IMPROVED BOREAL VEGETATION MAPPING USING IMAGING SPECTROSCOPY TO AID WILDFIRE MANAGEMENT, INTERIOR ALASKA

By

Anushree Badola, M.Sc.

A Dissertation Submitted in Partial Fulfillment of the Requirements

for the Degree of

Doctor of Philosophy

in

Geosciences

University of Alaska Fairbanks

August 2023

APPROVED:

Santosh K. Panda, Committee Chair
Uma S. Bhatt, Committee Co-Chair
Dar A. Roberts, Committee Member
Christine F. Waigl, Committee Member
Randi R. Jandt, Committee Member
Bernard J. Coakley, Chair
Department of Geosciences
Karsten Hueffer, Dean
College of Natural Science & Mathematics
Richard L. Collins, Director
Graduate School

Abstract

Wildfires are a natural and essential part of Alaska ecosystems, but excessive wildfires pose a risk to the ecosystem's health and diversity, as well as to human life and property. To manage wildfires effectively, vegetation/fuel maps play a critical role in identifying high-risk areas and allocating resources for prevention, suppression, and recovery efforts. Furthermore, vegetation/fuel maps are an important input for fire behavior models, along with weather and topography data. By predicting fire behavior, such as spread rate, intensity, and direction, fuel models allow fire managers to make informed decisions about wildfire suppression, management, and prevention. Traditionally used vegetation/fuel maps in Alaska are inadequate due to a lack of detailed information since they are primarily generated using coarser resolution (30m) multispectral data. Hyperspectral remote sensing offers an efficient approach for better characterization of forest vegetation due to the narrow bandwidth and finer spatial resolution. However, the high cost associated with data acquisition remains a significant challenge to the widespread application of hyperspectral data. The aim of this research is to create accurate and detailed vegetation maps and upscale them for the boreal region of Alaska. The study involves hyperspectral data simulation using Airborne Visible InfraRed Imaging Spectrometer - Next Generation (AVIRIS-NG) data and publicly available Sentinel-2 multispectral data, ground spectra convolved to Sentinel-2 and AVIRIS-NG using the spectral response function of each sensor. Simulated data captured the minute details found in the real AVIRIS-NG data and were classified to map vegetation. Using the ground data from Bonanza Creek Long-Term Ecological Research sites, we compared the new maps with the two existing map products (the LANDFIRE's Existing Vegetation Type (EVT) and Alaska Vegetation and Wetland Composite). The maps generated using simulated data showed an improvement of 33% in accuracy and are more detailed than existing map products. In addition to fuel maps, we performed sub-pixel level mapping to generate a needleleaf fraction map, which serves fire management needs since needleleaf species are highly flammable. However, validating the sub-pixel product was challenging. To overcome this, we devised a novel validation method incorporating high-resolution airborne hyperspectral data (1m) and ground data. The study addresses the limitations of traditional fuel/vegetation maps by providing a more detailed and accurate representation of vegetation/fuel in Alaska. The methods and findings advance fuel and vegetation mapping research in Alaska and offer a novel pathway to generate detailed fuel maps for boreal Alaska to aid wildfire management.

Table of Contents

List of Figures

List of Tables

Acknowledgements

I want to express my deepest gratitude and appreciation to the following individuals and organizations who have played a significant role in the completion of this thesis:

First and foremost, I would like to thank my supervisor, Santosh K. Panda, for his unwavering support, invaluable guidance, and insightful feedback. His expertise and dedication have been instrumental in shaping the direction of this research. I am truly grateful for his patience, encouragement, and the countless hours he has devoted to reviewing and refining my work.

I want to sincerely thank the members of my thesis committee, Uma S. Bhatt, Dar A. Roberts, Christine F. Waigl, and Randi R. Jandt, for their valuable insights and constructive feedback. Their expertise in their respective fields has greatly enriched this study, and their feedback has played a crucial role in shaping the research. My committee has been the pillar of support for this research.

I want to acknowledge the financial support provided by the Alaska Established Program to Stimulate Competitive Research (EPSCoR), AmericaView, and the College of Natural Science and Mathematics. This work is supported by the National Science Foundation under the award OIA-1757348 and by the State of Alaska and the U.S. Geological Survey under Grant/Cooperative Agreement No. G18AP00077. Their funding has been crucial in enabling me to carry out this research and pursue my academic goals.

I would like to extend my heartfelt gratitude to the dedicated members of the field crew who played a crucial role in the successful execution of this research. Their tireless efforts, expertise, and unwavering commitment were instrumental in collecting valuable data and overcoming challenges in the field.

I am grateful to the faculty members of the University of Alaska Fairbanks (UAF), whose lectures, seminars, and intellectual discussions have broadened my understanding and equipped me with the necessary tools to undertake this research endeavor. I want to thank the Alaska EPSCoR team for their continuous support throughout this journey and for providing me with a friendly platform to share my research. I am also thankful to the UAF Geophysical Institute (GI) writing group, as the regular writing sessions greatly helped to achieve my writing goals. I would like to express my gratitude to the UAF writing center for their help in reviewing my papers and thesis. Furthermore, I would like to thank the UAF International Student and Scholar Services (ISSS), especially Reija Shnoro, for always being nice and supportive during my Ph.D. duration.

I am grateful to the GI HR department for efficiently handling all the paperwork and to the Graduate School for their support.

I would like to express my gratitude to my colleagues, with a special mention of Chris Smith, who has fostered a stimulating academic environment and provided valuable perspectives. Their friendship and camaraderie have made this journey more enjoyable and rewarding. I extend my sincere thanks to my friends Utsav Soni, Abhisek Maiti, Debvrat Varshney, and Sayantan Majumdar for their unwavering support and invaluable brainstorming sessions, which have greatly enriched the research. Their contributions have been invaluable, and I am deeply grateful for their continuous support.

My heartfelt thanks go to my family for their unconditional love and support, and encouragement throughout my studies. Furthermore, I would like to express my heartfelt appreciation to Anupma Prakash, Snigdhamayee Mishra, and Rudi Gens for creating a supportive and welcoming environment during my time here. I consider myself fortunate to have had the privilege of their unwavering support and their continuous motivation and guidance, that have played a vital role in my personal and academic growth. Additionally, I want to extend my sincere thanks to my friends Brooke Kubby, Cody Dogah, Fahimeh Dehghani, Saurabh Bahuguna, Anjali Madhu, and Arpita Shukla for always being there for me as a pillar of moral support.

Finally, I would like to dedicate this work to Utsav Soni, who has been a perpetual source of inspiration and unwavering support. His profound impact on my academic journey and personal growth cannot be overstated. While Utsav is no longer with us, his memory continues to guide and motivate me.

Completing the thesis would not have been possible without the support and contributions of the individuals mentioned above. Their guidance, encouragement, and assistance have been invaluable, and I am deeply grateful for their involvement in this endeavor.

Chapter 1 Introduction

The Earth's average temperature has increased by approximately 1.1°C since the pre-industrial era and is projected to surpass the 1.5 ºC threshold in the next 20 years (IPCC, 2022). The warming of the Earth is due to the release of greenhouse gasses, like carbon dioxide, into the atmosphere from human activities, such as burning fossil fuels, deforestation, and industrial processes (IPCC, 2022). This warming has numerous impacts. One of the most widespread impacts is melting of land ice and the subsequent sea level rise which threatens to flood coastal cities and communities globally. The warming climate disrupts ecosystems leading to species extinction and food chain disruption, which can have far-reaching consequences for biodiversity and food security. The warming temperatures have been shown in attribution studies to increase the intensity of heat waves and severe weather events, such as hurricanes, droughts, and wildfires, that can damage homes, infrastructure, and economies.

The impact of rising temperatures is becoming increasingly visible and catastrophic, particularly in the form of wildfires occurring more frequently and with greater intensity across the globe (MacCarthy et al., 2022). Globally, 27% of the forest cover degradation is due to wildfire (Global Forest Watch, 2023). From 2001 to 2021, wildfires caused a total 119 million hectares reduction in tree cover worldwide (Global Forest Watch, 2023). The last two decades (2003-2022), have witnessed several major wildfire events worldwide that have burned vast land areas. These wildfire events have been particularly prevalent in the polar region, where the rate of warming has been amplified in comparison to other areas of the world, a phenomenon known as polar amplification (Manabe and Stouffer, 1980). Studies have shown that the Arctic has been warming at a rate approximately twice as fast as the rest of the world (Walsh, 2014). As a result of increasing warming, there has been an increase in the prevalence of wildfires in Arctic regions in recent decades (Bhatt et al., 2021; McCarty et al., 2021; Zheng et al., 2023). One of the most unforgettable Arctic wildfire events in human history occurred in 2003, when catastrophic wildfires ravaged the taiga forests of Eastern Siberia, burning more than 22 million hectares of forest. This event was likely exacerbated by excessively dry conditions and increased human activities like forest clearing and burning (Talucci et al., 2022). With the Arctic warming at an accelerated rate, the risk of wildfires in this region is expected to increase further, posing a significant threat to the ecosystem.

Alaska, the only state in the United States with an Arctic environment, is particularly vulnerable to the effects of polar amplification. Over the past few decades, Alaska has experienced a significant rise in temperatures, with an average increase of 0.3 °C per decade

between 1976 and 2016 (Bhatt et al., 2021; Partain et al., 2016). Projections suggest that this trend will continue, with an expected rise of 1 to 2 °C by 2050 (Chapin et al., 2014). As a consequence of the escalating temperatures, the frequency and intensity of lightning strikes in Alaska have also shown an upward trend (Bieniek et al., 2020). Recent research by Bieniek et al. (2020) revealed a 17 percent increase in lightning occurrences throughout the state over the past 30 years. Frequent lightning strikes and increasing temperatures stand as the predominant factors driving the occurrence and severity of wildfires in Alaska (Box et al., 2019; International Arctic Research Center, 2021).

In recent decades, there has been an increase in both the size and frequency of wildfires in Alaska (Figure 1.1). The state witnessed its largest wildfire season in 2004, which burned around 2.5 million hectares during July and August. The second largest fire year was in 2015, in which about 2 million hectares burned in mid-June and mid-July (primarily due to lightning) (Partain et al., 2016). The recent year 2022 was the seventh largest fire season recorded in the state; a total of 589 fires burned an area of approximately 1.3 million hectares (Alaska Interagency Coordination Center, 2022). According to a report published in Western Forester, the major cause of the 2022 fire was the intense drought followed by lightning strikes. Between 1940 and 2022, wildfires burned 34 million hectares of land in Alaska, causing extensive and devastating damage to the state's ecosystems and communities for over eight decades (Figure 1.1). Over the last two decades, wildfires have burned approximately 13.14 million hectares in Alaska (Figure 1.2), which is roughly thrice the area burned during the previous two decades (1983-2002: 4.16 million ha) (Alaska Interagency Coordination Center, 2022; Badola et al., 2022). This trend indicates a growing threat of wildfires, as global warming accelerates and creates conditions favorable for large and destructive fires. Therefore, it is essential to implement effective fire management practices to mitigate wildfire risk.

Effective fire management is crucial for protecting people, property, and the environment from the devastating effects of wildfires. Fire managers need a comprehensive understanding of the local climate and weather patterns, including seasonal climate variations, the likelihood of lightning strikes, and the fuel types to anticipate and mitigate wildfire risks and develop effective fire prevention, suppression, and post-fire restoration strategies (National Park Service, 2022). However, in Alaska, the fuel maps still lack sufficient details, making it difficult for fire managers to manage wildfires effectively (Ziel, 2019). Fire managers need to have access to detailed and accurate fuel maps to develop effective management strategies. Therefore, it is crucial to conduct further research and investigation in Alaska, to generate more detailed fuel maps, and to aid fire management.

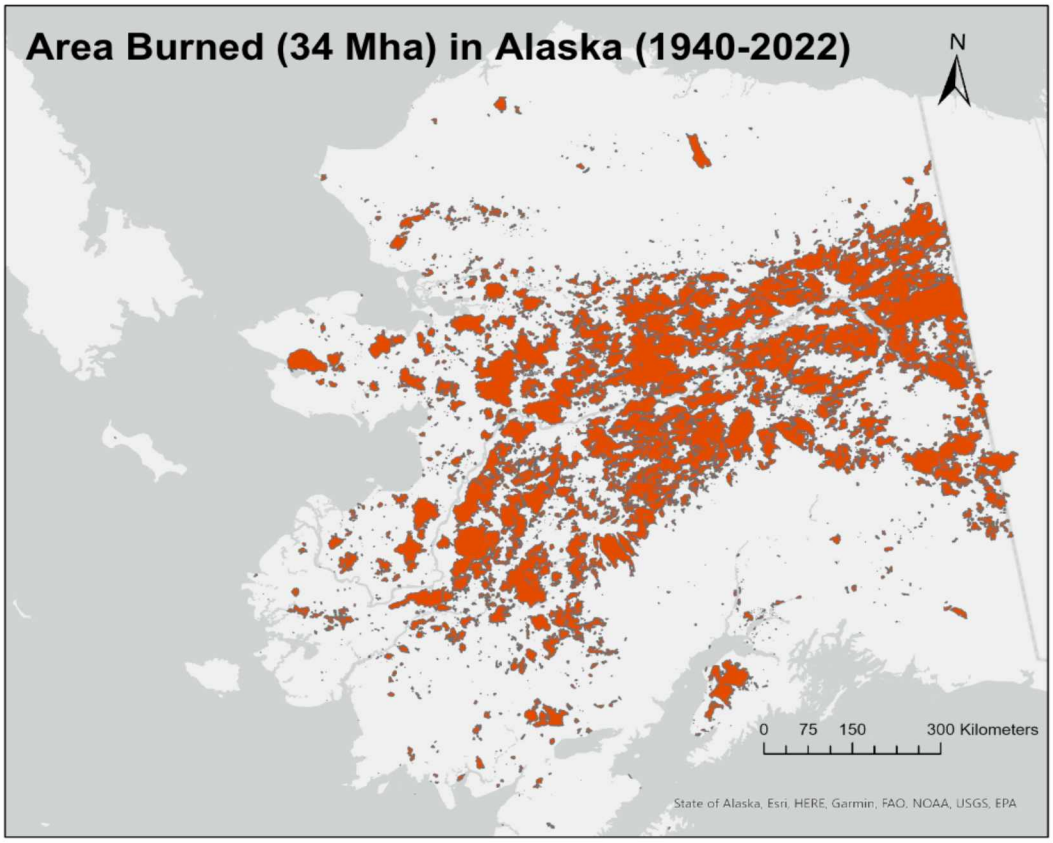


Figure 1.1: The total area burned in Alaska between 1940-2022. Most of the burned area is confined to the boreal region, interior Alaska. Data source: Fire Perimeters provided by Alaska Interagency Coordination Center using Sentinel-2 and Landsat Data.

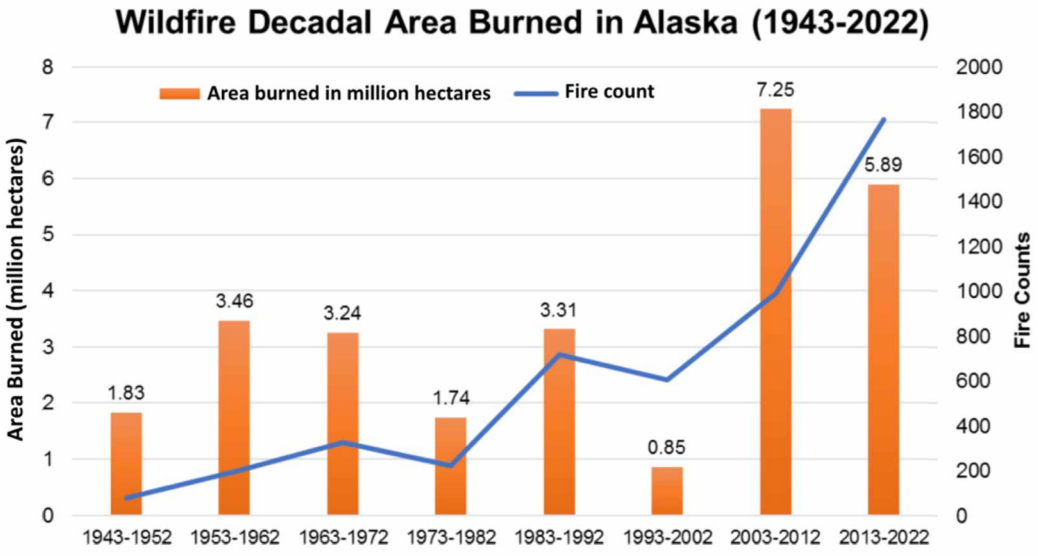


Figure 1.2: The decadal area burned in Alaska. In recent decades both the number of fires and area burned have increased. Data source: Fire Perimeters provided by Alaska Interagency Coordination Center.

Detailed fuel maps play a critical role in forest management practices to minimize the risk of both surface and crown fires. Standard fire management techniques to prevent surface fires (burns litter, lichens, moss, grasses, tree saplings, and shrubs) include prescribed (controlled) burning and removing downed dead trees and branches (slash) from the ground. Crown fires are extremely intense fires that burn tree branches and canopy and spread very quickly through the tree crowns. Increasing the distance between tree crowns (tree thinning) is a very common pre-fire management practice to check the crown fire spread, especially near developments. Pruning lower branches from tree crowns removes "ladder fuels" and thus helps prevent surface fire from reaching the crown (USDA, 2023; WWF, 2021). Vegetation and fuel maps are essential in developing effective fire management practices, as they provide critical information about the types and locations of vegetation and fuel sources in a given area (Government of Canada, 2021; Western Fire Chiefs Association, 2022). By using detailed fuel maps, fire and land managers can select the most appropriate fire management techniques and strategies, such as tree thinning, prescribed burning, creating fire breaks, and removing dead materials to protect communities and ecosystems from the devastating effects of wildfires.

To prepare for upcoming fire seasons, land managers rely on fire behavior models to predict fire intensity and spread based on input variables such as fuel characteristics, weather conditions, and topography. Forest growth and disturbance models such as the Forest Vegetation Simulator (Crookston and Dixon, 2005), and the Fire Dynamics Simulator (McGrattan et al., 2013) can be used to simulate future forest and disturbance conditions. Fuel maps are an important input variable for this type of planning as well as short-term fire behavior modeling during incidents. Unfortunately, the current fuel maps in Alaska lack detailed information, creating a significant gap in fire preparedness efforts. It is crucial to address this gap to mitigate the risk of wildfires in the region and improve our ability to manage future fire events effectively. Alaska has a vast expanse; hence, remote sensing technology is ideal for generating detailed fuel maps.

1.1. Remote sensing as a vegetation/fuel mapping tool

Remote sensing is gathering information about an object or feature without direct physical contact. It involves using various forms of electromagnetic radiation (EMR) to collect data about an object or surface, which is analyzed to obtain information. EMR includes radio waves, microwaves, visible light, X-rays, and gamma rays, and can interact with matter in different ways, such as being absorbed, reflected, or transmitted (Read & Torrado, 2009). Remote sensing works on the principle that different objects emit or reflect different amounts of energy in various electromagnetic wavelengths or bands, which is used to detect and discriminate between objects

or surface features (Aggarwal, 2004). Devices that detect this reflected or emitted EMR are called sensors, and the vehicle used to carry the sensor is called a platform. Remote sensing has many applications in various fields, including civil engineering, environmental monitoring, and disaster management. It allows repeated imaging of an object or surface at a finer spatial and temporal scale, and with different numbers of bands depending on the type of sensor. Sensors may be panchromatic, multispectral, or hyperspectral. A panchromatic sensor captures an image in a single broad spectral band. Multispectral sensors provide images in a few narrow bands in the visible and/or near-infrared (NIR) regions of the electromagnetic spectrum, and their bandwidth can range from 10-100 nm per band. Hyperspectral sensors capture images in hundreds of narrow and contiguous spectral bands, usually covering the visible to shortwave infrared (VSWIR) regions of the electromagnetic spectrum. The bandwidth of each band can be as narrow as 1-5 nm (Chuvieco & Kasischke 2007)

Multispectral data is freely available and has global coverage (Xie et al., 2008). NASA's Landsat program includes a series of Earth-observing satellites collecting multispectral images of the earth's surface since the 1970s. The Landsat 1 to 3 satellites collected images during 1972-1983 at 60 m spatial resolution. From Landsat 4 onwards, the data are available at 30 m spatial resolution making it the most popular image data for land cover mapping at regional and global scale (Wulder et al., 2012). The Arctic Boreal Vulnerability Experiment (ABoVE) land cover product is a notable example of a Landsat-derived dataset for the ABoVE core domain (Alaska and western Canada). They provide the annual dominant plant functional type maps at 30m spatial resolution, generated using Random Forests modeling with clustering and field photography (Wang et al., 2019). LANDFIRE (LANDscape Fire and Resource Management Planning Tools) is another program that produces maps and data products for the United States to support fire and resource management planning to state and federal fire management agencies for wildfire mitigation (LANDFIRE, 2016; Reeves et al., 2009; Rollins, 2009). The LANDFIRE Existing Vegetation Type (EVT) map product is traditionally used for fire management, fire spread modeling, and risk assessment in Alaska. This map product is derived from Landsat image data at a spatial resolution of 30 m. According to accuracy assessments conducted at three sites (Develice, 2012; Smith et al., 2021), the LANDFIRE EVT 2014 map product has an accuracy range of 20% to 45%, which is inadequate for studies requiring detailed analysis of vegetation patterns and characteristics. Sentinel is another Earth observation satellite developed by the European Space Agency (European Space Agency, 2023) as part of the Copernicus program, a European Union initiative to monitor the Earth's environment. Sentinel missions offer global data with high spatial resolution of 10m and a revisit time of 5-10 days. Descals et al. (2020) used

Sentinel 2A data to map the seasonal patterns in vegetation growth of the entire Arctic region, demonstrating the capability of Sentinel missions to study the Earth's surface at regional scale. The availability of multispectral data has revolutionized vegetation mapping and monitoring, making it possible to study the Earth's surface at a global and regional scale.

Imaging spectroscopy, also known as hyperspectral remote sensing, is a powerful technology that uses a high spectral resolution sensor to collect data across a wide range of wavelengths, providing detailed information about the physical and chemical properties of an object or surface. Imaging spectroscopy works by capturing the reflectance or emission of EMR across hundreds of narrow, contiguous spectral bands, which allows for the identification and characterization of materials based on their unique spectral signatures (Gupta, 1981; Xie et al., 2008). In recent years, there has been a substantial expansion in the use of hyperspectral data for detailed vegetation mapping due to its increasing availability and significant advancements in data processing and analysis techniques (Govender et al., 2019). In Alaska, NASA's Jet Propulsion Laboratory provides airborne hyperspectral data collected using the state-of-the-art Airborne Visible/Infrared Imaging Spectrometer-Next Generation (AVIRIS-NG) sensor as a part of Arctic-Boreal Vulnerability Experiment (ABoVE) airborne campaign (Miller et al., 2018). Goddard's LiDAR, Hyperspectral & Thermal Imager (G-LiHT), developed by NASA's Goddard Space Flight Center, is another airborne imaging system that offers a comprehensive range of data for terrestrial studies. It provides hyperspectral, LiDAR and thermal data, making it a powerful tool for researchers seeking to understand and analyze the Earth's surface and vegetation (Cook et al., 2013). One of the key advantages of imaging spectroscopy is its ability to differentiate between materials with similar visual appearances. For example, it can distinguish between different types of vegetation that may look similar in standard RGB imagery. This allows for a more detailed and accurate representation of the vegetation in an area and provides higher accuracy in discriminating between different tree species compared to traditional vegetation mapping methods using multispectral data (Badola et al., 2022; Fassnacht et al., 2016; Govender et al., 2019; Modzelewska et al., 2020; Smith et al., 2021). Accurate vegetation maps are crucial for proper fire management and planning. Studies have demonstrated the successful use of hyperspectral data for classifying forest tree species in tropical and subtropical forests (Badola et al., 2021; Clark & Roberts, 2012; Clark et al., 2005; Laurin et al., 2014), in the temperate forests (Heinzel & Koch, 2012) and mangrove forest studies (Hati et al., 2020).

Imaging spectroscopy has been used in boreal ecosystem studies (Cristóbal et al., 2021; Dalponte et al., 2014; Smith et al., 2021). Ustin & Xiao (2010) conducted a study at the Bonanza Creek Experimental Forest (BCEF), a long-term ecological research site located in interior Alaska,

that highlighted the potential of hyperspectral data for mapping forest ecosystems in boreal regions. The study utilized AVIRIS and a multispectral image data (the Satellite Pour l'Observation de la Terre, SPOT) to generate vegetation maps for BCEF. They reported higher accuracy of the maps generated using AVIRIS.

Zhang et al. (2020) combined LiDAR and hyperspectral data collected using the SpectIR hyperspectral sensor to map five major classes in the BCEF, including major vegetation classes: tussock tundra, moss spruce, and mixed forest. They performed Random Forest classification and assessed accuracy using an error matrix. They found that maps generated using combined data had 93% accuracy, while maps generated using hyperspectral data alone had 92.7% accuracy.

Smith et al. (2021) conducted a study on the BCEF using AVIRIS-NG hyperspectral data to generate highly detailed vegetation maps. The researchers mapped 20 dominant vegetation species with an accuracy of 80%, a significant improvement compared to the 8 dominant species mapped with an accuracy of 33% in the LANDFIRE EVT product. The use of AVIRIS-NG hyperspectral data allowed to capture fine-scale spectral information and discriminate between vegetation species with greater accuracy in BCEF. These studies demonstrate the potential of hyperspectral data for mapping boreal vegetation with high accuracy and detail.

Hyperspectral data is immensely valuable for vegetation mapping but the data are not widely available. For Alaska, the AVIRIS-NG and G-LiHT hyperspectral data were acquired only for limited sites and are not available for regional-scale vegetation and fuel mapping. A few studies have addressed the paucity of hyperspectral data by implementing a spectral reconstruction approach using multispectral data. Liu et al. (2009) and Tiwari et al. (2016) simulated Hyperion data from ALI multispectral data for Land-Use and Land-Cover (LULC) mapping using the Universal Pattern Decomposition Method (UPDM), a sensor-independent spectral unmixing technique (Zhang et al., 2007). This study aims to assess the performance of UPDM in simulating AVIRIS-NG data for the boreal region of Alaska. The simulated data is valuable because it provides the same spectral bands as the hyperspectral data collected by the sensor. The simulated data can be classified to create a detailed and accurate vegetation map at the pixel level, offering detailed insights into the distribution and composition of vegetation in the study area.

The maps generated using pixel-based algorithms like Random Forest, Decision Trees, Maximum Likelihood etc., classify each pixel of the image into a single class. But in reality, a pixel can cover more than one class on the ground hence we can model a pixel based on the

proportions of the different classes on ground, known as pixel/spectral unmixing (Adams et al., 1993). Spectral unmixing has been used in a few forest and wildfire studies to map green vegetation, soil, and non-photosynthetic vegetation. Roberts et al. (1998) utilized Multiple Endmember Spectral Mixture Analysis (MESMA), a linear spectral unmixing algorithm, to map vegetation in the Santa Monica mountains using AVIRIS data. Another study done by Lewis et al. (2011) used a similar approach to estimate forest floor consumption due to wildfire in the interior of Alaska. Their method involved comparing pre- and post-fire hyperspectral images and was effective in estimating forest floor consumption.

Interior Alaska forests are dominated by highly flammable needleleaf species, and mapping them is crucial for making informed decisions about wildfire and forest management. However, no work has been done in this area, creating an excellent opportunity to test spectral unmixing for needleleaf fraction cover mapping. By accurately identifying the needleleaf fraction cover, forest managers can better understand the vulnerability of the forest to wildfires and plan accordingly. Therefore, exploring the potential of spectral unmixing for needleleaf fraction mapping can have significant implications for effective forest management in interior Alaska.

1.2. Goal and objectives

The overarching goal of this dissertation is to develop a novel method to simulate hyperspectral data by utilizing Sentinel-2 multispectral data, which will enable us to generate accurate vegetation maps for the entire boreal region of Alaska. Additionally, we aim to develop a processing workflow to generate needleleaf fraction maps from hyperspectral data to aid wildfire management. The approach involves advanced machine learning techniques and state-of-the-art algorithms to map vegetation types, and needleleaf fractions with improved accuracy and detail to effectively monitor and manage boreal ecosystems in Alaska and mitigate the impacts of wildfire.

Research Questions:

Q1: Can high quality hyperspectral data be generated from widely available Sentinel-2 multispectral data to meet the need for greater spatial and temporal coverage of hyperspectral data for vegetation and fuel mapping at local scale in boreal Alaska?

Objectives:

1. Simulation of Hyperspectral image cube from Sentinel-2 data using AVIRIS-NG Spectral Response Function (SRF).
2. Validation of simulated data by means of visual interpretation, statistical analysis and by comparing spectral signature of tree/shrub species in the region.

3. Mapping of major tree/shrub classes by performing Random Forest Classification on simulated hyperspectral data and classification accuracy assessment.

Q2: How effectively can the simulated Hyperspectral data be used for identification of vegetation/fuel classes of the boreal region of Alaska at a regional scale?

Objectives:

1. Implement the Iterative Endmember Selection (IES) algorithm to derive the most representative endmember ground spectra for the boreal region of Alaska.
2. Improve upon the simulation algorithm to generate AVIRIS-NG hyperspectral image cube at Sentinel-2 scene scale (100 km X 100 km).
3. Derive detailed vegetation maps from the simulated hyperspectral data using a machine learning classifier and assess model accuracy and portability across space.
4. Implement the hyperspectral image simulation and vegetation classification algorithms in the Google cloud platform for efficient processing and ease of sharing with the research community.
5. Make the simulation algorithm, simulated hyperspectral data, and vegetation classification products available to land managers and research community for wider usage.

Q3: Can spectral unmixing of a pixel estimate the needleleaf fraction in a mixed boreal forest with reasonable accuracy and how do we validate needleleaf fraction estimates?

Objectives:

1. Implement spectral unmixing to estimate the needleleaf fraction in mixed boreal vegetation.
2. Assess the influence of spatial resolution of the hyperspectral data on the estimation of needleleaf fraction.
3. Validation of needleleaf fraction estimates at different spatial scales.

1.3. Study area

Alaska is situated between 66° 9' 37.8252" N and 153° 22' 8.9076" W, bordered by Canada to the east and the Arctic and Pacific Oceans to the north and south, respectively. It is the largest U.S. state encompassing 1.72 million square kilometers (Miller and Lynch, 2023), which comprises a vast and diverse landscape including boreal and temperate forests, tundra, coastal wetlands, mountains, and glaciers. This offers researchers unique opportunities to study a wide range of ecological processes and systems.

Boreal forests provide habitat to a wide range of wildlife, including iconic species such as grizzly bears, moose, caribou, and salmon (Juday, 2023; Kayes & Mallik, 2020). The state is also home to numerous indigenous communities that have relied on the land and its resources for millennia, providing a rich cultural context for research and study. Alaska is sparsely populated with approximately 730,000 residents (U.S. Census Bureau, 2022). Most of the population is concentrated in Anchorage, Fairbanks, and the state capital of Juneau (Brandt, 2009). The majority of Alaska's land is managed by the federal government, including the National Park Service, the Bureau of Land Management, and the US Forest Service, providing unique research and conservation opportunities.

Approximately 60-70% of the land area in Alaska is covered by the boreal forest (interior Alaska). As illustrated in Figure 1.1, most of the wildfires occur in this region. In this study, the majority of the analysis and ground data collection was conducted at the two major research sites located within interior Alaska: Bonanza Creek Experimental Forest (BCEF) and Caribou Poker Creek Research Watershed (CPCRW) (Figure 1.3). The National Science Foundation (NSF) established Bonanza Creek Long Term Ecological Research (LTER) in 1987 to conduct long-term ecological studies in interior Alaska. CPCRW is a site within the Bonanza Creek LTER (65.15 °N and 147.50 °W) (Bonanza Creek LTER, 2023). The following paragraphs provide detailed information about climate, geology, and vegetation of interior Alaska.

1.3.1. Climate

Interior Alaska lies between the Brooks Range to the north and Alaska Range to the south. The Alaska Range acts as a natural barrier blocking coastal air masses from reaching Interior Alaska (Bonanza Creek LTER, 2023). This gives rise to a strongly continental climate characterized by cold winters, and warm and dry summers in Bonanza creek LTER. The local climate in the study area is subarctic and is characterized by extreme weather conditions. Maximum summer air temperature can reach 33 ºC and minimum winter air temperature recorded is -50 ºC. The average annual temperature of the study area is -3 ºC and average annual precipitation is approximately 262 mm. The study area is covered in snow from October to late April; thus, the growing season lasts for less than 100 days (Juday, 2023; NEON, 2020; USDA, 2022).

1.3.2. Geology

Interior Alaska has a diverse and complex geologic history shaped by various geological processes: glaciation, volcanic activity, and tectonic activity. The parent material comprises Precambrian schist (bedrock), thick loess deposits derived from glaciation, and alluvial deposits

in flood plains. The soil is acidic due to the presence of conifer litter and is rich in organic matter (histosols) (Bonanza Creek LTER, 2023; Juday, 2023; USDA, 2022). The region's geology plays an essential role in shaping its ecology. The lithology and age of rocks, soils, and topography influence the distribution and abundance of plant and animal species. Soils are relatively immature, ranging from cold, shallow permafrost soils that are poorly drained in the lowlands to well-drained, warmer soils in the uplands. Permafrost formation is influenced by slope and aspect, with north-facing slopes typically underlain by permafrost, while south-facing slopes have warmer, well-drained soils. The region's geology is also crucial in determining the availability of mineral resources, including gold, copper, and other metals, which have played a vital role in the region's economic development (Gough et al., 2008). The geology of interior Alaska is a complex and dynamic system that continues to shape its ecology, economy, and cultural heritage.

1.3.3. Vegetation

The Boreal Forest is primarily composed of deciduous broad-leaved trees, evergreen needle-leaved trees, and deciduous needle-leaved trees. Among the needleleaf evergreen species, white spruce (*Picea glauca*) and black spruce (*Picea mariana*) are the most common. Larch (*Larix laricina*) is the only needleleaf deciduous tree species, and aspen (*Populus tremuloides*), birch (*Betula neoalaskana, B. papyrifera*), and poplar (*Populus balsamifera*) are major broad-leaved deciduous trees (Brandt, 2009). The vegetation in this region is diverse and varies depending on the geologic conditions and disturbance history. Uplands with well-drained soils support white spruce and deciduous species such as aspen and birch. In contrast, poorly drained flats that are mostly underlain by permafrost develop peaty wetlands and support the growth of black spruce in interior Alaska (Kayes & Mallik, 2020; USDA, 2022). The region also boasts a variety of shrub species, including lingonberry (*Vaccinium vitis-idaea*), bog Labrador tea (*Rhododendron groenlandicum*), several willow species (*Salix sp.*), alder *(Alnus viridis*), dwarf birch (Betula nana), blueberry (Vaccinium ovalifolium), and wild rose (Rosa acicularis). A thick layer of moss and lichens dominates the forest floor (Juday, 2023; NEON, 2020). The Boreal Forest is an important ecosystem with diverse flora, including tree species, shrubs, and ground cover. Its unique geologic conditions give rise to distinct habitats that support different plant species. Understanding and protecting this valuable resource is crucial for the preservation of biodiversity and the health of our planet.

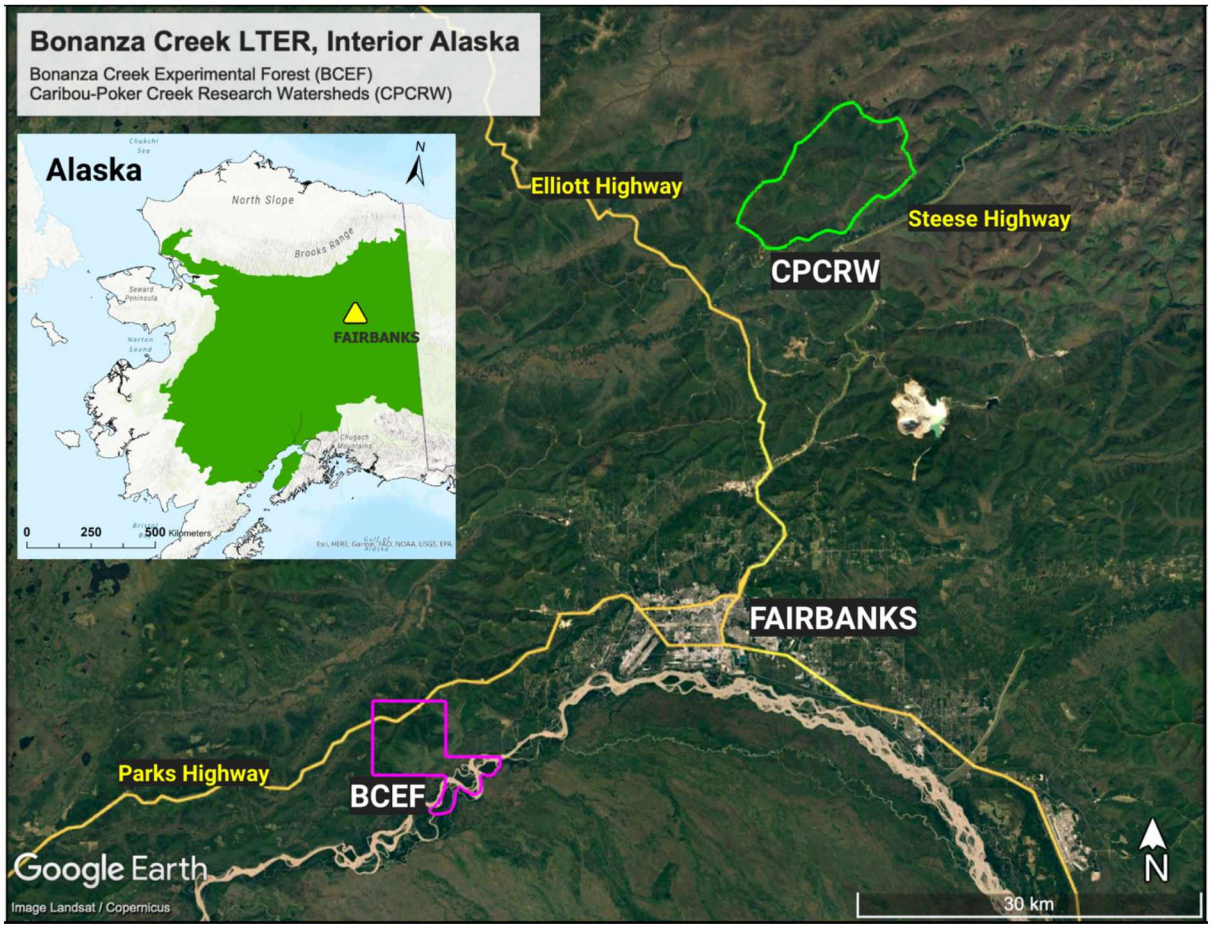


Figure 1.3: Study sites in Interior Alaska: Bonanza Creek Experimental Forest (marked in pink) and the Caribou-Poker Creek Research Watersheds (marked in green). The inset map displays the boreal region of Alaska (highlighted in green).

This thesis is structured around five chapters that provide a comprehensive overview of our research on hyperspectral data simulation and vegetation mapping in the Alaskan boreal region. Chapter one introduces the thesis and outlines the significance of this work, the study area, and research questions and objectives. Chapter two focuses on simulating AVIRIS-NG hyperspectral data from widely available Sentinel-2 multispectral data over a test site. We provide an in-depth analysis of different validation approaches and discuss their effectiveness in assessing the accuracy of the simulated data. In chapter three, we describe the approach to scaling up fuel maps and automating the simulation and vegetation classification process for generating vegetation maps at the regional scale. We compare the vegetation maps against ground data and vegetation products from other agencies. Chapter four presents an approach to map needleleaf fractions within each pixel, along with a novel validation method. Finally, in chapter five, we summarize the research findings and provide recommendations for future

research. Our study contributes to the ongoing efforts to understand and manage the impacts of wildfires in Alaska and provides valuable insights and tools for ecologists, land managers, and policymakers.

1.4. References

Adams, J. B., Smith, M. O., & Gillespie, A. R. (1993). Imaging Spectroscopy: Interpretation Based on Spectral Mixture Analysis. In C.M. Pieters and P.A.J. Englert (Eds.), *Remote Geochemical Analysis: Elemental and Mineralogical Composition (*pp. 145-166*)*. Cambridge University Press, New York.

Aggarwal, S. (2004). Principles of remote sensing. In M.V.K. Sivakumar, P.S. Roy, K. Harmsen and S.K. Saha (Eds.), *Satellite remote sensing and GIS applications in agricultural meteorology,* (pp. 23-28). World Meteorological Organisation, Geneva.

Alaska Interagency Coordination Center. (2022). *AICC - Incident Information - Alaska Large Fires*. https://fire.ak.blm.gov/incinfo/aklgfire.php

Badola, A., Padalia, H., Belgiu, M., & Verma, P. A. (2021, July 11-16). *Tree Species Mapping in Tropical Forests Using Hyperspectral Remote Sensing and Machine Learning* [Paper presentation]. 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, Brussels, Belgium. https://doi.org/10.1109/IGARSS47720.2021.9553549

Badola, A., Panda, S. K., Roberts, D. A., Waigl, C. F., Jandt, R. R., & Bhatt, U. S. (2022). A novel method to simulate AVIRIS-NG hyperspectral image from Sentinel-2 image for improved vegetation/wildfire fuel mapping, boreal Alaska. *International Journal of Applied Earth Observation and Geoinformation*, *112*, 102891. https://doi.org/10.1016/J.JAG.2022.102891

Bhatt, U. S., Lader, R. T., Walsh, J. E., Bieniek, P. A., Thoman, R., Berman, M., Borries-Strigle, C., Bulock, K., Chriest, J., Hahn, M., Hendricks, A. S., Jandt, R., Little, J., McEvoy, D., Moore, C., Rupp, T. S., Schmidt, J., Stevens, E., Strader, H., … Ziel, R. (2021). Emerging Anthropogenic Influences on the Southcentral Alaska Temperature and Precipitation Extremes and Related Fires in 2019. *Land 2021, 10(1),* 82. https://doi.org/10.3390/LAND10010082

Bieniek, P. A., Bhatt, U. S., York, A., Walsh, J. E., Lader, R., Strader, H., Ziel, R., Jandt, R. R., & Thoman, R. L. (2020). Lightning Variability in Dynamically Downscaled Simulations of Alaska's Present and Future Summer Climate. *Journal of Applied Meteorology and Climatology, 59(6),* 1139–1152. https://doi.org/10.1175/JAMC-D-19-0209.1

Bonanza Creek LTER. (2023). Bonanza Creek LTER, Institute of Arctic Biology, University of Alaska Fairbanks. https://www.lter.uaf.edu/boreal-forest/about

Box, J. E., Colgan, W. T., Christensen, T. R., Schmidt, N. M., Lund, M., Parmentier, F. J. W., Brown, R., Bhatt, U. S., Euskirchen, E. S., Romanovsky, V. E., Walsh, J. E., Overland, J. E., Wang, M., Corell, R. W., Meier, W. N., Wouters, B., Mernild, S., Mård, J., Pawlak, J., & Olsen, M. S. (2019). Key indicators of Arctic climate change: 1971-2017. *Environmental Research Letters, 14(*4*),* 045010. https://doi.org/10.1088/1748-9326/aafc1b

Brandt, J. P. (2009). The extent of the North American boreal zone. *Environmental Reviews*, *17*, 101–161. https://doi.org/10.1139/A09-004

Chapin, F. S., Trainor, S. F., Cochran, P., Huntington, H., Markon, C., McCammon, M., McGuire, A. D., Serreze, M. (2014). Ch. 22: Alaska. Climate Change Impacts in the United States: *The Third National Climate Assessment*. J. M. Melillo, Terese (T.C.) Richmond, and G. W. Yohe (Eds.), U.S. Global Change Research Program. https://doi.org/10.7930/J00Z7150

Chuvieco, E., & Kasischke, E. S. (2007). Remote sensing information for fire management and fire effects assessment. *Journal of Geophysical Research: Biogeosciences, 112(1)* G01S90. https://doi.org/10.1029/2006JG000230

Clark, M. L., & Roberts, D. A. (2012). Species-Level Differences in Hyperspectral Metrics among Tropical Rainforest Trees as Determined by a Tree-Based Classifier. *Remote Sensing*, *4*(6), 1820–1855. https://doi.org/10.3390/rs4061820

Clark, M. L., Roberts, D. A., & Clark, D. (2005). Hyperspectral discrimination of tropical rain forest tree species at leaf to crown scales. *Remote Sensing of Environment, 96(3–4)*, 375–398. https://doi.org/10.1016/j.rse.2005.03.009

Cook, B. D., Corp, L. A., Nelson, R. F., Middleton, E. M., Morton, D. C., McCorkel, J. T., Masek, J. G., Ranson, K. J., Ly, V., & Montesano, P. M. (2013). NASA Goddard's LiDAR, Hyperspectral and Thermal (G-LiHT) Airborne Imager. *Remote Sensing, 5(8),* 4045–4066. https://doi.org/10.3390/RS5084045

Crookston, N. L., & Dixon, G. E. (2005). The Forest Vegetation Simulator: A review of its structure, content, and applications. *Computers and Electronics in Agriculture. 49(1),* 60–80. https://doi.org/10.1016/j.compag.2005.02.003

Cristóbal, J., Graham, P., Prakash, A., Buchhorn, M., Gens, R., Guldager, N., & Bertram, M. (2021). Airborne Hyperspectral Data Acquisition and Processing in the Arctic: A Pilot Study Using the Hyspex Imaging Spectrometer for Wetland Mapping. *Remote Sensing, 13*(6), 1178. https://doi.org/10.3390/RS13061178

Dalponte, M., Ørka, H. O., Ene, L. T., Gobakken, T., & Næsset, E. (2014). Tree crown delineation and tree species classification in boreal forests using hyperspectral and ALS data. *Remote Sensing of Environment*, *140*, 306–317. https://doi.org/10.1016/J.RSE.2013.09.006

Descals, A., Verger, A., Yin, G., & Peñuelas, J. (2020). Improved estimates of arctic land surface phenology using sentinel-2 time series. *Remote Sensing, 12(22),* 1–13. https://doi.org/10.3390/RS12223738

DeVelice, R.L. (2012). *Accuracy of the LANDFIRE Alaska Existing Vegetation Map over the Chugach National Forest*. LANDFIRE Assessments, Volume 10. Retrieved Febuary 10, 202 https://landfire.cr.usgs.gov/documents/LANDFIRE_ak_110evt_accuracy_summary_013012.pdf

European Space Agency. (2023). *Copernicus Open Access Hub*. European Space Agency. Retrieved May 15, 2023 from https://scihub.copernicus.eu/dhus/#/home

Fassnacht, F. E., Latifi, H., Stereńczak, K., Modzelewska, A., Lefsky, M., Waser, L. T., Straub, C., & Ghosh, A. (2016). Review of studies on tree species classification from remotely sensed data. *Remote Sensing of Environment*, *186*, 64–87. https://doi.org/10.1016/J.RSE.2016.08.013

Filkov, A. I., Ngo, T., Matthews, S., Telfer, S., & Penman, T. D. (2020). Impact of Australia's catastrophic 2019/20 bushfire season on communities and environment. Retrospective analysis and current trends. *Journal of Safety Science and Resilience*, *1*(1), 44–56. https://doi.org/10.1016/J.JNLSSR.2020.06.009

Global Forest Watch. (2023). *Global Deforestation Rates & Statistics by Country | GFW*. https://www.globalforestwatch.org/dashboards/global/?category=fires&dashboardPrompts=eyJzaG93UHJvbXB0cyI6dHJ1ZSwicHJvbXB0c1ZpZXdlZCI6WyJkb3dubG9hZERhc2hib2FyZFN0YXRzIiwid2lkZ2V0U2V0dGluZ3MiXSwic2V0dGluZ3MiOnsic2hvd1Byb21wdHMiOnRydWUsInByb21wdHNWaWV3ZWQiOl

Gough, L. P., Crock, J. G., Wang, B., Day, W. C., Eberl, D. D., Sanzolone, R. F., & Lamothe, P. J. (2008). Substrate geochemistry and soil development in boreal forest and tundra ecosystems in the Yukon–Tanana Upland and Seward Peninsula, Alaska: U.S. Geological Survey Scientific Investigations Report 2008–5010, 18 p

Govender, M., Chetty, K., Naiken, V., & Bulcock, H. (2019). A comparison of satellite hyperspectral and multispectral remote sensing imagery for improved classification and mapping of vegetation. *Water SA*, *34*(2), https://doi.org/10.4314/wsa.v34i2.

Government of Canada. (2021, April 16). Fire behaviour. https://natural-resources.canada.ca/our-natural-resources/forests/wildland-fires-insects-disturbances/forest-fires/fire-behaviour/13145

Gupta, R.P. (2018). Imaging Spectroscopy. *Remote Sensing Geology*. Springer, Berlin, Heidelberg.

Hati, J. P., Goswami, S., Samanta, S., Pramanick, N., Majumdar, S. D., Chaube, N. R., Misra, A., & Hazra, S. (2020). Estimation of vegetation stress in the mangrove forest using AVIRIS-NG airborne hyperspectral data. *Modeling Earth Systems and Environment, 7,* 1–13. https://doi.org/10.1007/s40808-020-00916-5

Heinzel, J., & Koch, B. (2012). Investigating multiple data sources for tree species classification in temperate forest and use for single tree delineation. *International Journal of Applied Earth Observation and Geoinformation*, *18*(1), 101–110. https://doi.org/10.1016/J.JAG.2012.01.025

International Arctic Research Center. (2021). *Alaska's Changing Wildfire Environment | IARC*. https://uaf-iarc.org/alaskas-changing-wildfire-environment/

Intergovernmental Panel on Climate Change. (2022). Summary for Policymakers. In H. O. Pörtner, D. C. Roberts, M. Tignor, E. S. Poloczanska, K. Mintenbeck, A. Alegría, M. Craig, S. Langsdorf, S. Löschke, V. Möller, A. Okem, & B. Rama (Eds.), *Climate Change 2022: Impacts, Adaptation, and Vulnerability. Contribution of Working Group II to the Sixth Assessment Report of the Intergovernmental Panel on Climate Change (p. In Press).* Cambridge University Press.

Juday, G. P. (2023). taiga. In *Encyclopedia Britannica*. https://www.britannica.com/science/taiga

Kayes, I., Mallik, A. (2020). Boreal Forests: Distributions, Biodiversity, and Management. In: Leal Filho, W., Azul, A., Brandli, L., Lange Salvia, A., Wall, T. (eds) Life on Land. Encyclopedia of the UN Sustainable Development Goals. Springer, Cham. https://doi.org/10.1007/978-3-319-71065-5_17-1

LANDFIRE. (2016). LANDFIRE: Existing Vegetation Type. U.S. Department of Agriculture and U.S. Department of the Interior. Retrieved February 10, 2021, from http://www.landfire.gov

Laurin, G. V., Chan, J. C.-W., Chen, Q., Lindsell, J. A., Coomes, D. A., Guerriero, L., Frate, F. Del, Miglietta, F., & Valentini, R. (2014). Biodiversity Mapping in a Tropical West African Forest with Airborne Hyperspectral Data. *PLoS ONE, 9(*6*)*. https://doi.org/10.1371/journal.pone.0097910

Lewis, S. A., Hudak, A. T., Ottmar, R. D., Robichaud, P. R., Lentile, L. B., Hood, S. M., Cronan, J. B., Morgan, P., Lewis, S. A., Hudak, A. T., Ottmar, R. D., Robichaud, P. R., Lentile, L. B., Hood, S. M., Cronan, J. B., & Morgan, P. (2011). Using hyperspectral imagery to estimate forest floor consumption from wildfire in boreal forests of Alaska, USA. *International Journal of Wildland Fire, 20(*2*),* 255–271. https://doi.org/10.1071/WF09081

Liu, B., Zhang, L., Zhang, X., Zhang, B., & Tong, Q. (2009). Simulation of EO-1 Hyperion Data from ALI Multispectral Data Based on the Spectral Reconstruction Approach. *Sensors*, *9*(4), 3090–3108. https://doi.org/10.3390/s90403090

MacCarthy, J., Tyukavina, S., Weisse, M., & Harris, N. (2022, August 22). New Data Confirms: Forest Fires Are Getting Worse | World Resources Institute. https://www.wri.org/insights/global-trends-forest-fires

Manabe, S. and Stouffer, R.J. (1980). Sensitivity of a global climate model to an increase of CO2 concentration in the atmosphere. *Journal of Geophysical Research, 85,* 5529–5554. https://doi.org/10.1029/JC085iC10p05529

Mccarty, J. L., Aalto, J., Paunu, V. V., Arnold, S. R., Eckhardt, S., Klimont, Z., Fain, J. J., Evangeliou, N., Venäläinen, A., Tchebakova, N. M., Parfenova, E. I., Kupiainen, K., Soja, A. J., Huang, L., & Wilson, S. (2021). Reviews and syntheses: Arctic fire regimes and emissions in the 21st century. *Biogeosciences, 18*(18), 5053–5083. https://doi.org/10.5194/BG-18-5053-2021

McGrattan, K., McDermott, R., Weinschenk, C. and Forney, G. (2013), *Fire Dynamics Simulator, Technical Reference Guide, Sixth Edition, Special Publication (NIST SP)*. National Institute of Standards and Technology, Gaithersburg, MD.

Miller, C. E., Green, R. O., Thompson, D. R., Thorpe, A. K., Eastwood, M., Mccubbin, I. B., Olson-duvall, W., Bernas, M., Sarture*, C. M., Nolte, S., Rios, L. M., Hernandez, M. A., Bue, B. D., & Lundeen, S. R. (2018). ABoVE: Hyperspectral Imagery from AVIRIS-NG, Alaskan and Canadian Arctic, 2017-2019* (Version 1.2). ORNL Distributed Active Archive Center. https://doi.org/10.3334/ORNLDAAC/1569

Miller, M. M., & Lynch, D. (2023). Alaska. In *Encyclopedia Britannica*. https://www.britannica.com/place/Alaska

Modzelewska, A., Fassnacht, F. E., & Stereńczak, K. (2020). Tree species identification within an extensive forest area with diverse management regimes using airborne hyperspectral data. *International Journal of Applied Earth Observation and Geoinformation*, *84*. https://doi.org/10.1016/J.JAG.2019.101960

National Aeronautics and Space Administration Earth Observatory. (2005, March 30). The Rising Cost of Natural Hazards. https://earthobservatory.nasa.gov/features/RisingCost/rising_cost5.php

National Park Service. (2022). Wildland Fire Behavior. https://www.nps.gov/articles/wildland-fire-behavior.htm

National Ecological Observatory Network. (2020). Caribou-Poker Creeks Research Watershed NEON | NSF NEON | Open Data to Understand our Ecosystems. https://www.neonscience.org/field-sites/bona

Partain, J. L., Alden, S., Bhatt, U. S., Bieniek, P. A., Brettschneider, B. R., Lader, R. T., Olsson, P. Q., Rupp, T. S., Strader, H., Thoman, R. L., Walsh, J. E., York, A. D., & Ziel, R. H. (2016). An Assessment of the Role of Anthropogenic Climate Change in the Alaska Fire Season of 2015. *Bulletin of the American Meteorological Society, 97*(12), S14–S18. https://doi.org/10.1175/BAMS-D-16-0149.1

Read, J. M., & Torrado, M. (2009). Remote Sensing. R. Kitchin (Ed.), *International Encyclopedia of Human Geography* (pp. 335–346). Maxwell School of Syracuse University.

Reeves, M. C., Ryan, K. C., Rollins, M. G., & Thompson, T. G. (2009). Spatial fuel data products of the LANDFIRE Project. *International Journal of Wildland Fire, 18*, 250–267. https://doi.org/10.1071/WF08086

Roberts, D. A., Gardner, M., Church, R., Ustin, S., Scheer, G., & Green, R. O. (1998). Mapping Chaparral in the Santa Monica Mountains Using Multiple Endmember Spectral Mixture Models. *Remote Sensing of Environment*, *65*(3), 267–279. https://doi.org/10.1016/S0034-4257(98)00037-6

Rollins, M. (2009). LANDFIRE: A nationally consistent vegetation, wildland fire, and fuel assessment. *International Journal of Wildland Fire*, *18*, 235–249. https://doi.org/10.1071/WF08088

Smith, C. W., Panda, S. K., Bhatt, U. S., & Meyer, F. J. (2021). Improved Boreal Forest Wildfire Fuel Type Mapping in Interior Alaska using AVIRIS-NG Hyperspectral data. *Remote Sens.*, *13*(5), 897. https://doi.org/https://doi.org/10.3390/rs13050897

Society of American Foresters. (2022). *Western Forester*. https://forestry.org/wp-content/uploads/2022/11/WFOctNovDec2022-color.pdf

Talucci, A. C., Loranty, M. M., & Alexander, H. D. (2022). Siberian taiga and tundra fire regimes from 2001–2020. *Environmental Research Letters*, *17*(2), 025001. https://doi.org/10.1088/1748-9326/AC3F07

Tiwari, V., Kumar, V., Pandey, K., Ranade, R., & Agrawal, S. (2016, July 10-15). *Simulation of the hyperspectral data using Multispectral data* [Paper Presentation]. 2016 International Geoscience and Remote Sensing Symposium (IGARSS), Beijing, China 6157–6160. https://doi.org/10.1109/IGARSS.2016.7730608

U.S. Census Bureau. (2022, July 1). U.S. Census Bureau QuickFacts: Alaska. https://www.census.gov/quickfacts/AK

U.S. Department of Agriculture. (2022). Bonanza Creek Experimental Forest & Caribou-Poker Creeks Research Watershed | Pacific Northwest Research Station | PNW - US Forest Service. https://www.fs.usda.gov/pnw/experimental-forests-and-ranges/bonanza-creek-experimental-forest-caribou-poker-creeks-research

U.S. Department of Agriculture. (2023). Managing Fire | US Forest Service. Retrieved February 17, 2023, from https://www.fs.usda.gov/science-technology/fire

Ustin, S. L., & Xiao, Q. F. (2010). Mapping successional boreal forests in interior central Alaska. *International Journal of Remote Sensing, 22(9)*, 1779–1797. https://doi.org/10.1080/01431160118269

Walsh, J. E. (2014). Intensified warming of the Arctic: Causes and impacts on middle latitudes. *Global and Planetary Change*, *117*, 52–63. https://doi.org/10.1016/J.GLOPLACHA.2014.03.003

Western Fire Chiefs Association. (2022, July 5). *Understanding the Different Types of Wildfires*. https://wfca.com/articles/types-of-wildfire/

World Wildlife Fund. (2023). *Fire management*. https://www.worldwildlife.org/stories/fire-management

Wulder, M. A., Masek, J. G., Cohen, W. B., Loveland, T. R., & Woodcock, C. E. (2012). Opening the archive: How free data has enabled the science and monitoring promise of Landsat. *Remote Sensing of Environment*, *122*, 2–10. https://doi.org/https://doi.org/10.1016/j.rse.2012.01.010

Wulder, Michael A., Jeffrey G. Masek, Warren B. Cohen, Thomas R. Loveland, and Curtis E. Woodcock. (2012). Opening the Archive: How Free Data Has Enabled the Science and Monitoring Promise of Landsat. *Remote Sensing of Environment 122,* 2-10). doi: 10.1016/J.RSE.2012.01.010.

Xie, Y., Sha, Z., & Yu, M. (2008). Remote sensing imagery in vegetation mapping: a review. *Journal of Plant Ecology*, *1*(1), 9–23. https://doi.org/10.1093/jpe/rtm005

Zhang, C., Douglas, T. A., & Anderson, J. E. (2020, September 26- October 2). *Mapping Vegetation And Seasonal Thaw Depth In Central Alaska Using Airborne Hyperspectral And Lidar Data* [Paper Presentation]. IGARSS 2020 - 2020 IEEE International Geoscience and Remote Sensing Symposium.Waikoloa, HI, USA. https://doi.org/DOI:10.1109/IGARSS39084.2020.9323660

Zhang, L., Fujiwara, N., Furumi, S., Muramatsu, K., Daigo, M., & Zhang, L. (2007). Assessment of the universal pattern decomposition method using MODIS and ETM data. *International Journal of Remote Sensing*, *28*(1), 125–142. https://doi.org/10.1080/01431160600617228

Zheng, B., Ciais, P., Chevallier, F., Yang, H., Canadell, J. G., Chen, Y., van der Velde, I. R., Aben, I., Chuvieco, E., Davis, S. J., Deeter, M., Hong, C., Kong, Y., Li, H., Li, H., Lin, X., He, K., & Zhang, Q. (2023). Record-high CO2 emissions from boreal fires in 2021. *Science, 379(6635),* 912–917. DOI:10.1126/science.ade0805

Ziel, R. (2019). *Alaska's Fire Environment: Not an Average Place - International Association of Wildland Fire.* https://www.iawfonline.org/article/alaskas-fire-environment-not-an-average-place/

Chapter 2 Hyperspectral Data Simulation (Sentinel-2 To AVIRIS-NG) for Improved Wildfire Fuel Mapping, Boreal Alaska[1]

2.1. Abstract

Alaska has witnessed a significant increase in wildfire events in recent decades that have been linked to drier and warmer summers. Forest fuel maps play a vital role in wildfire management and risk assessment. Freely available multispectral datasets are widely used for land use and land cover mapping, but they have limited utility for fuel mapping due to their coarse spectral resolution. Hyperspectral datasets have a high spectral resolution, ideal for detailed fuel mapping, but they are limited and expensive to acquire. This study simulates hyperspectral data from Sentinel-2 multispectral data using the spectral response function of the Airborne Visible/Infrared Imaging Spectrometer-Next Generation (AVIRIS-NG) sensor, and normalized ground spectra of gravel, birch, and spruce. We used the Uniform Pattern Decomposition Method (UPDM) for spectral unmixing, which is a sensor-independent method, where each pixel is expressed as the linear sum of standard reference spectra. The simulated hyperspectral data have spectral characteristics of AVIRIS-NG and the reflectance properties of Sentinel-2 data. We validated the simulated spectra by visually and statistically comparing it with real AVIRIS-NG data. We observed a high correlation between the spectra of tree classes collected from AVIRIS-NG and simulated hyper-spectral data. Upon performing species level classification, we achieved a classification accuracy of 89% for the simulated hyperspectral data, which is better than the accuracy of Sentinel-2 data (77.8%). We generated a fuel map from the simulated hyperspectral image using the Random Forest classifier. Our study demonstrated that low-cost and high-quality hyperspectral data can be generated from Sentinel-2 data using UPDM for improved land cover and vegetation mapping in the boreal forest.

[1] Badola, A., Panda, S. K., Roberts, D. A., Waigl, C. F., Bhatt, U. S., Smith, C. W., & Jandt, R. R. (2021). Hyperspectral Data Simulation (Sentinel-2 to AVIRIS-NG) for Improved Wildfire Fuel Mapping, Boreal Alaska. Remote Sensing 2021, Vol. 13, Page 1693, 13(9), 1693. https://doi.org/10.3390/RS13091693

2.2. Introduction

Wildfires are of great importance when it comes to plant succession, natural regeneration, reducing debris accumulation, maintaining ecosystem health, diversity, nutrient cycle, and energy flow [1]. Since excess of anything causes harm, increase in wildfire frequency and area burned also poses a risk to the ecosystem's health and diversity. Severe wildfires are occurring globally every year, causing unprecedented ecological and economic damage. In 2019, a massive fire occurred in the Amazon rainforest, which attracted global attention. Again, in 2020, the Amazon forest suffered a severe loss from wildfires that burned an area of approximately 20,234 sq. km [2]. In the same year, Australia recorded a huge bushfire that burned an area of around 186,155 sq. km and nearly 3 billion animals were displaced [3]. In 2020, 17,230 sq. km in California burned from wildfires that spread over the West Coast of the United States, making 2020 the largest wildfire season recorded in California's modern history [4].

Alaska, the northernmost state of the US, has 509,904 sq. km of forested land [5]. Wildfires are a natural and essential part of Alaskan ecosystems. Nonetheless, wildfires in Alaska are increasing in frequency, area burned, and severity, mirroring the global increase in wildfire events [6,7]. In the last two decades (2001–2020: 127,671 sq. km), wildfires in Alaska have burned 2.5 times more forest than the previous two decades (1981–2000: 57,060 sq. km) [7]. In 2019, Alaska had 719 wildfires that burned nearly 10,500 sq. km of forest [8]. making it the 10th largest fire year in recorded history. Many of these fires were near major population centers along the Wildland Urban Interface (WUI). The societal impacts of WUI fires (i.e., risk to life and property, unhealthy air quality, and cost of suppression) can be reduced if fire managers have access to reliable fuel maps (that is, boreal vegetation maps) for the development of effective fuel and fire management strategies [9,10]. Enhanced fuel mapping is also essential for the strategic planning of wildfire mitigation [4].

Remote sensing is a viable approach to map the vegetation of the boreal forests, considering the region's remoteness and vastness [11,12,13,14,15]. The Landscape Fire and Resource Management Planning Tools Project (LANDFIRE) provides geospatial products to state and federal fire suppression agencies for wildfire mitigation [16,17]. The traditional map products provided by the LANDFIRE for Alaska's boreal domain lack granularity needed for fire management at the fire incident (meter) scale. LANDFIRE products are derived from Landsat 8 multispectral data, which has few spectral bands and moderate spatial resolution (30 m). Additionally, these products have classification accuracies in the range of 20% to 45%, leaving

considerable room for improvement [18]. In Alaska, effective management of fuels and active fire requires improved fuel maps at the species level.

Advancements in airborne hyperspectral remote sensing provide an efficient approach to retrieve essential information for better characterization of forest fuels [14,19,20,21]. A number of studies have shown that hyperspectral data is much more effective than multispectral data for detailed vegetation mapping at species or stand scales [14,22,23,24,25,26,27,28,29,30]. The narrower bandwidths and improved spatial resolution of airborne hyperspectral datasets makes them much more effective than multispectral datasets at distinguishing visually similar vegetation classes. However, one of the major challenges with airborne hyperspectral technology is the cost of data acquisition. Currently, available hyperspectral datasets collected as part of the NASA Arctic-Boreal Vulnerability Experiment (ABoVE) and Goddard's LiDAR, Hyperspectral, and Thermal Imager (G-LiHT) programs cover only a small portion of the boreal domain. There is a need for greater spatial coverage and frequency while providing detailed spectral information similar to hyperspectral datasets.

Few studies have attempted to address this need through the simulation of hyperspectral data using publicly available multispectral datasets [31,32,33]. Zhang et al. [33] proposed a spectral response approach that used the Universal Pattern Decomposition Method (UPDM) for hyperspectral simulation from Landsat 7 Enhanced Thematic Mapper Plus (ETM+) and Moderate Resolution Imaging Spectroradiometer (MODIS) data. Liu et al. [31] followed a similar approach in which they simulated 106 hyperspectral bands from EO-1 Advance Land Imager (ALI) multispectral bands using standard ground spectra of water, vegetation, and soil. They performed Land-Use and Land-Cover (LULC) classification using the Spectral Angle Mapper (SAM) classifier and obtained an overall accuracy of 87.6% from the simulated hyperspectral data compared to 86.8% from ALI data. Tiwari et al. [32] used a similar simulation technique to generate a LULC map for a site located in northern India. They simulated hyperspectral data from Landsat 8 Operational Land Imager (OLI) multispectral data using spectra of vegetation, water, and sand as the endmembers. Using a SAM classifier, they obtained an overall accuracy of 69.4% from simulated hyperspectral data compared to 63.0% accuracy from OLI data.

Airborne Visible/Infrared Imaging Spectrometer-Next Generation (AVIRIS-NG) is the most advanced imaging spectrometer developed by NASA's Jet Propulsion Laboratory (JPL). The AVIRIS-NG sensor offers a higher signal-to-noise ratio, excellent system calibration, and more accurate image geo-rectification [34]. The data are available at wavelengths ranging from 380 to 2510 nm with a 5 nm bandwidth, at spatial resolutions of a few meters (depending on flying height)

(Figure 2.1). Previous studies [31,32] attempted to simulate Hyperion data from EO-1 ALI and Landsat 8 OLI multispectral datasets in order to improve LULC classification. The Hyperion sensor flew on the EO-1 satellite from 2000 to 2017, and it has 242 spectral bands in the range of 400–2500 nm and 30 m spatial resolution [35]. Simulation of AVIRIS-NG data is as yet unexplored, and that offers an opportunity to explore AVIRIS-NG data simulation to generate low-cost hyperspectral data for improved vegetation and LULC mapping. Sentinel-2 is the most recent multispectral sensor with global coverage and open data access. It has 13 spectral bands (spatial resolution: 10 m for visible-near infrared bands, and 20 m for SWIR bands) (Figure 2.1), especially the presence of red edge, NIR, and SWIR bands, and higher spatial resolution makes it apt for hyperspectral simulation [36,37,38].

The overarching goal of this study is to generate low-cost and high-quality hyperspectral data from widely available Sentinel-2 data to meet the need for greater spatial and temporal coverage of hyperspectral data for improved vegetation and fuel mapping in the boreal forest. In this study, we simulated an AVIRIS-NG hyperspectral dataset from a Sentinel-2 multispectral dataset using the UPDM spectral reconstruction approach for the boreal forest of Alaska. Since birch (*Betula papyrifera*: a deciduous species) and spruce (*Picea mariana*: a coniferous species) are the dominant trees at the test site, and accurately distinguishing coniferous and deciduous forest is essential for fire behavior modeling, we used the spectra of birch, spruce, and gravel (bare ground and rocky areas) as the endmembers for simulation. We visually and statistically compared the results of the simulated hyperspectral dataset with the AVIRIS-NG dataset.

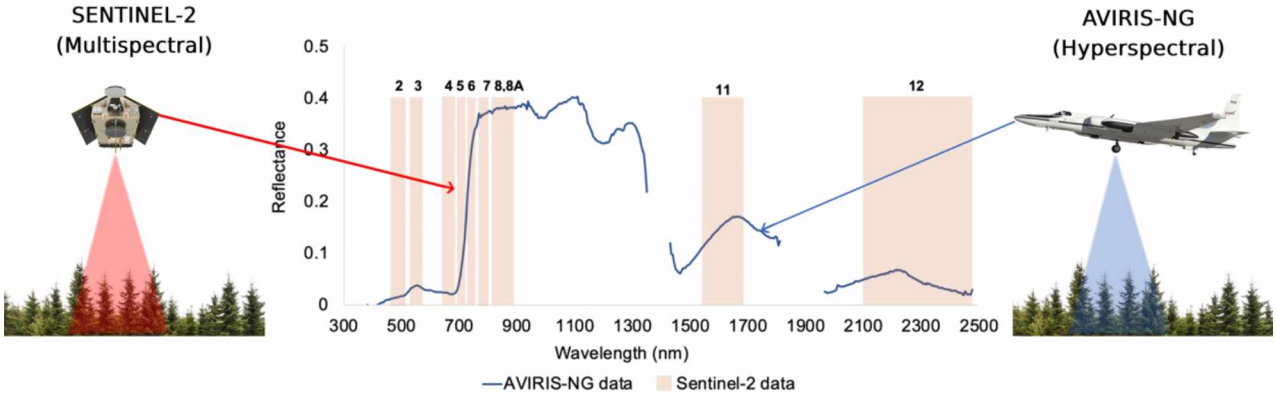


Figure 2.1: The graph at the center shows of AVIRIS-NG and Sentinel-2 bands. A reflectance profile of a vegetation pixel extracted from AVIRIS-NG (blue line). The columns represent Sentinel-2 bands (cream color); numbers at the top of the column are Sentinel-2 band numbers.

2.3. Materials and methods

2.3.1. Study area

The Caribou-Poker Creeks Research Watershed (CPCRW) is spread over a 104 square km area reserved for scientific study, including ecology, meteorology, and hydrological research. CPCRW is located in interior Alaska, 64 km northeast of Fairbanks (65.15° N, 147.50° W). We selected a test site within CPCRW for this study (Figure 2.2), where we had availability of an AVIRIS-NG scene. The air temperature varies from winter minima of –50 °C to summer peaks reaching 33 °C, with a long-term annual mean temperature of −3 °C. This area is typically under snow cover between October and April. The mean annual precipitation is about 262 mm, and 30% of it is in the form of snowfall [39].

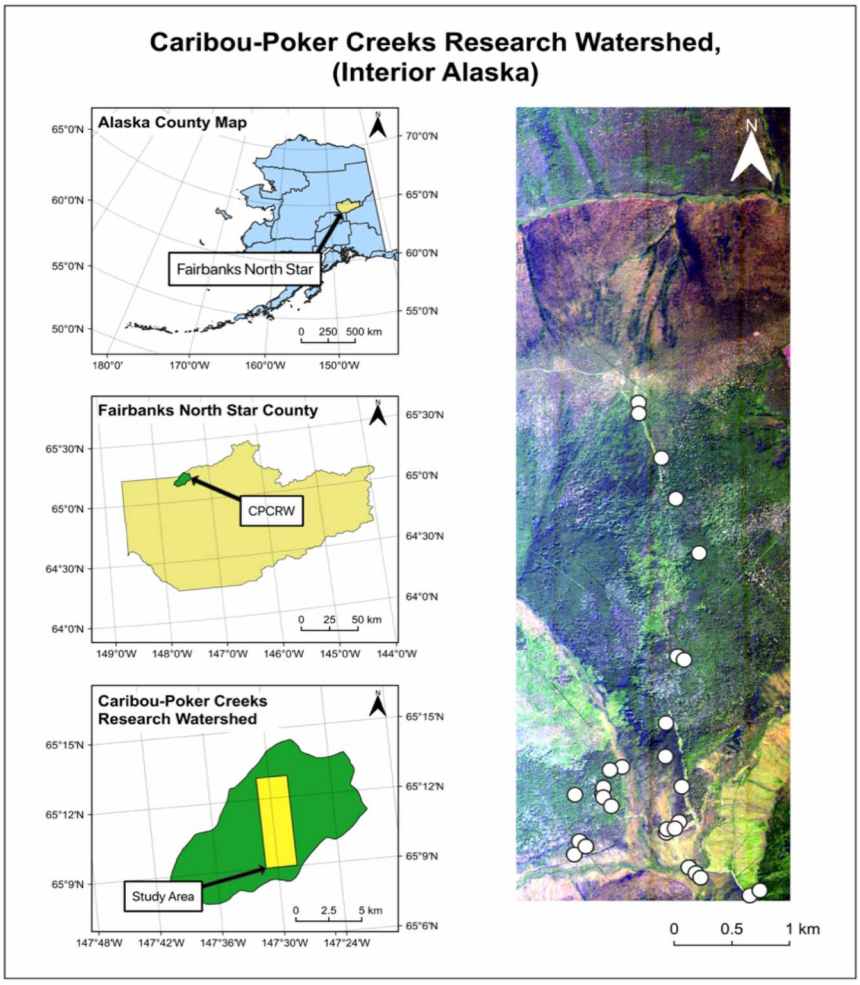


Figure 2.2: Study area: Caribou-Poker Creeks Research Watershed (CPCRW). Right: AVIRIS-NG subset (R:54, G:36, B:18; date acquired: 21 July 2018); white dots show the field survey locations.

2.3.2. Processing workflow

Figure 2.3 shows the processing workflow. The input data consists of Sentinel-2 multispectral imagery, the Spectral Response Function (SRF) of Sentinel-2 and AVIRIS-NG sensors, and spectra of birch, spruce, and gravel collected using the Spectral Evolution® PSR + 3500 hand-held spectroradiometer (Spectral Evolution Inc., Lawrence, MA, USA). The PSR + 3500 provides reflectance data in the range of 350–2500 nm at 1 nm spectral resolution for a total of 2151 channels.

The methodology is divided into four major phases: (1) field data collection, (2) remote sensing data preprocessing, (3) hyperspectral simulation, and (4) validation.

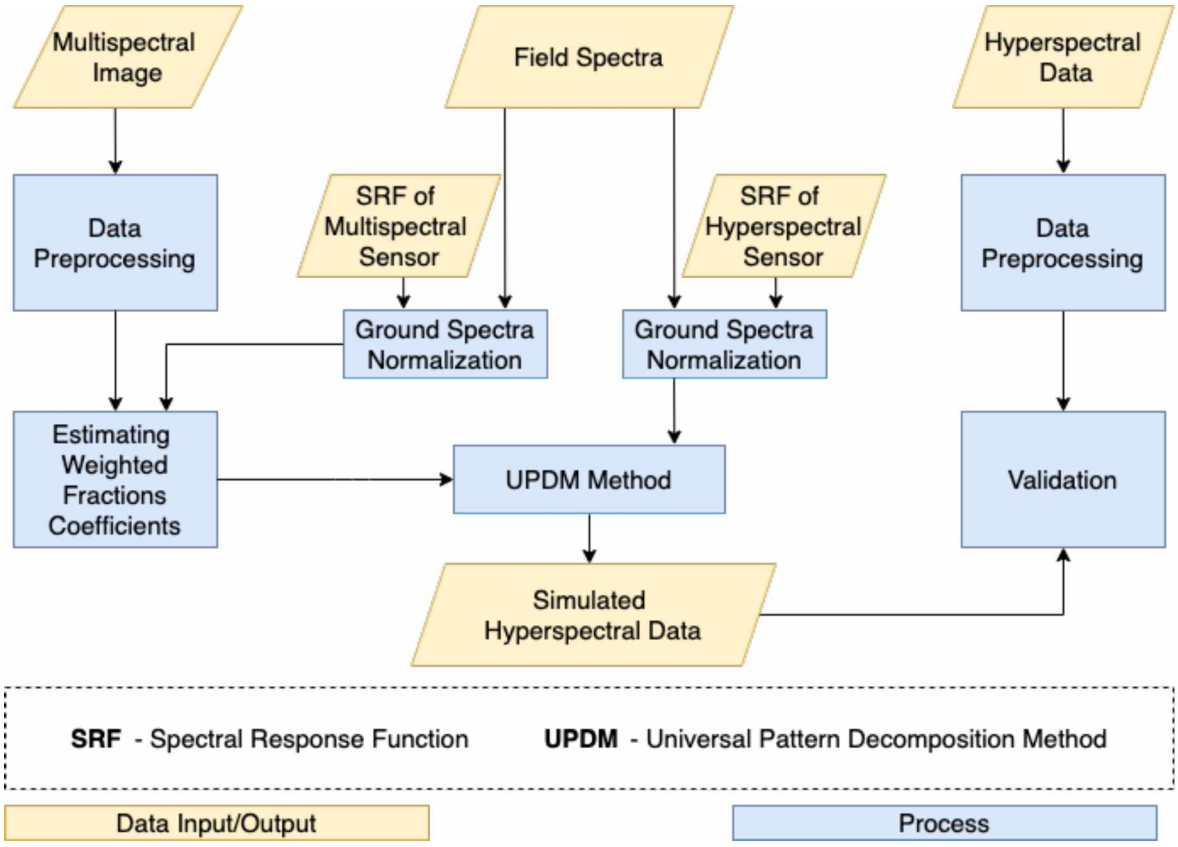


Figure 2.3: Flowchart showing processing workflow of hyperspectral simulation and validation.

2.3.3. Field data collection

We collected all field data during the summer of 2019 and 2020. We collected several leaf spectra samples for different tree/shrub species using a PSR + 3500 Field Spectroradiometer. We collected the field spectra on 17 August 2019 between 11:00 to 14:00 (weather: sunny with clear sky; solar noon: 14:06). We collected spectra holding the optic 2 inches away from leaves

and collected a minimum of 4 samples for each endmember. We used the mean endmember spectra in the simulation [20].

For the image classification, we recorded tree locations from stands where one type of tree species was present in clusters or groups. This enabled us to identify near to pure pixels for training and testing the image classifier as well as to reduce the background noise. In Figure 2.2, the white dots denote the locations of the sample sites. We surveyed sample sites using a Trimble Real-Time Kinematic (RTK) Global Positioning System (GPS) unit that offers millimeters positional accuracy. The study site (CPCRW) is part of protected state forests. The vegetation change at this site due to natural succession takes places at multiple decades to century time scales. However, dramatic vegetation change can occur due to wildfires or insect outbreaks. During the field survey, we did not observe any evidence of fire or insect outbreak within the study area. Also, we are not aware of any report of forest damage or change in the study areas since 2018 (when the AVIRIS-NG image was collected). So, we are certain that the use of field data collected in 2019 and 2020 for image classifier training and classification accuracy assessment are reasonable and resulted in accurate and reliable map products.

2.3.4. Remote sensing data preprocessing

2.3.4.1. Multispectral data preprocessing

We used atmospherically corrected Sentinel-2 Level-2A reflectance data available from the European Space Agency (ESA) Copernicus Open Access Hub [38] acquired on 24 July 2018. Sentinel-2 bands are available in different resolutions. The visible bands (band 2, 3, and 4) and the NIR band (band 8) have 10 m resolution, while the vegetation red edge bands (bands 5, 6, 7, and 8A) and the SWIR bands (band 11 and 12) have 20 m resolution. We resampled the pixels of all the bands with 20 m resolution to the lowest pixel resolution of 10 m to keep the pixel counts the same for all bands in the simulation. We removed coastal aerosol, water vapor, and cirrus bands from the data, and layer-stacked the remaining bands. From the stacked data, we clipped out the study area. Sentinel-2 data preprocessing was performed in the Quantum GIS (QGIS) software version 3.4 developed by the QGIS development team [40].

2.3.4.2. Hyperspectral data preprocessing

In this study, we used an AVIRIS-NG level 2 [41,42] product acquired on 21 July 2018, which covers a portion of CPCRW. The AVIRIS-NG scene has 425 bands and 5 m spatial resolution. Some of these bands were removed since they were from wavelengths dominated by water vapor and methane absorption and contained noise due to atmospheric scattering and poor radiometric correction. We refer to such bands as bad bands. All the bad bands were removed

from the original scene using the ENVI classic software [43]. We manually visualized each band and removed the noisy bands, resulting in a 332-band subset. Table 2.1 identifies all the bands which we removed from the original AVIRIS-NG data [44]. We used a spatial subset of the AVIRIS-NG scene for the study.

Table 2.1: List of bad bands removed from AVIRIS-NG.

Bands	Wavelength (nm)	Remarks
1–30	376.85–522.09985	Noise due to atmospheric scattering and poor sensor radiometric calibration
196–210	1353.55–1423.67	Water vapor absorption bands
288–317	1814.35–1959.60	Water vapor absorption bands
408–425	2415.39–2500.00	Noise due to poor radiometric calibration and strong water vapor and methane absorption

2.3.5. Hyperspectral simulation

The process of hyperspectral data simulation is divided into three steps: (1) ground spectra normalization, (2) calculation of weighted fractional coefficients, and (3) hyperspectral data simulation.

2.3.5.1. Ground spectra normalization

We used ground spectra from multiple locations for all three endmembers: birch, spruce, and gravel, and used their mean spectra in the simulation. We normalized each endmember spectrum by convolving it with the spectral response function (SRF) of both the multispectral and the hyperspectral sensors. The SRF is the probability that the sensor will detect a photon of a given frequency and it depends on the central wavelength and the bandwidth of the sensor [45]. The Sentinel-2 SRF was obtained from the Sentinel-2 document library [46]. The SRF of AVIRIS-NG was not directly available, but the Full Width at Half Maximum (FWHM) values were available. We used a Gaussian function to generate the AVIRIS-NG SRF [31], assuming that the peak of the Gaussian curve with respect to the central wavelength is at 1 (Equation (2.1)). We used Equation (2.2) to determine the bandwidth, σ.

$$g(\overline{\lambda}_i, \sigma_i) = exp^{-\frac{(\overline{\lambda}_i - \lambda)^2}{2\sigma^2}} \quad (2.1)$$

$$\sigma_i = \frac{FWHM_i}{2\sqrt{2ln2}} \quad (2.2)$$

where:

g = gaussian function

i = band number

$\overline{\lambda}$ = central wavelength

σ = bandwidth

λ = wavelength

$FWHM_i$ = Full Width at Half Maximum values for each band

Using the above Gaussian function, we constructed the SRF for all the bands of AVIRIS-NG.

2.3.5.2. Calculation of weighted fractional coefficients

In this step, we used the Universal Pattern Decomposition Method (UPDM), a linear unmixing method, used to model landcover in proportion to the endmember spectrum present in each pixel of the image [31,32,47]. This method uses normalized ground spectra and the reflectance from multispectral data to estimate weighted fractional coefficients. This method assumes that each pixel of the multispectral data is a linear mixture of normalized ground spectra in the image using Equation (2.3):

$$R_i = \Sigma_{j=1}^{n}(P_{ij}.\ C_j) \quad (2.3)$$

where:

i = Number of bands (1 to m)

j = Number of endmember or class (1 to n)

R_i = Reflectance value of i^{th} pixel in the image

P_{ij} = Field spectra of the j^{th} component, i.e., classes

C_j = Fraction of coefficient of the j^{th} component within the pixel

We can represent the linear unmixing equation for all the pixels in the image in matrix form using Equation (2.6):

$$R = PC \tag{2.4}$$

$$R = P_b C_b + P_s C_s + P_g C_g \tag{2.5}$$

$$\begin{pmatrix} R_1 \\ R_2 \\ \vdots \\ R_n \end{pmatrix} = \begin{pmatrix} P_{1b} & P_{1s} & P_{1g} \\ P_{2b} & P_{2s} & P_{2g} \\ \vdots & \vdots & \vdots \\ P_{nb} & P_{ns} & P_{ng} \end{pmatrix} \cdot \begin{pmatrix} C_b \\ C_s \\ C_g \end{pmatrix} \tag{2.6}$$

where:

R = total pixel reflectance

C = proportion of class

P = normalized ground reflectance

b = birch

s = spruce

g = gravel

n = number of bands

For a multispectral sensor, we can represent Equation (2.4) as:

$$R_M = P_M C_M \tag{2.7}$$

C_M can be calculated via inversion by applying the least squares method in Equation (2.7):

$$C_M = (P_M^T . P_M)^{-1} . P_M^T . R_M \tag{2.8}$$

We calculated C_M using the multispectral data and Equation (2.8). It is the fraction of each endmember in a pixel (i.e., fractional coefficient) in the form of a matrix for the whole image. R_M is the matrix with reflectance values from Sentinel-2 multispectral data and P_M is a matrix that contains the normalized ground spectra (birch, spruce, and gravel).

2.3.5.3. Hyperspectral data simulation

This step requires the fractional coefficient image of the multispectral data and the SRF of the hyperspectral sensor as inputs. For a pixel, the proportion occupied by an endmember will be the constant at a constant spatial resolution, irrespective of the sensor type. The simulated hyperspectral data will have the same spatial resolution as Sentinel-2 data. Therefore, the fractional coefficients (C_M) calculated using the multispectral data (Section 2.3.5.2.) will be the same. We also normalized the ground spectra of the three classes using SRF of hyperspectral data, as mentioned in Section 2.3.5.1. By using these two matrices, we calculated the simulated reflectance values using Equation (2.9):

$$R_H = P_H . C_H \tag{2.9}$$

Since $C_H = C_M$, we can replace C_H in Equation (9) with value C_M from Equation (2.8):

$$R_H = P_H . (P_M^T . P_M)^{-1} . P_M^T . R_M \tag{2.10}$$

Here, in Equation (2.10), R_H contains the reconstructed band values of the hyperspectral data, in the form of a matrix. This matrix was written as a raster file (GeoTiff format).

We performed hyperspectral data simulation in Python 3 [48] using Pandas library [49] to handle the data in a data frame format. Further, we used the Numpy library [50] to perform the matrix calculations. Finally, we used the GDAL library [51] to work with raster, especially to read and write the image data.

2.3.6. Validation

We validated the simulated hyperspectral data using visual interpretation, statistical analysis, and by comparing image classification results.

2.3.6.1. Visual and statistical analysis

We observed spectral signatures of different classes collected from AVIRIS-NG data, Sentinel-2 data, and simulated hyperspectral data, and further validated them using the field data. We compared the reflectance values and visually analyzed the pattern of the spectra. We also calculated the Pearson's correlation coefficient to evaluate the relationship between the spectra of simulated hyperspectral data and AVIRIS-NG data.

We performed a visual comparison using the Colored Infrared (CIR) image, also known as False-Color Composite (FCC) image, generated with bands 97, 56, and 36 as RGB for the

AVIRIS-NG and simulated hyperspectral image, and with bands 8, 4, and 3 as RGB for the Sentinel-2 image. We considered and analyzed different areas of interest based on how they differ visually in terms of the landcover pattern.

We computed the band-to-band correlation between the simulated hyperspectral data and the AVIRIS-NG data. This analysis indicated the degree of similarity to AVIRIS-NG bands and allows us to identify bands with low correlation values.

2.3.6.2. Classification

We classified the simulated hyperspectral data, AVIRIS-NG hyperspectral data, and Sentinel-2 data, and then compared results to validate the simulated hyperspectral data. Due to the presence of a large number of bands in both hyperspectral datasets, it was essential to select a suitable classifier. We chose a Random Forest (RF) classifier [52] to perform the classification due to its ability to deal with many features (bands). Another advantage of using RF was that there are only two user-defined parameters: the number of decision trees and the number of features per subset. RF produces each decision tree independently, and it splits each node of the decision tree using a number of features [53]. We performed RF classification using the 'RandomForestClassifier' function of the scikit-learn library [54] in Python 3, and both user-defined parameters were kept constant in all three cases. A low number of decision trees tend to create a bias in the result when dealing with multidimensional datasets, while with a high number of trees, the error gets stabilized. Hence, we took 500 decision trees for training the classifier [53]. We obtained the features per subset by calculating the square root of the total number of bands. Therefore, in our case, the number of features per subset will be $\sqrt{}$ (332) ≈18. We trained the RF classifier using the field survey locations as a guide and performed species-level classification in all three cases.

We surveyed vegetation at 29 plots in the field, of which 30% were used for testing the classification accuracy while the remaining plots were used to train the classifier. The total number of pixels surveyed on the ground for each class are presented in Table 2.2.

Table 2.2: Class-wise total number of pixels surveyed on the ground during fieldwork.

Class	Number of Pixels
Spruce	1847
Birch	426
Alder	302
Gravel	129

When using a machine learning classifier for LULC classification, it is preferable to have the same number of pixels in all the classes [55]. In our case, the number of pixels in the training and testing datasets for each class was different (Table 2.2), so to balance the pixels in all the classes, we applied the Synthetic Minority Oversampling Technique (SMOTE) [56]. SMOTE is an oversampling technique that duplicates the classes having fewer samples using the minority data population. While it increases the data, it does not add any new information to the machine learning model.

For accuracy assessment of the three classification outputs, we calculated confusion matrices [57], which indicate how many pixels are correctly identified. From the confusion matrix, we can evaluate the accuracy of each class in terms of producer accuracy, user accuracy, and kappa value. Producer accuracy identifies how often the real features on the ground are correctly shown on the map. Conversely, the user accuracy indicates how often the class on the map will be present on the ground.

2.3.7. Fuel type classification

We classified the simulated hyperspectral data using a Random Forest classifier to generate a fuel map of the study area. We identified different fuel classes from the ground data based on the fuel guide provided by the Alaska Wildland Fire Coordinating Group [58]. We used ground data from 58 surveyed field plots in 2019 and 2020 and were able to identify a total of 7 fuel classes.

2.4. Results

We simulated 332 bands of AVIRIS-NG based on the Sentinel-2 multispectral data and performed species-level as well as fuel-level classification. Figure 2.4 shows color infrared (CIR) images of the simulated hyperspectral data along with the AVIRIS-NG and Sentinel-2 data at the study site. Visual comparison of AVIRIS-NG and simulated hyperspectral data demonstrated high spatial and spectral similarity (Figure 2.4). Since these images are in CIR composition, broadleaf vegetation appears bright red. The central region of the study site mostly consists of deciduous forest and dense canopy. The top and the bottom region of the study site are dark green due to the dominance of needle-leaved species (mostly black spruce).

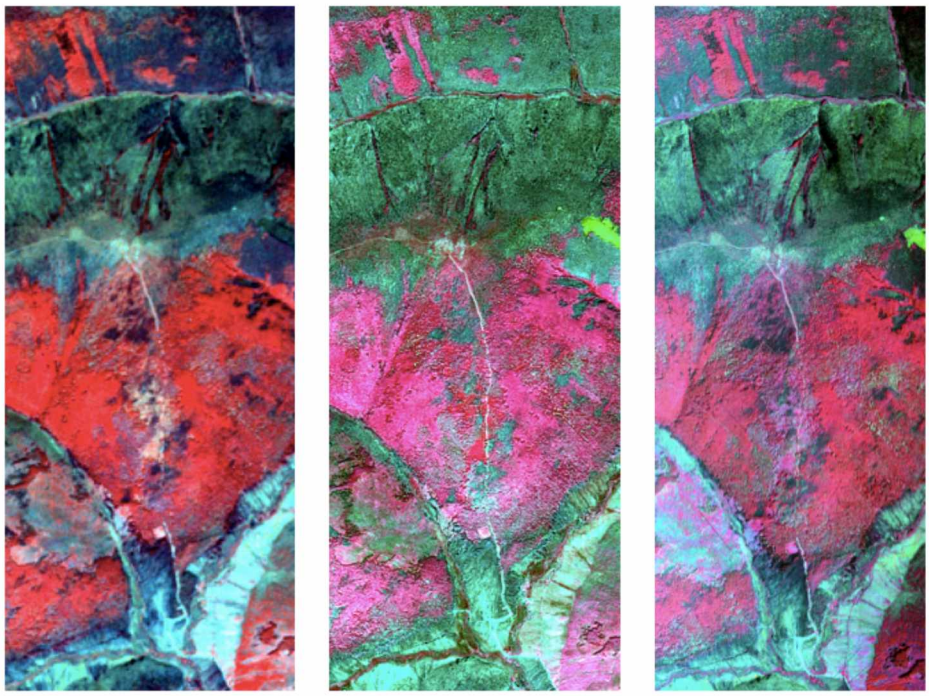


Figure 2.4: CIR image of Sentinel-2 (R: 8, G: 4, B: 3) data, simulated hyperspectral data (R: 97, G: 56, B: 36), and AVIRIS-NG data (R: 97, G: 56, B: 36).

2.4.1. Spectral profile comparison

The simulated hyperspectral data capture most of the absorption features and reflectance patterns present in the original AVIRIS-NG data. Figure 2.5 shows the comparison between spectral profiles of birch vs. spruce. The spectral signatures were selected from the regions where clusters of respective species were available on the ground.

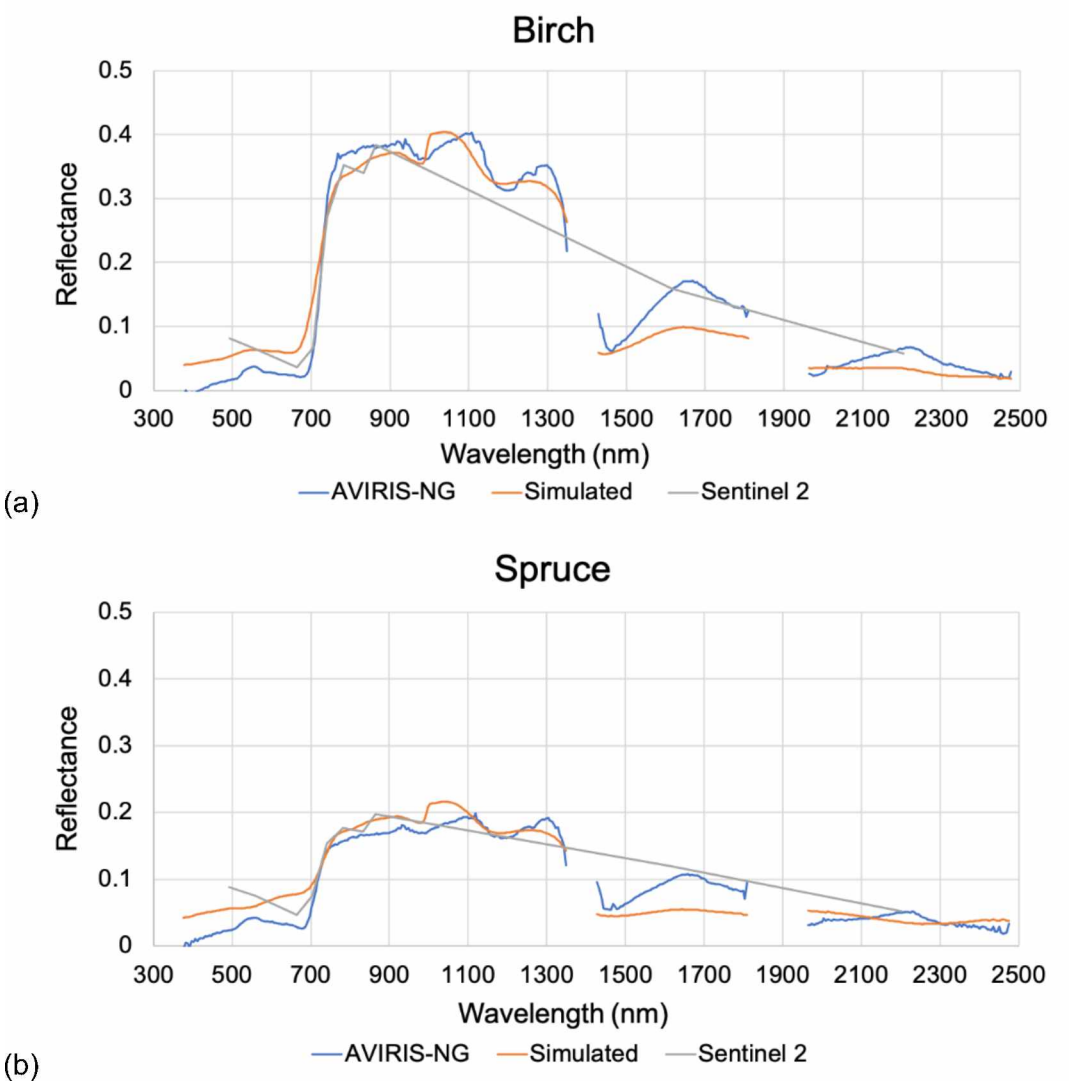


Figure 2.5: Comparison of spectral signature of (a) birch and (b) spruce for the three datasets.

We found correlation coefficients (r) of 0.97 and 0.92 between the reflectance values of the simulated hyperspectral data and the AVIRIS-NG data for birch and spruce, respectively. We also observed that for both cases, the spectra almost overlapped in the NIR region, while there were some minor deviations in the visible and the SWIR regions. The strong positive correlations confirm that the simulated hyperspectral data is capturing most of the absorption features and reflectance patterns present in the original AVIRIS-NG data.

2.4.2. Visual interpretation

The simulated hyperspectral data match very well with the actual hyperspectral data upon visual inspection (Figure 2.6). In Figure 2.6a, a trail can be identified in the middle of the study area. In the Sentinel-2 image, the trail was hardly visible, and it was difficult to discriminate

between the different vegetation classes, while in the case of the simulated hyperspectral image, the vegetation classes were easily differentiable, and the trail is clearly visible (enlarged in yellow circle). Indeed, the simulated hyperspectral image conveys a level of detail that looks similar to that of the original AVIRIS-NG image. In Figure 2.6b, we highlight a square patch of young alder and birch on the ground (in the yellow circle). In the simulated hyperspectral data and AVIRIS NG image, the features of the patch are easily distinguishable, but less so in the Sentinel-2 image. A third area with patches of low-growing vegetation including moss, cottongrass, tussock, and low shrub (blueberry and dwarf birch) was distinguished by the simulated hyperspectral image and AVIRIS-NG but not in the Sentinel-2 image (see yellow circle, Figure 2.6c). In the simulated hyperspectral image, more features and vegetation classes can be identified, similar to the AVIRIS-NG data. In contrast, in Sentinel-2, most of the area is covered by a single class.

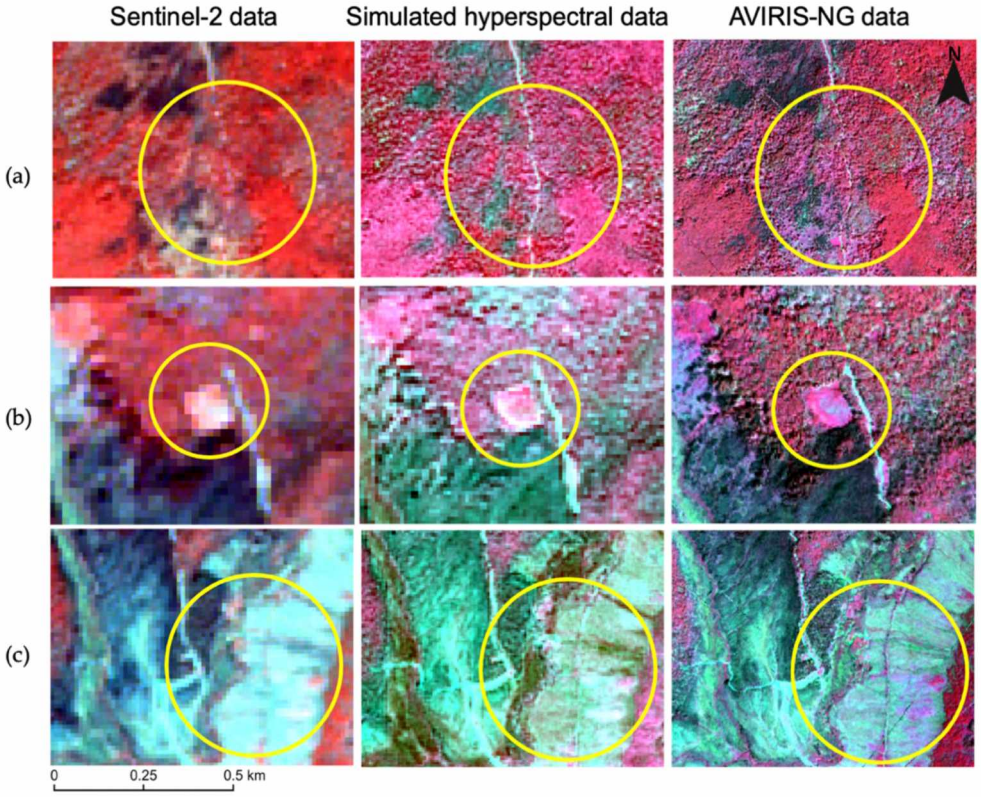


Figure 2.6: Visual analysis of the simulation result using CIR image composite for 3 areas: (a) central trail, (b) birch and alder patch, and (c) moss, blueberry, and dwarf birch.

2.4.3. Statistical analysis

In the simulated hyperspectral image, most bands showed good correlation with AVIRIS-NG, while a few showed a low correlation (Figure 2.7). There was high correlation in the NIR region, while correlation was poor in the visible and SWIR ranges.

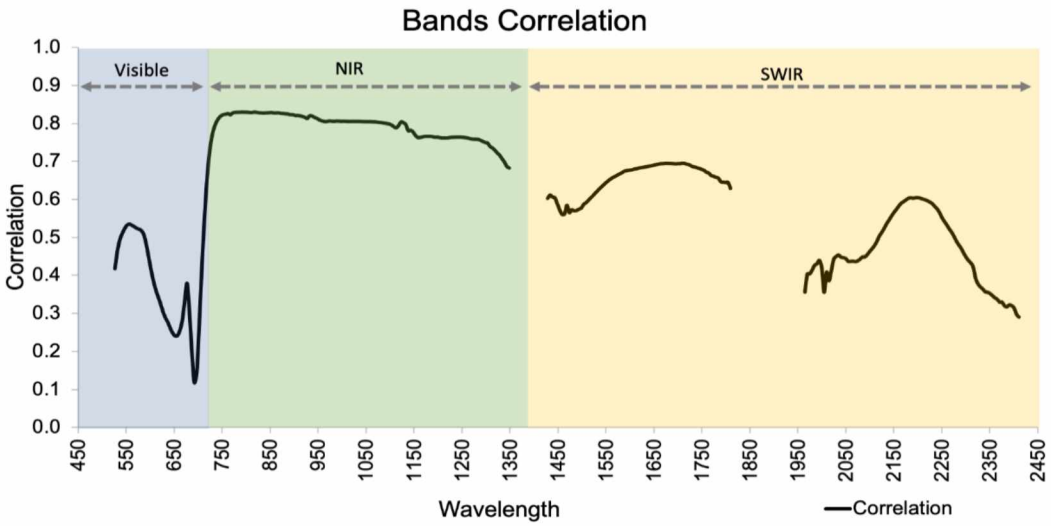


Figure 2.7: Band-to-band correlation between simulated hyperspectral and AVIRIS-NG data.

2.4.4. Image classification

Figure 2.8 highlights the results of species-level Random Forest classification. We performed the classification with four major classes: black spruce, birch, alder, and gravel. We obtained higher classification accuracy for simulated hyperspectral data than Sentinel-2 data. Table 2.3 shows the accuracy assessment of the three classification outputs. Since we considered only near to pure pixels for both training and testing, all three classes showed good classification accuracies. AVIRIS-NG performed the best with 94.6% accuracy and kappa = 0.93, followed by the simulated hyperspectral data showing 89% accuracy and a kappa value of 0.85, and finally Sentinel-2, with 77.8% accuracy and a 0.70 kappa value (Table 2.4).

For all the classes, the classified AVIRIS-NG dataset gave the best results for the user and the producer accuracy (Figure 2.9). Also, there was a substantial improvement in the accuracy of all the classes in the case of simulated hyperspectral data results when compared to the Sentinel-2 results.

To assess the effects of the different reflectance values on image classification accuracy, we reduced the reflectance of the original AVIRIS-NG data by 5% to 25% at an interval of 5% at each step and performed image classifications and accuracy assessments. We did not find any significant change in classification accuracy (Figure 2.10) due to a reduction in reflectance values. Based on these observations, we conclude that (up to 25%) differences in reflectance values (between original AVIRIS-NG and simulated hyperspectral data) have little or no impact on overall image classification accuracy.

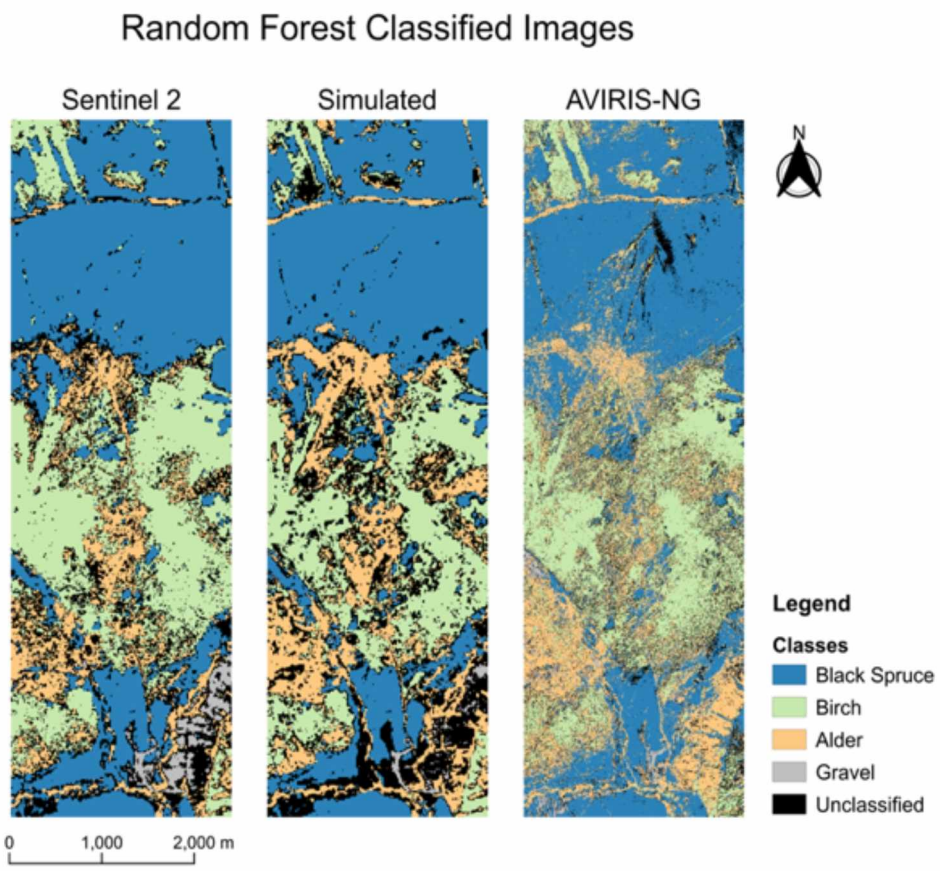
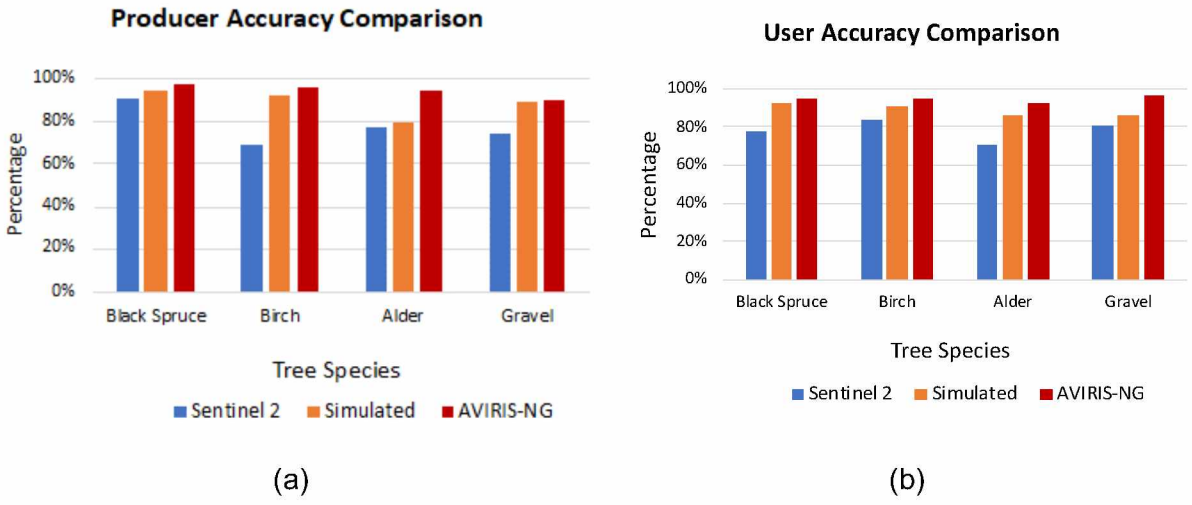


Figure 2.8: Tree species classification map generated using the Random Forest classifier for the three datasets.



(a) (b)

Figure 2.9: Class-wise comparison of (a) producer accuracy and (b) user accuracy obtained from the classification results for the three datasets.

Table 2.3: Confusion matrices of classification results for the three datasets.

Sentinel-2 Classification Confusion Matrix (Test Data)

		Reference Data				Total	Producer Accuracy (%)
		Black Spruce	Birch	Alder	Gravel		
Map Data	Black Spruce	642	33	9	22	706	90.9%
	Birch	0	488	218	0	706	69.1%
	Alder	0	61	543	102	706	76.9%
	Gravel	183	0	0	523	706	74.1%
Total		825	582	770	647	2824	
User Accuracy (%)		77.8%	83.8%	70.5%	80.8%		

Simulated Hyperspectral Classification Confusion Matrix (Test Data)

		Reference Data				Total	Producer Accuracy
		Black Spruce	Birch	Alder	Gravel		
Map Data	Black Spruce	666	0	40	0	706	94.3%
	Birch	53	653	0	0	706	92.5%
	Alder	0	42	563	101	706	79.7%
	Gravel	0	21	53	632	706	89.5%
Total		719	716	656	733	2824	
User Accuracy (%)		92.6%	91.2%	85.8%	86.2%		

AVIRIS-NG Classification Confusion Matrix (Test Data)

		Reference Data				Total	Producer Accuracy
		Black Spruce	Birch	Alder	Gravel		
Map Data	Black Spruce	688	0	17	1	706	97.5%
	Birch	0	679	5	22	706	96.2%
	Alder	0	39	667	0	706	94.5%
	Gravel	38	0	37	631	706	89.4%
Total		726	718	726	654	2824	
User Accuracy (%)		94.8%	94.6%	91.9%	96.5%		

Table 2.4. Overall accuracies of the classification results for the three datasets.

Data	Overall Accuracy
Sentinel-2	77.8%
Simulated hyperspectral data	89.0%
AVIRIS-NG data	94.4%

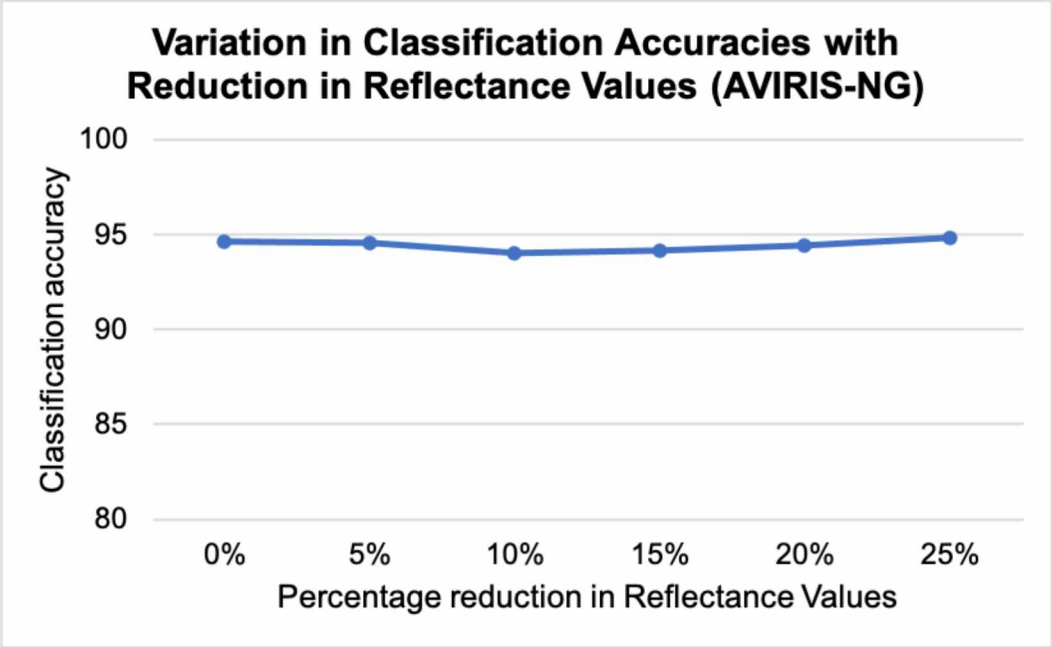


Figure 2.10: Variation of Accuracy with reduction in reflectance values of AVIRIS-NG data.

2.4.5. Fuel map

Upon fuel type classification, we found that the simulated hyperspectral data provided 65% overall accuracy, while classification accuracy of Sentinel-2 data was 56%. Figure 2.11 shows the fuel map, where we classified a total of 7 fuel types.

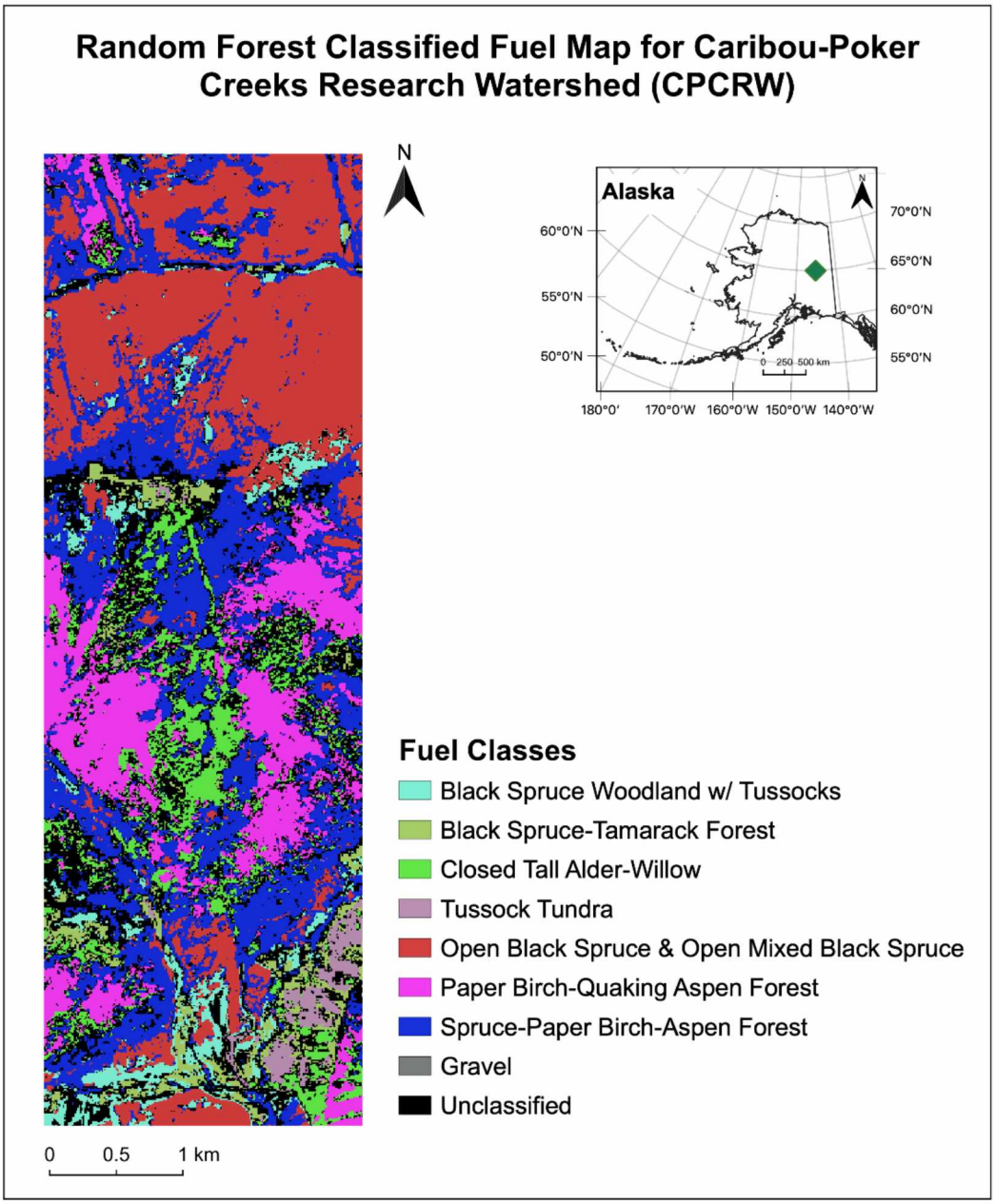


Figure 2.11: Fuel type map for study area generated using Random Forest classification on the simulated hyperspectral dataset.

2.5. Discussion

This study demonstrated the potential of simulated hyperspectral data for the purpose of forest fuel mapping. Visual inspection of RGB composites shows that the simulated hyperspectral image is similar to AVIRIS-NG image in texture, tone, and shading. The spectral comparison shows that the band-to-band correlations vary by wavelength, with highest correlations found in the NIR region, moderate in the SWIR region, low in the visible region, and very low along the

red-edge region (Figure 2.7). This is likely due to NIR scattering and non-linear mixing. In a study by Roberts et al. [19], non-linear mixing results in residual errors along the red-edge. These errors are present because plants do not scatter much in the visible region but do scatter in the NIR region. Since the NIR dominates the mixture, this results in high NIR correlation, but lower visible and SWIR correlation. We can minimize this problem by using field spectra collected at a scale that includes multiple scattering [20].

We found that the difference in reflectance values over the near infrared region (700–1400 nm) is relatively small, and the visual pattern of the spectra is also similar. Notable differences in the reflectance values in the SWIR region (1500–1800 nm) were observed. Zhang et al. [47] performed a similar simulation in which the simulated spectra showed little to no difference below 1000 nm, but a notable difference was found above 1000 nm wavelength when compared with the original spectra. This difference could be due to the variation in spatial resolution, especially in the SWIR region, 20 m for Sentinel-2 vs 5 m for AVIRIS-NG. The pixel resampling also contributed to the difference in reflectance value, where we resampled the 20 m pixel size of the Sentinel-2 SWIR region to a 5 m pixel size. The atmospheric corrections applied to Sentinel-2 data and AVIRIS-NG data were different due to the fact that Sentinel-2 data was captured from space while AVIRIS-NG data was captured from an aircraft at an altitude of 10.6 km, and that the data had different acquisition dates [59]. Therefore, the instantaneous field of view and the atmospheric corrections for these sensors are appreciably different, contributing to differences in reflectance values [31,60].

Visually, the simulated hyperspectral data appears similar to the AVIRIS-NG data, with minute spatial details preserved. The overall observation is that the simulated hyperspectral imagery provides an improved spectral resolution from Sentinel-2 imagery. We used three endmembers, and yet, areas of different vegetation cover types (moss, blueberry, and dwarf birch), which are not distinguishable in Sentinel-2 data, are clearly differentiable in the simulated hyperspectral data. In an open forest setting, woody materials such as downed logs, standing tree boles, dry grass, and leaf litter, together referred to as non-photosynthetic vegetation (NPV), can contribute to the reflectance of an image pixel [19]. In this study, we did not use NPV as an endmember. It would be interesting to further experiment with this simulation by adding a NPV variable in the UPDM equation as an endmember. Shade is another endmember that could be added to the equation, especially when working on the boreal forest where the canopy density is low.

In agreement with Liu et al. [31] and Tiwari et al. [16], we obtained higher classification accuracy from simulated hyperspectral data than the Sentinel-2 data (Table 2.4). The majority of misclassifications were gravel pixels. Gravel is mostly present on the narrow trails, and the young alder and birch patches present along the gravel trails were responsible for the misclassifications. Gravel was also misclassified with black spruce due to the open canopy structure, resulting in training pixels which included portions of ground reflectance reducing signal purity. In the case of Sentinel-2 results, birch was often misclassified with alder because of their spectral similarity, while simulated hyperspectral data performed better in discriminating these two species. This finding supports the notion that the simulated hyperspectral data can capture the minute spatial and spectral details of real hyperspectral data. The strength of this simulated dataset lies in providing spectrally enhanced data which can be used for detailed LULC classification. Tiwari et al. [32] used the UPDM technique to simulate Hyperion data for land cover classification at a test site in northern India, and obtained 6.45% improvement in mapping accuracy over ALI multispectral data. Likewise, in this study, we successfully simulated AVIRIS-NG hyperspectral data for species-level and fuel-level vegetation mapping at a test site in the boreal forest and obtained 11.2% improvement in mapping accuracy over Sentinel-2 data.

When we performed the fuel type classification, the simulated hyperspectral data achieved an overall classification accuracy of 65%. Smith et al. [14] carried out a detailed fuel type mapping from the original AVIRIS-NG data for the same study site and reported an accuracy of 61%. This suggests that simulated hyperspectral data can provide comparable mapping accuracy to real AVIRIS-NG data. Overall, these findings suggest that the generation of fuel maps from low-cost simulated hyperspectral data using the UPDM is feasible for Alaskan boreal forests.

2.6. Conclusions

The study aimed to simulate hyperspectral data from multispectral data and evaluate its utility compared to real hyperspectral data for fire fuel mapping. We found the universal pattern decomposition method (UPDM) to be a reliable algorithm for spectral unmixing. This algorithm requires ground measured spectra, and SRF from both multispectral and hyperspectral sensors. The algorithm is sensor-independent. Using UPDM, we successfully simulated 332 bands of AVIRIS-NG data from Sentinel-2 multispectral data. We validated the simulation results through visual interpretation, statistical comparison, and image classification. The visual inspection of simulated hyperspectral imagery reveals details of the vegetation fuel complex that are significant for predicting fire behavior but not discernible in the 30 m resolution multispectral imagery. There was a high correlation between the spectral signature of the tree species generated from actual

and the simulated hyperspectral data as well as high band-to-band correlation between both of the datasets. Finally, the classification results validated the improvement in fuel mapping accuracies for each class when compared with Sentinel-2 data. Our simulation results are encouraging and offer a path forward to generate a detailed fuel map for the entire boreal domain, which would be extremely useful for fire management and fuel treatment.

2.7. Acknowledgments

Thanks to Varun Tiwari for helping out during the initial stage of this work to help in understanding the basic concept. Special gratitude to the NASA-JPL team, including John W Chapman, Robert O. Green, and Sarah R Lundeen for providing the correct FWHM for AVIRIS-NG data. A heartfelt thanks to Colleen Haan, Robert Haan, and Brooke Kubby for assisting in fieldwork. Also, thanks to Utsav Soni for his support and help. Thanks to the NASA JPL and ESA for collecting and providing access to AVIRIS-NG scenes and Sentinel-2 data, respectively.
Funding: This material is based upon work supported by National Science Foundation and under the award OIA-1757348 and by the State of Alaska.

2.8. References

1. Leblon, B.; San-Miguel-Ayanz, J.; Bourgeau-Chavez, L.; Kong, M. Remote Sensing of Wildfires. In *Land Surface Remote Sensing: Environment and Risks*; Elsevier Inc.: Amsterdam, The Netherlands, 2016; pp. 55–95. ISBN 9780081012659.
2. NASA Earth Observatory Fires Raged in the Amazon Again in 2020. Available online: https://earthobservatory.nasa.gov/images/147946/fires-raged-in-the-amazon-again-in-2020 (accessed on 8 April 2021).
3. The Climate Reality Project Global Wildfires by the Numbers|Climate Reality. Available online: https://www.climaterealityproject.org/blog/global-wildfires-numbers (accessed on 8 April 2021).
4. CAL FIRE 2020 Fire Season|Welcome to CAL FIRE. Available online: https://www.fire.ca.gov/incidents/2020/ (accessed on 21 April 2021).
5. FS-R10-FHP. *Forest Health Conditions in Alaska 2019. A Forest Health Protection Report*; Publication R10-PR-45; U.S. Forest Service: Anchorage, AK, USA, 2019; 68p.
6. Box, J.E.; Colgan, W.T.; Christensen, T.R.; Schmidt, N.M.; Lund, M.; Parmentier, F.J.W.; Brown, R.; Bhatt, U.S.; Euskirchen, E.S.; Romanovsky, V.E.; et al. Key indicators of Arctic climate change: 1971–2017. *Environ. Res. Lett.* 2019, *14*, 045010.

7. Thoman, R.; Walsh, J.; Eicken, H.; Hartig, L.; Mccammon, M.; Bauer, N.; Carlo, N.; Rupp, S.; Buxbaum, T.; Bhatt, U.; et al. *Alaska's Changing Environment: Documenting Alaska's Physical and Biological Changes through Observations*; Review; University of Alaska Fairbanks: Fairbanks, AK, USA, 2019.
8. Alaska Department of Natural Resources Division of Forestry. *Alaska 2019 Fire Numbers*; Alaska Department of Natural Resources Division of Forestry: Anchorage, AK, USA, 2019.
9. Chuvieco, E.; Kasischke, E.S. Remote sensing information for fire management and fire effects assessment. *J. Geophys. Res. Biogeosci.* 2007, *112*.
10. Ziel, R. Alaska's Fire Environment: Not an Average Place—International Association of Wildland Fire. Available online: https://www.iawfonline.org/article/alaskas-fire-environment-not-an-average-place/ (accessed on 7 February 2021).
11. Fassnacht, F.E.; Latifi, H.; Stereńczak, K.; Modzelewska, A.; Lefsky, M.; Waser, L.T.; Straub, C.; Ghosh, A. Review of studies on tree species classification from remotely sensed data. *Remote Sens. Environ.* 2016, *186*, 64–87.
12. Xie, Y.; Sha, Z.; Yu, M. Remote sensing imagery in vegetation mapping: A review. *J. Plant Ecol.* 2008, *1*, 9–23.
13. Burai, P.; Deák, B.; Valkó, O.; Tomor, T.; Burai, P.; Deák, B.; Valkó, O.; Tomor, T. Classification of Herbaceous Vegetation Using Airborne Hyperspectral Imagery. *Remote Sens.* 2015, *7*, 2046–2066.
14. Smith, C.W.; Panda, S.K.; Bhatt, U.S.; Meyer, F.J. Improved Boreal Forest Wildfire Fuel Type Mapping in Interior Alaska using AVIRIS-NG Hyperspectral data. *Remote Sens.* 2021, *13*, 897.
15. Baldeck, C.A.; Asner, G.P.; Martin, R.E.; Anderson, C.B.; Knapp, D.E.; Kellner, J.R.; Wright, S.J. Operational Tree Species Mapping in a Diverse Tropical Forest with Airborne Imaging Spectroscopy. *PLoS ONE* 2015, *10*, e0118403.
16. Landfire: Existing Vegetation Type. Available online: http://www.landfire.gov (accessed on 10 February 2021).
17. Rollins, M. Landfire: A nationally consistent vegetation, wildland fire, and fuel assessment. *Int. J. Wildl. Fire* 2009, *18*, 235–249.
18. DeVelice, R.L. Accuracy of the LANDFIRE Alaska Existing Vegetation Map over the Chugach National Forest. 2012. Available online: https://landfire.cr.usgs.gov/documents/LANDFIRE_ak_110evt_accuracy_summary_013012.pdf (accessed on 26 April 2021).

19. Roberts, D.A.; Smith, M.O.; Adams, J.B.; Roberts, D.A. Green Vegetation, Nonphotosynthetic Vegetation, and Soils in AVIRIS Data. *Remote Sens. Environ.* 1993, *44*, 255–269.
20. Roberts, D.A.; Ustin, S.L.; Ogunjemiyo, S.; Greenberg, J.; Bobrowski, S.Z.; Chen, J.; Hinckley, T.M. Spectral and structural measures of northwest forest vegetation at leaf to landscape scales. *Ecosystems* 2004, *7*, 545–562.
21. Smith, C.W.; Panda, S.K.; Bhatt, U.S.; Meyer, F.J.; Haan, R.W. Improved Vegetation and Wildfire Fuel Type Mapping Using NASA AVIRIS-NG Hyperspectral Data, Interior AK. In Proceedings of the IGARSS 2020—2020 IEEE International Geoscience and Remote Sensing Symposium, Waikoloa, HI, USA, 26 September–2 October 2020; pp. 1307–1310.
22. Roberts, D.A.; Gardner, M.; Church, R.; Ustin, S.; Scheer, G.; Green, R.O. Mapping Chaparral in the Santa Monica Mountains Using Multiple Endmember Spectral Mixture Models. *Remote Sens. Environ.* 1998, *65*, 267–279.
23. Clark, M.L.; Roberts, D.A.; Clark, D. Hyperspectral discrimination of tropical rain forest tree species at leaf to crown scales. *Remote Sens. Environ.* 2005, *96*, 375–398.
24. Zhang, C. Combining hyperspectral and lidar data for vegetation mapping in the Florida everglades. *Photogramm. Eng. Remote Sens.* 2014, *80*, 733–743.
25. Singh, P.; Srivastava, P.K.; Malhi, R.K.M.; Chaudhary, S.K.; Verrelst, J.; Bhattacharya, B.K.; Raghubanshi, A.S. Denoising AVIRIS-NG data for generation of new chlorophyll indices. *IEEE Sens. J.* 2020, *21*, 6982–6989.
26. Salas, E.A.L.; Subburayalu, S.K.; Slater, B.; Zhao, K.; Bhattacharya, B.; Tripathy, R.; Das, A.; Nigam, R.; Dave, R.; Parekh, P. Mapping crop types in fragmented arable landscapes using AVIRIS-NG imagery and limited field data. *Int. J. Image Data Fusion* 2020, *11*, 33–56.
27. Hati, J.P.; Goswami, S.; Samanta, S.; Pramanick, N.; Majumdar, S.D.; Chaube, N.R.; Misra, A.; Hazra, S. Estimation of vegetation stress in the mangrove forest using AVIRIS-NG airborne hyperspectral data. *Model. Earth Syst. Environ.* 2020, 1–13.
28. Ahmad, S.; Pandey, A.C.; Kumar, A.; Lele, N.V. Potential of hyperspectral AVIRIS-NG data for vegetation characterization, species spectral separability, and mapping. *Appl. Geomat.* 2021, 1–12.
29. Badola, A.; Padalia, H.; Belgiu, M.; Prabhakar, M.; Verma, A. Mapping Tree Species Richness of Tropical Forest Using Airborne Hyperspectral Remote Sensing. Master's Thesis, University of Twente, Enschede, The Netherlands, 2019.

30. Varshney, P.K.; Arora, M.K. *Advanced Image Processing Techniques for Remotely Sensed Hyperspectral Data*; Springer: Berlin/Heidelberg, Germany, 2004.
31. Liu, B.; Zhang, L.; Zhang, X.; Zhang, B.; Tong, Q. Simulation of EO-1 Hyperion Data from ALI Multispectral Data Based on the Spectral Reconstruction Approach. *Sensors* 2009, *9*, 3090–3108.
32. Tiwari, V.; Kumar, V.; Pandey, K.; Ranade, R.; Agrawal, S. Simulation of the hyperspectral data using Multispectral data. In Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS), Institute of Electrical and Electronics Engineers Inc., Beijing, China, 10–15 July 2016; Volume 2016, pp. 6157–6160.
33. Zhang, L.; Fujiwara, N.; Furumi, S.; Muramatsu, K.; Daigo, M.; Zhang, L. Assessment of the universal pattern decomposition method using MODIS and ETM data. *Int. J. Remote Sens.* 2007, *28*, 125–142.
34. Townsend, P.A.; Foster, J.R. Comparison of EO-1 Hyperion to AVIRIS for mapping forest composition in the Appalachian Mountains, USA. In Proceedings of the International Geoscience and Remote Sensing Symposium (IGARSS), Toronto, ON, Canada, 24–28 June 2002; Volume 2, pp. 793–795.
35. USGS USGS EROS Archive—Earth Observing One (EO-1)—Hyperion. Available online: https://www.usgs.gov/centers/eros/science/usgs-eros-archive-earth-observing-one-eo-1-hyperion?qt-science_center_objects=0#qt-science_center_objects (accessed on 11 April 2021).
36. Grabska, E.; Hostert, P.; Pflugmacher, D.; Ostapowicz, K. Forest Stand Species Mapping Using the Sentinel-2 Time Series. *Remote Sens.* 2019, *11*, 1197.
37. Astola, H.; Häme, T.; Sirro, L.; Molinier, M.; Kilpi, J. Comparison of Sentinel-2 and Landsat 8 imagery for forest variable prediction in boreal region. *Remote Sens. Environ.* 2019, *223*, 257–273.
38. ESA Copernicus Open Access Hub. Available online: https://scihub.copernicus.eu/dhus/#/home (accessed on 23 November 2020).
39. NEON Caribou-Poker Creeks Research Watershed NEON|NSF NEON|Open Data to Understand our Ecosystems. Available online: https://www.neonscience.org/field-sites/bona (accessed on 3 March 2021).
40. QGIS Development Team. *QGIS Geographic Information System*; Version 3.14; Open Source Geospatial Foundation: Beaverton, OR, USA, 2020.
41. Gao, B.C.; Heidebrecht, K.H.; Goetz, A.F.H. Derivation of scaled surface reflectances from AVIRIS data. *Remote Sens. Environ.* 1993, *44*, 165–178.

42. NASA JPL AVIRIS-NG Data Portal. Available online: https://avirisng.jpl.nasa.gov/dataportal/ (accessed on 17 February 2021).
43. *Exelis Visual Information Solutions Version 5.3*; Exelis Visual Information Solutions Inc.: Boulder, CO, USA, 2010.
44. Harris Geospatial Solutions Preprocessing AVIRIS Data Tutorial. Available online: http://enviidl.com/help/Subsystems/envi/Content/Tutorials/Tools/PreprocessAVIRIS.htm (accessed on 17 November 2020).
45. Kim, D.S.; Pyeon, M.W. Aggregation of hyperion hyperspectral bands to ALI and ETM+ bands using spectral response information and the weighted sum method. *Int. J. Digit. Content Technol. Appl.* 2012, *6*, 189–199.
46. European Space Agency Sentinel-2 Spectral Response Functions (S2-SRF)—Sentinel-2 MSI Document Library—User Guides—Sentinel Online. Available online: https://sentinel.esa.int/web/sentinel/user-guides/sentinel-2-msi/document-library/-/asset_publisher/Wk0TKajiISaR/content/sentinel-2a-spectral-responses (accessed on 23 November 2020).
47. Zhang, L.; Furumi, S.; Muramatsu, K.; Fujiwara, N.; Daigo, M.; Zhang, L. Sensor-independent analysis method for hyperspectral data based on the pattern decomposition method. *Int. J. Remote Sens.* 2006, *27*, 4899–4910.
48. Python Core Team. *Python*; A Dynamic, Open Source Programming Language; Python Software Foundation: Wilmington, DE, USA, 2015.
49. McKinney, W. Data structures for statistical computing in Python. In Proceedings of the 9th Python in Science Conference, Austin, TX, USA, 28 June–3 July 2010; Volume 445, pp. 51–56.
50. Harris, C.R.; Millman, K.J.; van der Walt, S.J.; Gommers, R.; Virtanen, P.; Cournapeau, D.; Wieser, E.; Taylor, J.; Berg, S.; Smith, N.J.; et al. Array programming with {NumPy}. *Nature* 2020, *585*, 357–362.
51. GDAL/OGR contributors {GDAL/OGR} Geospatial Data Abstraction Software Library 2021. Available online: https://gdal.org/ (accessed on 26 April 2021).
52. Breiman, L. Random Forests. *Mach. Learn.* 2001, *45*, 5–32.
53. Belgiu, M.; Drăguţ, L. Random forest in remote sensing: A review of applications and future directions. *ISPRS J. Photogramm. Remote Sens.* 2016, *114*, 24–31.
54. Pedregosa, F.; Varoquaux, G.; Gramfort, A.; Michel, V.; Thirion, B.; Grisel, O.; Blondel, M.; Prettenhofer, P.; Weiss, R.; Dubourg, V.; et al. Scikit-learn: Machine Learning in {P}ython. *J. Mach. Learn. Res.* 2011, *12*, 2825–2830.

55. Douzas, G.; Bacao, F.; Fonseca, J.; Khudinyan, M. Imbalanced Learning in Land Cover Classification: Improving Minority Classes' Prediction Accuracy Using the Geometric SMOTE Algorithm. *Remote Sens.* 2019, *11*, 3040.
56. Chawla, N.V.; Bowyer, K.W.; Hall, L.O.; Kegelmeyer, W.P. SMOTE: Synthetic Minority Over-sampling Technique. *J. Artif. Intell. Res.* 2002, *16*, 321–357.
57. Congalton, R.G. A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sens. Environ.* 1991, *37*, 35–46.
58. Barnes, J.; Peter Butteri, N.; Robert DeVelice, F.; Kato Howard, U.; Jennifer Hrobak, B.; Rachel Loehman, N.; Nathan Lojewski, U.; Charley Martin, C.; Eric Miller, L.; Bobette Rowe, B.; et al. *Fuel Model Guide to Alaska Vegetation*; Alaska Wildland Fire Coordinating Group, Fire Modeling and Analysis Committee: Fairbanks, AK, USA, 2018.
59. NASA JPL AVIRIS-Next Generation. Available online: https://avirisng.jpl.nasa.gov/platform.html (accessed on 24 November 2020).
60. König, M.; Hieronymi, M.; Oppelt, N. Application of Sentinel-2 MSI in Arctic Research: Evaluating the Performance of Atmospheric Correction Approaches Over Arctic Sea Ice. *Front. Earth Sci.* 2019, *7*, 22.

Chapter 3 A Novel Method to Simulate AVIRIS-NG Hyperspectral Image from Sentinel-2 Image for Improved Vegetation/Wildfire Fuel Mapping, Boreal Alaska[2]

3.1. Abstract

Detailed vegetation maps are one of the primary inputs for forest and wildfire management. Hyperspectral remote sensing is a proven technique for detailed and accurate vegetation mapping. However, the availability of recent hyperspectral imagery in Alaska is limited because of the logistics and high cost involved in its acquisition. In this study, we simulated AVIRIS-NG (Airborne Visible InfraRed Imaging Spectrometer - Next Generation) hyperspectral data from widely available Sentinel-2 multispectral data using the Universal Pattern Decomposition Method (UPDM). The UPDM is a spectral unmixing technique that uses detailed ground spectra of vegetation classes and the Spectral Response Functions of AVIRIS-NG and Sentinel-2 sensors to simulate imagery with the same number of bands and spectral resolution as an AVIRIS-NG image. We simulated three images (each covering an area of 100 km × 100 km) from two ecoregions to test portability of the approach. We collected ground spectra of vegetation and bare ground during summers (2019–2021) using a PSR+ 3500 hand-held spectroradiometer and created a spectral library for this study. The Iterative Endmember Selection (IES) algorithm was used to optimize the spectral library and to select the most representative endmembers for simulation: birch, spruce, and gravel. We validated the simulated hyperspectral imagery by comparing it with available AVIRIS-NG images. The simulated image was visually and spectrally similar to the AVIRIS-NG image (RMSE of 0.03 and 0.02 for birch and spruce spectra, respectively). We applied the Random Forest image classification model to derive detailed vegetation maps from the simulated images. Our vegetation map showed an improvement of 33% in the map accuracy compared to the LANDFIRE EVT map. This study demonstrated an efficient and cost-effective approach to derive detailed vegetation maps at the Sentinel scene scale by simulating hyperspectral images in Google's cloud environment. It offers a novel pathway to generate detailed vegetation and fuel maps for the whole boreal region of Alaska to aid effective forest and fire management.

[2] Badola, A., Panda, S. K., Roberts, D. A., Waigl, C. F., Jandt, R. R., & Bhatt, U. S. (2022). A novel method to simulate AVIRIS-NG hyperspectral image from Sentinel-2 image for improved vegetation/wildfire fuel mapping, boreal Alaska. International Journal of Applied Earth Observation and Geoinformation, 112, 102891. https://doi.org/10.1016/J.JAG.2022.102891

3.2. Introduction

Alaska has an area of approximately 78 million hectares covered with boreal forests (Nowacki et al., 2003; U.S. Geological Survey, 2001). These forests extend from the Coast Range in the south to the Brooks Range in the north. Wildland fire is a ubiquitous feature of the boreal forest and in the past two decades (2001–2020) wildfires burned 12.7 million hectares of forest in Alaska (International Arctic Research Center, 2021). Alaska's boreal forest is highly flammable because of the dominance of black spruce, which is highly combustible, and its low-lying canopy structure, which serves as ladder fuel and promotes crown fires and rapid fire spread. In boreal forests, the ground surface is covered with feather moss, lichen, and fine fuels that ignite easily in dry conditions (National Park Service, 2021). Fire managers and the research community aim to improve fire management by generating and using improved fire spread models, climate and fuel inputs. Vegetation/fuel maps are one of the key inputs for fire risk assessment and fire spread modeling. Fire managers need accurate vegetation maps to constrain fire spread by locating the potential areas of risk, appropriately allocating suppression resources, and applying fuel treatments.

Remote sensing of vegetation and forest is a proven approach to mapping vegetation type and wildfire fuel distribution (Dudley et al., 2015, Wagner et al., 2018, Xie et al., 2008, Badola et al., 2019, Badola et al., 2021a, Smith et al., 2021). Specifically, multispectral sensors onboard Landsat, Sentinel 2A and 2B, and Terra/Aqua satellites image the entire globe, and their image data are heavily used for vegetation mapping (Dobrinić et al., 2021, Grabska et al., 2019, Mudele and Gamba, 2019). In the USA, Landscape Fire and Resource Management Planning Tools (LANDFIRE) is a shared program between the U.S. Department of Agriculture and the U.S. Department of the Interior (https://landfire.gov/about.php). It provides geospatial products, including vegetation and fuel maps, to state and federal fire management agencies for wildfire mitigation (Reeves et al., 2009). In Alaska, the LANDFIRE Existing Vegetation Type (EVT) map product derived from Landsat image data (9 bands) at 30 m spatial resolution is traditionally used for fire management, fire spread modeling, and risk assessment. The accuracy of the LANDFIRE EVT 2014 map product ranges between 20% and 45% as per accuracy assessment at three sites (Develice, 2012, Smith et al., 2021). The Alaska Center for Conservation Science (ACCS) offers the Alaska Vegetation and Wetland Composite (AVWC) map product, also generated from 30 m Landsat image data, which includes land cover, wetlands, and deep-water maps for Alaska to promote wetlands and deep-water habitat management (Alaska Vegetation and Wetland Composite, 2019).

Imaging spectroscopy or hyperspectral remote sensing provides an opportunity to generate improved vegetation and fuel maps. A hyperspectral remote sensing sensor images the landscapes in hundreds of narrow contiguous bands, making it more effective for vegetation mapping than a multispectral sensor. However, the high dimensionality of the hyperspectral data may reduce the map accuracy due to the Hughes phenomenon (Hsu, 2007). Despite the constraint posed by high data dimensionality, recent studies (Govender et al., 2019, Badola et al., 2021a, Smith et al., 2021) have shown that hyperspectral data provides more accurate and detailed vegetation/species maps than multispectral data for the boreal region of Alaska. The application of hyperspectral data in vegetation mapping is highly effective, but hyperspectral data are not readily available. NASA JPL provides airborne hyperspectral data collected using the state-of-the-art Airborne Visible/Infrared Imaging Spectrometer-Next Generation (AVIRIS-NG) sensor. These data are highly sought for vegetation/tree species mapping (Ahmad et al., 2021, Badola et al., 2021a, Clark et al., 2005, Hati et al., 2021, Salas et al., 2020, Singh et al., 2021, Smith et al., 2021, Zhang, 2014) due to their narrow bandwidth of 5 nm including a wavelength range of 400–2500 nm, meter-scale spatial resolution, and high signal-to-noise ratio. NASA's JPL team recently acquired AVIRIS-NG data over select sites in Alaska as part of the Arctic-Boreal Vulnerability Experiment (ABoVE) airborne campaign, but these acquisitions cover a fraction of Alaska's boreal forest, so the available hyperspectral data is insufficient for any regional scale vegetation/fuel mapping. In order to have hyperspectral image data for the whole boreal forest of Alaska, we conceived the idea of simulating AVIRIS-NG hyperspectral image data from widely available Sentinel-2 image data at Sentinel scene scale (100 km × 100 km) by modifying the approach developed by Badola et al. (2021b) and implementing it in Google's cloud environment for efficient processing.

Simulation of hyperspectral image data is an emerging research area in the field of remote sensing. A few studies have attempted to simulate hyperspectral data from multispectral data and ground spectra of vegetation and soil using a spectral reconstruction approach. Liu et al. (2009) simulated Hyperion data from ALI multispectral data using the Universal Pattern Decomposition Method (UPDM), a sensor-independent spectral unmixing technique (Zhang et al., 2007). The UPDM calculates the proportion of each class in a pixel and uses the spectral response function of the sensors. This approach was further tested by Tiwari et al. (2016) to simulate Hyperion data from ALI data for Land-Use and Land-Cover (LULC) mapping. They successfully simulated 70 Hyperion bands for a test site in Uttarakhand, India. Using UPDM, Badola et al. (2021b) successfully simulated AVIRIS-NG data (332 bands) from Sentinel-2 data for a boreal forest test site near Fairbanks, Alaska. They obtained higher classification accuracy from simulated data

(89%) than Sentinel 2 data (78%). In this study, our goal was to simulate the AVIRIS-NG data from the Sentinel-2 image data at Sentinel-2 scene scale for regional scale vegetation mapping and to test the simulation reproducibility across space and time. We had four research objectives:

A. Implement the Iterative Endmember Selection (IES) algorithm to derive the most representative endmember ground spectra for the boreal region of Alaska.
B. Improve upon the Badola et al. (2021b) simulation algorithm to generate AVIRIS-NG hyperspectral image at Sentinel-2 scene scale (100 km X 100 km).
C. Derive detailed vegetation maps from the simulated hyperspectral data using a machine learning classifier and assess model accuracy and portability across space.
D. Implement the hyperspectral image simulation and vegetation classification algorithms in the Google cloud platform for efficient processing and ease of sharing with the research community.

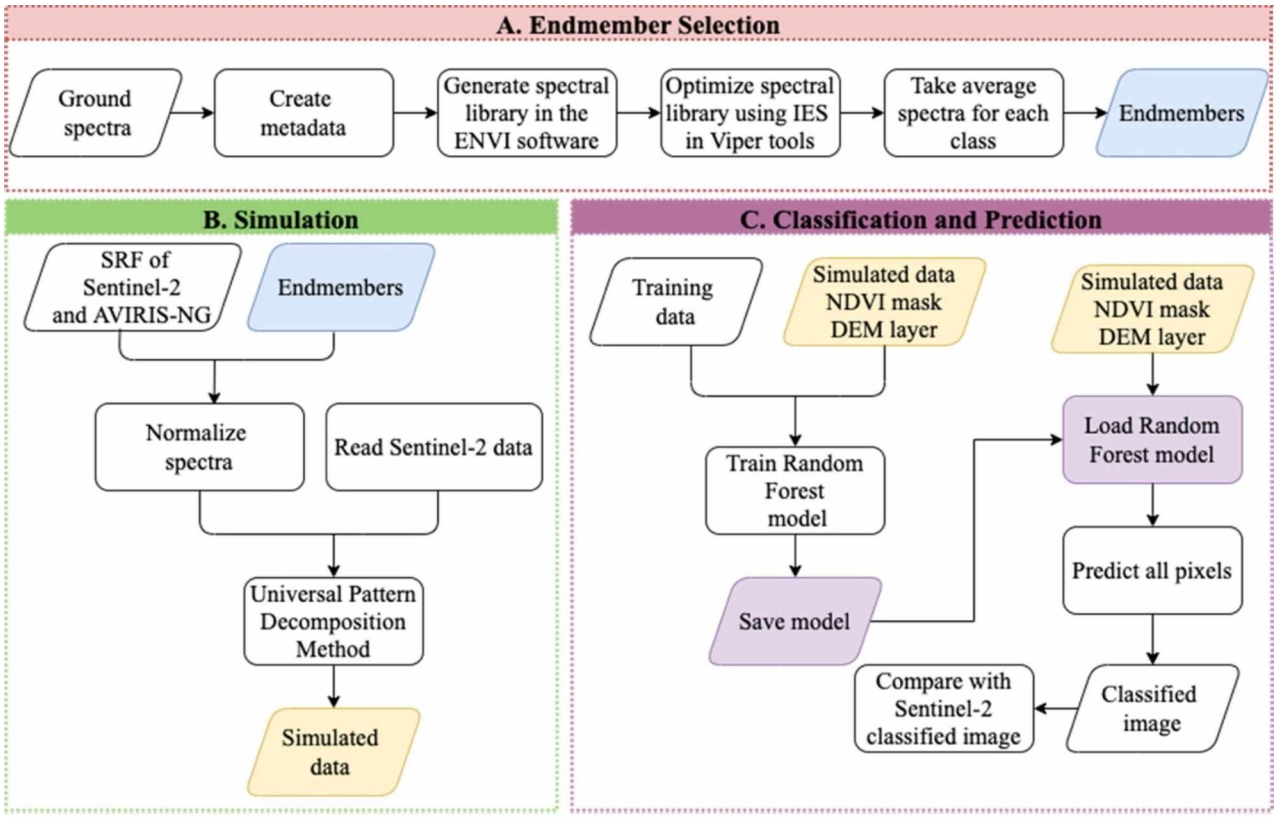
3.3. Materials and methods

The methodology is divided into three major phases: A) endmember selection, B) simulation, C) classification and prediction (Figure 3.1). We obtained endmembers as the output from the first phase. These were used as input for the simulation phase to simulate the AVIRIS-NG hyperspectral data. We applied a Normalized Differenced Vegetation Index (NDVI) mask and added a Digital Elevation Model (DEM) layer to the simulated data and trained a Random Forest (RF) model using training data collected from the field. Finally, we applied RF on simulated AVIRIS-NG to map vegetation class maps.

3.3.1. Field data collection

This research required extensive fieldwork to collect ground spectra of vegetation for hyperspectral data simulation as well as vegetation survey for classification of the simulated image. We collected field data (vegetation survey and leaf spectra of all major tree and tall shrub species) over three summers (2019 – 2021). We collected a total of 432 leaf spectra (15 – 20 spectra for each major tree/ tall shrub species) at three sites using a Spectral Evolution® PSR+ 3500 hand-held spectroradiometer (Spectral Evolution Inc., Lawrence, MA, USA). The PSR+ 3500 spectroradiometer provides reflectance data in the range of 350–2500 nm at 1 nm interval, comprising a total of 2151 channels. We collected leaf spectra by holding the optic at about 10 cm distance from the target between 11:00 to 16:00 (local time; local solar noon at 13:56) in sunny, clear-sky weather. We collected branch scale and leaf scale spectra for trees and canopy

scale spectra for shrubs (Figure 3.2). We also collected NPV spectra using a contact probe for tree bark. We targeted 12 public trails around the Fairbanks city and surveyed vegetation sites using a Garmin handheld GPS device that provides 3 m positional accuracy. At each site, we also recorded the information about canopy cover, vegetation composition, and understory vegetation that helped us in assigning vegetation class to a site. In Figure 3.3, the yellow triangles denote the locations of the collected ground data. Table 3.2 represents all the field data used in this study.



***IES**: Iterative Endmember Selection*
***VIPER**: Visualization and Image Processing for Environmental Research*
***SRF:** Spectral Response Function*
***NDVI**: Normalized Difference Vegetation Index*
***DEM**: Digital Elevation Model*
***AVIRIS-NG**: Airborne Visible InfraRed Imaging Spectrometer - Next Generation.*

Figure 3.1: Processing Workflow (A: Endmember Selection; B: Simulation; C: Classification and Prediction).

Figure 3.2: Vegetation spectra collection in the field; (a) branch scale spectra for birch using a bare fiber optic (b) branch scale spectra for black spruce using a bare fiber optic (c) leaf scale spectra for birch using a bare fiber optic (d) NPV spectra for birch bark using a contact probe.

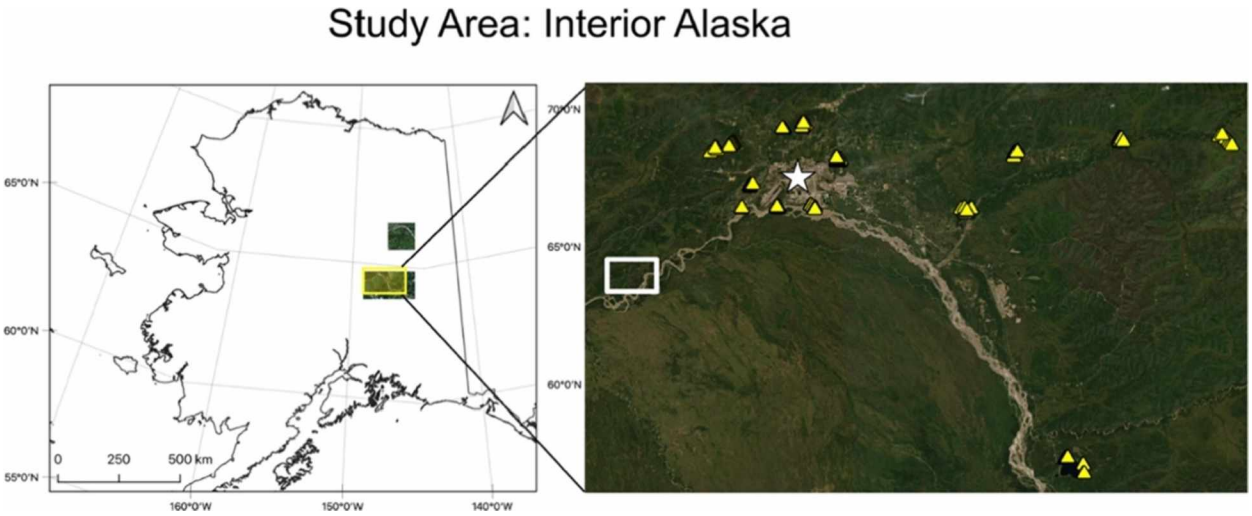


Figure 3.3: Study area: Showing the Sentinel-2 scenes used in this study (R: B8, G: B4, B: B3); yellow triangles mark the vegetation survey locations. The star shows the location of Fairbanks, Alaska and the rectangle shows the Bonanza Creek Experimental Forest site (BCEF).

Table 3.1: List of image datasets used in this study.

Data	Scene Identifier	Acquisition Date	Main area covered	Sub-ecoregion
Sentinel-2A	S2AT06WVS	July 22, 2018	Fairbanks	Tanana-Kuskokwim Lowlands
Sentinel-2B	S2BT06WWS	July 24, 2018	East of Fairbanks	Yukon-Tanana Uplands
Sentinel-2B	S2BT06WWU	July 01, 2021	Yukon flats	Yukon-Old Crow Basin
AVIRIS-NG	ang20180723t200207	July 23, 2018	Fairbanks	Tanana-Kuskokwim Lowlands
AVIRIS-NG	ang20170718t202618	July 18, 2017	South East of Fairbanks	Yukon-Tanana Uplands
AVIRIS-NG	ang20190705t192514	July 05, 2019	Yukon flats	Yukon-Old Crow Basin

Table 3.2: List of field data used in this study.

Data	Instrument	Location	Time of Data collection	Data collected
In-situ vegetation survey	Hand held Garmin GPS	12 public trails explored around Fairbanks	Summer, 2021	Vegetation composition, canopy cover, diameter and height
		University of Alaska, Fairbanks campus	Summer, 2019	
		BCEF	Summer, 2020	
Vegetation spectroscopy	PSR +3500 Spectroradiometer GPS	University of Alaska, Fairbanks campus	Summer, 2021	Spectra, sample location, vegetation type
		CPCRW	Summer, 2019 and 2021	
		BCEF	Summer, 2021	

The AVIRIS-NG scenes were acquired in 2018. Since we used these scenes to assess the spectral quality of simulated hyperspectral data, we used Sentinel scenes closer in date to the AVIRIS-NG scene acquisition dates (Table 3.1). In the Interior Alaska boreal forest, vegetation change can occur either due to natural succession, insect attack, wildfires, or anthropogenic disturbance such as timber harvesting. During fieldwork, we ensured that there was no evidence of insect outbreak or any major vegetation change at the study sites since 2018.

3.3.2. Data preprocessing

We used radiometrically and geometrically corrected Sentinel-2A level 1C (Top-of-atmosphere reflectance) data available from European Space Agency (ESA) Copernicus Open Access Hub (European Space Agency, 2014). Table 3.1 lists the datasets that were used in this study, including the sub-ecoregion where they belong (Nowacki et al., 2003). We used the Sen2cor processor (Louis et al., 2016) available in ESA's Sentinel Application Platform (SNAP) for atmospheric, terrain, and cirrus correction to obtain level 2A surface reflectance data. The size of each scene was 100 km × 100 km. Sentinel-2A data has 13 bands, from which we removed band 1 (coastal aerosol), band 9 (water vapor), and band 10 (SWIR-Cirrus). The visible bands (bands 2, 3, and 4) and NIR band (band 8) have 10 m resolution while the SWIR (bands 11 and 12) and vegetation red edge bands (bands 5, 6, 7, and 8A) have 20 m spatial resolution. We resampled all 20 m bands to 10 m to preserve the best possible spatial resolution and better match the 5 m resolution of AVIRIS-NG. We used atmospherically corrected level 2 AVIRIS-NG data (NASA JPL, 2018) with 425 bands and 5 m spatial resolution to validate the simulated data. We removed bands that contained excessive noise due to atmospheric scattering or are dominated by methane and water vapor absorption. We used the ASTER Global Digital Elevation Model (GDEM) Version 3 (EarthData, 2021) as an additional feature for image classification. It is available at a spatial resolution of 1 arc second (approximately 30 m) (Abrams et al., 2020).

3.3.3. Building a spectral library of boreal vegetation and endmember selection

We collected a total of 432 spectra (15–20 spectra for each species) of trees/shrubs from the boreal region of interior Alaska. We created a spectral library of all the collected ground spectra using ENVI classic software (Exelis Visual Information Solutions version 5.3, 2010). The PSR+ 3500 hand-held spectroradiometer also records the latitude, longitude, and elevation, along with each target spectrum. We extracted this information from all of the individual spectral files and created a separate metadata file for the spectral library. Furthermore, we did all spectral processing in the Visualization and Image Processing for Environmental Research (VIPER) Tools 2 (beta) software (Roberts et al., 2018).

Library pruning is an important step for creating a spectral library, as it reduces the size of a spectral library and provides the ideal spectra for each endmember. There are different library pruning techniques such as Endmember Average RMSE (EAR), Minimum Average Spectral Angle (MASA), Count-based Endmember Selection (CoB) and Iterative Endmember Selection (IES). These techniques rely on the square array (Roberts et al., 1997) that stores the information about how an endmember performs when used to unmix other spectra in the same library. The square array is an $n \times n$ grid of pixels where n is the total number of spectra and n along the row denotes the spectrum used for unmixing from other spectra. We can gather information about RMSE, shade fraction, and spectral angle from the square array. For more details on the square array refer to Roberts et al. (1997). In this study we used the IES (Roth et al., 2012, Schaaf et al., 2011) method for library pruning. It calculates the kappa coefficient (McHugh, 2012) to create a subset of spectra which provides the best class separability. IES classifies the entire spectral library using a subset from the original library. Endmembers are iteratively added and removed from the subset until kappa no longer improves. IES has been used for different applications. Roberts, et al. (2015) used IES to discriminate urban surface materials. IES was implemented to map vegetation species (Dudley et al., 2015, Roberts et al., 2015) and for improved burn severity mapping (Fernandez-Manso et al., 2016). The IES algorithm was implemented in the VIPER Tools 2 (beta) software. We ran IES in fully constrained mode (RMSE and fraction constrained) with default parameter settings, i.e., RMSE threshold of 0.025.

After pruning, we got 105 spectra out of 432 for 15 endmember classes (Table 3.3). We used the average of the spectra as the endmember for simulation. Hence, we ended up with 15 endmembers, including one for gravel. We initiated 15 endmembers as input for the simulation and compared our simulated product with AVIRIS-NG scene by collecting spectra from the known pixel (identified during field work) from both simulated and AVIRIS NG data. We reduced the number of endmembers in each iteration, tried different endmembers combinations (Table 3.4) for simulation, and simultaneously verified our results until we obtained similar spectra with low RMSE values. For the second iteration, we plotted all 15 endmembers (Figure 3.17), and removed similar endmember spectra. We ran the simulation model with 9 spectra. In the third iteration, we removed five more endmembers (alder, blueberry, larch, white spruce and asphalt). We kept one spectrum from each deciduous and coniferous class, gravel, and Non-Photosynthetic Vegetation (NPV) spectra (downed trunk). In the fourth iteration, we replaced downed trunk with asphalt and ran the simulation. In the fifth iteration, we used NPV spectra from the ECOSTRESS spectral library (Meerdink et al., 2019). We found a drastic discrepancy between spectra generated from simulated and AVIRIS-NG data in all the iterations. Finally, we used birch, black spruce and gravel

for simulation and found similar spectra with low RMSE values for spectra generated from simulated and AVIRIS-NG data.

Table 3.3: List of 15 endmembers obtained through Iterative Endmember Selection library pruning technique.

Endmembers		
Alder	Blueberry	Balsam poplar
Downed trunk	Black spruce	Carex
Dwarf birch	Green grass	Larch
Labrador tea	Birch	Gravel
Asphalt	Wild rose	White spruce

3.3.4. Simulation of hyperspectral data

The process of hyperspectral data simulation is divided into three steps: ground spectra normalization, weighted fraction coefficients, and hyperspectral data simulation.

3.3.4.1. Ground spectra normalization

We normalized three endmembers (black spruce, birch and gravel) by convolving them with the Spectral Response Function (SRF) (European Space Agency, 2017) of both Sentinel and AVIRIS-NG sensors. We obtained the SRF for Sentinel-2 data from the Sentinel 2 document library. We calculated the AVIRIS-NG SRF from Full Width at Half Maximum (FWHM) using Gaussian functions (Badola et al., 2021b; Liu et al., 2009).

3.3.4.2. Weighted fractional coefficient

We used the Universal Pattern Decomposition Method (UPDM) to estimate the proportion of each endmember in every pixel of the image (Badola et al., 2021b; Liu et al., 2009, Tiwari et al., 2016). UPDM is a linear unmixing method that is structured for satellite data analysis (Zhang et al., 2006). It assumes that reflectance at each pixel of an image is a linear mixture of normalized endmembers. The equation expressed in matrix form represents the linear unmixing of three endmembers (b: birch, s: spruce, g: gravel) (Equation (3.1)).

$$\begin{pmatrix} R_1 \\ R_2 \\ \vdots \\ R_n \end{pmatrix} = \begin{pmatrix} P_{1b} & P_{1s} & P_{1g} \\ P_{2b} & P_{2s} & P_{2g} \\ \vdots & \vdots & \vdots \\ P_{nb} & P_{ns} & P_{ng} \end{pmatrix} \cdot \begin{pmatrix} C_b \\ C_s \\ C_g \end{pmatrix} \tag{3.1}$$

Where R is the total pixel reflectance, C is the proportion of class, P is the normalized ground reflectance, and n is the band number.

For a multispectral sensor, we can represent Equation (3.1) as.

$$R_M = P_M C_M \tag{3.2}$$

C_M is the fraction of coefficients of each endmember in a pixel in the form of a matrix for the whole image. R_M is the matrix with reflectance values from Sentinel 2 multispectral data and P_M is a matrix that contains the reflectance values from the normalized endmembers.

C_M can be calculated from Equation (3.2) using reflectance from Sentinel 2 data by applying least square method:

$$C_M = (P_M^T . P_M)^{-1} . P_M^T . R_M \tag{3.3}$$

3.3.4.3. Hyperspectral data simulation

The spatial resolution of the simulated hyperspectral image will be same as in the Sentinel-2 image, therefore the fraction of coefficients (C_M) will remain the same. We normalized ground spectra (endmembers) using the SRF of the AVIRIS-NG sensor. Hence, we can calculate reflectance values using Equation (3.1) and Equation (3.3). The simulation method is discussed in more detail in Badola et al. (2021b). This simulated hyperspectral image has the same number of bands as AVIRIS-NG. Here, in Equation (3.4), R_H is the reconstructed reflectance values for the simulated hyperspectral image. We write out the simulated hyperspectral image file in GeoTiff file format.

$$R_H = P_H . (P_M^T . P_M)^{-1} . P_M^T . R_M \tag{3.4}$$

We implemented the hyperspectral image simulation in the Google cloud environment using Python 3 (Python Core Team, 2015) and a Jupyter notebook. We used the following libraries and packages: Pandas to handle the image data in a data frame; Numpy (Harris et al., 2020) to perform the matrix calculations; Rasterio (Gillies et al., 2013) to work with images, especially to read and write the image data. We implemented the algorithm by dividing a Sentinel scene covering an area of 100 km × 100 km into 36 square tiles of 2048 X 2048 pixels.

3.3.5. Simulated hyperspectral data validation

We validated the simulated data using spectral comparison, statistical analysis and visual interpretation. For spectral and statistical comparison, we extracted pixel spectra for birch and black spruce from AVIRIS-NG image data and the simulated hyperspectral image. These spectra were extracted from the pixels identified in the field. We compared the reflectance values and absorption peaks and visually analyzed the pattern of the spectra. We also calculated the Root Mean Square Error (RMSE) to evaluate the accuracy of the simulated birch and spruce pixel spectra.

We performed visual analysis by generating Colored Infrared (CIR) image using bands with wavelengths 843 nm, 662 nm, and 557 nm as RGB for the AVIRIS-NG and simulated hyperspectral image, and bands with wavelengths 842 nm, 665 nm and 560 nm as RGB for the Sentinel-2 image. We inspected and analyzed different areas of interest based on their visual appearance.

3.3.6. Image classification

We labeled each survey site to a vegetation class as per Viereck's Alaska Vegetation Classification (Viereck et al., 1992). For vegetation classes with similar spectra, we merged the classes (e.g., open birch forest and closed birch forest, Figure 3.18). Using an NDVI threshold of 0.3, we masked out the non-vegetated pixels. We identified 16 vegetation classes, including one 'Other' class that represents non-forested open vegetated areas such as grasslands.

We performed image classification using the 'RandomForestClassifier' function of the scikit-learn () in Python 3 with 500 decision trees and $\sqrt{(425)} \approx 20$ number of features per subset (Breiman, 2001, Pedregosa et al., 2011). We divided the survey data into two sets, a training set and testing set, and trained the classifier on a simulated hyperspectral scene (S2AT06WVS) that had the 16 vegetation classes representative of interior Alaska. The trained model was then used to classify two other simulated hyperspectral scenes to assess the model's portability to other sites.

3.3.7. Accuracy assessment

We performed the accuracy assessment using F1 scores and Intersection over Union (IoU). Each class has a different number of samples. The imbalance in classes can skew the results in favor of a more abundant class or classes with a greater number of samples and result in biased classification accuracy. In the event of imbalanced classes, classification accuracy is not enough to assess the classifier model performance. We used F1 score and IoU measures to evaluate the model performance. The F1 score is the harmonic mean of the precision and recall

of the model, with value ranging from 0 to 1. IoU (also known as the Jaccard index) measures the amount of overlap between the predicted and the actual label. A value of 0 means there is no overlap, while 1 denotes complete overlap. An IoU score greater than 0.5 is considered to be a good prediction.

We also compared our classified output with two available vegetation map products: LANDFIRE EVT (LANDFIRE, 2016) and the Alaska Vegetation and Wetland Composite (AKVWC) (Alaska Vegetation and Wetland Composite, 2019). We also compared the percentage cover of each species with the USDA Tanana Valley State Forest Pilot Inventory (Pattison et al., 2018).

3.4. Results

In this section, we present the spectral, statistical, and visual comparison of simulated hyperspectral image with AVIRIS-NG and Sentinel 2 image, image classification results, and comparison of our classified vegetation map with two other existing products.

3.4.1. Spectral and statistical comparison

Simulated spectra accurately captured the key absorption features that were available in AVIRIS-NG data. For spectral and statistical analysis, we removed the 93 bands with noise due to atmospheric scattering and poor radiometric correction, and bands dominated by water vapor and methane absorption (Badola et al., 2021b). Figure 3.4 shows the comparison of simulated and AVIRIS-NG spectra for birch and spruce vegetation. In spectra extracted from simulated data, the absorption features were similar to AVIRIS-NG spectra. These absorption features are caused by chlorophyll absorption in the red band (around 690 nm), water absorption in the NIR region, and lignin and cellulose absorption in the SWIR region. Around 1400 nm wavelength, a dip due to water absorption is present in both the spectra generated from simulated and AVIRIS-NG data. In both spectra, the difference in the reflectance values over the infrared region (700 nm − 1400 nm) is relatively small, and their pattern is similar. We achieved an RMSE of 0.03 and 0.02 for birch and spruce, respectively.

Figure 3.5 shows the band-to-band correlation between AVIRIS-NG and simulated hyperspectral images. A majority of the bands show high correlation, especially in the NIR and SWIR region. We also calculated the Coefficient of Variation (CoV = standard deviation/ mean) for the pixel difference between AVIRIS-NG and simulated hyperspectral images for 332 bands (Figure 3.6). The CoV is another metric to quantify and visualize the similarity between the two image data. A lower CoV value suggests higher similarity between the two image data, while a higher CoV value suggests less similarity. In our case, the CoV value ranges from 0.2 to 1.3, with

higher CoV values along trails and shaded areas where we expect lower simulation accuracy, as we did not include shade fraction as one of the end members. Both band-to-band correlation and the CoV of pixel differences between AVIRIS-NG and simulated hyperspectral images illustrate the quality of simulated hyperspectral image. The band-to-band correlation is higher than 0.6 for 252 bands, and 98% of the total pixels have a CoV less than 1 suggesting that the simulated image satisfactorily captures the spectral details at pixel scale.

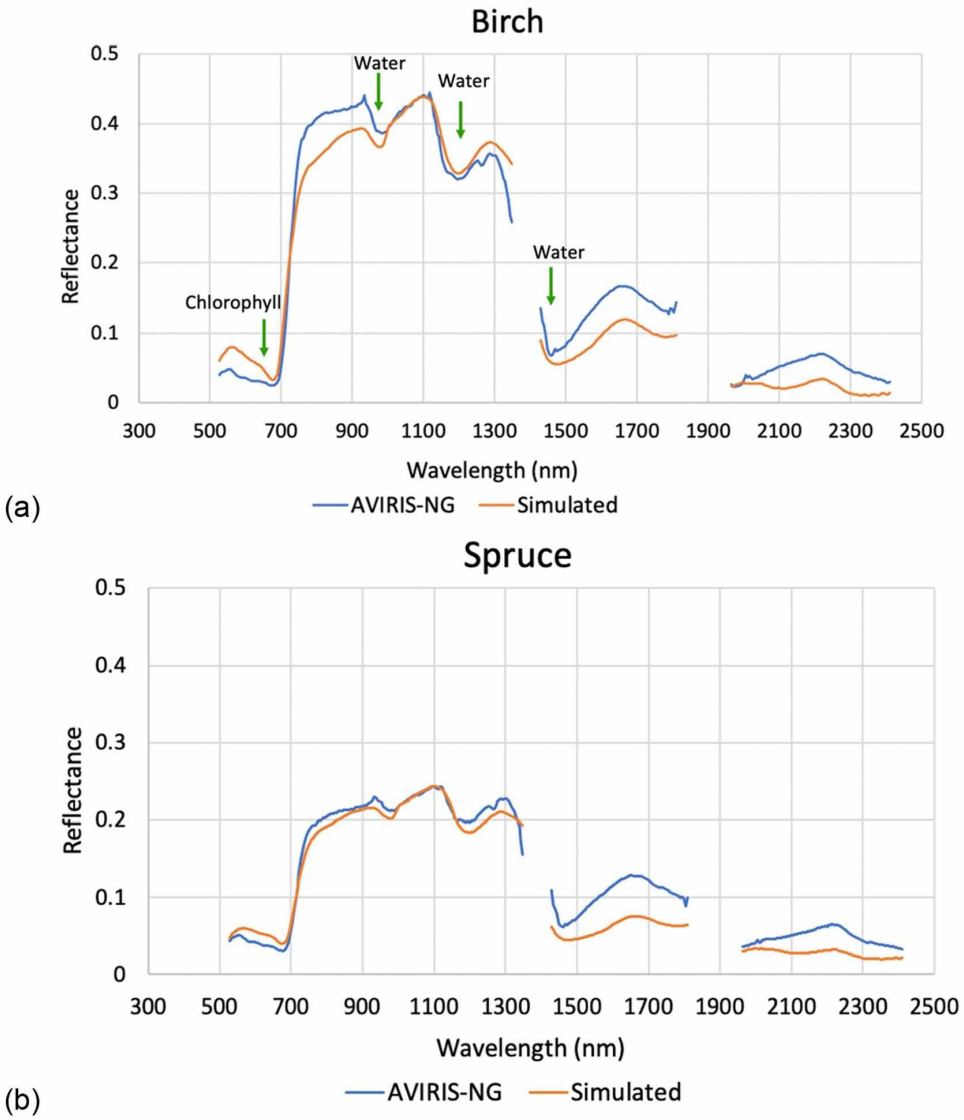


Figure 3.4: Spectral Signature of (a) Birch and (b) Spruce.

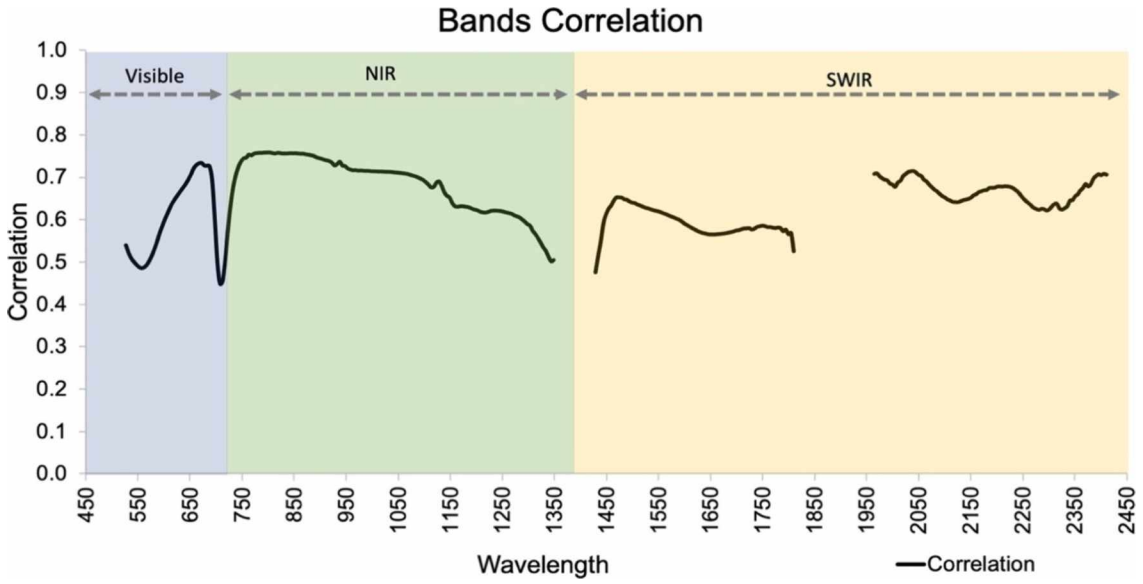


Figure 3.5: Band-to-band correlation between AVIRIS-NG and simulated hyperspectral images.

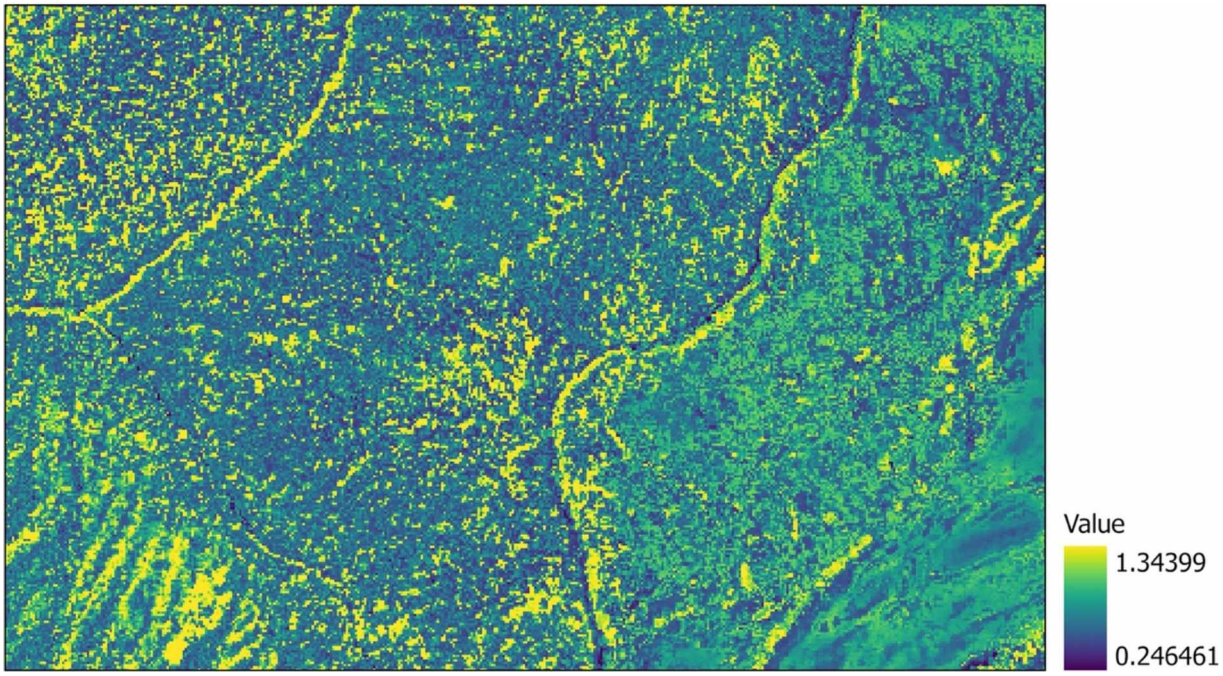


Figure 3.6: A map of the Coefficient of Variation (CoV) for the pixel difference between AVIRIS-NG and simulated hyperspectral data.

3.4.2. Visual analysis

The simulated image captures the minute details visible in the AVIRIS-NG image (Figure 3.7). AVIRIS-NG data were not available for the bottom left side of the area in Figure 3.7 (a) and (b), therefore the region is black. Figure 3.7 (a) is an image from Creamer's Field - Migratory Waterfowl Refuge area in Fairbanks. The region shown in a yellow circle is an open field that has

trees planted on the boundary. Simulated data captured the trees on the edges quite well. In Figure 3.7 (b), the small patches inside the yellow circle are water bodies. These water bodies are accurately captured in the simulated image. In Figure 3.7 (c), we overlaid ground surveyed points for birch and spruce on the three images. We can see a clear difference in the image tone and color saturation between the coniferous vegetation (spruce) and deciduous vegetation (birch) in all three images. The area dominated by birch has brighter pixels, while the area with spruce has darker pixels. The simulated image correctly captured this tonal and color differences between birch and spruce forests.

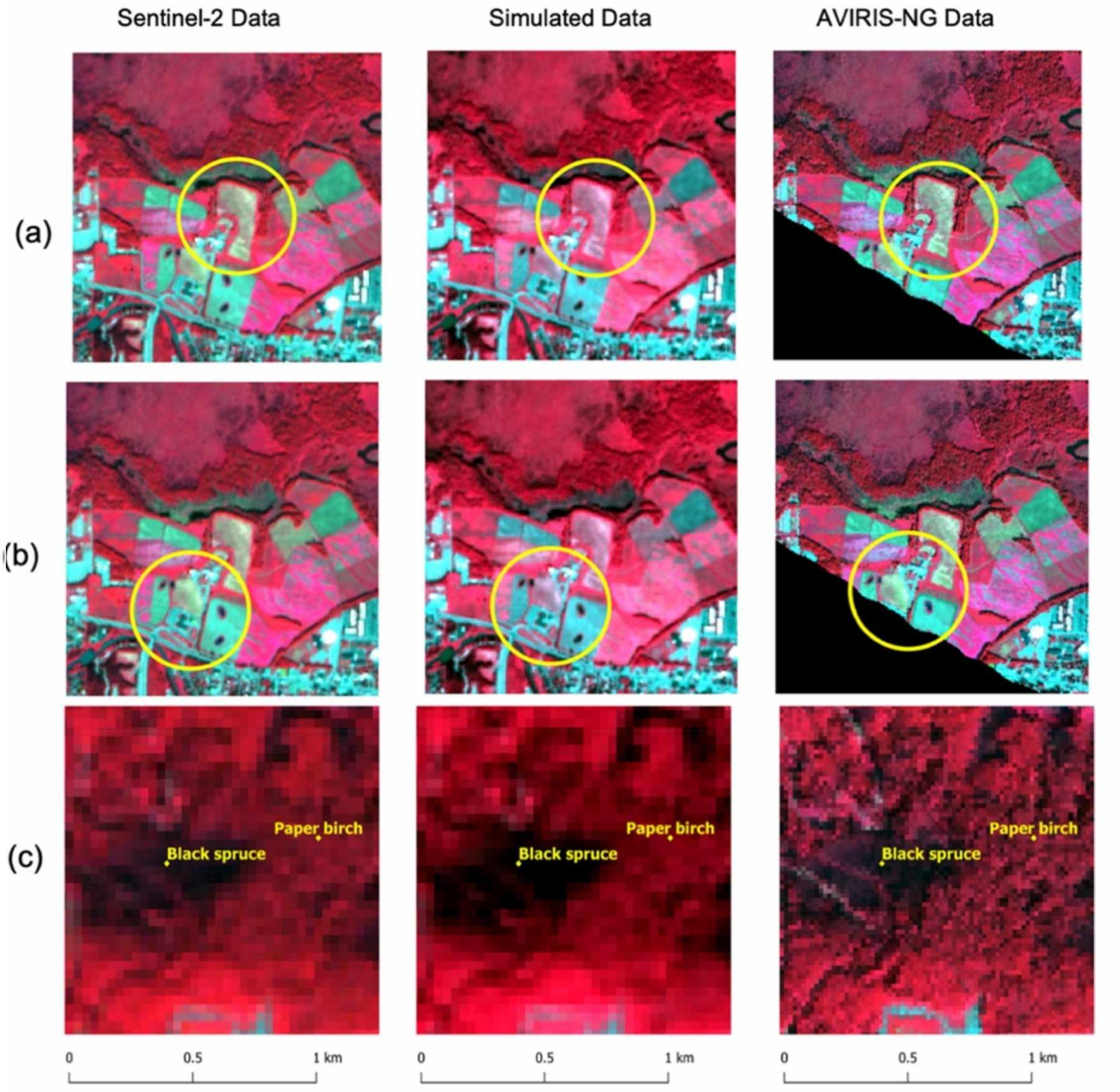


Figure 3.7: Visual comparison of the simulated Color Infra-Red image with AVIRIS-NG and Sentinel 2 for different areas: (a) open field (b) water bodies (c) Black spruce and paper birch.

3.4.3. Image classification

We performed vegetation classification on the simulated data according to the Alaska vegetation classification given by Viereck et al. (1992). Figure 3.8 highlights the results of vegetation mapping using the RF classifier. We used NDVI to mask out non-vegetation pixels. There is a separate class for the masked pixels in the legend. Mountains dominate the southern part of the scene; some parts of the mountain and high elevation areas are covered with alpine vegetation. We did not have any training data for alpine vegetation class. Therefore, we included a "Other" class to capture the alpine vegetation and other vegetation (mostly grasses) pixels that are not included in the training. The center of the scene is a lowland area. Most of the pixels are classified into vegetation classes dominated by black spruce.

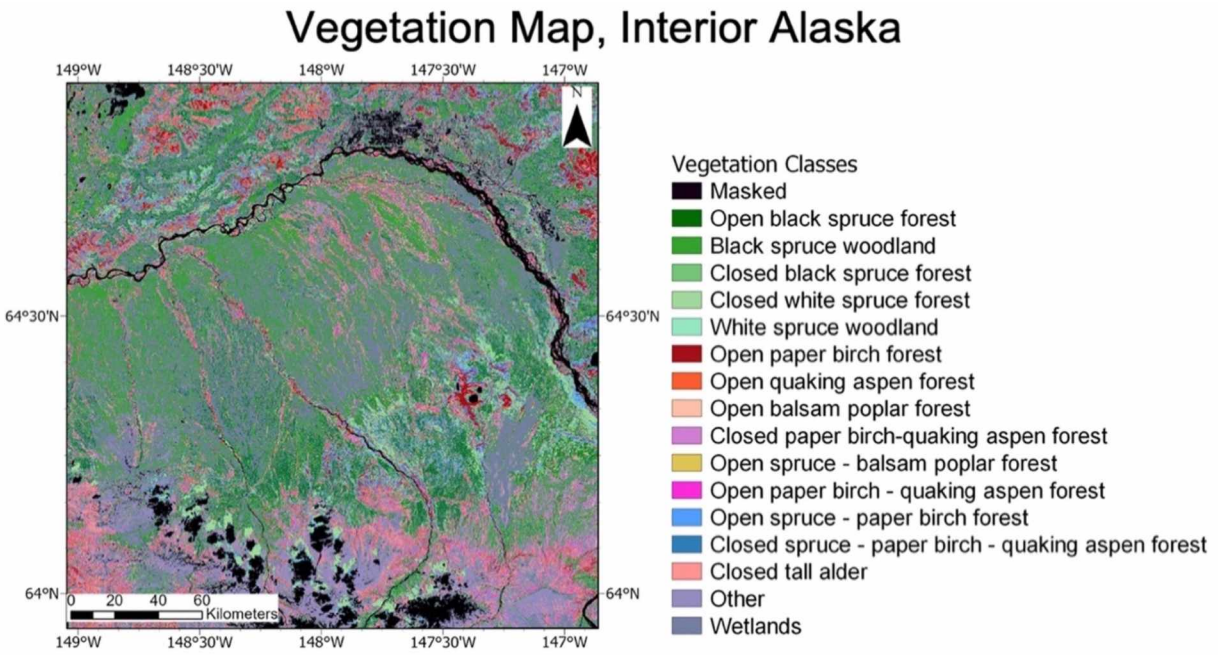


Figure 3.8: Vegetation map generated from simulated hyperspectral image (covering an area of 100 km × 100 km) using Random Forest classifier. The masked pixels shown in black include urban areas, water bodies, clouds and cloud shadows.

We had unbalanced test samples; different classes have different numbers of pixels. To overcome class imbalance, we calculated the F1 score and IoU shown in Figure 3.9 and Figure 3.10 respectively. A high F1 score means the class has performed well. "Open black spruce forest" and "Open balsam poplar forest" had the lowest F1 score, while most classes performed well. Additionally, if the IoU value is 0.5, the class has performed well. In our case, seven classes had an IoU value greater than 0.5, 3 classes had an IoU value very close to 0.5, and the remaining classes had a value between 0.3 and 0.4.

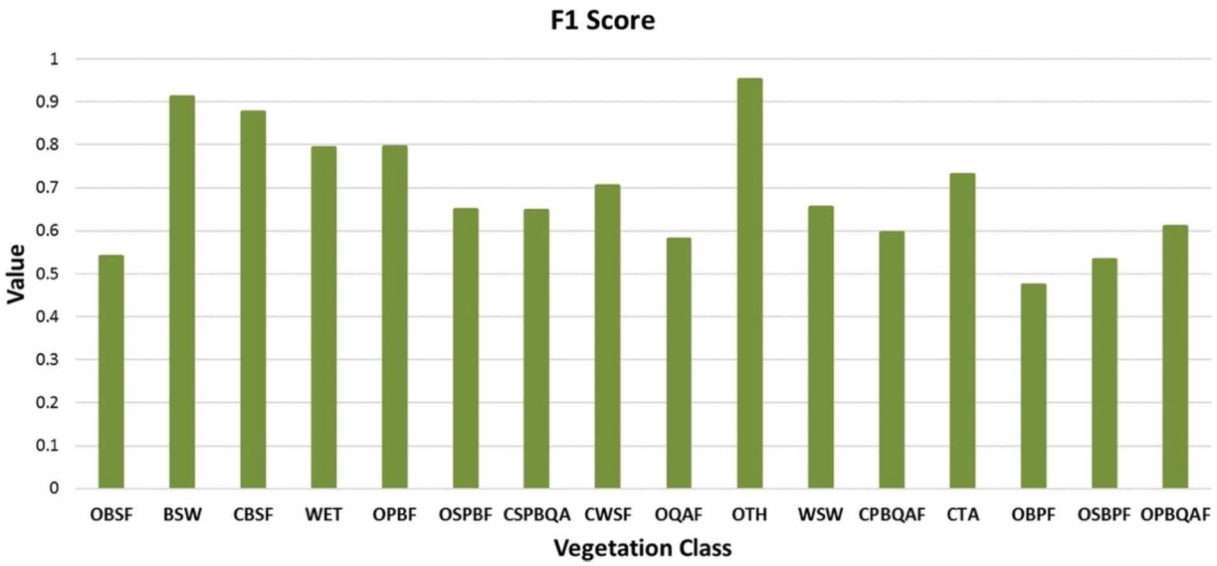


Figure 3.9: F1-score for each class.

Open black spruce forest: OBSF; Black spruce woodland: BSW; Closed black spruce forest: CBSF; Wetlands: WET; Open paper birch forest: OPBF; Open spruce - paper birch forest: OSPBF; Closed spruce - paper birch- quaking aspen forest: CSPBQA; Closed white spruce forest: CWSF; Open quaking aspen forest: OQAF; Other: OTH; White spruce woodland: WSW; Closed paper birch-quaking aspen forest: CPBQAF; Closed tall alder: CTA; Open balsam poplar forest: OBPF; Open spruce - balsam poplar forest: OSBPF; Open paper birch - quaking aspen forest: OPBQAF.

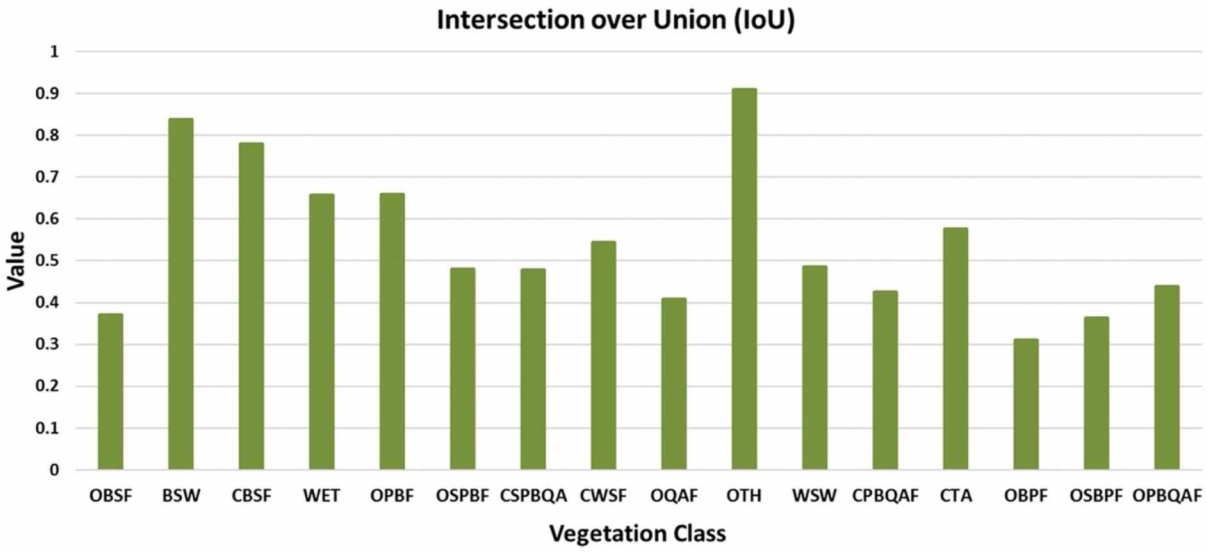


Figure 3.10: IoU score for each class.

For a test site in BCEF, we generated a vegetation map from Sentinel-2 data and assessed its accuracy using field validation plots. We did not use these plots for training. Out of 31 plots (Figure 3.11), 13 plots were correctly mapped with 42% accuracy in the case of Sentinel-

2 data, while 20 plots were correctly mapped in the case of simulated classified output, resulting in an accuracy of 65%.

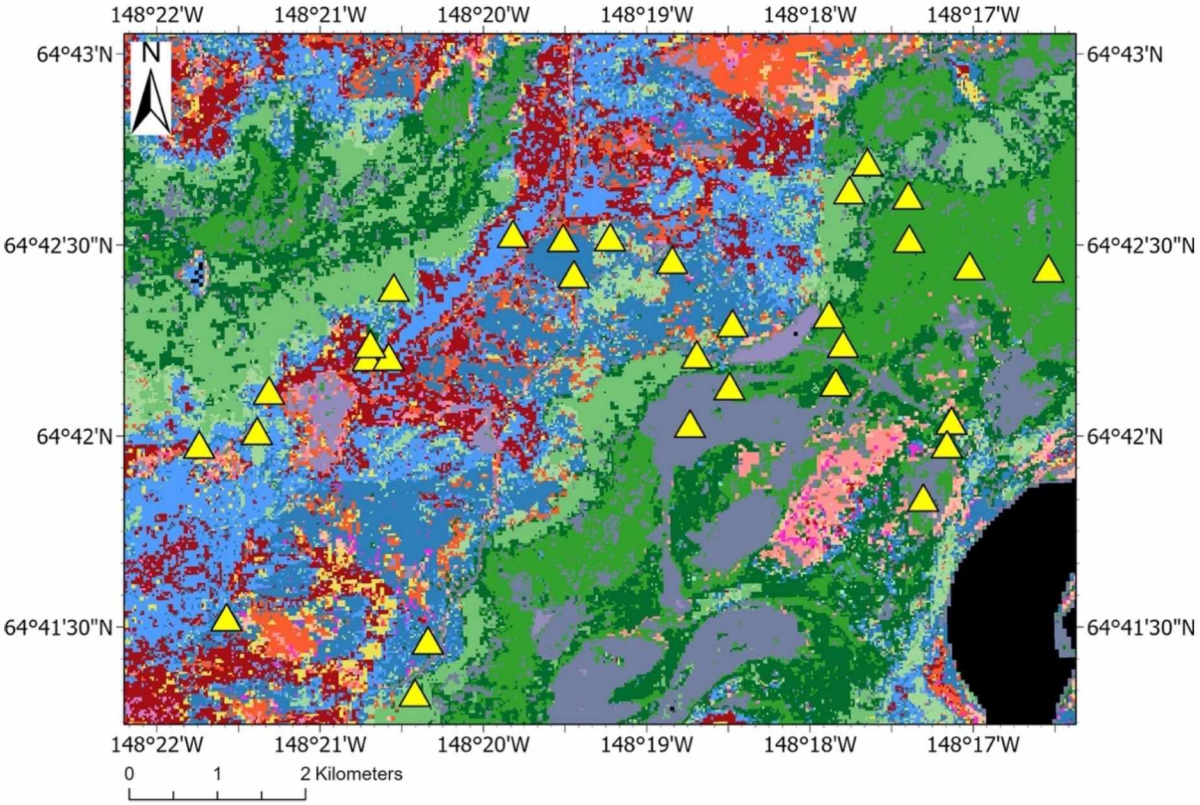


Figure 3.11: A vegetation map derived from the Sentinel-2 image for a test site within the Bonanza Creek Experimental Forest (BCEF); yellow triangles mark the field surveyed plot locations used as validation.

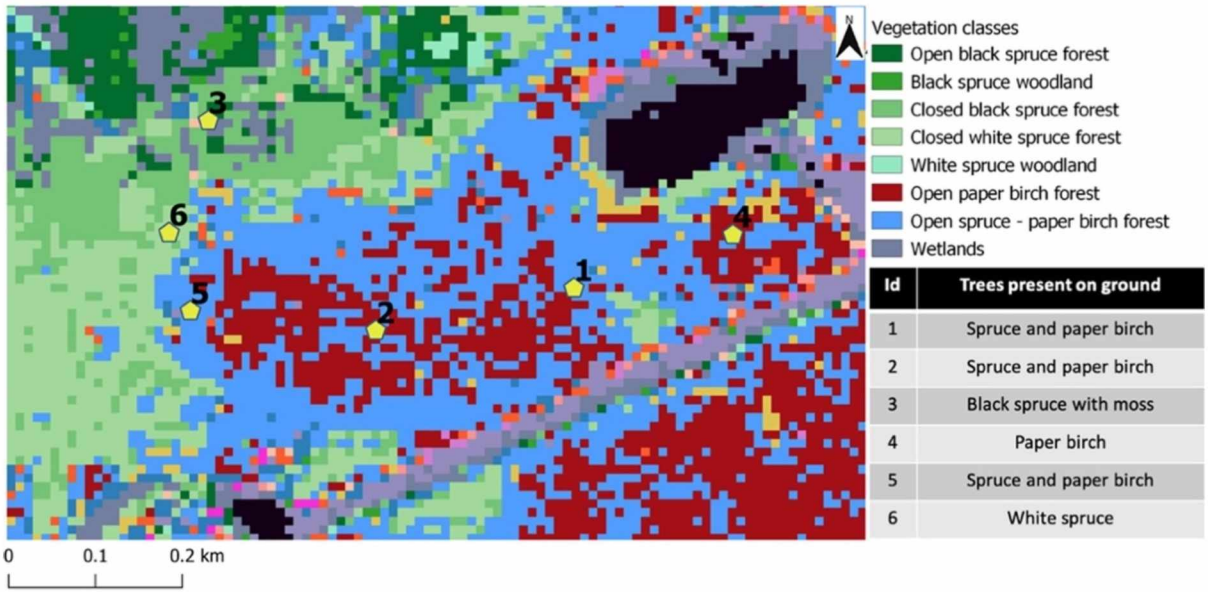
3.4.4. Process validation

To test the simulation method across space, i.e. for different geographic subregions, we simulated two more scenes (S2BT06WWS, S2BT06WWU) using the same ground spectra and classifier model (trained in scene 1) to predict pixel vegetation class. Scene 2 (S2BT06WWS) covers the area east of scene 1, and scene 3 (S2BT06WWU) covers a part of the Yukon flats region that lies to the north of Fairbanks. Locations of all three scenes are shown in Figure 3.3. Scene 2 (S2BT06WWS) falls under the Yukon-Tanana Uplands sub-ecoregion, and scene 3 (S2BT06WWU) covers the Yukon-Old Crow Basin sub-ecoregion. We had a smaller set of ground data for scene 2 that we used for validating our classified vegetation map output. We did not have any ground data for scene 3, so we only performed visual, statistical, and spectral comparison of the simulated image generated from scene 3.

We compared spectra collected from simulated image data and AVIRIS-NG image for scene 2 and scene 3. Figure 3.19 shows the comparison of the spectra collected from the same pixels in both images. We picked two different sites that represent deciduous and coniferous

species. The spectra showed a similar pattern and the same absorption features. The RMSE for deciduous vegetation (Figure 3.19 (a)) was 0.02, and 0.01 for coniferous vegetation (Figure 3.19 (b)). We visually compared the simulated data and AVIRIS-NG data. Simulated data has captured the trails and built-up area similar to AVIRIS-NG data (see the yellow circle, Figure 3.20 (a)). Coniferous and deciduous vegetation are easily distinguishable in simulated data (see the yellow circle, Figure 3.20 (b)).

Here we have discussed the classified vegetation map accuracy at select points for an area around Twin Bear Chena River to Ridge and Compeau trail east of Fairbanks. We overlaid the ground points over the classified map and visually assessed the classified map product accuracy (Figure 3.12). In Figure 3.12, points 1, 2, and 5 (from spruce and birch vegetation) are over the blue pixels; the blue pixels in the classified map represent open spruce-paper birch class. Point 4 has paper birch in the ground; the classifier mapped it as the open paper birch forest. Point 6 has white spruce, and the pixels around this point are mapped as the closed white spruce forest. Point 3 has black spruce and moss, and the classified map identified the pixels at and around this point as open black spruce forest and black spruce woodland. The dark gray pixels are wetlands and include all the vegetation growing on marshy areas with seasonal standing water. The gray pixels are mostly around the water bodies (shown as black pixels). Wetlands and vegetation classes present at all six points are correctly mapped in the classified vegetation map.



Id	Trees present on ground
1	Spruce and paper birch
2	Spruce and paper birch
3	Black spruce with moss
4	Paper birch
5	Spruce and paper birch
6	White spruce

Figure 3.12: A vegetation map derived from simulated hyperspectral image for an area around Twin Bear Chena River ridge and Compeau trail; ground data points are presented as yellow pentagon. Tree cover present at the ground point is listed in the table (bottom right).

Using our algorithms, one can simulate hyperspectral data and classify it to generate a vegetation map in 10 h (system configuration: Intel(R) Xeon(R) Gold 5222 CPU @ 3.80 GHz, 3801 MHz, 4 cores with 192 GB RAM). Each simulated hyperspectral image requires a minimum of 190 GB disk space (scene size: 100 X 100 km; pixel size: 10 m; 425 bands). To overcome the dependency on powerful computer systems with large storage, we implemented the entire processing in the Google Cloud Platform (GCP system configuration: 16 CPUs and 60 GB RAM).

3.5. Discussion

We developed a novel approach to efficiently simulate an AVIRIS-NG hyperspectral image cube from a Sentinel-2 image and subsequently derive a detailed vegetation map covering an area of 100 km × 100 km for the boreal region of Alaska. The simulated images accurately capture the minute landscape features present in the original AVIRIS-NG image. The spectral profile from the simulated image matches the original AVIRIS-NG image in the pattern but differs in reflectance magnitude in the SWIR region. Zhang et al., 2006, Badola et al., 2021 also reported a similar difference in reflectance over the SWIR region. Several factors contribute to the spectral differences: a) the difference in the spatial resolution of the two datasets, i.e. Sentinel-2 bands have a spatial resolution of 10 m and 20 m compared to 5 m for AVIRIS-NG, b) the difference in the sensor's altitude and sun and sensor geometry, and c) difference in date and time of acquisition. Also, the missing endmember, Non-Photosynthetic Vegetation (NPV), is contributing to the reflectance difference. NPV has higher SWIR reflectance (Asner, 1998), and depending on the NPV, a vegetation pixel can have lower reflectance in the visible region than in the SWIR region. However, when we added NPV to the model, it reduced model performance significantly. Hence, one of the major limitations of this model is that it cannot handle NPV as one of the endmembers without generating significant errors. However, by excluding NPV, we ended up with lower reflectance in the SWIR region. In reality, a pixel can have a significant NPV fraction, but this model can only handle distinct endmembers. Also, this model is built for the peak of the growing season and might not do well for images acquired during shoulder season; for example, when leaves start to senescence at the beginning of the fall season. Despite the above limitation, our simulation model can accurately capture the key absorption features and spectral details essential for detailed vegetation and landcover mapping, paving the path for on-demand hyperspectral data availability for boreal Alaska.

3.5.1. Comparison with other map products

We used 31 field surveyed plots in the Bonanza Creek Experimental Forest site, a Long-Term Ecological Research (BCEF LTER) site in Interior Alaska, to assess the accuracy of

LANDFIRE EVT and our vegetation map (Figure 3.13). Out of 31 plots, 10 plots were mapped correctly in the LANDFIRE EVT 2016 product, giving a product accuracy of 32%, whereas 20 plots were mapped correctly in our vegetation map product, giving a product accuracy of 65%. Table 3.5 shows the vegetation class present on the 11 plots that were misclassified. The majority of these misclassified plots have similar vegetation. Two plots of closed spruce-paper birch-quaking aspen were incorrectly mapped as open spruce-paper birch, while two other plots were mapped as open quaking aspen. One wetland plot was incorrectly mapped as Black spruce woodland, possibly due to the presence of isolated black spruce and understory vegetation found in black spruce woodland. We believe by merging similar classes and retraining the classifier, the map accuracy can be further improved.

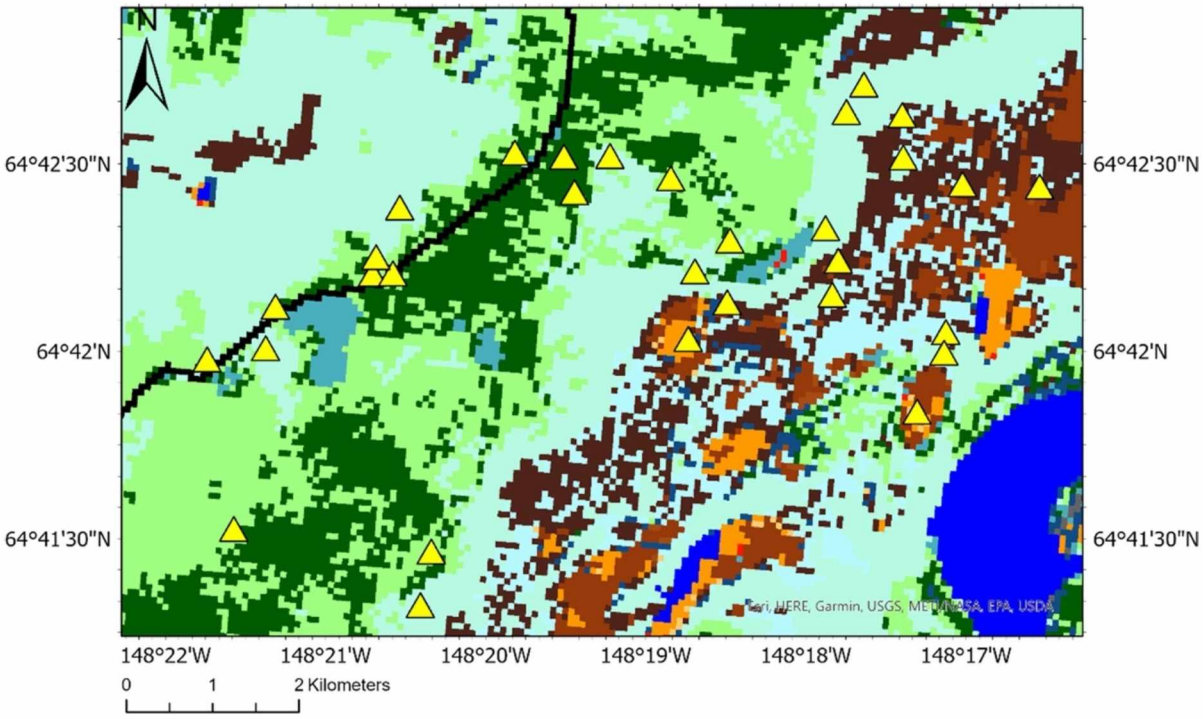


Figure 3.13: LANDFIRE EVT Product for the Bonanza Creek Experimental Forest (BCEF) site; yellow triangles mark the field surveyed plot locations.

Figure 3.14 demonstrates the comparison between our vegetation map and the available LANDFIRE EVT and Alaska Vegetation and Wetland Composite (AKVWC) maps. The rectangle, square, and circle in Figure 3.14 highlights the locations where we clearly see that the map generated using simulated data is more detailed and accurate in comparison to the other two maps. The region highlighted under the rectangle is dominated by single birch and aspen class in both AKVWC and LANDFIRE EVT maps, while four different birch and aspen vegetation

classes are mapped in the simulated classified map: open paper birch forest, open quaking aspen forest, open and closed paper birch, and quaking aspen forest. Similarly, the region inside the circle is dominated by black spruce and white spruce classes in all three cases, but in the case of a simulated classified map, the classes are more detailed. The pixels inside the square are mapped as wetlands in the case of AKVWC and simulated classified maps, while in the LANDFIRE EVT map, most of these pixels are mapped as black spruce-tamarack fen.

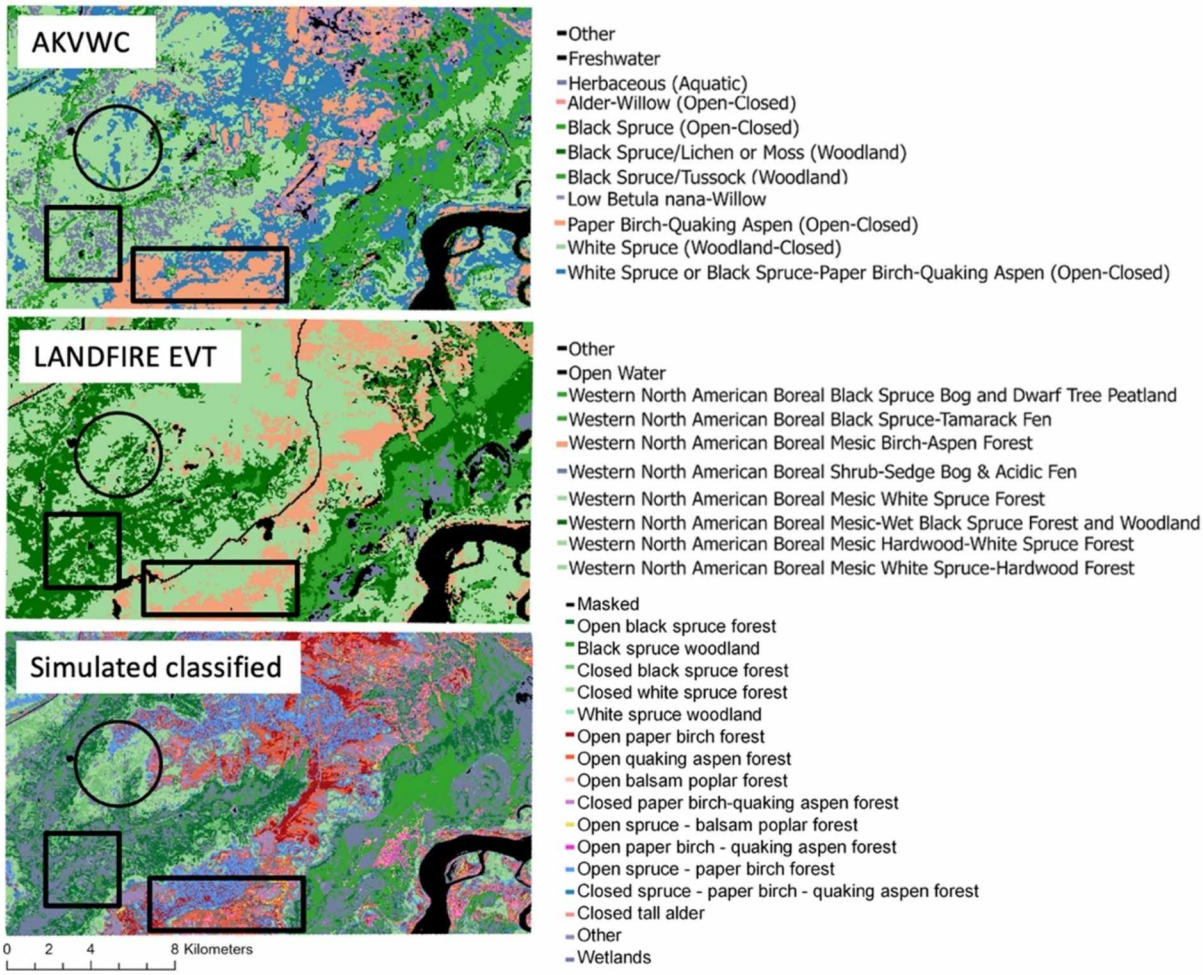


Figure 3.14: Comparison of simulated classified map with Alaska Vegetation and Wetland Composite (AKVWC) and LANDFIRE EVT Product.

The U.S. Department of Agriculture completed a Pilot Inventory for Tanana Valley State Forest, which is covered by the Sentinel scene used in this study. The survey for the Pilot Inventory began in 2014 and was fully implemented between 2016 and 2018. They surveyed 800 plots (around a 7.3 m radius) to get forest type acreage, extrapolating the observations to the entire Tanana Valley State Forest to get total acreage for forest types (Pattison et al., 2018, U.S. Forest Service, 2016). Figure 3.15 shows the comparison of forest type acreage estimated from

the simulated classified map with estimates from the Pilot Inventory program for five major tree species. For all five tree species, the acreage estimates from our vegetation map are comparable with the estimates from the Pilot Inventory program, and they follow the same pattern as well: i.e. black spruce is the most dominant tree species, followed by birch, with aspen and poplar being the least dominant tree species. Therefore, the forest type acreage we obtained from our map is comparable to the USDA pilot inventory forest acreage.

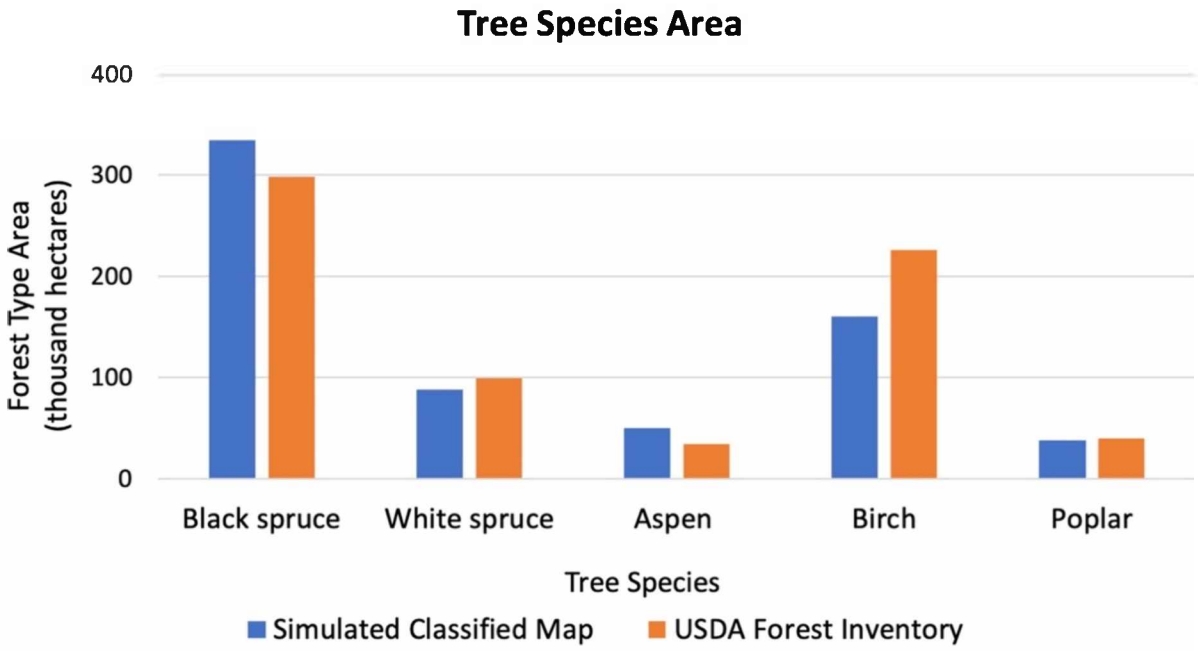


Figure 3.15: Comparison of area covered by different forest type reported by the USDA Tanana valley Pilot Inventory program and estimated from our classified map product.

We used a smaller set of ground data over scene 2 (covering the area east of scene 1) to assess the quality and accuracy of our vegetation map. These points were not used for training the classifier. Figure 3.16 compares the simulated classified map with LANDFIRE EVT and Alaska Vegetation and Wetland Composite (AKVWC) maps. In the LANDFIRE EVT map, the black spruce point is sitting on a pixel mapped as Western North American Boreal Mesic White Spruce-Hardwood Forest (bright green pixel); in the AKVWC product, the black spruce point is sitting on White Spruce or Black Spruce (Open) (purple pixels). In the simulated classified map, the black spruce point is correctly mapped as open black spruce forest (dark green pixel). For the wetland point, the LANDFIRE EVT product mapped the pixel as Western North American Boreal Black Spruce Bog and Dwarf-Tree Peatland (brown pixel) and AKVWC product mapped the pixel in the same way it mapped for black spruce, i.e. White Spruce or Black Spruce (Open) (purple pixel), while the simulated classified product mapped the area as wetlands (grey pixels). These

observations suggest that the simulated classified map correctly identified the open black spruce forest and wetlands pixels with better accuracy than the two existing maps.

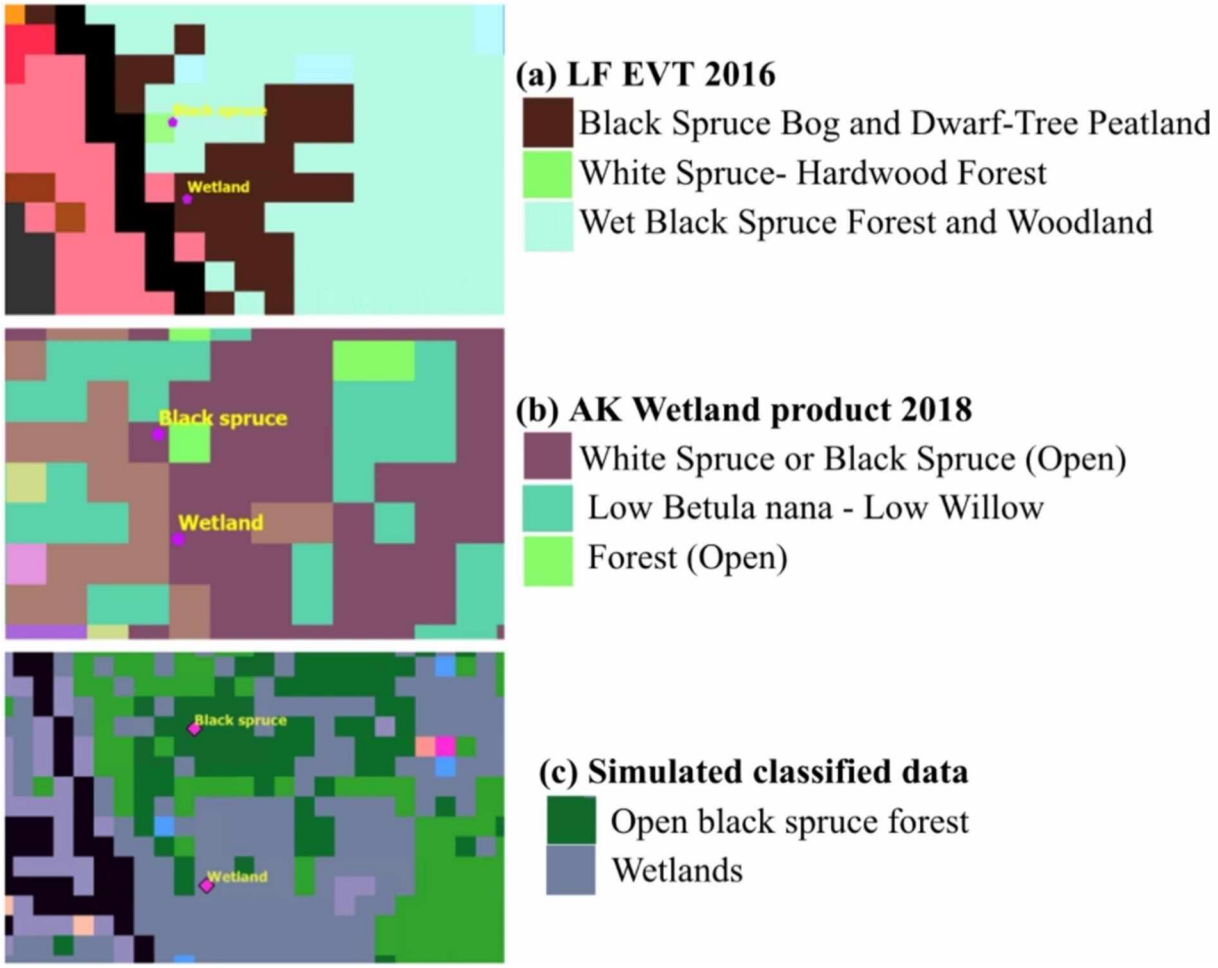


Figure 3.16: Comparison of classified map products; ground data points are presented in magenta color. Tree cover present at the ground point is shown on the image in yellow text (a) LANDFIRE EVT Product (b) Alaska Vegetation and Wetland Composite (AKVWC) (c) Simulated classified output.

We simulated a third scene (scene 3: S2BT06WWU) from a different ecoregion that covers a part of the Yukon Flats region. Figure 3.21 shows the spectral comparison between simulated data and AVIRIS-NG data. The simulated data accurately captured the absorption features and has a similar reflectance as AVIRIS-NG data: RMSE = 0.02 for deciduous vegetation (Figure 3.5 (a)) and RMSE = 0.01 for coniferous species (Figure 3.21 (b)).

The hyperspectral simulation of three different Sentinel 2 scenes and their evaluation demonstrated that we developed an efficient approach for obtaining on-demand hyperspectral data for vegetation/fuel mapping. Our vegetation maps are more detailed and their accuracies are on par with or better than the existing vegetation maps. To effectively implement this approach, it

is important to collect field spectra on a clear, sunny day. The field spectra collected on hazy days or cloudy days can degrade hyperspectral simulation results. We also learned that to process the entire Sentinel scene, one needs an expensive system with higher processing power and hardware, so a cloud computing platform like Google Cloud Platform (GCP) (Google, 2022) is a better alternative for efficient and cost-effective processing. GCP reduces the dependency on an expensive local system and the installation of libraries. The only requirements are a browser, a good internet connection, and a Google account to run the codes. More importantly, GCP makes it easy to share the codes with the research community. A simulated hyperspectral image cube covering an area of 100 km × 100 km requires 190 GB of disk space. GCP provides a considerable amount of storing capacity at a low cost to store such huge files and their efficient access for further processing. The hyperspectral nature of this data (5 nm bandwidth) makes it appropriate for generating a detailed vegetation map with improved accuracy for a variety of applications, including fire and forest management. Better vegetation/fuel maps are critical for effective fuel treatment, i.e. identifying areas for creating fire breaks or fire lines to check the spread of wildfires. The simulated hyperspectral image can potentially be used to extract other biophysical attributes of vegetation like chlorophyll, moisture, and nitrogen, expanding its applications to other areas of vegetation research.

3.6. Conclusions

In this study we developed and implemented a sophisticated workflow to generate simulated AVIRIS-NG hyperspectral image data from Sentinel-2 image at Sentinel scene scale (100 X100 km) in Google's cloud environment and tested simulation reproducibility across space and time. We employed a well-established endmember selection technique to make our method more consistent and reproducible across different geographies. By improving the simulation algorithm, we were able to turn a 4.5 GB Sentinel 2 data set into a full hyperspectral image cube of 190 GB in about 2 hrs. We developed a Random Forest image classification model using training signatures from one scene and tested the Random Forest model portability on two other scenes from different sub-ecoregions. The Random Forest vegetation classification model performed satisfactorily on these two scenes, suggesting that our model can be reliably used for improved vegetation mapping in boreal Alaska. The vegetation maps generated from simulated data were more detailed and accurate based on the available field data points than the two existing vegetation maps, LANDFIRE EVT and Alaska Vegetation and Wetland Composite (AKVWC). We assessed the mapping accuracy of the latest LANDFIRE EVT (32%) and our map product (65%) using ground survey data and observed an improvement of 33% in accuracy. This study developed a novel approach and algorithms to produce hyperspectral image from widely

available multispectral images and showed its applicability in mapping vegetation and fuel in the boreal forest of Alaska with improved accuracy, which will contribute to effective forest and fire management in Alaska.

3.7. Acknowledgments

A heartfelt thanks to Chris Smith, Malvika Shriwas and Brooke Kubby for assisting in fieldwork. Special thanks to Abhisek Maiti for his support with the programming part of this work. Thanks to the UAF Writing Center, especially Delcenia Cosman, for proofreading this manuscript and correcting grammatical errors. Also, we express our gratitude to the two anonymous reviewers and journal editor for their constructive feedback and suggestions that helped us improve this manuscript.

Funding: This material is based upon work supported by the National Science Foundation and under the award OIA-1757348, the State of Alaska and the U.S. Geological Survey under Grant/Cooperative agreement No. G18AP00077.

3.8. References

Abrams, M., Crippen, R., Fujisada, H., 2020. ASTER Global Digital Elevation Model (GDEM) and ASTER Global Water Body Dataset (ASTWBD). Remote Sens. 2020, Vol. 12, Page 1156 12, 1156. https://doi.org/10.3390/RS12071156.

Ahmad, S., Pandey, A.C., Kumar, A., Lele, N.V., 2021. Potential of hyperspectral AVIRISNG data for vegetation characterization, species spectral separability, and mapping. Appl. Geomatics 13 (3), 361–372.

Alaska Vegetation and Wetland Composite, 2019. Alaska Vegetation and Wetland Composite | Alaska Conservation Science Catalog [WWW Document]. URL https:// accscatalog.uaa.alaska.edu/dataset/alaska-vegetation-and-wetland-composite (accessed 11.4.21). Asner, G.P., 1998. Biophysical and Biochemical Sources of Variability in Canopy Reflectance. Remote Sens. Environ. 64, 234–253. https://doi.org/10.1016/S0034-4257(98)00014-5.

Badola, A., Padalia, H., Belgiu, M., Prabhakar, M., Verma, A., 2019. Mapping Tree Species Richness of Tropical Forest using Airborne Hyperspectral Remote Sensing.

Badola, A., Padalia, H., Belgiu, M., Verma, P.A., 2021a. Tree Species Mapping in Tropical Forests Using Hyperspectral Remote Sensing and Machine Learning 5421–5424. https://doi.org/10.1109/IGARSS47720.2021.9553549.

Badola, A., Panda, S.K., Roberts, D.A., Waigl, C.F., Bhatt, U.S., Smith, C.W., Jandt, R.R., 2021b. Hyperspectral Data Simulation (Sentinel-2 to AVIRIS-NG) for Improved Wildfire Fuel Mapping, Boreal Alaska. Remote Sens. 2021, Vol. 13, Page 1693 13, 1693. https://doi.org/10.3390/RS13091693.

Breiman, L., 2001. Random Forests. Mach. Learn. 45, 5–32. https://doi.org/10.1023/A:1010933404324.

Clark, M.L., Roberts, D.A., Clark, D., 2005. Hyperspectral discrimination of tropical rain forest tree species at leaf to crown scales. Remote Sens. Environ. 96, 375–398. https://doi.org/10.1016/j.rse.2005.03.009.

Develice, R.L., 2012. Accuracy of the LANDFIRE Alaska Existing Vegetation Map over the Chugach National Forest.

Dobrini´c, D., Ga˘sparovi´c, M., Medak, D., 2021. Sentinel-1 and 2 Time-Series for Vegetation Mapping Using Random Forest Classification: A Case Study of Northern Croatia. Remote Sens. 2021, Vol. 13, Page 2321 13, 2321. https://doi.org/10.3390/ RS13122321.

Dudley, K.L., Dennison, P.E., Roth, K.L., Roberts, D.A., Coates, A.R., 2015. A multitemporal spectral library approach for mapping vegetation species across spatial and temporal phenological gradients. Remote Sens. Environ. 167, 121–134. https://doi.org/10.1016/J.RSE.2015.05.004.

EarthData, 2021. Earthdata [WWW Document]. URL https://earthdata.nasa.gov/ (accessed 11.3.21).

European Space Agency, 2017. Sentinel-2 Spectral Response Functions (S2-SRF) - Sentinel-2 MSI Document Library - User Guides - Sentinel Online [WWW Document]. URL https://sentinel.esa.int/web/sentinel/user-guides/sentinel-2-msi/documentlibrary/-/asset_publisher/Wk0TKajiISaR/content/sentinel-2a-spectral-responses (accessed 11.23.20).

European Space Agency, 2014. Copernicus Open Access Hub [WWW Document]. URL (accessed 11.23.20).

Exelis Visual Information Solutions version 5.3, 2010. Exelis Visual Information Solutions.

Fernandez-Manso, A., Quintano, C., Roberts, D.A., 2016. Burn severity influence on postfire vegetation cover resilience from Landsat MESMA fraction images time series in Mediterranean forest ecosystems. Remote Sens. Environ. 184, 112–123. https://doi.org/10.1016/J.RSE.2016.06.015.

Gillies, S., others, 2013. Rasterio: geospatial raster I/O for {Python} programmers.

Google, 2022. Google Cloud documentation | Documentation [WWW Document]. URL https://cloud.google.com/docs (accessed 1.2.22).

Govender, M., Chetty, K., Naiken, V., Bulcock, H., 2019. A comparison of satellite hyperspectral and multispectral remote sensing imagery for improved classification and mapping of vegetation. Water SA 34, 147–154. https://doi.org/10.4314/wsa. v34i2.

Grabska, E., Hostert, P., Pflugmacher, D., Ostapowicz, K., 2019. Forest Stand Species Mapping Using the Sentinel-2 Time Series. Remote Sens. 11, 1197. https://doi.org/ 10.3390/rs11101197.

Harris, C.R., Millman, K.J., van der Walt, S.J., Gommers, R., Virtanen, P., Cournapeau, D., Wieser, E., Taylor, J., Berg, S., Smith, N.J., Kern, R., Picus, M., Hoyer, S., van Kerkwijk, M.H., Brett, M., Haldane, A., del Río, J.F., Wiebe, M., Peterson, P., G´erard-Marchant, P., Sheppard, K., Reddy, T., Weckesser, W., Abbasi, H., Gohlke, C., Oliphant, T.E., 2020. Array programming with {NumPy}. Nature 585 (7825), 357–362.

Hati, J.P., Goswami, S., Samanta, S., Pramanick, N., Majumdar, S.D., Chaube, N.R., Misra, A., Hazra, S., 2021. Estimation of vegetation stress in the mangrove forest using AVIRIS-NG airborne hyperspectral data. Model. Earth Syst. Environ. 7 (3), 1877–1889.

Hsu, P.H., 2007. Feature extraction of hyperspectral images using wavelet and matching pursuit. ISPRS Journal of Photogrammetry and Remote Sensing 62 (2), 78–92. https://doi.org/10.1016/J.ISPRSJPRS.2006.12.004.

International Arctic Research Center, 2021. Alaska's Changing Wildfire Environment | IARC [WWW Document]. URL https://uaf-iarc.org/alaskas-changing-wildfireenvironment/. (accessed 11.9.21).

Landfire, 2016. LANDFIRE: Existing Vegetation Type [WWW Document]. U.S. Dep. Agric. U.S. Dep, Inter http://www.landfire.gov. (accessed 2.10.21).

Liu, B., Zhang, L., Zhang, X., Zhang, B., Tong, Q., 2009. Simulation of EO-1 Hyperion Data from ALI Multispectral Data Based on the Spectral Reconstruction Approach. Sensors 9, 3090–3108. https://doi.org/10.3390/s90403090.

Louis, J., Debaecker, V., Pflug, B., Main-Knorn, M., Bieniarz, J., Mueller-Wilm, U., Cadau, E., Gascon, F., 2016. SENTINEL-2 SEN2COR: L2A Processor for Users.

McHugh, M.L., 2012. Interrater reliability: the kappa statistic. Biochem. Medica 22, 276. https://doi.org/10.11613/bm.2012.031.

Meerdink, S.K., Hook, S.J., Roberts, D.A., Abbott, E.A., 2019. The ECOSTRESS spectral library version 1.0. Remote Sens. Environ. 230, 111196.

Mudele, O., Gamba, P., 2019. Mapping vegetation in urban areas using Sentinel-2. 2019 Jt. Urban Remote Sens. Event, JURSE 2019. https://doi.org/10.1109/ JURSE.2019.8809019.

NASA JPL, 2018. AVIRIS-Next Generation [WWW Document]. URL https://avirisng.jpl. nasa.gov/platform.html (accessed 11.24.20).

National Park Service, 2021. Fire in Ecosystems: Boreal Forest (U.S. National Park Service) [WWW Document]. URL https://www.nps.gov/articles/000/fire-inecosystems-boreal-forest.htm (accessed 11.9.21).

Nowacki, G.J., Spencer, P., Fleming, M., Brock, T., Jorgenson, T., 2003. Unified Ecoregions of Alaska: 2001. Open-File Rep. https://doi.org/10.3133/OFR2002297.

Pattison, Robert; Andersen, Hans-Erik; Gray, Andrew; Schulz, Bethany; Smith, Robert J.; Jovan, Sarah, tech. coords., 2018. Forests of the Tanana Valley State Forest and Tetlin National Wildlife Refuge, Alaska: results of the 2014 pilot inventory. Gen. Tech. Rep. PNW-GTR-967. Portland, OR: U.S. Department of Agriculture, Forest Service, Pacific Northwest Research Station.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., Duchesnay, E., 2011. Scikit-learn: Machine Learning in {P}ython. J. Mach. Learn. Res. 12, 2825–2830.

Python Core Team, 2015. Python: A dynamic, open source programming language.

Reeves, M.C., Ryan, K.C., Rollins, M.G., Thompson, T.G., 2009. Spatial fuel data products of the LANDFIRE Project. Int. J. Wildland Fire 18 (3), 250.

Roberts, D. A., Halligan, K., Dennison, P., Dudley, K., Somers, B., Crabbe, A., 2018. Viper Tools User Manual, Version 2, 91 pp. [WWW Document].

Roberts, D.A., Dennison, P.E., Roth, K.L., Dudley, K., Hulley, G., 2015. Relationships between dominant plant species, fractional cover and Land Surface Temperature in a Mediterranean ecosystem. Remote Sens. Environ. 167, 152–167. https://doi.org/ 10.1016/J.RSE.2015.01.026.

Roberts, D.A., Gardner, M.E., Church, R., Ustin, S.L., Green, R.O., 1997. Optimum strategies for mapping vegetation using multiple-endmember spectral mixture models. https://doi.org/10.1117/12.278930 3118, 108-119. https://doi.org/ 10.1117/12.278930.

Roth, K.L., Dennison, P.E., Roberts, D.A., 2012. Comparing endmember selection techniques for accurate mapping of plant species and land cover using imaging spectrometer data. Remote Sens. Environ. 127, 139–152. https://doi.org/10.1016/J. RSE.2012.08.030.

Salas, E.A.L., Subburayalu, S.K., Slater, B., Zhao, K., Bhattacharya, B., Tripathy, R., Das, A., Nigam, R., Dave, R., Parekh, P., 2020. Mapping crop types in fragmented arable landscapes using AVIRIS-NG imagery and limited field data. Int. J. Image Data Fusion 11, 33–56. https://doi.org/10.1080/19479832.2019.1706646.

Schaaf, A.N., Dennison, P.E., Fryer, G.K., Roth, K.L., Roberts, D.A., 2011. Mapping Plant Functional Types at Multiple Spatial Resolutions Using Imaging Spectrometer Data. GIScience Remote Sens. 48, 324–344. https://doi.org/10.2747/1548- 1603.48.3.324.

Singh, P., Srivastava, P.K., Malhi, R.K.M., Chaudhary, S.K., Verrelst, J., Bhattacharya, B. K., Raghubanshi, A.S., 2021. Denoising AVIRIS-NG data for generation of new chlorophyll indices. IEEE Sens. J. 21 (5), 6982–6989.

Smith, C.W., Panda, S.K., Bhatt, U.S., Meyer, F.J., 2021. Improved Boreal Forest Wildfire Fuel Type Mapping in Interior Alaska using AVIRIS-NG Hyperspectral data. Remote Sens. 13, 897. https://doi.org/https://doi.org/10.3390/rs13050897.

Tiwari, V., Kumar, V., Pandey, K., Ranade, R., Agrawal, S., 2016. Simulation of the hyperspectral data using Multispectral data. In: in: International Geoscience and Remote Sensing Symposium (IGARSS). Institute of Electrical and Electronics Engineers Inc, pp. 6157–6160. https://doi.org/10.1109/IGARSS.2016.7730608.

U.S. Forest Service, 2016. PNW-FIA Interior Alaska Inventory | Pacific Northwest Research Station | PNW - US Forest Service [WWW Document]. URL https://www.fs.usda.gov/pnw/projects/pnw-fia-interior-alaska-inventory (accessed 12.12.21).

U.S. Geological Survey, 2001. BLM REA YKL 2011 Ecoregions of Alaska and Neighboring Territory [WWW Document]. URL http://agdc.usgs.gov/data/usgs/erosafo/ecoreg/ (accessed 1.1.22).

Viereck, L.A., Dyrness, C.T., Batten, A.R., Wenzlick, K.J., 1992. The Alaska vegetation classification. Gen. Tech. Rep. PNW-GTR-286. Portland, OR U.S. Dep. Agric. For. Serv. Pacific Northwest Res. Station. 278 p 286. https://doi.org/10.2737/PNW-GTR286.

Wagner, F.H., Ferreira, M.P., Sanchez, A., Hirye, M.C.M., Zortea, M., Gloor, E., Phillips, O.L., de Souza Filho, C.R., Shimabukuro, Y.E., Aragao, ˜ L.E.O.C., 2018. Individual tree crown delineation in a highly diverse tropical forest using very high resolution satellite images. ISPRS J. Photogramm. Remote Sens. 145, 362–377. https://doi.org/10.1016/j.isprsjprs.2018.09.013.

Xie, Y., Sha, Z., Yu, M., 2008. Remote sensing imagery in vegetation mapping: a review. J. Plant Ecol. 1, 9–23. https://doi.org/10.1093/jpe/rtm005.

Zhang, C., 2014. Combining hyperspectral and lidar data for vegetation mapping in the Florida everglades. Photogramm. Eng. Remote Sensing 80, 733–743. https://doi.org/10.14358/PERS.80.8.733.

Zhang, L., Fujiwara, N., Furumi, S., Muramatsu, K., Daigo, M., Zhang, L., 2007. Assessment of the universal pattern decomposition method using MODIS and ETM data. Int. J. Remote Sens. 28, 125–142. https://doi.org/10.1080/ 01431160600617228.

Zhang, L., Furumi, S., Muramatsu, K., Fujiwara, N., Daigo, M., Zhang, L., 2006. Sensorindependent analysis method for hyperspectral data based on the pattern decomposition method. Int. J. Remote Sens. 27, 4899–4910. https://doi.org/10.1080/01431160600702640.

Appendix

Table 3.4: Endmembers used in each iteration.

Iterations	Endmembers
1st iteration	All 15 endmembers
2nd iteration	Alder, Downed trunk, blueberry, black spruce, larch, birch, gravel, white spruce, asphalt
3rd iteration	Downed trunk, black spruce, birch, gravel
4th iteration	Asphalt, black spruce, birch, gravel
5th iteration	Black spruce, birch, gravel, NPV (from ECOSTRESS spectral library)
6th iteration	Black spruce, birch, gravel

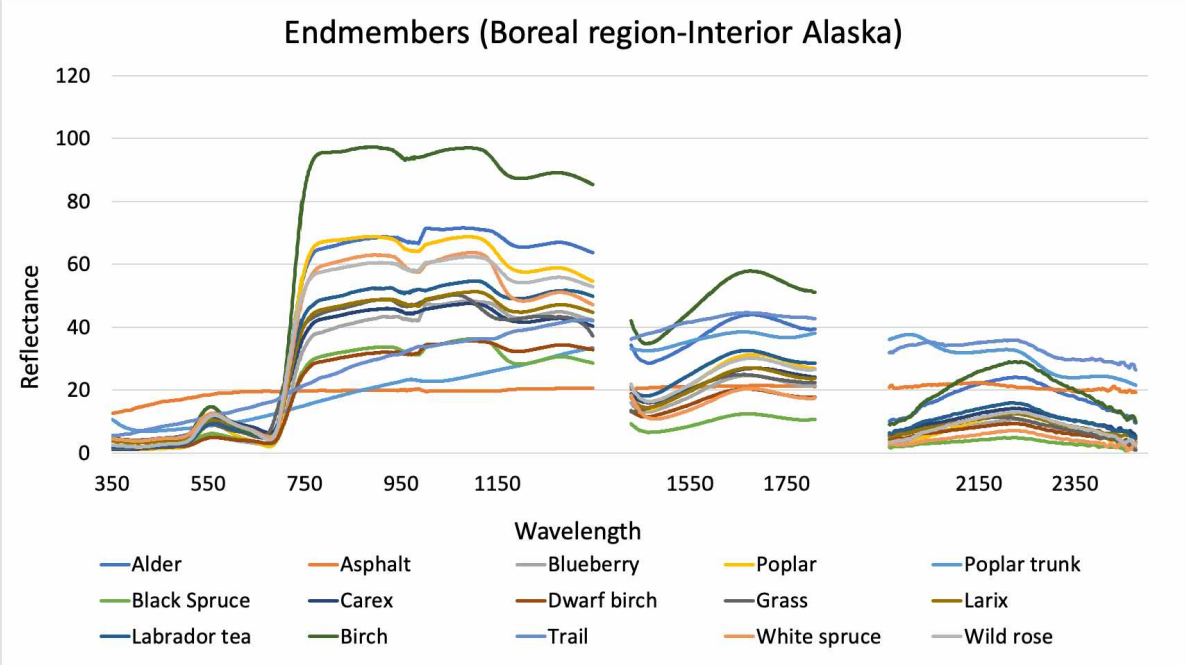
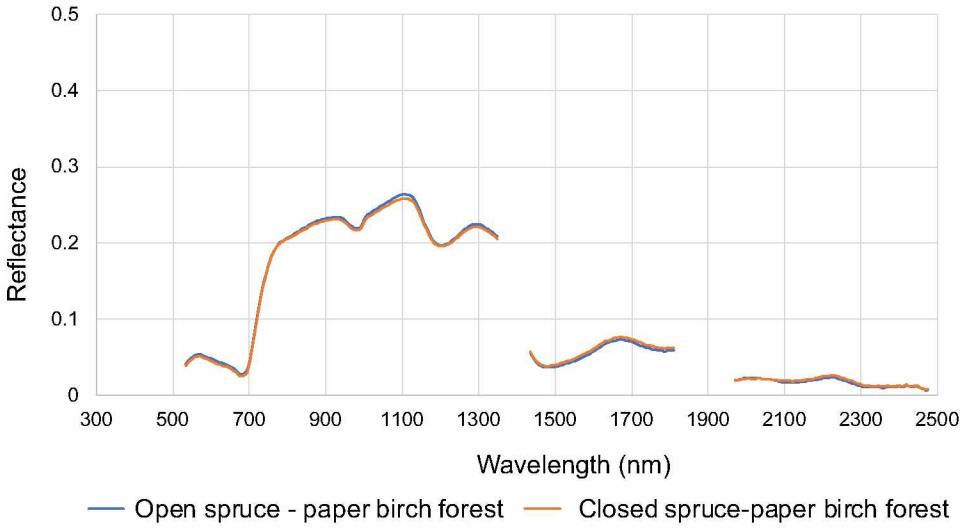
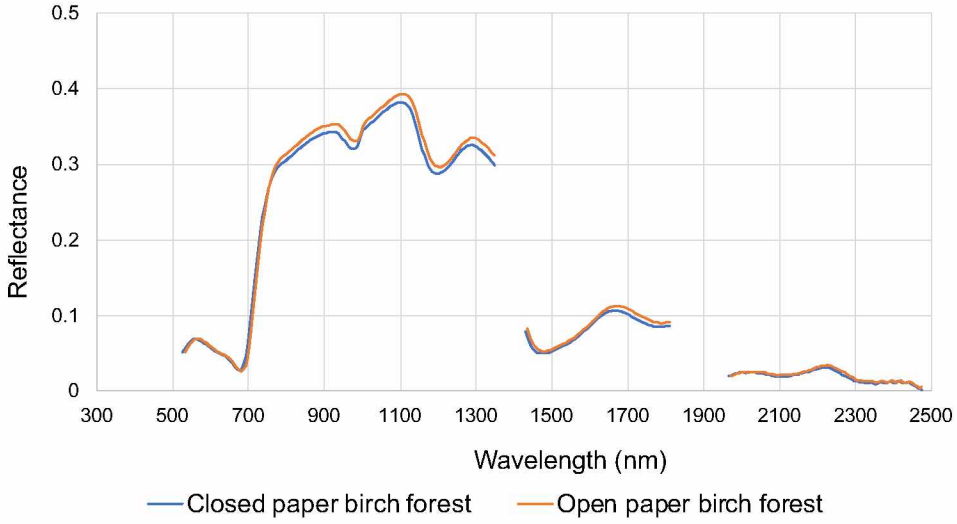


Figure 3.17: Spectral Signature of all the endmembers.

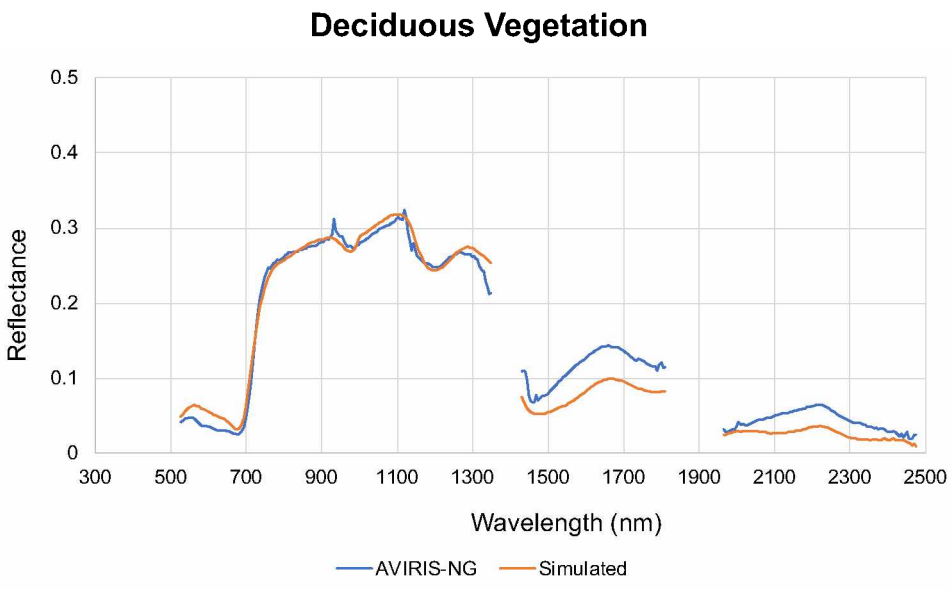


(a)

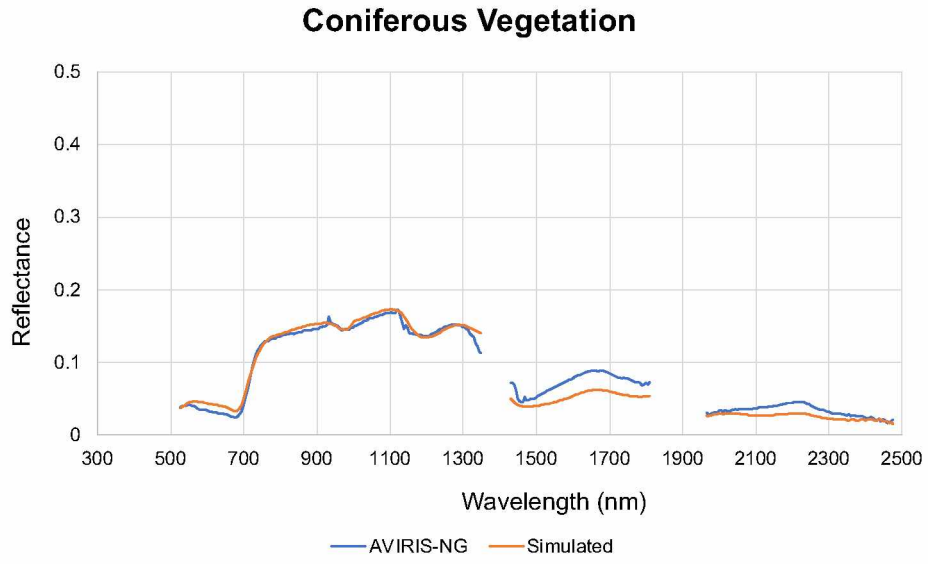


(b)

Figure 3.18: Spectral Signature comparison of different vegetation classes for creating training data.



(a)



(b)

Figure 3.19: Spectral Signature of (a) deciduous vegetation and (b) coniferous vegetation.

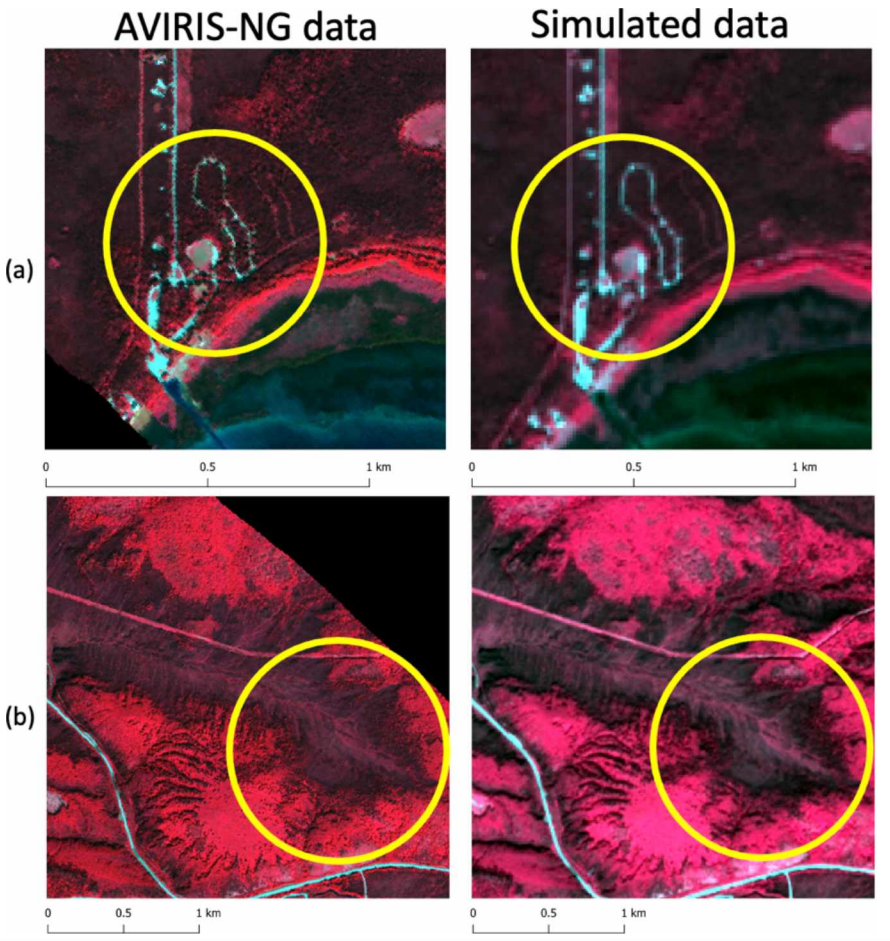
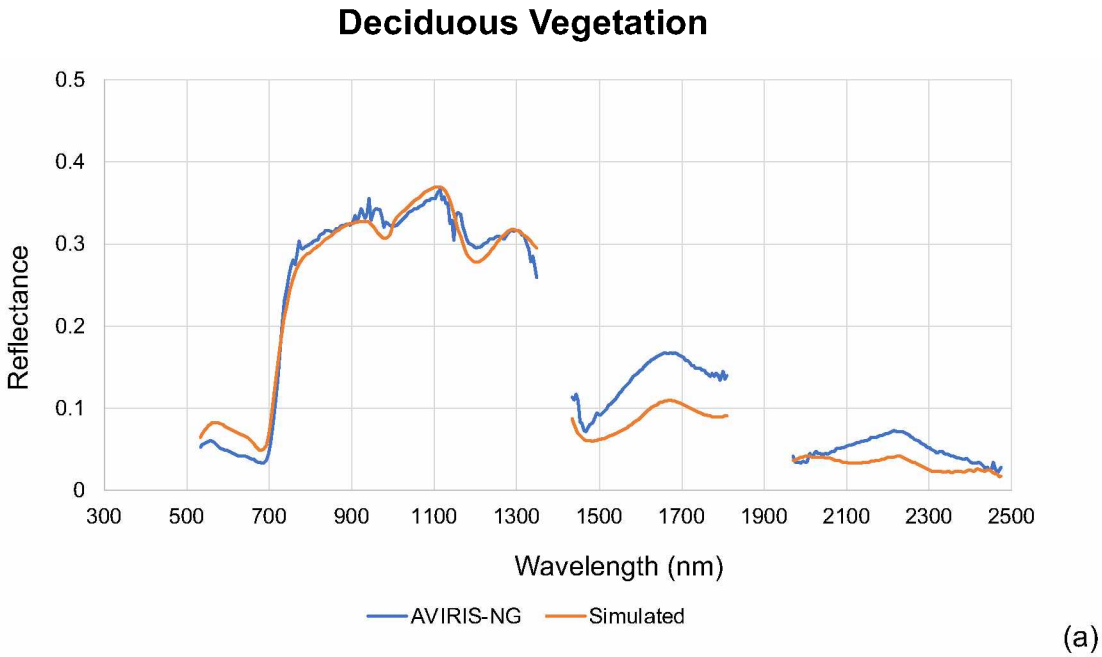


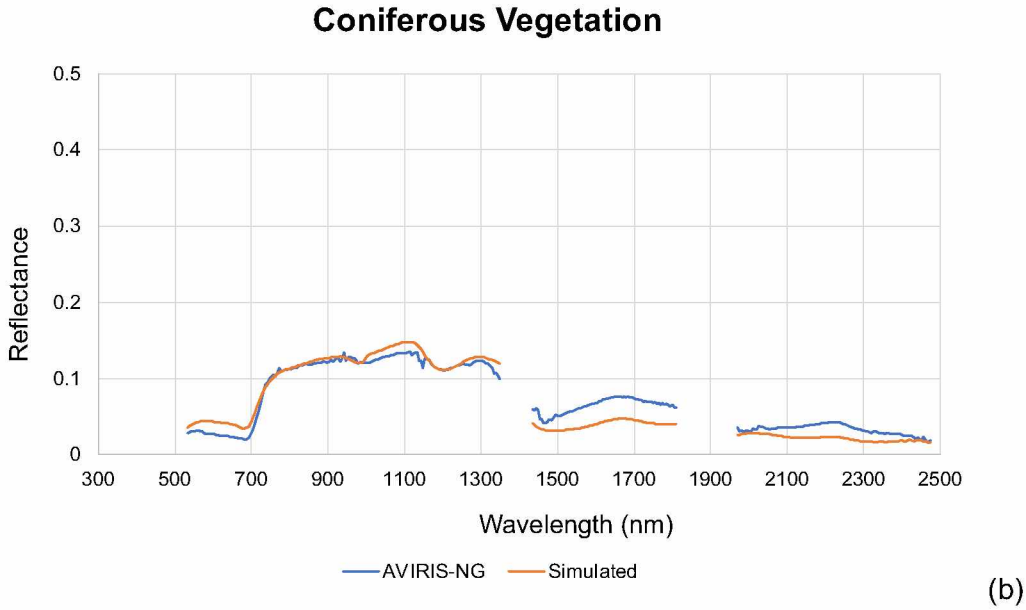
Figure 3.20: Visual Analysis of the simulated scene 2 using Colored InfraRed image for different areas: (a) trails and built-up area (b) coniferous and deciduous species.

Table 3.5: Misclassified Classes in Simulated Classified Product.

Ground data	**Simulated classified output**
Open black spruce (2 plots)	Open spruce - balsam poplar
Open spruce- paper birch (1 plot)	Black spruce woodland
Open spruce-paper birch (2 plots)	Open paper birch
Closed spruce-paper birch- quaking aspen (2 plots)	Open spruce - paper birch
Closed spruce- paper birch- quaking aspen (2 plots)	Open quaking aspen
Closed spruce- paper birch- quaking aspen (1 plot)	Other
Wetlands (1 plot)	Black spruce woodland



(a)



(b)

Figure 3.21: Spectral Signature of (a) deciduous vegetation and (b) coniferous.

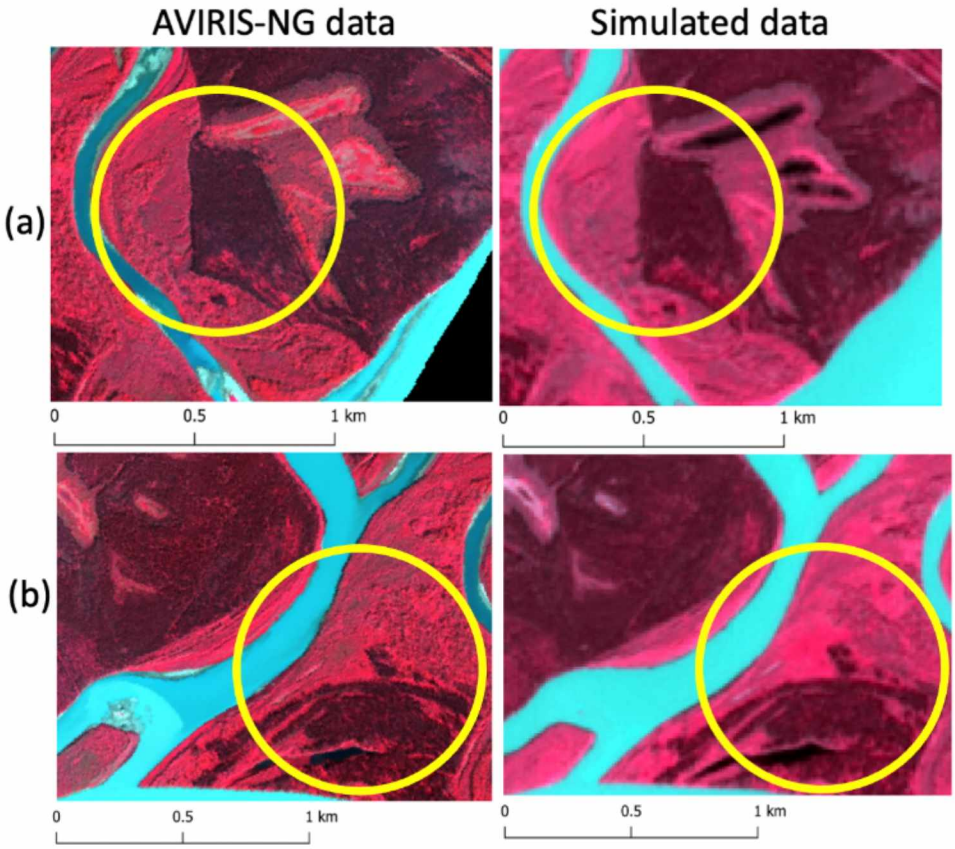


Figure 3.22: Visual Analysis of the simulated scene 3 using Colored InfraRed image for different areas: (a) different vegetation types (b) vegetation pattern.

Chapter 4 Estimation and Validation of Sub-Pixel Needleleaf Cover Fraction in the Boreal Forest of Alaska to Aid Fire Management[3]

4.1. Abstract

Wildfires, which are a natural part of the boreal ecosystem in Alaska, have recently increased in frequency and size. Environmental conditions (high temperature, low precipitation, and frequent lightning events) are becoming favorable for severe fire events. Fire releases greenhouse gasses such as carbon dioxide into the environment, creating a positive feedback loop for warming. Needleleaf species are the dominant vegetation in boreal Alaska and are highly flammable. They burn much faster due to the presence of resin, and their low-lying canopy structure facilitates the spread of fire from the ground to the canopy. Knowing the needleleaf vegetation distribution is crucial for better forest and wildfire management practices. Our study focuses on needleleaf fraction mapping using a well-documented spectral unmixing approach: multiple endmember spectral mixture analysis (MESMA). We used an AVIRIS-NG image (5 m), upscaled it to 10 m and 30 m spatial resolutions, and applied MESMA to all three images to assess the impact of spatial resolution on sub-pixel needleleaf fraction estimates. We tested a novel method to validate the fraction maps using field data and a high-resolution classified hyperspectral image. Our validation method produced needleleaf cover fraction estimates with accuracies of 73%, 79%, and 78% for 5 m, 10 m, and 30 m image data, respectively. To determine whether these accuracies varied significantly across different spatial scales, we used the McNemar statistical test and found no significant differences between the accuracies. The findings of this study enhance the toolset available to fire managers to manage wildfire and for understanding changes in forest demography in the boreal region of Alaska across the high-to-moderate resolution scale.

[3] Badola, A., Panda, S.K., Thompson, D.R., Roberts, D.A., Waigl, C.F., Bhatt, U.S., 2023. Estimation and Validation of Sub-Pixel Needleleaf Cover Fraction in the Boreal Forest of Alaska to Aid Fire Management. Remote Sens. 2023, Vol. 15, Page 2484 15, 2484. https://doi.org/10.3390/RS15102484

4.2. Introduction

Boreal forests in Alaska are found between the Brooks Range in the north and the Alaska Range in the south, covering an area of 43 million hectares [1,2]. Evergreen coniferous vegetation which mostly include black spruce (*Picea mariana*) and white spruce (*Picea glauca*) dominate boreal forests, particularly in interior Alaska which is the heartland of wildfires. Climate is the primary driver for wildfires in boreal Alaska [3]. Between 1976 and 2016, the annual temperature of Alaska increased by 0.3 degrees Celsius per decade and is expected to rise by 1–2 degrees Celsius by [4,5,6]. According to Bieniek et al. (2020), lightning has increased by 17% throughout Alaska over the last 30 years. These changes are influencing wildfire events and increasing the fire frequency, severity and burn acreage [7]. In the last two decades (2003–2022: 12.9 million hectares), wildfires in Alaska have burned around twice as many hectares than the previous two decades (1983–2002: 6.7 million hectares) [8,9]. Though wildfires are a natural part of the boreal ecosystem, recycling soil nutrients and renewing forest health, increasing fire frequency, severity and burned acreage have far-reaching environmental and societal impacts. Some of the adverse impacts include loss of habitat and subsistence resources, risk to communities at the wildland–urban interface, high cost of fire-fighting and restoration, and disproportionate carbon emissions [5,10]. To a large extent, fire spread and intensity are dependent on the vegetation or fuel types. Needleleaf vegetation/fuel are more flammable and spread fire more efficiently compared to broadleaf vegetation/fuel due to their resin content and low-lying canopy structure that serves as a ladder fuel leading to severe crown fires [11]. The maps of the needleleaf vegetation distribution are important at all three stages of fire management. Prior to the fire season, they can help land managers identify high fire risk areas to employ fuel management practices such as building fire breaks, tree thinning, removing dead fuels, etc. During an active fire, these maps can serve as input for fire spread modeling and forecasting near real-time fire spread and behavior. Post-fire, they can help understand the impacts of fire on the ecosystem, forest recovery and demography.

Remote sensing is a proven approach to map vegetation types or classes. Pixel-level mapping is very popular in the remote sensing community, where each pixel is mapped as a vegetation class [8,12,13,14,15,16]. In reality, a pixel can contain more than one class; in that case, the total pixel reflectance will be the combination of the reflectance from all classes present within the pixel. In a boreal landscape, a pixel can contain both needleleaf and broadleaf vegetation. However, a pixel-level vegetation product will map the pixel as one class that dominates in the pixel. For sub-pixel vegetation mapping, i.e., the estimation of the different vegetation fraction in a pixel, one can employ spectral unmixing, also known as spectral mixture analysis (SMA), an approach to map a vegetation fraction where the algorithm calculates the

proportion of each class within a pixel. SMA considers the spectrum of a single pixel as a weighted sum of the constituent spectra of classes or endmembers [17], providing sub-pixel level estimates of the vegetation class fraction. Multiple endmember spectral mixture analysis (MESMA) [18] is an advanced SMA method that assumes that an image is composed of large numbers of different endmembers or classes, but a pixel can be composed of a subset of endmembers. Hence, MESMA allows a large number of endmembers across the scene, but each pixel is modeled independently with a different number and type of endmembers. MESMA was used to map green vegetation, non-photosynthetic vegetation (NPV), and soil in Santa Monica Mountains California, USA, using AVIRIS data [18]. Fernández-García et al. (2021) applied MESMA to Landsat data to study habitat diversity over Cantabrian Mountains located in the northwest of the Iberian Peninsula [19]. Fernandez-Manso et al. (2016) used MESMA to map the burn severity using Landsat images over Sierra del Teleno in Northern Spain using green vegetation (GV), non-photosynthetic vegetation and ash (NPVA), and soil as endmembers [20]. In boreal Alaska, it can be challenging to distinguish a needleleaf pixel from a mixed pixel due to the lower spectral separability between the classes. This challenge provides an opportunity to test the MESMA algorithm for needleleaf mapping in boreal Alaska. Furthermore, there are several existing global hyperspectral space missions which pro-vide data at a coarser spatial resolution (30 m or 60 m). Examples of these missions include the Earth Surface Mineral Dust Source Investigation (EMIT) [21] which has 285 spectral bands (381–2493 nm) and 60 m spatial resolution; the Environmental Mapping and Analysis Program (EnMap) [22], which has 228 spectral bands (420–2450 nm) and 30 m spatial resolution; and the PRecursore IperSpettrale of the application mission (PRISMA) [23], which has 220 spectral bands (400–2500 nm) and 30 m spatial resolution. Additionally, a new mission called Surface Biology and Geology (SBG) [24] is planned to have 217 spectral bands (80–2500 nm) and 30 m spatial resolution. These missions will increase the availability of the hyper-spectral data for a variety of applications, including detailed vegetation mapping at the regional scale. These global sensors have a coarser spatial resolution compared to aerial hyperspectral images such as AVIRIS-NG; therefore, it is also important to evaluate the performance of MESMA in mapping major vegetation classes at different spatial resolutions [25]. This study focuses on sub-pixel level needleleaf vegetation mapping at different spatial scales using AVIRIS-NG data, addressing the following re-search questions:

A. Does MESMA have the potential to estimate the needleleaf fraction in a mixed boreal vegetation with reasonable accuracy?
B. Does the spatial resolution of a hyperspectral image influence the estimation of needleleaf fraction?

C. How can we validate spectral unmixing estimations at different spatial scales?

4.3. Materials and methods

The methodology comprises two main components: pixel unmixing and validation, as illustrated in Figure 4.1. We collected endmembers (needleleaf, broadleaf, and NPV) from hyperspectral imagery (AVIRIS-NG). Then, we performed spectral unmixing and validated the results using methods, elaborated in the following sections.

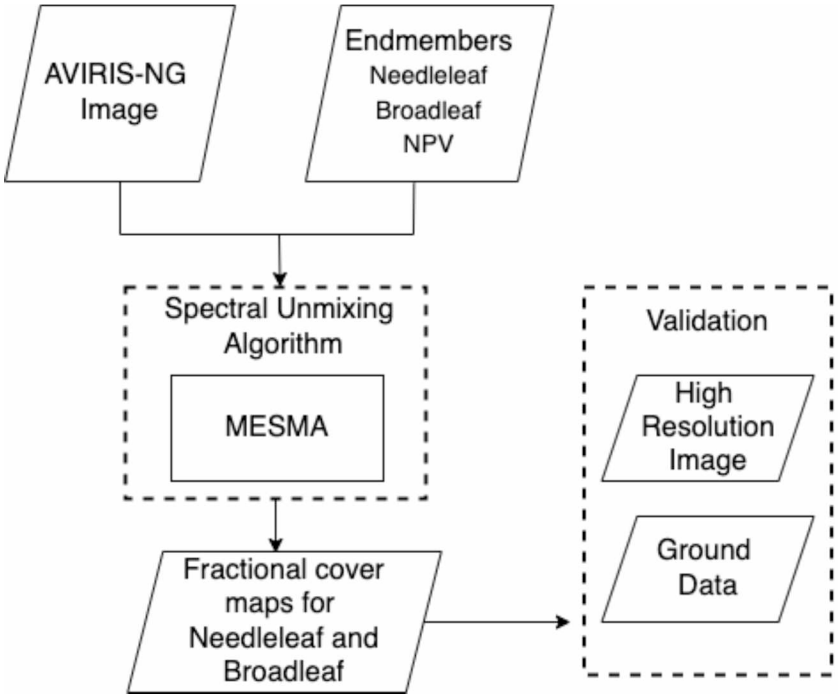


Figure 4.1: Processing workflow for needleleaf and broadleaf fraction mapping and validation.

4.3.1. Study area

The National Science Foundation (NSF) established the Long-Term Ecological Re-search Program (LTER) Network in 1980 to conduct ecological studies and collect long-term datasets to analyze environmental change. Bonanza Creek Experimental Forest (BCEF) is one of the LTER sites in Alaska. It is located in interior Alaska (64.70 °N, −148.30 °W), approximately 30 km southwest of Fairbanks, covering an area of 5053 ha. For this study, we selected a test site within BCEF (Figure 4.2), where the AVIRIS-NG scene was available. This region lies between the Brooks Range in the north and the Alaska Range in the south, which blocks the coastal air masses; hence, the area experiences cold winters as well as warm and dry summers. The study area experiences short growing seasons (100 days or less), and the air temperature ranges from

−50 °C in January to over +33 °C in July, with a long-term average annual temperature of −3 °C. The mean annual precipitation is approximately 269 mm, 30% of which is in the form of snowfall [1]. The study area includes both upland and lowland regions with a variety of vegetation types [26]. The soils are immature, ranging from cold soils with shallow permafrost in the lowlands to warm and well-drained soils in the uplands. Lowlands and north-facing slopes are covered by moss-dominated black spruce, while aspen, birch, and white spruce mainly grow in uplands and south-facing slopes.

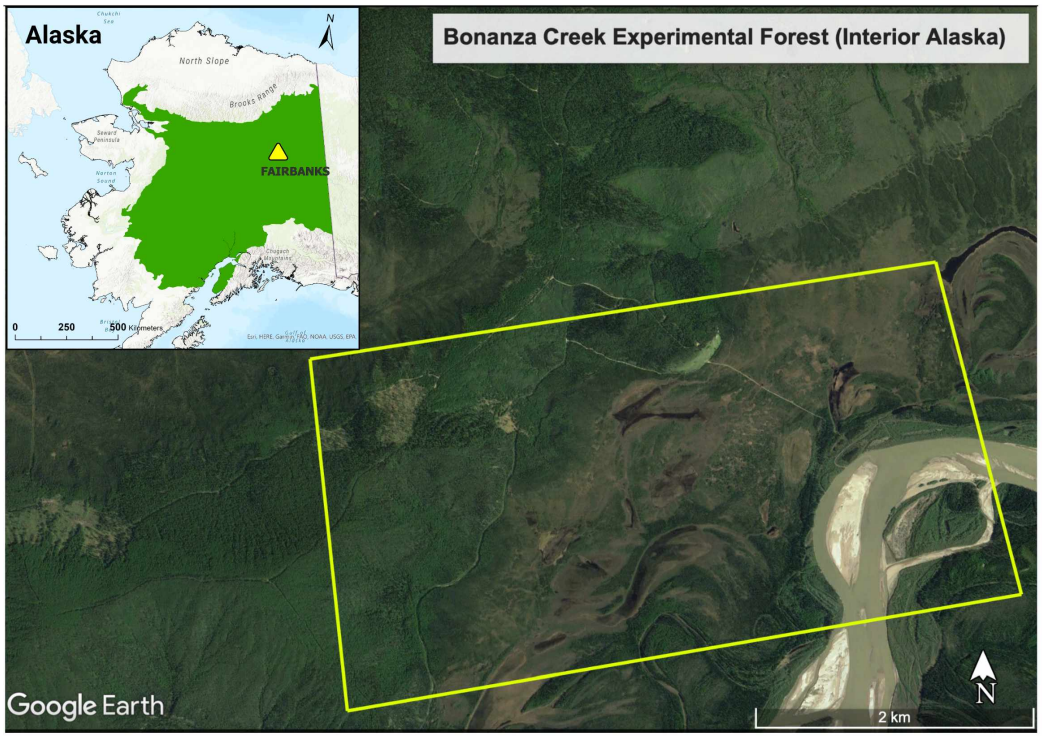


Figure 4.2: Study area: Bonanza Creek Experimental Forest (BCEF), located 30 km southwest of Fairbanks in interior Alaska. Boreal forest area represented in green on inset map. A yellow boundary delineates the study area within BCEF.

4.3.2. Field data collection

This study required extensive fieldwork to validate the fraction product generated in this study. It involved conducting on-site surveys to record the proportion and distribution of various species within the study area. Specifically, we collected the percentage cover and count of tree species within a plot. We surveyed 40 plots of 10 m × 10 m size during 2019 (Figure 4.3a, b) and two plots of ~1000 m2 in size in 2022 (Figure 4.3c) using a Trimble Real-Time Kinematic (RTK) Global Positioning System (GPS) unit that offers centimeter-level positional accuracy. In the field, we recorded the relative proportion of key vegetation species in 40 (10 m × 10 m) plots. We

subdivided the 2022 field plots into smaller plots and recorded the broadleaf and needleleaf tree counts. We also noted the tree species that dominate the top canopy. We did not see any evidence of forest damage or anthropogenic change such as timber harvesting since 2018 (image year), ensuring that the use of image data and field data collected during different times is reasonable (Tables 4.1 and 4.2).

Table 4.1: List of image datasets used in this study.

Data	Scene Identifier	Acquisition Date	Spatial resolution	Bands
AVIRIS-NG	ang20180723t200207	23 July 2018	5m	425
SkySat	20190629_002107_ssc10_u0002	29 June 2019	0.5m	4
HySpex	20210803_BC	3 August 2021	1m	459

Table 4.2: List of field data used in this study.

Plot Data	Instrument	Data Collection time	Data collected
10m x 10m (40 plots)	Trimble RTK-GPS	Summer 2020 and 2021	Vegetation composition, canopy cover, diameter and height
~1000 m^2 (2 plots)	Trimble RTK-GPS and Garmin Handheld GPS device	Summer 2022	Needleleaf tree count

4.3.3. Data preprocessing

We used the atmospherically and radiometrically corrected level-2 AVIRIS-NG data (Table 4.1) acquired as a part of the Arctic-Boreal Vulnerability Experiment (ABoVE) [27]. The data have 425 bands, 5 m pixel resolution, and were collected on 23 July 2018. We removed bands with excessive noise due to poor radiometric calibration and bands dominated by water vapor and methane absorption [28]. Table 4.3 lists the removed bands. For this study, we used a subset of the AVIRIS-NG flight line (Table 4.1). All the preprocessing was performed in ENVI classic 5.3 [29].

Table 4.3: Bands removed from AVIRIS-NG hyperspectral data

Bands	Wavelength (nm)	Remarks
196–210	1353.55–1423.67	Water vapor absorption bands
288–317	1814.35–1959.60	Water vapor absorption bands
408–425	2415.39–2500.00	Noise due to poor radiometric calibration and strong water vapor and methane absorption

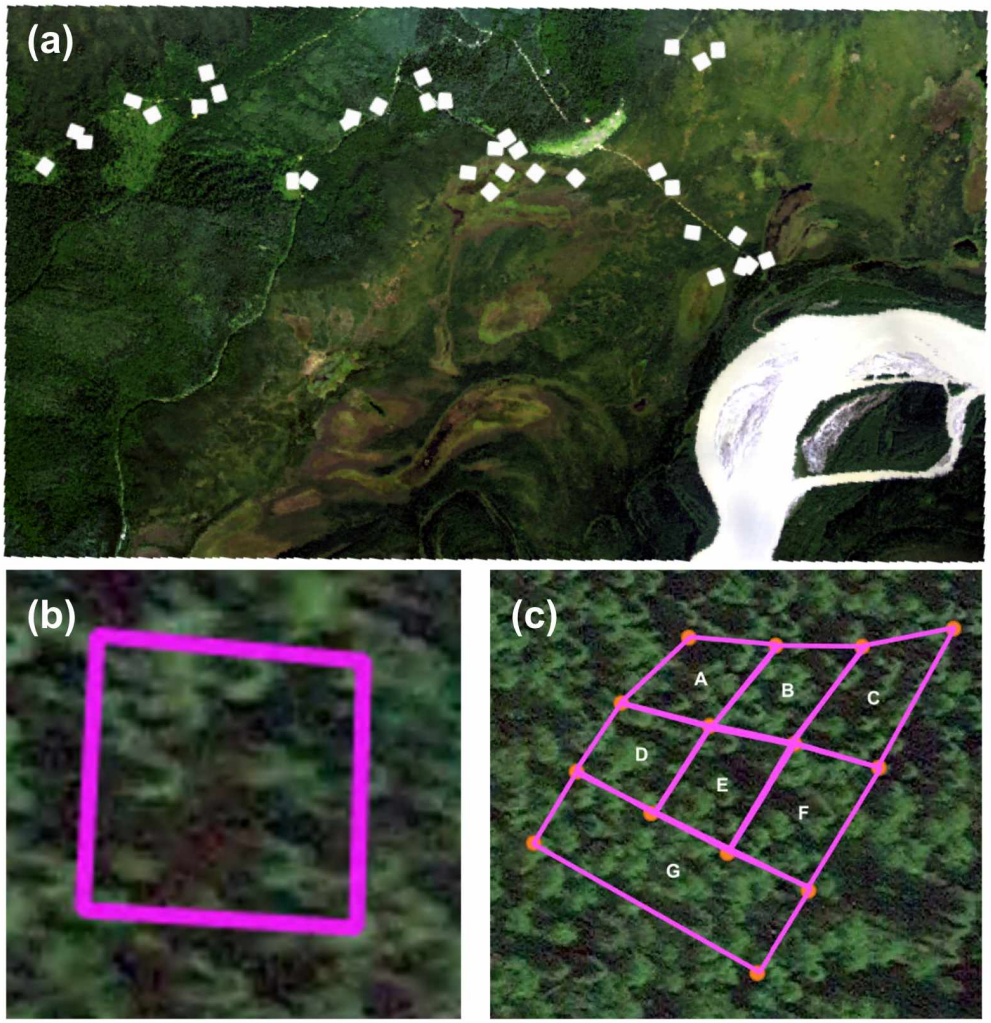


Figure 4.3: Ground data collected for validation: (a) white polygons show the location of 40 (10 m × 10 m) plots; (b) a single (10 m × 10 m) plot; (c) a ~1000 m^2 plot subdivided into 7 subplots (A–G). The fairly dense forest setting made it difficult to set up a 30 m × 30 m square plot in the field.

4.3.4. Endmember selection

Selecting appropriate endmembers is a critical step in spectral unmixing. In our study area, the primary vegetation classes consist of spruce (needleleaf) and birch (broadleaf), which serve as two crucial endmembers [8,15]. Additionally, we considered including other endmembers in our analysis, such as shrubs (including broadleaf and evergreen shrubs), non-photosynthetic vegetation (NPV), and soil. However, we found that the spectral contrast between broadleaf trees and shrubs was low, which made it difficult for the algorithm to distinguish between them. Our primary goal was to map needleleaf vegetation, which are highly flammable, so we combined the broadleaf trees and shrubs into a single class.

When we ran the spectral unmixing algorithm with the combined broadleaf/shrub class, needleleaf, NPV, and soil as endmembers, we found that the algorithm could not accurately distinguish the soil and NPV due to the low spectral contrast between them and could only map one of them. As NPV includes dead branches, leaf litter, and dry vegetation, which are highly flammable and important for wildfire management, we decided to drop soil from consideration and ultimately selected needleleaf, broadleaf, and NPV as our endmembers. Selecting these endmembers allowed us to focus on mapping the highly flammable fuels (needleleaf vegetation and NPV) that most impact fire spread. Spectral processing was performed in the Visualization and Image Processing for Environmental Research (VIPER) Tools 2 (beta) software as an extension of the ENVI software [30]. Figure 4.4 shows the endmember spectra selected for this study.

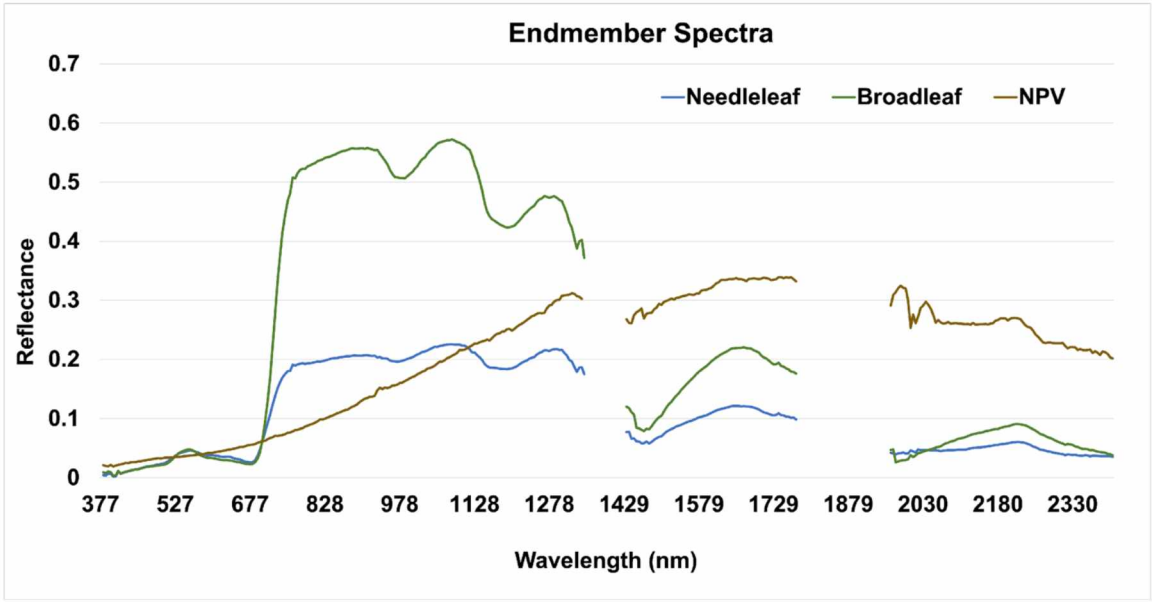


Figure 4.4: Endmember spectra of needleleaf, broadleaf and NPV. The graph shows the average spectra for all three endmembers used in the spectral unmixing model.

4.3.5. Spectral unmixing

This study focused on mapping the highly flammable fuels (needleleaf vegetation) at sub-pixel level. We assumed that every pixel in the AVIRIS-NG image could be modeled by a linear combination of two types of vegetation (broadleaf and needleleaf) and NPV (Figure 4.5). We applied multiple endmember spectral mixture analysis (MESMA) [31] by analyzing all of the potential endmember combinations for each pixel, starting with one endmember model: Needleleaf, Broadleaf, and NPV, and two endmember models: Needleleaf-Broadleaf, Needleleaf-NPV, and Broadleaf-NPV. Upon using the three-endmember model, we found an increase in the number of unclassified pixels to 25%, which indicated that the model was not effectively capturing the spectral variability in the data. As a result, we decided to use only one and two endmember models since the majority of the pixels (over 90%) were modeled by them. We constrained the minimum and maximum permitted endmember fractional values between 0.00 and 1.00, and a maximum allowable RMSE of 0.025.

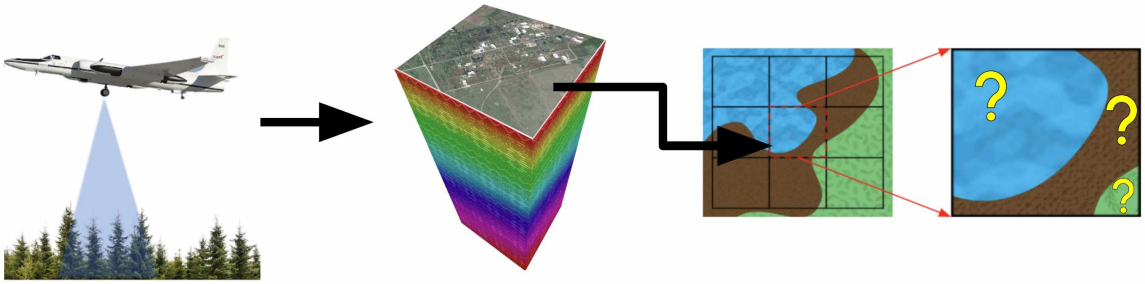


Figure 4.5: A pixel can contain more than one class at different proportions marked in yellow. Spectral unmixing approach is used to estimate the proportions of classes within a pixel.

In the MESMA algorithm, the selection of endmembers is critical for the accurate spectral unmixing of mixed pixels. However, in some cases, there may be low spectral contrast between the different endmember types, which can result in the algorithm picking one endmember and adding shade to compensate for brightness. To address this issue and improve the accuracy of spectral unmixing, a shade constraint is often applied in MESMA [32]. The shade constraint limits the maximum shade fraction in the spectral unmixing process, guiding the algorithm to select a bright endmember for bright stands and a dark endmember for dark stands. By doing so, it can effectively differentiate between different vegetation types and minimizes the misclassification caused by low spectral contrast [33]. In this study, we set the maximum allowable shade fraction

to 0.30, which means that the algorithm was optimized to select the needleleaf for darker vegetation and broadleaf for brighter vegetation.

We evaluated the performance of all endmember models using the RMSE and recorded the model with the lowest RMSE for each pixel. The output contained an RMSE band, a band with the models, and fraction bands containing broadleaf, needleleaf, and NPV fraction estimates.

Shade normalization [34] is a post-processing step in spectral unmixing algorithms to remove the effects of shadows caused by topography or other features in an image. The process involves normalizing the estimated fractions of each endmember by dividing them by the sum of fractions of all endmembers in a given pixel excluding shade. This normalization ensures that the total estimated fraction of all non-shade endmembers in a pixel equals 1. Shade normalization is particularly useful in cases where the contrast between endmembers is low, and the spectral unmixing algorithm tends to select one endmember over the others [31]. We performed shade normalization on the spectral unmixed output to improve the estimated fractions of broadleaf, needleleaf and NPV in each pixel. We used VIPER Tools, Version 2 (beta) software [30] as an ENVI plugin to run MESMA and Shade normalization. We resampled the 5 m AVIRIS-NG image to 10 m and 30 m pixel sizes using cubic resampling in the GDAL warp function [35], and re-ran MESMA keeping the same parameters. We used cubic resampling as it determines the pixel value through a weighted average of the 16 closest pixels, resulting in a more accurate representation of the original data [36].

4.3.6. Accuracy assessment

We used three methods to validate the fraction product: (1) visually using high-resolution SkySat image; (2) using 40 (10 m × 10 m) field plots; and (3) using 1 m resolution classified map derived from aerial HySpex image.

4.3.6.1. Visual assessment using high-resolution multispectral data

We visually compared the fraction output images with the high-resolution SkySat data (50 cm pixel resolution) provided by Planet Labs under NASA Commercial Smallsat Data Acquisition Program [37]. We analyzed different areas of interest based on the kind of vegetation class and compared them from the field photos as well.

4.3.6.2. Assessments using 10m X 10m field plots

We used the percentage cover information available from the 40 field plots (Figure 4.3) collected during fieldwork for assessing the performance of sub-pixel output from MESMA. In the MESMA output, we summed up the proportion of each pixel falling within the field plot boundary (ground data) and evaluated the needleleaf fraction. For this, we vectorized the pixels in the

fractional output image using the “Raster pixels to polygons” tool available in QGIS [38]. We then clipped field plots from the vectorized/polygon shapefile and obtained the final clipped shapefile, as shown in Figure 4.6b. We compared the needleleaf fraction from MESMA output and ground data and estimated RMSE. We used the Point Sampling Tool [39] in QGIS to extract the raster values (proportions) from the output fraction images using the clipped shapefile (Figure 4.6b).

We also considered the variability due to the positional inaccuracies in the field plot and the AVIRIS-NG image by taking a 10 m buffer from the plot centroid (Figure 4.6c). We summed up all the pixels whose center was located within the buffer boundary and evaluated the total proportion of needleleaf vegetation within the buffer. We compared the needleleaf proportion from the fraction output and the field plot data and calculated the RMSE. We also identified the dominant species on each field plot and compared it with the class with the highest proportion in the fraction images.

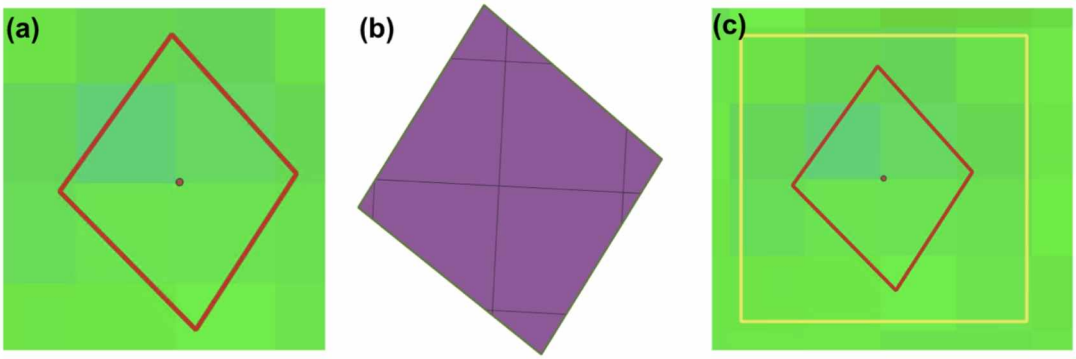


Figure 4.6: Estimating the proportion of each pixel within a (a) 10 m × 10 m field plot boundary (red polygon) overlaid on the fraction output raster point denotes the centroid of the plot; (b) pixels clipped (black boundary) that fall within the plot; and (c) a buffer of 10 m (yellow rectangle) created to include the variability of the vegetation around the plot.

To assess the impact of spatial resolution on fraction output, we resampled the AVIRIS-NG image to 10 m and 30 m and assigned each pixel in the fraction which outputs the class value of the dominant vegetation fraction within the pixel. We categorized 40 plots (10 m × 10 m) into three classes, needleleaf, broadleaf and mixed based on the dominant vegetation and then calculated the confusion matrix for fraction outputs (spatial resolutions: 5 m, 10 m, and 30 m). We then compared the user and producer accuracies evaluated using a confusion matrix for all three products as well as for all three classes.

4.3.6.3. Assessment using high-resolution (1 m) HySpex hyperspectral data

The accuracy assessment of fractional cover is challenging. We devised an approach where we used a ~1000 m^2 field plot and a 1 m resolution hyperspectral image (459 bands) acquired using the HySpex hyperspectral camera from a fixed-wing airplane [40]. Figure 4.7 shows the general methodology for the proposed fraction map vali-dation. We performed random forest classification on the HySpex image using 500 decision trees and 21 features per subset (square root of the total number of bands) [41]. We trained our model using pixels from three classes, needleleaf, broadleaf, and other, and validated the classified map against ~1000 m2 plot data with seven sub-blocks, as shown in Figure 4.3c. For each block, we recorded the needleleaf tree counts. We calculated the correlation coefficient for the needleleaf count from ground data and the pixels classified as needleleaf on a HySpex classified map. Once we validated the HySpex classified map, we generated 500 random points over the classified map and used the 104 points that fell on the needleleaf class. We then evaluated the needleleaf fraction maps at 5 m, 10 m and 30 m by comparing the points over a needleleaf class with the corresponding pixel in the MESMA outputs.

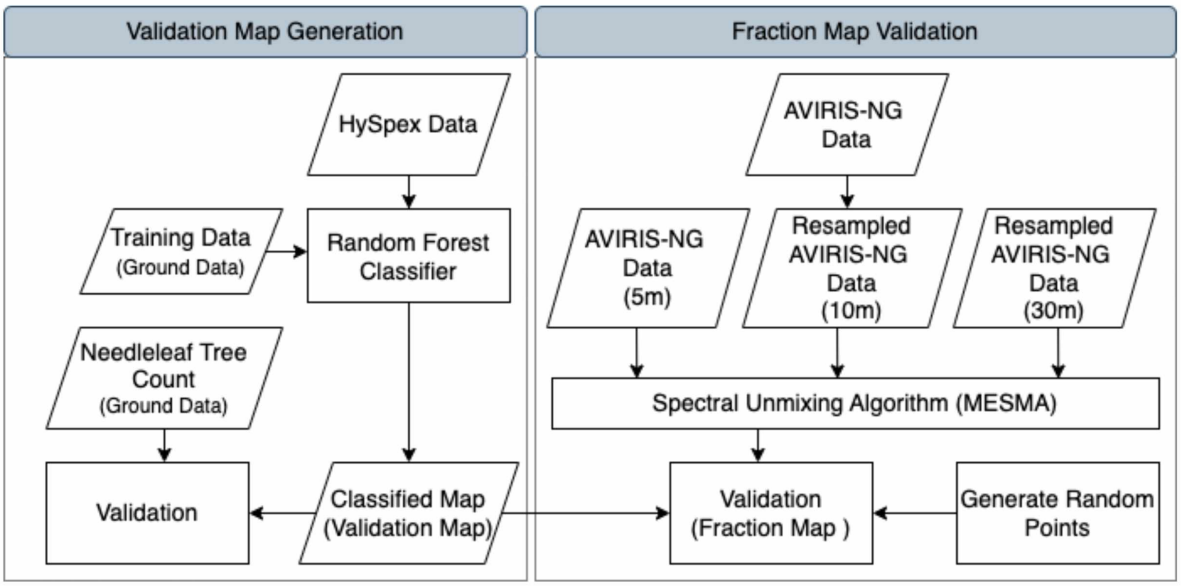


Figure 4.7: Workflow to validate the fraction cover output from AVIRIS-NG data at 5 m, 10 m, and 30 m resolution using 1m classified map derived from HySpex data.

4.3.6.4. Comparison of fraction outputs at different spatial scales.

We compared the fraction outputs at different spatial scales (5m, 10m, and 30m) using the McNemar test, a statistical method used to determine whether there is a significant difference

between the outputs [42]. The null hypothesis of the McNemar test is that the fraction outputs at different spatial scales are not significantly different. We used Equation 4.1 to calculate the z score and computed the *p*-value using the chi2.cdf function from the scipy.stats Python module [43] at a significance level of α = 0.05. If the *p* value is less than 0.05, we reject the null hypothesis, i.e., there is a significant difference in the fraction output across different spatial scales.
The test statistics (z score) for McNemar test is given by:

$$\mathbf{z} = \frac{(|\boldsymbol{a} - \boldsymbol{b}| - \mathbf{1})}{\sqrt{(\boldsymbol{a} + \boldsymbol{b})}} \tag{4.1}$$

where:

- a: number of pixels where test 1 (fraction output 1) is positive and test 2 (fraction output 2) is negative
- b: number of pixels where test 1 (fraction output 1) is negative and test 2 (fraction output 2) is positive

4.4. Results

We successfully generated fractional cover maps at 5 m, 10 m, and 30 m spatial resolution from the AVIRIS-NG image using the MESMA algorithm. Figure 4.8 shows an RGB image of the fractional cover maps (Red: broadleaf; Green: needleleaf; Blue: NPV), and Figure 4.9 shows the needleleaf fraction (in white–green shades) for the test site at a 5 m resolution.

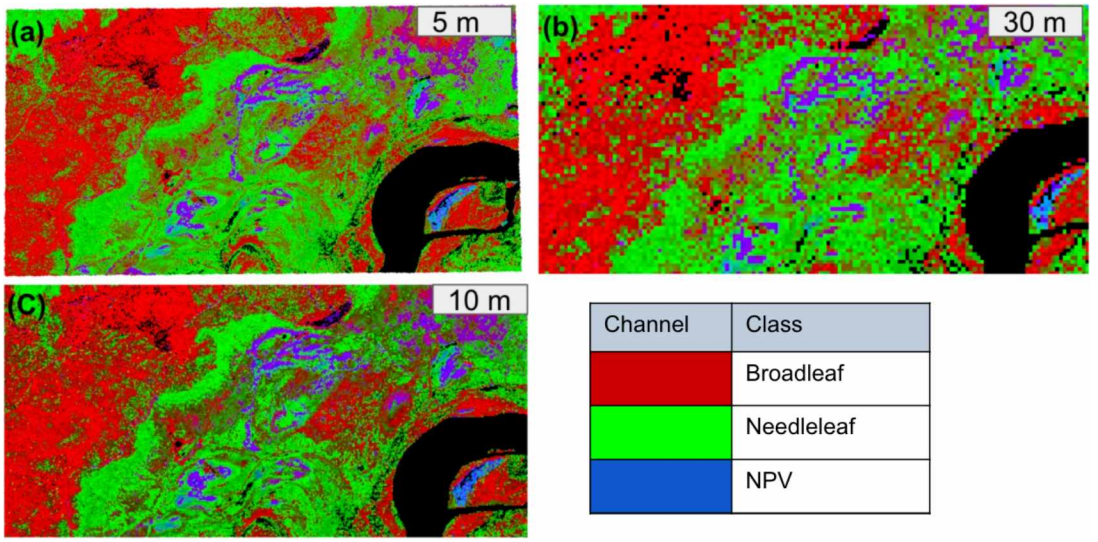


Figure 4.8: RGB images of fraction cover generated using MESMA (a) 5 m spatial resolution (b) 10 m spatial resolution (c) 30 m spatial resolution.

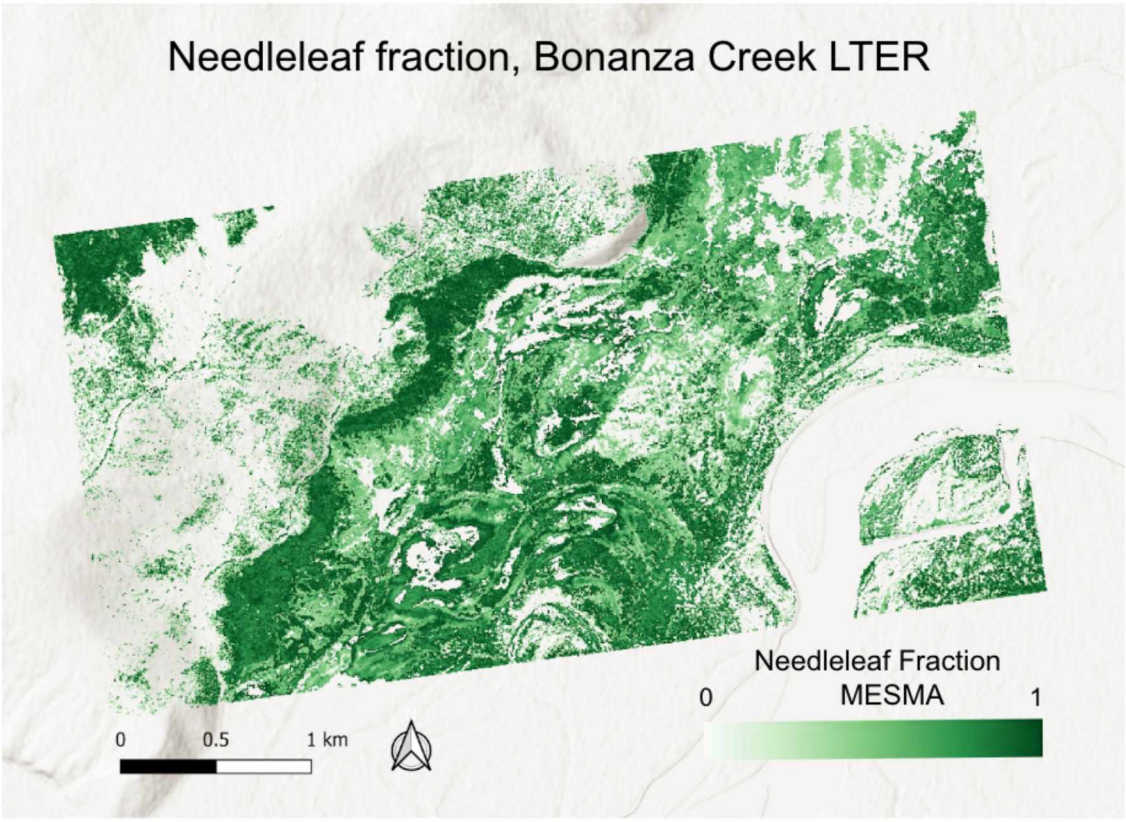


Figure 4.9: Needleleaf fraction cover map at 5 m spatial resolution generated using MESMA on AVIRIS-NG data.

4.4.1. Assessment using high-resolution multispectral data

We used a very high spatial resolution (50 cm) SkySat image and ground observations to visually compare the fractional map cover outputs and found that the MESMA per-formed well in capturing the distribution of needleleaf and broadleaf vegetation types. The area dominated by grass is unclassified in the fraction map since we did not use an endmember for grass (Figure 4.10a). In the color infrared (CIR) band combination, the needleleaf pixels are darker in color, and in the case of the fractional output map, a similar pattern of the needleleaf vegetation is present (in green), highlighted by a white boundary (Figure 4.10b).

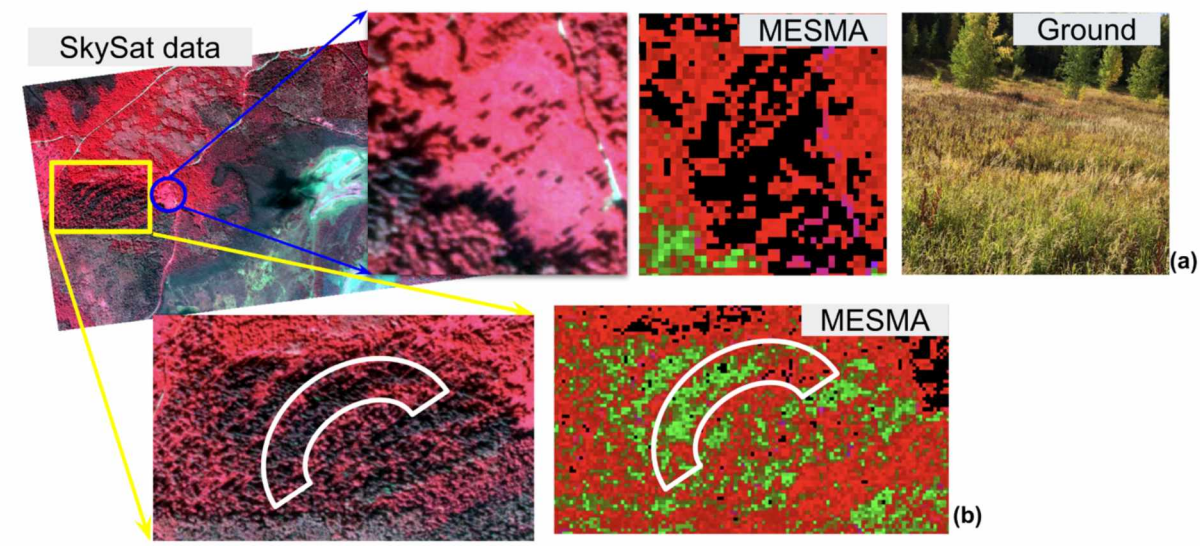
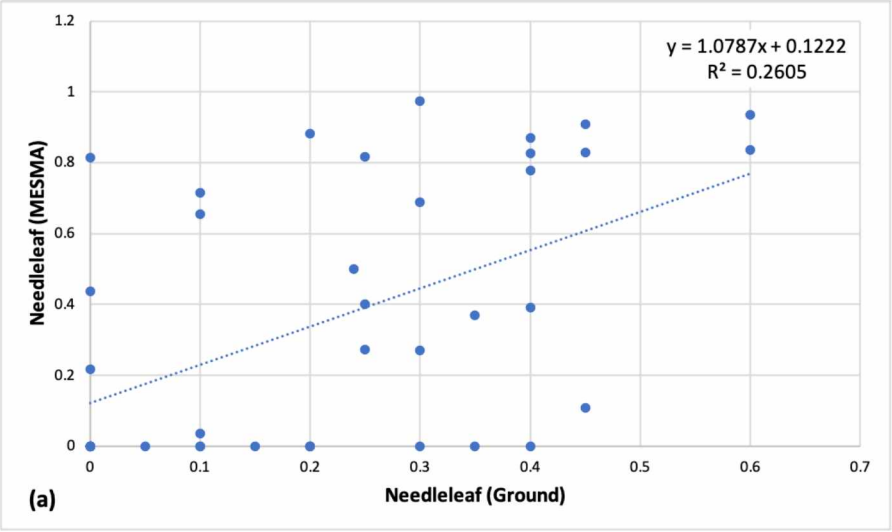


Figure 4.10: Visual comparison of MESMA fraction output using high resolution (50cm) multispectral data (SkySat data) (a) unclassified area where grass was present (b) similar pattern of needleleaf stands in SkySat image and fraction cover map.

4.4.2. Assessments using 10m X 10m field plots

The graph (Figure 4.11a) shows the comparison between the needleleaf proportion from fraction output and the ground observations (RMSE of 0.34). In the second case, we included a buffer of 10 m to compensate for positional inaccuracy in image data (RMSE reduced to 0.29) (Figure 4.11b).

Upon comparing the dominant species in each of the 40 plots and the fraction map, we found that 25 plots have the same dominant species as in the fraction map; hence, a total of 62% of plots were correctly mapped. Figure 4.12 shows the producer and user ac-curacy for the needleleaf, broadleaf, and other class at different scales using the 40 plots (10 m × 10 m).



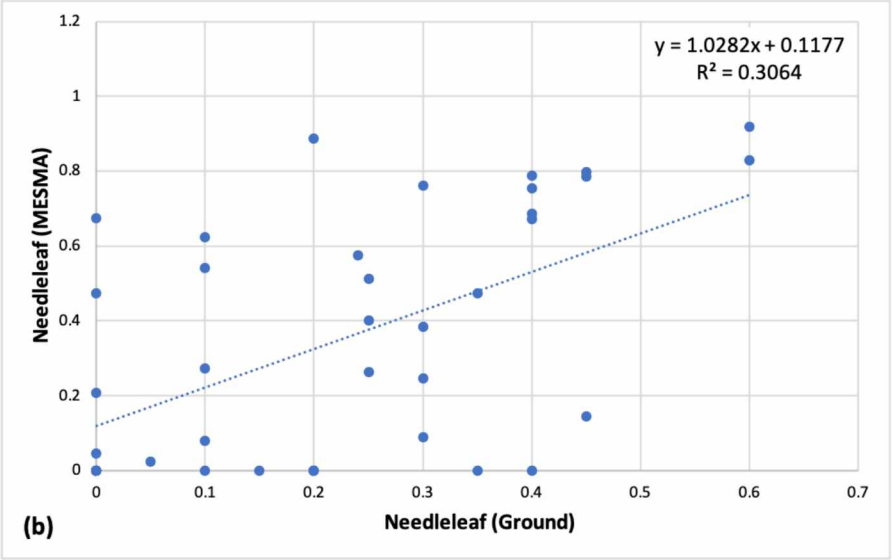


Figure 4.11: Comparison between the needleleaf proportion from fraction output and the ground data: (a) without buffer (b) with 10 m buffer.

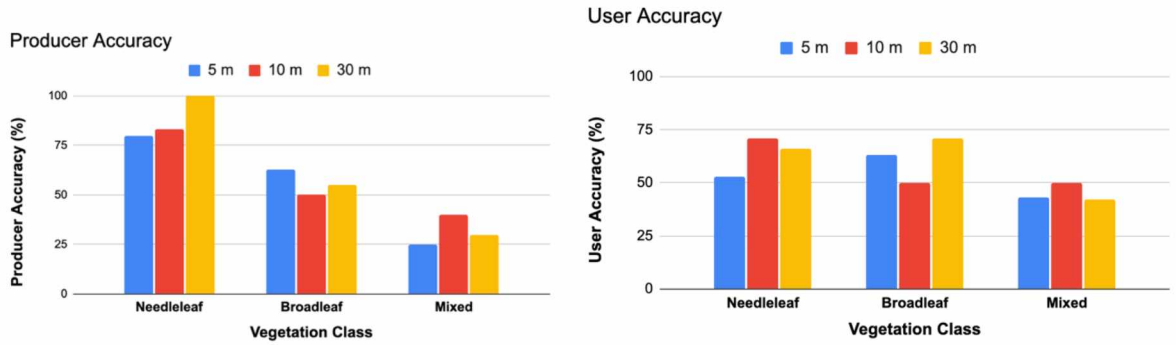


Figure 4.12: Individual class accuracy assessment of fraction cover maps at different spatial scales by assessing producer accuracy (left) and user accuracy (right).

4.4.3. Assessment using high-resolution (1 m) HySpex hyperspectral data

We generated a 1 m high-resolution classified map product using HySpex data to validate the fraction cover. Figure 4.13a shows the random forest classified map with three classes. Since we are interested in mapping the needleleaf fraction, we performed this validation for needleleaf vegetation. The blue pixels are the needleleaf pixels on the classified map. Figure 4.13b shows the correlation graph between the number of needleleaf trees based on the ground observation and the number of pixels classified as needleleaf vegetation represented in blue. We found a high positive correlation between them, with a coefficient of 0.9 and an r-squared value of 0.8. We validated the fraction cover map using random points (needleleaf class) and a classified map (validation map). In the case of the 5 m fraction map, 73% of the points were mapped correctly; for the 10 m fraction map, 79% of the points were mapped correctly; and for the 30 m spatial resolution product, 78% of the points were mapped correctly (Figure 4.14). There was no major difference in accuracy for needleleaf vegetation in all three spatial resolution outputs.

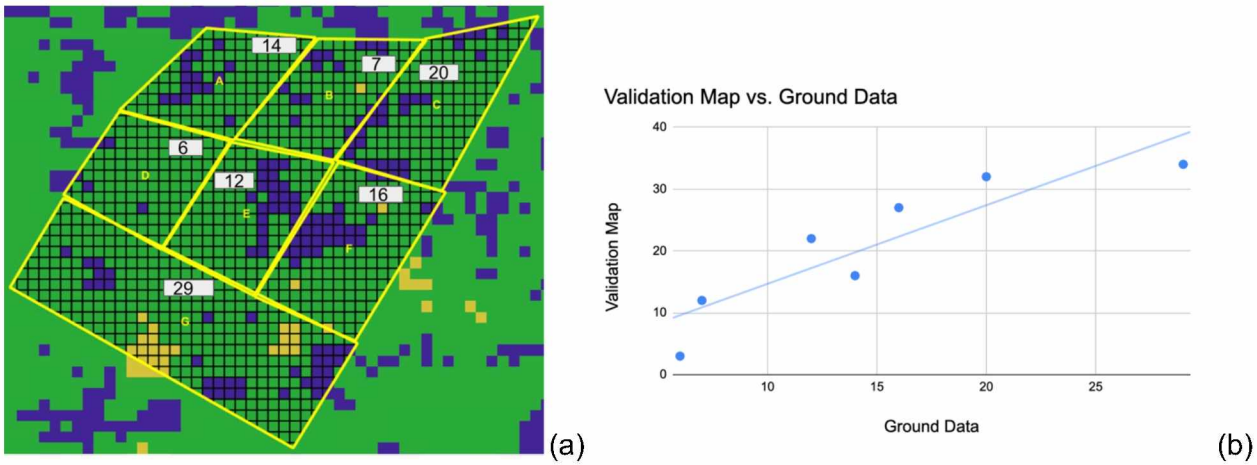


Figure 4.13: (a) Random Forest classified map (blue: needleleaf; green: broadleaf; yellow: other) of the test site using HySpex Data (1 m spatial resolution). Field plot boundary is shown in yellow, with subplots (A-G); (b) shows the correlation between the number of needleleaf trees based on the ground observation and the number of pixels classified as needleleaf class on HySpex Data.

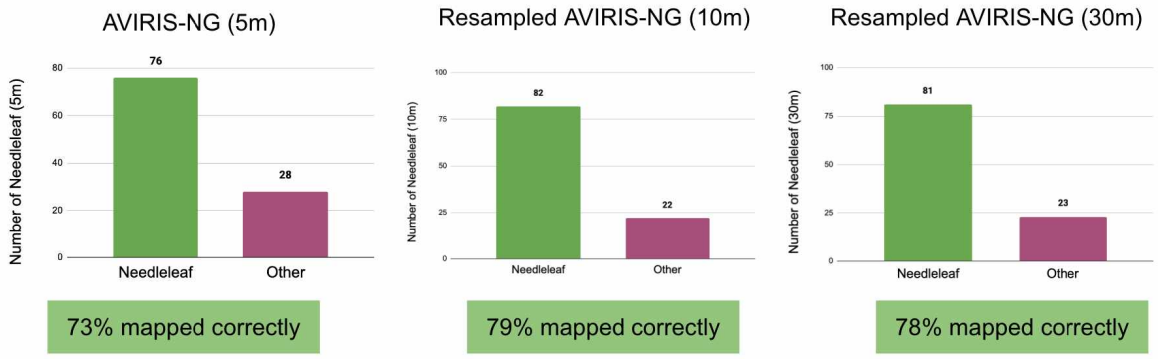


Figure 4.14: Validation of the MESMA fraction cover map (at 5m, 10m, and 30m) using random needleleaf pixels from the HySpex classified map.

4.4.4. Comparison of fraction outputs at different spatial scales.

We used the McNemar test to determine whether the fraction output results were significantly different from each other. Tables 4.4, Table 4.5 and Table 4.6 show the contingency matrices obtained by applying the McNemar test. Conifer fractions at 5 m and 10 m spatial resolution showed that 72 points were correctly mapped at both scales, resulting in a *z* score of 1.79 and a *p* value of 0.18 (Table 4.4). A comparison of fractions at 5 m and 30 m spatial resolutions, showed that 68 points were mapped correctly at both scales, resulting in a *z* score of 0.76 and a *p* value of 0.38 (Table 4.5). A comparison of fractions at 10 m and 30 m showed an agreement for 70 points, resulting in a *z* score of 0.00 and a *p* value of 1.00 (Table 4.6). In all three cases, the *z* scores were between −1.96 and 1.96, and the *p* values were greater than 0.05. These results are consistent with the null hypothesis, indicating that there is no significant difference in the fraction outputs across different spatial scales.

Table 4.4: Contingency matrix for McNemar test to compare fraction outputs (5m and 10m).

	5m Output (wrong)	5m Output (correct)	All
10m Output (wrong)	18	4	22
10m Output (correct)	10	72	82
All	28	76	104
McNemar	z score	1.79	
results	p value	0.18	

Table 4.5: Contingency matrix for McNemar test to compare fraction outputs (5m and 30m).

	5m Output (wrong)	5m Output (correct)	All
30m Output (wrong)	15	8	23
30m Output (correct)	13	68	81
All	28	76	104
McNemar	z score	0.76	
results	p value	0.38	

Table 4.6: Contingency matrix for McNemar test to compare fraction outputs (10m and 30m).

	10m Output (wrong)	10m Output (correct)	All
30m Output (wrong)	11	12	23
30m Output (correct)	11	70	81
All	22	82	104
McNemar	z score	0.00	
results	p value	1.00	

4.5. Discussion

Most space-borne hyperspectral data are available at a coarser spatial resolution (10 m and 30 m) so there is a need for sub-pixel estimates of the highly flammable fuels to aid in fire management. In this study, we tested the potential of MESMA to map the needleleaf fraction in the boreal region of Alaska and assessed the impact of spatial resolution on the MESMA output. A few studies have applied MESMA for mapping green vegetation, soil, and non-photosynthetic vegetation [18,44,45], but this is the first work in which MESMA has been used to map the needleleaf fraction in a mixed boreal forest of Alaska. The brightness of the needleleaf and broadleaf vegetation is crucial to distinguish between these two classes. In a previous study by

Wetherley et al. (2018), the MESMA shade factor was set to 20% to address the variations in brightness between trees and turfgrass in the urban environment of Los Angeles, California [33]. However, we found that limiting the MESMA shade factor to 30% is more effective to separate the needleleaf from broadleaf vegetation.

Validation at the pixel level commonly involves using the pixel value from the centroid of the field plot [19]. However, in mixed and highly diverse boreal settings, this method is not suitable. Therefore, we calculated the pro-portion of each pixel within a plot and summed up all the proportions for the ground-based validation (Figure 4.6). Assessing the fraction map accuracy using traditional ground-based methods has several challenges. One of the main challenges was accurately aligning the ground plot with the corresponding pixel, which was particularly difficult in a dense forest area. Furthermore, a precise estimation of the proportion of different vegetation species during the field surveys at the 10–30 m scale was challenging due to the inherent nature of the forested areas. Therefore, we approximated the proportion of the different species on the ground and performed the validation (Figure 4.11). The approach had limitations and proved ineffective for the needleleaf fraction accuracy assessment. Given the limitations of this validation method, we developed and tested an alternative approach, which is explained in Section 4.4.3.

For a qualitative assessment, we visually compared the fraction cover output with the 50 cm resolution SkySat image data and observed that the fraction map captured the patterns of needleleaf vegetation distribution reasonably well. Additionally, for quantitative assessment, we collected a ~1000 m^2 plot, divided it into seven sub-plots (Figure 4.3c), and counted trees by species in each subplot. A similar approach was used by Fernández-García et al. (2021), where they used the high-resolution aerial orthophoto in place of field plots to establish plots of 30 m × 30 m, and subdivided each plot into 100 cells of 3 m × 3 m [19]. However, collecting similar in situ plot data in a dense forest setting poses a significant challenge. Due to the unavailability of aerial orthophotos, we collected ground data and divided the ~1000 m^2 plot into seven subplots, counting the trees in each subplot. The trees of the same species were similar in shape and size due to belonging the same age group. These data were used to validate the high-resolution hyperspectral product (HySpex: 1 m spatial resolution). Assuming that one pixel of HySpex (1 m) corresponds to one tree, we computed Pearson's correlation coefficient which was very high (0.9), implying that the classified output (validation map) is reliable and can be used to validate the fraction output.

We assessed the accuracy of the MESMA output at different spatial resolutions using random points and found no major differences in accuracy at different spatial resolutions. This

suggests that MESMA can be used for improved sub-pixel cover estimates from the current and upcoming global space-borne hyperspectral data of coarser resolution. Roth et al. (2015) published a similar study to assess the impact of spatial resolution on the plant functional type classification over five different ecosystems in the USA, including three forest ecosystems: the Smithsonian Environmental Research Center (Maryland), the Wind River Experimental Forest (Washington), and the Sierra National Forest (California), a site with tidal marsh in the Gulf Coast (Louisiana), as well as the central coast region of Santa Barbara, the Santa Ynez Mountains, and the Santa Ynez Valley (California), which feature diverse habitats ranging from evergreen and deciduous shrublands to open grasslands and woodlands. They performed pixel-level classification and suggested that the plant functional type classification will be efficient in the current and upcoming (30 m and 60 m) coarser resolution hyperspectral data [25]. Similarly, we found that there was no significant difference in the fraction outputs at different spatial scales (5 m, 10 m, and 30 m), suggesting that a spectral unmixing technique (MESMA) is effective in estimating the sub-pixel needleleaf fraction from coarser spatial resolution data.

This study presents an effective approach to map and validate the sub-pixel needleleaf fraction in a boreal forest to aid fire management. Our validation approach introduces a novel methodology that will benefit future research on spectral unmixing validation at the sub-pixel level. This intermediate step can serve as a valuable tool for validating very-coarse-resolution orbital data from sensors such as EMIT and SBG.

The fraction map provides detailed information on needleleaf vegetation, which is essential for fire management in all three stages (pre-fire, active fire, and post-fire). It can help land managers and firefighters identify the location of high-risk fuels and employ pre-fire management practices: during an active fire, it will help prioritize the area that needs immediate attention to reduce the fire spread; post-fire, these maps can also help in studying the forest demography changes, especially the forest recovery. While we validated the fraction map in a typical boreal ecosystem, this approach can be used to validate the MESMA output in other ecosystems. The requirement is for detailed field surveys and high-resolution hyperspectral data over the same region.

4.6. Conclusions

Needleleaf vegetation is a high-risk fire fuel and responsible for rapid fire spread and high burn intensity. This study presents an effective approach to quantify the needleleaf fraction in each pixel of an AVIRIS-NG image using a well-documented pixel unmixing algorithm (MESMA) and validating the fraction estimates. We developed an approach to validate a fraction map

product using a high-resolution classified map product and needleleaf tree counts from the field. We found that MESMA has the potential to map a needleleaf fraction in a mixed boreal forest with reasonable accuracy. We applied MESMA on AVIRIS-NG data at different spatial resolutions and found no major difference in ac-curacies suggesting that spectral unmixing is effective in estimating the needleleaf fraction from coarse-resolution data. Future research should focus on the different unmixing techniques and compare their performance. The findings from this study supports the applications of the current and upcoming hyperspectral space missions for sub-pixel vegetation and landcover mapping for a variety of applications including wildfire management and ecosystem monitoring.

4.7. Acknowledgments

Part of this research was carried out at the Jet Propulsion Laboratory, California Institute of Technology, under contract with the National Aeronautics and Space Administration. We would like to acknowledge the NASA Commercial Smallsat Data Acquisition (CSDA) Program for providing the SkySAT data used in this study. We extend our gratitude to Tanya Harrison (Planet Labs) for her assistance in obtaining the SkySat data. We would also like to thank Francisco Ochoa and Philip G Brodrick from JPL for their guidance in understanding the unmixing algorithm. Lastly, we express our heartfelt appreciation to Chris Smith, Brooke Kubby, Colleen Haan, Malvika Shriwas, Soumitra Sakhalkar, Glen Woodworth, Naomi Hutchens, Michelle Q and Josh Jones for their invaluable help during the field work.

Funding: This material is based upon work supported by the National Science Foundation under the award OIA-1757348, the State of Alaska and the U.S. Geological Survey under Grant/Cooperative agreement No. G18AP00077.

4.8. References

1. Bonanza Creek LTER. Available online: https://www.lter.uaf.edu/boreal-forest/climate (accessed on 31 December 2022).
2. Alaska Department of Fish and Wildlife. *Wildlife Action Plan Appendix 5: Forest Habitats*; Alaska Department of Fish and Wildlife: Juneau, AK, USA, 2005.
3. International Arctic Research Center. Alaska's Changing Wildfire Environment. Available online: https://uaf-iarc.org/alaskas-changing-wildfire-environment/ (accessed on 9 November 2021).

4. Bhatt, U.S.; Lader, R.T.; Walsh, J.E.; Bieniek, P.A.; Thoman, R.; Berman, M.; Borries-Strigle, C.; Bulock, K.; Chriest, J.; Hahn, M.; et al. Emerging Anthropogenic Influences on the Southcentral Alaska Temperature and Precipitation Extremes and Related Fires in 2019. *Land* 2021, *10*, 82.
5. Partain, J.L.; Alden, S.; Bhatt, U.S.; Bieniek, P.A.; Brettschneider, B.R.; Lader, R.T.; Olsson, P.Q.; Rupp, T.S.; Strader, H.; Thoman, R.L.; et al. An Assessment of the Role of Anthropogenic Climate Change in the Alaska Fire Season of 2015. *Bull. Am. Meteorol. Soc.* 2016, *97*, S14–S18.
6. Thoman, R.; Walsh, J.; Eicken, H.; Hartig, L.; Mccammon, M.; Bauer, N.; Carlo, N.; Rupp, S.; Buxbaum, T.; Bhatt, U.; et al. *Alaska's Changing Environment: Documenting Alaska's Physical and Biological Changes through Observations*; Review; University of Alaska Fairbanks: Fairbanks, AK, USA, 2019.
7. Bieniek, P.A.; Bhatt, U.S.; York, A.; Walsh, J.E.; Lader, R.; Strader, H.; Ziel, R.; Jandt, R.R.; Thoman, R.L. Lightning Variability in Dynamically Downscaled Simulations of Alaska's Present and Future Summer Climate. *J. Appl. Meteorol. Climatol.* 2020, *59*, 1139–1152.
8. Badola, A.; Panda, S.K.; Roberts, D.A.; Waigl, C.F.; Jandt, R.R.; Bhatt, U.S. A novel method to simulate AVIRIS-NG hyperspectral image from Sentinel-2 image for improved vegetation/wildfire fuel mapping, boreal Alaska. *Int. J. Appl. Earth Obs. Geoinf.* 2022, *112*, 102891.
9. Alaska Interagency Coordination Center. AICC—Incident Information—Alaska Large Fires. Available online: https://fire.ak.blm.gov/incinfo/aklgfire.php (accessed on 26 December 2022).
10. Zheng, B.; Ciais, P.; Chevallier, F.; Yang, H.; Canadell, J.G.; Chen, Y.; van der Velde, I.R.; Aben, I.; Chuvieco, E.; Davis, S.J.; et al. Record-high CO_2 emissions from boreal fires in 2021. *Science* 2023, *379*, 912–917.
11. National Park Service Fire in Ecosystems: Boreal Forest (U.S. National Park Service). Available online: https://www.nps.gov/articles/000/fire-in-ecosystems-boreal-forest.htm (accessed on 9 November 2021).
12. Dudley, K.L.; Dennison, P.E.; Roth, K.L.; Roberts, D.A.; Coates, A.R. A multi-temporal spectral library approach for mapping vegetation species across spatial and temporal phenological gradients. *Remote Sens. Environ.* 2015, *167*, 121–134.

13. Smith, C.W.; Panda, S.K.; Bhatt, U.S.; Meyer, F.J. Improved Boreal Forest Wildfire Fuel Type Mapping in Interior Alaska using AVIRIS-NG Hyperspectral data. *Remote Sens.* 2021, *13*, 897.
14. Xie, Y.; Sha, Z.; Yu, M. Remote sensing imagery in vegetation mapping: A review. *J. Plant Ecol.* 2008, *1*, 9–23.
15. Badola, A.; Panda, S.K.; Roberts, D.A.; Waigl, C.F.; Bhatt, U.S.; Smith, C.W.; Jandt, R.R. Hyperspectral Data Simulation (Sentinel-2 to AVIRIS-NG) for Improved Wildfire Fuel Mapping, Boreal Alaska. *Remote Sens.* 2021, *13*, 1693.
16. Badola, A.; Padalia, H.; Belgiu, M.; Verma, P.A. Tree Species Mapping in Tropical Forests Using Hyperspectral Remote Sensing and Machine Learning. In Proceedings of the 2021 IEEE International Geoscience and Remote Sensing Symposium IGARSS, Brussels, Belgium, 11–16 July 2021; pp. 5421–5424.
17. Somers, B.; Asner, G.P.; Tits, L.; Coppin, P. Endmember variability in Spectral Mixture Analysis: A review. *Remote Sens. Environ.* 2011, *115*, 1603–1616.
18. Roberts, D.A.; Gardner, M.; Church, R.; Ustin, S.; Scheer, G.; Green, R.O. Mapping Chaparral in the Santa Monica Mountains Using Multiple Endmember Spectral Mixture Models. *Remote Sens. Environ.* 1998, *65*, 267–279.
19. Fernández-García, V.; Marcos, E.; Fernández-Guisuraga, J.M.; Fernández-Manso, A.; Quintano, C.; Suárez-Seoane, S.; Calvo, L. Multiple Endmember Spectral Mixture Analysis (MESMA) Applied to the Study of Habitat Diversity in the Fine-Grained Landscapes of the Cantabrian Mountains. *Remote Sens.* 2021, *13*, 979.
20. Fernandez-Manso, A.; Quintano, C.; Roberts, D.A. Burn severity influence on post-fire vegetation cover resilience from Landsat MESMA fraction images time series in Mediterranean forest ecosystems. *Remote Sens. Environ.* 2016, *184*, 112–123.
21. Connelly, D.S.; Thompson, D.R.; Mahowald, N.M.; Li, L.; Carmon, N.; Okin, G.S.; Green, R.O. The EMIT mission information yield for mineral dust radiative forcing. *Remote Sens. Environ.* 2021, *258*, 112380.
22. Guanter, L.; Kaufmann, H.; Segl, K.; Foerster, S.; Rogass, C.; Chabrillat, S.; Kuester, T.; Hollstein, A.; Rossner, G.; Chlebek, C.; et al. The EnMAP Spaceborne Imaging Spectroscopy Mission for Earth Observation. *Remote Sens.* 2015, *7*, 8830–8857.
23. Loizzo, R.; Guarini, R.; Longo, F.; Scopa, T.; Formaro, R.; Facchinetti, C.; Varacalli, G. Prisma: The Italian hyperspectral mission. In Proceedings of the 2018 IEEE International Geoscience and Remote Sensing Symposium, Valencia, Spain, 22–27 July 2018; pp. 175–178.

24. Thompson, D.R.; Basilio, R.; Brosnan, I.; Cawse-Nicholson, K.; Chadwick, K.D.; Guild, L.; Gierach, M.; Green, R.O.; Hook, S.; Horner, S.D.; et al. Ongoing Progress Toward NASA's Surface Biology and Geology Mission. In Proceedings of the 2022 IEEE International Geoscience and Remote Sensing Symposium, Kuala Lumpur, Malaysia, 17–22 July 2022; pp. 5007–5010.
25. Roth, K.L.; Roberts, D.A.; Dennison, P.E.; Peterson, S.H.; Alonzo, M. The impact of spatial resolution on the classification of plant species and functional types within imaging spectrometer data. *Remote Sens. Environ.* 2015, *171*, 45–57.
26. USDA. Bonanza Creek Experimental Forest & Caribou-Poker Creeks Research Watershed|Pacific Northwest Research Station|PNW—US Forest Service. Available online: https://www.fs.usda.gov/pnw/experimental-forests-and-ranges/bonanza-creek-experimental-forest-caribou-poker-creeks-research (accessed on 31 December 2022).
27. NASA JPL AVIRIS-Next Generation. Available online: https://avirisng.jpl.nasa.gov/platform.html (accessed on 6 May 2023).
28. L3Harris Geospatial Solutions Preprocessing AVIRIS Data Tutorial. Available online: https://www.l3harrisgeospatial.com/docs/preprocessaviris.html (accessed on 25 February 2023).
29. Exelis Visual Information Solutions Inc. *Exelis Visual Information Solutions Version 5.3*; Exelis Visual Information Solutions Inc.: Boulder, CO, USA, 2010.
30. Roberts, D.A.; Halligan, K.; Dennison, P.; Dudley, K.; Somers, B.; Crabbe, A. *Viper Tools User Manual*; Version 2; VIPER Lab UC Santa Barbara: Santa Barbara, CA, USA, 2018; 91p.
31. Crabbé, A.H.; Somers, B.; Roberts, D.A.; Halligan, K.; Dennison, P.; Dudley, K. MESMA ENVI Plugin (Version 1.0.7). 2020. Available online: https://bitbucket.org/kul-reseco/mesma (accessed on 25 December 2022).
32. Dennison, P.E.; Roberts, D.A. Endmember selection for multiple endmember spectral mixture analysis using endmember average RMSE. *Remote Sens. Environ.* 2003, *87*, 123–135.
33. Wetherley, E.B.; McFadden, J.P.; Roberts, D.A. Megacity-scale analysis of urban vegetation temperatures. *Remote Sens. Environ.* 2018, *213*, 18–33.
34. Clasen, A.; Somers, B.; Pipkins, K.; Tits, L.; Segl, K.; Brell, M.; Kleinschmit, B.; Spengler, D.; Lausch, A.; Förster, M. Spectral Unmixing of Forest Crown Components at Close Range, Airborne and Simulated Sentinel-2 and EnMAP Spectral Imaging Scale. *Remote Sens.* 2015, *7*, 15361–15387.

35. Warmerdam, F.; Rouault, E. Others Gdalwarp—GDAL Documentation. Available online: https://gdal.org/programs/gdalwarp.html (accessed on 27 April 2023).
36. GISGeography. Raster Resampling for Discrete and Continuous Data—GIS Geography. Available online: https://gisgeography.com/raster-resampling/ (accessed on 27 April 2023).
37. NASA CSDA. Commercial Smallsat Data Acquisition (CSDA) Program|Earthdata. Available online: https://www.earthdata.nasa.gov/esds/csda (accessed on 24 January 2023).
38. QGIS. *Development Team QGIS Geographic Information System*; Version 3.14; Open Source Geospatial Foundation: Chicago, IL, USA, 2020.
39. Jurgiel, B. Point Sampling Tool Plugin. Available online: https://github.com/borysiasty/pointsamplingtool (accessed on 24 January 2023).
40. Cristóbal, J.; Graham, P.; Prakash, A.; Buchhorn, M.; Gens, R.; Guldager, N.; Bertram, M. Airborne Hyperspectral Data Acquisition and Processing in the Arctic: A Pilot Study Using the Hyspex Imaging Spectrometer for Wetland Mapping. *Remote Sens.* 2021, *13*, 1178.
41. Belgiu, M.; Drăguţ, L. Random forest in remote sensing: A review of applications and future directions. *ISPRS J. Photogramm. Remote Sens.* 2016, *114*, 24–31.
42. Dietterich, T.G. Approximate Statistical Tests for Comparing Supervised Classification Learning Algorithms. *Neural Comput.* 1998, *10*, 1895–1923.
43. Virtanen, P.; Gommers, R.; Oliphant, T.E.; Haberland, M.; Reddy, T.; Cournapeau, D.; Burovski, E.; Peterson, P.; Weckesser, W.; Bright, J.; et al. {SciPy} 1.0: Fundamental Algorithms for Scientific Computing in Python. *Nat. Methods* 2020, *17*, 261–272.
44. Powell, R.L.; Roberts, D.A.; Dennison, P.E.; Hess, L.L. Sub-pixel mapping of urban land cover using multiple endmember spectral mixture analysis: Manaus, Brazil. *Remote Sens. Environ.* 2006, *106*, 253–267.
45. Brewer, W.L.; Lippitt, C.L.; Lippitt, C.D.; Litvak, M.E. Assessing drought-induced change in a piñon-juniper woodland with Landsat: A multiple endmember spectral mixture analysis approach. *Int. J. Remote Sens.* 2017, *38*, 4156–4176.

Chapter 5 Summary and Conclusions

5.1. Summary

Wildfires are a natural phenomenon in boreal Alaska. Recently, Alaska has experienced increased frequency and size of wildfires, leading to increased burned areas. Efficient fire management strategies are crucial for ensuring the safety of the public, protecting property, and preserving the environment from the catastrophic consequences of large wildfires. For effective wildfire management, a comprehensive understanding of weather parameters, such as the likelihood of lightning strikes, seasonal climate variations, and fuel types, is essential to develop fire prevention, suppression, and post-fire restoration strategies. However, the existing fuel maps in Alaska lack sufficient details and suffer from poor accuracy, posing challenges for fire managers in effectively managing wildfires. Therefore, it is crucial to improve the fuel maps. This study aims to advance fuel mapping to aid fire management efforts in Alaska by conducting research and investigation to produce detailed fuel maps with greater accuracy.

Detailed maps are crucial for accurate and effective wildfire prevention and suppression efforts. In Alaska, the LANDFIRE Existing Vegetation Type (EVT) (LANDFIRE, 2016) map product derived from Landsat image data (multispectral data) at 30m spatial resolution is traditionally used for fire management, fire spread modeling, and risk assessment. The LANDFIRE EVT 2014 map product's accuracy ranges between 20% and 45% (Develice, 2012; Smith et al., 2021). Hyperspectral data can meet the need to improve fuel maps due to its narrow bandwidth. Hyperspectral data is essential for accurate and detailed vegetation mapping in Alaska (Smith et al., 2021). However, it is scarce due to the high data acquisition cost, so alternate methods need to be developed to create such a map. This research has developed an automated workflow that uses 10 m Sentinel scene scale data to generate hyperspectral data and vegetation maps. This process resulted in maps with an accuracy that was 33% higher than the LANDFIRE EVT maps, as assessed on a test site at BCEF where ground data was available. The study also demonstrated that the forest-type acreage estimates for the five major tree species obtained from the map follow the same pattern as the estimates from the USDA Pilot Inventory program for Tanana Valley State Forest (Pattison et al., 2018; U.S. Forest Service, 2016), with black spruce being the most dominant species, followed by birch, aspen, and poplar being the least dominant. These highly detailed vegetation/fuel maps are valuable input for fuel models and useful for wildfire management and decision-making.

The study utilized the Universal Pattern Decomposition Method (UPDM) to simulate AVIRIS-NG data using remotely sensed Sentinel-2 data. The data were then classified using a Random Forest model to generate vegetation/fuel maps. Additionally, the study presents a well-designed processing workflow for mapping needleleaf fractions in mixed forest stands since it is of high interest to fire managers due to its high flammability. The study employed the spectral unmixing algorithm, Multiple Endmember Spectral Mixture Analysis (MESMA), to map needleleaf fraction in a test area in interior Alaska. Previous studies have tested MESMA for mapping green vegetation, soil, and non-photosynthetic vegetation (Brewer et al., 2017; Powell et al., 2006; Roberts et al., 1998). However, this is the first study to test MESMA for mapping needleleaf fraction in a mixed boreal forest in Alaska. Also, a comprehensive workflow was devised to validate the fraction cover map using high-resolution hyperspectral data (1m) and ground data.

5.2. Conclusions

RQ1: Can high-quality hyperspectral data be generated from widely available Sentinel - 2 data to meet the need for greater spatial and temporal coverage of hyperspectral data for vegetation and fuel mapping at a local scale in boreal Alaska?

The study demonstrated the reliability of UPDM for spectral unmixing. This sensor-independent algorithm requires ground-measured spectra and spectral response functions from multispectral and hyperspectral sensors.

1. By utilizing UPDM, we successfully simulated 332 bands of AVIRIS-NG data from Sentinel-2 multispectral data. The resulting hyperspectral data exhibited the spectral characteristics of AVIRIS-NG and the spatial resolution of Sentinel-2 data.
2. On visual comparison of simulated, AVIRIS-NG, and Sentinel-2 images, we found that the simulated image was able to capture the minute details and features present in the AVIRIS-NG image.
3. High band-to-band correlation between simulated and AVIRIS-NG data resulting in a strong statistical correlation between the spectral signature extracted from them for the major tree species (birch and spruce), with correlation coefficients (r) of 0.97 for birch and 0.92 for spruce.
4. On performing species-level classification, we were able to achieve an 89% accuracy for the simulated hyperspectral data, which surpassed the accuracy of Sentinel-2 data (77.8%). Our findings are consistent with previous studies by Liu et al. (2009) and Tiwari et al. (2016), who also reported significant improvements in classification accuracy when using simulated Hyperion data.

The results indicate that the simulation process successfully generated hyperspectral data that closely resembles an AVIRIS-NG image, which can be used to generate improved vegetation and land cover maps for various applications. The ability to generate such data sets is particularly valuable in areas where acquiring real hyperspectral imagery is expensive and logistically challenging.

RQ2: How effectively can the simulated Hyperspectral data be used for the identification of vegetation/fuel classes of the boreal region of Alaska at a regional scale?

This study presents a novel approach for generating hyperspectral images from commonly available multispectral images at the Sentinel-2 scene scale (i.e., 100 km x 100 km). The results demonstrate the effectiveness of this approach for detailed and improved vegetation and fuel mapping at a regional scale in Alaska's boreal forest, contributing to the effort for effective forests and fire management in the region.

1. The study identified birch, spruce, and gravel as the most suitable endmembers for the study area. We collected ground spectra of 13 major vegetation (including shrubs and trees), gravel, and asphalt during summers (2019–2021) using a PSR+ 3500 hand-held spectroradiometer and created a spectral library for this study. We employed the Iterative Endmember Selection (IES) algorithm to optimize the spectral library and select the most representative endmembers for simulation.
2. The approach addresses the paucity of hyperspectral data in Alaska. We automated and optimized algorithms so that a 425-band AVIRIS-NG hyperspectral image cube covering an area of 100 km x 100 km can be generated in 2 hours.
3. The results showed RMSE of 0.03 and 0.02 for two major vegetation classes, birch and spruce, respectively, indicating a high level of spectral similarity between the simulated and AVIRIS-NG images.
4. The study developed a Random Forest image classification model using training data from one scene and tested its portability on two other scenes from different sub-ecoregions. The model performed well in these two scenes, indicating its reliability and reproducibility across space in boreal Alaska.
5. As per the assessment using available field data, the vegetation maps generated from simulated data were more detailed and accurate compared to the existing maps (LANDFIRE EVT and Alaska Vegetation and Wetland Composite (AKVWC)). The accuracy of the vegetation map generated using simulated data was 65%, and the LANDFIRE EVT map was 32%, suggesting an improvement of 33% in accuracy.

6. The study presents an automated workflow for the hyperspectral image simulation and vegetation classification in Google's cloud platform (GCP system configuration: 16 CPUs and 60 GB RAM). By improving and optimizing the simulation and classification algorithms, we were able to transform a 4.5 GB Sentinel-2 (100 X100 km) dataset into a complete hyperspectral image cube of 190 GB in two hours and generate a vegetation map in 6 hours.

RQ3: Can spectral unmixing of a pixel estimate the needleleaf fraction in a mixed boreal forest with reasonable accuracy, and how do we validate needleleaf fraction estimates?

The study successfully mapped needleleaf fraction using MESMA, a widely used spectral unmixing algorithm. However, validation of the resulting fraction map was a major challenge. To address this issue, we proposed a novel approach to validate the needleleaf fraction map.

1. Needleleaf species in mixed boreal forests can be effectively mapped with reasonable accuracy by adjusting the brightness parameter in the MESMA algorithm.
2. We developed a novel approach to validate a fraction map product using a high-resolution classified map product generated using a 1m spatial resolution HySpex image and field-based needleleaf tree counts.
3. There was no significant difference in the accuracy of needleleaf fraction maps at different spatial resolutions (73%, 79%, and 78% for 5 m, 10m, and 30 m spatial resolution images, respectively). Hence, the findings suggest that MESMA can potentially be a valuable tool for needleleaf fraction mapping at a regional scale using coarser resolution space-borne hyperspectral data from current and upcoming missions.

5.3. Broader impacts

The research produces three significant products: simulated hyperspectral data and vegetation/fuel maps for the entire boreal region of Alaska and a needleleaf fraction cover map for a test site in interior Alaska. These products are important for wildfire management practices and extend their impact beyond fire management. Further details on the broader implications and knowledge/product sharing are discussed in the following paragraphs.

The availability of Hyperspectral data for Alaska has multiple benefits: monitoring land use changes, assessing wildfire impacts, mapping invasive species, and estimating biomass and carbon stocks. It paves the way for several other arctic applications, including sea ice extent mapping and monitoring permafrost degradation. Additionally, hyperspectral data can aid in mineral exploration and mapping, providing valuable information for mining and resource extraction industries. The availability of hyperspectral data can lead to a better understanding of

Alaska's environment and support sustainable development and management of its natural resources while advancing scientific research and decision-making.

Vegetation/fuel maps play a critical role as input to fuel models such as the Forest Vegetation Simulator and the Fire Dynamics Simulator. These models can be used to simulate future forest and fire disturbance conditions based on the fuel type, moisture content, and other variables, including topography and climate. The accuracy and reliability of the models depend heavily on the quality and resolution of the vegetation/fuel maps. At the wildland-urban interface, the detailed maps can assist in identifying areas at a higher risk of fire and determining appropriate fire management strategies such as prescribed burns, fuel reduction treatments, and fire suppression efforts to reduce the amount of available fuel and minimize the risk of fire spread.

The new vegetation maps can be accessed through the ArcGIS web portal, embedded in the story map available at https://tinyurl.com/HySim-storymap. This story map provides essential information on the method used for hyperspectral data simulation and classification. To minimize access barriers and enhance the self-sufficiency of fire managers, we are sharing Python codes on GitHub at https://github.com/abadola21/hysim. These codes will empower fire and land managers by equipping them with the knowledge and resources to independently perform simulations and generate vegetation maps to make informed decisions. We have also planned to organize a workshop, with the help of the Alaska Fire Science Consortium (AFSC), on running simulations and classification using Google Colab, an open-source platform. This workshop aims to provide fire and land managers with hands-on experience and practical guidance. The workshop recording will be publicly available through the published story map, allowing anyone to access and benefit from the technology.

The study also involved the collection of ground spectra of 13 major vegetation types, including shrubs and trees, and samples of gravel and asphalt, using a PSR+ 3500 hand-held spectroradiometer during the summers of 2019 to 2021. The spectral library and the ground data that includes vegetation cover and location data for interior Alaska are available at https://github.com/abadola21/hysim/tree/main/SpectralLib_GroundData. Ground data and the comprehensive spectral library generated from this study are valuable resources for future research in the boreal domain of Alaska.

Engaging in discussions with fire and land managers, in collaboration with the AFSC, has provided valuable insights into the critical importance of needleleaf fraction mapping for effective fire management. These insights have motivated us to develop a reproducible data processing workflow for needleleaf fraction mapping in interior Alaska. The needleleaf fraction map and the

ground data for this study are available at https://tinyurl.com/ycxnns7t. Needleleaf fraction cover maps are particularly useful for modeling fire spread and predicting fire behavior to support effective fire suppression strategies. For example, when a wildfire is approaching an area with a high needleleaf species, managers may deploy more resources and take more aggressive suppression actions, such as creating firebreaks or conducting controlled burns, to prevent the fire from spreading. Moreover, needleleaf fraction cover maps can aid in identifying areas with high biomass and carbon stocks, which can be used to prioritize conservation efforts and support programs aimed at offsetting carbon emissions. The needleleaf fraction maps are also crucial for Alaska's forestry industry. The map can be used to identify areas with high potential for timber harvesting and aid in forest resource management and planning.

5.4. Future work and recommendations

Hyperspectral data is highly valuable for detailed studies, but its limited availability has constrained its use. However, with the advent of simulated hyperspectral data, new possibilities for research and collaborations have emerged. In a collaborative effort, we worked with researchers from the Indian Institute of Technology, Varanasi, India, to study the chlorophyll content in agricultural crops in India. Our methodology was successfully implemented, leading to the joint publication of a paper (Verma et al., 2022). The study demonstrated the method's potential in simulating AVIRIS-NG data to estimate chlorophyll levels in crops.

The study has shown the advantages of utilizing a pixel unmixing approach and a Random Forest classification model for improved vegetation/fuel mapping. However, there is potential to enhance the quality of these maps further. One effective strategy is incorporating more training data into the classification model, including data from different seasons and regions to capture the maximum variability in the spectral signatures over space and time. In addition to increasing the diversity and quantity of training data, incorporating more advanced machine learning algorithms, such as deep learning algorithms, has the potential to further improve the accuracy and precision of vegetation/fuel mapping.

The study demonstrates the potential of MESMA in mapping needleleaf fraction cover across varying spatial resolutions, from high to low. As current and upcoming space-borne hyperspectral satellite missions (e.g., EMIT, EnMap, PRISMA, and SBG) increase the availability of hyperspectral data at coarser resolutions, it creates an opportunity to scale up needleleaf maps at a continental scale. A significant contribution of this research is developing a robust validation method for needleleaf fraction estimates, which can serve as a valuable tool for validating products generated by these hyperspectral space missions. The utilization of MESMA for

mapping needleleaf fraction in the study area has shown promising results. However, future research should explore different unmixing techniques and compare their performance with MESMA.

5.5. References

Brewer, W. L., Lippitt, C. L., Lippitt, C. D., & Litvak, M. E. (2017). Assessing drought-induced change in a piñon-juniper woodland with Landsat: a multiple endmember spectral mixture analysis approach. *International Journal of Remote Sensing, 38(14),* 4156–4176. https://doi.org/10.1080/01431161.2017.1317940

DeVelice, R.L. (2012) Accuracy of the LANDFIRE Alaska Existing Vegetation Map over the Chugach National Forest; LANDFIRE Assessments. Volume 10. https://landfire.cr.usgs.gov/documents/LANDFIRE_ak_110evt_accuracy_summary_013012.pdf

LANDFIRE. (2016). LANDFIRE: Existing Vegetation Type. Retrieved February 10, 2021, from U.S. Department of Agriculture and U.S. Department of the Interior website: http://www.landfire.gov

Liu, B., Zhang, L., Zhang, X., Zhang, B., & Tong, Q. (2009). Simulation of EO-1 Hyperion Data from ALI Multispectral Data Based on the Spectral Reconstruction Approach. *Sensors*, *9*(4), 3090–3108. https://doi.org/10.3390/s90403090

Pattison, Robert; Andersen, Hans-Erik; Gray, Andrew; Schulz, Bethany; Smith, Robert J.; Jovan, Sarah, tech. coords. (2018). *Forests of the Tanana Valley State Forest and Tetlin National Wildlife Refuge, Alaska: results of the 2014 pilot inventory. Gen. Tech. Rep. PNW-GTR-967.* Retrieved from https://accscatalog.uaa.alaska.edu/dataset/alaska-vegetation-and-wetland-composite

Powell, R. L., Roberts, D. A., Dennison, P. E., & Hess, L. L. (2007). Sub-pixel mapping of urban land cover using multiple endmember spectral mixture analysis: Manaus, Brazil. *Remote Sensing of Environment, 106(2),* 253-267. https://doi.org/10.1016/j.rse.2006.09.005

Roberts, D. A., Gardner, M., Church, R., Ustin, S., Scheer, G., & Green, R. O. (1998). Mapping Chaparral in the Santa Monica Mountains Using Multiple Endmember Spectral Mixture Models. *Remote Sensing of Environment*, *65*(3), 267–279. https://doi.org/10.1016/S0034-4257(98)00037-6

Smith, C. W., Panda, S. K., Bhatt, U. S., & Meyer, F. J. (2021). Improved Boreal Forest Wildfire Fuel Type Mapping in Interior Alaska using AVIRIS-NG Hyperspectral data. *Remote Sensing*, *13*(5), 897. https://doi.org/https://doi.org/10.3390/rs13050897

Tiwari, V., Kumar, V., Pandey, K., Ranade, R., & Agrawal, S. (2016, July 10-15). *Simulation of the hyperspectral data using Multispectral data* [Paper Presentation]. 2016 International Geoscience and Remote Sensing Symposium (IGARSS), Beijing, China 6157–6160. https://doi.org/10.1109/IGARSS.2016.7730608

U.S. Forest Service. (2016). PNW-FIA Interior Alaska Inventory | Pacific Northwest Research Station | PNW - US Forest Service. Retrieved December 12, 2021, from https://www.fs.usda.gov/pnw/projects/pnw-fia-interior-alaska-inventory

Verma, B., Prasad, R., Srivastava, P. K., Singh, P., Badola, A., & Sharma, J. (2022). Evaluation of Simulated AVIRIS-NG Imagery Using a Spectral Reconstruction Method for the Retrieval of Leaf Chlorophyll Content. *Remote Sensing, 14*(15), 3560. https://doi.org/10.3390/RS14153560