs and uses.

Table 2 Data requirements

Dataset	Name	Accuracy/ Precision	Source	Collection Date	Data Size	Data Format
Ground truth sample	N/A	Only used points collected under 4.3 m accuracy	Self-collected	9/10/17-9/24/17	56.02 MB	Shp
MN Forest Stand Inventory	MNDNR Forest Stand Inventory	Community based with good accuracy	MN DNR	9/17/2009	170 MB	Shp
AVIRIS	f120930t01p00r11 and f061002t01p00r11	4.3 m and 16.5 m	NASA-JPL	9/30/12 and 10/2/06	3.67 GB and 0.99 GB	Binary
FSA NAIP Imagery	ortho_1-1_1n_s_mn031(75)_2013_1	1m GSD	FSA through NAIP	2013	3.83 GB zip	Geotiff

Table 3 Programs & uses

Programs	**Uses**
TerrSet	Image classification, atmospheric correction
ArcGIS	Spatial Analyst, figure creation
Pathfinder	Data transfer, conversion, correction, and dictionary
TerraSync	Data collection
Excel	Plot spectra

3.1.2. Bandwidth Selection

In identification of plants remotely this project took advantage of the seasonal variation between different species. Those differences are seen in the autumn as the leaf color changes due to reducing chlorophyll concentrations. Not all bands can identify the change in spectra due to leaf color change. A selective range of bands were exploited in this project to identify species based on that difference. Band selection was based on part of the VIS and the “red edge “ of the NIR regions. The EM spectrum within those regions is where chlorophyll is both absorbed and reflects incoming radiation. The header file accompanying the downloaded AVIRIS image dataset had band center points descriptions. When the header file was converted into a text file it could be read. Only bands covering 530 to 745 nm were selected.

3.1.3 Data Correction

This project used a variety of methods to maintain high data quality standards. It was important to the success of this thesis to georeference the image dataset and process the data for noise. Addressing these factors is fundamental to a favorable outcome for the classifying procedures tested. Data was corrected into the same orthorectified datum. Images used were corrected geographically and atmospherically to allow for information extraction about ground features. This allowed for the two different AVIRIS images to be compared with limited variability. Preprocessing of atmospheric corrections was needed to accurately quantify ground characteristics. TerreSet’s SCREEN procedure was utilized to rule out bandwidths that displayed a high threshold of autocorrelation and low threshold due to particle scattering. This procedure will find and eliminate bands that have significant atmospheric attenuation. Threshold 0.99 and 0.6 were used in this project. Bands 15-43 were grouped in a raster group file (RGF) and input into the procedure.

Image classification required georeferencing, so pixels could be related to ground features and would allow comparison of surveyed reference data. The images were corrected for earth distortions with respect to shape. The raster data image can be manipulated based on control points to correct the image to true orthometric. NASA's JPL packages the AVIRIS data as orthorectified, but for this project it is important to verify pixel position before analysis. Data for this thesis was corrected in TerrSet using RESAMPLE. The nearest neighbor option was used to resample the images. The nearest neighbor resample was chosen because it doesn't change the value of the input cells recorded at the sensor during image acquisition. NAIP tiled mosaic images were used as the output reference file. This increased the number of pixels within the small training data sample size. The NAIP images were clipped into ArcMap to the extent of the 4.3 m and 16.5 m images. This reduced working file size. The clipped extents were exported in TIF format and imported into TerrSet with GDAL in RST format. Band 43 was used for the input reference file in both image resolutions. The resample file specifications used the RGF of Bands 15-43 and the output reference parameters were set to study area clipped. Once images were georeferenced the data frame was registered to an appropriate datum. AVIRIS data was projected in World Geodetic System 1984 and was reprojected into NAD 83 for this project. All datasets used the reference coordinate system of UTM Zone 15 N NAD 83. Both field collected vector data and AVIRIS raster data was validated against the continually operating reference station (CORS) nearest to the study area. The NAIP imagery was tested against the Grand Marais, MN CORS base station (GDMA). This location has a continuous position update every second. Its ground position was: 47 44 54.75712 N, 090 20 28.47142 W and in NAD83.

Ground control points (GCP) were established to resample the AVIRIS images. The AVIRIS and NAIP scenes were georeferenced based on easily identifiable ground features, such as road crossings, field corners, and lake points. Figure 8 shows an example of ground feature utilized to georeferenced the AVIRIS imagery. Rocks along the lake shore that were prominent

Figure 8 Georeferenced GCP location example

across historical images were used. Past glacial events removed much of the top soil within this region exposing prominent bedrock outcroppings. Those outcroppings tend to remain static over time and were used as GCPs. Lakeshores themselves change over time and were not used as GCPs. The GCPs used were spread throughout the image capturing all four quadrants of the image. The final selection of GCPs was based on trial and error. Points that produced the lowest root mean squared value were kept.

The 4.3 m image established 12 GCPs to georeference the image. Readily identifiable objects within both scenes were matched. From the 12, only GCPs with residual errors under 0.5 (image resolution (4.3)) were used. The AVIRIS 16.5 m image used 5 GCPs. Table 4 shows the x and y locations used to register the AVIRIS images. The clipped output reference area can be seen in Table 5. The AVIRIS images were clipped using ArcMap's image analysis tool to a smaller extent and saved as a .tif, which could be imported into TerrSet for use. The study area extent was chosen based on MN Forest Stand Inventory polygons within the 16.5 m AVIRIS scene and inside the NSH region. The clipped images served to reduce function processing time and limit file size. The resampling type used was a nearest neighbor so to not alter original values contained within the images. Furthermore, a first order linear map function was used to reduce the time to process and reduce the final number of GCPs needed. The 16.5 m scene was clipped before images were georeferenced.

Table 4 Resampled images GCPs used for the 4.3 m/ 16.5 m scenes

Pixel Resolution/ ID	Input X	Input Y	Output X	Output Y
4.3 / 1	622666.7988	5237020.9083	617885.5656	5237317.0642
4.3 / 2	622092.4844	5235131.7176	615934.8548	5237618.5533
4.3 / 3	623456.5614	5234542.7600	615547.6336	5236206.7594
4.3 / 6	622915.8458	5237848.4642	619633.4516	5237309.8924
16.5 / 1	596556.6298	5271691.7982	603985.7767	5283454.6238
16.5 / 3	593447.6751	5228298.3777	643381.8895	5265005.5165
16.5 / 4	602752.2926	5230394.3880	644763.9992	5274440.4801
16.5 / 5	600109.3798	5260353.2769	615848.3665	5282707.9713

Table 5 Output reference parameters

Pixel resolution	4.3 m	16.5 m
Number of columns	1022	785
Number of rows	1410	4305
Minimum X coordinate	620441.2	638528.8
Maximum X coordinate	624835.8	646094.2
Minimum Y coordinate	5233857.1	5265209.2
Maximum Y coordinate	5239920.1	5271638.0

3.1.3. Data Integration

Vector data collected in TerraSync was transferred to shapefile to pull the layer into ArcMap for use in the classification of the raster image. Once downloaded the imaged flight data from the AVIRIS platform had to be uncompressed from the .tar zipped file format. Once the file is unzipped the metadata could be explored. The data needed a new header and extension file to allow import of AVIRIS binary raster data into ArcMap.

All self-collected tree data was cleaned on return from the field. Only points collected with positive identification and horizontal accuracies under pixel resolution were used in this work. The collected tree data was split into training and validation sample points to avoid over fitting of the classification model. This study used a simple split of points collected to assess initial accuracies (70% training and 30% as validation). This was accomplished by creating a new ascending number column joined to the selected paper birch data. A Random number set generator was used to select the 70% training data vs. the 30% validation data. Surveyed Forest Inventory was assessed to ground truth species present at validation locations. These crosssections of birch habitat aligned with MNDNR forest inventory birch habitat. Areas within the 16.5 m image were delineated to validate present or absences of vegetation community type stated in that layer. Areas within the 16.5 m stated as “Paper Birch” were bisected to record species present. Only polygons with paper birch communities were used to perform image analysis.

The MN Forest Stand Inventory layer was processed to reduce the size of the working layer. Attributes were selected in Lake County, 2006 or later acquisition, cover type of paper birch, specie type of paper birch, DBH category of 4 or greater, and areas within clipped 16.5 m scene.

The North Dakota Tree Handbook states that paper birch canopies at maturity can be observed from 6 to 12 m. A 9 m average canopy width was assumed and buffered to create polygons to establish areas of training data pixels. For each birch tree location, a 4.5 m buffer was used to account for the 9 m diameter canopy coverage to establish training pixel area. The buffer of 4.5 m was applied to the 57 validation tree locations. Then the buffered tree canopy area was rasterized. The produced raster was compared to the linear hyperunmix classification coverage for analysis. Figure 9 shows an inset area of the 4.3 m scene with buffers applied, then buffered areas rasterized.

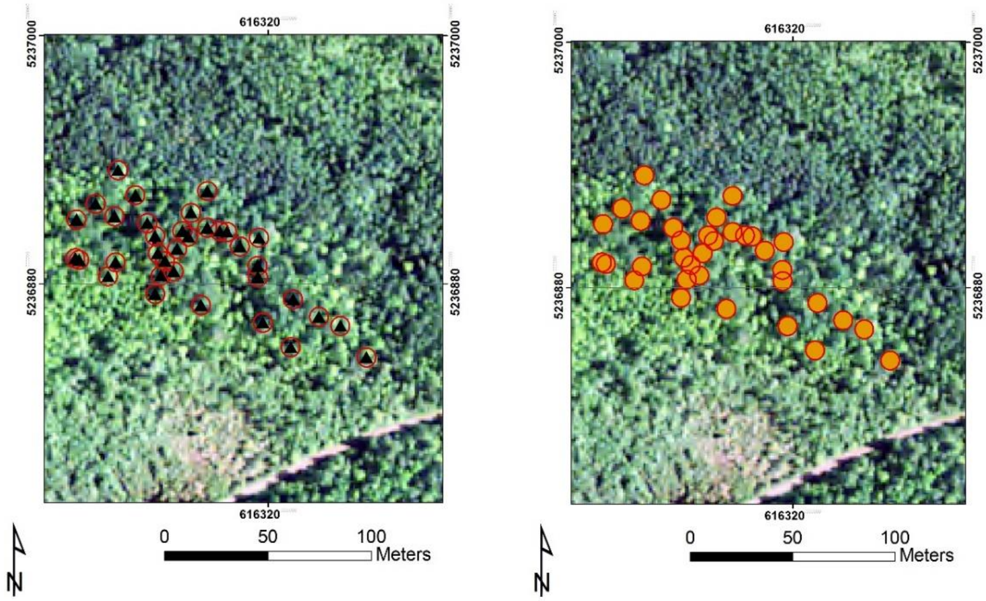

Figure 9 Paper birch locations with buffer (left image), then canopy area rasterized (right image).

3.2. Image Classification

This project grouped pixels in an image by similar spectral signatures. Using the ground truth samples at the 4.3 m site, paper birch signatures were selected for areas with birch present. Supervised classification utilized training points collected in the field. The supervised classifications were performed across the entire 4.3 m raster image. The supervised class produced served to find the optimal spectral signature to identify birch on the landscape. This is the simplification or generalization of the true complexity of the natural scene.

Pure signatures of paper birch at this resolution were expected to be rare and most pixels to exhibit a mixture of flora. Buffered polygons corresponding to the 70% training data was imported as a vectored polygon into TerrSet. The training data polygon was used as the input to base pixel selection in TerrSet's HYPERUNMIX hyperspectral image classifier. All 4.3 m band scenes were grouped as an RGF and input into the classification. Other readily identifiable habitats were selected in TerrSet's MAKESIG to reduce residual areas of scene.

Image interpretation of community type was performed and utilized for this experiment. Easily identifiable informational groups were chosen to develop a paper birch spectral signature. A minimum of 100 pixels was needed to produce each in the selected informational group polygons. Selected groupings included: evergreens, upland grasses, wetland/ marsh, open water, sparse vegetation and shadows. These signatures were grouped with the produced paper birch vector signature and was run in the HYPERUNMIX function. The produced paper birch signature was used against the 16.5 m image. Only a clipped extent of the 16.5 m was classified for this project based on output references.

The linear unmixing model assumes that pixels can be modeled as a linear function. The linear function requires other classes to project classes against one another as a linear function. Harrison et al. (1991) provide a simplistic calculation of the linear unmixing model. The multispectral dataset in n layers (26 bands), with y endmember types (7 classes), x is the observed reflectance (n *1), and f is the unknown endmember proportions (y *1). The goal is to calculate f(x):

$$\mathbf{x=M*f+e}$$

M is the endmember matrix (n*y), and e is recorded noise. Calculate f by standard least squares fit:

$$\mathbf{(x - M*f)*(x - M*f)}$$

In this system one cannot have more classes than bands. HYPERUNMIX performs these calculations for each pixel to classify endmember proportions.

3.3. Spatial Analysis

To perform spatial analysis on both the 4.3 m and 16.5 m AVIRIS scene the validation polygons needed to be rasterized. This allowed pixels to be directly compared. The 30 % of ground truth data collected was input through the zonal statistics tool in ArcMap. This produced area of pixels the birch trees occupied. The same was done for the selected MNDNR forest stand inventory polygons. Reclassification of the produced classified raster images was performed to group by percent of signature into ranges. Table 6 shows reclassified ranges in TerrSet.

Table 6 Reclassified ranges categories

1	0-0.75
2	0.75-0.85
3	0.85-0.95
4	0.95-1.1

Accuracy analysis for the classified AVIRIS images assess area of classifications coverage compared to validation areas. A zonal statistic was again performed to produce pixel coverage. Validation trees were also averaged for distance from 0.95 signature classification. Validation results of the test polygons on both scene resolutions were based on whether a 95% signature similarity was classified within the test polygon.

A confusion matrix was established for the 4.3 m validation site. This matrix compared predicted classes with > 95 % similar signature as the training data to < 95 % signatures. Stratified random points were sampled throughout the image. 50 points were randomly sampled in both areas above and below 95 % similar signature. These points were the compared to known reference data to complete the confusion matrix. Confusion matrix statistics were calculated based on the completed matrix.

Chapter 4 Results

This chapter outlines the results of the linear unmixing classifier to predict known paper birch locations. Once the spectral signatures were developed based on individual tree canopies the focus of this project was to overlay the classifications and compare the resulting spatial coverages. The results produced categories representing percent spectral signature similarities and the corresponding coverages of the MN DNR forest stand inventory layer. Paper birch classifications were based on both individual tree spectra captured from hyperspectral data and field collected vector data. The methods were used to find outcomes when using individual tree spectral signatures to predict known community level specie occurrence. This project hypothesized these methods would produce elevated levels (90%) of classified pixel coverage of known validation polygons based on literature review. The following sections of this chapter give insight into produced results.

After all the methods were applied the linear unmixing classification produced low coverage on analysis. Both coverage of the training data 4.3 m site and the 16.5 m validation site experienced low coverages results ranging roughly from 7% - 30%. Validation results of the test polygons were based on whether a 95% signature similarity was classified within the test polygon. This led to higher levels of validation polygons identified as classifying the informational group. This approach to polygon validation identified 44% of the individual tree locations and an assumed higher rate of 93% of community level polygons identified. These results show the ability of the methods to produce generalized coverage maps of birch trees.

Processing and analysis required substantial interactions from users and further automation is needed to expedite these methods for use. These results can only be interpreted through statistical analysis. Direct association between pixel results and ground features are not assumed by the spatial results, but contributions of individual paper birch are assumed by pixel mixtures produced. The resulting linear unmixing model classifier produced is shown in Figure 10. This map is of the 16.5 m resolution scene clipped to the working extent. The map shows the coverage of the 16.5 m classification and canopy validation areas.

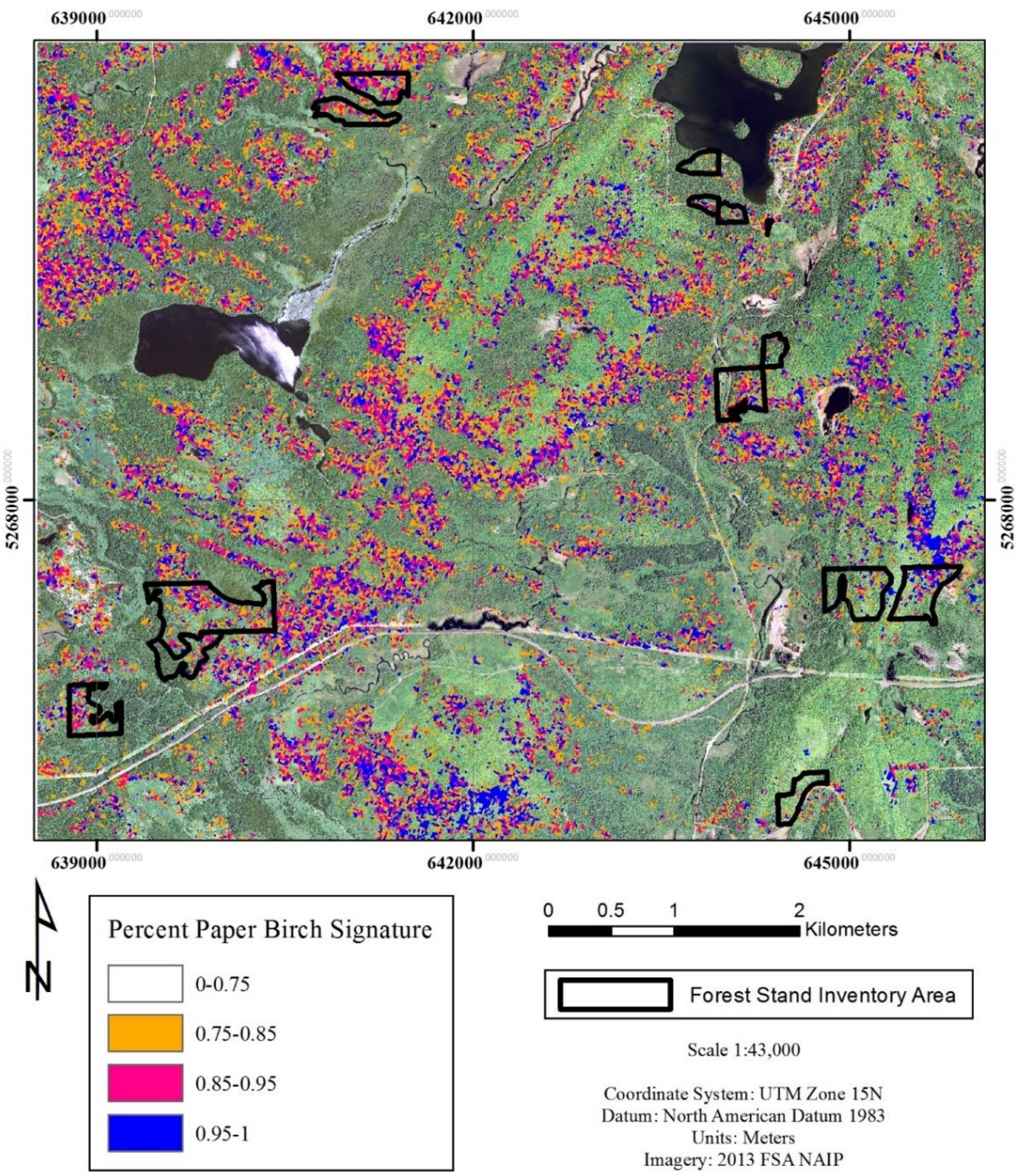

Figure 10 Linear unmixing model classification of the 16.5 m AVIRIS image with validation forest plots

4.1. Data Inspection & Correction Results

AVIRIS imagery, NAIP imagery, and validation vector layers were needed to predict paper birch coverage maps. All data for this project would require exploration and correction when necessary.

4.1.1. Image Georeferencing

Alignment of AVIRIS data (4.3 m and 16.5 m pixel resolution) was fundamental to assure classified pixels represented ground truth data. The AVIRIS scenes were georeferenced to NAIP imagery. Ideal GCPs would have easily identifiable locations i.e. (road crossings, structures, and monuments) within the image. The nature of this study area was remote. Furthermore, this led to sparse numbers of permanent and well-defined GCP locations. Figure 11 show a typical GCP used at the 4.3 m training data site.

Figure 11 Intersection of dirt road and mowed tree line used as GCP.

Ground control points were established and described in section (3.1.2.). Out of the 12 points identified in the 4.3 m scene, 4 points were used (1, 2, 3, and 6). The set of GCPs had a residual error of less than 0.5*(4.3 m) and the Root-Mean Square (RMS) of this set was 0.836196. The RMS is an estimate of the average error within the points that have been selected. The RMS does not average the error of the entire image. To achieve an RMS of equal to or less than half the scenes spatial resolution, GCPs with high residual errors were not selected for the transformation.

The 16.5 m scene used GCPs 1, 3, 4, and 5. These GCPs had residual errors of less than 0.5*(16.5 m) and an RMS of 0.498495. Table 5 in section (3.1.2.) showed the GCPs and output parameters used for both 4.3 m and 16.5 m images. The distribution of points for the resample were chosen based on the lowest produced RMS. The images were then georeferenced into geographic coordinate system UTM 15N, and datum North American Datum1983. Only bands 18 through 43 were resampled.

4.1.2. Atmospheric Scattering and Spatial Autocorrelation

Image classification can be altered due to atmospheric distortion; thus, it was important to process the data for possible distortion. The screen procedures were used to remove bandwidths with autocorrelation above 0.99 and lower than 0.6 due to atmospheric scattering. The screen function was used on bands within the EM spectral range in question (bands 15-43). The function resulted in 26 bands kept and 3 bands (bands 15- 17) removed for autocorrelation. This procedure was run on only bands 18-43 of the 16.5 m AVIRIS because 15-17 were removed. Results (see, Figure 12) found no bands were removed due to scattering, but 3 bands were removed on the 4.3 m scene due to autocorrelation.

Both scenes images were acquired under clear conditions with no clouds or haze. Though both scenes had a lack of clouds during image acquisition atmospheric distortions may still crop up and are recovered through the autocorrelation testing performed. Those distortions can be caused by particulate matter in the atmosphere.

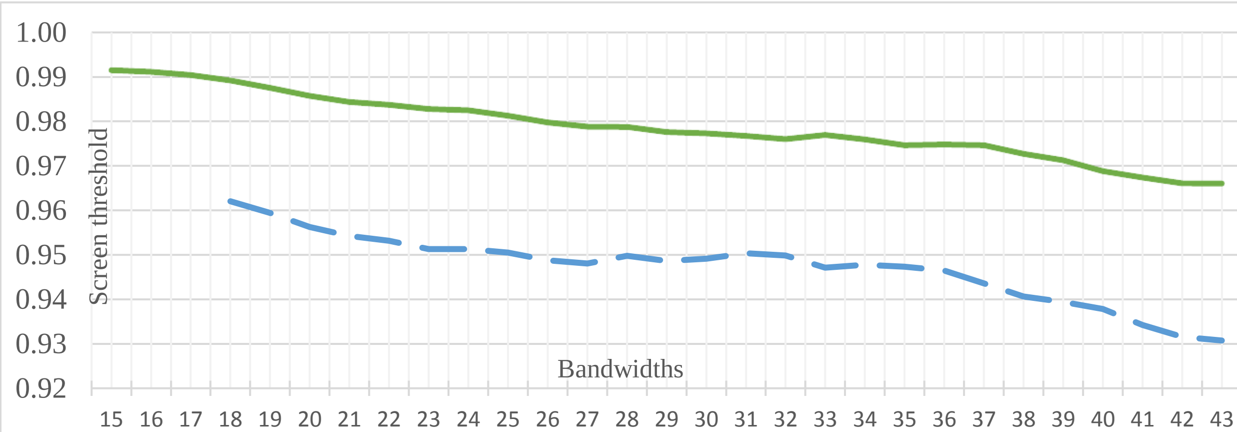

Figure 12 Screen procedure results for bands 15 – 43 in AVIRIS dataset, solid green line (4.3 m dataset), dash blue line (16.5 m dataset)

4.2. Classification Results

Classification was split into two major groupings: the 4.3 m training site and the 16.5 m validation site. Both resolution images were classified and assessed for coverage of known tree locations.

4.2.1. Training Site 4.3 m Resolution

Data selection played an integral part in the project's methodology. The training data selection process produced 189 valid paper birch locations within the 4.3 m resolution image. These 189-birch tree locations were used to develop signature files and validation data. This was after selecting all trees with a DBH of greater or equal to 20 cm and GPS accuracies under pixel resolution. A simple split of training and validation data was performed. Out of the 189-tree sampled 132 trees were used for training data vector locations, which equals ~ 70% of total birch samples. These were the points used to establish a supervised classification of paper birch and create a spectral signature group file. The other ~30% or 57 paper birch locations were used for validation points. The total number of pixels in the buffered validation area was 3338.

The linear unmixing model returned pixels with percent similar signature as the training dataset. The classification was reclassed to see areas of stratified pixels. Pixels with a match of less than 75% similar signature were not investigated in this project.

The initial 4.3 m scene classification identified 25 out of 57 (44%) validation trees as a 95% signature match. Zonal statistics of the 4.3 m image produced 29.1% canopy coverage with a signature match of 75% or greater. Lower classified coverages were produced as the percent signature match increased. The 4.3 m scene had low percent coverage seen in Table 7.

Table 7 Training data site (4.3 m) image analysis results.

Signature match (%)	Number of pixel	Percent coverage
>75	971	29.1
>85	673	20.2
>95	428	12.8

The linear unmixing classification produces a residual map seen in Figure 13. The higher the value displayed in the residual map the less likely the pixels can be produced by the training data. The residual map shows the NW corner of the unmix classification had a lower signature match than some of the central regions of the image. The NW corner was not accessed during ground truth data collection and could be a reason for the results displayed.

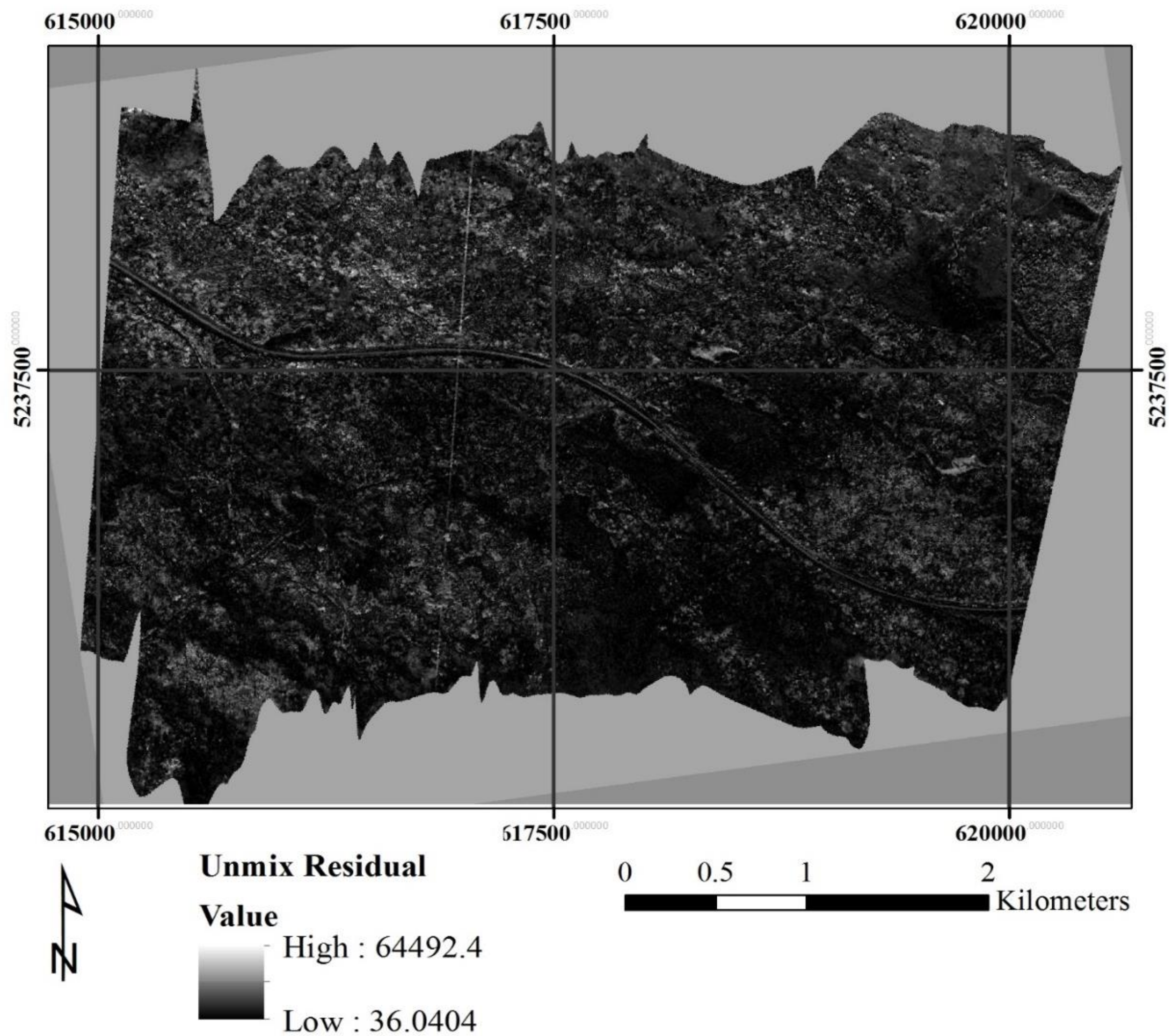

Figure 13 Residual map of 4.3 m AVIRIS image

The 4.3 m linear unmixing classifciation produced three categories of spectral similarities and each category was compared to pixel coverage of the 9 m diameter paper birch canopy buffer (Figure 14).

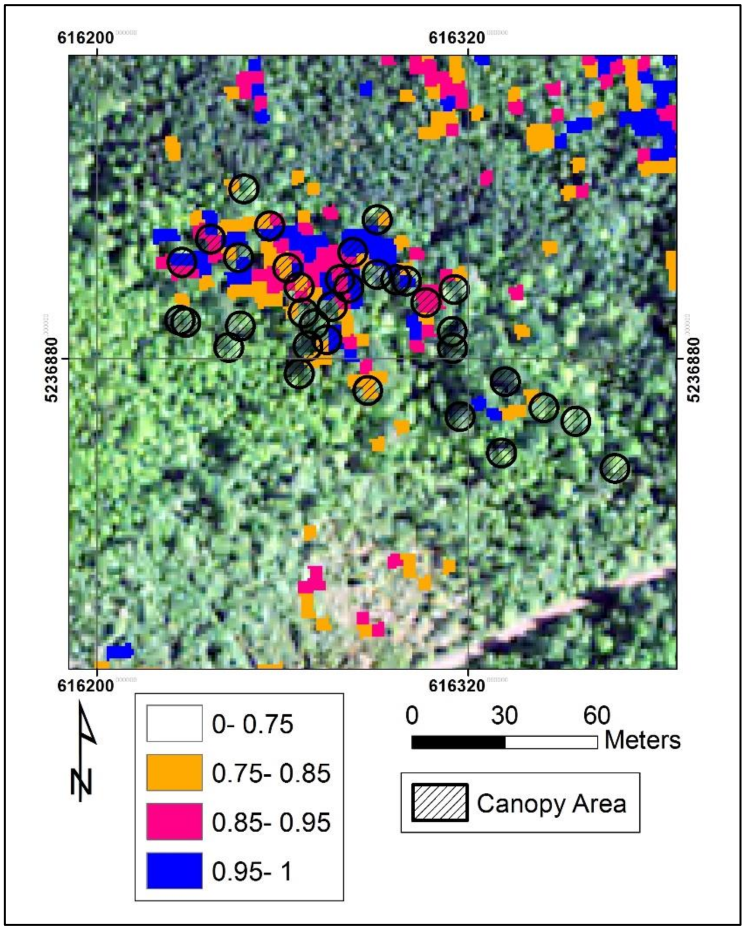

Figure 14 Insert location in the 4.3 m training data site classification results and canopy coverage

Only percent polygon coverages were used in this project result due to only paper birch locations sampled. Analysis was exclusively performed on pixels within paper birch canopies. Furthermore, the actual pixels classified as paper birch inside the validation buffer should have been 100%. Table 8 below shows a confusion matrix and statistics on accuracies. It shows a non-significant kappa value of 0.13. This is repressed as a low agreement of predicted classification compared to the referenced ground truth data.

Table 8 Confusion matrix for the 4.3 validation canopies with statistics

		Actual Cover Class		
		> 95 %	< 95 %	Total
Predicted Cover Class	> 95 %	10	2	12
	< 95 %	44	44	88
	Total	54	46	100

Overall accuracy = 54 %
Error of omission = 81.5 %
Error of commission = 16.7 %
Producers accuracy = 18.5 %
User accuracy = 83.3 %
Kappa = 0.13

4.2.2. Validation Site 16.5 m Resolution

Validation polygons were obtained from the MN DNR with similar proximity to Lake Superior as the training site and similar time of year image acquisition date. There were 14 forest stand inventory plots used as validation polygons of paper birch communities. These validation polygons covered an area of 128.1 ha. The developed spectral signatures were applied to the community level polygons to validate previously sampled field data.

The 4.3 m paper birch spectral signature was applied to the 16.5 m scene and validated 13 out of 14 (93%) forest plots as having some picture elements with 95% similar signature. Though a high percent of the forest plots were identified as having similar spectral signatures as paper birch, the coverage of the validation plots was low as seen in Table 9. The total number of pixels used was 5015 for the 16.5 m image. The overall forest plot coverage of the 16.5 m classification was 30.1% with a signature match 75% or more.

Table 9 Validation data site (16.5 m) image analysis results.

Signature match (%)	Number of 16.5(m) pixel	Percent coverage
> 75	1510	30.1
> 85	877	17.5
> 95	365	7.3

The produced residual map shown in Figure 15 is of the paper birch signature to emulate similar pixel signatures. The low values represented in black are pixels with a good ability to match those pixels and higher values in grey are a poor pixel match to represent that signature. The map illustrates low ability to characterize cover classes: open water, spare vegetation, evergreens, and upland grasses. This was an expected result. The resulting percent coverage was mapped by mean area covered from low-high and shown in Figure 16. Then an insert map is shown in Figure 17 that illustrates two validation polygons with differencing levels of cover classification.

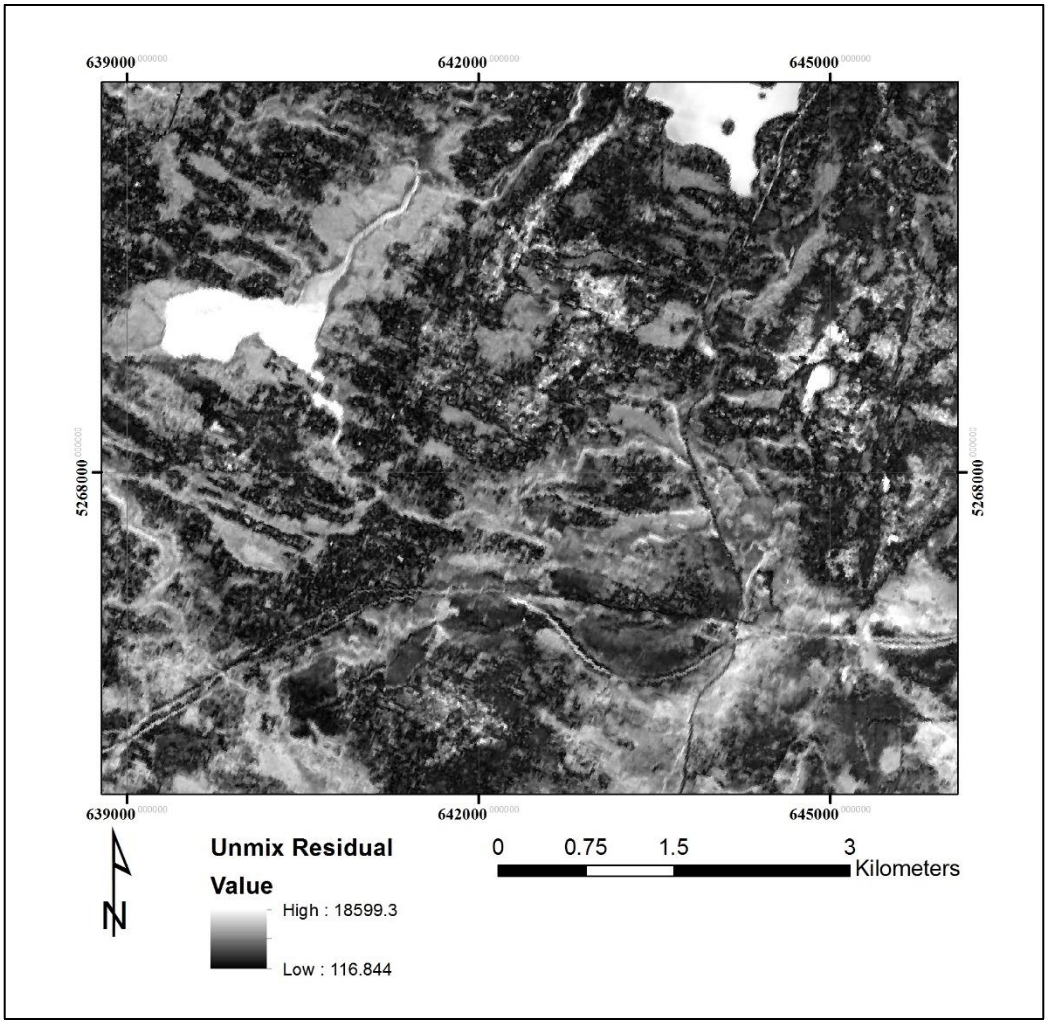

Figure 15 Residual map of 16.5 m AVIRIS image.

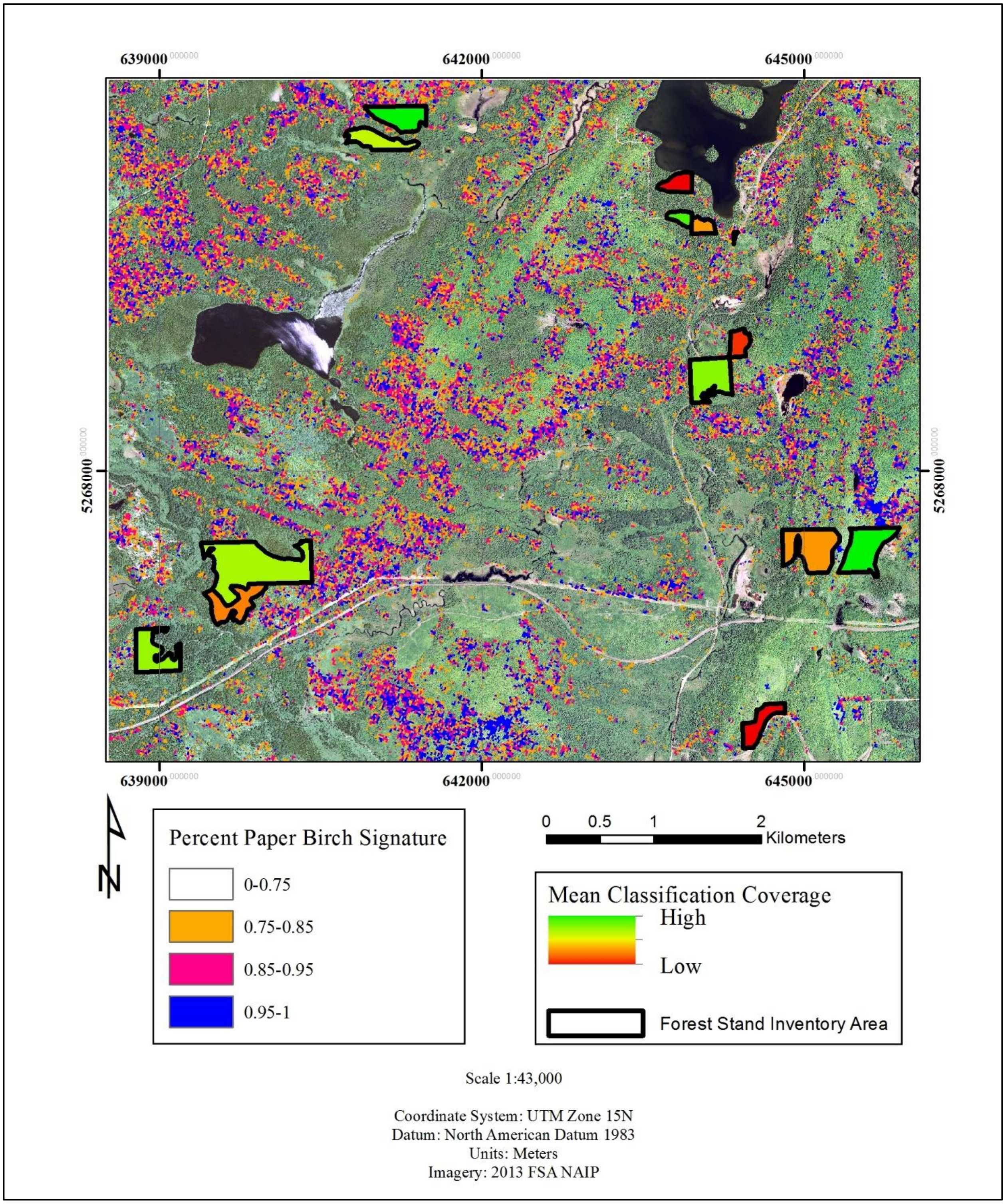

Figure 16 Validation site (16.5 m) with average coverage of forest stand inventory plots.

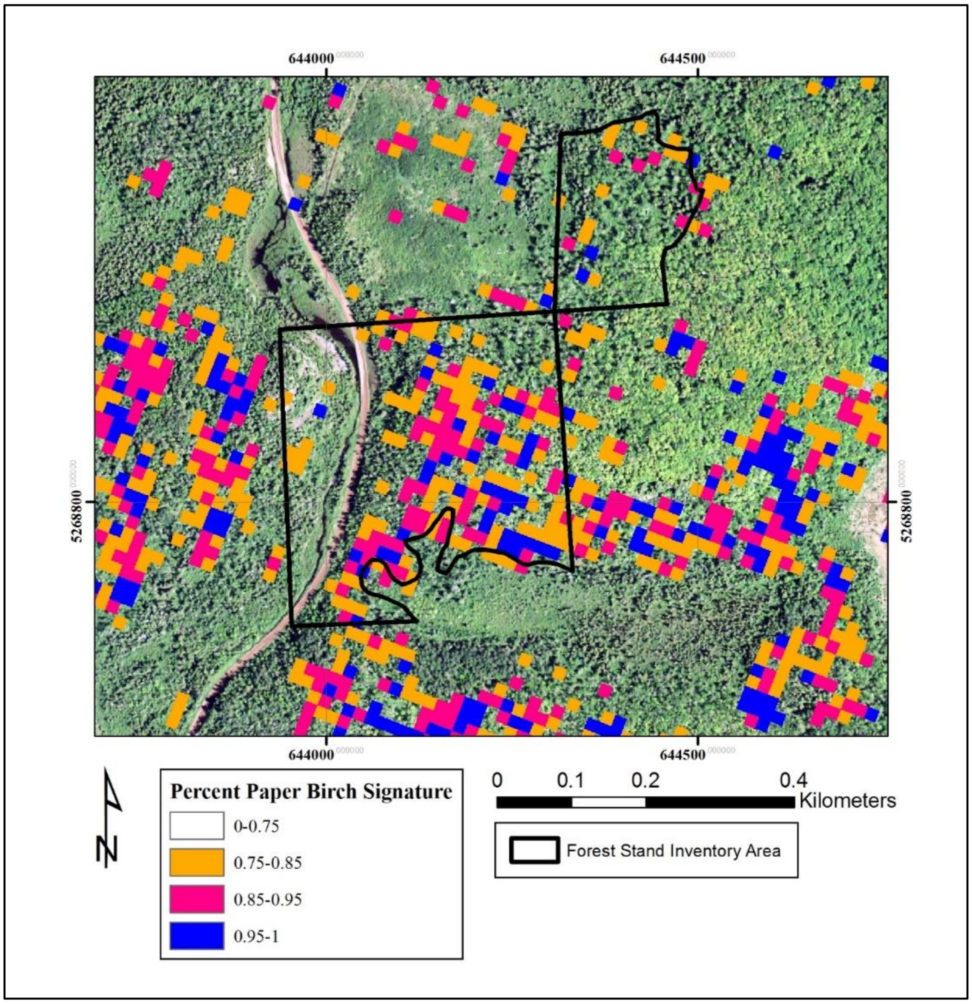

Figure 17 Insert map of the 16.5 m validation site displays two forest stand plots with differencing levels of pixel coverage.

False returns were seen throughout the 16.5 m image by photo interpretation. Paper birch signature was seen in pixels ranging from roads to wetlands. The occurrence of false pixel classification was low and assumed to be similar to the 4.3 m classification due to the confusion matrix. This needed to be noted and will need to be investigated further. Figure 18 shows paper birch classification signature in a wetland West of Paccini Lake.

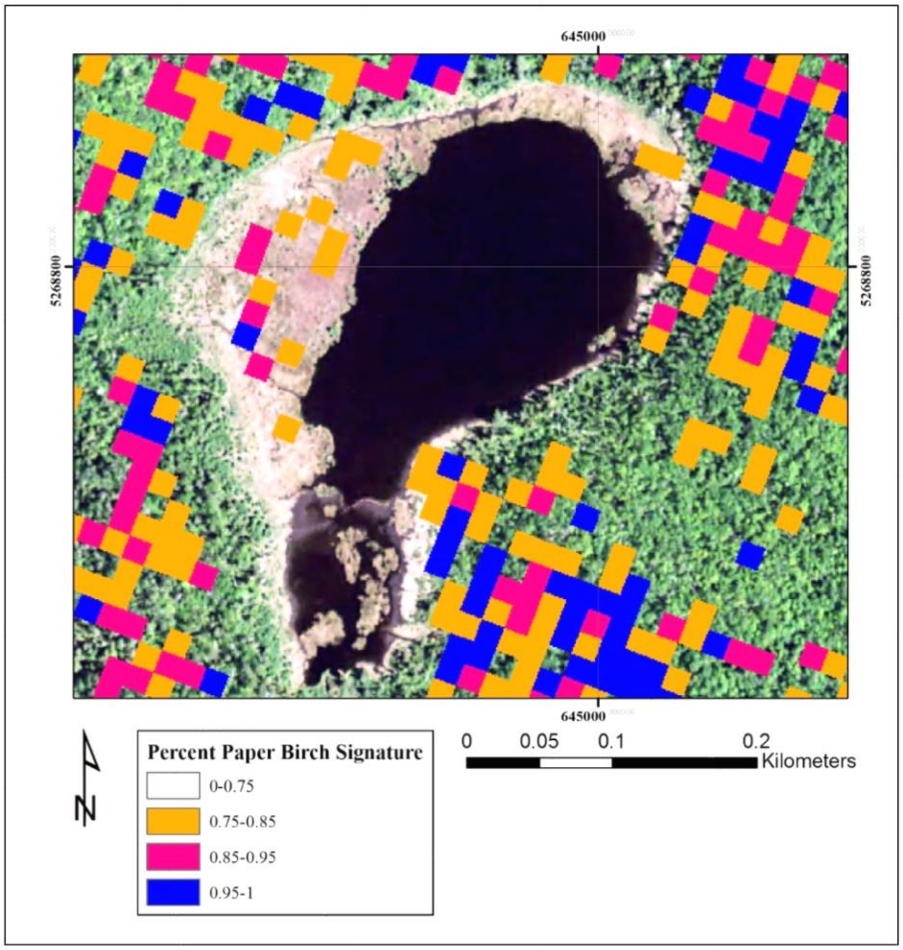

Figure 18 Image interpretation shows false return of paper birch spectral signature within wetland on Westside of Paccini Lake.

Chapter 5 Discussion and Conclusions

Traditionally, large scale species delineation demands intensive field work. This includes ancillary data collection and visual estimation. That process can be costly and time-consuming due to accessibility (Lee et al. 1996). Remote sensing delivers economical means to discriminate and estimate the physical properties of species. Due to the losses in paper birch population regeneration it is important to utilize every reasonable measure to understand and monitor this problem. Mapping specie distribution, quality, and quantity are critical tasks for management of forests. Furthermore, it is necessary to regularly update spatial information about the extent and quality of vegetation to manage these ecosystems effectively (He et al. 2005).

This project had both strengths and limitations. This chapter assesses the results of the work and applications. The results should be viewed as exploratory and accuracies of methods could be improved with further research that is described in this chapter.

5.1. Method's Strengths

This study's methods resulted in overall low accuracies for the attempted identification. These methodologies have not been applied to such a specific task before and results going into analysis were unknown. The input considered in this project is an example of specific sets of circumstances. The classification was produced to give the best likelihood of accurately classifying paper birch. Though the produced accuracies were lower than sought there was much to takeaway form this project.

The resulting paper birch signatures produced classifications of >95% membership that on average were within 6 m of ground truth samples. Also, the user accuracy was calculated at 83 %. This shows how often the predicted class will be present on the ground. These methods provide forest manager or researcher the ability to verify general paper birch locations. The resulting data is usable to update current MN Forest Stand Inventory data.

Complex topography and convoluted mosaic of plants can affect sampling of field data. The study are had areas inaccessible due to complex topography, but for standardizing spectra responses flat terrain was needed. Avoidance of steep terrain was easy due to the flat nature of the 4.3m sample sites. These methods were also able to exclude autocorrelation of pixels when screened. This was a quick check for data quality. No scattering was observed in screening for atmospheric interactions. This may be due to low altitude image acquisition flight path and no reported haze on acquisition date. The preparation of data for this project was sound and could setup a future project with different classification methodologies.

This study used limited training data to produce paper birch signatures. The process for finding paper birch in the higher resolution AVIRIS image was random. If forest management knew of paper birch locations prior to field work, they could quickly obtain GPS locations and create a signature group. The other informational grouping used for classification were simple to identify intentionally. This makes for rapid characterization of the imaged scene. Reduced datasets used for this experiment increased turnaround time for forest managers too. The Gitelson (1994) experiment showed that several functions of reflectance were directly proportional to chlorophyll concentrations. Therefore, a reduced spectral resolution was used to investigate the signature created by chlorophyll. This smaller dataset reduced file size and then processing time for this project. This is an advantage to quickly producing a cover classification.

This project provided a learning opportunity to study the many facets of remote sensing. It added variables that together were complicit in low accuracies. This shows that future work needs to do more to limit the scope of paper birch classification to get a better handle on what variables cause the most variation in the results.

5.2. Method's Limitations

Classification error can arise from several sources including: orbital position, clock error, atmospheric affects, receiver noise, and multipath error (Campbell 2011). Limitations must be considered as part of the cost and benefit analysis that all remote sensing specialists must acknowledge. This project focused on limited variables to test methodologies. The reality is that this classification methodology is an extreme simplification of the natural phenomena. Limitations to this project are presented by both a short research window and level of data that can be obtained.

The methods used in this project obtained lower than sought classification goals for paper birch on the North Shore. Many factors have influenced the results of this project. The forest stand inventory was not homogeneous in paper birch. This dataset identifies the plots of dominate species and densities of that species. The forest areas sampled in the 4.3 m site exhibited high species diversity with few homogenous groupings of paper birch. Terrestrial plants will occur in large saturation compared to other community groupings (Elgadi 2010). Largest samples of paper birch training data utilized were dense stand samples. This may underestimate the true variability of paper birch on the NSH, due to spatial autocorrelation. Furthermore, autocorrelation of trees was inevitable because paper birch can arise from asexual parental rhizomes. This leads to pockets of paper birch groves.

This project utilized both reprojecting and resampling procedures. Further research into AVIRIS images used showed that a simple rotation of the image was all that was needed and not a reprojection. The resample procedure added error into this project considering the NAIP images were used as the output standard. This increased pixel numbers within the canopy area which increased variation of a “pure spectra”. These functions were used because of the availability of the products within the Idrisi TerrSet software. They were not necessary for this project and introduced error into ultimate results.

Crown density was not reviewed in this research and could introduce background noise into signal. Figure 19 illustrates the typical paper birch canopy seen at the 4.3 m AVIRIS site. It shows maple trees in the understory and a loose canopy morphology of paper birch trees.

Figure 19 Typical paper birch ground truth sample.

Tree selection was important to the overall success of this project. In researching the historic aerial photos of the 4.3 m scene primitive roads were identified as access points to gather data throughout the study area. The actual nature of the ground sampling was not as accessible. Trees selected may not represent the true scene variation due to accessibility issues. Also, the extent of the study area and the number of specific class used will affect the results of a classification (Woodcock 1987). Sparse number of samples collected were usable due to quality standards. The small number of training data samples led to not being able to stratify data and identify best samples to base classification. The random number generator could have produced different results by choice. Some trees will have a better or worst paper birch signature. Utilizing a random number generator, the quality of the signature is chosen at random and accuracies will be affected.

This project had an assumption of tree canopy buffer size. The diameter of tree could add error to the classification. Tree canopies most likely had a variety of diameters. Averaging tree canopy size most likely captures non-paper birch spectra. The application of these buffers could have increased mixed pixels in the scene. Understory attributes were noted in a field notebook, but not applied to methods. Mixing of vegetation and soil reflectance will occur with some mixed pixels. Furthermore, the timing of this project's data acquisition is narrow. Acquisition of data just before or after leaf-off will lead to similarities between ground reference and dead leaf reflectance. This project's method is only applicable with fall image acquisition and need to be considered in further work. Mixed pixels will diversify the true signature of birch on the landscape and affect resulting classifications. Each pixel's spectral response or "signature" can be unique identifiers of pixels under ideal condition, but natural variation in the plants can vary the spectral response (Campbell 2011).

Gradient and age of stands were not used, and the selection of a diverse age group of birch introduced spectral variability into the experiment. All polygons used in the 16.5 m validation scene had an average tree age of 70.7 years old. They also were noted to have paper birch in decline at survey date. This influenced spectral output of specimens, but not examined.

The final classification methodologies used only 26 bands. Even with increased spectral bands and reduced widths it is difficult to classify growth forms, because they also increase fluctuation within and between spectral groups (Ustin 2010). Utilizing a reduced number of bands was beneficial to data processing time, but with less bandwidth range was not appropriate to resolve all endmembers involved in classifying. Further work utilizing the full range of bandwidths may prove favorable for classifications accuracies. Many studies will remove bands centered on both atmospheric water absorption and at the end of the EM spectrum due to noise (Fassnacht et al., 2016). Spectral curves across species do not change that much, the level of reflectance is the driver of specie contrast and can be exploited. This contrast between species can be greatest in the NIR. The reflectance in the NIR is primarily controlled by leaf morphology. This would need to be discussed if used for further work in this domain.

5.3. Further Research

This project shows inherent problems in classifying specie type on a landscape. That means there is much opportunity to increase accuracies through adapting methodologies. Further investigation into this domain will likely resolve part of the limitations. The use of remote sensing tools like ArcGIS will have to be explored further. More research will be needed on ground truthing surveys prior to going into the field.

Investigation of intra-specie variation is needed to be assessed if future work is done on a subpixel level. Soft classifiers could be used to see if these classifiers do not assume homogeny. Adding similar broadleaf species data to the linear modelling function could help narrow signature to assess inter-species differences. Or comprehensive spectral libraries are needed to delineate plant species under different conditions. Ground spectrometers could then be added to project to capture best spectra for training data. The understory vegetation was not reviewed in this project, but future classification could take into effect known understory vegetation.

The USGS Spectroscopy lab studies methods for classification through spectroscopic remote sensing. There they have full spectra libraries of spectral signatures of minerals, vegetation, and more. The cover classification produced in this project could be compared to groupings with spectral curves mapped. Paper birch is not in their digital spectral library yet, but species with similar morphology to birch like aspen have been investigated. They have recorded spectra for all major seasonal variations within the specie. This could be used as training data for future similar work given the timing of image acquisition.

Finally, future assessments could place more stratified random points. The strata would be based on the collected ground truth refence data or prior known surveys. A traditional confusion matrix would then be applied to the harden classifications. This could be produced within the MN Forest Stand Inventory data. These stratified points would be buffered to the 9 m diameter, turned into raster, and compared directly to canopy coverages. Future work could use the soft classification of paper birch in the multiple image format and automate the process as a hard classification. This would assign the pixels with the highest percentage of paper birch signature into one membership grouping.

5.4. Applications of Research

Northeastern Minnesota is seen by the NFSC as a potential area of refuge for boreal species. Adapting to climate change will need action plans that can anticipate and respond to specie decline. There are many partners involved in managing and protecting these forests. Both the MN DNR and the USDA Forest Service have some of the largest tracks of land within the NSH. This project would have the greatest impact if utilize by them due to the resources involved for both of those public organizations. This research helps to identify general paper birch stand locations. This project recommends that those stand locations should be crossreferenced against the best sites to retain paper birch populations for long-term protection e.g. topography, hydrology, soil type. Pure spectral signature of paper birch is hard to come by within the NSH. This project showed there is still a dominate percent of this habitat occupied by paper birch. Areas classified as greater than 95 percent paper birch signature should be considered as dense habitat and identified as priority locations for paper birch refuge. This would facilitate the protection of paper birch by prioritizing management allocation on the landscape.

Forest communities are traditionally grouped into dominant specie type. This linear unmixing approach assumes that large pixel resolution is diverse with multiple additive spectral responses. The species boundaries are gradual and can be present in more than one community type. The soft classifier is a valid approach to establishing methods for classifying the diverse forest landscapes. This project's method shows a way in which not to obtain high accuracies. This knowledge will help future researchers with methods to more appropriately predict mixed forest tree species.

Further GIS analysis and overlay would help forest managers to assess paper birch change over time. The land cover change over time is detected by the comparison of two dates of imagery. Birch tree loss overlay mapping could compare variables that affect paper birch i.e. drought, soil compactness, historically logged areas, and deer population browsing. A "hot spot" analysis could be produced with percent coverage of classification. Such a map would be significant to forest restoration groups for funding opportunities and land cover change monitoring. Being able to produce accurate specie locations based on spectral signatures would greatly reduce cost of monitoring and resource exploration.

References

Abrams, Marc and Mark Kubiske. 1990. “Leaf structural characteristics of 31 hardwood and conifer tree species in central Wisconsin: Influence of light regime and shade-tolerance rank.” *Forest Ecology and Management* 31, no. 4 (April): 245-253.

Atkinson, P. M., M. E. J. Cutler and H. Lewis. 1997. “Mapping sub-pixel proportional land cover with AVHRR imagery.” *International Journal of Remote Sensing* 18, no. 4 (November): 917-935.

Blackburn, George A. 2007. “Hyperspectral remote sensing of plant pigment.” *Journal of Experimental Botany* 58, no. 4 (March): 855-867. doi: 10.1093/jxb/erl123.

Campbell, James B. and Randolph H. Wynne. 2011. *Introduction to Remote Sensing.* 5th ed. New York: the Guildford Press.

Cochrane, M.A. 2000. “Using vegetation reflectance variability for species level classification of hyperspectral data”. *International Journal of Remote Sensing* 21, no. 10 (November): 2075-2087.

Conklin, James G. 1969. “Insect enemies of birch.” In Proceedings, birch symposium. 151-154. USDA Northeastern Forest Experiment Station, Upper Darby, PA.

Coops, N.C., J.L. Lovell, D.L.B. Jupp, and D.S. Culvenor. 2003. “Using airborne and ground-based ranging lidar to measure canopy structure in Australian forests.” *Canadian Journal of Remote Sensing* 29, no. 5 (June), 607-622.

Crapper, P. F. 1984. “An estimate of the number of boundary cells in a mapped landscape coded to grid cells.” *Photogrammetric Engineering and Remote Sensing* 50, no. 10: 1497-1503.

Danson, F. M. 1995. “Developments in the remote sensing of forest canopy structure”. *In Advances in Environmental Remote Sensing.* edited by F.M. Danson and S. E. Plummer (Chichester: John Wiley & Sons): 53–69.

Day, R. J., and E. M. Harvey. 1981. “Forest dynamics in boreal mixedwood.” Lecture, Thunder Bay, Ontario, September 16.

Elhadi, Adam, Mutanga Onisimo, and Denis Rugege. 2010. “Multispectral and hyperspectral remote sensing for identification and mapping of wetland vegetation: a review.” *Wetland Ecol Manage* 18, (December): 281-296.

Fassnacht, Fabian Ewald, Hooman Latifi, Kryzstof Sterenczak, Aneta Modzelewska, Michael Lefsky, Lars T. Waser, Christopher Straub, and Aniruddha Ghosh. 2016. “Review of studies on tree species classification from remotely sensed data.” *Remote Sensing of Environment* 186, (December): 64-87.

Foody, G.M., and D. P. Cox. 1994. “Sub-pixel land cover composition estimation using a linear mixture model and fuzzy membership functions.” *Remote Sensing* 15, no. 3 (March): 619-631.

Franklin, S.E., and M.A. Wulder. 2002. “Remote sensing methods in medium spatial resolution satellite data land cover classification of large areas.” *Progress in Physical Geography* 26, no. 2 (June): 173–205.

Franklin, S.E., B. Titus, and R. Gillespie. 1994. “Remote sensing of vegetation cover at forest regeneration site.” *Global Ecology and Biogeography Letters* 4, no. 2 (March): 40-46.

Gao, Bo-Cai, Alexander Goetz, Marcos J. Montes, and Curtiss O. Davis. 2009 “Atmospheric correction algorithms for hyperspectral remote sensing data of land and ocean.” *Remote Sensing in the Environment* 113, no. 1 (September): S17-S24.

Gartzia, Maite, Concepcion Alados, Fernado Perez-Cabello, and C. Guillermo. 2013 “Improving the Accuracy of Vegetation Classifications in Mountainous Areas: A Case Study in the Spanish Central Pyrenees.” *Mountain Research and Development* 33, no. 1 (December): 63-74.

Gaussman, H.W. 1977. “Reflectance of leaf components.” *Remote Sensing in the Environment* 6, no. 1: 1-9.

Gitelson, Anatoly and Mark N. Merzlyak. 1994. “Spectral Reflectance Changes Associated with Autumn Senescence of *Aesculus hippocastanum* L. and *Acer platanoides* L. Leaves. Spectral Features and Relation to Chlorophyll Estimation.” *Journal of Plant Physiology* 143, no. 3 (March): 286-292.

Giannico, Chiara 2004. “Remote sensing of vegetation in the Calabrian region.” *Acta Astronaut* 60, no. 2 (January): 119-131.

Goetz, A.F.H., and V. Srivastava. 1985. “Mineralogical Mapping in the Cuprite Mining District, Nevada,” Proc. of the Airborne Imaging Spectrometer Data Anal. Workshop; 22-31 conference paper Jet Propulsion Lab. Pasadena, CA

Harrison, A. R., S. J. Hurcom, M. Taberner. 1991. “Examining spectral mixture modelling using in situ spectroradiometric measurements.” *Spatial Data 2000* (Nottingham: Remote Sensing Society), 242-249.

He, C., Q. Zhang, Y. Li, X. Li, and P. Shi. 2005. "Zoning grassland protection area using remote sensing and cellular automata modeling—A case study in Xilingol steppe grassland in northern China." *Journal of Arid Environments* 63, no. 4 (December): 814-826.

Hoffer, R. M., and Chris J. Johannsen. 1969. "Ecological potentials in spectral signature analysis." In P. L. Johnson ed., *Remote sensing in ecology*, University of Georgia, Athens, Georgia.

Ichoku, Charles, and Arnon Karnieli. 2009. "A review of mixture modeling techniques for sub-pixel land cover estimation." *Remote Sensing Reviews* 13, no. 3-4 (October): 161-186.

Irwin, Larry L. 1985. "Foods of moose, Alces, and white-tailed deer, Odocoileus virginianus, on a burn in boreal forest." *Canadian Field-Naturalist* 99, (February): 240-245.

Jensen, J.R., 1996. *Introduction to Digital Image Processing: A remote sensing perspective*. 2nd ed. Piscataway, NJ: Prentice Hall.

Jordan, James S. and Francis M. Rushmore. 1969. "Report on animal damage to birch." In Proceedings, birch symposium. 155-163. USDA Northeastern Forest Experiment Station, Upper Darby, PA.

Key, Thomas, Timothy A. Warner, James B. McGraw, and Mary Ann Fajvan. 2001. "A Comparison of Multispectral and Multitemporal Information in High Spatial Resolution Imagery for Classification of Individual Tree Species in a Temperate Hardwood Forest". Remote Sensing of Environment 75, no. 1 (January): 100-112.

Lauver, Chris. 1997. "Mapping species diversity patterns in Kansa shortgrass region by integrating remote sensing and vegetation analysis." *Journal of Vegetation Sciences* 8, no. 3 (February): 387-394.

Leckie, Donald G., Sally Tinis, Trisalyn Nelsen, Charles Burnett, Francois A. Gougeon, Ed Cloney, and Dennis Paradine. 2005. "Issues in species classification of trees in old growth conifer stands." *Canaidian Journal of Remote Sensing* 31, no. 2 (June): 175-190.

Lee, K.H. and R.S. Lunetta. 1996. Wetland Detection Methods. In: Lyon J.G. and McCarthy J. (eds). *Wetland and Environmental Applications of GIS*. Lewis Publishers, New York, 249–284.

Lu, D., and Q. Weng, 2007. "Survey of image classification methods and techniques for improving classification performance." *International Journal of Remote Sensing* 28, no. 5 (March): 823–870.

Malgorzata V.W., and N. Egyetem, 2010. "Data acquisition and integration, Remote Sensing." University of West Hungary Faculty of Geoinformatics.

Marshall, Arbust. Amer. 19. 1785 “*Betula papyrifera*,” Flora of North America 3, Accessed June 23, 2017 http://www.efloras.org/florataxon.aspx?flora_id=1&taxon_id= 233500260.

Mehner, Henny, Mark Cutlert, David Fairbairn, and Gillian Thompson. 2004. “Remote sensing of upland vegetation: the potential of high spatial resolution satellite sensors.” *Global Ecology and Biogeography* 13, no. 4 (June): 359-369.

North Dakota Tree Handbook. North Dakota State University, Centennial Tree Program project.

MN DNR. 2006. *North Shore Highlands; Subsection Profile, an Action plan for Minnesota Wildlife.* Accessed June 16, 2017. http://files.dnr.state.mn.us/assistance/nrplanning/bigpicture/cwcs/profiles/north_shore_highlands.pdf.

North Shore Forest Collaborative. “Why are the Birch Dying?” Accessed June 16, 2017. http://northshoreforest.org/what-happened-to-the-forest/.

Ollinger, S.V. 2011. “Source of variability in canopy reflectance and the convergent properties of plants.” *New Phytologist* 189, no. 2 (November): 375-394.

Parker, D.C., and M.F. Wolff. 1965. “Remote Sensing*” International Science and Technology* 43, (July): 20-31.

Reed, Bradley C., Jesslyn F. Brown, Darrel VanderZee, Thomas R. Loveland, James W. Merchant, and Donald O. Ohlen. 1994. “Measuring phenological variability from satellite imagery” *Journal of Vegetation Science* 5, no. 5 (October): 703-714.

Ruiliang, Pu, 2009. “Broadleaf species recognition with in situ hyperspectral data.” *International Journal of Remote Sensing* 30, no. 11 (June): 2759-2779. Accessed September 18, 2017. DOI: 10.1080/01431160802555820.

Safford, L. O., John C. Bjorkbom, and John C. Zasada. 1990. Betula papyrifera Marsh. paper birch In: Burns. Silvics of North America. Hardwoods. Agric. Handbook. U.S. Department of Agriculture, Forest Service. Washington, DC. 654 (2): 158-171.

Sayn-Wittgenstein, L. 1978. Recognition of Tree Species on Aerial Photographs. Canadian Forest Service Forest Management Institute Information Report, FMR-X-118, Environment Canada. 97.

Uchytil, Ronald J. 1991. “*Betula papyrifera.*” In: Fire Effects Information System. USDA, Rocky Mountain Research Station, Fire Sciences Laboratory.

Ustin, Susan, and John Gamon. 2010. “Remote sensing of plant functional groups.” *New Phytologist* 186, no. 4 (May): 795-816.

Woodcock, Crutis E., and Alan H. Straher. 1987. "The factor of scale in remote sensing*."* *Remote Sensing of the Environment* 21, no. 3 (April): 311-332.

Zarco-Tejada, P.J., A. Berjon, and J.R. Miller. 2000. "Stress Detection in Crops with Hyperspectral Remote Sensing and Physical Simulation Models." Institute for Sustainable Agriculture Universidad de Valladolid, Spain Dept. Earth and Space Science & Engineering, York University.

Appendix A: Dataset Metadata

AVIRIS Metadata:

Abstract: The AVIRIS sensor collects data that can be used for characterizing Earth's surface and atmosphere from geometrically coherent spectro-radiometric measurements. The AVIRIS instrument has 224 spectral bands with wavelengths from 400 to 2500 nanometers (visible, near infrared and mid-infrared portions of the spectrum). The AVIRIS instrument is flown on a NASA ER-2 aircraft. Data archived at the JPL AVIRIS Data Facility are available in both reflectance and radiance units. Scale and resolution depend on flight statistics.

Table 10 AVIRIS flight data

Flight Name	F120930t01p00r11	F061002t01p00r11
Date/ Time	9/30/2012 UTC 18:33	10/2/2006 UTC 18:45
Site Name	Aspen 4, MN	Arrowhead 1, MN
Investigator	Phil Townsend	Phil Townsend
Comments	Alt = 17.5 kft, SOG = 100 kts, Clouds = Clear, Direction = 91.6	100% clear
# Samples	1022	758
# Lines	1410	4305
Pixel Size	4.3 m	16.5 m
Solar Elevation	38.84	37.59
Solar Azimuth	192.05	195.94
Rotation	-82	69
File Size	0.71 GB	1.67 GB

FSA NAIP Imagery:

National Agriculture Imagery Program (NAIP) natural color .6-meter pixel resolution. The imagery was collected statewide. This data set contains imagery from the National Agriculture Imagery Program (NAIP). The NAIP acquires digital ortho imagery during the agricultural growing seasons in the continental U.S. A primary goal of the NAIP program is to enable availability of ortho imagery within one year of acquisition. The NAIP provides two main products: 1-meter ground sample distance (GSD) ortho imagery rectified to a horizontal accuracy within +/- 5 meters of reference digital ortho quarter quads (DOQQ's) from the National Digital Ortho Program (NDOP) or from the National Agriculture Imagery Program (NAIP); 1 meter or 60cm GSD ortho imagery rectified within +/- 6 meters to true ground. The tiling format of NAIP imagery is based on a 3.75' x 3.75' quarter quadrangle with a 300-pixel buffer on all four sides. The NAIP imagery is formatted to the UTM coordinate system using the North American Datum of 1983 (NAD83). The NAIP imagery may contain as much as 10% cloud cover per tile. Constraints: Not to be used for navigation, for informational purposes only.

Ground Truth Dataset:

This data was collected by Kyle Uhler. Data collected is an inventory of paper birch locations near roads within the AVIRIS image F120930t01p00r11. Areas accessed are shown in figure 7. Only paper birch with DBH greater or equal to 20 cm were sampled. Horizontal accuracies above pixel resolution of 4.3 m were not kept for projects analysis. This data was collected for the methodologies outlined in this thesis document. Data is up-to-date as 09/24/2017. All data was collected in UTM Zone 15 N coordinate system and North American Datum 1983 (meters).

MN Forest Stand Inventory Metadata:

This dataset is operated by Minnesota Department of Natural Resources (MNDNR) - Division of Forestry. This layer is a digital inventory of individual forest stands. The data are collected by MNDNR Foresters in each MNDNR Forestry Administrative Area, and is updated on a continuous basis, as needed. There are 50 classification categories in this layer. Most stands are field checked, and their characteristics described. The MN DNR uses internal MNDNR classification schema. This data originates from the MNDNR's "Forest Inventory Management" system (also referred to as FIM). The data are collected for Forest Resource Planning, harvest plans, treatment plans, wildlife habitat assessment, biotic community mapping support, historical vegetation studies. The data are up-to-date as of 07/14/2017. Content date indicates the date which the user can be confident of accuracy and completeness of the dataset. All polygons are in UTM Zone 15 N NAD83 (meters).