Applications of Imaging Spectroscopy in Forest Ecosystems at Multiple Scales

Beth R. Stein

Dissertation submitted to the faculty of the Virginia Polytechnic Institute and State University in partial fulfillment of the requirements for the degree of

Doctor of Philosophy
In
Forestry

Valerie A. Thomas, Chair
Randolph H. Wynne
Philip J. Radtke
Mary E. Martin

August 24, 2015
Blacksburg, VA

Keywords: imaging spectroscopy, nutrients, invasive plants, species diversity

Applications of Imaging Spectroscopy in Forest Ecosystems at Multiple Scales

Beth R. Stein

ABSTRACT

Forests provide a number of ecosystem services which sustain and enrich the wildlife, human societies, and the environment. However, many disturbances threaten forest ecosystems, making it necessary to monitor their health for optimal management and conservation. Although there are many indicators of forest health, changes in biogeochemical cycling, loss of species diversity, and invasive plants are particularly useful due to their vulnerability to the effects of climate change and intensive agricultural land use. Thus, this work evaluates the use of imaging spectroscopy to monitor forest nutrient status, species diversity, and plant invasions in the Mid-Atlantic region. The research is divided into four separate studies, each of which evaluated a unique application for imaging spectroscopy data at a different scale within the forest.

The first two studies examined loblolly pine nutrient status at the leaf and canopy scales, respectively. The first study determined that loblolly pine foliar macronutrient concentrations can be successfully modeled across the Southeastern US (R^2=0.39-0.74). Following on these results, the second study focused on the relationship between physical characteristics, reflectance, and nutrients. Reflectance values and W scattering coefficients produced successful nitrogen models across loblolly pine plots at the canopy scale. Regression models showed similar explanatory power for nitrogen, although W scattering coefficients were significantly correlated with nitrogen at multiple wavelengths and reflectance variables were not. However, the direction of some of the correlations with W and the unusually high directional area scattering factor values indicate a need for further experimentation. The third study found that several imaging spectroscopy algorithms were moderately successful in identifying wavyleaf basketgrass invasions in mixed deciduous forests (overall accuracy=0.35-0.78; kappa=0.41-0.53). Lastly, the fourth study used a novel imaging spectroscopy/lidar fusion to identify canopy gaps and measure species diversity of understory vegetation. The lidar algorithm identified 29 of 34 canopy gaps, and regression models explained 49 percent of the variance in gap species diversity. In conclusion, imaging spectroscopy can be used to evaluate ecosystem health through forest nutrient status, nitrogen models, species diversity estimates, and identification of invasive plant species.

Acknowledgements

There are so many people who I wish to acknowledge for their assistance and support through the course of this project. First and foremost, my committee—Randy Wynne, Phil Radtke, and Mary Martin—has helped at every stage of the way by offering their advice and expertise to advance my research and enable my growth as a scientist. My deepest thanks go to Valerie Thomas, my advisor and mentor. None of this would have been possible without her encouragement, knowledge, and most importantly, common sense. On numerous occasions, I frantically rushed to my advisor with insurmountable problems—only to leave the meeting with the problem resolved, my optimism renewed, and ready to face the next challenge. For her calmness, patience, and willingness to hear me out on my crazy ideas over the course of the past 5 years, I would like to say "thank you."

I would also like to thank Les Fuller, Sue Snow, and Stacey Kuhar for their technical and administrative help which enabled me to glide easily through the logistical hoops and focus on my research. Their hard work and helpfulness is greatly appreciated. Similarly, Janaki Alavalapati provided support and encouragement during my graduate studies, and Dean Stauffer gave me guidance and useful feedback throughout my MS research.

My sincere appreciation is also due to the people who helped all the details of this research come together. I would specifically like to thank Ranjith Gopalakrishan, Marco Yanez, David Britt, Matthew Sumnall, Matt Dempsey, Joel Stein, and Adam Epstein for their help in the field. Katie Britt deserves special acknowledgement for her help in many aspects of this project, including research development, field and labwork, image analysis, and good humor. I would also like to thank our colleagues from the NASA and ESA Joint Campaign who generously provided their data and allowed us to participate in their research study. The experience was

invaluable, and I greatly enjoyed the opportunity to collaborate. Thanks to Bruce Cook, Elizabeth Middleton, Petya Campbell, Fred Huemmerich, Lawrence Ong, Francisco Pinto, Tomasso Julitta, Micol Rossini, and many others in the group. Special thanks goes to Sara Childs at Duke University and Asko Noormets at North Carolina State University, who assisted with field arrangements and logistics. In addition, for their help with data analysis and invaluable wisdom, thanks to John Seiler, John Peterson, David Mitchem, Jay Raymond, Laura Lorentz, Tom Wieboldt, Katie Trozzo, Amy Werner, Evan Brooks, Will Guthrie, and LISA Statistical Consulting.

And finally, thank you to all my friends, colleagues, and professors who enriched my time at Virginia Tech both academically and socially. Similarly, my family has been behind me every step of the way, and it is to them who I dedicate this work.

Table of Contents

List of Figures

List of Tables

1 Chapter 1—Introduction and Objectives

Introduction

Definition and Importance of Forest Health

Forest ecosystem health refers to the ability of the forest to conduct natural processes that are characteristic of a forest. Forests are expected to produce plants, store carbon, and cycle nutrients in rates and quantities that are similar to other healthy forests (Naeem et al. 1999). Ecosystems under stress may be prone to lower primary productivity; nutrient depletion; reductions in biodiversity; unstable populations of key species; reversal of successional patterns; and increases in disease (Rapport 1992). Although the validity of the "ecosystem health" concept is controversial due to the lack of a steady state within ecosystems, the concept is particularly useful in management contexts. The notion of a healthy forest offers a standard to strive for through management activities and to compare against current conditions through monitoring efforts (O'Laughlin et al. 1994).

Forest health is generally determined through an assessment of the overall ecosystem functioning, generally across multiple criteria. Commissioned in 1993, the Montreal Process Criteria and Indicators are the internationally recognized standards for forest health. The working group established the concepts of criteria and indicators, where the former are defined as preservation values and the latter represent "measurable aspects" to evaluate these criteria. The selected criteria are as follows: "(1) conservation of biological diversity; (2) maintenance of the productive capacity of forest ecosystems; (3) maintenance of forest ecosystem health and vitality; (4) conservation and maintenance of soil and water resources; (5) maintenance of forest contributions to global carbon cycles; (6) maintenance and enhancement of long-term socio-economic benefits; and (7) the legal, institutional and economic framework for forest

conservation and sustainable management" (Coulombe 1995; Woodall et al. 2011). These criteria helped shape forest monitoring protocols in the US. The US Forest Service developed a number of different indicators to assess these criteria in their sampling protocols, including diversity of vegetation, crown properties, down woody materials, lichen communities, ozone effects, and soil characteristics (Woodall et al. 2011). Other countries have adopted similar indicators to assess in their monitoring programs (Gillis, Omule, & Brierley 2005; Woodall et al. 2009).

Of the numerous indicators to evaluate the forest health criteria, three attributes are particularly important due to their vulnerability to the effects of climate change and intensive agricultural land use. These indicators are changes in biogeochemical cycling, loss of species diversity, and invasive plants (Chapin et al. 2000). There is potential to use remote sensing to monitor the status of each of these indicators. This dissertation investigates remote sensing of these indicators at multiple scales. Consequently, the following sections will describe these three critical issues of concern and explain their ability to gauge forest health in the face of climate change and soil degradation.

Selected Ecological Concerns

1.1.1.1 Nutrient Status

Nutrients are chemical elements necessary for the nourishment of organisms. Trees require multiple nutrients in order to grow and meet their physiological demands. These elements are divided into the primary nutrients--nitrogen (N), phosphorus (P), and potassium (K)—and the secondary nutrients—calcium (Ca), magnesium (Mg), and sulfur (S). Elements fulfill a variety of necessary plant functions, such as photosynthesis, respiration, cell division, and cell wall

production (Epstein & Bloom 2004; Taiz & Zeiger 2010). Although necessary quantities vary by nutrient based on species and site-specific needs, insufficient amounts can lead to reduced growth and health (Pallardy 2007).

Long-term agriculture and poor harvesting practices have depleted soil nutrients around the world (Dupouey et al. 2002; Flinn & Vellend 2005). In the US, secondary forests are often located in those lands where nutrients are too poor for agriculture. Many soils in the Southeast show signs of nutrient depletion, particularly with respect to nitrogen (Albaugh et al. 2010). Furthermore, global climate models predict increased rainfall, higher rates of evaporation, and a greater frequency of droughts in the region (Bracho et al. 2012). These changes in precipitation could affect forest nutrients in several ways. Potential effects include reductions in nitrogen uptake and leaching due to greater drought; slower decomposition rates; and changes in soil and plant responses to fertilization (Zhang et al. 2013).

1.1.1.2 Understory Species Diversity

The role of plant diversity in ecosystem functioning has been hotly debated over the years by both experimental and theoretical biologists (Hooper et al. 2005; Huston 1997; Isbell et al. 2011). Despite a clear consensus whether few or many species are necessary, there are many benefits of a diversity of species for ecosystem services (Isbell et al. 2011). The understory, in particular, has an important role in the forest ecosystem. Understory species diversity is important for ecosystem functioning, such as nutrient cycling, overstory feedbacks, forest regeneration, and energy flow (Gilliam 2007; Kooch et al. 2012; Roberts 2004). For humans and wildlife, understory plants provide ecosystem services such as food, material resources (e.g. medicines, dyes), and landscape aesthetics. They also serve as indicators of forest health and

condition, such as invasive plants, underlying soil nutrients, forest recovery after a disturbance (Duffy & Meier 1992). Herbaceous plant diversity may also be indicative of underlying soil properties, particularly nitrogen availability (Gilliam 2007; Small & McCarthy 2005).

In some ecosystems, there is evidence that increases in species diversity greatly increased productivity and biomass (Tilman, Reich, & Isbell 2012). This relationship, known as the diversity-productivity relationship, is based on the niche complementarity hypothesis that purports that greater diversity leads to more productivity as a result of greater resource and nutrient use. However, studies have reported mixed results regarding herbaceous communities in forests (Adler et al. 2011; Zhang, Chen, & Reich 2012). In general, though, species loss has major effects on productivity and decomposition (Hooper et al. 2012).

Species diversity is affected by shifting species distributions due to changes in resources, community dynamics, and disturbance regimes that result from climate change. In addition, land use also has a significant effect on the composition and richness of communities. Understory plant communities in lands cleared for agriculture or resource extraction have lower species diversity and different composition than those in primary forests. These effects can be seen centuries, and even millennia, after the changes in land cover occurred (Flinn & Vellend 2005). Conversion of forests to agriculture can improve or decrease plant functionality. In many sites, a reduction in habitat connectivity and area prevents species dispersals and limits their distribution. Forest herbaceous plants may take centuries to recover (Flinn & Vellend 2005).

1.1.1.3 Invasive Plants

Invasive plants have the potential to affect all levels of ecology: individuals, plant species and communities, animal species and communities, and ecosystems. They influence basic

processes such as fitness, growth, diversity and abundance, nutrient pools, and soil properties. These effects occur both directly and indirectly through interactive feedbacks and cycles (Vilà et al. 2011). Consequently, invasive plants are themselves an agent of great change, capable of causing changes in species distributions and nutrient availability (Lockwood, Hoopes, & Marchetti 2013). Invasive plants have potential to affect population dynamics of other species, which in turn can affect the entire plant community and inter-specific ecosystem dynamics (Vilà et al. 2011).

While causes of plant introductions run the gamut from the intentional transport by humans to the accidental or biological, they are mediated by underlying ecological conditions that either encourage or discourage the species to survive in their new environment (Lockwood et al. 2013). Any species can be introduced, but a successful invasion requires adaptation and dispersal in the new environment. Habitats can be more or less susceptible to plant invasions, depending on factors such as disturbance, resource availability, and stability. Changing conditions in soil nutrients and climate may open the door for new species to thrive and take over new niches (Sardans & Peñuelas 2012).

Potential Applications of Remote Sensing

While each of these issues can be most accurately identified on the ground, they often occur at landscape-level scales that necessitate large areal mapping efforts. To conduct this type of analysis, remote sensing methods offer great potential. For several decades, remote sensing has been routinely employed for Earth observation and ecological analysis. These observations provide valuable information on the spatial and temporal dynamics of the world's ecosystems with high levels of detail (Cohen & Goward 2004; Gökkaya 2012).

Remote sensing is the ability to obtain information about a surface from a distance through the measurement of electromagnetic radiation. In the case of Earth observation, the surface of interest is often an area on the Earth being studied from an airborne or satellite-based sensor. Sensors can be classified into passive and active technology, depending on their role in the transmittance of electromagnetic radiation. Active sensors include radar and lidar, which provide their own energy source to transmit to the target. Alternatively, passive sensors rely on an external source of radiation (typically the sun) and receive the incoming reflectance. They can be further classified according to the type of radiation (i.e. visible, infrared, microwave, or radiowave) and the spectral resolution. Multispectral or broadband sensors, such as Landsat, receive radiation in multiple broad ranges of the spectrum. On the other hand, narrowband sensors measure radiation in many bands of very high spectral resolution. This type of remote sensing is known as imaging spectroscopy, and will be the focus of much of this work.

Imaging spectroscopy is the use of spectral properties to study the chemical and physical composition of the landscape across the geographic span of an image (Green et al. 1998). Imaging spectroscopy has been used to examine a multitude of ecological issues, including ecosystem health, land cover change, and disturbance. One of the benefits of imaging spectroscopy over conventional broadband sensors is its ability to discern very fine resolution differences in visible and infrared radiation across space and time. The sensors can identify details that extend much beyond our capacity to see with the human eye, both in terms of the spectral range and resolution. As a result, in the field of vegetation analysis, they are particularly informative for nutrient/moisture concentrations and stress diagnostics; species discrimination; and biophysical measurements (Thenkabail et al. 2014). Imaging spectrometers can measure foliar chemical constituents, elements, and nutrients. Laboratory measurements of foliage show

unique reflectance patterns in the visible and near-infrared wavelengths from different elements and compounds (Curran 1989). Thus, researchers can use these spectral signatures to describe leaf composition, properties, and chemical concentrations. Nitrogen has been studied in numerous species, and other nutrients have also gained attention (Al-Abbas et al. 1974; Ferwerda & Skidmore 2007; Peterson et al. 1988). The introduction of partial least squares regression as a statistical approach for spectroscopy brought about a new wave of research that better utilized the entire spectrum in nutrient-reflectance relationships (Martin & Aber 1997).

Imaging spectroscopy is also commonly used to discriminate and classify individual plant species through the use of basic statistical techniques. Once species are identified, methods can be used to map species abundance and distributions across the landscape. With the use of more advanced methods, multiple species can be discerned and species richness can be estimated. The "spectral variation" hypothesis purports that greater diversity in reflectance measurements can be related to higher species richness (Rocchini et al. 2010). The underlying theory is that spectral diversity serves as a proxy for habitat heterogeneity, which may be evidence of species diversity or a greater number of niches and, thus, species. Support for this hypothesis comes from recent studies across several ecosystems, including wetlands (Heumann, Hackett, & Monfils 2015) and rainforests in Hawaii (Carlson et al. 2007) and Peru (Féret & Asner 2014).

Invasive plant detection through imaging spectroscopy employs many of the same principles and techniques as species discrimination (Brown & Noble 2005; He et al. 2011; Lawrence, Wood, & Sheley 2006). Classification and unmixing algorithms are used to identify spectral components based on training data or user-specified endmembers. Classification methods operate at the pixel level to "match" the reflectance spectrum with the closest reference spectrum, or to group pixels based on similar reflectance properties. Unmixing techniques are

based on the premise that each pixel is composed of multiple endmembers, and the goal is to determine the relative fraction of each material within the pixels. Some methods produce "rule" images along with a classification map of the dominant pixel types. Overall accuracy varies depends on many factors, including plant characteristics, the nature of the invasion, ecosystem complexity, image resolution, and timing (Dehaan et al. 2007). Greater success has been attributed to unique physical characteristics in the invasive plant that enable its detection, imagery with sufficient spectral and spatial resolution to "see" the invasive plant, large spatial coverage of the invasion, and few obstructions between the plant and the sensor. Often this depends on plant phenology and appropriate image timing, as well.

Focus and Objectives

Objectives

This dissertation evaluates the use of remote sensing to monitor forest nutrient status, species diversity, and plant invasions in the Mid-Atlantic region. These attributes were selected for two main reasons. First, these characteristics face major impacts from global climate and land use change, so an ability to monitor their present and future condition remains critical to assessing forest health around the world. Second, these attributes are representative of multiple scales within the nested hierarchy of the ecosystem. They reflect a broad range of interactions both within and between species. As a result, the current state of each attribute provides a snapshot of ecosystem health and functioning from a different perspective and hierarchical level. I wanted to evaluate the ability to use remote sensing to study ecological concerns at multiple scales, in a similar manner to field-based assessments. The main research objectives are as follows:

1. **Objective 1:** Develop loblolly pine macronutrient models across local and regional scales to determine the relationship between spectral reflectance and nutrient concentrations. A secondary goal is to evaluate the importance of geographic scale and fertilization treatments in model development.
2. **Objective 2:** Investigate the relationship between physical characteristics, reflectance, and nutrients within loblolly pine at the canopy-scale. Based on these results, the secondary goal is to compare nitrogen models derived from reflectance and W scattering coefficients, which are the structurally-corrected component of the bidirectional reflectance factor.
3. **Objective 3:** Evaluate the use of remote sensing for species discrimination and mapping of an understory invader, wavyleaf basketgrass.
4. **Objective 4:** Determine the potential use of imaging spectroscopy/lidar fusion to measure species diversity across the forest understory.

One chapter is dedicated to each objective.

Forest Ecosystems in the Mid-Atlantic Region

For the purposes of the forthcoming analysis, we are defining the "Mid-Atlantic region" as including Maryland, Virginia, and North Carolina. Although certain chapters may venture into other parts of the country (notably, Chapter 2 and Southeastern US), the focus of this work is on forested ecosystems within Maryland, Virginia, and North Carolina. We would venture so far as to propose that this work may represent types of analysis and modeling approaches that could be done in the surrounding areas, as well, particularly in similar ecosystems.

These forests are composed of a mixture of "natural" and planted forests. The planted forests serve a variety of functions, including parks and recreation, experimental forests, riparian buffers, and timber production. Many of these forests had been cleared for agriculture in the early days of North American settlement, and remained in use until recent years. In many of the forests we studied, agriculture was a recent phenomenon. For example, both Duke Forest and Patapsco Valley State Park were reforested and set aside for their current land use in the early-to-mid 20th century (Maryland Department of Natural Resources 2014). Thus, effects of their historical legacies may still carry over into the current day.

Various Scales of Analysis

The term "scale" can encompass many definitions. This work examines forest ecosystems across various ecological scales: leaves and foliar constituents (suborganism), an individual tree canopy (organism), an understory invasive plant species (population), and the diversity of understory plants (community). These units of analysis are not inclusive of the entire scale; for instance, only a small geographic area of the wavyleaf basketgrass population in Patapsco Valley State Park is studied. However, they are still representative of the particular level of ecological scale (suborganism-organism-species-community) under investigation.

The second chapter begins the analysis by exploring nutrient status within an individual species. The chapter explores intra-specific nutrient variation across a geographic gradient, fertilization treatments, and even within the trees themselves, through various leaf flushes and canopy levels (upper and lower). Since models are drawn between trees, the largest unit of analysis is the individual tree in all cases; nutrient data is never aggregated across trees. In the third chapter, the analysis also occurs at the tree-level, but this time at the canopy scale. Due to

variations in tree size and the use of an airborne sensor, the plots no longer represent leaves and trees as discretized units; instead, each pixel is an amalgamation of one or more trees, various branching patterns, soil and understory background, remnant atmospheric effects, and other factors. The level of complexity grows in the fourth chapter, where subpixel algorithms are used to disentangle an understory invasive plant from the soil background and overstory forest canopy, which may dominate the reflectance signature. In this chapter, the focus remains on an individual species, but there is now no attempt to examine individual plants. Due to the spatial scale of the imagery and the plant size, this chapter attempts to benefit from the close intermingling of plants within the wavyleaf basketgrass population. Finally, the fifth chapter looks at the understory vegetation community based on species richness.

Scale also varies in terms of the spatial and temporal resolution of the sensor technology. For the analysis in the second chapter, I used a handheld spectrometer. The third and fifth chapters involved imaging spectroscopy and lidar data from Goddard's Lidar Hyperspectral and Thermal Imager (G-LiHT) and HyPlant sensors, which were flown simultaneously (Cook et al. 2013; Rossini et al. 2015). Meanwhile, the Hyperion satellite was used in the fourth chapter. These sensors fly at various heights, have different imaging specifications, and are even in different stages of development. Consequently, each sensor has its own issues and required separate workflows. With the ASD spectrometer, we were responsible for conducting the white reference calibrations, properly positioning the sensor, setting the measurement specifications, and smoothing the reflectance curves. Measurements could be collected at any time interval of our choosing. In theory, the airborne nature of G-LiHT and HyPlant offered us greater temporal resolution and flexibility in terms of flight timing; however, in reality, logistical issues regarding flight regulations, funding, and weather conditions created overwhelming limitations. These

sensors were also still in their early phases of development, so there were many issues in terms of preprocessing (i.e. radiometric correction, geometric correction, and sensor calibration). With Hyperion, the temporal resolution was restricted by its infrequent satellite overpasses, availability for scheduling, and sensor functionality; at the time of image acquisition, Hyperion was plagued by numerous technical problems that degraded image quality (such as bad bands, pixels, and image misalignment).

Through this exploration, we examine some of the key topics in ecology today: nutrients, invasive plants, and community biodiversity. The use of various scales and topics enables me to incorporate a number of different imaging spectroscopy applications into the analysis. In the second chapter, nutrient analysis is conducted through leaf-level models. The third chapter "scales up" to the use of airborne sensors and radiative transfer techniques to measure nutrients. The fourth chapter applies sub-pixel classification algorithms for imaging spectroscopy, previously used in simple ecosystems, to a relatively recent invader in the temperate deciduous forest. Finally, the fifth chapter offered the opportunity for data fusion by combining imaging spectroscopy with lidar data to model understory species diversity using texture metrics.

References

Adler, Peter B, Seabloom, Eric W, Borer, Elizabeth T, Hillebrand, Helmut, Hautier, Yann, Hector, Andy, . . . Anderson, T Michael. (2011). Productivity is a poor predictor of plant species richness. *Science, 333*(6050), 1750-1753.

Al-Abbas, AH, Barr, R, Hall, JD, Crane, FL, & Baumgardner, MF. (1974). Spectra of normal and nutrient-deficient maize leaves. *Agronomy Journal, 66*(1), 16-20.

Albaugh, Janine M, Blevins, Leandra, Allen, H Lee, Albaugh, Timothy J, Fox, Thomas R, Stape, José L, & Rubilar, Rafael A. (2010). Characterization of foliar macro-and micronutrient concentrations and ratios in loblolly pine plantations in the southeastern United States. *Southern Journal of Applied Forestry, 34*(2), 53-64.

Bracho, Rosvel, Starr, Gregory, Gholz, Henry L, Martin, Timothy A, Cropper, Wendell P, & Loescher, Henry W. (2012). Controls on carbon dynamics by ecosystem structure and climate for southeastern US slash pine plantations. *Ecological Monographs, 82*(1), 101-128.

Brown, Ralph B, & Noble, Scott D. (2005). Site-specific weed management: sensing requirements-what do we need to see? *Weed Science, 53*(2), 252-258.

Carlson, Kimberly M, Asner, Gregory P, Hughes, R Flint, Ostertag, Rebecca, & Martin, Roberta E. (2007). Hyperspectral remote sensing of canopy biodiversity in Hawaiian lowland rainforests. *Ecosystems, 10*(4), 536-549.

Chapin, F Stuart, III, Zavaleta, Erika S, Eviner, Valerie T, Naylor, Rosamond L, Vitousek, Peter M, Reynolds, Heather L, . . . Hobbie, Sarah E. (2000). Consequences of changing biodiversity. *Nature, 405*(6783), 234-242.

Cohen, Warren B, & Goward, Samuel N. (2004). Landsat's role in ecological applications of remote sensing. *Bioscience, 54*(6), 535-545.

Cook, Bruce D, Nelson, Ross F, Middleton, Elizabeth M, Morton, Douglas C, McCorkel, Joel T, Masek, Jeffrey G, . . . Montesano, Paul M. (2013). NASA Goddard's LiDAR, Hyperspectral and Thermal (G-LiHT) Airborne Imager. *Remote Sensing, 5*(8), 4045-4066.

Coulombe, Mary J. (1995). Sustaining the world's forests: the Santiago Agreement. *Journal of Forestry, 93*(4), 18-21.

Curran, Paul J. (1989). Remote sensing of foliar chemistry. *Remote sensing of Environment, 30*(3), 271-278.

Dehaan, Remy, Louis, John, Wilson, Andrea, Hall, Andrew, & Rumbachs, Rod. (2007). Discrimination of blackberry (Rubus fruticosus sp. agg.) using hyperspectral imagery in Kosciuszko National Park, NSW, Australia. *ISPRS journal of photogrammetry and remote sensing, 62*(1), 13-24.

Duffy, David Cameron, & Meier, Albert J. (1992). Do Appalachian herbaceous understories ever recover from clearcutting? *Conservation Biology, 6*(2), 196-201.

Dupouey, Jean-Luc, Dambrine, E, Laffite, Jean-Denis, & Moares, Concha. (2002). Irreversible impact of past land use on forest soils and biodiversity. *Ecology, 83*(11), 2978-2984.

Epstein, Emanuel, & Bloom, Arnold J. (2004). *Mineral Nutrition of Plants: Principles and Perspectives* (2 ed.): Sinauer Associates.

Féret, Jean-Baptiste, & Asner, Gregory P. (2014). Mapping tropical forest canopy diversity using high-fidelity imaging spectroscopy. *Ecological Applications, 24*(6), 1289-1296.

Ferwerda, Jelle G, & Skidmore, Andrew K. (2007). Can nutrient status of four woody plant species be predicted using field spectrometry? *ISPRS Journal of Photogrammetry and Remote Sensing, 62*(6), 406-414.

Flinn, Kathryn M, & Vellend, Mark. (2005). Recovery of forest plant communities in post-agricultural landscapes. *Frontiers in Ecology and the Environment, 3*(5), 243-250.

Gilliam, Frank S. (2007). The ecological significance of the herbaceous layer in temperate forest ecosystems. *Bioscience, 57*(10), 845-858.

Gillis, MD, Omule, AY, & Brierley, T. (2005). Monitoring Canada's forests: the national forest inventory. *The Forestry Chronicle, 81*(2), 214-221.

Gökkaya, Kemal. (2012). *Prediction of Foliar Biochemistry in a Boreal Forest Canopy Using Imaging Spectroscopy and LiDAR Data.* Doctoral dissertation, Virginia Tech. 128 pp.

Green, Robert O, Eastwood, Michael L, Sarture, Charles M, Chrien, Thomas G, Aronsson, Mikael, Chippendale, Bruce J, . . . Solis, Manuel. (1998). Imaging spectroscopy and the airborne visible/infrared imaging spectrometer (AVIRIS). *Remote Sensing of Environment, 65*(3), 227-248.

He, Kate S, Rocchini, Duccio, Neteler, Markus, & Nagendra, Harini. (2011). Benefits of hyperspectral remote sensing for tracking plant invasions. *Diversity and Distributions, 17*(3), 381-392.

Heumann, Benjamin W, Hackett, Rachel A, & Monfils, Anna K. (2015). Testing the spectral diversity hypothesis using spectroscopy data in a simulated wetland community. *Ecological Informatics, 25*, 29-34.

Hooper, David U, Adair, E Carol, Cardinale, Bradley J, Byrnes, Jarrett EK, Hungate, Bruce A, Matulich, Kristin L, . . . O'Connor, Mary I. (2012). A global synthesis reveals biodiversity loss as a major driver of ecosystem change. *Nature, 486*(7401), 105-108.

Hooper, David U, Chapin, FS, III, Ewel, JJ, Hector, A, Inchausti, Pablo, Lavorel, Sandra, . . . Naeem, S. (2005). Effects of biodiversity on ecosystem functioning: a consensus of current knowledge. *Ecological monographs, 75*(1), 3-35.

Huston, Michael A. (1997). Hidden treatments in ecological experiments: re-evaluating the ecosystem function of biodiversity. *Oecologia, 110*(4), 449-460.

Isbell, Forest, Calcagno, Vincent, Hector, Andy, Connolly, John, Harpole, W Stanley, Reich, Peter B, . . . van Ruijven, Jasper. (2011). High plant diversity is needed to maintain ecosystem services. *Nature, 477*(7363), 199-202.

Kooch, Y, Hosseini, SM, Mohammadi, J, & Hojjati, SM. (2012). Effects of uprooting tree on herbaceous species diversity, woody species regeneration status and soil physical characteristics in a temperate mixed forest of Iran. *Journal of Forestry Research, 23*(1), 81-86.

Lawrence, Rick L, Wood, Shana D, & Sheley, Roger L. (2006). Mapping invasive plants using hyperspectral imagery and Breiman Cutler classifications (RandomForest). *Remote Sensing of Environment, 100*(3), 356-362.

Lockwood, Julie L, Hoopes, Martha F, & Marchetti, Michael P. (2013). *Invasion* ecology(2nd ed.) West Sussex, UK: Wiley-Blackwell.

Martin, Mary E, & Aber, John D. (1997). High spectral resolution remote sensing of forest canopy lignin, nitrogen, and ecosystem processes. *Ecological applications, 7*(2), 431-443.

Maryland Department of Natural Resources. (2014). Patapsco Valley State Park History. from http://dnr.maryland.gov/publiclands/patapscohistory.asp

Naeem, Shahid, Chapin III, FS, Costanza, Robert, Ehrlich, Paul R, Golley, Frank B, Hooper, David U, . . . Sala, Osvaldo E. (1999). Biodiversity and ecosystem functioning: maintaining natural life support processes. *Issues in Ecology Series, No. 4.* Ecological Society of America, Washington, DC.

O'Laughlin, Jay, Livingston, R Ladd, Thier, Ralph, Thornton, John P, Toweill, Dale E, & Morelan, Lyn. (1994). Defining and measuring forest health. *Journal of Sustainable Forestry, 2*(1-2), 65-85.

Pallardy, Stephen G. (2007). *Physiology of Woody Plants* (3rd ed.). Orlando, FL, US: Academic Press.

Peterson, David L, Aber, John D, Matson, Pamela A, Card, Don H, Swanberg, Nancy, Wessman, Carol, & Spanner, Michael. (1988). Remote sensing of forest canopy and leaf biochemical contents. *Remote Sensing of Environment, 24*(1), 85-108.

Rapport, David J. (1992). What is clinical ecology. *Ecosystem health: new goals for environmental management*, 144-156.

Roberts, Mark R. (2004). Response of the herbaceous layer to natural disturbance in North American forests. *Canadian Journal of Botany, 82*(9), 1273-1283.

Rocchini, Duccio, Balkenhol, Niko, Carter, Gregory A, Foody, Giles M, Gillespie, Thomas W, He, Kate S, . . . Luoto, Miska. (2010). Remotely sensed spectral heterogeneity as a proxy of species diversity: recent advances and open challenges. *Ecological Informatics, 5*(5), 318-329.

Rossini, M, Nedbal, L, Guanter, L, Ač, A, Alonso, L, Burkart, A, . . . Drusch, M. (2015). Red and far-red sun-induced chlorophyll fluorescence as a measure of plant photosynthesis. *Geophysical Research Letters*, *42*(6), 1632-1639.

Sardans, Jordi, & Peñuelas, Josep. (2012). The role of plants in the effects of global change on nutrient availability and stoichiometry in the plant-soil system. *Plant Physiology, 160*(4), 1741-1761.

Small, Christine J, & McCarthy, Brian C. (2005). Relationship of understory diversity to soil nitrogen, topographic variation, and stand age in an eastern oak forest, USA. *Forest Ecology and Management, 217*(2), 229-243.

Taiz, Lincoln, & Zeiger, Eduardo. (2010). *Plant Physiology* (5th ed.). Sunderland, MA: Sinauer Associates.

Thenkabail, Prasad S, Gumma, Murali Krishna, Teluguntla, Pardhasaradhi, & Mohammed, Irshad A. (2014). Hyperspectral remote sensing of vegetation and agricultural crops. *Photogrammetric Engineering and Remote Sensing, 80*(8), 697-709.

Tilman, David, Reich, Peter B, & Isbell, Forest. (2012). Biodiversity impacts ecosystem productivity as much as resources, disturbance, or herbivory. *Proceedings of the National Academy of Sciences, 109*(26), 10394-10397.

Vilà, Montserrat, Espinar, José L, Hejda, Martin, Hulme, Philip E, Jarošík, Vojtěch, Maron, John L, . . . Pyšek, Petr. (2011). Ecological impacts of invasive alien plants: a meta-analysis of their effects on species, communities and ecosystems. *Ecology letters, 14*(7), 702-708.

Woodall, Christopher W, Rondeux, Jacques, Verkerk, Pieter J, & Ståhl, Göran. (2009). Estimating dead wood during national forest inventories: a review of inventory methodologies and suggestions for harmonization. *Environmental management, 44*(4), 624-631.

Woodall, Christopher W, Amacher, Michael C, Bechtold, William A, Coulston, John W, Jovan, Sarah, Perry, Charles H, . . . Tkacz, Borys. (2011). Status and future of the forest health indicators program of the USA. *Environmental Monitoring and Assessment, 177*(1-4), 419-436.

Zhang, Y, West, JB, Will, RE, & Vogel, JG. (2013). *Effects of fertilization and drought on substrate decomposition and inorganic nitrogen concentration in a managed loblolly pine forest.* AGU Meeting, San Francisco, CA.

Zhang, Yu, Chen, Han YH, & Reich, Peter B. (2012). Forest productivity increases with evenness, species richness and trait variation: a global meta-analysis. *Journal of ecology*, *100*(3), 742-749.

2 Chapter 2—Predicting Macronutrient Concentrations From Loblolly Pine Leaf Reflectance Across Local and Regional Scales

Beth R. Stein, Valerie A. Thomas, Laura J. Lorentz, and Brian D. Strahm
Department of Forest Resources and Environmental Conservation, Virginia Tech

Published in *GIScience and Remote Sensing*, Volume 51, Issue 3, May 2014, Pages 269-287, available via http://www.tandfonline.com/doi/abs/10.1080/15481603.2014.912875

Attribution
BRS developed the research plan, conducted all analyses, and wrote the paper; VAT assisted with research development, analysis and writing; LJL conducted the field work for the regional study; and BDS assisted with the research plan, fieldwork methodology, and writing the paper.

GIScience & Remote Sensing, 2014
Vol. 51, No. 3, 269–287, http://dx.doi.org/10.1080/15481603.2014.912875

Taylor & Francis
Taylor & Francis Group

Predicting macronutrient concentrations from loblolly pine leaf reflectance across local and regional scales

Beth R. Stein[a]*, Valerie A. Thomas[b], Laura J. Lorentz[a] and Brian D. Strahm[c]

[a]*Department of Forest Resources & Environmental Conservation, College of Natural Resources and Environment, Virginia Tech, 305 Cheatham Hall, Blacksburg, 24061 VA, USA;* [b]*Department of Forest Resources & Environmental Conservation, College of Natural Resources and Environment, Virginia Tech, 307A Cheatham Hall, Blacksburg, 24061 VA, USA;* [c]*Department of Forest Resources & Environmental Conservation, College of Natural Resources and Environment, Virginia Tech, 228 Cheatham Hall, Blacksburg, 24061 VA, USA*

(*Received 3 December 2013; accepted 24 March 2014*)

Given the economic importance of loblolly pine (*Pinus taeda*) in the southeastern US, there is a need to establish efficient methods of detecting potential nutrient deficiencies that may limit productivity. This study evaluated the use of remote sensing for macronutrient assessment in loblolly pine. Reflectance-based models were developed at two spatial scales: (1) a natural nutrient gradient across the species' range, and (2) localized fertilization and genotype treatments in North Carolina and Virginia. Fascicles were collected regionally from 237 samples of 3 flushes at 18 sites, and locally from 72 trees with 2 fertilization treatments and 6 genotypes. Sample spectral reflectance was calculated using a spectroradiometer, and nutrient concentrations were measured with dry combustion and wet chemical digestion. Results were analyzed statistically using nutrient correlations with reflectance and common vegetation indices, and partial least squares regression (PLSR). PLSR performed well at the regional scale, with R^2 values for nitrogen, phosphorus, potassium, calcium, and magnesium of 0.81, 0.70, 0.68, 0.42, and 0.51, respectively. No model successfully predicted nutrients at local sites for any treatment or canopy stratum. This discrepancy implies that a large nutrient range and/or spatial scale may be necessary to model loblolly pine nutrients with spectral reflectance.

Keywords: remote sensing; spectroradiometer; nutrients; loblolly pine; partial least squares regression; spatial scale

Introduction

Loblolly pine (*Pinus taeda*) is the most commercially productive species for timber in the southeastern US (SE US), covering over 13 million hectares in the region (Susaeta et al. 2012). However, soil nutrient limitations are common, particularly in nitrogen (N) and phosphorus (P). Along with potassium (K), these nutrients form the primary nutrient requirements of plants; secondary nutrients are calcium (Ca), magnesium (Mg), and sulfur (S). N, P, K, and Mg are required in photosynthesis and respiration, while Ca is used in cell division and cell walls (Taiz and Zeiger 2010). Insufficient macronutrients can stunt plant growth, injure organs, and lead to mortality (Pallardy 2008).

*Corresponding author. Email: bstein2@vt.edu

When nutrient stress is detected early, forest managers can reduce the likelihood of damage by applying fertilizer or other silvicultural treatments (Svotwa et al. 2012). However, field measurements to determine nutrient concentrations are time- and labor-intensive (Gong et al. 2012). With such a vast area allocated to loblolly pine production, remote sensing may offer advantages over current field methods for nutrient assessment (Im et al. 2009; Svotwa et al. 2012). Satellite hyperspectral sensors, such as NASA's Hyperion sensor and ESA's Compact High Resolution Imaging Spectrometer (CHRIS), collect high resolution spectral data that is sensitive to biophysical characteristics in vegetation. The upcoming launches of the Environmental Mapping and Analysis Program (EnMAP) satellite in Germany and the NASA Hyperspectral Infrared Imager (HyspIRI) mission necessitate a greater understanding of the relationship between macronutrients and reflectance across species' ranges to maximize data benefits.

Effective use of hyperspectral data relies on field-derived relationships between spectral reflectance patterns and chemical composition. The chemical constituents of foliage reflect unique amounts of radiation in the visible and NIR portions of the spectrum. Since the 1970s, researchers have used spectral signatures to identify foliar biochemicals and nutrients (Al-Abbas et al. 1974). After Hinzman, Bauer, and Daughtry (1986) found differences in visible, near infrared (NIR), and mid-infrared reflectance from fertilizer treatments, Peterson et al. (1988) predicted leaf nitrogen concentrations from reflectance measurements. Studies of other nutrients followed a similar trajectory, as researchers used observed spectral patterns to predict nutrient quantities (Curran 1989; Ferwerda and Skidmore 2007; Gholz et al. 1997).

Relationships between reflectance and nutrients vary significantly by element and plant species. For this reason, interspecific models are often quite successful at predicting nutrients across diverse areas (e.g. Asner et al. 2011; Ollinger et al. 2008). Im et al. (2009) used airborne imaging spectroscopy to quantify macronutrient concentrations in several species, including loblolly pine. However, interspecific models may not be appropriate for individual species, which often have a much narrower range in nutrient concentrations. Research on other plants suggest that relatively low intra-site variability necessitates a large environmental gradient for monospecific models (Ferwerda and Skidmore 2007; Pimstein et al. 2011). Currently, few studies have examined loblolly pine reflectance beyond nitrogen or leaf area index (Gong et al. 2012; Nelson, Gjerstad, and Glover 1986; Tsay, Gjerstad, and Glover 1982); so, the necessary nutrient range is undefined. Even given the known nutrient limitations, there are no remote-sensing models specifically for loblolly pine macronutrients.

The objective of this study is to use reflectance measurements to develop models for loblolly pine macronutrient concentrations (N, P, K, Ca, and Mg) across two settings: (1) a natural nutrient gradient across the SE US, and (2) localized experimental fertilization treatments in Virginia and North Carolina. The goal is twofold: to determine the relationship between spectral reflectance and nutrient concentration in loblolly pine, and to investigate the role of geographic scale in model accuracy. To achieve these aims, the study considers several treatments and leaf characteristics known to affect nutrient concentration, and in some cases, the spectral signature. These variables include fertilization, genotype, canopy stratum, and flush (Im et al. 2009; Zhang and Allen 1996). Determination of the appropriate spatial scale for loblolly pine nutrient models can improve current methods for modeling and mapping nutrient quantities in the SE US. Relationships between flush and spectra will demonstrate the effectiveness of representing total foliage, both nutritionally and spectrally, with a given flush. In addition, results may provide insight into the possibility and potential limitations of scaling algorithms up to landscape-level analysis using satellites.

Methodology

Study sites

Regional analysis

The 18 regional study sites are located across the SE US. These sites range latitudinally from Virginia to Florida, with a longitudinal range to East Texas and Oklahoma (Figure 1). They consist of loblolly pine plantations across the species' environmental habitat gradient. Sites represent 5 US Environmental Protection Agency Level II ecoregions and 13 Level III ecoregions (US Environmental Protection Agency 2010). Site conditions are extremely varied, including differences in soils, stress, and climate. Stands range in soil properties, tree age, level of thinning, and amount of understory. Mean annual daily average temperatures are 12 to 21°C, and mean annual precipitation is 101 to 178 cm (National Climatic Data Center 2005).

Local analysis

The two local sites are within the Reynolds Homestead Forestry Research Center in Patrick County, Virginia, US (36°40′ N, 80°10′ W), and Bladen Lakes State Forest in Bladen County, North Carolina, US (34°41′ N, 78°36′ W, Figure 1). Reynolds Homestead is in the Virginia Piedmont and has well-drained Fairview series soils. For a more detailed description of the topography, climate, and soils at Reynolds Homestead, please refer to Stovall et al. (2011). Bladen Lakes, in the North Carolina Coastal Plain, has poorly drained soils of the Rains series. The NSF Center for Advanced Forestry Systems established the genotype and fertilization plots in 2009. Preliminary results in the first three years of the experiment indicate silvicultural and site differences in loblolly pine growth (Yanez, forthcoming).

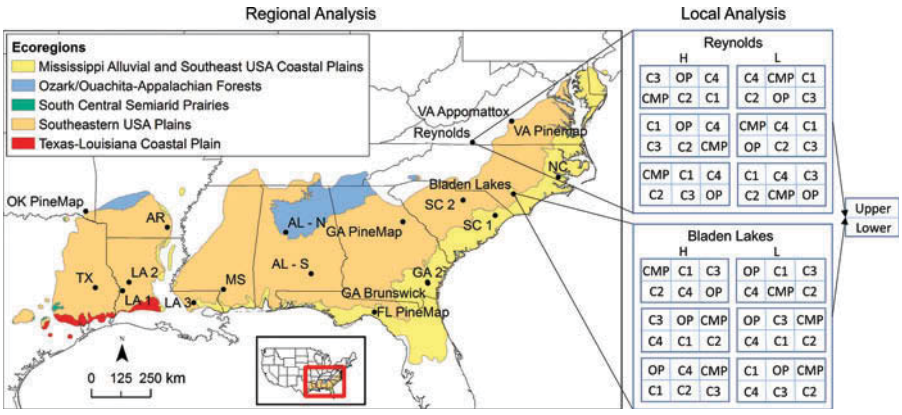


Figure 1. Study site locations in the southeastern US for analyses at the regional and local scales, with the local sampling scheme. Sites span the geographic range of loblolly pine (Little 1971). Regional sites represent a diversity of environmental characteristics and cover multiple ecoregions, as characterized by the US Environmental Protection Agency's Level II ecoregional classification (US EPA 2010). At the local scale, an in-depth analysis was conducted across canopy levels ("Lower" and "Upper"), fertilization treatments ("H" = High, "L" = Low), and genotypes (C1-C4 = distinct genotypes, "CMP" = controlled mass pollination, "OP" = open pollination).

Field and laboratory measurements

Data collection and sampling procedures differed by the scale of analysis. Separate field campaigns were conducted for the regional analysis and localized genotype and fertilization experiments.

Regional sampling across loblolly range

A FieldSpec 3 Rapid and Portable Spectroradiometer (ASD Inc., Boulder, CO) was used to take measurements at the SE US study sites. The spectroradiometer has a spectral range of 350–2500 nm and records 2150 spectral channels at 1.4–2 nm resolution. At each site, branches were shot down from the upper canopy of one or two loblolly pine trees. One to six foliar samples consisting of multiple adjacent fascicles were collected from the most recent flush, previous flush, and previous year's flush. Prior to data collection, the spectroradiometer was calibrated with a white reference panel (Spectralon Labsphere Inc., North Sutton, NH, US). Consequently, the entire view surface of the plant probe was covered with fascicles. An average of 10 spectral readings for each wavelength in the range were collected for each sample. The presence and quantity of particular flushes varied across branches. Measurements were taken in the summer of 2012.

Local sampling for fertilization treatment and genotype

At Reynolds Homestead and Bladen Lakes, leaf samples and reflectance measurements (n = 144) were collected in the upper and lower canopies of 72 trees. The experimental design consists of three replicates of two levels of silvicultural treatments (high and low intensity) and six loblolly genotypes: C1, C2, C3, C4, Controlled Mass Pollination (CMP), and Open Pollination (OP). At Reynolds Homestead, high silvicultural plots were fertilized with a phosphate-coated urea fertilizer (CUF: 39% N, 4% P, 0.16% B; 239 g/tree) in March, 2010, and herbicide was applied in April 2010 and 2011 (4 oz Arsenal AC + 2 oz Oust XP + 0.25 oz Escort per ac). At Bladen Lakes, herbicide and fertilizer were applied in March 2010, and herbicide was re-applied in May 2011 (Yanez, Forthcoming). Measurements were taken in September and October, 2012, in the fourth growing season.

Reflectance at the needle-level for each canopy stratum were measured with the aforementioned spectroradiometer. Three sunlit branches in the lower and upper canopies were selected for sampling. One fascicle was detached from the latest flush on each branch in the canopy stratum, producing two sets of three fascicles per tree. The spectroradiometer was calibrated with a white reference (Spectralon Labsphere Inc., North Sutton, NH), and then each set of fascicles was arranged on the plant probe. An average of 10 reflectance measurements was immediately recorded for samples. Additionally, 25 fascicles were collected from the same branches for nutrient analysis.

Laboratory procedures

Fascicles were immediately refrigerated and stored in a dry room prior to conducting laboratory analyses. Macronutrient concentrations were assessed for all foliage samples through the dry ash procedure using a spectrometer (Jones and Steyn 1973). Samples were ground with a Wiley mill and oven-dried for 24 hours. Desiccated samples of 0.5 g were poured into ignition tubes and heated in a muffle furnace at gradually increasing temperatures up to 500°C. Samples were mixed with 10 mL 6N HCl and 40 mL distilled water, shaken, and analyzed with an inductively

coupled plasma-optical emission spectrometer (ICP-OES) on a Varion Vista-Max (Varion, Palo Alto, CA, US). For C and N concentrations, oven-dried samples were measured by dry combustion with a vario MAX CN Analyzer (Elementar, Hanau, Germany).

Statistical approach

First derivatives and 28 common reflectance indices for vegetation analysis were calculated from the reflectance data (Table 1). These indices were selected due to

Table 1. Common vegetation reflectance indices and their formulas calculated for correlation with and prediction of foliar macronutrient concentrations (Gökkaya et al. 2014; Schlemmer et al. 2013; Stagakis et al. 2010; Ustin et al. 2009). "R" denotes reflectance and "D" the derivative for a given wavelength (nm).

Index	Abbreviation	Formula
Chlorophyll Absorption In Reflectance Index	CARI	$(R_{700} - \rho_{670}) - 0.2(R_{700} - R_{550})$
Modified CARI	MCARI	$(R_{700} - R_{670}) - 0.2(R_{700} - R_{550})(R_{700}/R_{670})$
Red Edge Chlorophyll Index	CI RE	$R_{NIR}/R_{RE} - 1$
Photochemical Reflectance Index	PRI	$(R_{531} - R_{570})/(R_{531} + R_{570})$
Normalized Difference Vegetation Index	NDVI	$(R_{NIR} - R_{Red})/(R_{NIR} + R_{Red})$
Modified NDVI	mNDVI	$(R_{750} - R_{705})/(R_{750} + R_{705})$
Simple Ratio Index	SR	R_{NIR}/R_{Red}
Modified Red Edge Simple Ratio Index	mSR705	$(R_{750} - R_{445})/(R_{705} - R_{445})$
Enhanced Vegetation Index	EVI	$2.5(R_{NIR} - R_{Red})/(1 + R_{NIR} + 6R_{Red} - 7.5R_{Blue})$
Greenness Index	GI	R_{554}/R_{677}
Structure Insensitive Pigment Index	SIPI	$(R_{800} - R_{445})/(R_{800} - R_{680})$
Pigment Specific Simple Ratio 675	PSSR 675	R_{800}/R_{675}
Pigment Specific Simple Ratio 650	PSSR 650	R_{800}/R_{650}
Pigment Specific Normalized Difference 675	PSND 675	$(R_{800} - R_{675})/(R_{800} + R_{675})$
Pigment Specific Normalized Difference 650	PSND 650	$(R_{800} - R_{650})/(R_{800} + R_{650})$
Plant Senescence Reflectance Index	PSRI	$(R_{680} - R_{500})/R_{750}$
Carotenoid Reflectance Index 1	CRI1	$(1/R_{510}) - (1/R_{550})$
Carotenoid Reflectance Index 2	CRI2	$(1/R_{510}) - (1/R_{700})$
Anthocyanin Reflectance Index 1	ARI1	$(1/R_{550}) - (1/R_{700})$
Anthocyanin Reflectance Index 2	ARI2	$((1/R_{550}) - (1/R_{700}))R_{800}$
Red Green Ratio Index	RGRI	R_{Red}/R_{Green}
Gitelson & Merzlyak 2	G&M2	R_{750}/R_{700}
Lichtenthaler Index 1	Lic1	$(R_{800} - R_{680})/(R_{800} + R_{680})$
Vogelmann Index 1	Vog1	R_{740}/R_{720}
Vogelmann Index 2	Vog2	$(R_{734} - R_{747})/(R_{715} + R_{726})$
Derivative Chlorophyll Index	DCI	D_{705}/D_{722}
Maximum Derivative of Red-Edge	DmaxRE	$Dmax_{RE}$
Maximum Derivative of Red-Edge Divided by Derivative at 703 nm	Dmax703	$Dmax_{RE}/D_{703}$

their repeated use in the related literature. Data was tested for normality and determined to require non-parametric statistics. Three correlation-based statistical approaches were conducted: (1) correlation analysis using reflectance, derivatives, and absorbance transformations (Asner 2008; Smith et al. 2003); (2) correlation analysis with indices (Im et al. 2009); and (3) partial least squares regression (PLSR) with reflectance, derivative, and absorbance values (Asner et al. 2011; Martin et al. 2008). The purpose of the correlational analyses was to determine any regions of the spectrum (i.e. real features across multiple wavelengths) or indices that are consistently correlated with loblolly pine nutrient status. As a powerful modeling technique increasingly used in spectroscopy and remote sensing, PLSR offers the ability to model nutrient concentration across the full spectrum. Reflectance (R) values were transformed to absorbance (A) using the equation, $A = \log(1/R)$ (Smith et al. 2003).

Spearman's rank correlation coefficients (r_s) were calculated between nutrient concentrations and wavelength reflectance values and indices. Next, PLSR was conducted with centered and scaled variables using leave-one-out cross-validation (LOOCV). LOOCV is effective on small data sets and has low bias, variance, and error rates compared to other cross-validation techniques (Cawley and Talbot 2003; Mevik and Cederkvist 2004). For each nutrient, the PLSR model with the minimum root mean PRESS was selected. For flush model comparison, the model with the least number of latent factors with a van der Voet T^2 significance level greater than 0.10 was selected (Townsend et al. 2003). The van der Voet T^2 is a test statistic that indicates whether model residuals are significantly greater than residuals in the model with the lowest error (Tobias 1995). For models with 15 latent factors and no minimum PRESS, the model with 1 latent factor was selected, indicating no valid model. PLSR models were also calculated using major indices (Table 1); and various subsets of the spectrum, consisting of different numbers and combinations of wavelengths derived from the variable importance for the projection (VIP) and correlation analyses. Differences between the various strata – flush, canopy level, genotype, and fertilization treatment – were examined to assess the value in developing treatment-specific models (Figure 2). Tukey–Kramer HSD groupings indicated significant differences between strata (Table 2). Consequently, models were also calculated across each stratum. For all models, centered and scaled regression coefficients and VIP were calculated. VIP indicate the contribution of each variable to the model fit across all latent factors for the predictors and response variables. Wold's criterion of 0.8 was used as the threshold for variable significance (Wold 1994). Predicted versus observed and residuals versus predicted plots were evaluated for model fit.

Results

Nutrients and reflectance

Foliar nutrient concentrations varied between sites within and across ecoregions (Figure 3). Correlation coefficients of the nutrient concentrations and spectra also varied substantially, as shown in Figure 4. There were many significant correlations in the regional samples ($p < 0.05$). Ca has the strongest relationship with reflectance (730 nm, $r_s = -0.62$), followed by K (733 nm, $r_s = 0.45$). The NIR plateau (750–930 nm) was moderately correlated with N, K, and Ca (average $r_s = 0.41$, 0.42, –0.57); this was also the best-performing region for P, but correlations were weaker (average $r_s = 0.28$). K and Ca were also strongly related to wavelengths in the red-edge and green regions. Mg was best, albeit weakly, correlated with mid-IR wavelengths (1380–2500 nm, average $r_s = -0.23$).

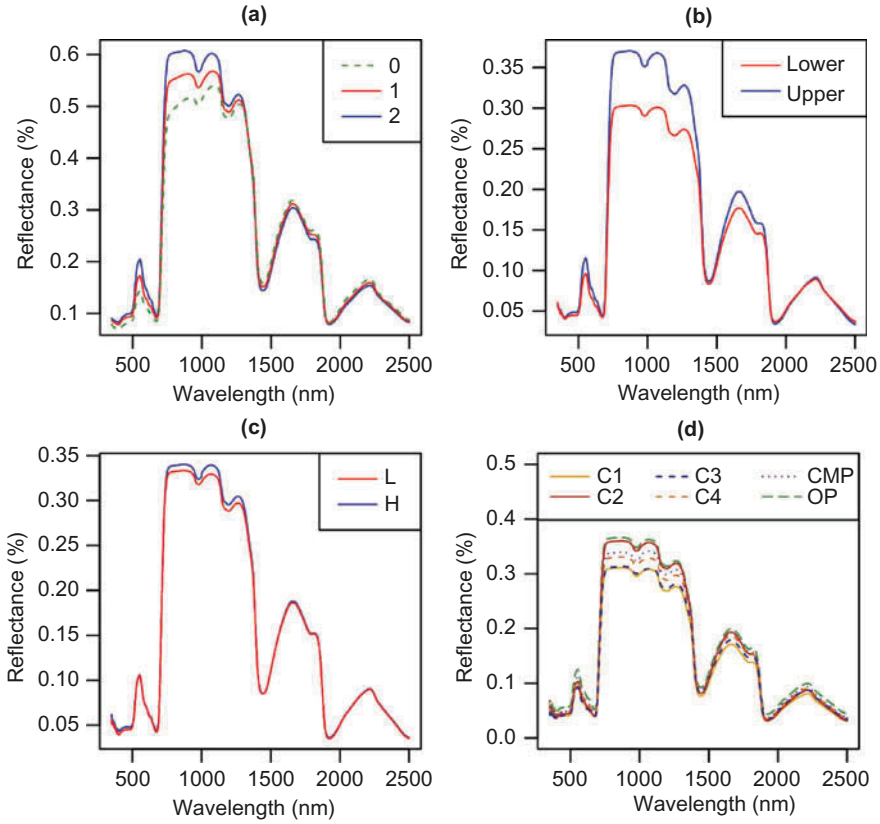


Figure 2. Significant differences exist in average reflectance across strata: (a) flush ("0" = older flush, "1" = previous flush, "2" = most recent flush), (b) canopy level, (c) genotype (C1-C4 = distinct genotypes, "CMP" = controlled mass pollination, "OP" = open pollination), (d) fertilization treatment ("H" = high, "L" = low).

Table 2. Tukey–Kramer groupings of average foliar nutrient content by flush in loblolly pine samples across the southeastern US. Statistically significant nutrient differences ($p < 0.05$) are shown by different groups and exist between flushes for nearly all nutrients.

	Grouping (A, B or C), Mean (%)				
Flush	Nitrogen	Phosphorus	Potassium	Calcium	Magnesium
0	A, 1.08	A, 0.11	A, 0.62	A, 0.27	A, 0.11
1	B, 1.21	B, 0.10	B, 0.51	B, 0.19	A, 0.11
2	B, 1.25	C, 0.09	C, 0.36	C, 0.14	A, 0.10

Correlations have inverse signs for reflectance and absorbance. Derivative correlations fluctuated greatly across the spectrum, with peaks in similar locations across nutrients. K and Ca best correlated with the derivative spectra in the NIR plateau (average $r_s = -0.44$ and 0.56, respectively). The derivative correlation with N peaked at 740 nm ($r_s = 0.68$, $p < 0.001$).

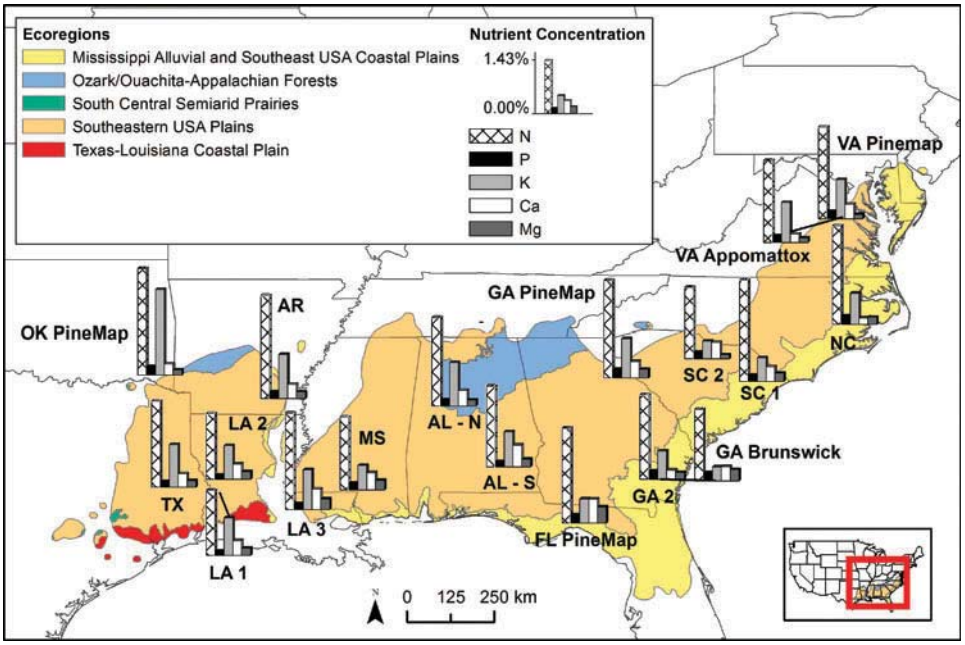


Figure 3. Map of the spatial distribution of loblolly pine foliar macronutrient concentrations (%) across sites in the southeastern US. High variability in nutrient content exists between and within the US Environmental Protection Agency's Level II ecoregions (US EPA 2010).

There were many significant correlations between nutrients and indices ($p < 0.001$). Figure 5 shows the top correlations between each nutrient and a common vegetation index. As Table 3 depicts, no one index was best for all five nutrients; however, several indices and wavelengths have moderate relationships with multiple nutrient concentrations.

Nutrient models

Regional analysis

Partial least squares regression models range in their predictive capabilities by nutrient (Table 4). The r^2 values for N, P, K, Ca, and Mg are 0.81, 0.70, 0.68, 0.42, and 0.51, respectively. Figures 6 and 7 depict model fit using the predicted versus observed values and residuals versus predicted values. PLSR models for the SE sites by flush are shown in Table 5. According to O'Brien's test for equal variances, there is no significant difference between the flush models for N or P ($p = 0.96$ and $p = 0.07$), so the model with the lowest PRESS was selected. For nutrients with a significant difference between the models ($p < 0.001$, $p = 0.02$, and $p = 0.004$ for K, Ca, and Mg, respectively), the model with the least residual variability was selected.

Predictive capabilities of the models generally improve when stratified by flush. However, flush model selection varies by nutrient. The most recent flush produced the selected model for N, Ca, and Mg, while the oldest flush produced models for P and K. Table 4 also displays the PLSR models for the derivatives. Derivative models for all nutrients except Ca have five factors, with a mean R^2 of 0.70. PLSR models with selected

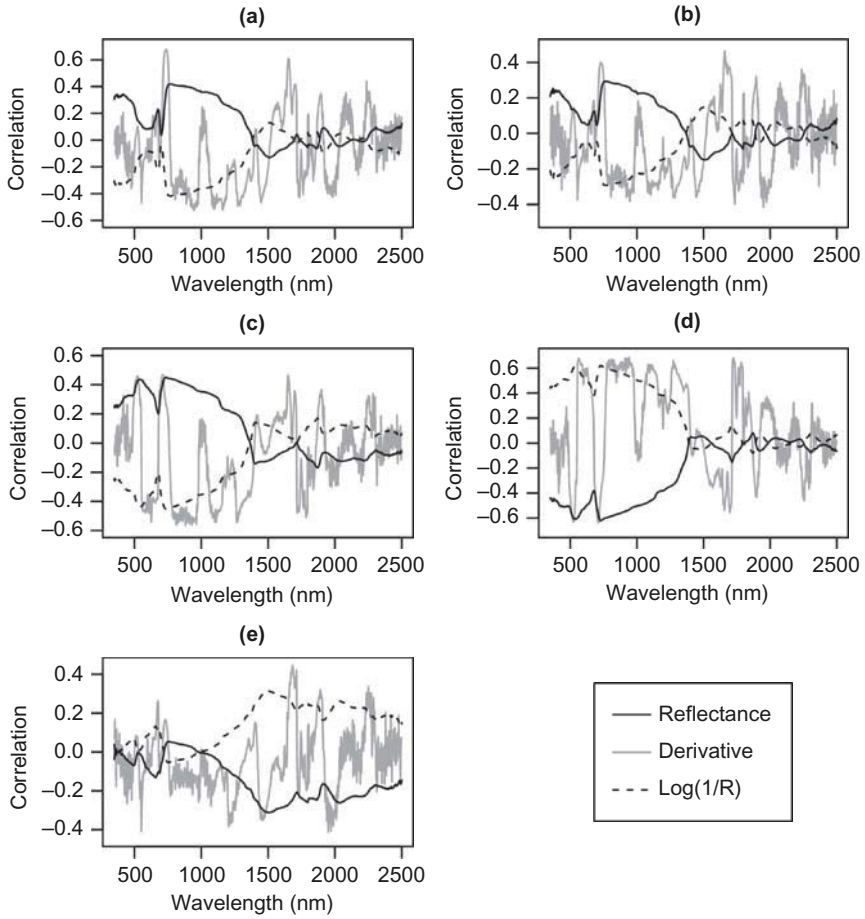


Figure 4. Spearman's rank correlation coefficients (r_s) at the southeastern US sites between original reflectance (R), derivative reflectance, and log(1/R) and nutrient concentrations: (a) nitrogen, (b) phosphorus, (c) potassium, (d) calcium, and (e) magnesium.

indices or wavelengths generally have lower predictive capabilities than the original models.

Similarities and differences in VIP and regression coefficients can be seen in Figures 8 and 9. Flush coefficients generally have similar peaks throughout the spectrum. All nutrients have a peak in the VIP in the visible spectrum between 500–600 nm and near the red edge, around 700 nm.

Local analysis

No stratification produces valid PLSR models consistently across all nutrients or strata, regardless of site, canopy level, genotype, and fertilization treatment. Use of the absorbance transformation, indices, derivatives, and logarithms did not significantly improve models. All derivative models had 0 factors as the lowest PRESS value. This finding is likely due to relatively low variability at the local scale (Figure 10).

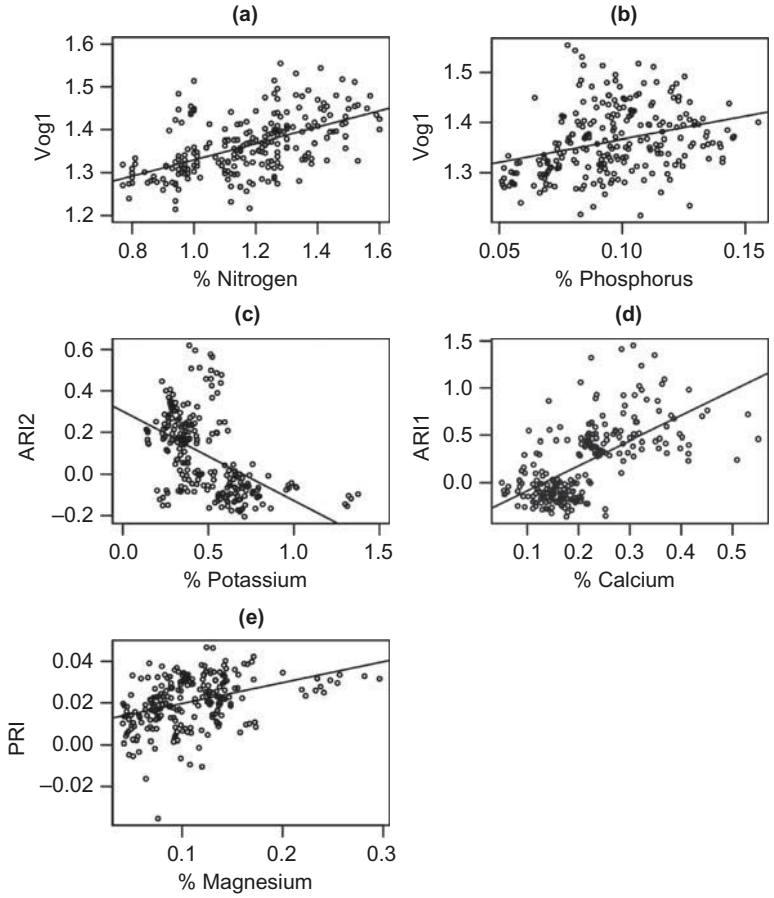


Figure 5. Index versus nutrient plots across all samples at the southeastern US loblolly pine sites: (a) nitrogen and Vogelmann Index 1 (Vog1), (b) phosphorus and Vog1, (c) potassium and the Anthocyanin Reflectance Index 2 (ARI2), (d) calcium and the Anthocyanin Reflectance Index 1 (ARI1), and (e) magnesium and the Photochemical Reflectance Index (PRI).

Discussion

Regional analysis

The PLSR models explain a comparable amount of nutrient variation to prior monospecific studies (Christensen et al. 2004; Pimstein et al. 2011), as well as the related study by Im et al. (2009). However, many studies that model nutrients across multiple species have larger nutrient ranges and produce stronger models. Asner et al. (2011) used PLSR on imaging spectroscopy data to predict macronutrients in tropical forest canopies consisting of several thousand plant species; they obtained R^2 values for N, P, K, Ca, and Mg of 0.77, 0.63, 0.51, 0.65, and 0.57, respectively. Similarly, Petisco et al. (2005) found much higher R^2 values for N, P, and Ca (R^2 = 0.99, 0.94, and 0.95) across 18 different woody plant species in the Iberian Peninsula.

The absorbance transformations, major indices, and subsets of selected wavelengths did not considerably improve the explanatory power of the models. The logarithm transformation produced weaker relationships in the correlation analysis and PLSR.

Table 3. Top correlations between macronutrients and indices for loblolly pine reflectance data from 18 southeastern US sites based on Spearman's rank correlation coefficients ($p < 0.001$). Common vegetation reflectance indices are compiled from prior studies (Table 2).

Nutrient	Index	Correlation coefficient (r_s)
N	Vog1	0.58
	mSR705	0.57
	Vog2	−0.56
	mNDVI	0.55
	G&M2	0.53
P	Vog1	0.37
	mNDVI	0.34
	Vog2	−0.34
	PRI	0.34
	SIPI	−0.33
K	ARI2	−0.55
	ARI1	−0.54
	SIPI	−0.54
	PSRI	−0.54
	EVI	0.45
Ca	ARI1	0.64
	ARI2	0.63
	PSRI	0.63
	SIPI	0.62
	EVI	−0.59
Mg	PRI	0.43
	PSSR 650	0.29
	PSND 650	0.29
	SR	0.27
	NDVI	0.27

Table 4. Partial least squares regression models of reflectance and derivatives by nutrient for the southeastern US sites ($n = 235$) using LOOCV.

Model	Nutrient	Number of factors	R^2	Root mean PRESS	RMSEP
Reflectance	N	14	0.81	0.55	0.08
	P	15	0.70	0.71	0.01
	K	11	0.68	0.64	0.13
	Ca	3	0.42	0.78	0.07
	Mg	10	0.51	0.81	0.11
Derivative	N	5	0.83	0.60	0.08
	P	5	0.77	0.76	0.01
	K	5	0.76	0.70	0.11
	Ca	1	0.44	0.76	0.07
	Mg	5	0.71	0.82	0.02

Derivative models improved predictions and parsimony across all nutrients. This result agrees with a similar study by Petisco et al. (2005) that obtained better results using first and second derivative transformations for N, P, and Ca than the logarithm transformation. Furthermore, Karnieli, Karnieli, and Bonfil (2007) found greater success with derivatives and wavelength selection than the full spectrum of reflectance measurements for N predictions. In this study, selected indices and wavelengths may not have been successful

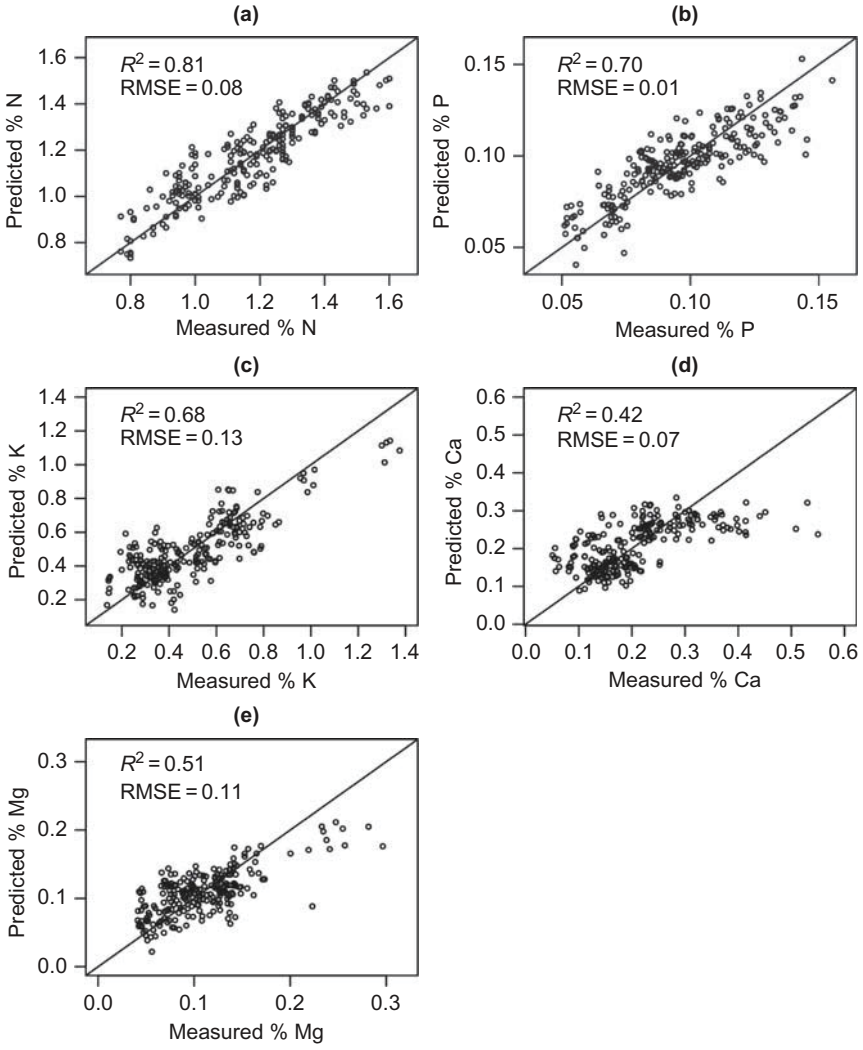


Figure 6. Measured versus predicted plots of partial least squares regression models for 5 nutrients in loblolly pine foliage: (a) nitrogen, (b) phosphorus, (c) potassium, (d) calcium, and (e) magnesium. Models use wavelengths from 350–2500 nm measured with a field spectroradiometer. Samples were collected across the southeastern US.

in the PLSR models because individual wavelengths and/or regions were not strongly correlated with nutrient concentrations.

Flush stratification generally improved parsimony and/or increased explanatory power of nutrient models. The wide range in R^2 values between flushes indicates the importance of flush consideration in sampling. For all nutrients except K, models for the oldest flush used more latent factors and explained more variance than the most recent flush. Likewise, Christensen et al. (2004) used more components in their phosphorus models for older leaves, which they attribute to the greater visibility of nutrient stress in younger leaves. Both Asner (1998) and Christensen et al. (2004) also obtained better N predictions with leaves of later growth stages.

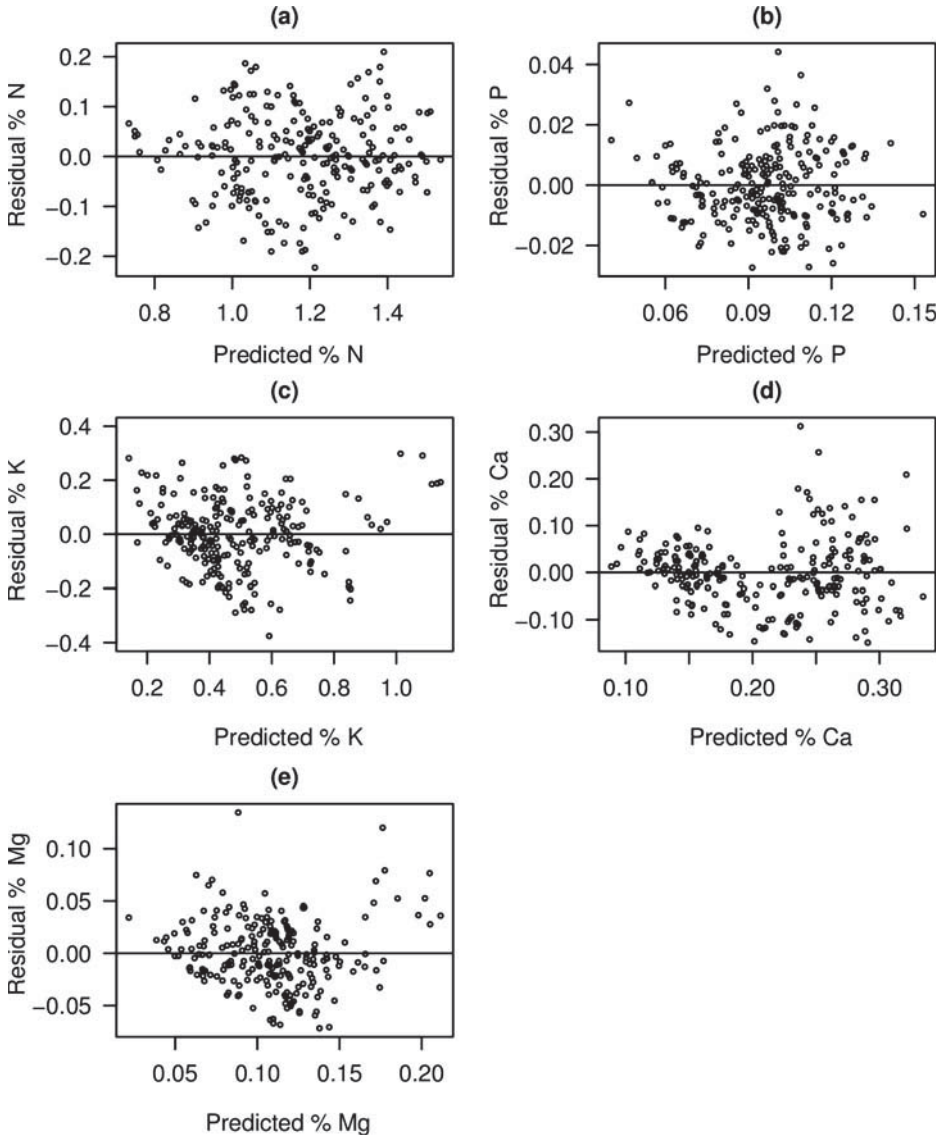


Figure 7. Residual versus predicted plots for 5 nutrients in loblolly pine foliage: (a) nitrogen, (b) phosphorus, (c) potassium, (d) calcium, and (e) magnesium.

Model components

The wavelengths of importance in the PLSR reflectance models are consistent with those in the selected flush models, as well as in prior studies. Given that the flush models use a subset of the data in the reflectance models, this result suggests model stability. The Ca model is the one exception, perhaps due to the relatively low predictive power of the selected flush model. For leaf N status, the N VIP shows the visible, red edge, and NIR regions as the most important reflectance bands (Coops et al. 2003). Wavelengths around 690 nm are particularly related to N due to the absorption of chlorophyll (Christensen et al. 2004).

Table 5. Partial least squares regression models by nutrient and flush for southeastern US sites using LOOCV. Flush 0 indicates the last flush of 2011 (n = 79), Flush 1 indicates the first flush of 2012 (n = 103), and Flush 2 refers to the second flush of 2012 (n = 53).

Nutrient	Flush	Number of factors	R^2	Root mean PRESS	RMSEP
N	0	8	0.82	0.54	0.08
N	1	11	0.78	0.68	0.08
N	2	6	0.81	0.54	0.08
P	0	12	0.78	0.74	0.01
P	1	8	0.54	0.83	0.01
P	2	1	0.18	0.94	0.02
K	0	4	0.34	0.93	0.08
K	1	9	0.74	0.65	0.11
K	2	6	0.73	0.68	0.14
Ca	0	12	0.76	0.83	0.04
Ca	1	7	0.55	0.78	0.06
Ca	2	5	0.46	0.93	0.04
Mg	0	12	0.86	0.63	0.02
Mg	1	10	0.69	0.75	0.00
Mg	2	8	0.73	0.80	0.02

P, K, and Mg showed similar wavelengths of importance in the PLSR models. The P VIP indicates significant contributions from wavelengths near the visible and red-edge regions, 2300 nm, and the 2400–2500 nm range linked to protein, starch, cellulose, and sugar (Curran 1989). The latter connection may be due to the inverse relationship between P content and cellulose or the role of P in protein synthesis (Sawan, Hafez, and Basyony 2001; Specht and Rundel 1990). Visible wavelengths, which are known to indicate K deficiency, had a significant effect on K predictions (Zhai et al. 2013). Furthermore, wavelengths in the chlorophyll absorption features and red-edge are highly related to Mg, which is a chlorophyll component (Gökkaya et al. 2014).

The spectral regions with high correlation coefficients encompass many of the same wavelengths of importance in the PLSR models. However, the degree of similarity between regression coefficients for the reflectance and selected flush models vary by nutrient. Reflectance and flush coefficients for K and Mg models highly match, while P coefficients match moderately. Coefficients for N and Ca models exhibit substantial differences, with high variability in coefficient amplitude and patterns. This disagreement in the Ca models, in conjunction with the discrepancies in the variable importance plots, raise doubts regarding the stability of the Ca model and its physiological underpinnings. Unfortunately, PLSR models are not easily interpretable, particularly with such a large number of variables.

Local analysis

The inability to develop models at Reynolds Homestead or Bladen Lakes is most likely a result of an insufficient environmental gradient. The low nutrient ranges and weak relationships between nutrients and reflectance hinder model development. This finding is consistent with Gong et al's (2012) development of strong local models for sycamore and sweetgum, but inability to do so for loblolly pine. The researchers attribute their difficulty to use red-edge wavelengths exclusively; however, the inclusion of a wide-range

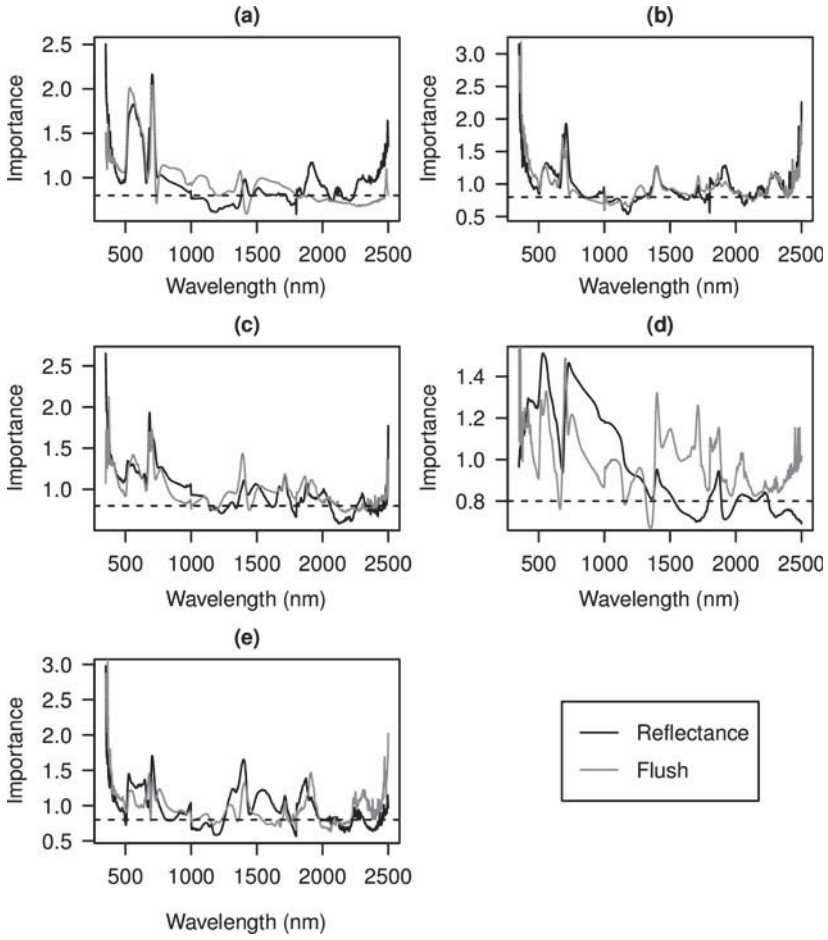


Figure 8. Variable Importance for the Projection (VIP) plots for reflectance and selected partial least squares regression flush models for the following nutrients: (a) nitrogen, (b) phosphorus, (c) potassium, (d) calcium, and (e) magnesium. Dashed line indicates Wold's criterion (0.8). VIP plots indicate the contribution of each variable to model fit across all latent factors in partial least squares regression.

of wavelengths did not resolve the problem. It is possible that the tendency of loblolly pine trees to increase leaf area index when fertilized, rather than foliar N, may confound model development (Gough, Seiler, and Maier 2004). Patterns across the 18 sites in the SE US were apparent because the nutrient ranges were much greater, much as other studies have shown success with interspecific models (Gökkaya et al. 2014; Im et al. 2009).

Spectral differences exist between genotypes and hybrids of different species, as well as in the Bladen Lakes and Reynolds Homestead loblolly pine varieties (Espinoza, Hodge, and Dvorak 2012; Im et al. 2009). However, nutrient differences across genotypes did not match differences in reflectance across genotypes, thus precluding the use of genotype-specific models at Bladen Lakes and Reynolds. Therefore, it is evident that nutrients are not the main driver of spectral differences in these genotypes.

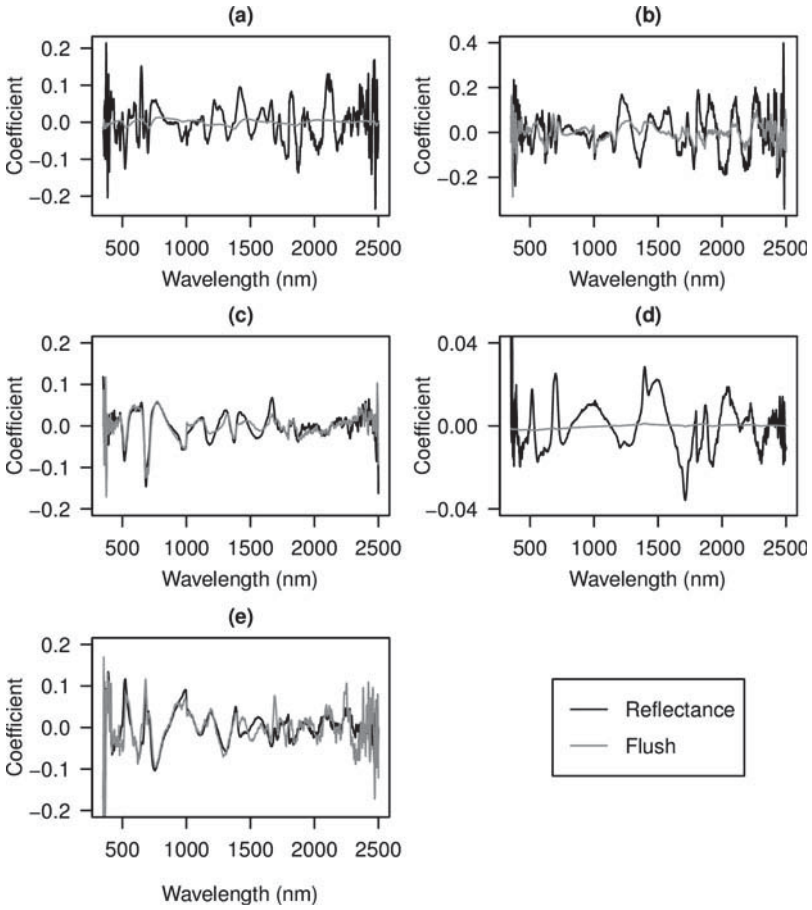


Figure 9. Regression coefficients for reflectance and selected flush partial least squares regression models for nutrients: (a) nitrogen, (b) phosphorus, (c) potassium, (d) calcium, and (e) magnesium. Model coefficients are centered and scaled to improve comparability.

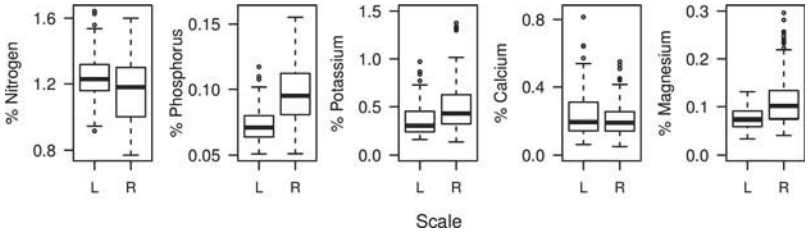


Figure 10. Nutrient distributions of loblolly pine foliar samples across the 18 southeastern US regional ("R") study sites and the localized ("L") study sites in Virginia and North Carolina. Most nutrient ranges are significantly greater at the regional scale.

Conclusion

This study is among the first to successfully identify foliar macronutrients (e.g. N, P, K, Ca, Mg) with spectral reflectance in loblolly pine plantations at the leaf-level scale. Model

accuracies are comparable to those of other species, and the models explain a significant amount of variation in macronutrient concentrations across sites in the SE US. Results indicate that localized loblolly pine nutrient studies, even with fertilization treatments, are less likely to produce successful remote-sensing models than studies across a large geographic region. In addition, the study emphasized the importance of flush in nutrient models using remote sensing. Future work will need to encompass other methods of assessing spectral reflectance, such as continuum-removal or machine-learning techniques. Furthermore, it is critical to determine the relationship between leaf- and canopy-level analyses for future assessment with satellite data.

Acknowledgment

The authors would like to thank Marco Yanez, Jay Raymond, John Peterson, David Mitchem, Evan Brooks, Ranjith Gopalakrishnan, and Will Guthrie for their help with this paper.

Funding

This research was funded by a grant through the USDA National Needs Fellowship, USDA-NIFA-NNF-2010-03349, the McIntire-Stennis Cooperative Forestry Research Program through the USDA CSREES under Project VA-136614, and the Department of Forest Resources and Environmental Conservation at Virginia Tech.

References

Al-Abbas, A. H., R. Barr, J. D. Hall, F. L. Crane, and M. F. Baumgardner. 1974. "Spectra of Normal and Nutrient Deficient Maize Leaves." *Agronomy Journal* 66 (1): 16–20. doi:10.2134/agronj1974.00021962006600010005x.

Asner, G. P. 1998. "Biophysical and Biochemical Sources of Variability in Canopy Reflectance." *Remote Sensing of Environment* 64 (3): 234–253. doi:10.1016/S0034-4257(98)00014-5.

Asner, G. P. 2008. "12 Hyperspectral Remote Sensing of Canopy Chemistry, Physiology, and Biodiversity in Tropical Rainforests." In *Hyperspectral Remote Sensing of Tropical and Sub-Tropical Forests*, edited by M. Kalacska and G. Arturo Sanchez-Azofeifa, 261–296. Boca Raton, FL: CRC Press.

Asner, G. P., R. E. Martin, D. E. Knapp, R. Tupayachi, C. Anderson, L. Carranza, P. Martinez, M. Houcheime, F. Sinca, and P. Weiss. 2011. "Spectroscopy of Canopy Chemicals in Humid Tropical Forests." *Remote Sensing of Environment* 115 (12): 3587–3598. doi:10.1016/j.rse.2011.08.020.

Cawley, G. C., and N. L. C. Talbot. 2003. "Efficient Leave-One-Out Cross-Validation of Kernel Fisher Discriminant Classifiers." *Pattern Recognition* 36 (11): 2585–2592. doi:10.1016/S0031-3203(03)00136-5.

Christensen, L. K., B. S. Bennedsen, R. N. Jørgensen, and H. Nielsen. 2004. "Modelling Nitrogen and Phosphorus Content at Early Growth Stages in Spring Barley Using Hyperspectral Line Scanning." *Biosystems Engineering* 88 (1): 19–24. doi:10.1016/j.biosystemseng.2004.02.006.

Coops, N. C., M. L. Smith, M. E. Martin, and S. V. Ollinger. 2003. "Prediction of Eucalypt Foliage Nitrogen Content from Satellite-Derived Hyperspectral Data." *IEEE Transactions on Geoscience and Remote Sensing* 41 (6): 1338–1346. doi:10.1109/TGRS.2003.813135.

Curran, P. J. 1989. "Remote Sensing of Foliar Chemistry." *Remote Sensing of Environment* 30 (3): 271–278. doi:10.1016/0034-4257(89)90069-2.

Espinoza, J. A., G. R. Hodge, and W. S. Dvorak. 2012. "The Potential Use of Near Infrared Spectroscopy to Discriminate between Different Pine Species and their Hybrids." *Journal of Near Infrared Spectroscopy* 20 (4): 437–447. doi:10.1255/jnirs.1006.

Ferwerda, J. G., and A. K. Skidmore. 2007. "Can Nutrient Status of Four Woody Plant Species be Predicted Using Field Spectrometry?" *ISPRS Journal of Photogrammetry and Remote Sensing* 62 (6): 406–414. doi:10.1016/j.isprsjprs.2007.07.004.

Gholz, H. L., P. J. Curran, J. A. Kupiec, and G. M. Smith. 1997. "Assessing Leaf Area and Canopy Biochemistry of Florida Pine Plantations Using Remote Sensing." In *The Use of Remote Sensing in the Modeling of Forest Productivity*, edited by H. L. Gholz, K. Nakane, and H. Shimoda, 3–22. Netherlands: Springer.

Gökkaya, K., V. Thomas, T. Noland, H. McCaughey, and P. Treitz. 2014. "Testing the Robustness of Predictive Models for Chlorophyll Generated from Spaceborne Imaging Spectroscopy Data for a Mixedwood Boreal Forest Canopy." *International Journal of Remote Sensing* 35 (1): 218–233. doi:10.1080/01431161.2013.866291.

Gong, B., J. Im, J. R. Jensen, M. Coleman, J. Rhee, and E. Nelson. 2012. "Characterization of Forest Crops with a Range of Nutrient and Water Treatments Using AISA Hyperspectral Imagery." *GIScience & Remote Sensing* 49 (4): 463–491. doi:10.2747/1548-1603.49.4.463.

Gough, C. M., J. R. Seiler, and C. A. Maier. 2004. "Short-Term Effects of Fertilization on Loblolly Pine (*Pinus Taeda* L.) Physiology." *Plant, Cell and Environment* 27 (7): 876–886. doi:10.1111/j.1365-3040.2004.01193.x.

Hinzman, L. D., M. E. Bauer, and C. S. T. Daughtry. 1986. "Effects of Nitrogen Fertilization on Growth and Reflectance Characteristics of Winter Wheat." *Remote Sensing of Environment* 19 (1): 47–61. doi:10.1016/0034-4257(86)90040-4.

Im, J., J. R. Jensen, M. Coleman, and E. Nelson. 2009. "Hyperspectral Remote Sensing Analysis of Short Rotation Woody Crops Grown with Controlled Nutrient and Irrigation Treatments." *Geocarto International* 24 (4): 293–312. doi:10.1080/10106040802556207.

Jones Jr., J. B., and W. J. A. Steyn. 1973. "Sampling, Handling, and Analyzing Plant Tissue Samples." In *Soil Testing and Plant Analysis*, edited by L. M. Walsh and J. B. Beaton, 249–270. Madison, WI: Soil Science Society of America.

Karnieli, A., A. Karnieli, and D. J. Bonfil. 2007. "Wheat and Maize Monitoring Based on Ground Spectral Measurements and Multivariate Data Analysis." *Journal of Applied Remote Sensing* 1 (1): 013530. doi:10.1117/1.2784799.

Little Jr., E. L. 1971. *Conifers and Important Hardwoods. Vol. 1 of Atlas of United States Trees*. Washington, DC: U.S. Department of Agriculture.

Martin, M. E., L. C. Plourde, S. V. Ollinger, M. L. Smith, and B. E. McNeil. 2008. "A Generalizable Method for Remote Sensing of Canopy Nitrogen across a Wide Range of Forest Ecosystems." *Remote Sensing of Environment* 112 (9): 3511–3519. doi:10.1016/j.rse.2008.04.008.

Mevik, B. H., and H. R. Cederkvist. 2004. "Mean Squared Error of Prediction (MSEP) Estimates for Principal Component Regression (PCR) and Partial Least Squares Regression (PLSR)." *Journal of Chemometrics* 18 (9): 422–429. doi:10.1002/cem.887.

National Climate Data Center. 2005. "Climate Maps of the United States (CLIMAPS)." US Department of Commerce. http://cdo.ncdc.noaa.gov/cgi-bin/climaps/climaps.pl

Nelson, V. L., D. H. Gjerstad, and G. R. Glover. 1986. "Determining Nitrogen Status of Young Loblolly Pine by Leaf Reflectance." *Tree Physiology* 1 (3): 333–339. doi:10.1093/treephys/1.3.333.

Ollinger, S. V., A. D. Richardson, M. E. Martin, D. Y. Hollinger, S. E. Frolking, P. B. Reich, L. C. Plourde, et al. 2008. "Canopy Nitrogen, Carbon Assimilation, and Albedo in Temperate and Boreal Forests: Functional Relations and Potential Climate Feedbacks." *Proceedings of the National Academy of Sciences* 105 (49): 19336–19341. doi:10.1073/pnas.0810021105.

Pallardy, S. G. 2008. *Physiology of Woody Plants*. 3rd ed. Burlington, MA: Academic Press.

Peterson, D. L., J. D. Aber, P. A. Matson, D. H. Card, N. Swanberg, C. Wessman, and M. Spanner. 1988. "Remote Sensing of Forest Canopy and Leaf Biochemical Contents." *Remote Sensing of Environment* 24 (1): 85–108. doi:10.1016/0034-4257(88)90007-7.

Petisco, C., B. García-Criado, B. R. Vázquez de Aldana, I. Zabalgogeazcoa, S. Mediavilla, and A. García-Ciudad. 2005. "Use of Near-Infrared Reflectance Spectroscopy in Predicting Nitrogen, Phosphorus and Calcium Contents in Heterogeneous Woody Plant Species." *Analytical and Bioanalytical Chemistry* 382 (2): 458–465. doi:10.1007/s00216-004-3046-7.

Pimstein, A., A. Karnieli, S. K. Bansal, and D. J. Bonfil. 2011. "Exploring Remotely Sensed Technologies for Monitoring Wheat Potassium and Phosphorus Using Field Spectroscopy." *Field Crops Research* 121 (1): 125–135. doi:10.1016/j.fcr.2010.12.001.

Sawan, Z. M., S. A. Hafez, and A. E. Basyony. 2001. "Effect of Phosphorus Fertilization and Foliar Application of Chelated Zinc and Calcium on Seed, Protein and Oil Yields and Oil Properties of Cotton." *The Journal of Agricultural Science* 136 (2): 191–198. doi:10.1017/S0021859601008644.

Schlemmer, M., A. Gitelson, J. Schepers, R. Ferguson, Y. Peng, J. Shanahan, and D. Rundquist. 2013. "Remote Estimation of Nitrogen and Chlorophyll Contents in Maize at Leaf and Canopy Levels." *International Journal of Applied Earth Observation and Geoinformation* 25: 47–54. doi:10.1016/j.jag.2013.04.003.

Smith, M. L., M. E. Martin, L. Plourde, and S. V. Ollinger. 2003. "Analysis of Hyperspectral Data for Estimation of Temperate Forest Canopy Nitrogen Concentration: Comparison between an Airborne (AVIRIS) and a Spaceborne (Hyperion) Sensor." *IEEE Transactions on Geoscience and Remote Sensing* 41 (6): 1332–1337. doi:10.1109/TGRS.2003.813128.

Specht, R. L., and P. W. Rundel. 1990. "Sclerophylly and Foliar Nutrient Status of Mediterranean-Climate Plant Communities in Southern Australia." *Australian Journal of Botany* 38 (5): 459–474. doi:10.1071/BT9900459.

Stagakis, S., N. Markos, O. Sykioti, and A. Kyparissis. 2010. "Monitoring Canopy Biophysical and Biochemical Parameters in Ecosystem Scale Using Satellite Hyperspectral Imagery: An Application on a *Phlomis Fruticosa* Mediterranean Ecosystem Using Multiangular CHRIS/PROBA Observations." *Remote Sensing of Environment* 114 (5): 977–994. doi:10.1016/j.rse.2009.12.006.

Stovall, J. P., C. A. Carlson, J. R. Seiler, T. R. Fox, and M. A. Yanez. 2011. "Growth and Stem Quality Responses to Fertilizer Application by 21 Loblolly Pine Clones in the Virginia Piedmont." *Forest Ecology and Management* 261 (3): 362–372. doi:10.1016/j.foreco.2010.10.018.

Susaeta, A., P. Lal, J. Alavalapati, E. Mercer, and D. Carter. 2012. "Economics of Intercropping Loblolly Pine and Switchgrass for Bioenergy Markets in the Southeastern United States." *Agroforestry Systems* 86 (2): 287–298. doi:10.1007/s10457-011-9475-3.

Svotwa, E., B. Maasdorp, A. Murwira, and A. Masuka. 2012. "Selection of Optimum Vegetative Indices for the Assessment of Tobacco Float Seedlings Response to Fertilizer Management." *ISRN Agronomy* 1–10. doi:10.5402/2012/450473.

Taiz, L., and E. Zeiger. 2010. *Plant Physiology.* 5th ed. Sunderland: Sinauer Associates.

Tobias, R. D. 1995. "An Introduction to Partial Least Squares Regression." In *Proceedings of the Twentieth Annual SAS Users Group International Conference*, 2–5. Orlando, FL: SAS Institute.

Townsend, P. A., J. R. Foster, R. A. Chastain Jr., and W. S. Currie. 2003. "Application of Imaging Spectroscopy to Mapping Canopy Nitrogen in the Forests of the Central Appalachian Mountains Using Hyperion and AVIRIS." *IEEE Transactions on Geoscience and Remote Sensing.* doi:10.1109/TGRS.2003.813205.

Tsay, M. L., D. H. Gjerstad, and G. R. Glover. 1982. "Tree Leaf Reflectance: A Promising Technique to Rapidly Determine Nitrogen and Chlorophyll Content." *Canadian Journal of Forest Research* 12 (4): 788–792. doi:10.1139/x82-118.

US Environmental Protection Agency. 2010. *Ecoregions of North America*. http://www.epa.gov/wed/pages/ecoregions/level_iii_iv.htm

Ustin, S. L., A. A. Gitelson, S. Jacquemoud, M. Schaepman, G. P. Asner, J. A. Gamon, and P. Zarco-Tejada. 2009. "Retrieval of Foliar Information about Plant Pigment Systems from High Resolution Spectroscopy." *Remote Sensing of Environment* 113: S67–S77. doi:10.1016/j.rse.2008.10.019.

Wold, S. 1994. "PLS For Multivariate Linear Modeling." In *QSAR: Chemometric Methods in Molecular Design. Methods and Principles in Medicinal Chemistry*, edited by H. de Vaterbeemd, 195–218. Weinheim: Verlag-Chemie.

Yanez, M. Forthcoming. "Effect of Intensive Silviculture on Leaf-Level Physiology of Young Loblolly Pine Varieties in the Virginia Piedmont and North Carolina Coastal Plain." PhD diss., Virginia Tech.

Zhai, Y., L. Cui, X. Zhou, Y. Gao, T. Fei, and W. Gao. 2013. "Estimation of Nitrogen, Phosphorus, and Potassium Contents in the Leaves of Different Plants Using Laboratory-Based Visible and Near-Infrared Reflectance Spectroscopy: Comparison of Partial Least-Square Regression and Support Vector Machine Regression Methods." *International Journal of Remote Sensing* 34 (7): 2502–2518. doi:10.1080/01431161.2012.746484.

Zhang, S., and H. L. Allen. 1996. "Foliar Nutrient Dynamics of 11-Year-Old Loblolly Pine (*Pinus Taeda*) Following Nitrogen Fertilization." *Canadian Journal of Forest Research* 26 (8): 1426–1439. doi:10.1139/x26-159.

3 Chapter 3—Incorporation of a Canopy Structural Correction Parameter in Loblolly Pine Nitrogen Models

Abstract

Recent studies have called into question the previously established relationships between reflectance and nitrogen concentrations, based on the confounding role of structure in mixed forest stands. In our study, we compared nitrogen models derived from reflectance with those derived using a canopy structural correction parameter within the context of a loblolly pine (*Pinus taeda L.*) monoculture. We collected foliar samples from 41 plots in Duke Forest in Chapel Hill, NC, and measured their nitrogen concentration in the laboratory. We also acquired coincident imaging spectroscopy and lidar data from the ESA HyPlant sensor and NASA Goddard's Lidar, Hyperspectral, and Thermal airborne imager. We used the radiative transfer method developed by Knyazikhin et al. (2013) to remove the effect of structure from reflectance measurements, thus obtaining W scattering coefficients. Using several regression-based approaches, we compared the relationships between reflectance and W scattering coefficients with nitrogen. While we found no significant correlations between reflectance and %N, there were many significant correlations between W and %N in the visible and near-infrared wavelengths. However, multiple linear and partial least squares regression models of foliar nitrogen concentrations had similar explanatory power for both variables. These results indicate the need for greater research into these two approaches, particularly in other monospecific stands.

Keywords: nitrogen, canopy structure, W scattering coefficients, radiative transfer,DASF

Introduction

Remote sensing links between spectral reflectance and canopy nutrients have been used for decades to measure forest nitrogen status (Coops et al. 2003; Curran, Kupiec, & Smith 1997; Wessman, Aber, & Peterson 1989). Researchers using regression-based approaches have drawn relationships between reflectance and nitrogen across various landscapes (Martin & Aber 1997; Smith et al. 2003; Townsend et al. 2003). Direct correlations between the variables resulted in significant positive associations (Hollinger et al. 2010; Ollinger et al. 2008). However, recent studies have led to widespread agreement amongst scientists that the positive relationships are actually the result of a confounding variable: forest structural differences between plant functional types (Bartlett et al. 2011; Knyazikhin et al. 2013; Ollinger 2011; Ollinger et al. 2013; Townsend et al. 2013; Wicklein et al. 2012). Many of the previous studies took place in mixed forest stands, leading to nitrogen covariation with forest structure. Consequently, the strongest driver for the aforementioned relationships is broadleaf fraction of leaf area (Knyazikhin et al. 2013). Broadleaf trees have higher levels of nitrogen than needleleaf trees, and also have higher levels of reflectance. Thus, studies with varying percentages of broadleaf and needleleaf trees find a positive relationship at the canopy scale, regardless of the actual relationship between reflectance and nitrogen. This finding recouples canopy and leaf scale relationships, as leaf-level nitrogen consistently correlates negatively with reflectance (Bartlett et al. 2011; Wicklein et al. 2012). Radiative transfer theory also dictates that there should be a negative relationship, since nitrogen absorbs radiation within the leaves (Knyazikhin et al. 2013).

Despite consensus regarding the effect of structure, there is disagreement regarding the implications. Townsend et al. (2013) and Ollinger et al. (2013) state that the previously established relationships provide valuable information that can be used to map and predict

nutrients across variable landscapes. Ollinger et al. emphasizes that nitrogen may still play a causal role in the structure-reflectance relationship, although perhaps through a more indirect route: nitrogen availability affects all levels of plant and canopy traits, as well as site conditions that may determine broadleaf vs. needleleaf composition. Alternatively, Knyazikhin et al. concluded that the prior relationships were spurious and could not be used for nitrogen estimation. Instead, they developed a radiative-transfer approach to quantify the amount of reflectance attributable to foliar biochemical properties (S6.2; Knyazikhin et al. 2013). These values, known as the W scattering coefficients, could then be directly linked to nitrogen concentrations.

The Knyazikhin et al. method divides canopy reflectance into two components: a structural component and a biochemical component. The structural component, known as the directional area scattering factor (DASF), is based on the photon recollision probability theory, which states that photon interactions within the canopy depend exclusively on canopy structure (Panferov et al. 2001; Yáñez Rausell 2014). By removing the effects of structure from the bidirectional reflectance factor (BRF) at each wavelength, the researchers obtained W scattering coefficients. These values quantified the biochemical component of the canopy reflectance, such that any change in W at a given wavelength between plots indicates a difference in foliar absorption properties. Using this methodology, Knyazikhin et al. (2013) found a negative relationship between structurally-corrected reflectance (423-855nm) and %N.

Both perspectives rely on the role of mixed forest stands for nutrient mapping through the use of their structural and/or nutrient gradients. However, to the best of our knowledge, how these methods compare within a single species has yet to be examined. Through this study, we would like to extend the methods developed by Knyazikhin et al. (2013) to learn more about the

relationship between physical characteristics, reflectance, and nutrients within a loblolly pine monoculture. Within monospecific stands, the effect of broadleaf fraction of leaf area on the reflectance-nitrogen relationship becomes irrelevant. Although other structural gradients exist, this situation will remove the main source of confusion. In this context, we hope to compare the accuracy of nitrogen models derived from reflectance with those derived from W scattering coefficients. We wish to investigate the nature of reflectance patterns in the absence of a major structural gradient, the loblolly pine range in DASF values, and how well the structurally-corrected reflectance values predict nitrogen within a single species.

Our goal is to incorporate the W scattering coefficients into the common regression-based methods typically used to predict nitrogen concentrations. Our research questions are as follows:

(1) What is the relationship between forest structural characteristics (i.e., age, stem density, height, crown depth, etc.), nitrogen, and canopy reflectance in loblolly pine stands?

(2) How does the accuracy of nitrogen predictions change when derived from the original reflectance values and the structurally-corrected reflectance (W scattering coefficients)?

Results from this research may have applications in the fields of climatology, forestry, and remote sensing. Specific applications will depend on the predictive capabilities of the various models. Results will determine if the W scattering coefficients have the potential to improve albedo estimates in climate and other ecological process models for loblolly pine. Furthermore, forest nitrogen concentrations derived from remote sensing could be used in photosynthesis and net primary production estimates (Ollinger & Smith 2005; Smith et al. 2002). These techniques

could improve our ability to assess forest health across large landscapes in an efficient and cost-effective manner.

Methods

Study Area

The study area is in the Blackwood Division of the Duke Forest Teaching and Research Laboratory in central North Carolina, US (35.58′ N, 79.8’ W). The site consists of loblolly pine stands planted in various years ranging from 1948 to 2009, and approximately 50 understory species (Luo et al. 2001). Selected trees range in age and environmental characteristics (e.g. soil properties, hydrology, and spatial distribution). Two trees are located within the Duke Free-air Carbon Enrichment (FACE) site, including one fertilized tree. There are 9 soil series, including Enon loam, Herndon silt loam, and Tarrus silt loam. Slopes range from 0 to 25 percent (Duke University 2013). The climate is temperate, with average air temperatures ranging between 5 and 30 degrees C in the winter and summer, respectively. Mean precipitation is 1.06 m (Luo et al. 2001). Loblolly pine site indices ranges from 67 to 95, where higher site indices indicate greater productivity (Amateis & Burkhart 1985; Duke University 2013).

At Duke Forest, we established 36 circular plots with radius of 24 ft (7.31 m) to cover an expansive spatial area with a large environmental gradient, consisting of a wide spectrum of soil characteristics, loblolly pine age classes, and site disturbance properties. Figure 3.1 shows the sampling locations overlaid on a map of soil classes. Plots were located within environmental strata, while considering ease of access and sample collection.

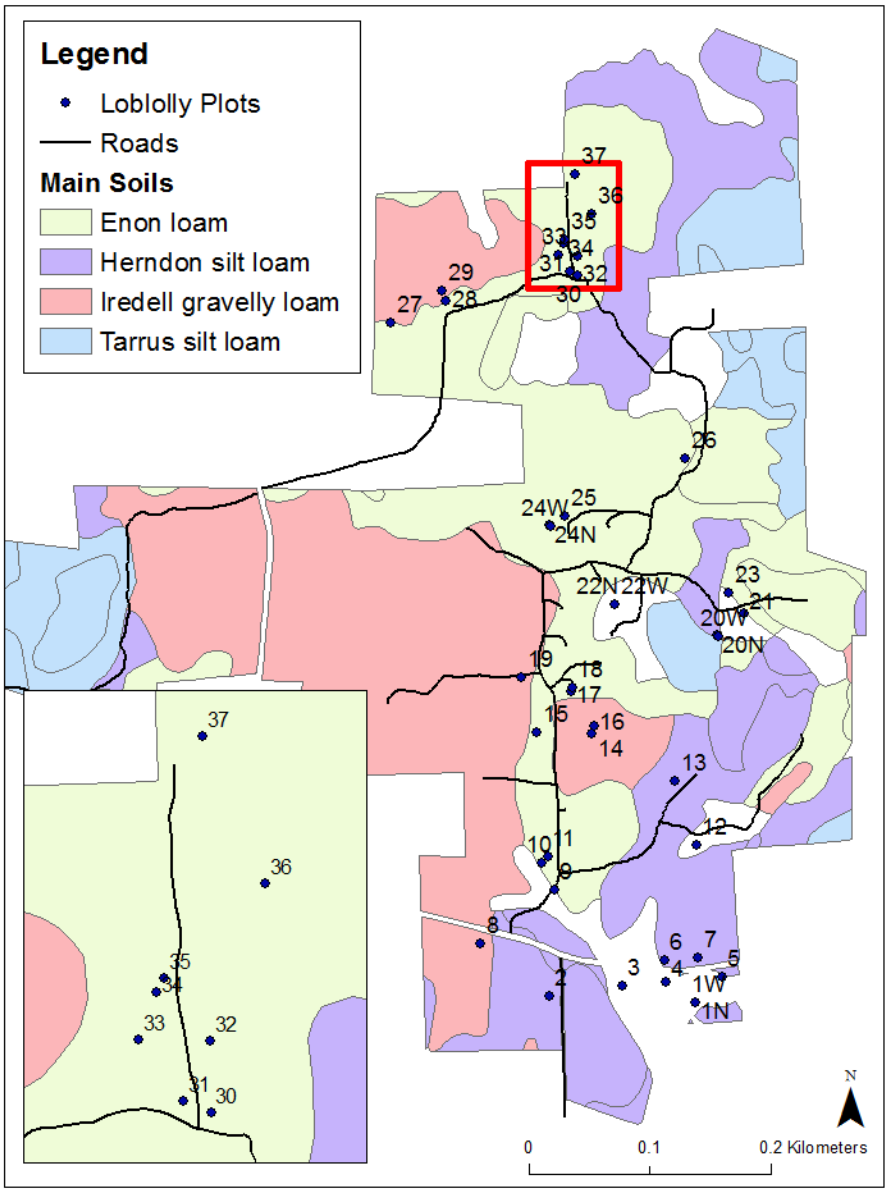


Figure 3.1. Map of Duke Forest sampling plots spanning across range of soil conditions (Duke University 2013).

Field Sampling

Within each plot, 1-2 trees were selected for sampling. Foliar samples were collected from 42 trees at Duke Forest on Sept. 24-26, 2014. Trees were distributed across the geographic and environmental range of the study site, encompassing a variety of age classes, heights, and

soil characteristics. One to three branches per plot were shot down from the sunlit, upper canopy of each tree using a shotgun. All fascicles per branch were removed and grouped by flush across all branches, resulting in sample combinations from as many as three branches and four flushes per tree.

Structural data from each site were collected in the field for the purposes of ground validation and the derivation of additional characteristics (Table 3.1, Appendix A). Tree-level measurements included height, DBH, crown width, and crown depth. Number of trees (live and dead), top of crown height, bottom of crown height, crown width, and DBH for trees over 12.7 cm in diameter were measured for each plot. We used a GPS to record the center point for each tree.

Table 3.1. Descriptive statistics for sample trees at all plots in the Duke Forest.

Descriptive statistics	Value
Average site index	76 (7.15)
Mean height (standard deviation)[1]	17.64 m (8.41 m)
Mean DBH (standard deviation)	0.92 m (0.49 m)
Mean canopy length (standard deviation)[2]	5.33 m (2.96 m)

[1] Derived from lidar data
[2] Measured N-S and E-W

Laboratory Work

We immediately stored foliar samples in a cooler with dry ice while in transit to a deep freezer (-80 °C). Samples remained frozen until laboratory analysis, at which point we transferred them to a drying oven (60 °C) for 48 hours. Samples were ground in a Wiley™ mill (Thomas Scientific, Swedesboro, NJ) and then returned to the drying oven for 24 hours. Nitrogen concentration was measured using dry combustion in a vario Max CN Analyzer (Elementar, Hanau, Germany)(Kiser, Fox, & Carlson 2013). Nitrogen was calculated for all needle age

classes present. However, due to the lack of older age classes on many of the branches, we decided to limit our analysis to the two most recent flushes. We equally weighted the two most recent flushes across each tree since their actual proportions were unknown to us and could not be determined from our small sample size per tree.

Image Acquisition and Processing

Imagery from the HyPlant sensor was collected on Sept. 30, 2014. HyPlant is a push-broom hyperspectral imaging system designed by Forschungszentrum Jülich in Germany (Rossini et al. 2015). The spectral range is from 370 to 2500 nm, with 3 nm spectral resolution in the visible and near-infrared wavelengths, and 10 nm resolution in the short-wave infrared (Forschungszentrum Jülich 2015). Data was delivered in units of reflectance. Three flight lines covered the majority of our loblolly pine plots with 1 m spatial resolution. Due to radiance calibration differences in one of the flight lines, reflectance data from only two of the flight lines were aggregated for modeling. For those plots located within overlapping flight lines, reflectance data from the higher quality flight line was prioritized, with quality defined as the percent of unshadowed imagery. Due to the high prevalence of cloud and intra-canopy gap shadows in the image, we also conducted a maximum likelihood supervised classification for shadow presence/absence. The resulting classification was then used to mask shadows within the image. This process produced what is henceforth referred to as "shadow-masked" images, in comparison with the "original" images. Shadows covered approximately 30 percent of the pixels in the images. For both the original and shadow-masked images, we extracted the mean spectral reflectance from 3 m buffers surrounding the GPS point for each sample tree.

We followed the steps outlined in Knyazikhin et al. (2013) to calculate canopy scattering coefficients, W, for each of the plots. First, we used the Lewis and Disney (2007) approximation of the PROSPECT model as our reference spectrum, with the same input parameters as Knyazikhin et al. (2013) (Figure 3.2). Then we found the linear relationship between the reference spectrum and the extracted plot spectrum between 710 to 790 nm. From the relationship, we determined the slope, intercept, and R^2. We divided the intercept by 1 minus the slope to calculate the DASF at each plot. By dividing the reflectance at each wavelength by the DASF, we obtained the approximate canopy scattering coefficients (W) for all wavelengths at each plot. Figure 3.3 outlines this procedure and the following statistical analysis.

Lidar data were obtained from Goddard's Lidar, Hyperspectral, and Thermal airborne imager (G-LiHT), which was flown concurrently with the HyPlant sensor in September and October of 2013. G-LiHT is an airborne system "that permits simultaneous measurements of vegetation structure, foliar spectra, and surface temperatures" for terrestrial ecosystems (Cook et al. 2013). The instruments collect high resolution (<1m) lidar, optical, and thermal data using lasers, a GPS-INS, up/downwelling spectrometers, and non-ITAR instruments (Cook et al. 2013). We calculated common lidar-derived metrics, such as mean canopy height, maximum canopy height, height percentiles, and percent returns per 1 m voxel.

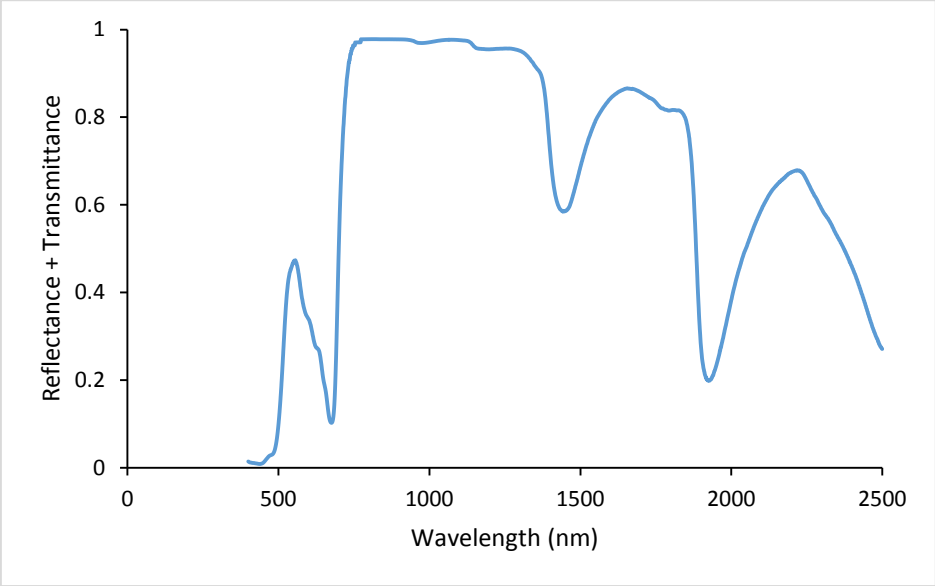


Figure 3.2. Leaf-level reflectance + transmittance spectrum from Lewis-Disney model of PROSPECT.

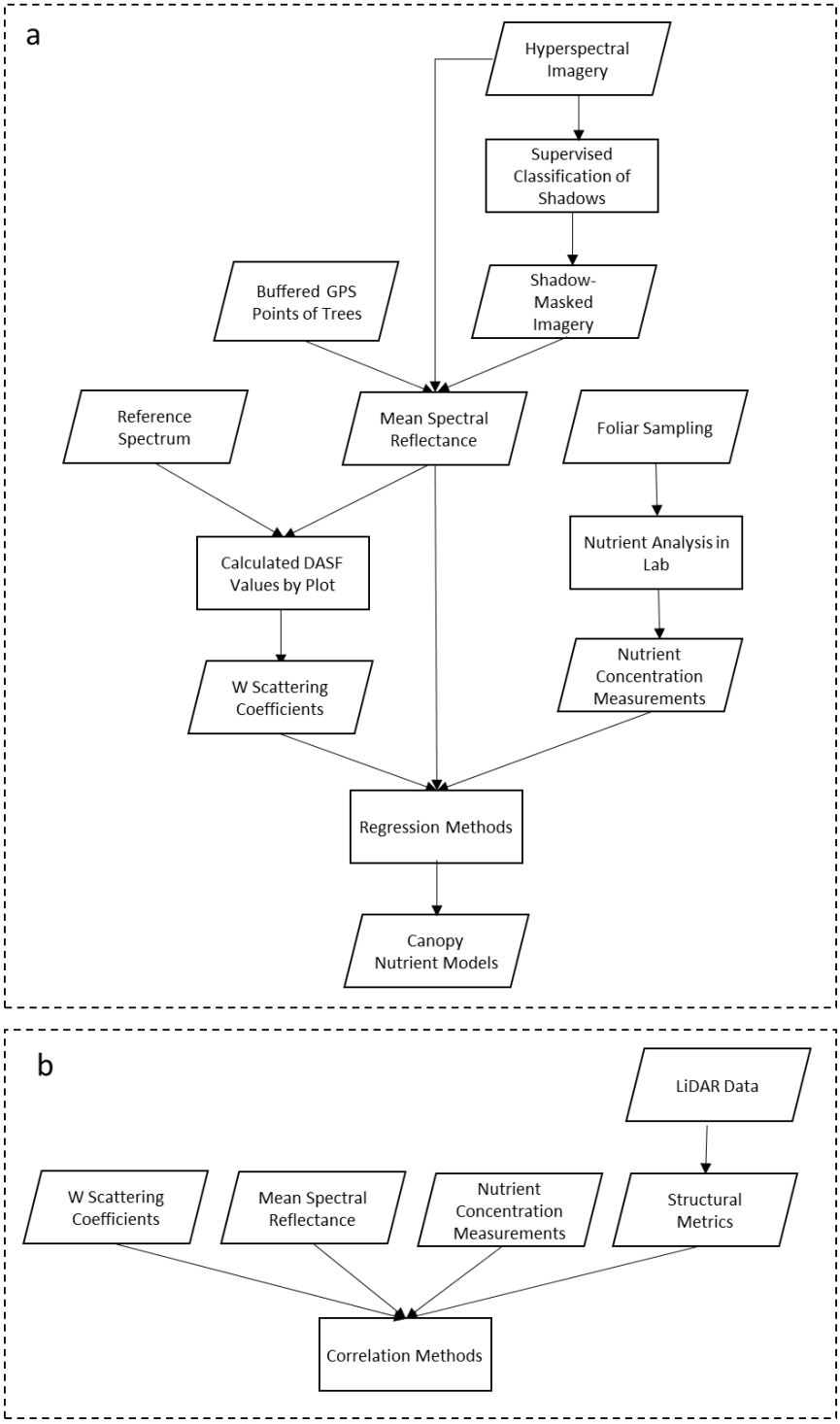


Figure 3.3. Methods involved a two-tiered approach: (a) the development of nutrient models, and (b) the assessment of the relationship between structural variables, spectral reflectance, W scattering coefficients, and nitrogen.

Statistical Analysis

Data were analyzed using Spearman rank correlation coefficients, multiple linear regression, and partial least squares regression. For the correlation analysis, the confidence level was reduced to $\alpha=0.01$ to take into account the increased potential of Type I error from multiple hypothesis testing. Previous work with loblolly pine showed statistically significant differences in foliar nitrogen concentration between flushes (Stein et al. 2014). Thus, models were created to predict Flush 1 N, and Mean Flush 1 & Flush 2 N from reflectance and W values using the original and shadow-masked images. Flush 2 N was not modeled because the older needles are located closer to the trunk on the branch, and are less likely to be visible from the satellite alone without the younger needles. For all models, various subsets of wavelengths and indices were evaluated to determine the optimal combinations. In a recent paper, Thenkabail et al. (2014) proposed the use of numerous non-redundant hyperspectral wavelengths and several new indices to study vegetation. We selected a subset of these wavelengths with linkages to nitrogen to use in our statistical analysis, along with the indices shown in Table 3.2. These wavelengths will henceforth be referred to as "Thenkabail's best wavelengths" (Thenkabail et al. 2014, Table 3.2). Although many of these indices are directly linked to nitrogen, the Enhanced Vegetation Index was selected due to its relationship with leaf area index (LAI), and the tendency of loblolly pine trees to increase their LAI when fertilized with nitrogen (Gough, Seiler, & Maier 2004; Hatfield & Prueger 2010; Xiao et al. 2005). For the partial least squares regression models, reflectance and W values across the entire spectrum were also assessed.

Best subset multiple linear regression models were selected using the minimum Akaike Information Criterion (AIC) and fewest number of variables (Pope & Treitz 2013). When possible, models were limited to a maximum of 3 variables. Residuals were examined for

normality, homoscedasticity, and bias. Multicollinearity was evaluated by ensuring that variance inflation factors (VIFs) did not exceed 10 (Darvishzadeh et al. 2008). In addition, the Shapiro-Wilk Goodness-of-Fit test was used to confirm the validity of this approach; nitrogen values at all plots, excluding outliers, do not reject the null hypothesis of normality (Flush 1 N: $p=0.17$, Flush 1 & 2 N: $p=0.55$). Partial least squares regression models were computed using a maximum of 10 latent factors. Leave-one-out cross-validation was conducted with minimum RMSE used for model selection (Geladi & Kowalski 1986). One sample was excluded from analysis because the tree was dying when the samples were taken, and another sample was excluded from Mean Flush 1 & 2 N analysis because of an unrealistic nitrogen concentration. Several other points were found to be outliers in a few wavelengths, but due to the small sample size and lack of consistency, they were retained in the analysis.

We also evaluated the relationship between structure and nitrogen, and then structure and reflectance/scattering. First, we conducted a correlational analysis using the lidar-derived structural metrics and nitrogen. Next, we conducted a correlational analysis with the structural metrics, the reflectance measurements, and the reflectance-derived variables (e.g. DASF, W, and DASF R^2). Finally, we evaluated the relationship between the DASF and forest field measurements. As DASF is normal according to the Shapiro-Wilk Goodness-of-Fit test ($p=0.63$), we used Pearson correlations.

Table 3.2. Selected imaging spectroscopy indices with their abbreviations, formulas, and sources documenting their links to nitrogen.

Index	Abbrev.	Formula	Source
Hyperspectral Biochemical Index 8	HBCI8	(550-515)/(550+515)	Thenkabail et al. 2014
Hyperspectral Biochemical Index 9	HBCI9	(550-490)/(550+490)	Thenkabail et al. 2014
Hyperspectral Biochemical Index 10	HBCI10	(720-550)/(720+550)	Thenkabail et al. 2014
Hyperspectral Biochemical Index 11	HBCI11	(550-375)/(550+375)	Thenkabail et al. 2014
Hyperspectral Biochemical Index 12	HBCI12	(855-550)/(855+550)	Thenkabail et al. 2014
Hyperspectral Biochemical Index 13	HBCI13	(550-682)/(550+682)	Thenkabail et al. 2014
Normalized Difference Nitrogen Index	NDNI	(log(1/1510)-log(1/1680))/ (log(1/1510)+log(1/1680))	Serrano et al. 2002
Normalized Difference Vegetation Index	NDVI	(800-670)/(800+670)	Ferwerda et al. 2005
Simple Ratio	SR	800/670	Daughtry et al. 2000
Enhanced Vegetation Index	EVI	2.5*(800-670)/(800+6*670-7.5*470 + 1)	Hatfield & Prueger 2010

Results

Nitrogen and Reflectance

Mean nitrogen values at the site were within one standard deviation of previously reported N amounts for loblolly pine in the southeastern US (Albaugh et al. 2010). Table 3.3 summarizes the nitrogen concentrations across the 41 plots within the hyperspectral images.

Table 3.3. Nitrogen summary statistics across the loblolly pine plots (n=41) sampled at Duke Forest. Parentheses denote value after outlier removal.

Nitrogen (%)	Flush 1	Flush 2	Mean Flush 1 & 2	All Flushes
Mean	1.08	1.03 (1.00)	1.06 (1.05)	1.06 (1.04)
Min	0.90	0.79	0.85	0.85
Max	1.50	1.72 (1.33)	1.48 (1.40)	1.48 (1.40)
SD	0.13	0.19 (0.14)	0.14 (0.12)	0.15 (0.13)

Shadow-masked images performed better than original reflectance in the statistical analyses, so only the results from the shadow-masked images are presented here. Reflectance values from the 3m-buffered plots are shown in Figure 3.4. The majority of spectral curves follow the typical shape of a vegetation curve, with reflectance values in the near-infrared region ranging from 0.02 to 0.30. DASF values for the shadow-masked plots ranged from 0.01 to 0.28, with a mean of 0.14 and a standard deviation of 0.07. Lower DASF values resulted from the use of shadowed images, with a lower range (0.04 to 0.15), mean (0.09), and standard deviation (0.03). Retrieval accuracy followed a similar trend. Average R^2 values for shadow-masked and shadowed images were 1.00 and 0.99, respectively, both with a standard deviation of 0.00. As shown in Figure 3.5, our W scattering coefficients are in the expected range of 0 to 1 for wavelengths from 400 to 800 nm.

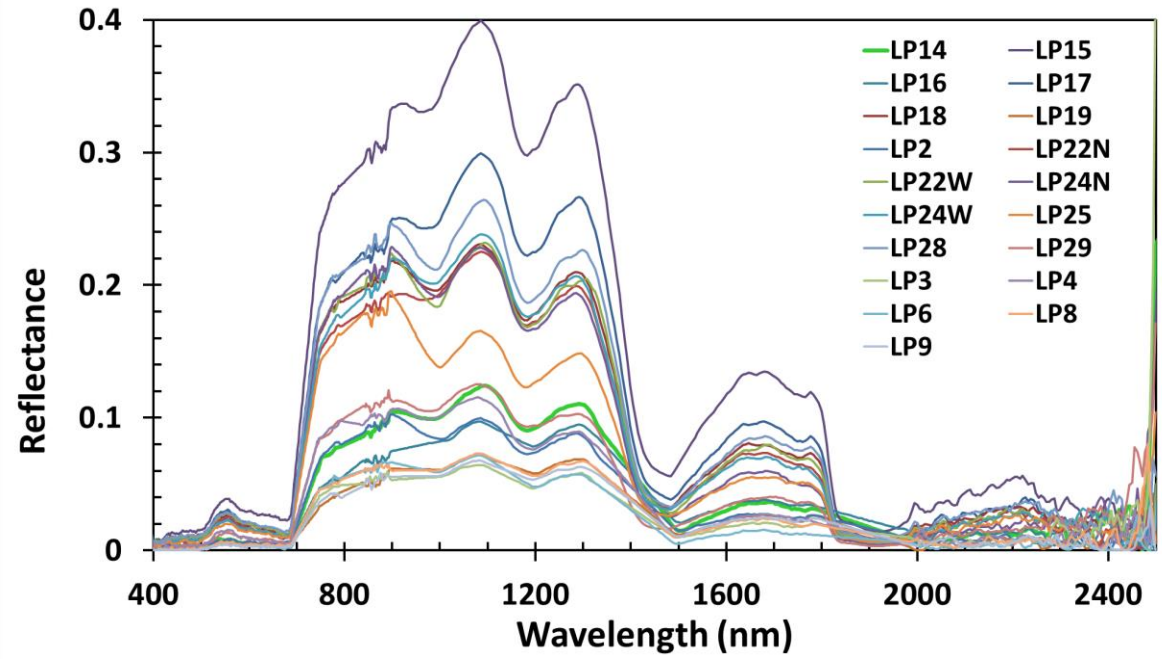


Figure 3.4. Reflectance values from 3m-buffered plots within selected flight lines.

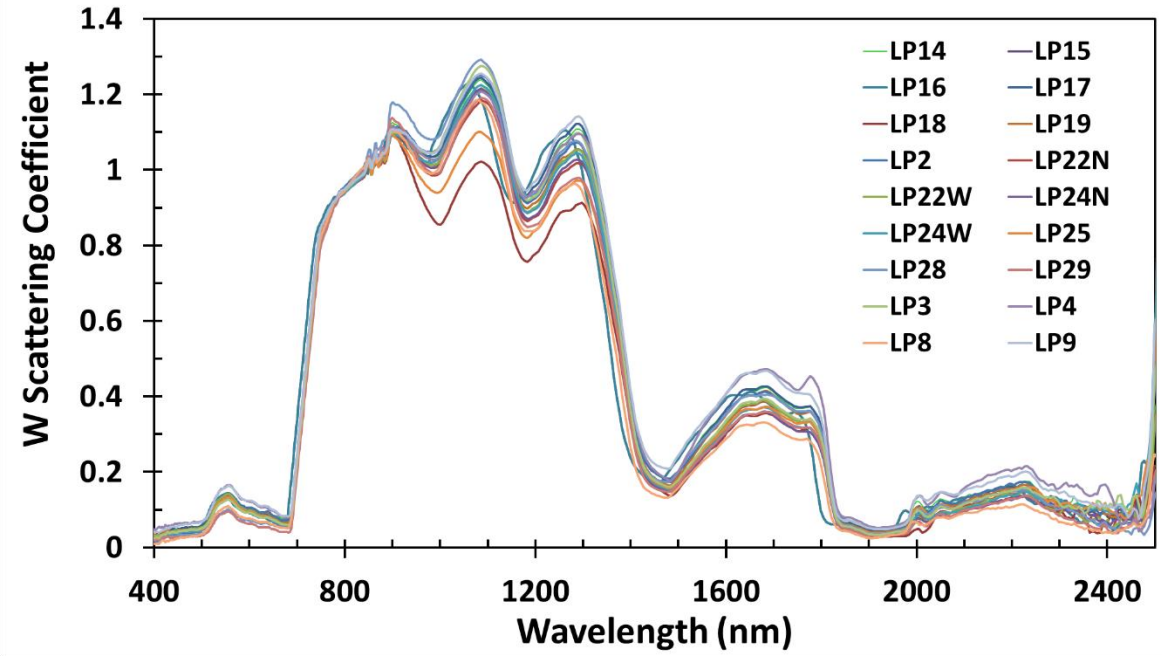


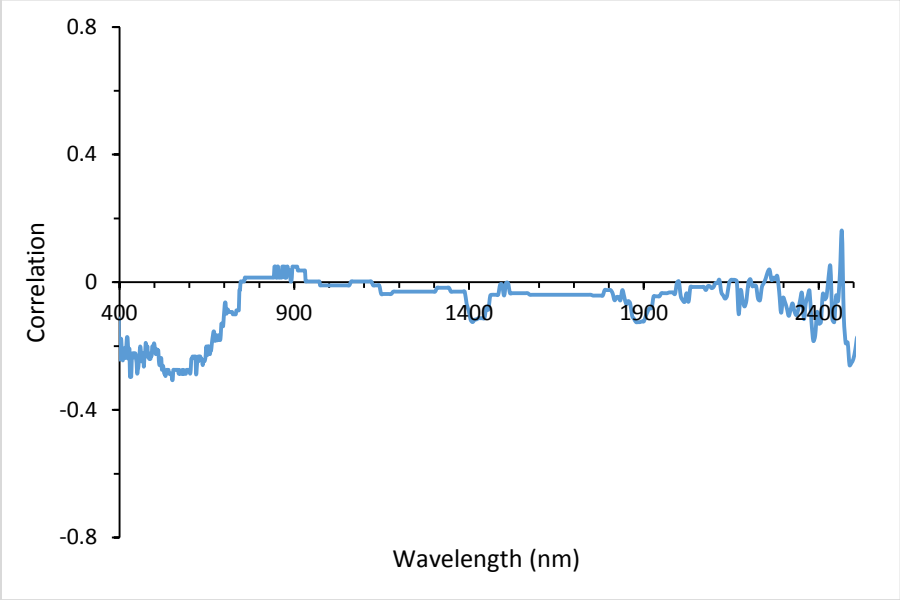
Figure 3.5. Canopy Scattering Coefficients (W) for 3m buffer plots.

Statistical Results

Spearman correlations between Flush 1 N and reflectance or W vary across the spectrum (Figure 3.6). There were no significant correlations for Flush 1 (α=0.01) and reflectance, but significant correlations for W were observed at the following wavelength ranges: 526-530, 533-535, and 544-549 nm. Other wavelengths with strong correlations between foliar N and W include 518-587, 593-595, 717-771, 845, 850, 862-866, 881, and 890-895 nm. Reflectance correlations depict almost an exclusively negative association between Flush 1 N and reflectance, whereas the W correlations are more variable between spectral regions. For the Mean Flush 1 & 2 N values, there were no significant correlations (α=0.01) for either Reflectance or W (Figure 3.7). However, similar wavelengths as above were significant at a reduced level of confidence (α=0.05).

We did not find any significant correlations between nitrogen and the lidar-derived metrics, reflectance-derived structural variables (e.g. DASF and DASF R^2), or W. DASF for the shadow-masked plots were significantly correlated with the number of trees at each plot (R=0.87, p=0.01) and average crown width (R=-0.77, p=0.04). However, an effort to incorporate lidar metrics in the best subset regressions produced little to no improvement in the models. In the few cases where the metrics did appear, they were not particularly intuitive.

(a)



(b)

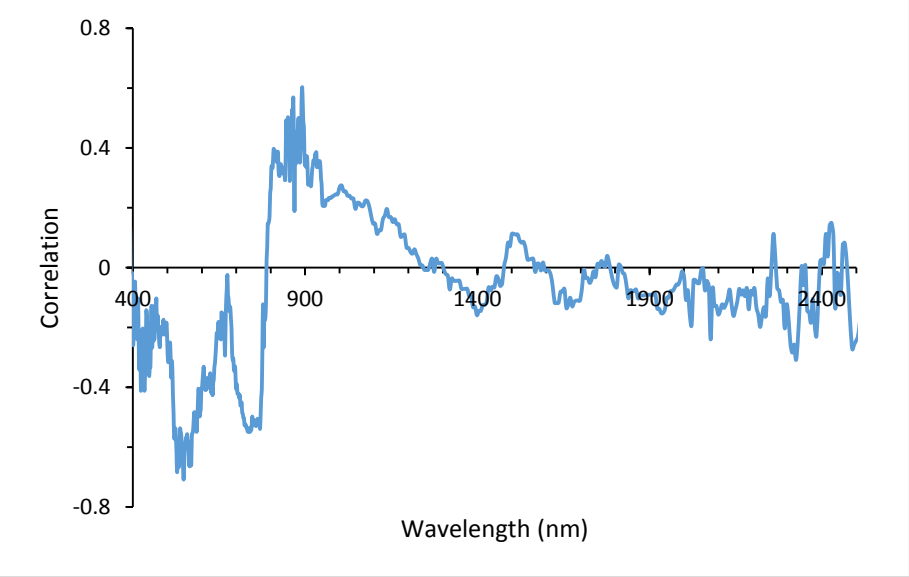
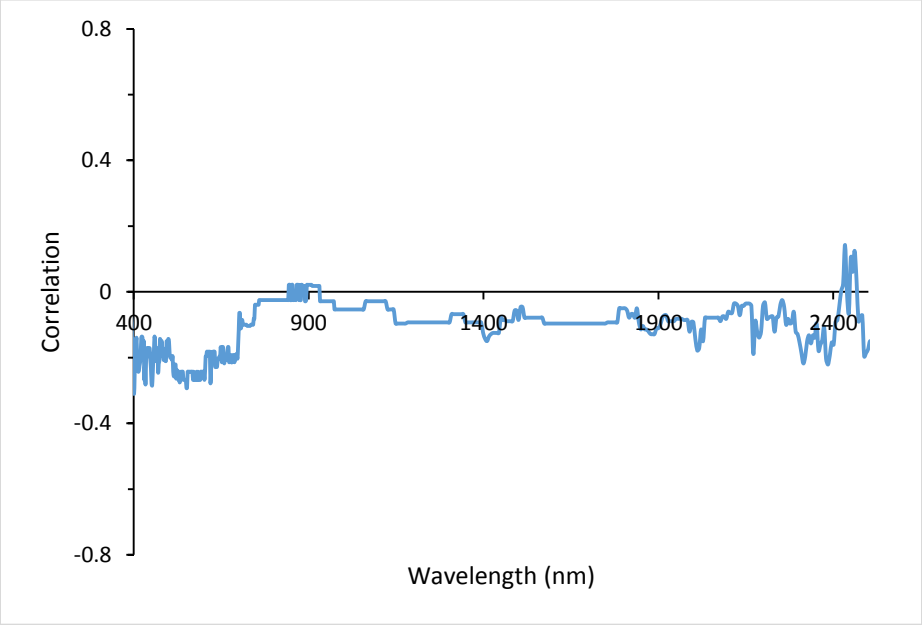


Figure 3.6. Spearman correlation coefficients between (a) reflectance and Flush 1 N and (b) W and Flush 1 N at loblolly pine plots (n=18).

(a)



(b)

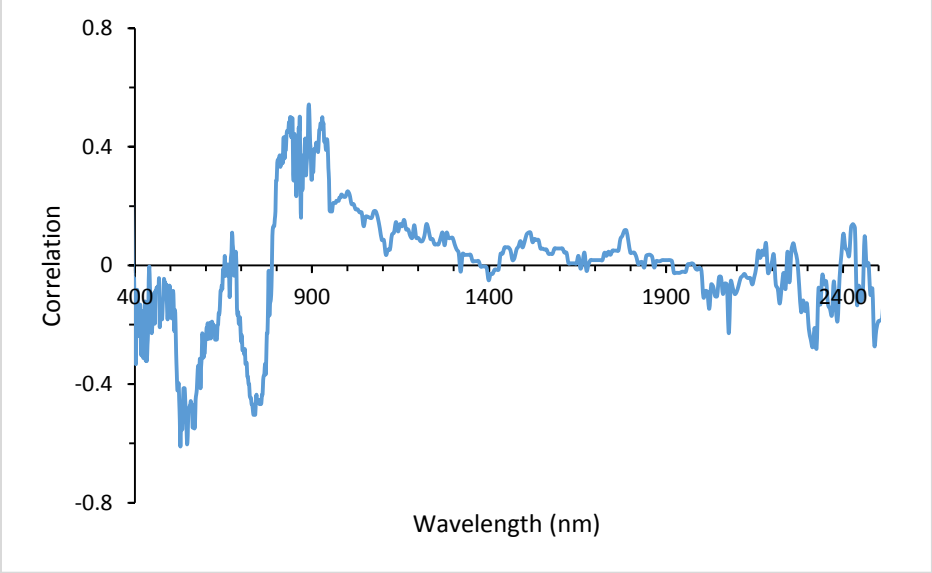


Figure 3.7. Spearman correlation coefficients between (a) reflectance and Mean Flush 1&2 N and (b) W and Mean Flush 1&2 N at loblolly pine plots (n=17).

Multiple linear regression results differed between reflectance and W scattering coefficients for wavelengths and indices. Table 3.4 depicts the selected models for Flush 1 N and Mean Flush 1 & 2 N. No regression model could be produced using Thenkabail's wavelengths for the reflectance values, due to multicollinearity and lack of variable significance in models. Although the ratio indices produced the same values for reflectance and W, the other indices affected model selection. Consequently, models using reflectance values for indices produced more accurate predictions of foliar nitrogen than those using W values (e.g. Flush 1 RMSE=0.07 vs 0.08; R2=0.78 vs. 0.66). Regression analyses on the original images, without the shadow-masking, produced low-accuracy models with indices and/or Thenkabail's wavelengths for Flush 1 N (RMSE>0.12). A separate regression analysis for the 10 shadowed plots did not improve error rates for Flush 1. Decent Mean Flush 1 & 2 regression models exist for the shadowed plots by themselves; however, the relatively small number of observations raises questions regarding their viability (n=8).

Table 3.4. Multiple linear regression models of Flush 1 N (n=17) and Mean Flush 1 & 2 N (n=16) with hyperspectral indices and reflectance at selected wavelengths (nm, indicated by "R" followed by the wavelength). Model type refers to the use of reflectance values (R) or W scattering coefficients in the models.

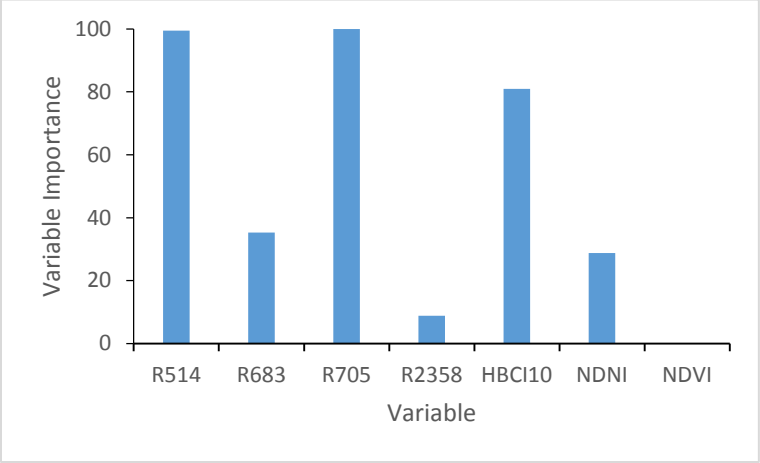
Flush	Model Type	Variable Type	Predictors	R^2	Adj R^2	RMSE	PRESS, PRESS RMSE
1	R	Indices	HBCI12, NDNI, NDVI	0.78	0.73	0.07	0.08, 0.07
1	W	Thenkabail Wavelengths	R683, R705	0.58	0.52	0.09	0.15, 0.09
1	W	Indices	HBCI13, NDNI, EVI	0.66	0.58	0.08	0.14, 0.09
1 & 2	R	Indices	HBCI12, NDNI, NDVI	0.79	0.74	0.06	0.05, 0.06
1 & 2	W	Thenkabail Wavelengths	R683, R705	0.69	0.64	0.07	0.08, 0.07
1 & 2	W	Indices	HBCI10, HBCI13	0.57	0.49	0.08	0.10, 0.08

Partial least squares regression models were developed using all wavelengths, Thenkabail's best wavelengths, indices, DASF, and various combinations of these subsets. PLSR models using all wavelengths and Thenkabail's best wavelengths produced models of low predictive capabilities (RMSE=0.10-0.14, Validation R^2=0.01-0.42). Indices and Thenkabail's best wavelengths produced stronger models (RMSE=0.06-0.09, Validation R^2=0.53-0.72). However, models with the least error were obtained using subsets of Thenkabail's best wavelengths and indices. These optimal models, presented in Table 3.5, outperformed the MLR models. They had lower RMSE values (0.04-0.05) and higher cross-validated coefficients of determination for most models (R^2=0.78-0.87). DASF was also included to incorporate a structural variable, and in some cases, this led to improvement. Figure 3.8 shows the variable importance for the optimal PLSR reflectance and W models for Flush 1 N. No viable PLSR models were produced with the 10 shadowed plots, as RMSE values were high (RMSE>0.12).

Table 3.5. Partial least squares regression models selected by first local minimum root mean square error (RMSE) for Flush 1 and Mean Flush 1 & 2 N. Variables include wavelengths and indices selected from Thenkabail et al. (2014), and several other common vegetation indices.

Type	Variable	Wavelengths and/or Indices	N Comp	RMSE	Validation R^2
Reflectance	Flush 1 N	R514, R683, R705, R2358, HBCI10, NDNI, NDVI	6	0.05	0.87
Reflectance	Flush 1 & 2 N	R514, R552, R683, R705, R722, HBCI13, NDNI, SR	6	0.04	0.86
W	Flush 1 N	R404, R514, R532, R552, R569, R683, R2358, DASF, NDNI, NDVI, EVI	8	0.05	0.87
W	Flush 1 & 2 N	R406, R514, R533, R552, R569, R683, DASF, NDNI	4	0.05	0.78

(a)



(b)

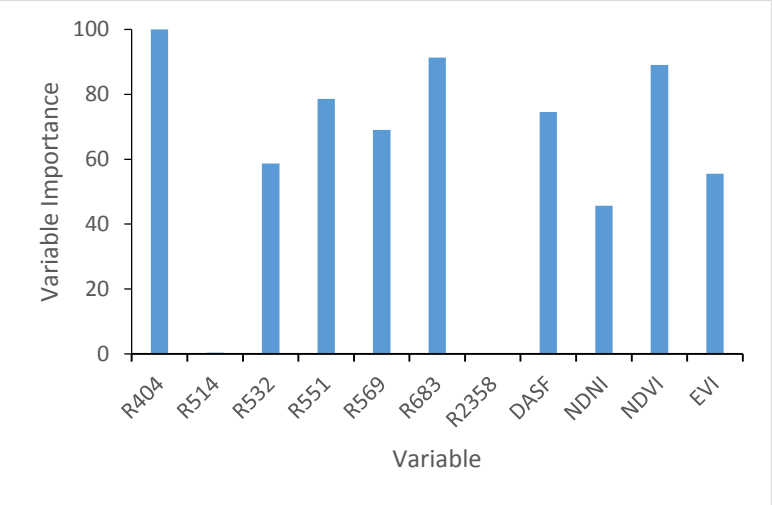


Figure 3.8. Variable importance for the selected partial least squares regression models of Flush 1 N using (a) reflectance and (b) W scattering coefficients. Many models produced results of similar accuracy and predictive capabilities; however, these graphs provide examples of variable weightings in two models with the lowest first local minimum RMSE.

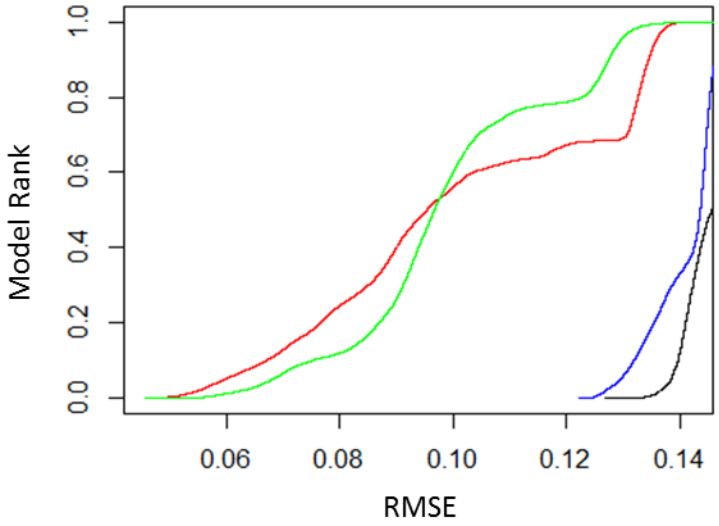


Figure 3.9. Root mean square error (RMSE)-rank distributions for partial least squares regression models selected based on the lowest first local minimum RMSE. Graph depicts the distributions for models produced using reflectance values and W scattering coefficients for both the original and shadow-masked images (black: original reflectance, blue: original W scattering coefficients, red: shadow-masked reflectance, green: shadow-masked W scattering coefficients).

For MLR, the same variables are in both Flush 1 N and Mean Flush 1 & 2 N models for W. Similar variables were selected between Reflectance and W models. In the selected PLSR models, different indices and wavelengths were selected than those in the MLR models. W MLR models used reflectance at 514 and 531 nm, whereas PLSR models used more of the wavelengths and put less emphasis on those two bands in particular. Similarly, the reflectance MLR used fewer variables than the reflectance PLSR models, which used 7-8 variables and 6 components.

Figure 3.9 compares the various model types: original reflectance, original W, shadow-masked reflectance, and shadow-masked W. The graph shows a high disparity between the original and shadow-masked plots, with the latter greatly outperforming the former. A

comparison of Reflectance and W methods depicts lower RMSE with the best W models than the best Reflectance models, although exact results depend on model selection processes.

Discussion

Correlational Analysis

The correlational results depict a clear difference in the relationships between reflectance and the W scattering coefficients with nitrogen. While we found no significant correlations between reflectance and %N, there were many significant correlations between W and %N. This finding supports the use of the W scattering coefficients by Knyazikhin et al. (2013). Significant correlations for W contain Thenkabail et al.'s specified 531 and 550 nm features. The former is related to light use efficiency, chlorophyll, and stress; the latter is mainly attributed to chlorophyll. Although these are not specifically directed at nitrogen, nitrogen concentrations are known to affect each of these attributes (Peltoniemi et al. 2012). Similarly, Knyazikhin et al. (2013) present moderate relationships between W and nitrogen at 557 nm (R^2=0.49). Previous studies also found strong relationships between reflectance at 550 nm and nitrogen concentrations, leading to its frequent use in indices (Blackmer et al. 1996; Moran et al. 2000; Thenkabail et al. 2014).

Reflectance and W correlations with nitrogen in the NIR wavelengths differ from those in prior remote sensing literature in both strength and direction. Correlations between foliar nitrogen and reflectance in this region of the spectrum have repeatedly been positive (Martin et al. 2008; Ollinger et al. 2008). Contrarily, our results showed insignificant positive and negative relationships with nitrogen for reflectance in the NIR wavelengths. Also, whereas Knyazikhin and colleagues found negative associations between the scattering coefficient and %N at every

wavelength from 423-855 nm, our nitrogen relationships with W coefficients were insignificantly positive between 788 and 855 nm. The lack of significance at a 95% confidence level may indicate a superficial relationship, especially given the strong negative trend for the majority of wavelengths in the range. However, as few other studies to date have examined the W coefficients, it is difficult to ascertain why the correlations exhibited a positive trend. Possibilities include radiometric errors from the sensor, an insufficient number of observation plots, low nutrient ranges, and inaccurate DASF values. A discussion of the accuracy of the DASF values is included below.

In the visible wavelengths, reflectance correlation plots show a prevalence of slightly negative relationships, whereas the W plot shows a stronger negative trend. This finding is in agreement with Knyazikhin et al. (2013), who found insignificant and variable relationships between %N and reflectance across the visible wavelengths, but moderate, negative relationships with W. Leaf-level results throughout the literature show a similar reflectance pattern (e.g. Ollinger et al. 2008, Ollinger 2011). While most of our wavelengths had negative correlations, there were several "outliers" with insignificant positive relationships to %N. The relationships were so low that they do not warrant concern, but we did seek a reasonable explanation. This difference could be due to our use of a different sensor or the possibility of minor radiometric errors. Since Knyaziknin et al.'s plots represented the average reflectance across a much larger area than our own plots, their data might have also smoothed out any anomalous wavelengths, canopy shadows, or other sources of error in reflectance measurements.

The W correlations also differed from the study by Knyazikhin et al. (2013) in terms of their strength. Knyazikhin and colleagues found no relationship between BRF and canopy nitrogen (R^2=0.01), but a moderate one with W (R^2=0.43). The strength of this association is a

far cry from our own, which peaked with R^2 values approximately half of those they achieved. However, to obtain their strong linear relationships, Knyazikhin et al. (2013) used a larger geographic range and many different tree species. Our study attempted to use 42 plots of monospecific trees within Duke Forest, indicating that either a larger nutrient or structural gradient may be necessary to obtain better models. We did produce strong predictive models, but they involved much more extensive statistical analysis and use of variable combinations through multivariate approaches. Model development in monospecific stands within a small geographic region is often difficult due to the low natural nutrient ranges (Stein et al. 2014).

Regression Models

Multiple linear regression and partial least squares regression models were moderately predictive for the small number of samples. MLR models explained much of the variation in foliar nitrogen using three or fewer variables. The inability to develop models with reflectance wavelengths supports the lack of significant wavelengths in the reflectance correlations. However, these results lead us to question the robustness of the index-derived models. PLSR models were comparable to those in other studies for nitrogen with respect to their leave-one-out cross-validation R^2 and RMSE values (Asner et al. 2011; Im et al. 2009; Petisco et al. 2005; Pimstein, Karnieli, & Bonfil 2007). Using the first local minimum RMSE as the criterion for model selection, we obtained models with similar numbers of components as in other studies (Darvishzadeh et al. 2011; Im et al. 2009).

PLSR models worked best with a subset of variables and indices linked to nitrogen. This finding is in agreement with the studies by Pimstein et al. (2007) and Darvishzadeh et al. (2011), which found that variable subsets outperformed the entire spectrum. Similarly, Thulin et al.

(2012) concluded that PLSR processing methods, including subsetting, can have a major effect on the outcome of the models. Our PLSR models using wavelengths across the entire spectrum produced poor results. The use of indices greatly improved the models, particularly in combination with Thenkabail band combinations. Although we presented the optimal PLSR models and their associated variable importance plots, we would like to emphasize that many good models could be produced using various subsets of Thenkabail's wavelengths and indices, as shown in Figure 3.9. The shallow slope of the top-ranked variable combinations in the graph indicates that there is a considerable difference in RMSE between the highly ranked variable combinations, as opposed to some of the mid-ranked combinations, which increase in rank at a faster rate relative to their RMSE. However, hundreds of model combinations produced comparable error rates and coefficients of determination. Consequently, we place little importance on the precise combination of variables in the optimal models and their variable importance weights.

Reflectance vs. W Scattering Coefficients

According to the Knyazikhin et al. (2013) study, the relationships that we draw using the reflectance data are structural artifacts. However, their argument is based on the use of different percentages of broadleaf and needleleaf trees in forest nitrogen models. Consequently, the question arises as to what extent structural differences in monospecific plots would drive reflectance-nitrogen relationships. Across the broadleaf to needleleaf spectrum, the effect of shoot self-shading and increased multiple collisions within needleleaf trees leads to decreased reflectance and nitrogen. However, within an individual species, these effects should be reduced, along with a decrease in the strength of the relationship. Thus, we would expect that the

convolution of the nitrogen-reflectance relationship would not be as affected as in other studies with multiple species. While it is possible that our models with nitrogen and reflectance in a monospecific stand are indicative of different levels of self-shading, shoot-to-area clumping, etc., our reflectance models may be more defensible than those with multiple species. The intraspecific tree geometry of loblolly pine does not appear to affect reflectance in a straight-forward manner. Unlike in the three plots used by Knyazikhin, our DASF values are not correlated with nitrogen, indicating that increased nitrogen does not augment tree structural components in such a way as to consistently affect reflectance. We clearly are not seeing much correlation between nitrogen on reflectance without taking structure into account, either.

Knyazikhin et al. (2013) argue that the canopy-level reflectance may not be directly linked to nitrogen due to the confounding nature of structure. Based on this study, we do not have the ability to evaluate the merit of these claims; we can only provide a comparison of the reflectance and W relationships using advanced statistical methods. Results from the correlational analysis seem to indicate no relationship between reflectance measurements at the canopy scale and foliar nitrogen, whereas W scattering coefficients at multiple wavelengths are significantly related. This finding lends support to the theory that without a large structural gradient, relationships between reflectance and nitrogen may be unclear. However, regression models do not present a clear conclusion. With the vegetation indices, reflectance variables produced slightly stronger nitrogen models than W coefficients in the top-ranked models using similar variable subsets. Yet in other cases, reflectance variables were unable to produce valid models with wavelengths supported by previous leaf-level studies and the W coefficients.

Nutrients, Tree Structure and DASF

The lack of relationship between nitrogen and structure was consistent with both the lidar- and reflectance-derived structural metrics, such as DASF. This finding indicates one of several possibilities: (1) an insufficient structural gradient between loblolly pine plots, (2) an inability to describe structural differences with the selected lidar metrics or DASF, and/or (3) an insufficient or inconsistent nitrogen gradient between plots. Although Knyazikhin et al. (2013) found distinct differences in structure between needleleaf, mixed, and broadleaf plots, it is likely that the structural gradient within loblolly pine plots is notably less significant than the inter-specific structural gradient. The accepted differences in tree geometry and architecture that occur between species, and their corresponding effects on canopy reflectance, is a plausible cause (Ollinger 2011). However, there are definitely intraspecific differences between the sample trees, even if they are less obvious. These characteristics include traits that may not be adequately described by our lidar metrics, such as the amount of leaf clumping or number of branches. Yet DASF was intended to discern these types of traits, and our large range in DASF values is indicative of structural differences between plots. Thus, perhaps the observed nitrogen patterns are related to other confounding factors. Although nitrogen variability between our plots is large enough to cause structural differences in loblolly pine trees based on evidence at other sites (Vose & Allen 1988), age differences between trees may have superseded any impacts of nitrogen differences (Will et al. 2002). However, the effect of age would have resulted in a strong correlation between height and DASF, which did not happen. Thus, an alternative hypothesis is that foliage undergoing seasonal nutrient translocation in the fall may not accurately indicate nutrient concentrations present at the peak of the growing season, when structural differences occur (Zhang & Allen 1996).

Our large range in DASF values suggests a large intra-specific species gradient in structure that was not expected. While the upper range of our DASF values is similar to the needleleaf plot in the Knyazikhin study (DASF=0.27), the majority of our plots fell well below this value. Many plots had very low DASF values, indicating little effect of tree structure on reflectance in the absence of biochemical constituents. However, as few studies have examined DASF ranges by species, we have little means for comparison (Knyazikhin et al. 2013; Latorre-Camona et al. 2014). One hypothesis is that DASF is not exclusively representative of structure in the case of our loblolly pine plots. We are not sure what caused the significant negative correlation between DASF and crown width. The relationship with number of trees at each plot may be confounded by crown width, or could indicate that the 3m buffer included trees beyond the sample tree, increasing both the amount of structural variation and unaccounted nitrogen. Alternatively, the small sample size of field measurements may also have affected the DASF correlations.

Additional factors may also influence these relationships, including the internal leaf structure, other biochemical traits, and external variables. Some of Thenkabail's best wavelengths that were selected are tied to biophysical characteristics, rather than biochemical. For example, the reflectance and W values at 683 nm appeared in multiple of the selected models; however, this wavelength is linked to "biophysical quantities and yield" such as LAI, biomass, tree height, and crop type (Thenkabail et al. 2014). Furthermore, the remaining variation in DASF retrieval quality can possibly be attributed to background reflectance from the understory (Stenberg, Mottus, & Rautiainen 2008) and leaf cellular structure (Ollinger et al. 2011). The DASF is believed to take into account tree geometric factors, such as leaf orientation, leaf angle, leaf area index, crown shape, crown size, and clumping. However, sources of

uncertainty may also have an effect, including atmospheric correction issues, mixed pixels, shadow effects, and georegistration. The retrieval quality of the DASF from the relationship between our reference and plot spectrums indicate a good fit, although some plots may be slightly affected by background noise.

Conclusion

In this study across loblolly pine plots, we used W scattering coefficients to successfully model foliar nitrogen concentrations using multiple linear and partial least squares regression. However, the W scattering coefficient is not a perfect solution for modeling nitrogen in monospecific stands. Whereas W variables previously explained over 40 percent of the variation across needleleaf, mixed, and broadleaf forests, we found that they explained only half that amount of variation in our plots. Furthermore, the range in DASF values was abnormally high for a monospecific stand, indicating the influence of other additional factors or that the presupposition of DASF as the equivalent to structure may not hold under these conditions. Regardless, the correlation results produced much stronger relationships with the W scattering coefficients than reflectance values across the spectrum, although the direction of the trends in the near-infrared spectrum raised questions regarding their validity. In spite of this discrepancy, the other statistical approaches showed a moderate-to-strong ability to model foliar nitrogen with both the reflectance values and W scattering coefficients. We believe more experimentation in the future with larger sample sizes, cross-validation across multiple sites, and an examination of the relationships between W scattering coefficients and nitrogen would be useful. In addition, a closer look at the connection with other nutrients may also be beneficial.

Acknowledgements

The authors would like to thank our colleagues in the NASA/ESA Joint Campaign and the Duke Forest staff for their assistance with fieldwork, image acquisition, and data processing. This research was funded by a grant through the USDA National Needs Fellowship, USDA-NIFANNF-2010-03349, the McIntire-Stennis Cooperative Forestry Research Program through the USDA CSREES under Project VA-136614, and the Department of Forest Resources and Environmental Conservation at Virginia Tech.

References

Albaugh, Janine M, Blevins, Leandra, Allen, H Lee, Albaugh, Timothy J, Fox, Thomas R, Stape, José L, & Rubilar, Rafael A. (2010). Characterization of foliar macro-and micronutrient concentrations and ratios in loblolly pine plantations in the southeastern United States. *Southern Journal of Applied Forestry*, *34*(2), 53-64.

Amateis, Ralph L, & Burkhart, Harold E. (1985). Site index curves for loblolly pine plantations on cutover site-prepared lands. *Southern Journal of Applied Forestry*, *9*(3), 166-169.

Asner, Gregory P, Martin, Roberta E, Knapp, David E, Tupayachi, Raul, Anderson, Christopher, Carranza, Loreli, . . . Weiss, Parker. (2011). Spectroscopy of canopy chemicals in humid tropical forests. *Remote Sensing of Environment*, *115*(12), 3587-3598.

Bartlett, Megan K, Ollinger, Scott V, Hollinger, David Y, Wicklein, Haley F, & Richardson, Andrew D. (2011). Canopy-scale relationships between foliar nitrogen and albedo are not observed in leaf reflectance and transmittance within temperate deciduous tree species. *Botany*, *89*(7), 491-497.

Blackmer, Tracy M, Schepers, James S, Varvel, Gary E, & Walter-Shea, Elizabeth A. (1996). Nitrogen deficiency detection using reflected shortwave radiation from irrigated corn canopies. *Agronomy journal, 88*(1), 1-5.

Cook, Bruce D, Nelson, Ross F, Middleton, Elizabeth M, Morton, Douglas C, McCorkel, Joel T, Masek, Jeffrey G, . . . Montesano, Paul M. (2013). NASA Goddard's LiDAR, Hyperspectral and Thermal (G-LiHT) Airborne Imager. *Remote Sensing, 5*(8), 4045-4066.

Coops, Nicholas C, Smith, M-L, Martin, Mary E, & Ollinger, Scott V. (2003). Prediction of eucalypt foliage nitrogen content from satellite-derived hyperspectral data. *Geoscience and Remote Sensing, IEEE Transactions on, 41*(6), 1338-1346.

Curran, Paul J, Kupiec, John A, & Smith, Geoffrey M. (1997). Remote sensing the biochemical composition of a slash pine canopy. *Geoscience and Remote Sensing, IEEE Transactions on, 35*(2), 415-420.

Darvishzadeh, Roshanak, Atzberger, Clement, Skidmore, Andrew, & Schlerf, Martin. (2011). Mapping grassland leaf area index with airborne hyperspectral imagery: A comparison study of statistical approaches and inversion of radiative transfer models. *ISPRS Journal of Photogrammetry and Remote Sensing, 66*(6), 894-906.

Darvishzadeh, Roshanak, Skidmore, Andrew, Schlerf, Martin, Atzberger, Clement, Corsi, Fabio, & Cho, Moses. (2008). LAI and chlorophyll estimation for a heterogeneous grassland using hyperspectral measurements. *ISPRS journal of photogrammetry and remote sensing, 63*(4), 409-426.

Daughtry, C. S. T., Walthall, C. L., Kim, M. S., De Colstoun, E. B., & McMurtrey, J. E. (2000). Estimating corn leaf chlorophyll concentration from leaf and canopy reflectance. *Remote Sensing of Environment*, *74*(2), 229-239.

Duke University. (2013). *Blackwood Division GIS Data*.

Ferwerda, Jelle G, & Skidmore, Andrew K. (2007). Can nutrient status of four woody plant species be predicted using field spectrometry? *ISPRS Journal of Photogrammetry and Remote Sensing*, *62*(6), 406-414.

Forschungszentrum Jülich. (2015). HyPlant. Retrieved Feb. 24, 2015, from http://www.fz-juelich.de/ibg/ibg-2/EN/methods_jppc/HyPlant/_node.html

Geladi, Paul, & Kowalski, Bruce R. (1986). Partial least-squares regression: a tutorial. *Analytica chimica acta*, *185*, 1-17.

Gough, CM, Seiler, JR, & Maier, Chris A. (2004). Short-term effects of fertilization on loblolly pine (Pinus taeda L.) physiology. *Plant, Cell & Environment*, *27*(7), 876-886.

Hatfield, Jerry L, & Prueger, John H. (2010). Value of using different vegetative indices to quantify agricultural crop characteristics at different growth stages under varying management practices. *Remote Sensing*, *2*(2), 562-578.

Hollinger, David Y, Ollinger, SV, Richardson, AD, Meyers, TP, Dail, DB, Martin, ME, . . . Clark, KL. (2010). Albedo estimates for land surface models and support for a new paradigm based on foliage nitrogen concentration. *Global Change Biology*, *16*(2), 696-710.

Im, Jungho, Jensen, John R, Coleman, Mark, & Nelson, Eric. (2009). Hyperspectral remote sensing analysis of short rotation woody crops grown with controlled nutrient and irrigation treatments. *Geocarto International*, *24*(4), 293-312.

Kiser, L Chris, Fox, Thomas R, & Carlson, Colleen A. (2013). Foliage and litter chemistry, decomposition, and nutrient release in Pinus taeda. *Forests, 4*(3), 595-612.

Knyazikhin, Yuri, Schull, Mitchell A, Stenberg, Pauline, Mõttus, Matti, Rautiainen, Miina, Yang, Yan, . . . Lewis, Philip. (2013). Hyperspectral remote sensing of foliar nitrogen content. *Proceedings of the National Academy of Sciences, 110*(3), E185-E192.

Latorre-Carmona, Pedro, Knyazikhin, Yuri, Alonso, Luis, Moreno, Jose F, Pla, Filiberto, & Yan, Yang. (2014). On hyperspectral remote sensing of leaf biophysical constituents: decoupling vegetation structure and leaf optics using CHRIS-PROBA data over crops in Barrax. *Geoscience and Remote Sensing Letters, IEEE, 11*(9), 1579-1583.

Lewis, P, & Disney, M. (2007). Spectral invariants and scattering across multiple scales from within-leaf to canopy. *Remote Sensing of Environment, 109*(2), 196-206.

Luo, Yiqi, Medlyn, Belinda, Hui, Dafeng, Ellsworth, David, Reynolds, James, & Katul, Gabriel. (2001). Gross primary productivity in Duke Forest: modeling synthesis of CO2 experiment and eddy-flux data. *Ecological Applications, 11*(1), 239-252.

Martin, Mary E, & Aber, John D. (1997). High spectral resolution remote sensing of forest canopy lignin, nitrogen, and ecosystem processes. *Ecological applications, 7*(2), 431-443.

Martin, ME, Plourde, LC, Ollinger, SV, Smith, M-L, & McNeil, BE. (2008). A generalizable method for remote sensing of canopy nitrogen across a wide range of forest ecosystems. *Remote Sensing of Environment, 112*(9), 3511-3519.

Moran, Jonathan A, Mitchell, Alan K, Goodmanson, Graeme, & Stockburger, Keri A. (2000). Differentiation among effects of nitrogen fertilization treatments on conifer seedlings by foliar reflectance: a comparison of methods. *Tree physiology, 20*(16), 1113-1120.

Ollinger, Scott V. (2011). Sources of variability in canopy reflectance and the convergent properties of plants. *New Phytologist, 189*(2), 375-394.

Ollinger, Scott V, Reich, Peter B, Frolking, Steve, Lepine, Lucie C, Hollinger, David Y, & Richardson, Andrew D. (2013). Nitrogen cycling, forest canopy reflectance, and emergent properties of ecosystems. *Proceedings of the National Academy of Sciences, 110*(27), E2437-E2437.

Ollinger, SV, Richardson, AD, Martin, ME, Hollinger, DY, Frolking, SE, Reich, PB, . . . Oren, R. (2008). Canopy nitrogen, carbon assimilation, and albedo in temperate and boreal forests: Functional relations and potential climate feedbacks. *Proceedings of the National Academy of Sciences, 105*(49), 19336-19341.

Ollinger, Scott V, Smith, Marie-Louise. (2005). Net primary production and canopy nitrogen in a temperate forest landscape: an analysis using imaging spectroscopy, modeling and field data. *Ecosystems, 8*(7), 760-778.

Panferov, Oleg, Knyazikhin, Yuri, Myneni, Ranga B, Szarzynski, Jörg, Engwald, Stefan, Schnitzler, Karl G, & Gravenhorst, Gode. (2001). The role of canopy structure in the spectral variation of transmission and absorption of solar radiation in vegetation canopies. *Geoscience and Remote Sensing, IEEE Transactions on, 39*(2), 241-253.

Peltoniemi, Mikko, Pulkkinen, Minna, Kolari, Pasi, Duursma, Remko A, Montagnani, Leonardo, Wharton, Sonia, . . . Christensen, Torben. (2012). Does canopy mean nitrogen concentration explain variation in canopy light use efficiency across 14 contrasting forest sites? *Tree physiology, 32*(2), 200-218.

Petisco, C, Garcia-Criado, B, de Aldana, BR Vazquez, Zabalgogeazcoa, I, & Mediavilla, S. (2005). Use of near-infrared reflectance spectroscopy in predicting nitrogen, phosphorus

and calcium contents in heterogeneous woody plant species. *Analytical and bioanalytical chemistry, 382*(2), 458-465.

Pimstein, Agustin, Karnieli, Arnon, & Bonfil, David J. (2007). Wheat and maize monitoring based on ground spectral measurements and multivariate data analysis. *Journal of Applied Remote Sensing, 1*(1), 013530-013530-013516.

Pope, Graham, & Treitz, Paul. (2013). Leaf Area Index (LAI) Estimation in Boreal Mixedwood Forest of Ontario, Canada Using Light Detection and Ranging (LiDAR) and WorldView-2 Imagery. *Remote Sensing, 5*(10), 5040-5063.

Rossini, M, Nedbal, L, Guanter, L, Ač, A, Alonso, L, Burkart, A, . . . Drusch, M. (2015). Red and far-red sun-induced chlorophyll fluorescence as a measure of plant photosynthesis. *Geophysical Research Letters*.

Serrano, L., Penuelas, J., & Ustin, S. (2002). Remote sensing of nitrogen and lignin in Mediterranean vegetation from AVIRIS data: Decomposing biochemical from structural signals. Remote Sensing of Environment, *81*(2), 355–364.

Smith, M-L, Martin, Mary E, Plourde, Lucie, & Ollinger, Scott V. (2003). Analysis of hyperspectral data for estimation of temperate forest canopy nitrogen concentration: comparison between an airborne (AVIRIS) and a spaceborne (Hyperion) sensor. *Geoscience and Remote Sensing, IEEE Transactions on, 41*(6), 1332-1337.

Stein, Beth R, Thomas, Valerie A, Lorentz, Laura J, & Strahm, Brian D. (2014). Predicting macronutrient concentrations from loblolly pine leaf reflectance across local and regional scales. *GIScience & Remote Sensing, 51*(3), 1-19.

Stenberg, Pauline, Mottus, Matti, & Rautiainen, Miina. (2008). Modeling the spectral signature of forests: application of remote sensing models to coniferous canopies. *Advances in Land Remote Sensing*. Ed. S. Liang. Netherlands: Springer. 147-171.

Thenkabail, Prasad S, Gumma, Murali Krishna, Teluguntla, Pardhasaradhi, & Mohammed, Irshad A. (2014). Hyperspectral remote sensing of vegetation and agricultural crops. *Photogrammetric Engineering and Remote Sensing, 80*(8), 697-709.

Thulin, Susanne, Hill, Michael J, Held, Alex, Jones, Simon, & Woodgate, Peter. (2012). Hyperspectral determination of feed quality constituents in temperate pastures: Effect of processing methods on predictive relationships from partial least squares regression. *International Journal of Applied Earth Observation and Geoinformation, 19*, 322-334.

Townsend, Philip A, Foster, Jane R, Chastain Jr, Robert A, & Currie, William S. (2003). Application of imaging spectroscopy to mapping canopy nitrogen in the forests of the central Appalachian Mountains using Hyperion and AVIRIS. *Geoscience and Remote Sensing, IEEE Transactions on, 41*(6), 1347-1354.

Townsend, Philip A, Serbin, Shawn P, Kruger, Eric L, & Gamon, John A. (2013). Disentangling the contribution of biological and physical properties of leaves and canopies in imaging spectroscopy data. *Proceedings of the National Academy of Sciences, 110*(12), E1074.

Vose, James M, & Allen, H Lee. (1988). Leaf area, stemwood growth, and nutrition relationships in loblolly pine. *Forest Science, 34*(3), 547-563.

Wessman, Carol A, Aber, John D, & Peterson, David L. (1989). An evaluation of imaging spectrometry for estimating forest canopy chemistry. *International Journal of Remote Sensing, 10*(8), 1293-1316.

Wicklein, Haley F, Ollinger, Scott V, Martin, Mary E, Hollinger, David Y, Lepine, Lucie C, Day, Michelle C, . . . Norby, Richard J. (2012). Variation in foliar nitrogen and albedo in response to nitrogen fertilization and elevated CO2. *Oecologia*, *169*(4), 915-925.

Will, Rodney E, Munger, Gregory T, Zhang, Yujia, & Borders, Bruce E. (2002). Effects of annual fertilization and complete competition control on current annual increment, foliar development, and growth efficiency of different aged Pinus taeda stands. *Canadian journal of forest research*, *32*(10), 1728-1740.

Xiao, Xiangming, Zhang, Qingyuan, Hollinger, David, Aber, John, & Moore III, Berrien. (2005). Modeling gross primary production of an evergreen needleleaf forest using MODIS and climate data. *Ecological Applications*, *15*(3), 954-969.

Yáñez Rausell, L. (2014). *Coniferous needle-leaves, shots and canopies: a remote sensing approach*. Doctoral dissertation, Wageningen University.

Zhang, Shangshan, & Allen, H Lee. (1996). Foliar nutrient dynamics of 11-year-old loblolly pine (Pinus taeda) following nitrogen fertilization. *Canadian Journal of Forest Research*, *26*(8), 1426-1439.

4 Chapter 4—Subcanopy Invasive Grass Detection with Hyperion

Abstract

Wavyleaf basketgrass (*Oplismenus hirtellus* ssp. *undulatifolius*, wavyleaf basketgrass) is a non-native perennial that has recently begun invading the deciduous forests of Maryland and Virginia. Despite its discovery in 1995, relatively little is known regarding wavyleaf basketgrass invasion geography. This study aims to investigate the use of several imaging spectroscopy algorithms in mapping wavyleaf basketgrass distribution. Our goals are two-fold: (1) to produce maps of the extent and coverage of wavyleaf basketgrass in the study area, and (2) to determine the effect of percent coverage on mapping accuracy of this understory species. The study area consists of two sites in Patapsco Valley State Park, MD. At each site, we mapped presence and absence plots with various percent coverages of wavyleaf basketgrass and collected coverage estimates at 100-150 validation points. We acquired Hyperion images from various autumn months in 2013 and 2014. Using multiple imaging spectroscopy algorithms, we classified the images as presence or absence and conducted point- and area-based accuracy assessments. We calculated error matrices, overall accuracy, and the kappa coefficients for each classified image. We produced several maps of the wavyleaf basketgrass distribution with moderate accuracy for management purposes (overall accuracy=0.41-0.78; kappa<0.53); however, the relationship between percent coverage and seasonality remains unclear. Greater knowledge of the target species would enhance our ability to predict and prevent further spread of this invader.

Keywords: Invasive plants, imaging spectroscopy, Hyperion, spectral unmixing, classification accuracy

Introduction

Invasive plants pose a threat to ecosystems through their effect on wildlife food sources, habitat, soil nutrient pools, and biodiversity (Dehaan et al. 2007; Lockwood et al. 2013). Detecting invasive plants early before they increase their coverage and density increases the probability of successful removal and keeps costs low. Traditional methods for surveying weeds are labor-intensive, expensive, and often limited in their spatial and temporal coverage (Dehaan et al. 2007). Thus, remote sensing increases the likelihood of detection in a timely and cost-effective manner (Lass et al. 2005). With its ability to discern differences in reflectance at narrow bandwidths, imaging spectroscopy imagery has a unique ability to detect sub-pixel plant infestations and discriminate between different species, which is critical for invasive plant detection (Brown & Noble 2005; He et al. 2011; Lawrence, Wood, & Sheley 2006). Multispectral imagery does not offer such success with species discrimination in diverse, highly vegetated areas, without multitemporal images or unique phenology (e.g. Resasco et al. 2007, Peterson et al. 2007). As a result, many researchers have investigated the use of imaging spectroscopy data for species discrimination of multiple invaders (Glenn et al. 2005; Lass et al. 2005; Miao et al. 2006; Mundt, Streutker, & Glenn 2007; Underwood, Ustin, & DiPietro 2003; Williams & Hunt 2002).

Various processing methods have been used to map weed infestations with imaging spectroscopy imagery, and methods vary in success based on plant characteristics (Dehaan et al. 2007). Supervised classification methods can be divided into two types: distance-based or unmixing-based. Distance classifiers classify unknown pixels based on their distance from known reference values. These algorithms include spectral angle mapper (SAM), maximum likelihood, and minimum distance. They assume a pure pixel with complete plant coverage (Lass

et al. 2005). SAM calculates the angle formed between the vectors of the target reflectance and pixel value and using a maximum angle threshold to determine the classification. The more narrow the angle, the more confidence the algorithm has in the classification. Greater angles classify a larger percent of the image, but the confidence of the classifications is lower. This may result in high rates of commission error (Lass et al. 2005). Lass et al. (2005) used the SAM algorithm at various angles to identify 57 percent of knapweed and 97 percent of babysbreath infestations with 2m imaging spectroscopy imagery. These infestations occurred within relatively simple land cover conditions for image classification: rangelands and pastures. Underwood et al. (2003) used AVIRIS imaging spectroscopy imagery to successfully map iceplant and jubata grass with users' accuracies of 94 and 89 percent (Dehaan et al. 2007).

Since image reflectance is the amalgamation of radiation from all materials within the sensor's field of view, images often contain many mixed pixels (Quintano et al. 2012). Unmixing classifiers are a set of algorithms designed to separate the reflectance spectrum into the proportion estimates of known components. The most common method is linear spectral unmixing, which applies a least squares approach to linearly unmix endmembers from the entire image (Quintano et al. 2012). These classifiers work best when the user has endmembers for all existing vegetation (Dópido & Plaza 2011; Lass et al. 2005). However, since they map several endmembers per pixel, they may not always give accurate estimates of pixel composition (Dehaan et al. 2007). When not all endmembers are known, partial unmixing methods such as mixture-tuned matched filtering (MTMF) can quantify abundance fractions per pixel (Dópido & Plaza 2011). MTMF matches pixel values to the target spectrum while suppressing the reflectance noise through the use of an infeasibility estimate (Mundt et al. 2007). Consequently, the MTMF algorithm quantifies sub-pixel abundance of an individual endmember with scores of

0 (absent) to 1 (complete coverage) (Dópido & Plaza 2011). There are multiple examples of successful vegetation mapping with both types of algorithms (Miao et al. 2006; Mundt et al. 2007; Williams & Hunt 2002).

In the case of invasive plants, the assumption of a pure endmember is often an issue of concern with these techniques. A minimum plant density is generally required for detection and classification by remote sensing techniques (Brown & Noble 2005). Brown & Noble (2005) report that the necessary "threshold of detection" depends on sensor resolution, plant species, size, growth stage, and soil/vegetation background. Consequently, most studies set a cover threshold for both training and ground truth points (e.g. de Castro et al. 2012, Lamb & Weedon 1998). Lawrence, Wood and Sheley (2006) selected plots with a minimum of 5 percent infestation. Other studies have examined the classification accuracies of imaging spectroscopy image analysis using various coverage delineations. Ustin et al. (2002) used SAM to classify iceplant and jubata presence in California's sage scrub ecosystems into multiple density categories. The geographic trends were generally correct; however, iceplant densities were sometimes misclassified. Kloppenburg (2014) compared leafy spurge detection accuracy in grasslands across three density classes (high, medium, and low stems/m^2). She found that the ability to map leafy spurge greatly improved with the increase in plant density, as few low density patches were successfully detected. Similarly, Williams and Hunt (2004) detected leafy spurge invasions of at least 10 percent canopy cover with relatively high accuracy using MTMF (Mundt et al. 2007; Williams & Hunt 2002).

Forested environments represent a particular challenge for endmember selection and detection of understory plants. The classification accuracy of understory plants is often affected by mixed pixels and low visibility by overhead sensors. Even when the target species has nearly

complete ground coverage, the overstory crown closure will obscure the understory. Likewise, the pixels to be classified often have similar conditions. Due to these difficulties, few studies have attempted to map plants that grow in the diverse and complex understory of the temperate deciduous forest. Thus, the extent to which plant ground coverage affects our ability to map a forest understory plant remains unknown. We would like to investigate this problem in the context of mapping an invasive plant in the temperate deciduous understory.

Wavyleaf basketgrass (*Oplismenus hirtellus* ssp. *undulatifolius*) is a relatively new invader in the Mid-Atlantic region of the United States. Since the invasion's discovery in 1995, the plant has spread into multiple counties within Maryland and Virginia (EDDMapS 2013). However, little is known about the species ecology, invasion history, or its level of threat (Beauchamp et al. 2013). For the purposes of this study, wavyleaf basketgrass has several traits that make it a good candidate to assess the remote detection of understory grasses. First, wavyleaf basketgrass grows in dense, monospecific patches that can span multiple acres. Second, plants have a unique appearance and phenology that may aid in their detection. The wavy, hairy texture of their leaves offers a distinct contrast to Japanese stiltgrass. In addition, wavyleaf basketgrass seeds and senesces later than the adjacent vegetation. Consequently, there is reason to assume that mature plants in large patches could be detected with remote sensing.

The specific objectives of this work are to produce maps of the wavyleaf basketgrass invasion in the study area, and to evaluate the mapping accuracy of wavyleaf basketgrass using imaging spectroscopy methods. This second objective focuses on two factors that affect pixel composition: wavyleaf basketgrass ground coverage and overstory seasonality. With respect to these characteristics, the study addresses two hypotheses: (1) A minimum level of ground coverage is a necessary precondition for successfully mapping wavyleaf basketgrass, with

greater levels of coverage producing marginal increases in accuracy up to a certain threshold; and (2) Images taken after overstory leaves have begun falling but before wavyleaf basketgrass senescence will have better classification accuracies than those taken earlier in the season. The latter hypothesis is based on the assumption that the positive effect of overstory leaf fall on classification accuracy is greater than the obscuring effect of the litter on the understory plants.

To evaluate these hypotheses, we assess the accuracy of wavyleaf basketgrass classification maps across a range of ground coverage rates and image acquisition dates using field measurements. Through this study, we hope to improve our understanding of remote sensing methods for mapping invasive understory grasses in mixed deciduous forests. Our long-term research goal is to provide land managers with better knowledge of the current wavyleaf basketgrass distribution and the utility of imaging spectroscopy applications in future detection.

Methods

Study Area

The study area consists of two sections of Patapsco Valley State Park, McKeldin Recreation Area and Southeast Patapsco. McKeldin Recreation Area (39.36 N, 76.88 W) is located approximately 30 km northwest of Baltimore, MD, US (Figure 4.1). Southeast Patapsco is several kilometers to the southeast of McKeldin. The first documented occurrences of wavyleaf basketgrass in the mid-Atlantic region of the US occurred at McKeldin Recreation Area in 1995. Since then, the invasion has spread to cover acres of the forest (Beauchamp et al. 2013). The Southeast Patapsco area contains smaller patches and plants in lower densities than McKeldin (EDDMapS 2013).

Patapsco Valley State Park is located along the Patapsco River, in the Piedmont Uplands ecoregion of Maryland (Woods, Omernik, & Brown 1999). Originally established as the Patapsco Forest Reserve in 1920, today the park covers an area of over 16 thousand acres (Maryland Department of Natural Resources 2014). The terrain consists of rolling hills and streams characteristics of the Piedmont Physiographic Province, with an average elevation of 96 m (Maryland Department of Natural Resources,1998; Nowak & Heisler 2010). Soils in the region are typically "deep, well-drained" loams. Ultisols and inceptisols are common, including soils from the Manor, Glenelg, and Chester series (Beauchamp et al. 2013). The average annual temperature ranges from 1.7 °C to 22.9 °C, with average precipitation of 104.4 cm (Beauchamp et al. 2013). Forests dominate the landscape, with tree canopy cover of 68 percent (Woods et al. 1999). Within Patapsco Valley State Park, there are five general vegetation classes: upland mixed hardwoods (e.g. tulip poplar, white oak, scarlet oak), bottomland mixed hardwoods (e.g. sycamore, tulip poplar), marsh, cultivated fields, and pine plantations (Maryland Department of Natural Resources 1998). The park lies within the Appalachian Oak Forest ecoregion, which is dominated by white and red oaks (Woods et al. 1999). Common recreational activities include hiking, horseback riding, and mountain biking along the 270 km of trails, and disc golf (Maryland Department of Natural Resources 2014).

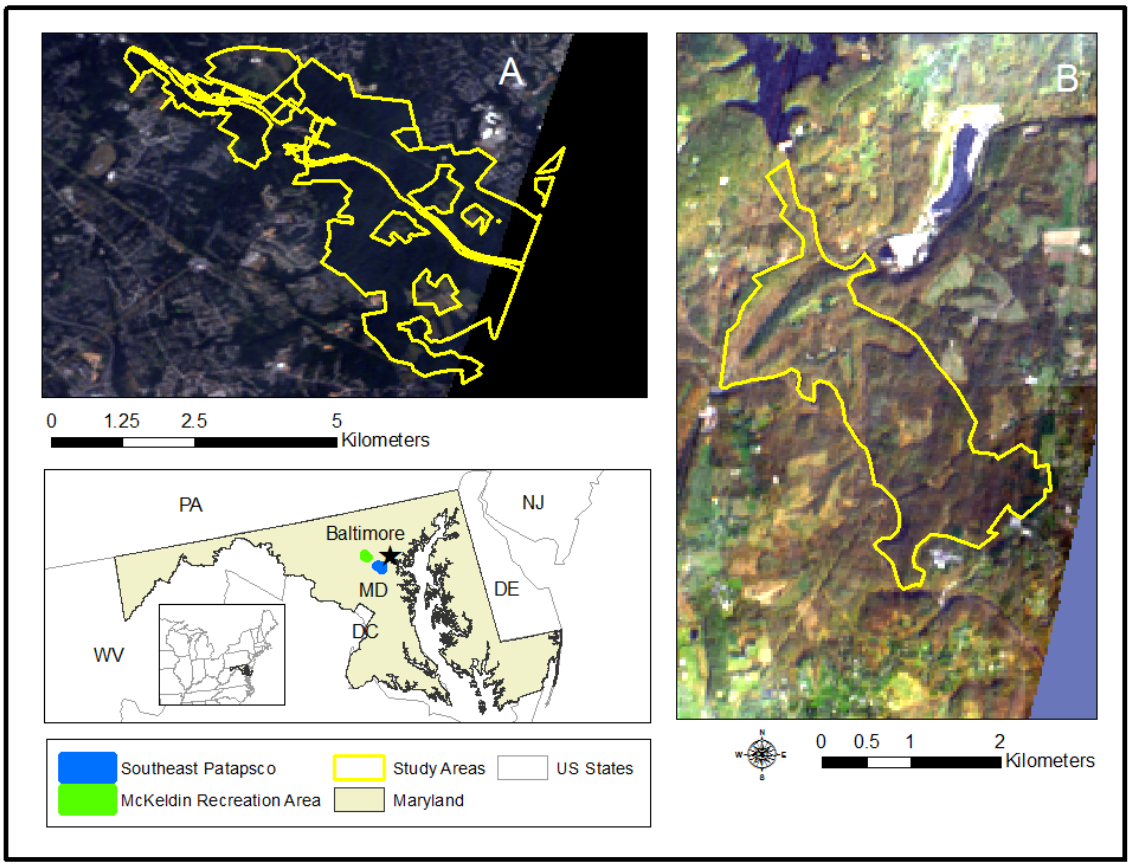


Figure 4.1. Maps of (A) Southeast Patapsco and (B) McKeldin Recreation Area study area in Patapsco Valley State Park, MD. Images overlaid on preprocessed Hyperion imagery from August 2013, and October 2014, respectively. The study area locations near Baltimore, Maryland are shown in the map on the left.

Field Sampling

Fieldwork was conducted in the summer and fall of 2013 and 2014. Due to evident site differences, a separate methodology was used for field sampling at each site. At the McKeldin study area, we selected 28 large polygonal areas to represent complete coverage and absence plots of wavyleaf basketgrass within the forested areas. Plots were located in known areas of infestation or near trails. Visual inspection of the entire plot area confirmed complete coverage of wavyleaf basketgrass on the 'presence' plots. In the presence plots, wavyleaf basketgrass covered the entire ground in very high densities. With the exception of trees, small patches of other native and non-native grasses, and some bushes, wavyleaf basketgrass excluded all other vegetation. We delineated 13 presence plots and 15 absence plots across a large geographic area using a handheld GPS. Absence plots consisted predominately of soil and leaves, or other low-lying vegetation (e.g. Japanese stiltgrass, *Microstegium vimineum*). For each absence plot, we also estimated the percent coverage for other vegetation types, with a particular focus on Japanese stiltgrass, another common invader in the region.

At Southeast Patapsco, we delineated 4 absence plots and 14 presence plots within the forested areas. We defined absence plots as containing few or no small patches of wavyleaf basketgrass, covering an area of less than 1 percent of the polygon. We selected presence plots to include a range of wavyleaf basketgrass percent coverages, as well as many locations with relatively high abundances of wavyleaf basketgrass. We used a minimal plot size of 60 m x 60 m in order to fully encompass at least one pixel, even if plots were located along pixel edges or there was high georegistration error. Plots ranged in area from 1,066 to 57,744 m^2, with a mean area of 11,585 m^2 and standard deviation of 11,476 m^2. At each presence plot, we estimated

wavyleaf basketgrass density within randomly located 2m x 2m quadrats. The number of quadrats increased with plot size, ranging from 9 to 12 quadrats.

Following data collection, we examined the GPS data for topological errors and agreement between several handheld units. When possible, we also evaluated the ground accuracy on the computer and in comparison to reference Landsat data. Fewer than half of all data points were accurate within 5m.

Image Preprocessing

We acquired several cloud-free Hyperion images over the two study areas in the fall of 2013 and 2014. The Hyperion satellite is a push-broom sensor that consists of 220 bands with a spectral range from 400-2500 nm. Spectral and spatial resolutions are 10 nm and 30 m, respectively. We obtained images of McKeldin Recreation Area on November 20, 2013, and October 24, 2014; and of Southeast Patapsco images on August 17, 2013, and September 15, 2013. At the time of image acquisition, plants were fully grown and seeding. The October and November images were taken under partially leaf-off conditions in order to increase the satellite's view of the forest understory; however, the fallen leaves partly obscured the understory.

Images were acquired with geometric correction and systematic terrain correction (Level 1T). However, Hyperion is experiencing serious technical problems that affect image quality. Consequently, major preprocessing steps were required prior to analysis. We used Hyperion Tools for image destriping, bad band removal, and to calculate scale factors for Fast Line-of-sight Hypercubes (FLAASH) (Datt et al. 2003). We processed the images in FLAASH for atmospheric correction, using the US standard atmospheric model and urban aerosol model, due

to the study area's location in the suburbs of Baltimore, Maryland. Due to the continued presence of striping, we conducted a forward Principal Components Analysis (PCA) on the image (Datt et al. 2003). PCA is used to reduce data dimensionality, while retaining maximum information. Since the first three principal components contained the greatest amount of information, only those components were used in the inverse PCA. Next, we conducted a spatial subset on the images to reduce the geographic coverage.

Image Analysis

Images were processed using Minimum Noise Fraction Transform (MNF) to further improve the noise-to-signal quality of the data (Boardman & Kruse 1994). To determine the presence and coverage of wavyleaf basketgrass within each pixel, we used various imaging spectroscopy classification techniques: maximum likelihood supervised classification, mixture-tuned matched filtering (MTMF), linear unmixing, and the Spectral Angle Mapper (SAM). We selected presence and absence endmembers from pixels in field-mapped plots with particularly high or low amounts of wavyleaf basketgrass. To compensate for any geographic registration errors, we located endmembers within the plots' interior and visually assessed the spectra to verify sufficient differences. Endmembers were derived from the image rather than the field-measured spectra because the former are more consistently illuminated and better match the scale and atmospheric conditions of the image (Dehaan et al. 2007; Plaza et al. 2004). For the SAM analysis, several maximum angles were evaluated prior to selecting the optimal threshold for each image. The optimum threshold was defined as that which classified nearly all of the pixels within the study area and also maintained a reasonable level of accuracy. Figure 4.2 shows the methods for the image processing and accuracy assessment.

Accuracy Assessment

We conducted separate accuracy assessments for each study area using two validation approaches. First, we generated 100-150 random points within the forested portion of the study area. We determined forest cover based on a visual interpretation of aerial imagery. Then, we visited each point and walked in a 15m-diameter circle in order to estimate the percentage of wavyleaf basketgrass coverage. Due to the ambiguous definitions of presence and absence, we classified the coverage amounts into several different categories. At McKeldin Recreation Area, we defined presence as at least 25, 50, or 75 percent coverage. Due to the lower density of the invasion at Southeast Patapsco, we used classes of 0, 1 and 5 percent coverage at the validation points.

The second validation approach used the previously delineated presence and absence plots to conduct an area-based accuracy assessment. We calculated zonal statistics and used the majority raster value to determine whether the algorithm characterized the plot as presence or absence. This area-based accuracy estimate is described in detail by Stehman and Wickham (2011). At McKeldin, presence plots represented nearly complete (100 percent) ground coverage by wavyleaf basketgrass, excluding tree basal area. At Southeast Patapsco, we used plots with various wavyleaf basketgrass coverage amounts, ranging from 6 to 48 percent. We evaluated the accuracy of the plot classifications for wavyleaf basketgrass coverage in increments of 10 percent.

Error matrices were used to calculate commission and omission error, as well as overall accuracy (Card 1982, Congalton 1991). We also calculated the kappa coefficient to determine how classification accuracy compared to a random classification (Dehaan et al. 2007).

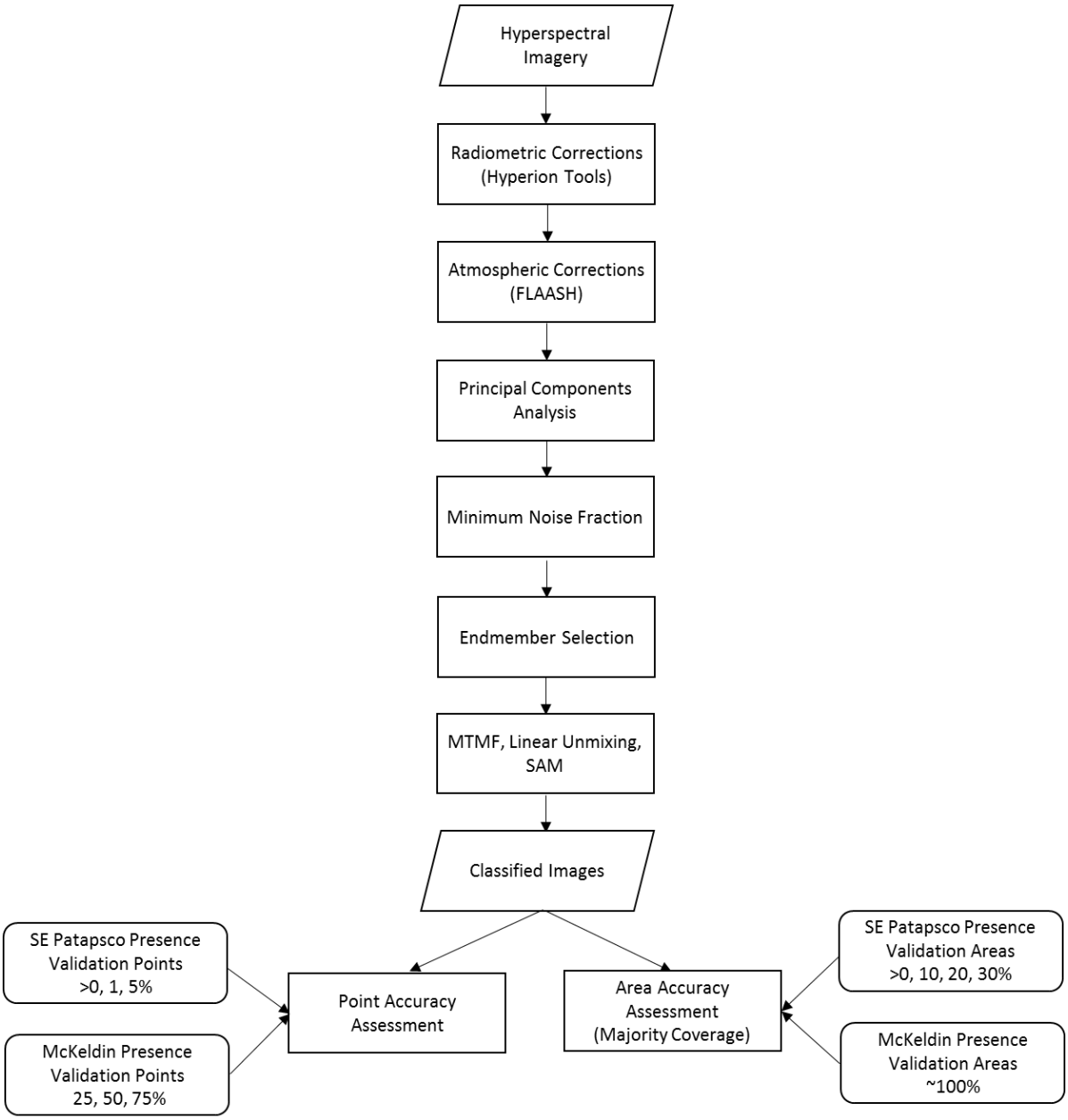


Figure 4.2. Procedural flowchart depicting the image processing, classification, and accuracy assessment methods taken to conduct this analysis. Percentages listed in the accuracy assessment boxes denote classification threshold categories for wavyleaf basketgrass coverage within the pixel or area.

Results

Wavyleaf Basketgrass Reflectance Spectrum

The wavyleaf basketgrass reflectance spectrum for the October 24, 2014 image of McKeldin Recreation Area can be seen in Figure 4.3, along with spectra of several selected land cover classes. The gaps in all spectra between 1336 and 1488 nm and 1800 to 1972 nm are due to removal of strong water vapor absorption bands. The presence spectrum has greater reflectance in the orange, red, and near infrared plateau wavelengths (e.g. 690 to 1300 nm) than the absence spectrum. The absence spectrum, obtained from a forested area with no wavyleaf basketgrass present, has deeper water absorption features around 1000 and 1200 nm. Neither the presence nor absence spectrum shows much reflectance in the blue or green wavelengths.

The water spectrum of the selected endmember has higher reflectance than the typical water spectrum, with greater peaks and valleys throughout the visible and near-infrared (NIR) wavelengths. Water reflectance typically peaks in the blue and declines in the near infrared (NIR) wavelengths. This difference could be due to the shallow, turbid nature of the water, and the narrow size of the river in the area. With the narrow river and the canopy cover along the floodplains and going into the water, the water features are typically smaller than a pixel and are mixed. Likewise, urban spectra are generally mixed and can represent a number of different features. The urban spectra clearly lack the absorption features seen in the vegetative spectra from the leaf pigments, structure, and moisture content. The increasing reflectance values as the spectra goes from absence to presence to fields may be indicative of increasing leaf area. Since this image was taken under partly leaf-off conditions, reflectance could be primarily attributed to the greenness and coverage of the vegetation. Consequently, the non-vegetated land cover classes have lower reflectance in the near infrared portion of the spectrum.

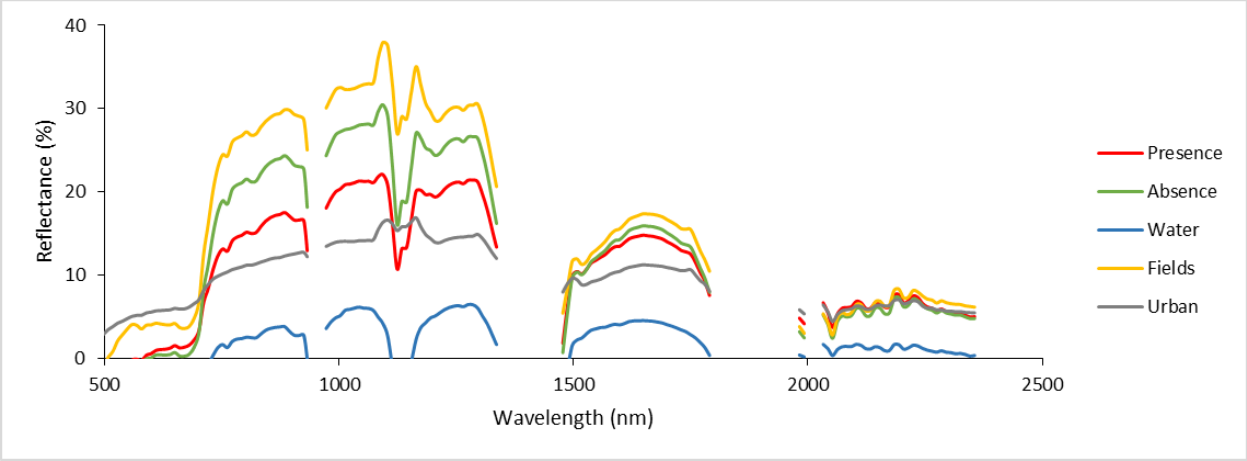


Figure 4.3. Endmember spectral plots at McKeldin Recreation Area across Hyperion bands for selected land cover classes.

Image Classification Results

Out of the imaging spectroscopy algorithms we evaluated, SAM and linear unmixing produced the best results. For this reason, we focus the remainder of this section on the SAM and unmixing results; when SAM and linear spectral unmixing classifications were similar, we exclusively present SAM.

A visual assessment of the SAM land cover classification based on the reference Landsat image indicated a high degree of accuracy. (Presence and absence could not be evaluated from the Landsat image, but the other land cover classes could be evaluated.) Although the algorithm was unable to classify large waterbodies from endmembers selected within the study area, portions of the Patapsco River were correctly classified. The forest land cover class, represented by the combination of presence and absence plots, is correctly differentiated from fields throughout the image. The classification map with all land cover classes, and the corresponding Landsat image, can be seen in Figure 4.4. The successful differentiation of broad land cover classes using a imaging spectroscopy algorithm gives us confidence in our decision to evaluate

their use in classifying more specific land cover classes, i.e. wavyleaf basketgrass presence and absence. Consequently, the SAM and linear spectral unmixing classifications for McKeldin Recreation Area are shown in Figure 4.5. A comparison of the classified maps in Figures 4.4 and 4.5 with the Landsat reference image in Figure 4.4 shows the correct classification of water, urban, and fields as "absence" when the other land cover classes are not specified as options. Those pixels classified as "presence" remain the same across both figures. The Patapsco River, dam, and some fields remain unclassified.

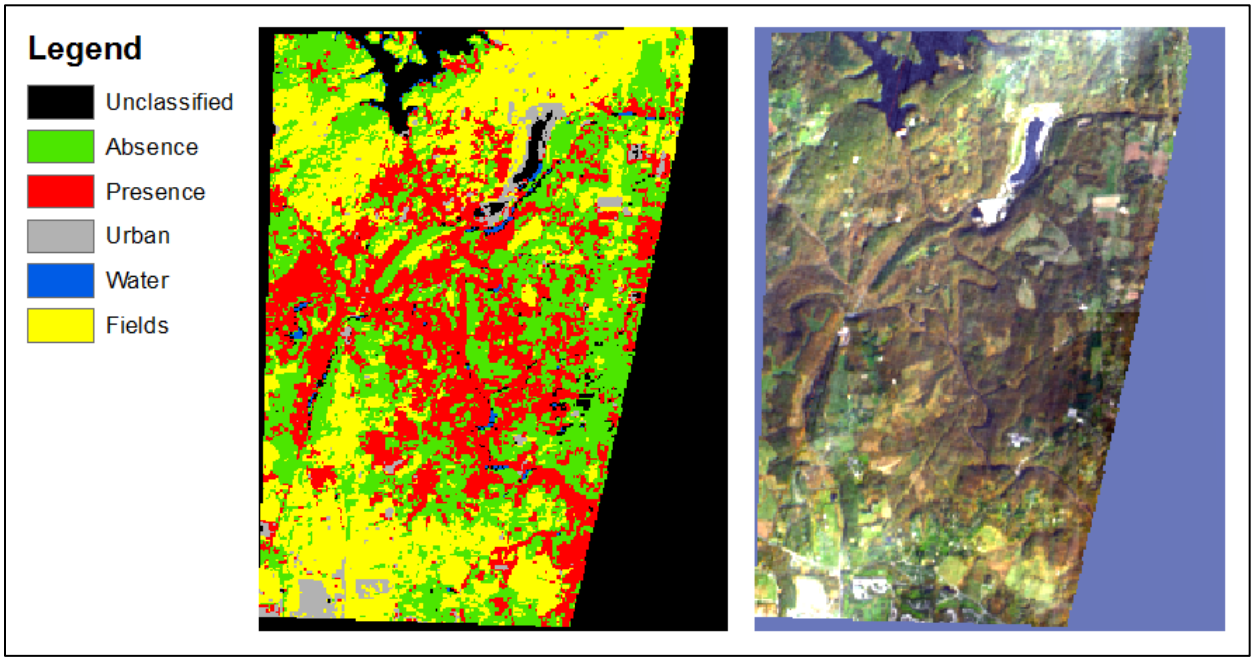


Figure 4.4. SAM classification of wavyleaf basketgrass presence and absence along with several other land cover classes in McKeldin Recreation Area, with a maximum angle of 0.20. A Landsat image of the study area is shown as a reference.

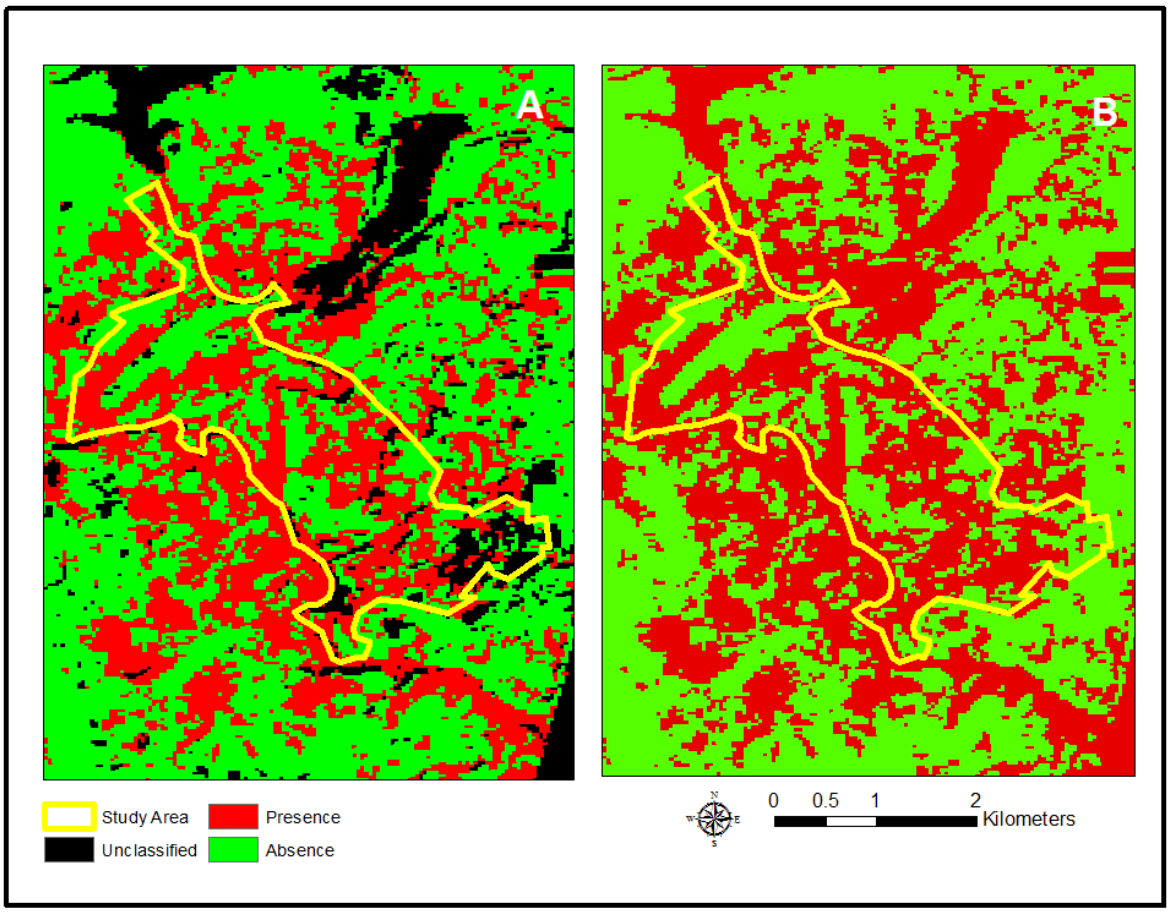


Figure 4.5. Classification of wavyleaf basketgrass presence and absence in McKeldin Recreation Area with (A) SAM and (B) linear spectral unmixing. The SAM algorithm used a maximum angle of 0.20.

Wavyleaf basketgrass presence in Figure 4.5 shows agreement with the general areas and suitable habitat features where wavyleaf basketgrass is known to occur. For instance, most of the infestation occurs along the Patapsco River and within the floodplain, as depicted in the image. Similarly, some of the areas classified as "absence" have steep slopes, which do not typically show signs of wavyleaf basketgrass invasion. Classification maps of the Southeast Patapsco study area also depict realistic patterns of wavyleaf basketgrass infestation as the prior image. Figure 4.6 shows the SAM and unmixing results using the image from August 2013. A

maximum angle of 0.10 was selected as producing the best results and optimal balance between number of classified pixels and accuracy. The image shows high levels of presence in the eastern half of the study area, which is where the wavyleaf basketgrass tends to be most abundant. Sections of the Patapsco River remain unclassified, which is realistic given the difference in reflectance between land cover classes.

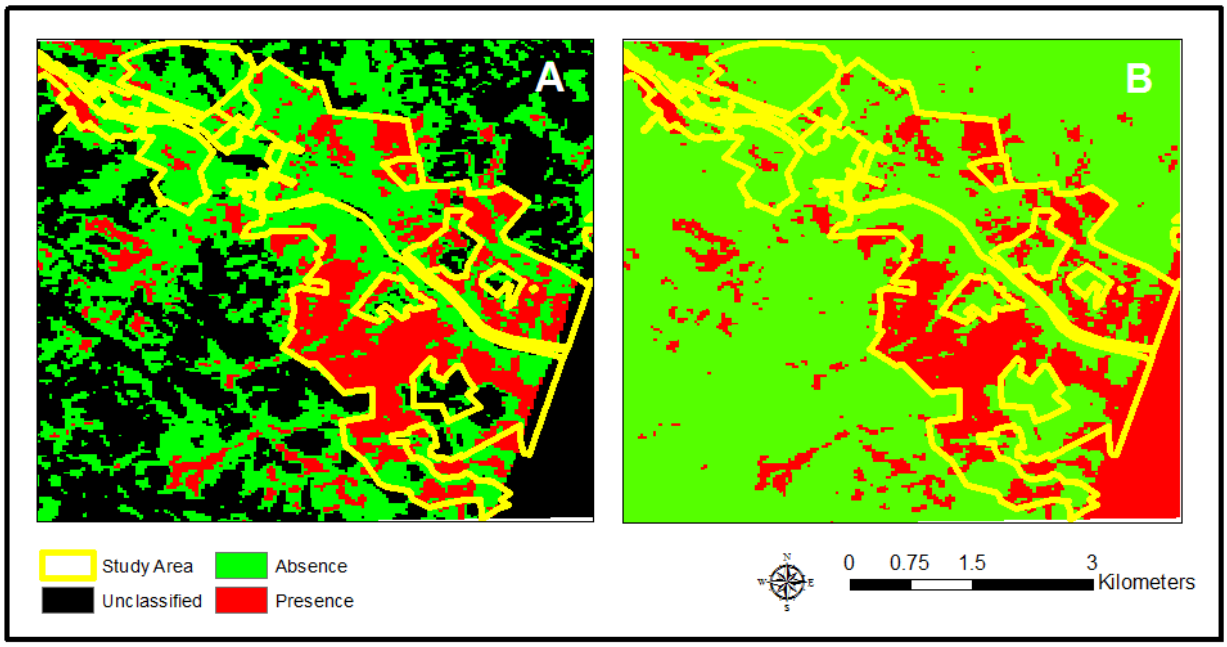


Figure 4.6. (A) SAM and (B) linear spectral unmixing classifications of wavyleaf basketgrass presence and absence in Southeast Patapsco.

Tables 4.1 and 4.2 show the effect of validation percent plant coverage on classification accuracy at the two sites. As the percent of wavyleaf basketgrass at the validation points increased, omission error and overall accuracy improved, although commission error also increased. This signifies that fewer presence points were omitted, but more points with less than the threshold amount of wavyleaf basketgrass were classified as present. A similar pattern was also seen with the validation points at Southeast Patapsco, and reinforced by the area-based

accuracy assessments at both sites. At McKeldin, the validation plots in the accuracy assessment represent areas with nearly complete wavyleaf basketgrass coverage. Thus, the higher accuracy obtained follows the trend of the validation points and indicates better mapping accuracy with higher percent coverage. At Southeast Patapsco, overall classification accuracy was also high, despite significantly lower densities of wavyleaf basketgrass. However, overall accuracy decreased as wavyleaf basketgrass coverage increased within the presence validation areas, due to high rates of commission error.

Table 4.1. Percent coverage in validation points and area greatly affect the error and accuracy of the SAM results of McKeldin Recreation Area from the October 2014 image. The presence areas consist of nearly complete wavyleaf basketgrass coverage across the forest floor.

Accuracy Assessment Method	Omission Error	Commission Error	Overall Accuracy	Kappa
Points (Absence)	7/17=0.41	45/55=0.82	46/98=0.47	0.02
Points (Presence>0%)	45/81=0.56	7/43=0.16	46/98=0.47	0.02
Points (Presence>=25%)	29/49=0.59	23/43=0.53	46/98=0.47	-0.06
Points (Presence>=50%)	14/28=0.50	29/43=0.67	55/98=0.56	0.07
Points (Presence>=75%)	6/14=0.43	35/43=0.81	57/98=0.58	0.08
Area (Presence~100%)	3/11=0.27	7/15=0.47	17/27=0.63	0.27

Table 4.2. Percent coverage in validation points and plots greatly affects the error and accuracy of the results in Southeast Patapsco (August 17, 2013). Since Southeast Patapsco contained lower densities of wavyleaf basketgrass than McKeldin Recreation Area, the presence validation plots consisted of areas with lower wavyleaf basketgrass coverage.

Accuracy Assessment Method	Omission Error	Commission Error	Overall Accuracy	Kappa
Points (Absence)	29/81=0.36	15/67=0.22	61/105=0.58	0.01
Points (Presence>0%)	15/24=0.63	29/38=0.76	61/105=0.58	0.01
Points (Presence>=1%)	5/7=0.71	36/38=0.95	64/105=0.61	-0.03
Points (Presence>=5%)	2/3=0.67	37/38=0.97	66/105=0.63	0.00
Area (Plots with Presence>0%)	2/9=0.18	2/9=0.18	14/18=0.78	0.53
Area (Plots with Presence>=10%)	1/8=0.12	4/11=0.36	13/18=0.72	0.46
Area (Plots with Presence>=20%)	0/5=0.00	6/11=0.55	12/18=0.67	0.39
Area (Plots with Presence>=30%)	0/7=0.00	10/11=0.91	8/18=0.44	0.07

Supervised classification and MTMF results show low correspondence to our field measurements. From our fieldwork, we saw that wavyleaf basketgrass was generally located along the river, in relatively flat terrain, and in certain regions of the park. Supervised classification and MTMF results displayed confusion between the spectral signature of the wavyleaf basketgrass and other endmembers in these areas. The supervised classification results (not shown) contained very little wavyleaf basketgrass at the sites, even at our presence validation points. However, the overall accuracy of the minimum distance classification was unrealistically high due to chance, since few validation points contained wavyleaf basketgrass and the minimum distance algorithm overestimated absence. Likewise, we found no relationship between the MTMF scores and field measurements. We tried removing extreme and/or infeasible MTMF scores and reclassifying the ground coverage as present or absent, but neither method produced significant results. A binary classification of the MTMF scores as presence and absence produced a speckled image, with little basis in reality or the environmental features in

the landscape. The MTMF classified image and validation assessment can be seen in Figures 4.7 and 4.8. In the MTMF accuracy assessment graph, higher MTMF presence scores are associated with higher infeasibility scores.

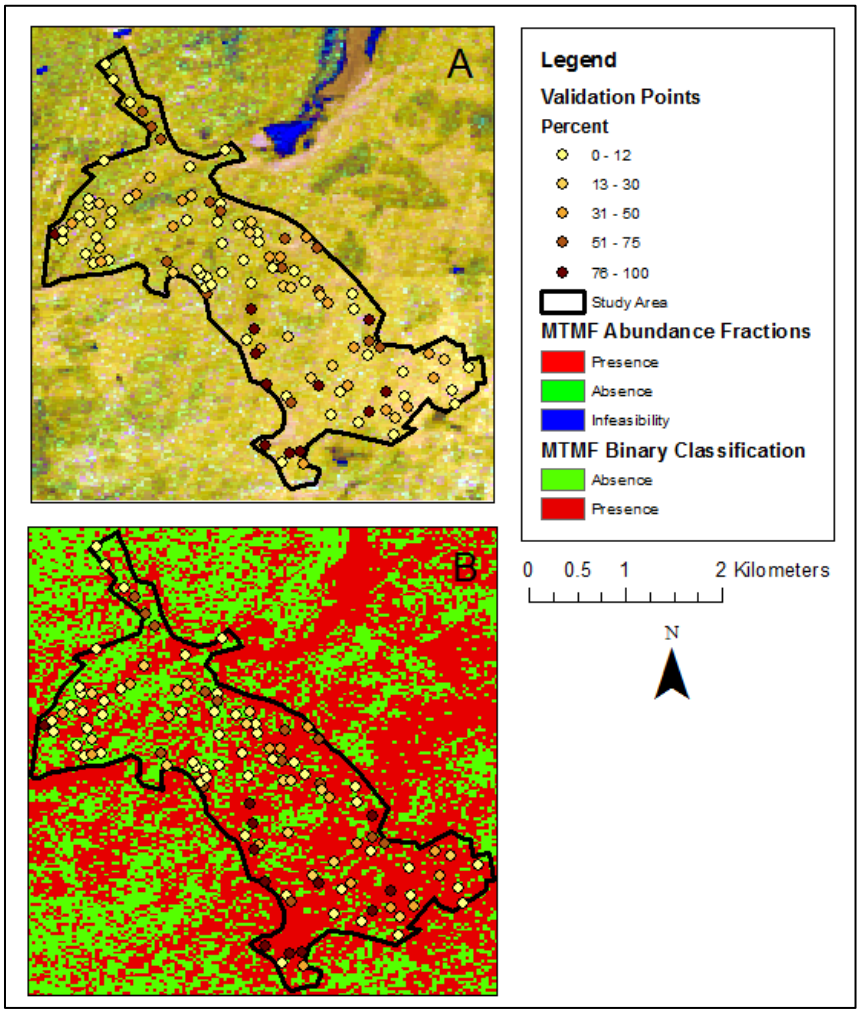


Figure 4.7. (A) MTMF score and (B) MTMF binary classification maps of wavyleaf basketgrass presence and absence. Random validation points are overlaid on the maps, color-coded by ground coverage class.

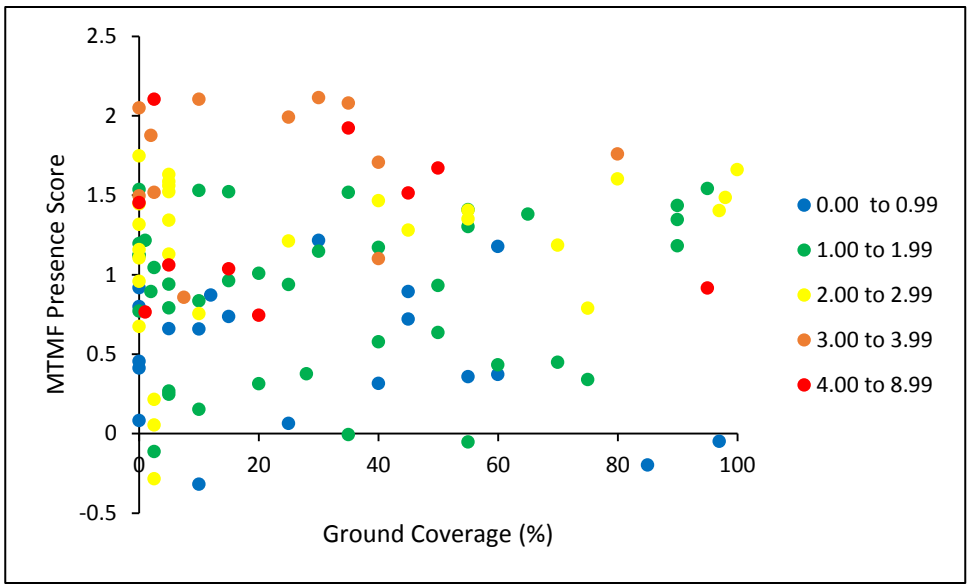


Figure 4.8. Accuracy assessment results from the MTMF presence endmember shows no linear relationship between field-mapped wavyleaf basketgrass ground coverage and the presence MF score (R^2=0.00). Points are color-coded according to presence infeasibility.

To improve our results, we investigated a potential cause of confusion in the classifications. Many of the wavyleaf basketgrass absence plots at McKeldin contained Japanese stiltgrass in high abundance. Thus, we decided to exclude those plots from analysis and evaluate the results. There was a noticeable improvement in error rates and overall accuracy. Omission and commission errors declined from 0.35 and 0.36 to 0.28, overall accuracy increased from 0.63 to 0.71; and the kappa coefficient improved to 0.43. At Southeast Patapsco, results of the accuracy assessment using the random points were low due to the relatively few points with any wavyleaf basketgrass presence, and the small patch size/low density of the patches present. This resulted in low accuracy for both map and reality presence classifications, although absence classifications were moderately accurate. As the percent of ground cover required to define presence increased, so did the accuracy for map absence when the plant was actually absent.

We also evaluated the role of seasonality in image classification through image acquisition date. We compared overall accuracy at both sites for several image dates, as shown in Table 4.3. Kappa values for the majority of these classifications indicate poor agreement with reality. The only exceptions are the October 2014 and August 2013 images, which show moderate agreement in their areal accuracy assessments. These classifications avoid over half of errors from a completely random classification. However, all point accuracy assessments had poor results.

Table 4.3. Temporal comparison of SAM and linear spectral unmixing overall accuracies across sites, with Kappa coefficient in parentheses. For both methods of accuracy, presence is defined as any amount of wavyleaf basketgrass.

		SAM Accuracy Assessment		Unmixing Accuracy Assessment	
Site	Date	Points	Area	Points	Area
McKeldin	Oct. 24, 2014	0.47 (0.02)	0.63 (0.27)	0.46 (0.00)	0.75 (0.51)
	Nov. 20, 2013	0.62 (0.03)	0.50 (0.00)	0.62 (-0.01)	0.50 (0.02)
Southeast	Aug. 17, 2013	0.58 (0.01)	0.78 (0.53)	0.59 (0.03)	0.58 (0.41)
Patapsco	Sept. 15, 2013	0.73 (0.19)	0.41 (-0.11)	0.70 (0.08)	0.44 (-0.06)

Discussion

Overview of Results

The spectral angle mapper and linear spectral unmixing were only moderately successful at mapping wavyleaf basketgrass in Patapsco Valley State Park. The overall accuracies for several images indicated moderate agreement between the classification and validation data; however, the kappa coefficients showed that many of the results were no better than chance. Out of all the classifications, only the August 2013 image from Southeast Patapsco and the October

2014 image from McKeldin produced useful classifications according to the kappa values. This finding was contrary to our expectations, as Southeast Patapsco had lower densities of wavyleaf basketgrass than McKeldin Recreation Area, and the images were taken under leaf-on conditions. Furthermore, this result reflected inconsistencies between image dates, sites, and accuracy assessment methods. No classified images of Southeast Patapsco or the alternative validation method produced such strong results. The reason for this discrepancy is not immediately clear to us.

The MTMF and SAM algorithms were surprisingly unsuccessful with images of McKeldin Recreation Area, even with the late fall imagery and nearly complete coverage of wavyleaf basketgrass. Linear spectral unmixing performed better, classifying the October image with moderate accuracy (overall accuracy=0.58, kappa=0.41). A key factor in this result could be the seasonality and other vegetation. While the understory in the October image was obscured by overstory foliage, wavyleaf basketgrass in the November image was obscured by leaves on the ground. Thus, neither the endmember nor the pixels truly represented a pure spectrum. Furthermore, results indicate that there is some confusion between wavyleaf basketgrass and other vegetation species. When Japanese stiltgrass is excluded from the analysis, SAM was moderately successful for large, high density plots of wavyleaf basketgrass in late fall imagery. However, given that the two plants typically grow together, confusion between the two species is likely to continue unless better methods are developed. Some options to improve the classification accuracy include the addition of Japanese stiltgrass as an endmember, the use of ancillary species-habitat data to differentiate the two species, or the classification of both species together.

Difficulty mapping wavyleaf basketgrass impeded our ability to evaluate both hypotheses. With respect to the first hypothesis, we were unable to determine the effect of plant coverage on classification accuracy due to inconsistencies in within-site and across-site results. While producer's and overall accuracy rose with increasing percent coverage of wavyleaf basketgrass at the validation points at both sites, the same pattern did not occur with the area-based accuracy assessment at Southeast Patapsco. Furthermore, the insignificant kappa coefficients call the validity of these results into question. No minimum threshold of ground coverage for accurate classifications was apparent across images, perhaps due to changing environmental conditions between images (i.e. overstory seasonality, site conditions, litterfall).

Inconsistencies between sites also complicated the evaluation of this hypothesis. Since the validation points at Southeast Patapsco had lower wavyleaf basketgrass percent coverage than those at McKeldin, we expected lower classification accuracy. Instead, results were comparable, which indicates one of two possibilities. First, the results could be similar because the endmember spectrum used was site-specific, such that lower density wavyleaf basketgrass was used at the lower density site. As a result, the algorithm was able to match the spectrum in the image. This finding would indicate that there is no minimum threshold for detection. The second option is that higher wavyleaf basketgrass coverage does improve classification accuracy, but the reduced presence of wavyleaf basketgrass at Southeast Patapsco led to lower rates of omission, and thus similar overall accuracies. Many more validation points were classified as absence at Southeast Patapsco than at McKeldin, and these tended to be classified correctly due to the distinct spectral signature of absence compared to the confusion of other vegetation classes with the presence spectra.

The lack of significant conclusions regarding the second hypothesis stemmed from issues with image timing. Despite our intent to acquire imagery at the moment when leaves were off the trees and the wavyleaf basketgrass was still green while other vegetation senesced, we were relatively unsuccessful. The cloud-prone nature of our study site greatly impeded this effort, although our final images were relatively cloud-free. Some leaves remained in the trees in the September and October images, and some areas had more leaves still in trees than others due to concurrent effects of topography and species composition. As typical for the ecoregion, the site also had a large range in species, with some areas having different mixed species composition than others. In addition, the leaves covered large portions of the target species. Resasco et al. (2007) only found success in linking their target species with Landsat images in November. Much like wavyleaf basketgrass, Amur honeysuckle (*Lonicera maackii*) retains its leaves until late in the season. Thus, the researchers found that images acquired in the beginning to middle of November are optimal for understory detection, although the best date depends on annual environmental conditions (Resasco et al. 2007). Under this logic, it is possible that our late October image was too early and the late November image was too late. Classification results at both sites may have improved with more appropriately timed images.

Comparison with Prior Literature

Many of the kappa coefficients were low and indicated poor agreement between the reference image and map classification, particularly in comparison to coefficients reported in other studies. Lawrence, Wood, and Sheley (2006) found kappa coefficients of 0.56 for spotted knapweed and 0.62 for leafy spurge, also used variable density plots, some of which had very low concentrations of the target species. They found that low positional accuracy was especially

problematic for small infestations, leading to difficulties in classifying individual pixels. This issue could be the same problem that we faced at Southeast Patapsco, and may also be partly responsible for the improvement in accuracy from point to areal accuracy assessment in some images. Kappa coefficients in other studies included 0.72 (Dehaan et al. 2007) and a range from 0.44-0.63 (Underwood et al. 2003). Dehaan et al. observed blackberry, which is a much larger plant than wavyleaf basketgrass, while Underwood at al. examined various scrub and iceplant densities amongst less diverse vegetation. Mundt et al. (2007) had similar producer's accuracy, although their user's and overall accuracies were slightly higher.

According to Lawrence, Wood, and Sheley (2006), classification accuracies below 70 percent are not useful for invasive plant management (Driscoll 2002). Although this criterion for success is debatable, a relatively high accuracy is necessary to ensure cost-effective and efficient management practices. Many of our overall accuracies were just below the 70 percent threshold; however, our area accuracy assessments show low omission errors at both sites, which is critical for successful detection. As a result, this approach may be useful to natural resource managers with a large area to inventory, as it successfully narrows down the potential locations of wavyleaf basketgrass. The high number of false positives indicates that many of the prospective locations will not actually have any invasions. However, the omission error was relatively good, although the small number of invasion sites may affect the ability to assess accuracy.

Potential Sources of Error and Future Improvements

In this study, many potential sources of error are systematic and stem from technical specifications of the instruments. For example, the Hyperion image is only accurate to a 30 m spatial resolution, which raises the problem of mixed pixels and the effect of other vegetation.

Since the species of interest is subcanopy, trees were present in all pixels with various amounts of leaves. Other vegetation, including similar grasses such as Japanese stiltgrass, was also located in both the presence and absence plots. Dudek et al. (2004) reported similar issues in their study mapping leafy spurge with imaging spectroscopy data of 20 m resolution; they attributed their low accuracies to the small, fragmented nature of the invasions (Mundt et al. 2007). In this study, the accuracy of the GPS units may also have affected the mapping effort and accuracy assessment points. Additionally, the Hyperion satellite has serious technical failures, including issues such as striping, a spectral smile, and poor geometric registration (Datt et al. 2003). Consequently, the measures that we took to correct these issues could introduce sources of error. The use of PCA in preprocessing, for instance, most likely resulted in a loss of information.

There are multiple other potential sources of error in this study. Human errors include inaccurate coverage estimates at the accuracy assessment points, subjective definitions for presence/absence with random points, and endmember selection (Mundt et al. 2007). Decisions MNF dimensionality and maximum angle thresholds also affected the classification outcomes. In addition, environmental variability within the study area further complicated the classification. Topographic shadows in the hilly terrain were common, and delineated areas of infestation did not have uniform coverage of weeds. Furthermore, differences in tree species, understory vegetation, and amount of leaf litter on the ground were also apparent.

We expected several of the imaging spectroscopy classification methods to produce successful results of the high density wavyleaf basketgrass areas. The poor results using mixture-tuned matched filtering were particularly surprising. Previous studies have shown that MTMF may produce better results than linear spectral unmixing when not all endmembers are known

(Dópido & Plaza 2011). However, Mundt et al. (2011) found weak relationships (R^2=0.32) between their MTMF scores and field abundance estimates. We attribute the low classification accuracies to many of the same reasons discussed above. In addition, these methods perform best when endmembers are available for all vegetation types present in the image. One possibility is that the diversity of tree species confounded the classification, especially since we were not using "pure" wavyleaf basketgrass endmembers. Our results may have improved if we included endmembers for the different tree species in our analysis. Furthermore, methods exist to generate "virtual endmembers" when endmembers are impure (Plaza et al. 2011).

In the future, we recommend the inclusion of environmental geospatial data layers and the use of a higher spatial resolution image that has a red-edge band. Beauchamp et al. (2013) found elevation and soil type to be among the best predictors for suitable habitat, and our own observations indicated a link between wavyleaf basketgrass presence and distance to water and trails. In addition, greater spatial resolution would improve the purity of endmembers and enable us to better differentiate between land cover classes. Many small patches of the invasion are currently sub-pixel, so reductions in pixel size would provide more options for pixel-level analysis. Similarly, many pixels on the boundary of wavyleaf basketgrass-covered areas have mixed composition, due to the overstory trees or adjacent patches of Japanese stiltgrass. Although mixed pixels are unavoidable, smaller pixels will enable more accurate vegetation classification of the understory as well as forest overstory. Smaller pixels also have the potential to see through gaps in the canopy to the wavyleaf basketgrass covering the forest floor.

Conclusions

Overall, this study shows that the SAM and linear spectral unmixing algorithms are moderately successful in mapping wavyleaf basketgrass in mixed deciduous forests. Imaging spectroscopy techniques can identify probable areas of wavyleaf basketgrass infestation, which could provide cost- and time-efficient direction for ground survey crews (Lass et al. 2005). However, the accuracy of these maps remains a major weakness in terms of their usefulness for natural resource decision-making. Furthermore, the applicability of these methods to sites beyond the study area has not yet been evaluated.

Through this work, we hope to improve image acquisition timing and field methods for measuring other invasive, understory grasses. Currently, little is known about wavyleaf basketgrass in its new environment. As more is learned about its ecology, we would like to build upon these methods to develop better options for tracking invasions. We hope this research has paved the way for future analysis with Germany's EnMAP satellite when it is launched. Although data acquisition is almost at a standstill as the remaining satellites fail from instrumentation issues, this type of research will likely gain speed with the launch of the next sensor.

Acknowledgements

The authors would like to thank J. Stein, M. Dempsey, and A. Epstein for their assistance in the field, and L. Ong for his help with image acquisition.

References

Beauchamp, Vanessa B, Koontz, Stephanie M, Suss, Christine, Hawkins, Chad, Kyde, Kerrie L, & Schnase, John L. (2013). An introduction to Oplismenus undulatifolius (Ard.) Roem. & Schult.(wavyleaf basketgrass), a recent invader in Mid-Atlantic forest understories 1, 2. *Journal of the Torrey Botanical Society, 140*(4), 391-413.

Boardman, Joseph W, & Kruse, Fred A. (1994). *Automated spectral analysis: a geological example using AVIRIS data, north Grapevine Mountains, Nevada.* Paper presented at the Proceedings of the Thematic Conference on Geologic Remote Sensing.

Brown, Ralph B, & Noble, Scott D. (2005). Site-specific weed management: sensing requirements-what do we need to see? *Weed Science, 53*(2), 252-258.

Card, Don H. (1982). Using known map category marginal frequencies to improve estimates of thematic map accuracy. *Photogrammetric Engineering and Remote Sensing, 48*(3), 431-439.

Congalton, Russell G. (1991). A review of assessing the accuracy of classifications of remotely sensed data. *Remote Sensing of Environment, 37*(1), 35-46.

Datt, Bisun, McVicar, Tim R, Van Niel, Tom G, Jupp, David LB, & Pearlman, Jay S. (2003). Preprocessing EO-1 Hyperion imaging spectroscopy data to support the application of agricultural indexes. *Geoscience and Remote Sensing, IEEE Transactions on, 41*(6), 1246-1259.

de Castro, Ana Isabel, Jurado-Expósito, Monterrat, Peña-Barragán, José M, & López-Granados, Francisca. (2012). Airborne multi-spectral imagery for mapping cruciferous weeds in cereal and legume crops. *Precision Agriculture, 13*(3), 302-321.

Dehaan, Remy, Louis, John, Wilson, Andrea, Hall, Andrew, & Rumbachs, Rod. (2007). Discrimination of blackberry (Rubus fruticosus sp. agg.) using imaging spectroscopy imagery in Kosciuszko National Park, NSW, Australia. *ISPRS journal of photogrammetry and remote sensing, 62*(1), 13-24.

Dópido, Inmaculada, & Plaza, Antonio. (2011). *Unmixing prior to supervised classification of urban imaging spectroscopy images.* Paper presented at the Urban Remote Sensing Event (JURSE), 2011 Joint.

Driscoll, Shana Grace. (2002). *Detecting and mapping leafy spurge (Euphorbia esula) and spotted knapweed (Centaurea maculosa) in rangeland ecosystems using airborne digital imagery.* Montana State University--Bozeman.

Dudek, Kathleen Burke, Root, Ralph R, Kokaly, Raymond F, & Anderson, Gerald L. (2004). *Increased spatial and temporal consistency of leafy spurge maps from multidate AVIRIS imagery: A modified, hybrid linear spectral mixture analysis/mixture-tuned matched filtering approach.* Paper presented at the Proceedings of the 13th JPL Airborne Earth Sciences Workshop, RO Green. Ed.,(ftp://popo. jpl. nasa. gov/pub/docs/workshops/04_docs/Dudek_aviris_2004_web. pdf), NASA Jet Propulsion Laboratory, Pasadena, California, USA.

EDDMapS. (2013). Early Detection & Distribution Mapping System. Retrieved April 26, 2015, from http://www.eddmaps.org

Glenn, Nancy F, Mundt, Jacob T, Weber, Keith T, Prather, Timothy S, Lass, Lawrence W, & Pettingill, Jeffrey. (2005). Imaging spectroscopy data processing for repeat detection of small infestations of leafy spurge. *Remote Sensing of Environment, 95*(3), 399-412.

He, Kate S, Rocchini, Duccio, Neteler, Markus, & Nagendra, Harini. (2011). Benefits of imaging spectroscopy remote sensing for tracking plant invasions. *Diversity and Distributions, 17*(3), 381-392.

Kloppenburg, Catherine. (2014). *Detecting leafy spurge in native grassland using imaging spectroscopy image analysis.* Lethbridge, Alta.: University of Lethbridge, Dept. of Geography, 2014.

Lamb, DW, & Weedon, M. (1998). Evaluating the accuracy of mapping weeds in fallow fields using airborne digital imaging: Panicum effusum in oilseed rape stubble. *Weed Research, 38*(6): 443-452.

Lass, Lawrence W, Prather, Timothy S, Glenn, Nancy F, Weber, Keith T, Mundt, Jacob T, & Pettingill, Jeffery. (2005). A review of remote sensing of invasive weeds and example of the early detection of spotted knapweed (Centaurea maculosa) and babysbreath (Gypsophila paniculata) with a imaging spectroscopy sensor. *Weed Science, 53*, 242-251.

Lawrence, Rick L, Wood, Shana D, & Sheley, Roger L. (2006). Mapping invasive plants using imaging spectroscopy imagery and Breiman Cutler classifications (RandomForest). *Remote Sensing of Environment, 100*(3), 356-362.

Lockwood, Julie L, Hoopes, Martha F, & Marchetti, Michael P. (2013). *Invasion* ecology(2nd ed.) West Sussex, UK: Wiley-Blackwell.

Maryland Department of Natural Resources. (1998). Trail Management Plan: Patapsco Valley State Park.

Maryland Department of Natural Resources. (2014). Patapsco Valley State Park History. from http://dnr.maryland.gov/publiclands/patapscohistory.asp

Miao, Xin, Gong, Peng, Swope, Sarah, Pu, Ruiliang, Carruthers, Raymond, Anderson, Gerald L, . . . Tracy, CR. (2006). Estimation of yellow starthistle abundance through CASI-2 imaging spectroscopy imagery using linear spectral mixture models. *Remote Sensing of Environment, 101*(3), 329-341.

Mundt, Jacob T, Streutker, David R, & Glenn, Nancy F. (2007). *Partial unmixing of imaging spectroscopy imagery: theory and methods.* Paper presented at the Proceedings of the American Society of Photogrammetry and Remote Sensing.

Nowak, David J, & Heisler, Gordon M. (2010). Air Quality Effects of Urban Trees and Parks. *National Recreation and Park Association*.

Peterson, EB. (2005). Estimating cover of an invasive grass (Bromus tectorum) using tobit regression and phenology derived from two dates of Landsat ETM+ data. *International Journal of Remote Sensing, 26*(12), 2491-2507.

Plaza, Antonio, Du, Qian, Bioucas-Dias, José M., & Jia, Xiuping. (2011). Foreword to the special issue on linear spectral unmixing of remotely sensed data. *IEEE Transactions on Geoscience and Remote Sensing, 49*(11), 4103.

Plaza, Antonio, Martínez, Pablo, Pérez, Rosa, & Plaza, Javier. (2004). A quantitative and comparative analysis of endmember extraction algorithms from imaging spectroscopy data. *Geoscience and Remote Sensing, IEEE Transactions on, 42*(3), 650-663.

Quintano, Carmen, Fernández-Manso, Alfonso, Shimabukuro, Yosio E, & Pereira, Gabriel. (2012). Linear spectral unmixing. *International Journal of Remote Sensing, 33*(17), 5307-5340.

Resasco, J, Hale, AN, Henry, MC, & Gorchov, DL. (2007). Detecting an invasive shrub in a deciduous forest understory using late-fall Landsat sensor imagery. *International Journal of Remote Sensing, 28*(16), 3739-3745.

Stehman, Stephen V, & Wickham, James D. (2011). Pixels, blocks of pixels, and polygons: Choosing a spatial unit for thematic accuracy assessment. *Remote Sensing of Environment, 115*(12), 3044-3055.

Underwood, Emma, Ustin, Susan, & DiPietro, Deanne. (2003). Mapping nonnative plants using imaging spectroscopy imagery. *Remote Sensing of Environment, 86*(2), 150-161.

Ustin, Susan L, DiPietro, D, Olmstead, K, Underwood, Emma, & Scheer, GJ. (2002). *Imaging spectroscopy remote sensing for invasive species detection and mapping.* Paper presented at the Geoscience and Remote Sensing Symposium 2002. IGARSS'02. 2002 IEEE International.

Williams, Amy Parker, & Hunt, E Raymond. (2002). Estimation of leafy spurge cover from imaging spectroscopy imagery using mixture tuned matched filtering. *Remote Sensing of Environment, 82*(2), 446-456.

Woods, A.J., Omernik, J.M., & Brown, D.D. (1999). *Level III and IV Ecoregions of Delaware, Maryland, Pennsylania, Virginia, and West Virginia*. Corvallis, OR.

5 Chapter 5—Assessing Understory Diversity Using a Lidar/Imaging Spectroscopy Fusion Approach

Abstract

Small canopy gaps in the temperate deciduous forest offer a unique opportunity to examine the understory vegetation using remote sensing. Prior studies have found mixed results regarding how representative the gap vegetation is of the adjacent understory. Thus, we propose a novel lidar/imaging spectroscopy fusion approach to identify canopy gaps and relate their species diversity to spectral reflectance. In this study, we measured vascular species diversity at 34 canopy gaps in Duke Forest, North Carolina. We acquired concurrent imaging spectroscopy data from ESA's HyPlant sensor and lidar data from NASA Goddard's LiDAR, Hyperspectral and Thermal Imager (G-LiHT). We used a lidar algorithm to extract the imaging spectroscopy data within the canopy gaps, then produced statistical models to relate species diversity to reflectance and lidar metrics. Reflectance, vegetation indices, and lidar metrics did not produce strong models, but texture-based approaches successfully related diversity to reflectance (R^2=0.49). Then we used traditional regression and spatial statistics approaches to relate vegetation diversity within gaps to diversity outside of gaps. Although these methods proved unsuccessful, outside gap diversity was successfully modeled from both inside and outside gap texture and lidar metrics.

Keywords: *Canopy gaps, understory, biodiversity, lidar, imaging spectroscopy, data fusion*

Introduction

The forest understory, which consists of vascular plants under 1 m in height, provides numerous essential ecosystem services. The understory is important for forest biodiversity, influences the trajectory of forest regeneration, interacts with the overstory, and cycles essential nutrients (Gilliam 2007). In addition, the understory is a valuable food source for wildlife, and the diversity of species improves the aesthetic value of the landscape. From a management perspective, knowledge of understory plant diversity can serve as indicators for forest health following a disturbance (Duffy & Meier 1992). Consequently, the importance of the understory vegetation leads to a clear need to quantify species diversity. Estimates of tree species richness are generally done with field measurements or remote sensing. However, below the forest canopy, the understory is scarcely visible with air- or spaceborne sensors (Singh, Davis, & Meentemeyer 2015). As a result, few studies have attempted to study understory plants or estimate diversity using remote sensing.

We propose a novel lidar/imaging spectroscopy fusion approach to identify canopy gaps and relate their species diversity to spectral reflectance. NASA Goddard's LiDAR, Hyperspectral, and Thermal (G-LiHT) high-resolution imager offers a unique opportunity to achieve this goal. We hypothesize that gaps in the canopy can provide sensors with a glimpse of small pockets of underlying vegetation, if the spatial resolution is roughly the size of, or finer than, the gaps. Use of lidar to map large canopy gaps is well-established (Asner et al. 2013), and several recent studies have successfully modeled species diversity in other vegetation types using imaging spectroscopy. These studies include simulated wetlands (Heumann, Hackett, & Monfils 2015), rainforests in Hawaii (Carlson et al. 2007), and the Peruvian Amazon rainforest (Féret & Asner 2014).

However, the viability of using lidar/imaging spectroscopy fusion to identify small gaps and estimate their species diversity is unknown. Thus, the following question arises: How does the forest understory in canopy gaps differ from the rest of the understory, and what is the effect of the understory composition on reflectance? Our objective is to determine how well a fusion of imaging spectroscopy and lidar technology can describe understory vegetation diversity, both within gaps and beyond. To achieve this end, we conduct extensive modelling using reflectance- and lidar-derived variables, spatial statistics, and regression approaches.

Understory species diversity is affected by many environmental factors that are often influenced by the surrounding vegetation. The availability and quality of resources such as light, water, and nutrients in the soil, have a major effect on the composition and number of species (Barbier, Gosselin, & Balandier 2008). Understory gap vegetation grow under a very different micro-climate than their more shaded counterparts (Denslow & Hartshorn 1994). Sunlit plants tend to have different levels of nitrogen, chlorophyll, and photosynthesis than shaded plants. As a result, the question arises as to whether the "gap vegetation" can provide any information on the other plants nearby.

Many studies have examined the relationship between gap dynamics and understory since the 1980s, although they have mainly investigated larger gaps (>40m^2) in temperate and tropical forests (Barton, Fetcher, & Redhead 1989; Canham et al. 1990; Denslow 1980). These gaps tend to be the result of tree fall and produce a variety of environmental effects, including increases in light availability, soil disturbance, soil moisture, and nutrient levels. In the tropics, the ecosystem-level consequences of canopy gaps affect species turnover, seedling establishment, and sapling growth (Denslow & Hartshorn 1994). However, this large discrepancy between gap and non-gap environmental conditions may be much less evident in other ecosystems, where

canopy cover is lower, natural disturbances occur more often, and seasonal patterns increase intra-annual variation.

Previous studies in non-tropical environments also tend to focus on larger gaps. In a longleaf pine stand, experimentally produced gaps of 1100 to 16300m^2 led to increased understory and seedling growth within the gaps (McGuire et al. 2001). Another study in the mixed-oak stands at Coweeta Hydrologic Lab examined rhododendron gaps of approximately 300m^2 on the understory microclimate. Although the researchers found a gap response in the amount of photosynthetically active radiation, there was no response in water content, soil temperature, or air temperature (Clinton 2003). Likewise, the gaps produced no increase in understory species diversity or light levels (Beckage et al. 2000). Even the few studies that analyze small gap vegetation suggest differences in plant abundance rather than composition compared to non-gaps (Goldblum 1997; Mladenoff 1990). Larger gaps may be necessary to cause more significant effects. A recent study by Yao et al. (2014) reached a similar conclusion in their examination of gaps with a mean size of 10 m^2. They found no significant difference in light availability and seedling community composition between gaps and non-gaps (Yao et al. 2014).

Material and Methods

Study Area

The study area consists of a 0.5 km x 1 km region in the northern Blackwood Division of the Duke Forest Teaching and Research Laboratory. According to Duke University, there are 984 plant species in 459 genera and 134 families in the Duke Forest. Dominant tree species in the mixed hardwood forest include red maple (*Acer rubrum*), green ash (*Fraxinus*

pennsylvanica), yellow-poplar (*Liriodendron tulipifera*), and sweetgum (*Liquidambar styraciflua*). During the field campaign, 34 canopy gaps with a range of sizes were selected for analysis based on a field assessment (Table 5.1, Figure 5.1). Although overstory species remain similar across the area, gaps were distributed across a topographic and moisture gradient to encompass a large number of species and abundance patterns.

Table 5.1. Canopy gap summary statistics (n=34).

Summary statistic	Area (m^2)
Mean area	13.15
Minimum area	0.32
Maximum area	62.95
Standard deviation	16.16
Mean area of correctly identified gaps	14.33
Mean area of incorrectly identified gaps	1.04

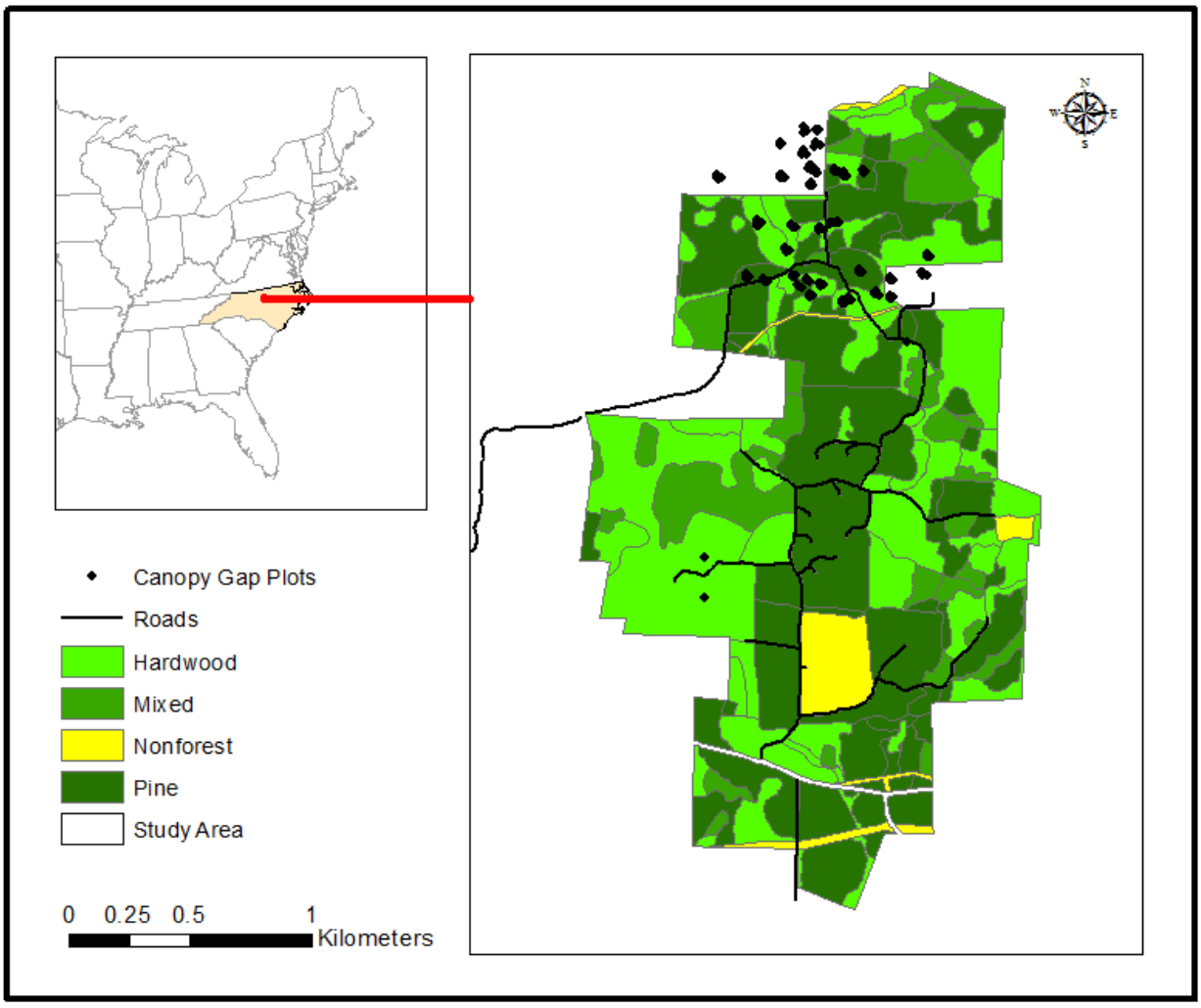


Figure 5.1. Map of sampled canopy gaps within several hardwood stands in the Duke Forest of North Carolina, US. Each cluster of points represents the six quadrat locations within and outside of each canopy gap.

Field Measurements

Quadrats of 0.5 m x 0.5 m were randomly placed at three locations within and three locations outside of each gap: 5m N, 10m N, and 10m E. These locations include areas with previously determined differences in canopy gap light regimes (Canham et al. 1990). Photographs were taken of each gap, each quadrat, and each unique plant species within a gap.

The overstory tree species and probable cause of gap creation, such as treefall or natural lack of canopy closure, were recorded. The percent cover of twigs, leaves, soil, and rock within each gap was also estimated along with the percent of living, green vegetation and total number of plants within each quadrat. Geographic coordinates at the approximate center of each gap and quadrat were acquired using a handheld GPS unit.

Ground locations within and outside of gaps were identified using a forestry sighting tube. The sighting tube was used to look up at the canopy and determine whether a given ground location was within the gap, and a handheld leveling instrument was used to ensure that the viewer was looking at the canopy with an approximately 180° angle. Biases occur when the sighting tube is angled greater than 5° from vertical (Jennings, Brown, & Sheil 1999). This method of sighting was based on the description by Korhonen et al. (2011) for measuring vertical canopy closure. Decisions regarding precise estimates of crown edges were based on the crosshair in the center of the sighting tube. Gaps were measured by selecting the longest transect as the length and the longest transect that intersected the length at a right angle as the width (Yao et al. 2014). Gap lengths and widths were measured and multiplied to approximate area.

The number of unique vascular plant species per plot and inside and outside each gap were counted and identified with the help of outside experts. Several species number counts were conducted: the number of known species, number of known genuses, and number of unique species. This latter count is the total number of different species, including those which could not be identified to the species-level but were verifiably different from those species already counted within the quadrat or gap. Those species that could not be identified to the species or genus level (due to small size, unidentifiable attributes, or unknown to examiners) were compared to the

other species present. If they could not be confirmed to be different, then they were not counted as separate species.

Image Acquisition and Processing

G-LiHT and the European Space Agency's HyPlant sensor acquired lidar and imaging spectroscopy data in September and October, 2014. The G-LiHT system includes an airborne scanning lidar instrument with 1 m spatial resolution and a 10 cm footprint at an altitude of 335 m, with 0.3 mrad beam divergence angle (Cook et al. 2013). The HyPlant sensor has a spatial resolution of 1 m, spectral resolution of 3-10 nm, and wavelength range of 370-2500 nm (Forschungszentrum Jülich 2015; Rossini et al. 2015). At the time of image acquisition, deciduous trees were beginning to show signs of senescence.

Significant processing of the lidar data was required prior to analysis. Individual datasets were mosaicked and reclassified into ground and non-ground classes. A digital elevation model (DEM) and digital surface model (DSM) with 1 m spatial resolution were generated from the ground returns and all returns, respectively (Evans et al. 2009). Whereas a DEM is a map of the ground elevation, a DSM is a height map of all ground and above-ground components (i.e., trees or buildings) in a region. Height above ground was calculated for the forest canopy using linear interpolation (Rosette et al. 2009). To determine the percentage of points hitting the understory, the number of lidar first returns below a 1.2 m threshold was divided by the total number of points in each pixel. These fractions were then extracted at the gap GPS points and used to evaluate the percent distributions for gap and non-gap points. A t-test verified a statistical difference in distributions.

Canopy gaps were located based on the density of lidar points below the DSM. An understory density threshold of 50 percent was used to delineate gap pixels. Success of the lidar method at identifying small canopy gaps was evaluated by comparing lidar gap shape and sizes with field records.

Mean reflectance values within each gap were extracted from the imaging spectroscopy data. Due to the large number of related wavelengths in the HyPlant images, spectral analysis was restricted exclusively to a set of wavelengths determined by Thenkabail et al. as the optimal combination for maximizing non-redundant information (Thenkabail et al. 2014) (Table 5.2). These wavelengths were then used to calculate image texture, which is the amount of tonal variation in the grayscale image based on spatial scale (Haralick, Shanmugam, & Dinstein 1973). To calculate texture per gap, we used co-occurrence matrices to apply a texture filter across 3x3 pixel windows with an eight pixel neighborhood. Texture metrics were generated using the same subset of wavelengths described above; variables include mean, variance, homogeneity, contrast, dissimilarity, entropy, second moment, and correlation. Equations were based on those by Haralick et al. (1973). Based on Pearson correlation analysis, the texture metrics most closely related to species diversity were then included as predictors in best subset linear regression models of species diversity.

Table 5.2. Selected wavelengths from Thenkabail et al. (2014) and reason for including each wavelength.

Wavelength (nm)	Band	Justification
550 +/- 5	100	Thenkabail recommendation: vegetation classification and discrimination
682 +/- 5	177	Thenkabail recommendation: crop type, crop discrimination
855 +/- 20	277	Thenkabail recommendation: crop type, crop discrimination
970	344	Thenkabail "Best 8 Bands"
1075 +/- 5	367	Thenkabail recommendation: crop type, crop discrimination
1180	386	Thenkabail "Best 8 Bands"
1450 +/- 5	433	Thenkabail recommendation: vegetation classification and discrimination
1725 +/- 5	482	Thenkabail recommendation: vegetation and crop discrimination
2025 +/- 5	536	Thenkabail recommendation: litter-soil differentiation
2133 +/- 5	555	Thenkabail recommendation: litter-soil differentiation
2205	568	Thenkabail "Best 8 Bands"
2260 +/- 5	578	Thenkabail recommendation: vegetation type

Statistical Analysis

A three-tiered statistical approach was used to examine understory diversity within canopy gaps, as shown in Figure 5.2. The first tier of analysis examined understory diversity within gaps, outside of gaps, and between gaps. Summary and frequency statistics were generated for gap diversity and quadrat composition based on selected ground coverage classes. A T-test and Kruskal-Wallis test by ranks were used to determine any significant differences in vegetation within gaps and across the sampling distance. Sørensen similarity index for beta diversity was calculated between inside and outside gap diversity (Magurran 2004). The index estimates beta diversity, or species turnover, across two locations based on the following formula: 2C/(A+B), where A represents the number of species in location A, B is the number of species in location B, and C is the number of species in common. Values range from 0 to 1, where 0 indicates complete turnover and 1 indicates perfect agreement in species.

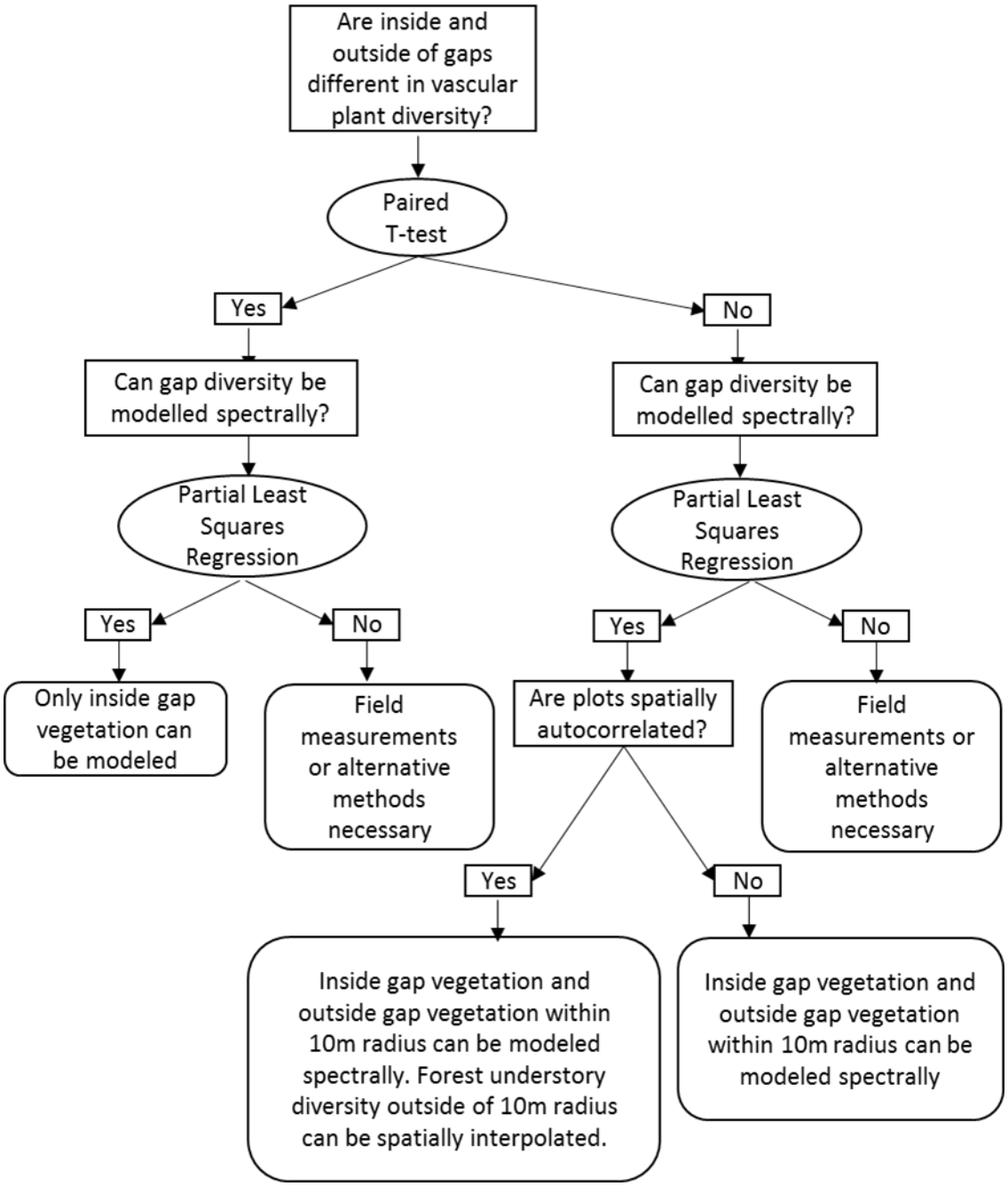


Figure 5.2. Flowchart outlining the research questions and basic statistical methods to conduct this study. Implications at each level of analysis are also shown.

The second tier of analysis involved modelling gap diversity using imaging spectroscopy reflectance and reflectance-derived metrics. This process is shown in Figures 5.3 and 5.4. First, the relationship between mean gap reflectance and inner-gap ground composition, defined as percent soil and percent live vegetation, was evaluated using Pearson correlation coefficients. Significant relationships would indicate good visibility of the ground and understory within the gap to the sensor, while poor relationships would imply the opposite. Next, inner gap species diversity was related to mean gap reflectance using Pearson correlation coefficients, best subsets multiple linear regression, and partial least squares regression (e.g. Asner et al. 2013). Regression approaches produced predictive equations for diversity using reflectance, vegetation indices, texture variables, and lidar metrics. Selected lidar metrics are shown in Table 5.3. Criteria for multiple linear regression model selection included number of variables and Bayesian Information Criterion. Model selection required variance inflation factors less than 5 to reduce multicollinearity (Ficetola & Padoa-Schioppa 2009). For partial least squares regression, we selected models by lowest overall cross-validation RMSE.

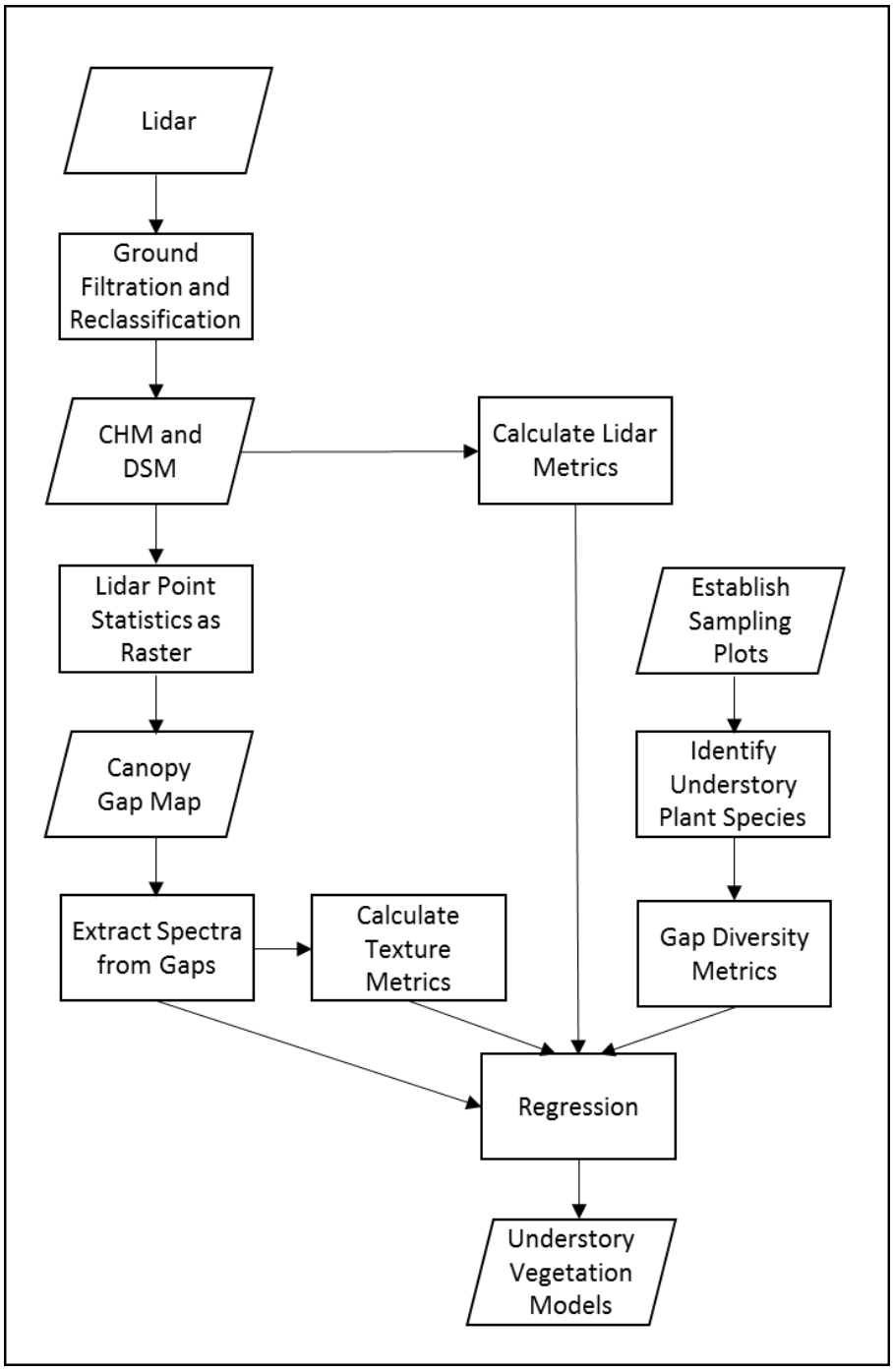


Figure 5.3. Process flowchart for gap identification and modelling with reflectance and texture metrics. Lidar data is used to produce a canopy gap, which is subsequently used to extract spectra from imaging spectroscopy data. Regression techniques are then applied to the spectra and understory plant species diversity.

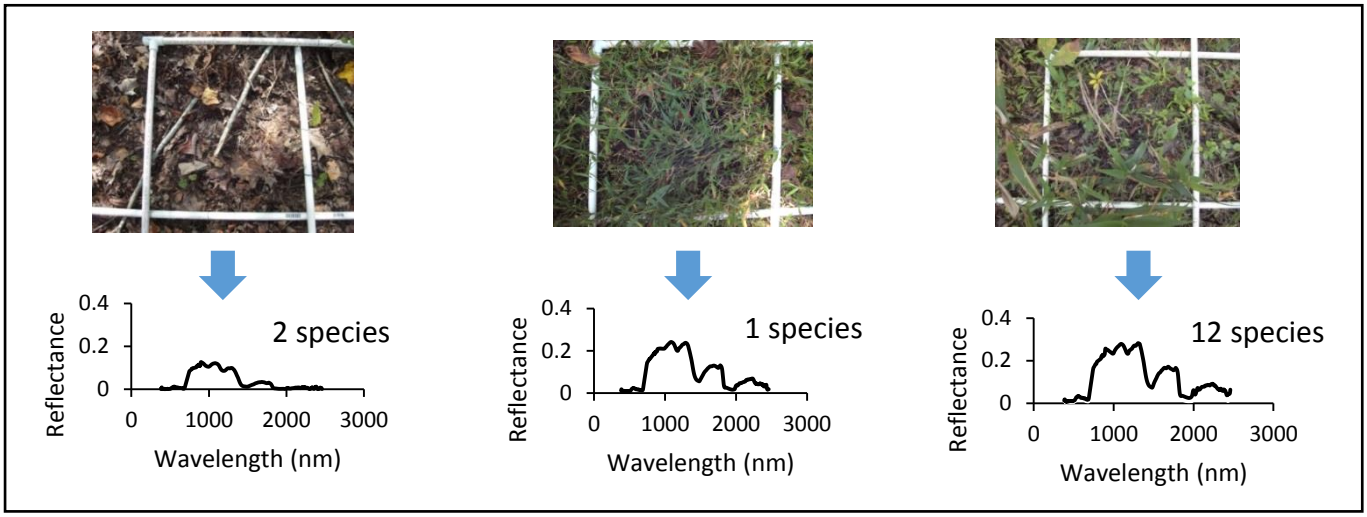


Figure 5.4. Selected canopy gap quadrat photos, their extracted spectra, and number of species that represent the diversity of vegetation conditions within the canopy gaps. The diversity of compositions inside canopy gaps—in terms of their soil/vegetation cover, species appearance and abundance, and number of unique species—produced unique spectral signatures. These characteristics are used in the second tier of statistical analysis to relate the number of species measured within canopy gaps to the spectral diversity measured from imaging spectroscopy data.

Table 5.3. Relevant lidar metrics included in multiple linear regression models (reprinted from the G-LiHT website (NASA, Undated).

Returns	Metric	Description	Units
All	Pulse_density	Laser pulse density	Pulses m^{-2}
	Pulse_scan_angle	Mean laser pulse scan angle (from nadir)	Degrees
	Returns_per_pulse	Mean number of returns per laser pulse	Counts
	All_mean	Mean of all return height	Meters
	All_qmean	Quadratic mean of all return heights	Meters
	All_stdev	Standard deviation of all return heights	Meters
	All_skew	Skewness of all return heights	Meters
	All_kurt	Kurtosis of all return heights	Meters
	All_p10….all_p100	Height percentiles (10% increments) of all returns	Meters
	All_d0…all_d9	Density deciles (10% increments) of all returns	Fraction
	All_refl_max	Maximum relative reflectance of all returns	Fraction
Shrub	Shrub_mean	Mean return height of shrub returns	Meters
	Shrub_stdev	Standard deviation of shrub heights	Meters
	Shrub_refl_max	Maximum relative reflectance of shrub single returns	Fraction
Tree	Tree_mean	Mean of tree return heights	Meters
	Tree_qmean	Quadratic mean of tree return heights	Meters
	Tree_stdev	Standard deviation of tree return heights	Meters
	Tree_rugosity	Standard deviation of gridded CHM values	Meters
	Tree_skew	Skewness of tree return heights	Meters
	Tree_kurt	Kurtosis of tree return heights	Meters
	Tree_qmean	Quadratic mean of tree return heights	Meters
	Tree_fcover	Fraction of first returns intercepted by tree	Fraction
	Tree_fract_all	Fraction of all returns classified as tree	Fraction
	Tree_p10…Tree_p100	Height percentiles (10% increments) of tree returns	Meters
	Tree_d0…Tree_d100	Density deciles (10% increments) of tree returns	Fraction
	Tree_iqr	Interquartile range of tree returns	Meters
	Tree_vdr	Vertical distribution ratio of tree returns	Unitless
	Tree_mad	Median absolute deviation of tree returns	Meters
	Tree_aad	Mean absolute deviation of tree returns	Meters
	Tree_crr	Canopy Relief Ratio	Unitless
	Tree_refl_max	Maximum relative reflectance of tree single returns	Fraction
Ground	Ground_elev_mean	Mean of ground return elevations	Meters
	Ground_slope	Mean slope derived from gridded DTM	Degrees
	Ground_aspect	Aspect derived from gridded DTM	Degrees
	Ground_refl_max	Maximum relative reflectance of ground single returns	Fraction

The third tier of analysis involved the use of spatial statistics and regression to model outside and total gap species diversity. Global Moran's I was calculated to examine the spatial distribution of the number of understory species (i.e., whether the points were clustered, dispersed, or randomly distributed), and its level of significance. Global Moran's I tests the null hypothesis that species diversity is randomly distributed across the canopy gap plots (Perry et al. 2010). Next, semivariograms were produced to describe the relationship between understory species diversity and plot distance. Based on the range, sill and nugget, semivariograms depict the range of distances at which points are spatially autocorrelated, and at what point they are no longer correlated (Scheller & Mladenoff 2002). Vegetation diversity was evaluated across the 10 m transect inside gaps and outside gaps, as well as between plots.

Best subset multiple linear regression models were used to predict outside gap and total gap diversity using the aforementioned model selection criteria. Spectral and lidar characteristics of the entire gap area were extracted from 15 m buffered polygons surrounding the approximate gap centroid and then used as predictor variables in the models.

Results

The lidar method for gap identification successfully found 29 of 34 gaps measured in the field. Gaps were typically located within 3 m of their field GPS location. Problems with gap identification were generally due to missing returns in the lidar data. In the few cases where the lidar algorithm showed gaps as completely closed canopy, the gaps were smaller than 2 m in diameter. Figure 5.5 shows a comparison of the actual gap photos and the lidar-derived gaps with the actual GPS points from each quadrat overlaid. Note the high correspondence between actual quadrat locations and the lidar-derived gap locations in Figure 5.5C.

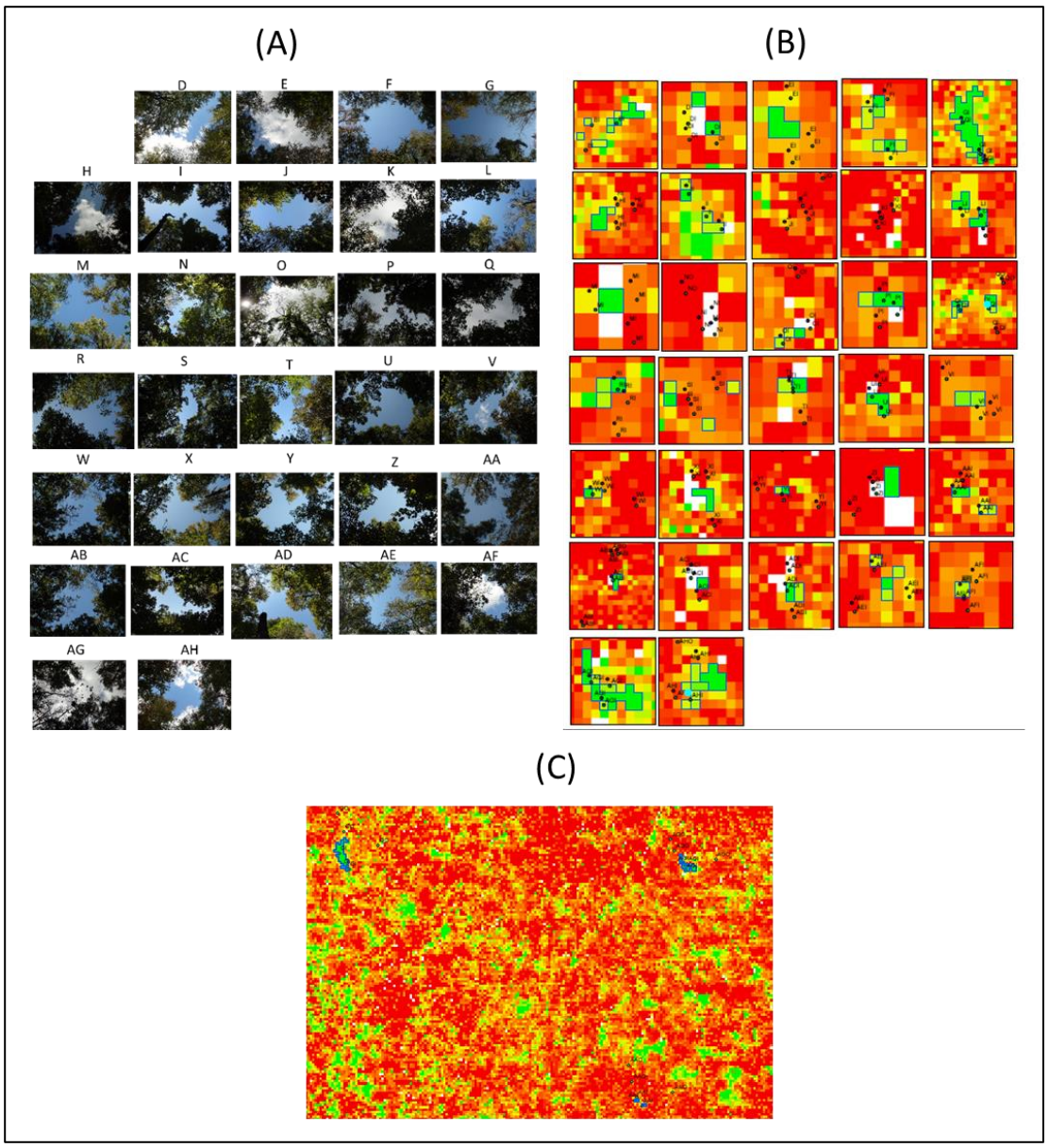


Figure 5.5. Comparison of (A) gap photos from the field and (B) lidar-derived gap delineations, color-coded by percentage of canopy closure (green represents open gap, red represents closed canopy). Gaps are in corresponding location in both images. Neither photos nor raster images are drawn to scale, and photo orientation is unknown. (C) shows a landscape-level view of several large canopy gaps with the quadrat GPS points ("AA", "AG", and "G" represent gap IDs, and the "O" or "I" at the end denotes inside or outside gap, respectively).

Over 500 understory plants were identified to the species. Figure 5.6 shows species frequency for the most common 20 species. The vegetation comprised 86 different species; the most common species were two non-natives, *Microstegium vimineum* and *Lonicera japonica*, and *Acer rubrum*. The mean Sørensen similarity index for inside and outside gap species diversity is 0.22, with a standard deviation of 0.07 and a range from 0.10 to 0.38. Based on the quadrat samples, the average inner gap composition was 26 percent green, 6 percent woody material, 57 percent leaves, 8 percent soil, and 2 percent rock. Thus, understory vegetation only composed a small portion of the gaps.

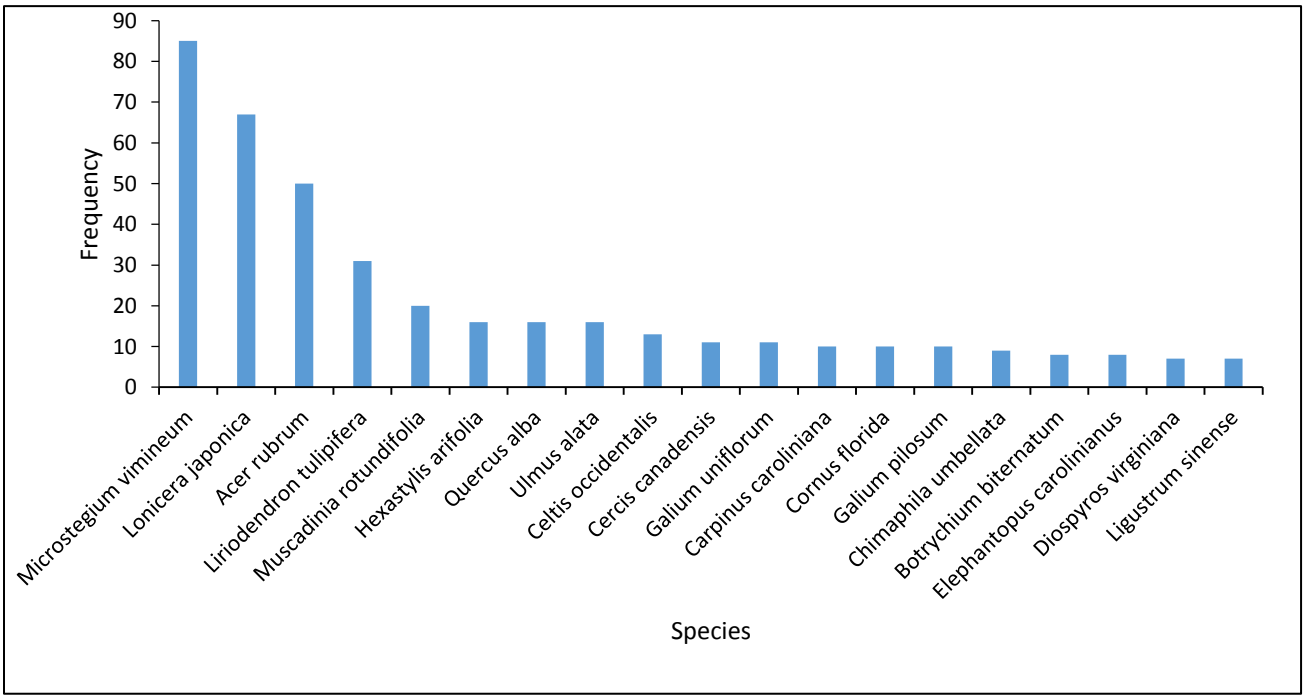


Figure 5.6. Species frequency chart for the most common 20 species found within the canopy gaps at Duke Forest.

Kruskal-Wallis test results indicate insignificant differences in the number of species between plot locations (χ^2 = 0.11, p-value = 0.99). There is no evidence of a statistical difference in the mean rank of species at plot locations inside and outside of canopy gaps. Figure 5.7 shows the medians and ranges of number of species by sampling location relative to the gap. A matched pairs test between total number of species inside and outside each gap also found no significant differences in species diversity (t=-0.74, p=0.47).

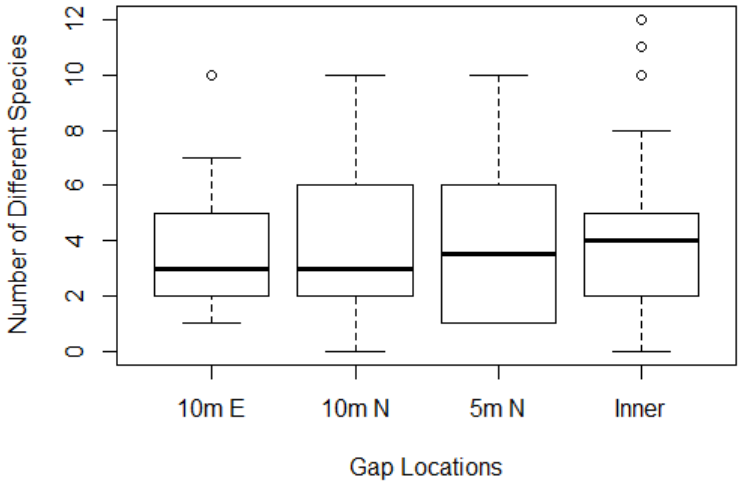


Figure 5.7. Boxplot depicting medians and ranges of species diversity at different gap locations across 34 plots.

Moran's I between all plots indicated significant spatial autocorrelation in species diversity (observed=0.18, expected=-0.00, p=0.00). A separate Moran's I analysis with only plots inside gaps reinforced these findings that inside-gap plots are spatially autocorrelated (observed=0.38, expected=-0.01, p=0.00). However, the total species diversity inside each gap (as defined by number of different species across all three inside quadrats) is not spatially autocorrelated (observed=-0.04, expected=-0.04, p=0.91). This finding implies that some method

is necessary to estimate the number of species within each gap, but then the number found between quadrats can be interpolated.

Mean gap reflectance produced typical vegetation spectral curves (Figure 5.8). The range in reflectance across the gaps is quite large given the consistency in overstory species. In addition, due to the high percentage of soil compared to understory plants in many of the gaps but the lack of signs of a soil curve in the spectra, it is possible that the gap reflectance curves are seeing more of the overstory than understory. This conclusion is supported by the lack of significant relationships between inside-gap reflectance and gap composition data.

There were several significant correlations between mean gap reflectance and number of different species ($p<0.05$). Significant wavelengths include 1939-2000 nm and 2439-2477 nm. The selected multiple linear regression model with 3 variables had a predictive capability of $R^2=0.46$, Adjusted $R^2=0.39$, and RMSE=3.41. The intercept and two variables were significant at $p<0.05$; the third variable was significant at $p<0.10$. The selected partial least squares regression model used 5 variables, 6 latent factors, and had RMSE=4.09 and $R^2=0.27$.

The use of texture-based approaches also produced significant results (Figure 5.8). Table 5.4 shows the significant correlations ($p<0.05$) for selected wavelengths out of all texture metrics. Figure 5.4 displays the multiple linear regression results for the selected two-variable model (2 variable model: $R^2=0.49$, Adjusted $R^2=0.44$, $p=0.00$; 3 variable model: $R^2=0.57$, Adjusted $R^2=0.51$, $p=0.00$). Wavelengths used in the two-variable model are Dissimilarity 1726 and Dissimilarity 2133; the three-variable model also includes Mean 2260.

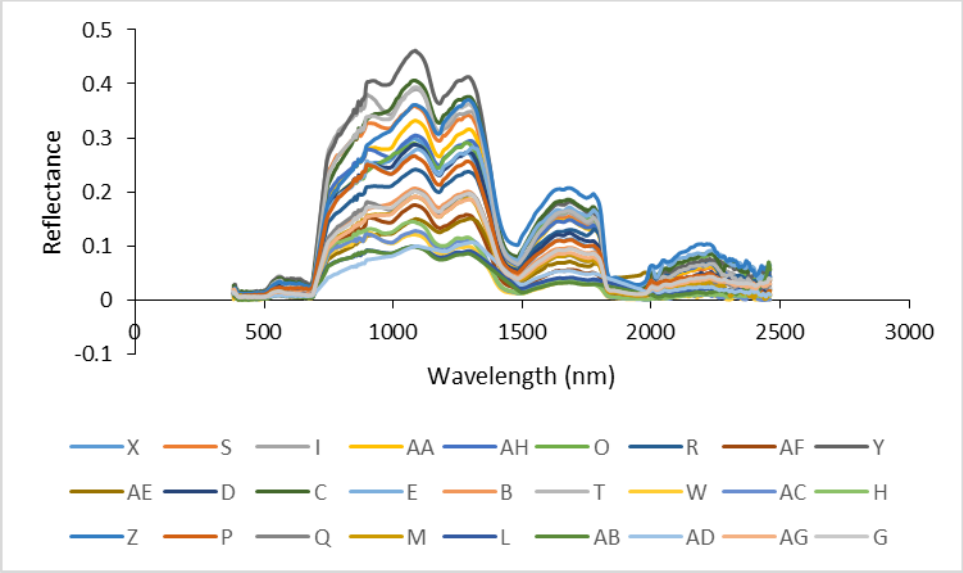


Figure 5.8. Spectral curves extracted from HyPlant imaging spectroscopy data based on canopy gap pixels delineated using G-LiHT lidar. Each spectrum represents the mean reflectance within each gap.

Table 5.4. Significant correlations ($p<0.05$) for selected wavelengths out of all co-occurrence texture metrics at gaps (n=26). Co-occurrence texture metrics are as follows: mean, variance, homogeneity, contrast, dissimilarity, entropy, second moment, and correlation.

Wavelength feature (nm)	Metric	Pearson coefficient (r)	p-value
1726	Dissimilarity	0.41	0.04
2028	Variance	0.46	0.02
	Contrast	0.47	0.02
	Dissimilarity	0.48	0.01
2133	Mean	0.39	0.05
	Variance	0.48	0.01
	Homogeneity	-0.40	0.04
	Contrast	0.56	0.00
	Dissimilarity	0.56	0.00
2205	Variance	0.46	0.02
	Contrast	0.52	0.01
	Dissimilarity	0.51	0.01
2260	Mean	0.41	0.04
	Variance	0.51	0.01
	Homogeneity	-0.40	0.04
	Contrast	0.52	0.01
	Dissimilarity	0.50	0.01

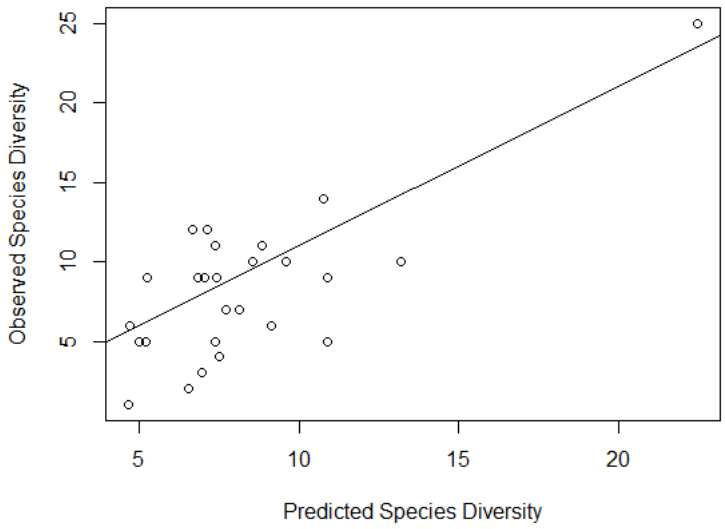


Figure 5.9. Multiple linear regression results for number of different species in canopy gaps and co-occurrence texture variables at selected wavelengths. Line Y=X shown for scale.

Multiple linear regression models for inside gap, outside gap, and total species diversity are shown in Table 5.5. Linear regression cannot be used to predict outside gap diversity from inside gap diversity successfully. However, outside gap and total gap diversity can be modeled using similar sets of variables as inside gap diversity. Due to large differences in texture values across flight lines in the buffered outside metrics, only canopy gaps occurring within one flight line were used to generate outside metric models. Consequently, these models are based on only half the gaps (df=12-13) as the other models (df=26).

All selected models and predictors were significant at a=0.05. Texture metrics produced stronger models than lidar metrics both inside and outside gaps. For total gap diversity, some lidar metrics produced higher coefficients of determination than texture metrics in selected models. However, only texture metrics were selected in all combined models using inside gap lidar and texture variables as predictors. Alternatively, models selected from outside gap texture

and lidar metric predictors used a combination of both variable types. These models also showed stronger predictive capabilities; however, they were based on fewer observations. No significant models could be produced using tree percentages for inside or outside gap diversity, or all returns percentages for total diversity.

Table 5.5. Multiple linear regression models for species diversity at three different plot locations relative to canopy gaps: inside gaps, outside gaps, and both. Species diversity outside gaps is pooled across three quadrats located at 5 m and 10 m north of the gap, and 10 m east. Predictors labelled "15 m Buffer" were generated from spectral and lidar extractions from a 15 m buffered region around the approximate gap centroid. All other predictors are based on spectral and lidar extractions within representative inside-gap pixels.

Response Diversity	Predictors	Selected Model	R^2	Adj R^2	P
Inside	All returns percentages	all_p10 + all_p70	0.21	0.14	0.05
Inside	All returns densities	all_d3	0.24	0.21	0.01
Inside	All returns	tree_d1 + tree_d2	0.34	0.29	0.00
Inside	Texture Inside	Dissimilarity 1726 + Dissimilarity 2133	0.49	0.44	0.00
Inside	All returns and Texture	Dissimilarity 2205 + Contrast 1183	0.47	0.43	0.00
Inside	15m Buffer: All returns and Texture	ground_elev_mean + Correlation 2028	0.77	0.73	0.00
Outside	Tree percentages	tree_p70 + tree_p40 + tree_mad	0.29	0.20	0.05
Outside	All returns percentages	all_p100 + all_p90	0.25	0.19	0.04
Outside	All returns densities	all_d3	0.31	0.28	0.00
Outside	All returns	all_d0 + tree_fract_all + ground_elev_mean	0.56	0.50	0.00
Outside	Texture Inside	Homogeneity 2260 + Homogeneity 1074 + Correlation 855	0.65	0.60	0.00
Outside	All returns + Texture	Dissimilarity 551 + Homogeneity 1074	0.60	0.56	0.00
Outside	15m Buffer: All returns and Texture	all_p10 + Second Moment 2028	0.63	0.57	0.00
Total	Tree percentages	tree_d3	0.27	0.24	0.00
Total	All returns densities	all_d3	0.33	0.30	0.00
Total	All returns	shrub_mean + shrub_refl_max + ground_elev_mean	0.65	0.61	0.00
Total	Texture Inside	Homogeneity 1074 + Dissimilarity 2260	0.51	0.47	0.00
Total	All returns + texture	Homogeneity 1074 + Dissimilarity 2260	0.51	0.47	0.00
Total	15m Buffer: All returns and Texture	tree_refl_max	0.39	0.35	0.01

Discussion

Gap Identification

The lidar algorithm identified the majority of gaps found in the field, and only missed a few of the smaller gaps. This result was expected; difficulty measuring gap size in the field led the field team to mistakenly select several gaps that were much too small for the resolution of the imagery. Gap size and shape varied quite a bit between measurements in the field and those derived from the sensor, thus leading to some difficulty in delineating the complete gap from the lidar data. Many gaps contained non-contiguous sections that we chose not to include because they would have involved inclusion of relatively closed sections of canopy. Furthermore, our gap threshold of 50 percent was relatively arbitrarily selected, and it is possible that a different value would have selected slightly different gap pixels.

Diversity Measurements

Field measurements indicate differences in the composition of species inside and outside of small gaps at the study site, but no significant differences in the number of species. The mean rank of species diversity differed insignificantly between the inside gap locations and each of the outside gap locations. However, the Sørensen similarity index measured little overlap in species composition inside and outside of gaps. Given the small size of the gaps and the relative canopy openness in many plot locations outside of the gaps, this result was particularly surprising. Prior studies attributed differences in species composition at gap and non-gap locations to variations in microclimate, photosynthetically active radiation, nutrients, and species adaptations (Brokaw & Scheiner 1989; Denslow & Hartshorn 1994). However, it is also possible that the beta diversity values are artificially low, as many species could not be identified to the species level. In the

future, an assessment of plant turnover at the genus level may provide a more accurate assessment of community dynamics than the current approach.

Species Diversity Models

The best remote sensing models for understory species diversity were obtained using image textural metrics, rather than reflectance or lidar variables. The success of texture-based regression variables is in agreement with previous studies in a variety of ecosystems (Heumann, Hackett, & Monfils 2015; Wood et al. 2013). With the exception of invasion locations, quadrats typically had few plants per square meter. Consequently, the sensor was predominately viewing high percentage of soil, mixed with a small amount of plant coverage. As a result, any textural differences between pixels were most likely due to changes in species composition, either as an individual plant or mixture of plants. We believe that this is the reason the textural variables outperformed reflectance or vegetation indices.

Although models vary substantially in their selected variables, some variables occur repeatedly. Mean ground elevation and all_d3 occur in all three model classes, and can easily be explained. In this study area, the link between ground elevation and species diversity may be related to water, moisture, and microclimate patterns, as many of the plots at lower elevations are located near streams. Riparian forests may experience more frequent disturbances and contain greater heterogeneity than adjacent upland forests, which could increase niche diversity (Araujo Calçada et al. 2015). Streams offer seed dispersal mechanisms for many plants, and also promote favourable site conditions by providing water, nutrients, and organic matter (Araujo Calçada et al. 2015). In addition, the canopy tended to be more open along the stream than elsewhere, which could increase light availability and promote greater species richness at the local level (Frelich,

Machado, & Reich 2003). Depending on the height of the canopy, all_d3 is likely measuring the height of the subdominant trees or shrubs, which also have a tendency to affect the understory light regime and plant diversity (Brantley & Young 2009).

Multiple texture metrics and wavelengths occurred repeatedly as variables of significance in the correlations and/or regression models. Texture metrics at 2260 and 1074 nm wavelengths were selected most frequently, along with the Dissimilarity metric at various wavelengths. The success of these particular metrics in the correlations and/or regression models (e.g. variance, contrast, and dissimilarity) is likely due to their emphasis on differences between pixels, rather than similarities. Furthermore, a previous study related variance and contrast metrics to a diversity of foliage heights (Wood et al. 2013). All three variables were significantly correlated with mature forest biomass in the Brazilian Amazon, and variance alone was significantly related to successional forest biomass (Lu & Batistella 2005). Although species diversity is a much different characteristic, the biomass of the surrounding forest canopy may affect light levels, which could serve as a proxy for understory diversity. Similarly, the understory vegetation in the gaps spanned a wide range of biomass levels, despite all studied plants being below the 1 m height threshold. Plants with high amounts of Japanese stiltgrass tended to have nearly complete plant coverage, while those with native plants ranged from little to moderate coverage.

Neither lidar variables nor spectral reflectance independently produced strong models. Lidar metrics produced models with lower predictive capabilities, and they generally did little to improve the models apart from the texture variables. This finding could be due to multipath errors, difficulty differentiating the understory from the ground with lidar, poor canopy penetration, insufficient lidar pulse density, or lack of relationship between species diversity and understory height patterns (Aguilar et al. 2010; Evans & Hudak 2007; Riaño et al. 2007).

Similarly, efforts to model understory diversity with spectral reflectance were only moderately successful. Species diversity can be estimated using various regression approaches, but the predictions had high error. These findings are consistent with some earlier research on species diversity models for other vegetation types (Nagendra 2001). While spectral signatures within a pixel may be linked to a target species or physiological condition (e.g. nutrient, structure), and algorithms exist for spectral unmixing, there is additional complexity to estimate species richness. Greater inter- than intra-specific spectral diversity may be required (Asner & Martin 2008). Alternatively, canopy shadows inside gaps may have also obscured model predictions from reflectance (Lu & Batistella 2005).

Despite strong relationships with texture, one issue raises concerns regarding validity of the models. The lack of clear relationships between greenness and soil composition within gaps and the respective spectra does not support our ability to model species diversity inside gaps using remote sensing. However, there are several possible reasons: (1) composition estimates are incorrect; (2) relationships with understory species diversity use a surrogate environmental characteristic that is visible by the overhead sensor (i.e. overstory species diversity); (3) selected vegetation indices were not representative of the dominant cover type (e.g. leaves covered a much larger area than soil, and they may be confused with plants); (4) correlations may not be the proper way of modelling mixed pixels (e.g. a spectral unmixing approach may work better); or (5) modelling approaches are not valid. Additionally, this scenario represents an even more complex modelling exercise than originally expected, since the mixed pixels consist of relatively low percentages of plants. The high concentration of leaves on the forest floor may explain why the understory spectral curve resembles an overstory curve, rather than looking more like a mixture of soil and vegetation.

Outside Gap Models

Although understory vegetation within gaps can be predicted from imaging spectroscopy data, no relationships could be drawn directly between inside- and outside-gap diversity. Spatial autocorrelation results were inconsistent across locational classes. The absence of any spatial autocorrelation in the inside-gap diversity measurements is confusing given the significant spatial autocorrelation of both total and outside gap diversities. Furthermore, the variogram shows an unusual trend, where furthest apart are more similar than plots at a range of closer distances. We hypothesize that this result may be due to underlying environmental variability. In a future study, we would like to further examine these variables, such as temperature (e.g. G-LiHT thermal data), moisture, and tree species. In addition, the next step is to evaluate whether gap size has any effect on inside-outside gap models.

The underlying assumption of this study was that canopy gaps would provide an invaluable glimpse of the understory that would enable successful modelling of the understory vegetation diversity. While gap vegetation was modeled with fairly high accuracy, our unexpected ability to model outside gap diversity with similar variables and explanatory power contradicted the premise of the study. This study showed that both inside and outside metrics offer the potential for inside, outside, and total gap diversity estimates. Consequently, the use of gaps to view the understory may not be necessary, as the outside gap area may provide similar information as the inside gap area. Reasons for this result are not immediately clear. Two possible causes are that the outside gap area also allows views of the understory through small openings, or that linkages exist between overstory characteristics and understory species diversity. Either way, this finding obviates the use of spatial autocorrelation or regression-based approaches to derive outside gap diversity as a function of inside gap diversity and/or location.

However, we caution readers to keep in mind that the large disparity in number of observations between models using inside and outside metrics make a direct comparison between the two approaches difficult.

Applications and Future Directions

Understory diversity predictive maps provide natural resource managers with a starting point for understory inventories, as they identify potential biodiversity "hotspots" and areas with particularly low diversity. While the former may indicate good wildlife habitat or potential conservation areas, the latter are also informative. At many gaps, extremely low species diversity indicated the presence of a dominant invasive species. Knowledge of these locations is critical to map invasions and monitor spatial patterns. Thus, understory diversity estimates at both extremes could be useful for a variety of purposes.

Despite strong predictive capabilities of the models, it should be noted that these models are location-specific and apply to only a small area within Duke Forest. Care must be taken not to apply these relationships elsewhere. The variables of significance may not hold up in other locations or time periods, where there are different levels of disturbance, light regimes, and geographic patterns (Bartels & Chen 2010). However, this paper provides a unique proof-of-concept for testing elsewhere. Furthermore, it must be remembered that only approximately 50 percent of the percent variation was explained, and other factors influence the spatial patterns of understory diversity. Model error rates are particularly problematic if they form the basis for regression models of the non-gap understory. Error from spectral models may propagate through the regression and produce highly inaccurate estimates for the remainder of the understory.

There are several options for improving model relationships. The incorporation of other related variables, such as topography and LAI, may improve models due to their correlation with productivity, light interception, and overstory diversity. Another possibility is to model diversity using genus or family, since related species may resemble each other both physically and spectrally. This option would also lead to a larger range in diversity values, since many species were unidentifiable at this taxonomic level. Furthermore, we also recommend modelling diversity using a greater number of bands, which may be necessary to identify the spectral variability between multiple species (Thenkabail et al. 2014).

Unfortunately, the high spatial and spectral resolution of G-LiHT is currently uncommon in remote sensing technology. Many sensors have high spatial resolution, but fewer have high spectral resolution, and even fewer have both. Thus, this study may only be repeatable with similar sensors, or may require larger gaps to compensate for the coarser resolution. We should note, however, that we focused on small gaps because prior literature has produced mixed results regarding differences between inside- and outside-gap vegetation. As gaps increase in size, the environmental differences within and outside of the gap typically become more pronounced. Consequently, lower resolution sensors may impede the ability to relate gap diversity to the surrounding understory.

Conclusion

Our study shows that lidar/imaging spectroscopy fusion can be a useful technique for studying understory vegetation diversity within canopy gaps. Although spectral modelling is limited to areas seen by the overhead sensor, regression-based approaches can provide diversity estimates for the understory both inside and outside of gaps. While users must be mindful of

prediction error, there are many applications for predictive maps generated with these methods. Potential applications of this work include locating areas of high and low biodiversity. Knowledge of high biodiversity is useful for identifying areas with great potential for ecosystem services, such as food and material resources, landscape aesthetics, and nutrient cycling. They may also be indicative of good forest health, as many species can fill various niches and potentially increase productivity, particularly in the event of a disturbance. Alternatively, sites with low diversity may indicate a disturbance or invasive plant, which could benefit from human involvement. Furthermore, understory diversity is important for forest regeneration, as many seedlings grow up to form the forest canopy.

Acknowledgments

We are grateful to our colleagues in the NASA/ESA Joint Campaign and the Duke Forest staff for their assistance with fieldwork, image acquisition, and data processing. We would also like to thank J. Peterson, J. Seiler, and T. Wieboldt for their assistance with plant identification. This research was funded by a grant through the USDA National Needs Fellowship, USDA-NIFANNF-2010-03349, the McIntire-Stennis Cooperative Forestry Research Program through the USDA CSREES under Project VA-136614, and the Department of Forest Resources and Environmental Conservation at Virginia Tech.

References

Aguilar, Fernando J, Mills, Jon P, Delgado, Jorge, Aguilar, Manuel A, Negreiros, JG, & Pérez, José L. (2010). Modelling vertical error in LiDAR-derived digital elevation models. *ISPRS Journal of Photogrammetry and Remote Sensing, 65*(1), 103-110.

Araujo Calçada, Emmanuelle, Lenoir, Jonathan, Plue, Jan, Broeckx, Laura S, Closset-Kopp, Déborah, Hermy, Martin, & Decocq, Guillaume. (2015). Spatial patterns of water-deposited seeds control plant species richness and composition in riparian forest landscapes. *Landscape Ecology*, 1-14.

Asner, Gregory P, Kellner, James R, Kennedy-Bowdoin, Ty, Knapp, David E, Anderson, Christopher, & Martin, Roberta E. (2013). Forest canopy gap distributions in the southern Peruvian Amazon. *PloS one, 8*(4), e60875.

Asner, Gregory P, & Martin, Roberta E. (2008). Airborne spectranomics: mapping canopy chemical and taxonomic diversity in tropical forests. *Frontiers in Ecology and the Environment, 7*(5), 269-276.

Barbier, Stéphane, Gosselin, Frédéric, & Balandier, Philippe. (2008). Influence of tree species on understory vegetation diversity and mechanisms involved—a critical review for temperate and boreal forests. *Forest Ecology and Management, 254*(1), 1-15.

Bartels, Samuel F, & Chen, Han YH. (2010). Is understory plant species diversity driven by resource quantity or resource heterogeneity? *Ecology, 91*(7), 1931-1938.

Barton, Andrew M, Fetcher, Ned, & Redhead, Steven. (1989). The relationship between treefall gap size and light flux in a Neotropical rain forest in Costa Rica. *Journal of Tropical Ecology, 5*(4), 437-439.

Beckage, Brian, Clark, James S, Clinton, Barton D, & Haines, Bruce L. (2000). A long-term study of tree seedling recruitment in southern Appalachian forests: the effects of canopy gaps and shrub understories. *Canadian Journal of Forest Research, 30*(10), 1617-1631.

Brantley, Steven T, & Young, Donald R. (2009). Contribution of sunflecks is minimal in expanding shrub thickets compared to temperate forest. *Ecology, 90*(4), 1021-1029.

Brokaw, Nicholas VL, & Scheiner, Samuel M. (1989). Species composition in gaps and structure of a tropical forest. *Ecology*, 538-541.

Canham, Charles D, Denslow, Julie S, Platt, William J, Runkle, James R, Spies, Tom A, & White, Peter S. (1990). Light regimes beneath closed canopies and tree-fall gaps in temperate and tropical forests. *Canadian Journal of Forest Research, 20*(5), 620-631.

Carlson, Kimberly M, Asner, Gregory P, Hughes, R Flint, Ostertag, Rebecca, & Martin, Roberta E. (2007). Hyperspectral remote sensing of canopy biodiversity in Hawaiian lowland rainforests. *Ecosystems, 10*(4), 536-549.

Clinton, Barton D. (2003). Light, temperature, and soil moisture responses to elevation, evergreen understory, and small canopy gaps in the southern Appalachians. *Forest Ecology and Management, 186*(1), 243-255.

Cook, B.D., Nelson, R.F., Middleton, E.M., Morton, D.C., McCorkel, J.T., Masek, J.G., . . . Montesano, P.M. (2013). NASA Goddard's LiDAR, Hyperspectral and Thermal (G-LiHT) Airborne Imager. *Remote Sensing, 5*(8), 4045-4066.

Denslow, Julie Sloan. (1980). Gap partitioning among tropical rainforest trees. *Biotropica,* 47-55.

Denslow, Julie Sloan, & Hartshorn, Gary S. (1994). Tree-fall gap environments and forest dynamic processes. *La Selva: ecology and natural history of a neotropical rain forest*, 120-127.

Duffy, David Cameron, & Meier, Albert J. (1992). Do Appalachian herbaceous understories ever recover from clearcutting? *Conservation Biology, 6*(2), 196-201.

Evans, Jeffrey S, & Hudak, Andrew T. (2007). A multiscale curvature algorithm for classifying discrete return lidar in forested environments. *Geoscience and Remote Sensing, IEEE Transactions on, 45*(4), 1029-1038.

Evans, Jeffrey S, Hudak, Andrew T, Faux, Russ, & Smith, Alistair. (2009). Discrete return lidar in natural resources: Recommendations for project planning, data processing, and deliverables. *Remote Sensing, 1*(4), 776-794.

Féret, Jean-Baptiste, & Asner, Gregory P. (2014). Mapping tropical forest canopy diversity using high-fidelity imaging spectroscopy. *Ecological Applications, 24*(6), 1289-1296.

Ficetola, Gentile Francesco, & Padoa-Schioppa, Emilio. (2009). Human activities alter biogeographical patterns of reptiles on Mediterranean islands. *Global Ecology and Biogeography, 18*(2), 214-222.

Forschungszentrum Jülich. (2015). HyPlant. Retrieved Feb. 24, 2015, from http://www.fz-juelich.de/ibg/ibg-2/EN/methods_jppc/HyPlant/node.html

Frelich, Lee E, Machado, José-Luis, & Reich, Peter B. (2003). Fine-scale environmental variation and structure of understorey plant communities in two old-growth pine forests. *Journal of Ecology, 91*(2), 283-293.

Gilliam, Frank S. (2007). The ecological significance of the herbaceous layer in temperate forest ecosystems. *Bioscience, 57*(10), 845-858.

Goldblum, David. (1997). The effects of treefall gaps on understory vegetation in New York State. *Journal of Vegetation Science, 8*(1), 125-132.

Haralick, Robert M, Shanmugam, Karthikeyan, & Dinstein, Its' Hak. (1973). Textural features for image classification. *Systems, Man and Cybernetics, IEEE Transactions on*(6), 610-621.

Heumann, Benjamin W, Hackett, Rachel A, & Monfils, Anna K. (2015). Testing the spectral diversity hypothesis using spectroscopy data in a simulated wetland community. *Ecological Informatics, 25*, 29-34.

Jennings, SB, Brown, ND, & Sheil, D. (1999). Assessing forest canopies and understorey illumination: canopy closure, canopy cover and other measures. *Forestry, 72*(1), 59-74.

Korhonen, Lauri, Korpela, Ilkka, Heiskanen, Janne, & Maltamo, Matti. (2011). Airborne discrete-return LIDAR data in the estimation of vertical canopy cover, angular canopy closure and leaf area index. *Remote Sensing of Environment, 115*(4), 1065-1080.

Lu, Dengsheng, & Batistella, Mateus. (2005). Exploring TM image texture and its relationships with biomass estimation in Rondônia, Brazilian Amazon. *Acta Amazonica, 35*(2), 249-257.

Magurran, Anne E. (2004). *Measuring biological diversity*. Oxford, UK: Blackwell.

McGuire, John P, Mitchell, Robert J, Moser, E Barry, Pecot, Stephen D, Gjerstad, Dean H, & Hedman, Craig W. (2001). Gaps in a gappy forest: plant resources, longleaf pine regeneration, and understory response to tree removal in longleaf pine savannas. *Canadian Journal of Forest Research, 31*(5), 765-778.

Mladenoff, David J. (1990). The relationship of the soil seed bank and understory vegetation in old-growth northern hardwood-hemlock treefall gaps. *Canadian Journal of Botany, 68*(12), 2714-2721.

Nagendra, Harini. (2001). Using remote sensing to assess biodiversity. *International journal of remote sensing, 22*(12), 2377-2400.

NASA. (Undated). G-LiHT LiDAR Plot-Scale Metrics. Retrieved Jan. 25, 2015, from ftp://fusionftp.gsfc.nasa.gov/multimedia/docs/metrics_readme.pdf

Perry, Eileen M, Dezzani, Raymond J, Seavert, Clark F, & Pierce, Francis J. (2010). Spatial variation in tree characteristics and yield in a pear orchard. *Precision agriculture, 11*(1), 42-60.

Riaño, David, Chuvieco, Emilio, Ustin, Susan L, Salas, Javier, Rodríguez-Pérez, José R, Ribeiro, Luis M, . . . Fernández, Helena. (2007). Estimation of shrub height for fuel-type mapping combining airborne LiDAR and simultaneous color infrared ortho imaging. *International Journal of Wildland Fire, 16*(3), 341-348.

Rosette, J.A., North, P.R.J., Suárez, J.C., & Armston, J.D. (2009). A comparison of biophysical parameter retrieval for forestry using airborne and satellite LiDAR. *International Journal of Remote Sensing, 30*(19), 5229-5237.

Rossini, M, Nedbal, L, Guanter, L, Ač, A, Alonso, L, Burkart, A, . . . Drusch, M. (2015). Red and far-red sun-induced chlorophyll fluorescence as a measure of plant photosynthesis. *Geophysical Research Letters*.

Scheller, Robert M, & Mladenoff, David J. (2002). Understory species patterns and diversity in old-growth and managed northern hardwood forests. *Ecological Applications, 12*(5), 1329-1343.

Singh, Kunwar K, Davis, Amy J, & Meentemeyer, Ross K. (2015). Detecting understory plant invasion in urban forests using LiDAR. *International Journal of Applied Earth Observation and Geoinformation, 38*, 267-279.

Thenkabail, Prasad S, Gumma, Murali Krishna, Teluguntla, Pardhasaradhi, & Mohammed, Irshad A. (2014). Hyperspectral remote sensing of vegetation and agricultural crops. *Photogrammetric Engineering and Remote Sensing, 80*(8), 697-709.

Wood, Eric M, Pidgeon, Anna M, Radeloff, Volker C, & Keuler, Nicholas S. (2013). Image texture predicts avian density and species richness. *PloS one*, *8*(5), e63211.

Yao, Ai-Wen, Chiang, Jyh-Min, McEwan, Ryan, & Lin, Teng-Chiu. (2014). The effect of typhoon-related defoliation on the ecology of gap dynamics in a subtropical rain forest of Taiwan. *Journal of Vegetation Science*.

6 Chapter 6—Conclusion

Summary and Conclusions

Results from this work conclude that imaging spectroscopy can be used to evaluate ecosystem health through forest nutrient status, nitrogen models, species diversity estimates, and identification of invasive plant species. Specifically, this work addressed each of the objectives described earlier.

1. **Objective 1:** Loblolly pine foliar macronutrient concentrations can be successfully modeled at the regional scale across the Southeastern US. Partial least squares regression models produced R^2 values for the regional models as follows: nitrogen: 0.74, phosphorus: 0.58, potassium: 0.41, calcium: 0.39, and magnesium: 0.51. No models could be produced at the local scale, even with the inclusion of several fertilization treatments and canopy levels. This result suggests that the local sites had an insufficient nutrient gradient for intraspecific models based on reflectance.
2. **Objective 2:** Reflectance and W scattering coefficients produced successful nitrogen models across loblolly pine plots at the canopy scale. Models showed similar explanatory power for nitrogen. W scattering coefficients were significantly correlated with nitrogen at multiple wavelengths in the visible and near-infrared regions, although reflectance variables were not. W scattering coefficients explained approximately 20 percent of the nitrogen variation across the plots. The direction of some of the correlations with W and unusually high DASF values indicate a need for greater experimentation, perhaps across more plots, multiple sites, and with better quality images.

3. **Objective 3:** Imaging spectroscopy algorithms, such as SAM and linear spectral unmixing, were moderately successful in identifying wavyleaf basketgrass invasions in mixed deciduous forests. Overall classification accuracy ranged from 0.35-0.78 for SAM, and 0.44-0.75 for linear spectral unmixing, depending on the particular image and accuracy assessment method. Kappa coefficients were generally low, although several values indicated better-than-chance classifications (0.41-0.53). Due to poor classification results for some images, the relationship between percent coverage and accuracy is not clear.
4. **Objective 4:** The lidar algorithm identified 29 of 34 canopy gaps measured in the field. In conjunction with lidar, the imaging spectroscopy data produced relatively accurate models for understory species diversity within the gaps using texture metrics (R^2=0.49). Outside gap diversity was successfully modeled from both inside and outside gap texture and lidar metrics (R^2=0.65 and 0.63).

Future Directions

Specific recommendations for future analysis are covered within each chapter, but there are some general ideas interwoven across multiple chapters. First, many stages of the analysis could have benefitted from a larger study area or greater number of study sites. With the exception of Chapter 2, plots were generally localized within a particular area. Thus, the question arises as to how these results would hold up across a larger area. Similarly, these chapters call for future analysis of other nutrients, species, and locations. Do the results in chapter 4 hold true of other invasive species? How do the results and predictive accuracy change when the models developed in chapter 3 are applied to other nutrients, besides nitrogen? Can species diversity estimates from

chapter 2 work in pine forests as well, which are naturally more open, and may not require canopy gaps? These are all valid follow-up questions which warrant further analysis in the future.

Finally, airborne imaging spectroscopy sensors, such as G-LiHT and HyPlant, that are relatively inexpensive and offer opportunities for data fusion or flexible flight lines, may be the next great technology. However, despite their many advantages in terms of spatial and temporal resolution, they do not offer the same overpass frequency and predictability as their satellite counterparts. As such, for the methods investigated in this dissertation to become more widely used, there is an urgent need for greater imaging spectroscopy technology to be freely and readily available. With the addition of new satellites, there will also be a need to test these methods under the new sensors' specifications.

Appendix A—Table of tree plots

Plot	DBH (m)	Crown Width (m, EW x NS)	CHM 1m pixel	CHM 3m pixel	CHM 5m pixel	CHM 7m pixel
LP 10	1.02	8.2 x 9.2	22.02	22.79	22.02	23.14
LP 11	0.99	7.1 x 6.1	20.84	22.42	23.81	23.95
LP 12	1.03	6.8 x 7.1	18.10	20.20	21.28	21.08
LP 13	1.09	7.6 x 8.1	21.39	21.39	24.47	24.47
LP 14	0.68	3.2 x 3.8	16.49	16.56	16.94	17.43
LP 15	0.9	6.5 x 5.9	22.68	22.72	22.72	23.62
LP 16	0.91	4.2 x 4.9	18.53	20.43	19.23	20.43
LP 17	0.76	3.4 x 4.3	22.19	23.32	23.32	24.72
LP 18	0.84	4.1 x 3.9	2.24	2.29	21.22	21.30
LP 19	1.24	8.6 x 9.1	22.84	23.69	23.69	23.56
LP 1N	0.6	4.1 x 4.8	20.09	20.89	22.72	22.72
LP 1W	1.07	3.3 x 4.6	21.07	22.72	22.72	22.72
LP 2	0.68	4.1 x 4.9	9.75	12.91	12.84	12.91
LP 20 N	0.53	2.3 x 2.2	14.85	18.62	19.38	19.42
LP 20 W	0.77	5.2 x 4.4	10.88	19.70	20.33	23.56
LP 21	0.54	2.2 x 2.5	12.81	15.13	15.13	16.59
LP 22 N	0.64	2.9 x 3.1	21.89	22.08	22.23	23.56
LP 22 W	0.94	4.4 x 4.0	21.77	22.00	22.23	23.51
LP 23	1.19	5.4 x 3.7	28.69	28.71	28.71	28.71
LP 24 N	0.49	1.5 x 2.1	16.52	18.74	22.53	21.25
LP 24 W	0.83	5.3 x 5.4	18.53	20.31	21.56	21.25
LP 25	1.12	3.5 x 5.2	24.25	25.00	25.00	25.00
LP 26	0.82	4.3 x 4.4	19.63	20.61	21.58	21.11
LP 27	0.38	2.7 x 2.9	8.84	8.84	8.84	9.02
LP 28	0.59	5.4 x 4.3	7.46	8.56	8.85	8.85
LP 29	0.51	3.7 x 3.5	1.02	5.96	8.37	9.20
LP 3	1.1	4.0 x 5.4	21.93	22.25	23.86	23.86
LP 30	0.11	1.1 x 1.4	2.67	4.15	5.08	5.08
LP 31	1.23	6.2 x 4.1	21.72	24.33	28.68	29.83
LP 32	0.57	4.6 x 3.9	8.04	8.91	9.67	10.25
LP 33	0.68	3.8 x 4.0	11.05	13.02	13.02	13.02
LP 34	2.05	14.1 x 13.6	30.34	32.04	32.04	32.04
LP 35	1.79	10.7 x 9.1	6.38	32.51	33.88	33.68
LP 36	0.17	1.2 x 1.6	2.90	5.09	5.09	5.09
LP 37	1.93	12.2 x 12.2	34.61	35.54	35.54	35.54
LP 38	0.02	0.7 x 1.1				
LP 4	0.96	5.4 x 7.2	18.11	20.66	21.48	21.87
LP 5	0.9	5.7 x 3.9	20.94	22.55	22.77	22.77
LP 6	1.71	10.6 x 9.2	30.31	31.50	31.50	31.18
LP 7	1.84	10.6 x 8.9	31.96	32.49	33.18	34.08

LP 8	0.54	3.3 x 3.6	9.57	11.20	11.20	11.20
LP 9	2.05	10.4 x 9.3	27.41	28.36	28.74	28.74