

A COMPARISON OF METHODS FOR SCALING FIELD DATA FOR USE IN MAPPING  
DRYLAND ECOSYSTEM VEGETATION WITH AIRBORNE IMAGING  
SPECTROSCOPY

A Thesis  
by  
MEGAN MALONEY

Submitted to the Graduate School  
at Appalachian State University  
in partial fulfillment of the requirements for the degree of  
MASTER OF ARTS

August, 2017  
Department of Geography and Planning

A COMPARISON OF METHODS FOR SCALING FIELD DATA FOR USE IN MAPPING  
DRYLAND ECOSYSTEM VEGETATION WITH AIRBORNE IMAGING  
SPECTROSCOPY

A Thesis  
by  
MEGAN MALONEY  
August, 2017

APPROVED BY:

---

Jessica Mitchell, Ph.D.  
Chairperson, Thesis Committee

---

Jeffrey Colby, Ph.D.  
Member, Thesis Committee

---

Howard S. Neufeld, Ph.D.  
Member, Thesis Committee

---

Kathleen Schroeder, Ph.D.  
Chairperson, Department of Geography

---

Max C. Poole, Ph.D.  
Dean, Cratis D. Williams School of Graduate Studies

Copyright by Megan Maloney, 2017  
All Rights Reserved

## **Abstract**

### **A COMPARISON OF METHODS FOR SCALING FIELD DATA FOR USE IN MAPPING DRYLAND ECOSYSTEM VEGETATION WITH AIRBORNE IMAGING SPECTROSCOPY**

Megan Maloney  
B.A., Sweet Briar College  
M.A., Appalachian State University

Chairperson: Jessica Mitchell, Ph.D.

Dryland ecosystems cover 41% of Earth's terrestrial surface. Globally, these lands house a third of our growing human population as well as many endangered and listed species. Drylands provide essential ecosystem services such as rangeland, water filtration, soil genesis, wildlife habitat, and carbon sequestration. Drylands store twice the organic carbon of forest ecosystems due to their large area and high soil organic carbon pool. However, recent research shows interannual variation in drylands is responsible for 39% of the variability in global carbon sequestration rates. Interconnected pressures with poorly understood feedback interactions degrade land and limit ecosystem services.

Foliar nitrogen (N) is commonly used as an indicator of vegetative growth due to its role in photosynthetic processes, which relates to the ability of dryland ecosystems to provide these services. This research investigates scaling methods of field data to interpret aerial surveys for N estimation using imaging spectroscopy. I compared the performance of four field-based methods to scale sagebrush foliar N estimates from shrub to the plot level (10 m x

10 m) for 21 plots collected in a dryland ecosystem in 2014 and 2015. Partial least squares regression was used to relate the four series of foliar N plot estimates to imaging spectroscopy variables to determine which field data collection variables and scaling methods provided strong relationships between the foliar N estimates and the remote sensing data. The regression models were ranked using adjusted  $R^2$  and RMSE. Results showed sensitivity to scaling method; pretreatment of imaging spectroscopy signals; subdividing the dataset into years; reducing predictor variables to reduce noise; and the number of model iterations. The best performing scaling methods used biomass allometry with density counts or cover estimates with leaf thickness with a log transformation and Savitzky-Golay smoothing method. Regression models selected different sets of wavelengths as significant predictors, with several relying on wavelengths in the visual range associated with chlorophyll absorbance and few relying on wavelengths in the "red edge" of 800-850 nm. The best performing model ( $R^2 = .88$ ; RMSE = 0.14 g/m<sup>2</sup>) used biomass allometry to scale from leaf N to plot-level N and a subset of wavelengths that consistently performed well across model iterations: 677 (near the 660 absorption feature for chlorophyll a), 992 (near the 990 absorption feature for starch), 1133, 1208 (near the 1200 absorption feature for water, cellulose, starch, lignin), 1213, 1218, 1223, and 1263 nm. The coefficients of this model were applied to imaging spectroscopy data at a 1 m<sup>2</sup> resolution across the study area, the Reynolds Creek Experimental Watershed in Idaho, USA, to create a wall to wall map of predicted foliar N values. The methods reported here can generate foliar N maps to 1) inform rangeland and conservation managers on forage quality, 2) investigate patterns in weed invasions and fire regimes, and 3) parameterize Ecosystem Demographics Models to predict future ecosystem scale structural dynamics, including carbon sequestration.

## **Acknowledgments**

I would like to thank my thesis adviser, Dr. Jessica Mitchell, for her support of the project and my interests and for tirelessly helping me to learn remote sensing concepts and instrumentation through numerous amazing opportunities. Thank you to my committee members, Dr. Howie Neufeld and Dr. Jeff Colby for providing encouragement as well as insight and guidance on the project design and manuscripts. Thank you to Dr. Peter Olsoy, Dr. Jill Thomley, and Dr. Holly Hirst for their aptitude and patience in explaining statistical methods; to Dr. Mitch Parry and Amanda Smith for their essential roles in the leaf area image processing, and Dr. Parry's insights on applying the PLSR results; and to Kaitlyn Finan for her instruction; and to Brielle Davidson, Lauren Andersen, Charles Robinson, and Chris Ely for their vital assistance in data wrangling. Thank you to Andrew Poley and Hamid Dashti, top picks for my apocalypse team; to Luke Spaete for the coffee and rescue; to Nayani Ilangakoon for her expertise and help; to Dr. Yi Qi for his expertise and good judgement; and to Dr. Nancy Glenn for her welcome, guidance, and putting together an amazing team. It's a privilege and joy to work with all of you.

This work is supported by the National Aeronautics and Space Administration Terrestrial Ecology Grant NNX14AD81G, the Graduate Research Assistant Mentorship program, the Research Institute for Environment, Energy, Economics (RIEEE), and Appalachian State University's Department of Geography, as well as through contributions

by Mr. Steve Vacendak in honor of his son Stephen's passion for environmental science.  
Their generosity supports our research.

Finally, thank you to my family and friends. This is only possible through your  
unfaltering support and encouragement. I will always be grateful for you.

## Table of Contents

Abstract .....	iv
Acknowledgments .....	vi
Abbreviations .....	x
Foreword .....	xi
1. Introduction and background.....	1
1.2 Semi-arid rangelands in the Great Basin .....	4
1.3 Sagebrush structure and foliar nitrogen.....	5
1.4 Remote sensing.....	6
1.4.1 Imaging spectroscopy .....	7
1.4.2 The influence of canopy structure on foliar Nitrogen detection.....	10
2. Materials and methods.....	16
2.1 Site selection.....	16
2.2 Field and remote sensing data collection.....	18
2.2.1 Vegetation sampling .....	18
2.2.2 Sagebrush biomass allometry .....	19
2.2.3 Leaf area estimation .....	19
2.2.4 Plot level scaling of nitrogen content .....	20
2.3 Remote sensing data .....	23
2.3.1 Imaging spectroscopy pre-processing .....	24
2.4. Data Analysis .....	25
2.4.1 Partial least squares regression .....	25
2.5 Nitrogen mapping for Reynolds Creek Experimental Watershed .....	26
3 Results .....	26

3.1 Field data collection results .....	26
3.1.2 Biomass allometry .....	30
3.2 Partial least squares regression .....	33
3.3 Plot level scaling of foliar N content .....	39
3.4 Spectral transformation .....	41
3.5 Wavelength selection using model repetitions .....	43
3.6 SLA and Cover results.....	44
3.7 Map of predicted foliar N values.....	46
4. Discussion .....	48
4.1 Partial least squares regression.....	48
4.2 Field collection discussion .....	52
4.3 Transformations and subsetting.....	54
4.4 Mapping predicted foliar N .....	54
5. Conclusions .....	55
References .....	57
Appendices .....	66
Appendix A. Leaf area estimation.....	66
Appendix B. Leaf area code .....	70
Appendix C. Lidar fusion.....	75
Appendix D. Code for foliar N prediction using PLSR coefficients.....	78
Appendix E. Bands from Curran (1989) .....	80
Appendix F. Annotated graphs of model coefficients along the spectrum.....	81
Vita.....	83

## Abbreviations

ACCP: Accelerated Canopy Chemistry Program  
AGB: Above ground biomass, the combined woody and green biomass of destructively sampled shrubs, provided in grams  
AIS : Airborne Imaging Spectrometer  
AVIRIS-NG: Airborne Visible/InfraRed Imaging Spectrometer (AVIRIS) Next Generation instrument  
CA: Crown area, calculated from the major and perpendicular minor canopy widths  
CO<sub>2</sub>: Carbon dioxide  
CVHt: Crown volume, calculated from crown area and height  
FIAT: Fire and Invasive Assessment Tool  
GB: Green biomass; the leaves and stems of destructively sampled shrubs, provided in grams  
LAI: Leaf area index  
Landsat: Sensor providing long-term moderate-resolution land remote sensing data  
Lidar: Light Detection and Ranging  
LMA: Leaf mass per unit area, the dry weight of leaves divided by the area of those leaves, provided in grams per square centimeter  
MODIS: Moderate Resolution Imaging Spectroradiometer sensor  
N: Nitrogen  
NIFC: National Interagency Fire Center  
NIR: Near infrared spectral range from 700-1300 nm  
NPP: Net primary productivity  
NPV: Non-photosynthetic vegetation such as woody biomass and leaf litter  
PLSR: Partial least squares regression  
RCEW: Reynolds Creek Experimental Watershed in Idaho, USA; the collection site for this study.  
Red edge: Range of the spectral from 800-850 nm where vegetative signals show a sharp increase in reflectance  
RMSE: Root mean square error  
SG: Savitzky-Golay; a smoothing algorithm used in this study to reduce signal noise by applying a first derivative transformation to reflectance measurement using a polynomial fitting and variable window for its moving average  
SLA: Specific leaf area, the inverse of LMA  
SNV: Standard Normal Variate; a function that transforms reflectance measurements to normalize readings across plots  
SOC: Soil organic carbon  
SWIR: Short wave infrared 1500-2500 nm  
UV: Ultra violet  
VIS: Visible range of the spectrum from 400-700nm in which chlorophyll absorbance features occur  
WB: Woody biomass; the branches and trunks of destructively sampled shrubs, provided in grams

## Foreword

Chapters 1 through 6 of this thesis will be submitted to *Ecological Indicators*, a peer-reviewed journal. These chapters and their references are formatted according to the style guide of the journal. The introduction and conclusions of the thesis are incorporated into the article.

## **1. Introduction and background**

Dryland ecosystems, which include semi-arid lands in the Western US, cover 41% of Earth's terrestrial surface, yet their sensitivity and ability to drive climate change are poorly understood (Adeel and World Resources Institute, 2005; Ahlström et al., 2015; United Nations Environment Management Group, 2011). Globally, these lands house 35.5% of our growing human population, 90% of whom live in developing countries and may be particularly vulnerable to environmental changes and dependent on ecosystem services (Adeel and World Resources Institute, 2005; Gilbert, 2011; United Nations Environment Management Group, 2011). Drylands are habitat for important species, many of which are endangered, listed, or endemic; these species provide for pollination, ecotourism and recreation activity, forage for grazing, pharmaceuticals and medicinal research, and genetic resources that are important to adaptation and survival in a changing climate (United Nations Environment Management Group, 2011). This biodiversity supports essential ecosystem services such as provisioning of food and rangeland, erosion control and soil genesis, carbon fixation, water purification and provisioning, recreation, and cultural resources (United Nations Environment Management Group, 2011). Interconnected pressures with poorly understood feedback interactions are resulting in land degradation and the subsequent limitation of sustainable ecosystem services. To maintain these ecosystem services, research is needed to quantify and monitor how drylands respond to and drive environmental variables.

Drylands store roughly 15% of Earth's total soil organic carbon (SOC) due to their large size and high SOC pools (Noojipady et al., 2015). This is twice the organic carbon of forest ecosystems despite the sparse vegetation and low carbon sequestration rates of

drylands (sequestration is estimated at  $703 \pm 44$  g/m<sup>2</sup> for drylands as opposed to  $869 \pm 34$  g/m<sup>2</sup> for forests and woodlands in one global model derived from satellite observations) (Cao et al., 2004; Noojipady et al., 2015; Safriel et al., 2006). New research is finding that drylands exert major influence on the interannual variability of the global carbon sink (Ahlström et al., 2015; Poulter et al., 2014). This influence is caused by their large area and sensitivity to rainfall, temperature, and disturbances, which alter dryland ecosystem net primary productivity (NPP), carbon fixation, and emission (Ahlström et al., 2015; Poulter et al., 2014). For example, increases in precipitation cause greater plant growth and carbon fixation and suppress fire; drier conditions reduce plant growth and carbon fixation and make fire related emissions more likely (Poulter et al., 2014). Dryland ecosystems are historically subject to natural disturbance regimes including wildfire and drought (Field et al., 2014) and vulnerable to desertification and wind erosion; between 10-20% are estimated to already be degraded by desertification and a larger amount is at risk (Adeel and World Resources Institute, 2005). Additional changes which alter carbon emissions and fixation rates include anthropogenic climate change and associated disturbances in rainfall patterns, increasing atmospheric carbon dioxide (CO<sub>2</sub>) concentrations, and changes in temperature averages and extremes (Ahlström et al., 2015; Field et al., 2014). Direct land use pressures and development also alter carbon emissions and fixation rates; for example, conversion to grazing and crops, suppression of regional rainfall through changes in albedo and evapotranspiration from altering surfaces, and increases in sources of fire ignition (Field et al., 2014). All of these factors are expected to exacerbate pressure on ecosystem services and increase the rate of habitat shift and loss and the frequency and intensity of wildfires, droughts, floods, and erosion (Adeel and World Resources Institute, 2005; Field et al., 2014).

Changes in vegetative growth and wildfires are two important processes to consider when examining the interaction of drylands and climate. Recent literature hypothesized that observed "greening" of drylands was attributable to rising atmospheric carbon dioxide (CO<sub>2</sub>) allowing more efficient carbon fixation (Poulter et al., 2014). This allows plants to keep their stomata closed more often and retain more water, reserving soil water levels without decreasing productivity (Donahue et al., 2013). In water limited biomes like drylands, increased soil water, an effect of efficient water use due to carbon fertilization, is hypothesized to increase vegetative growth (Donahue et al., 2013). In contrast, in situ studies have shown that small increases in temperatures in dryland sites, such as those expected to occur as CO<sub>2</sub> rises, may reduce soil water and limit growth and productivity, possibly also favoring certain species and affecting invasion patterns of non-natives (Wertin et al., 2017). Vegetative presence and structure affects hydrology as well as resistance to wind erosion in drylands, creating of microclimates with lower temperatures and evaporative rates under the canopy (Breshears and Nyhan, 1998). Woody vegetation may create heterogeneous aeolian transport of sediments which changes patterns of nutrient distribution and may drive land degradation and desertification (Mueller et al., 2007; Okin, 2008; Sankey et al., 2012). A changing climate can affect species composition and productivity, which in turn drives climate change through CO<sub>2</sub> emission rates.

Fire, in turn, increases wind erosion but allows redistribution of soil nutrients (Ravi et al., 2009). Climate change driven variation in rainfall patterns affects species distribution and abundance by altering seasonal water balance (Field et al., 2014). Increasing temperatures may extend the growing season or shorten it by changing water availability and evaporative demand (Field et al., 2014). Together, changing climate and vegetation affects nitrogen

mineralization (Bobbink et al., 2010), soil structure, and soil microbial communities (Field et al., 2014; Hu et al., 2016). Despite variation in greening and water retention, globally the changes in climate and land use have resulted in overall increased aridity, drought, and wind erosion (Field et al., 2014).

In short, relationships that describe how dryland ecosystems drive and react to climate change are complex and interactive. While vegetation is relatively sparse compared to other ecosystems, the immense size of these ecosystems and our heavy reliance on the services they provide make these feedbacks important to quantify when evaluating management choices. High resolution synoptic mapping will improve our understanding of complex multi-scale processes and inform management decisions.

## **1.2 Semi-arid rangelands in the Great Basin**

The Great Basin region in the western US provides a representative, national study site to examine global dryland dynamics related to desertification, land use change, grazing pressure, biological invasion, fire regimes, nutrient cycling, and climate change interactions. Sagebrush-dominant communities are characteristic of semiarid landscapes in the Western US (West and Young, 2000) and are part of a larger global distribution of dryland systems that collectively exert significant influence on interannual variability in carbon stocks (Ahlström et al., 2015). They may act as a substantial carbon sink on their own in years of vegetative growth (Svejcar et al., 2008) or substantial sources of carbon in years of vegetative loss (Poulter et al., 2014). Sagebrush and other woody plants drive rates of evapotranspiration, erosion (Breshears and Nyhan, 1998; Prater et al., 2006), and carbon and nutrient cycling (Breshears et al., 2006, Yang et al., 2012) for arid and semi-arid regions. These processes indirectly drive forage production, habitat quality, and aforementioned

ecosystems services, all of which will be further impacted by climate change (Polley et al., 2013). The variation between locations in species composition, environmental variables, and disturbance events (Hasselquist et al., 2011) illuminates the need for real time monitoring and frequent re-sampling over these prohibitively large and difficult to access areas to quantify and monitor climate change impacts and feedbacks over seasonal and annual changes (Olsoy et al., 2016).

Finally, there is recent public acknowledgement of the importance of geospatial science applications for semi-arid systems in the US. A mandate was issued to develop research to protect and restore these areas to maintain the ecosystem services they provide, including control of extreme megafires (Barret et al., 2016). The Department of the Interior Secretary Sally Jewell, through Secretarial Order 3336 (S.O. 3336), "Rangeland Fire Prevention, Management, and Restoration," created task forces to ensure to science-based policies and strategies would be made available to the public for use in fire management, restoration, and conservation (Barret et al., 2016). The order emphasizes the development of geospatial science and distribution through mobile devices in coordination with the National Interagency Fire Center's (NIFC) Fire and Invasive Assessment Tool (FIAT) to support and speed up management decisions to monitor resource conditions and reduce likelihood of habitat loss in fire events (Barret et al., 2016).

### **1.3 Sagebrush structure and foliar nitrogen**

Foliar nitrogen (N) influences the quantity and quality of forage (Frye et al., 2013; Skidmore et al., 2010). Climate pressures such as elevated ultra violet (UV) light, temperature, and drought may alter the dietary quality, palatability, and amount of forage available to herbivores by influencing the amount of constituents such as crude protein

(mostly N) and defensive chemicals, such as monoterpenes and phenolics (Forbey et al., 2013). Remote sensing of foliar N can therefore be used to infer and monitor rangeland health and herbivory patterns (Forbey et al., 2013). Foliar N data can also be applied to studies of atmospheric N deposition, productivity patterns, and invasive species invasion patterns (Ollinger, 2011).

Additionally, foliar N is used to monitor ecosystem processes because it is a critical growth element in plants due to its role in carbon fixation during photosynthesis (Field and Mooney, 1986; Lepine et al., 2016; Mattson, 1980). Dryland ecosystems are co-limited by water and N (Hooper and Johnson, 1999; Sinsabaugh et al., 2015), and absence of N limits the advantages of increasing atmospheric CO<sub>2</sub> for net carbon storage (Moorcraft, 2006). Photosynthetic capacity scales with the investment of N in foliar biochemical compounds and components central to photosynthetic function such as thylakoids, chlorophyll, and soluble proteins (Evans, 1989; Field and Mooney, 1986; Wright et al., 2004). Foliar N is thus an important indicator of ecosystem metabolism and health as it relates to net primary productivity, rate of photosynthetic capacity, and light use efficiency (Lepine et al., 2016; Ollinger and Smith, 2005; Ollinger et al., 2008). Foliar N also influences availability of N in soils and carbon assimilation through influencing litter decay, competition between plants and microbial communities, net mineralization, and plant N uptake mechanisms (Hu et al., 2016; Lepine et al., 2016).

#### **1.4 Remote sensing**

Remote sensing is a cost effective alternative that provides rapid quantitative mapping for regions in which large-scale sampling is prohibitively expensive and labor-intensive (Hunt et al., 2003). Remote sensing refers to the use of sensors without direct contact

with the target, such as when data are recorded at a distance using a terrestrial (e.g. handheld), aerial (plane or unmanned aerial vehicle mounted), or satellite platform (Campbell and Wynne, 2011). Remote sensing data such as imaging spectroscopy can be used to passively measure biochemical content, such as foliar N, while lidar (Light Detection and Ranging) uses the active emission of laser beams and can measure vegetative structure, including biomass, across a landscape. As foliar N and vegetative structure are important indicators of ecosystem health, remotely monitoring these variables supports carbon flux research and land management decision-making.

#### **1.4.1 Imaging spectroscopy**

Imaging spectroscopy refers to the analysis of wavelength intervals along the electromagnetic spectrum to infer information about a target. Imaging spectroscopy data can be used to detect and map vegetation through proxy indicators such as N content. As shown in Figure 1, vegetation commonly displays spectral features which can be used for identification and characterization. These features include high reflectance in the near infrared (NIR) range and an extreme slope around 700 nm which is referred to as the “red edge” and caused by leaf structures that transmit non-photosynthetic wavelengths to avoid overheating (Curran, 1989). Additionally, water-related absorption points in vegetation occur at 970, 1200, 1400, and 1940 nm, and proximity of a band measurement to a water absorption range will increase measurement error (Curran, 1989). Chlorophyll absorption points for photosynthesis occur in the visible range of the spectrum (VIS) of red and blue light from 400 to 700 nm (Curran, 1989).

Drylands present particular challenges to remote sensing methods developed for other biomes due to their sparse vegetation, high albedo soils, and the nonlinear mixing that occurs

between these and other background materials such as rock and grasses (Borel and Gerstl, 1994; Ray and Murray, 1996). The reflectance signal from sagebrush shrubs is relatively weak and contains a high degree of mixing as cover decreases (Borel and Gerstl, 1994; Ray and Murray, 1996). As an illustration, note that in Figure 1, an imaging spectroscopy pixel that is associated with low sagebrush cover (8%) does not display chlorophyll absorption or NIR reflectance as strongly as a pixel associated with high sagebrush cover (58%). Open shrub canopies with low percent cover mix with bright soil reflectance that dominates the background. Plant evolutionary adaptations to dryland light and heat conditions, such as small or vertically oriented leaves that reduce sun exposure, gray trichomes (leaf hairs) that trap moist air near the leaf surface, or reflective waxy coatings, influence expected vegetation spectral responses, such as leaf pigment absorptions in portions of the visible region of the electromagnetic spectrum and a strong "red edge" in the near infrared, and are present in other spectral features such as wax absorbance at 1720 nm (Mitchell et al., 2012b; Okin et al., 2001). Discrimination of non-photosynthetic vegetation (NPV) in drylands appears spectrally similar to soil and additionally complicates interpretation of spectral signals (Roberts et al., 1993). These difficulties impede direct mapping of vegetation.

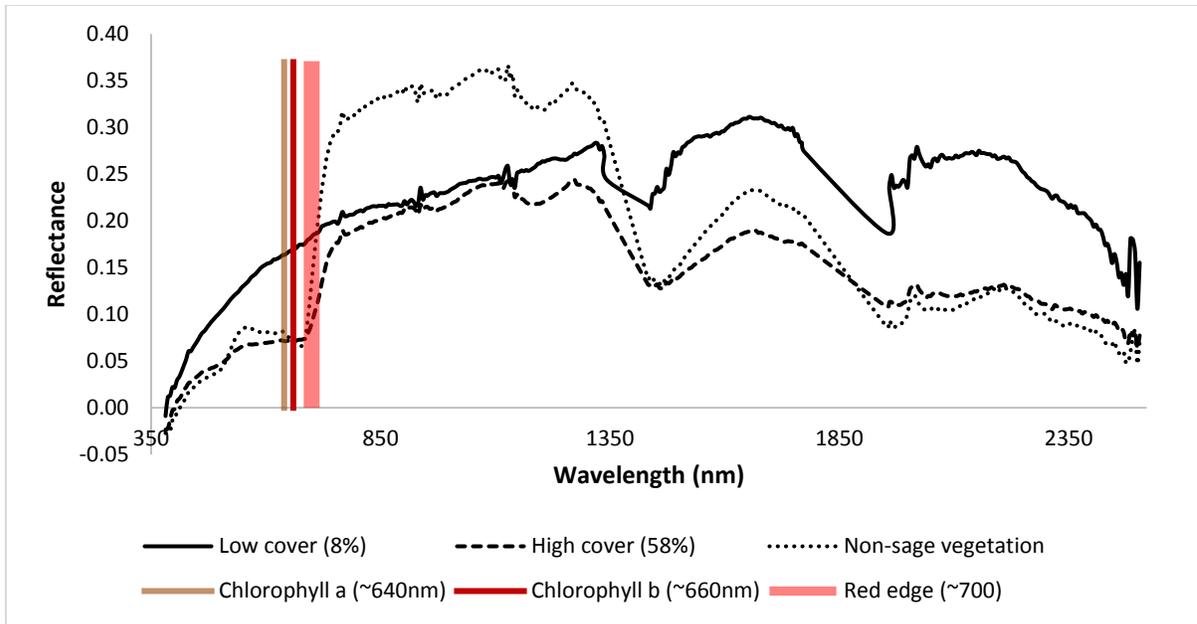


Figure 1. Reflectance signatures for imaging spectroscopy pixels extracted from low and high sagebrush canopy cover are graphed beside a nearly pure vegetation pixel with characteristic vegetation features, such as a strong red edge near 700 nm and chlorophyll absorbance in the visible region.

The process of relating biochemical characteristics to remote sensing data using imaging spectroscopy is an ongoing field of study that became established with NASA’s Accelerated Canopy Chemistry Program (ACCP) in 1991–1992 (ACCP, 2004; Ustin, 2013). Imaging spectroscopy is used to identify and quantify vegetative biochemistry such as pigment, water, nitrogen, and carbon (Asner et al., 2007; Ustin et al., 2004). Nitrogen content is indicated by the shape and depth of the signal at wavelengths: 1020, 1510, 1980, 2060, 2130, 2180, and 2300 nm (Curran, 1989). Near infrared reflectance spectroscopy has been used successfully to measure nitrogen-based crude protein estimates in agriculture and food science generally and specifically for sagebrush in a lab setting by Olsoy et al. (2016), using wavelengths 1000-2500 nm.

Imaging spectroscopy is also used to estimate canopy chemistry such as foliar N and N-containing chlorophyll and proteins (Asner and Martin, 2008; Lepine et al., 2016; Martin et al., 2008; Ollinger and Smith, 2005; Ollinger et al., 2008; Ustin et al., 2013; Wang et al., 2016). Detecting these signatures ideally requires narrow 10 nm intervals and a signal to noise ratio high enough to distinguish target features (Curran, 1989). This is provided in imaging spectroscopy sensors such as the Airborne Visible/InfraRed Imaging Spectrometer (AVIRIS) Next Generation instrument (Chrien et al., 1990). Imaging spectroscopy records a wide range (400-2500 nm) of wavelengths in near-continuous narrow (10-20 nm) band intervals (Curran, 1989).

The imaging spectroscopy sensor AVIRIS was used in detection of N in early efforts by Martin and Aber (1997) which related AVIRIS data to field measurements of foliar N first-difference reflectance bands at 950 and 2290 nm. Recently, aggregated AVIRIS data were found to improve detection of canopy N content compared to satellite platforms such as MODIS and Landsat (Lepine et al., 2016). Asner and Lobell (2000) found that modeling using the 2100-2400 nm range successfully mapped cover and distinguished between signatures of bright, reflective soils, green canopy, and NPV in arid regions. The narrow bandwidth of imaging spectrometers was identified as particularly necessary to discriminate between soil and vegetation (Borel and Gerstl, 1994).

#### **1.4.2 The influence of canopy structure on foliar Nitrogen detection**

Investigating imaging spectroscopy responses associated with foliar N addresses a current discussion in the literature. The depths and shapes of absorption features are highly correlated with foliar chemistry and caused by vibrations of chemical bonds and overtones which overlap and interact (Kokaly and Clark, 1999). Different wavelength regions contain

absorption features that are indicative of foliar N content and associated with constituents such as proteins in chlorophyll (Curran, 1989).

Past research has drawn attention to the strong correlation between laboratory-measurements of foliar N in specimens collected from the field and variations in NIR (700-1300 nm) reflectance obtained from airborne platforms (e.g., Wessman et al., 1988; Martin et al., 2008). Airborne studies in a range of ecosystems have identified high reflectance in wavelengths in the NIR region of the electromagnetic spectrum to be strong predictors of foliar N (e.g., Asner et al., 2008; Martin et al., 2008; Ollinger et al., 2008). Strong reflectance in NIR is presumably an indicator of vegetation as foliar structural characteristics that scatter these wavelengths; NIR wavelengths are not used in photosynthesis, and it is advantageous for leaves to reflect or transmit them to avoid overheating (Curran, 1989; Knyazikhin et al., 2013). In other words, spectral response in the NIR region is largely driven by plant structure rather than physical mechanisms such as known absorption features such as those associated with, for example, protein bonds (Kokaly and Clark, 1999). The extent to which foliar N has a direct physical connection to plant structure is debatable (Knyazikhin et al., 2013; Ollinger et al., 2013; Townsend et al., 2013; Ustin, 2013) and studies designed to decouple the two variables are limited (Knyazikhin et al., 2013; Latorre-Carmona et al., 2013). Knyazikhin et al. (2013) isolated structural variables from a subset of the data used in Ollinger et al. (2008) and concluded that the distribution of canopy gaps, not N absorption, drove the correlation of foliar N content and NIR reflectance, and that subtracting structural effects removed correlation of foliar N content to high NIR reflectance. Knyazikhin et al. (2013) suggested that without a mechanism relating foliar N content to structure, the predictive relationship is not necessarily reliable and that confounding effects associated with structure should be

removed before analysis. Ollinger et al. (2013) responded that the strong correlation which emerged across biomes in Ollinger et al. (2008) likely indicated that the importance of N in plant function, particularly in N limited ecosystems, drives structural phenology, and that reflectance may be indirectly measuring a consistent relationship of N that covaries with plant functional type.

In systems with sparse canopies, such as sagebrush-dominated arid and semi-arid grasslands with low stem density and lower LAI, NIR may contain important wavelengths that detect structure. By extension these may be used to predict foliar N content, assuming physical mechanisms relate foliar N to plant functions that determine structure, such as the influence of metabolism on growth. However, the relatively low stature and sparse canopy structure of dryland vegetation, including interspersed NPV, may complicate NIR detection of structure. Plants like sagebrush and bitterbrush have leaves that are narrow, small, and vertically oriented. This may alter structural correlation between foliar N and NIR observed in other ecosystems. Conversely, these plants may display higher albedo due to light colored trichomes, creating similar or alternate structural signals. While N limitations and relationship to growth and structure of sagebrush and other dryland vegetation may provide a predictive relationship, it is unlikely to be the same relationship defined in forest areas. In forests, some of the defined differences have been broad high N deciduous leaves with high reflectance versus vertical and narrow low N pine needles with lower reflectance over continuous cover. In drylands the differences may be between bright high reflectance zero N bare ground and litter versus bright light colored vertical and narrow low N sagebrush leaves. Consequently, inclusion of VIS in analysis may boost interpretation due to the correlation between N and chlorophyll (Lepine et al., 2016). Additionally, short wave infrared (SWIR,

1500-2500 nm) reflectance is useful to detect N bonds (Curran, 1989). Mitchell et al. (2012b) identified predictive wavelengths associated with  $\beta$ -carotene and chlorophyll features in the VIS region and shifted N absorption features present throughout the infrared region.

The study presented herein follows previous imaging spectroscopy research methods in dryland ecosystems which successfully use partial least squares regression (PLSR) to identify potential biochemical absorption features and structural patterns from dominant dryland species to investigate relationships between reflectance signals and foliar N (Kokaly et al., 2003; Mitchell et al., 2012a; Mitchell et al., 2012b; Yi et al., in review; Dashti et al., in review). In order to address remote sensing challenges associated with sparse canopy structure and high degrees of spectral mixing, which may present plant functional types and indirect relationships and spectral features that are different from ecosystems in other studies (Townsend et al., 2013), the full narrowband spectrum from 400 to 2500 nm is retained for analysis. This includes chlorophyll contributions in the VIS region (Mitchell et al., 2012a; Mitchell et al., 2012b; Lepine et al., 2016), NIR/SWIR regions (>1,100 nm) where leaf level foliar nutrients are captured (Asner et al., 2011; Kokaly et al., 2009; Ollinger, 2011) as well as other N absorbance features in longer wavelengths (Curran, 1989).

This thesis aims to more efficiently address challenges associated with remote sensing of vegetative characteristics in dryland ecosystems by identifying an optimal field data collection method for scaling foliar N content from the leaf to plot scale for high resolution airborne N mapping at the watershed scale. Unique remote sensing challenges include nonlinear mixing and detection of low stature desert adapted vegetation with open canopies against a high albedo background. Foliar N tends to have a strong signal and can be a useful indicator of ecosystem services, such as forage quality for conservation and

rangeland, wildfire mitigation, and carbon sequestration (Forbey et al., 2013; Frye et al., 2013; Ollinger, 2011; Skidmore et al., 2010). However, the structural variables that contribute to complex scattering are numerous and include low open canopy mixing with bare ground and litter. Vegetative adaptations to dryland ecosystems at leaf and canopy levels are very different from other ecosystems, and structural variation is expected to strongly influence spectral response (Asner, 1998; Borel and Gerstl, 1994; Knyazikhin et al., 2013; Mitchell et al., 2015; Ray and Murray, 1996).

This work ranked four methods for scaling ground reference samples from shrub level to plot level in order to identify which field data variables are most useful to inform remote sensing detection of foliar N and to identify the best performing scaling method. The selected method was used to develop PLSR models that rely on important wavelength ranges to predict foliar N content. The best performing model coefficients were used to map vegetative foliar N across the Reynolds Creek Experimental Watershed study area.

Comparing the model performance of different scaling methods can inform researchers and land managers on which vegetation characteristics to sample in the field to remotely predict sagebrush foliar N. Narrowing down which field variables are necessary will minimize the field time and resources needed to parameterize models. Scaling methods that include variables for deriving more precise estimates of shrub volume on a mass basis (e.g., biomass, specific leaf area (SLA), shrub density, LAI) are expected to outperform scaling methods based on coarser estimates of shrub volume (e.g., shrub cover and height)

In addition to identify optimal scaling methods, PLSR model performance can also compare the relative importance of different remote sensing wavelengths regions on foliar N prediction. It is anticipated that wavelengths in the VIS associated with chlorophyll pigments,

wavelengths in NIR associated with the red edge, and wavelengths in the NIR/SWIR associated with known N absorption features will be consistently selected by the model as significantly correlated to N (specific wavelengths identified by Curran, 1989, are listed in the appendices). It is also expected that the selection of important wavelength predictors will be sensitive to the type of transformation that are applied to the reflectance spectra prior to analysis (e.g., log transform, first derivative).

Models using PLSR generate regression coefficients that can be applied to imaging spectroscopy raster data to predictively estimate foliar N estimates across the watershed. Quantifiable uncertainty (RMSE) can be reported for the mapping. The high resolution N mapping generated as part of this study is expected to consistently and accurately identify features on the landscape associated with high foliar N content (e.g., riparian corridors). As land cover types diverge and fine scale heterogeneity increases with decreased shrub cover and the mixing of grasses, bare ground, and litter, N estimations are expected to become more uncertain and less reliable.

This work extends N prediction methods by testing various combinations of field data that represent structural variables, evaluating them for their ability to successfully scale leaf level N to airborne imaging spectroscopy data. Results were used to predictively map foliar N across a watershed with the goal of identifying spectral regions that are strong predictors of foliar N at the airborne scale for use in the context of regional upscaling.

## 2. Materials and methods

### 2.1 Site selection

Data were collected at the Reynolds Creek Experimental Watershed (RCEW) (Figure 2), a 238 km<sup>2</sup> cold semi-arid desert sagebrush-steppe ecosystem in the Owyhee Mountains of southern Idaho, USA with significant climate, elevation, and precipitation gradients (Li et al., 2015). The average air temperature is 11.7° C and varies from between 1.4° C minimum to 16.2° C maximum (WRCC, 2009). Elevation ranges from 1049 to 2245 meters. Mean annual precipitation varies from ~250 mm in the north to >1100 mm in the southwest (Marks et al., 2007). Vegetative cover is predominantly grassland and sagebrush species such as low sagebrush (*Artemisia arbuscula* Nutt.), big sagebrush (*Artemisia tridentata* Nutt. subsp. *vaseyana* [Rydb.] Beetle and subsp. *wyomingensis*) and bitterbrush (*Purshia tridentata* [Pursh] DC), which respectively average 50 cm, 50–100 cm, and 60–185 cm in height (Mitchell et al., 2015). Twenty one field sampling locations within the study site were selected in which sagebrush species were the dominant vegetative cover in order to minimize potential mixing of signal responses and to create homogenous samples to correlate to spectral and lidar variation (Mitchell et al., 2015). Lidar and AVIRIS imagery coverage were obtained for the entire watershed for both 2014 and 2015 (Mitchell et al., 2015).

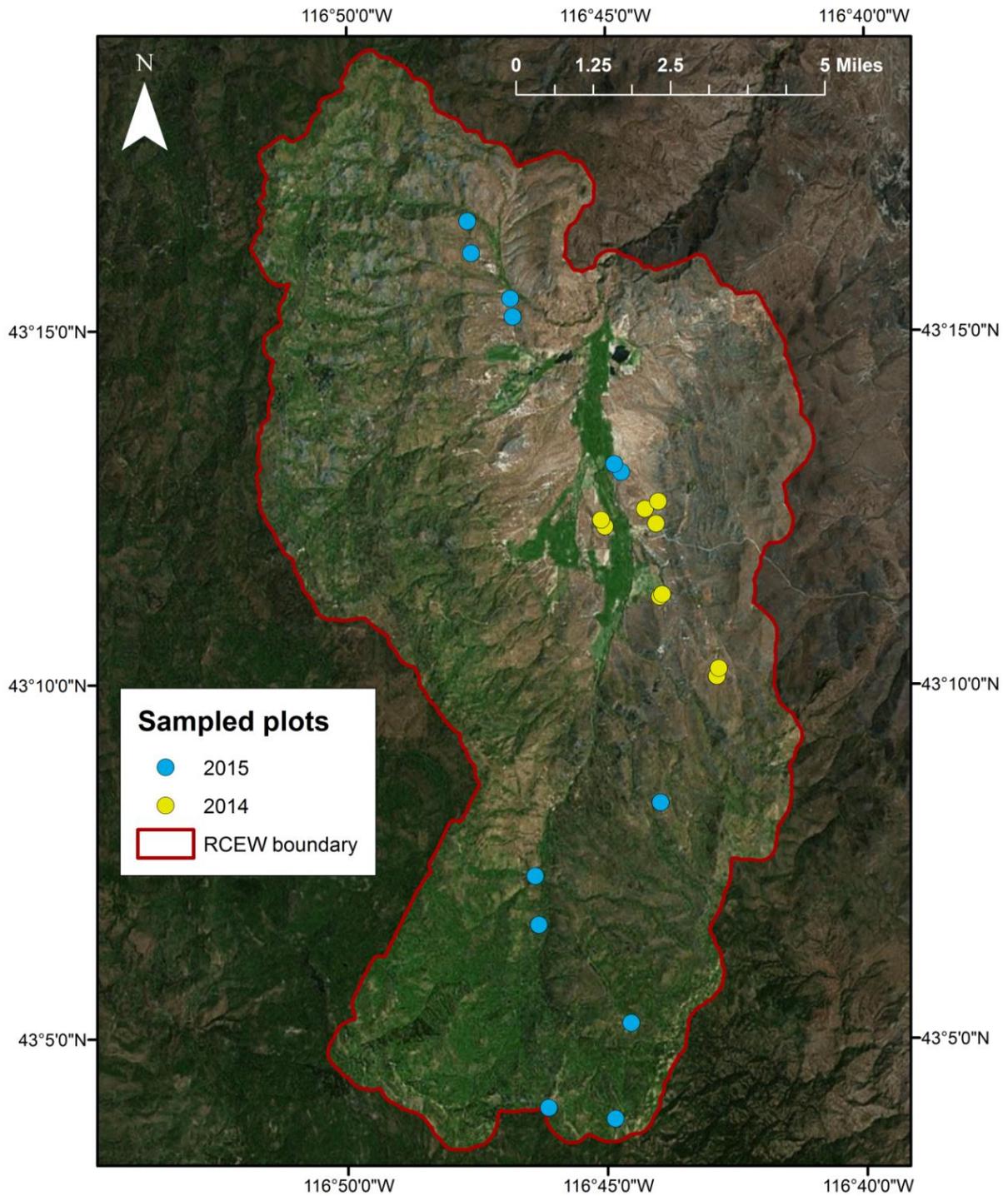


Figure 2. Reynolds Creek Experimental Watershed (RCEW) where airborne AVIRIS-NG and lidar data were acquired. Spectral data were acquired on September 14, 2014, and June 11, 2015. Twenty one 10 x 10 m plots were sampled, nine in 2014 from September 16<sup>th</sup>

through October 3<sup>rd</sup>, and twelve from May 11<sup>th</sup> through the 28<sup>th</sup>, 2015. Lidar were collected in 2015.

## **2.2 Field and remote sensing data collection**

### **2.2.1 Vegetation sampling**

Twenty-one 10 x 10 m plots were established wherein ground reference data were collected to develop scalable vegetation products. Nine plots were sampled in September, 2014 and 11 plots were sampled from May, 2015. Ground reference data variables included: LAI for selected shrubs; leaf sample dry weight, area, and percent N content; height and major and minor canopy widths measured from terrestrial lidar scans of selected shrubs; plot percent cover estimate; and plot density count of live sagebrush species individuals. LAI was collected for 20 shrubs in each plot, along transects set at 1, 3, 5, 7, and 9 meters, and at 2, 4, 6, and 8 meters down each transect.

Allometric biomass equations were developed for scaling sagebrush canopy measurements using a relationship between terrestrial laser scanning measurements and weights from destructive sampling conducted in summer and fall 2012 at RCEW, Idaho National Laboratory, and Hollister, Idaho (Olsoy et al., 2014).

Phenological differences may occur between the two collection years which were taken in different seasons. Sagebrush grow two crops of leaves: ephemeral leaves, which have larger surface area to take advantage of water available for evaporative cooling earlier in the season; and perennial leaves, which are smaller and more resilient to hot, dry conditions (Miller and Shultz, 1987). In our September collection and late summer collection of imaging spectroscopy data, ephemeral leaves may have dropped and more resilient, smaller leaves may have become dominant. Thus our sampling for May showed roughly

double the leaf area of the September leaf samples and as a result a higher LAI. In previous survey work, shrubs which shed ephemeral leaves during the summer drought lost two thirds of their leaf weight (Miller and Shultz, 1987).

### **2.2.2 Sagebrush biomass allometry**

A method using allometry of canopy was developed to investigate the use of terrestrial lidar in measurement and the use of canopy width and biomass estimates in scaling products. The predictor variables of canopy major and minor width and plant height were measured using RiScan software (RIEGL Laser Measurement Systems GmbH, Horn, Austria) with terrestrial lidar scans of the shrubs that were later destructively sampled to produce biomass measurements (Olsoy et al., 2015). The training dataset included three types of biomass measurements taken through destructive sampling: green biomass (GB) consisting of foliage, woody biomass (WB) consisting of stems and branches, and total aboveground biomass (AGB). These were used individually and aggregated into area and volume metrics as predictors. The predictors and independent data were linearly regressed and then back-transformed to express predictions in grams. The crown area method was based on the formula for elliptical area, using half of the major and minor canopy width measurements for inputs. The crown volume method additionally incorporated height using an ellipsoid volume estimate.

### **2.2.3 Leaf area estimation**

Leaf mass area (LMA), the dry weight of the leaf per unit of leaf surface, was used in three of the scaling methods tested. Leaf samples from six randomly selected shrubs of representative size were collected and weighed twice, first fresh to determine wet weight, and then after drying. Prior to drying, the leaf samples were scanned in groups and processed for

area estimates using pixel selection. Selections were counted twice and averaged using an automated technique developed by the author (Appendix A). Multiplying the selected leaf pixels by DPI provided leaf area metrics for use in scaling methods.

#### **2.2.4 Plot level scaling of nitrogen content**

Four methods were tested. Each method estimated plot level foliar nitrogen by combining ground reference data. The four methods included: a green biomass and density method based on Cleary et al. (2008); a leaf area index (LAI; a characterization of canopy density) and species density method based on Smith et al. (2001); a percent species cover and specific area method based on Serbin et al. (2014); and a percent species cover and plant height method based on Serrano and Ustin (2002).

The first method used the allometric equations derived from the Olsoy TLS dataset in combination with leaf area and weight measurements to first estimate total sagebrush leaf area per plot. The percent nitrogen was then combined to estimate foliar nitrogen per plot in grams. Three density measurements were tested: all shrubs, live shrubs only, and live sagebrush only.

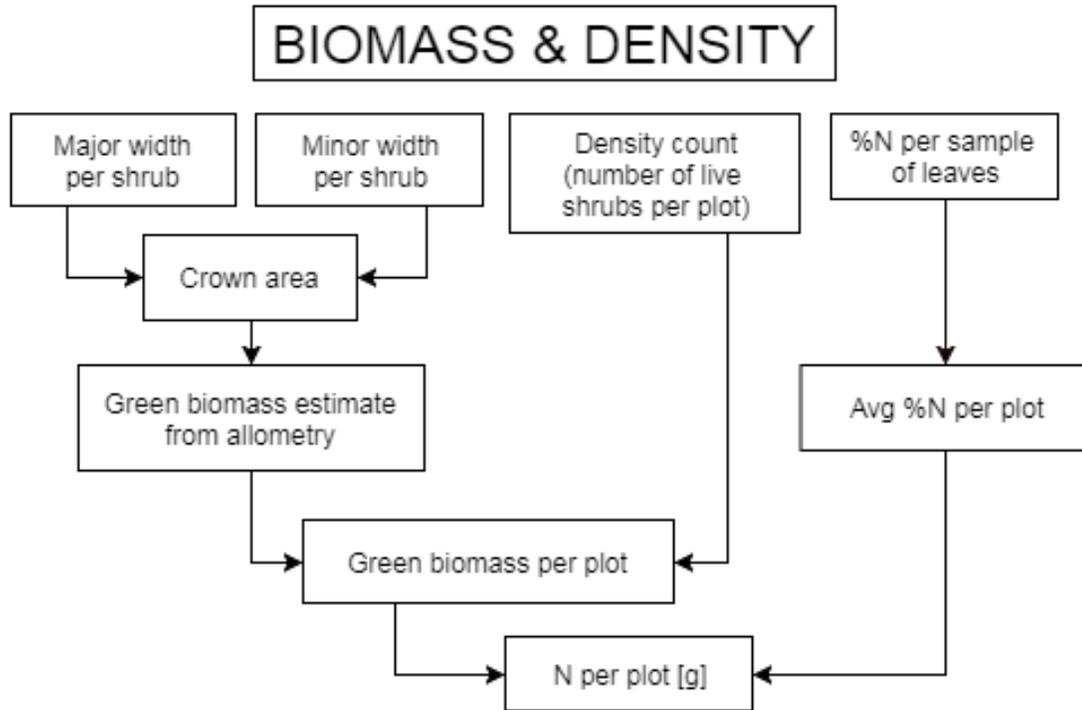


Figure 3. The method based on Cleary et al. (2008) uses canopy measurements and density (number of shrubs per plot).

The second method (LAI/Density/LMA) depended on a plot level density count and LAI. All live shrubs and all live sagebrush shrubs were tested. Density was adjusted by the LAI per shrub and averaged to plot. This estimate was adjusted by plot specific leaf nitrogen to provide foliar nitrogen in grams per square meter.

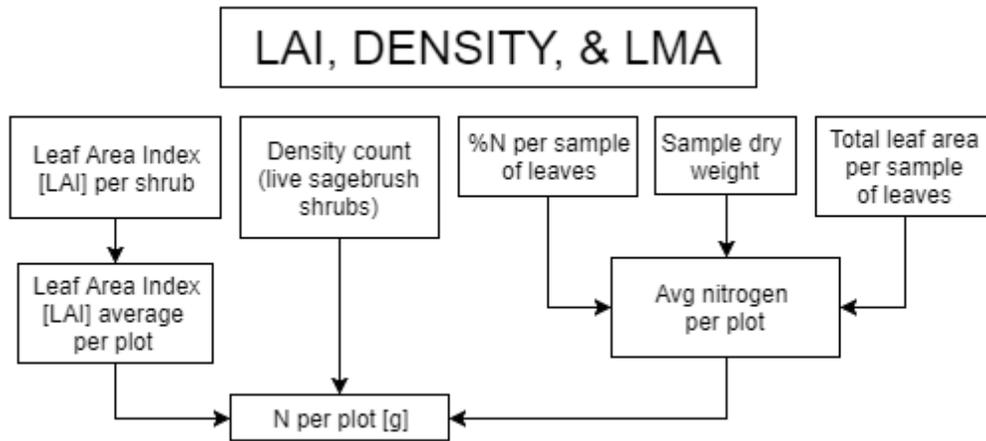


Figure 4. The method based on Smith et al. (2001) uses LAI and density count.

The third method (Cover/LMA) used a percent cover measurement determined by counting the proportion of transect measurements which cross above sagebrush canopy. This plot level cover measurement was adjusted by weight, leaf area, and then foliar nitrogen to determine grams of nitrogen per square meter for the plot.

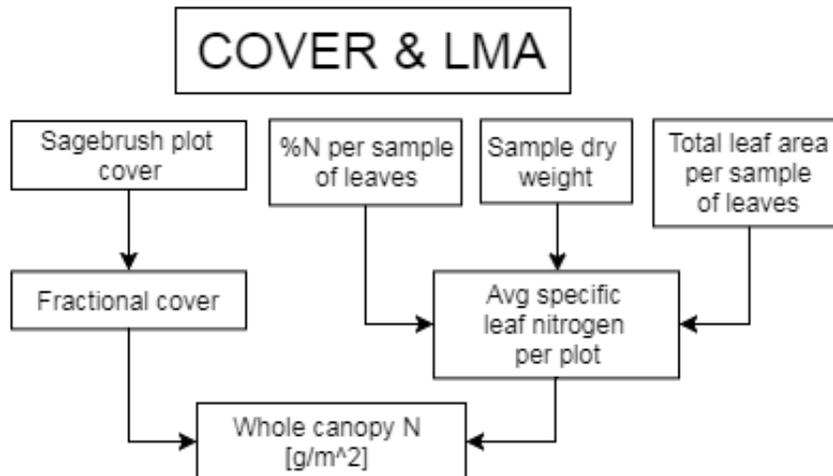


Figure 5. The method based on Serbin et al. (2014) uses percent cover to scale specific leaf area.

The fourth method (Height/Cover) uses height as a proxy for LAI measurements. Combined with percent foliar nitrogen and plot level sagebrush cover, this method estimates a concentration of nitrogen expected in the plot canopy.

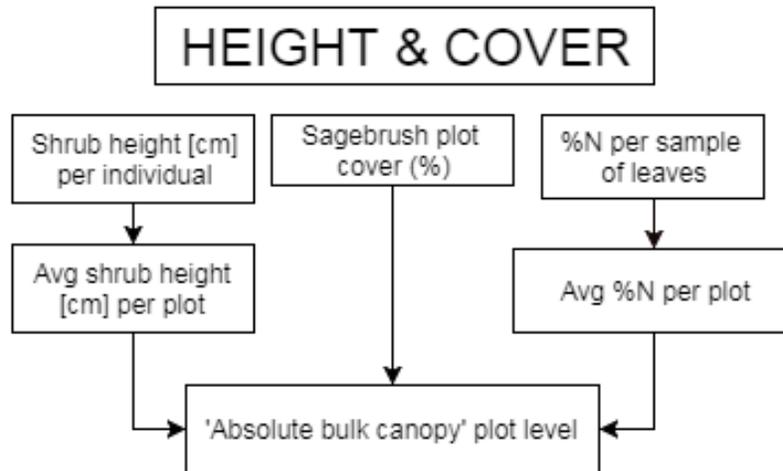


Figure 6. The method based on Serrano and Ustin (2002) uses the percent species cover and plant height.

### 2.3 Remote sensing data

Imaging spectroscopy data were obtained from the AVIRIS-NG instrument at an aerial level for the entire RCEW site in 2014 and 2015. The AVIRIS instrument was flown in a Twin Otter aircraft. Data in 2014 were collected in September, which was selected to be advantageous for model parameterization as grasses have senesced by this time and to complement the timing of existing datasets. Data were later collected in June, 2015, which was selected to explore the advantage of sampling while ephemeral leaves were present. Atmospheric corrections were applied, as recommended in previous work (Martin and Aber, 1997).

Imaging spectroscopy and lidar data were extracted using polygonal site boundaries. The data were imported into ENVI imagery analysis software (Exelis Visual Information Solutions, ENVI/IDL 5.2.1, Boulder, Colorado) and compared to imported vector data of site boundaries. Values encompassed in the site boundaries were extracted and averaged. If multiple flight lines overlay the same plot, those values were also averaged. These values were exported for use in the PLSR analysis.

### **2.3.1 Imaging spectroscopy pre-processing**

Imaging spectroscopy data were pre-processed to remove bad bands. Bands were examined visually in ENVI using band animation, and bands with notable distortions were excluded from analysis. Removed bands included: 381-416, 722, 907, 947, 1123, 1228, 1343, 1453, 1954, 1999, 2004-2024, and 2470-2505 nm.

The imaging spectroscopy values were transformed prior to the PLSR analysis. Analysis was attempted on original reflectance data, as well as data transformed to absorbance and then transformed with either a gap derivative or Savitzky-Golay smoothing (Savitzky and Golay, 1964). The data were transformed to absorbance using the Compute General function in Unscrambler software (version 10.4, CAMO ASA, Norway) using the formula:  $\log_{10}(1/X)$  (Ollinger and Smith, 2005). For the subsequent gap derivative transformation, a gap of 3 and 5 were tried, but model performance results showed low  $R^2$  values. In the Savitzky-Golay smoothing formula, symmetrical points of 1, 2, and 3 were chosen, and combined with polynomial orders of 1, 2, 3, and 4.

Additionally, a Standard Normal Variate (SNV) and detrending transformation were modeled. Results were promising, though these models were not selected for mapping due to processing limitations.

## 2.4. Data Analysis

### 2.4.1 Partial least squares regression

The pre-processed dependent predictor variables (the imaging spectroscopy wavelengths extracted by plot boundaries and averaged) and independent measured variables (the plot level foliar N estimates from the four scaling methods) were input into the PLSR model in Unscrambler for data exploration of model sensitivity to different transformation settings. R code was then used to compare the model output by calculating adjusted  $R^2$  values from the predicted  $R^2$  provided by Unscrambler.

Variation in sample subsetting was tested. To avoid gaps produced by large water absorption ranges, data were broken into intervals: 451-1338, 1433-1773, and 1959-2464 nm. Subsets using only samples with 15% and 20% cover were run. Additionally, a bare ground sample was used to explore bias associated with a heterogeneous sample set, such as concern that the model was relying on creating a y-intercept that approximated the average N of the sites. A visually identified bare ground area was extracted, processed similarly, and inserted into the sample set while given a foliar nitrogen content of zero, replicated either 1, 10, or in numbers equivalent to the selection of vegetated plots.

Model RMSE, predicted and adjusted  $R^2$ , and number of factors were compared. Beta coefficients were graphed and compared to the 42 absorption features associated with vegetation listed in Curran (1989). Influential segments of the imaging spectroscopy data indicated by very high or low beta coefficients were considered relevant to predicting foliar nitrogen. Beta coefficients which were marked as significant at a  $p \leq .05$  alpha level were marked and the model was rerun with only these.

## **2.5 Nitrogen mapping for Reynolds Creek Experimental Watershed**

Beta coefficients produced by the model were applied to imaging spectroscopy data using R programming in R Studio. The code is provided in the appendix. Prior to import, the imaging spectroscopy data were transformed using absorbance and Savitzky-Golay (1,3,1) smoothing in ENVI software. As the imaging spectroscopy data used were resampled down to 1 m<sup>2</sup> from 2-3 m<sup>2</sup> pixels, beta coefficients were produced at a unit measurement of foliar N per 1 m<sup>2</sup> rather than plot level to ease interpretation. Blocks of the imaging spectroscopy data were subset, and then values per cell were extracted and beta coefficients were used to calculate foliar N estimates. The code creates a raster identical in size to the input imaging spectroscopy data. The foliar N estimate for that cell is calculated and written to the new raster, which is saved.

## **3 Results**

### **3.1 Field data collection results**

Field data were averaged to the plot scale (Table 1). Green biomass estimates using live shrubs were estimated to be an average of 5107 grams per plot. There was an average of 162 live shrubs per plot. LAI was low, averaging 0.68, and variable, with a range of 0.22-1.54. Average sagebrush shrub height was 50 cm. Though plots were chosen to contain dominant sagebrush cover, overall sagebrush canopy cover was low, an average of 22%. Leaf area averaged 3.5 cm<sup>2</sup>. Average dry weight of leaf samples was 0.048 grams. Average LMA was 0.015 g/cm<sup>2</sup>. N values, averaged to plot level, ranged from 1.2-2.3% with an average of 1.8%, similar to previous sagebrush sampling (1.5-2.8% in Mitchell et al., 2012b; 1.6-2.4% in Dashti et al., in review; 2.3% average in Yi et al., in review) and mixed canopy forest from other studies (0.9-1.8% in Knyazikhin et al., 2013). Shrubs in 2014 had slightly higher

average green biomass, despite having few shrubs per plot with less dense canopies, as indicated by LAI measurements. Overall, the plots were variable, for example, plot N-Sage15 had one of the lowest density values but relatively high green biomass predicted from canopy widths and allometry, which was nearly twice the biomass predicted for a plot with similar density: N-Sage07. A number of the plots were unusual when compared to the group and appeared as outliers in some PLSR models. For example, Sagebrush9 and N-Sage17 had approximately four times the density of other plots.

Plots (2014, 2015)	Green biomass [g]	Density (number of live shrubs)	LAI	Average sagebrush height per plot [cm]	% Cover of sagebrush per plot	Leaf area [cm <sup>2</sup> ]	Average sample dry weight [g]	LMA [g/cm <sup>2</sup> ]	Foliar N content
Sagebrush2	4401	57	0.86	67	32%	2.6	0.058	0.022	2.1%
Sagebrush3	2961	92	0.53	47	17%	2.0	0.044	0.022	2.2%
Sagebrush4	2302	54	0.22	48	15%	1.6	0.033	0.020	2.3%
Sagebrush5	2973	56	0.47	48	8%	2.5	0.047	0.019	2.0%
Sagebrush6	3777	84	0.34	44	25%	2.4	0.048	0.021	1.9%
Sagebrush7	5604	174	0.42	42	25%	2.4	0.047	0.020	2.3%
Sagebrush8	5146	101	0.44	47	22%	1.9	0.043	0.022	2.0%
Sagebrush9	11391	422	1.36	56	36%	2.4	0.062	0.025	1.9%
Sagebrush10	6730	103	0.90	55	28%	1.9	0.040	0.020	1.9%
N-Sage01	4462	128	0.44	54	13%	4.7	0.071	0.016	1.7%
N-Sage02	1742	154	0.33	28	5%	3.4	0.037	0.011	1.9%
N-Sage03	7889	59	1.54	77	21%	4.0	0.046	0.012	1.2%
N-Sage06	2945	135	0.45	37	14%	5.2	0.068	0.013	1.3%
N-Sage07	3478	42	0.49	52	10%	5.1	0.059	0.012	1.4%
N-Sage11	3739	794	0.25	20	19%	2.3	0.021	0.009	1.6%
N-Sage12	5079	135	0.49	27	26%	5.2	0.037	0.007	2.2%
N-Sage13	3144	111	0.47	42	14%	5.3	0.044	0.008	2.0%
N-Sage14	6962	98	1.12	48	19%	5.2	0.051	0.010	1.7%
N-Sage15	6578	44	1.44	110	19%	5.0	0.047	0.009	2.0%
N-Sage17	10208	468	1.05	26	58%	2.9	0.030	0.011	1.5%
H-Sage01	5728	81	0.61	78	33%	6.0	0.066	0.011	1.5%
Total avg	5107	162	0.68	50	22%	3.5	0.048	0.015	1.84%
2014 avg	5032	127	0.62	50	23%	2.2	0.047	0.021	2.07%
2015 avg	5163	187	0.72	50	21%	4.5	0.048	0.011	1.67%

Table 1. Descriptive statistics for the 21 sites sampled. Green biomass is calculated at plot level using a density count of live shrubs with allometric equations (leaf area is not included). Averages between the year subsets are listed below.

### 3.1.2 Biomass allometry

The simpler crown area method consistently outperformed the crown volume method, as seen in the following regression  $R^2$  table (Table 2). The crown area formulas that determine predicted weights are reported below each graph, which show the relationship of the predicted weights to actual weights (Figures 7-9). As size increased and there were fewer representative samples, the model becomes less accurate (Figures 7-9).

Compared measurements - y axis	$R^2$ using Ln(CVHt): Ln( $\pi*a*b*ht$ ) as x axis	$R^2$ using Ln(CA): Ln( $\pi*a*b$ ) as x axis
Above Ground Biomass	0.84	0.88
Green Biomass	0.79	0.80
Woody Biomass	0.84	0.88

Table 2. The  $R^2$  values here describe the relationship of the biomass measurements to their estimates using two different equations. Including height in the estimation formula consistently decreased correlation. The improved  $R^2$  values where height was omitted are in the third column.

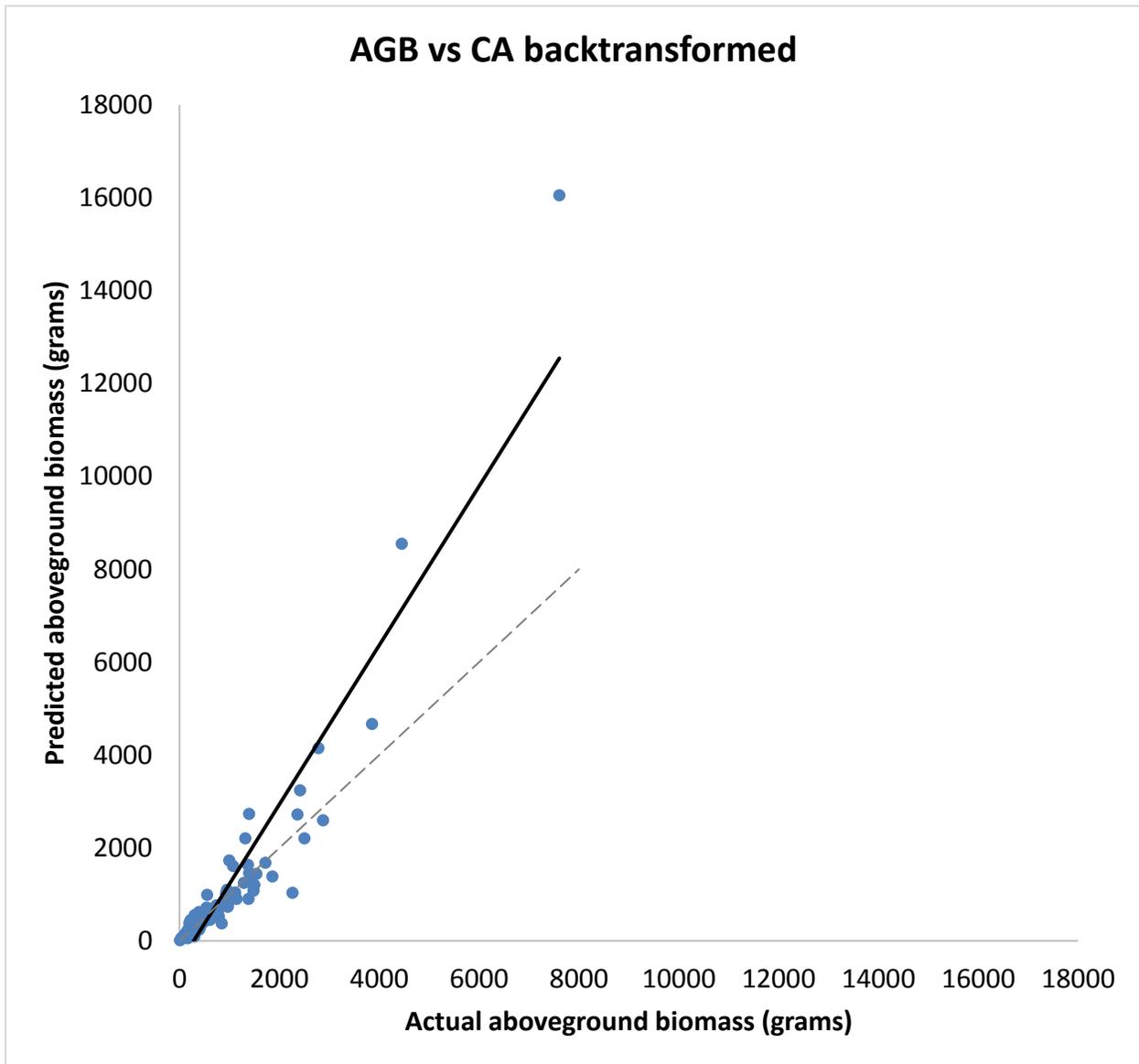


Figure 7. Aboveground biomass results. The aboveground biomass (AGB) prediction (y) using the crown area (CA) (x) yields the equation:  $AGB = e^{[1.208 * \ln(CA) + 7.7226]}$  ( $r^2=0.88$ ,  $P \leq 0.01$ ,  $n=91$ ). A 1:1 line is drawn in gray to show ideal prediction.

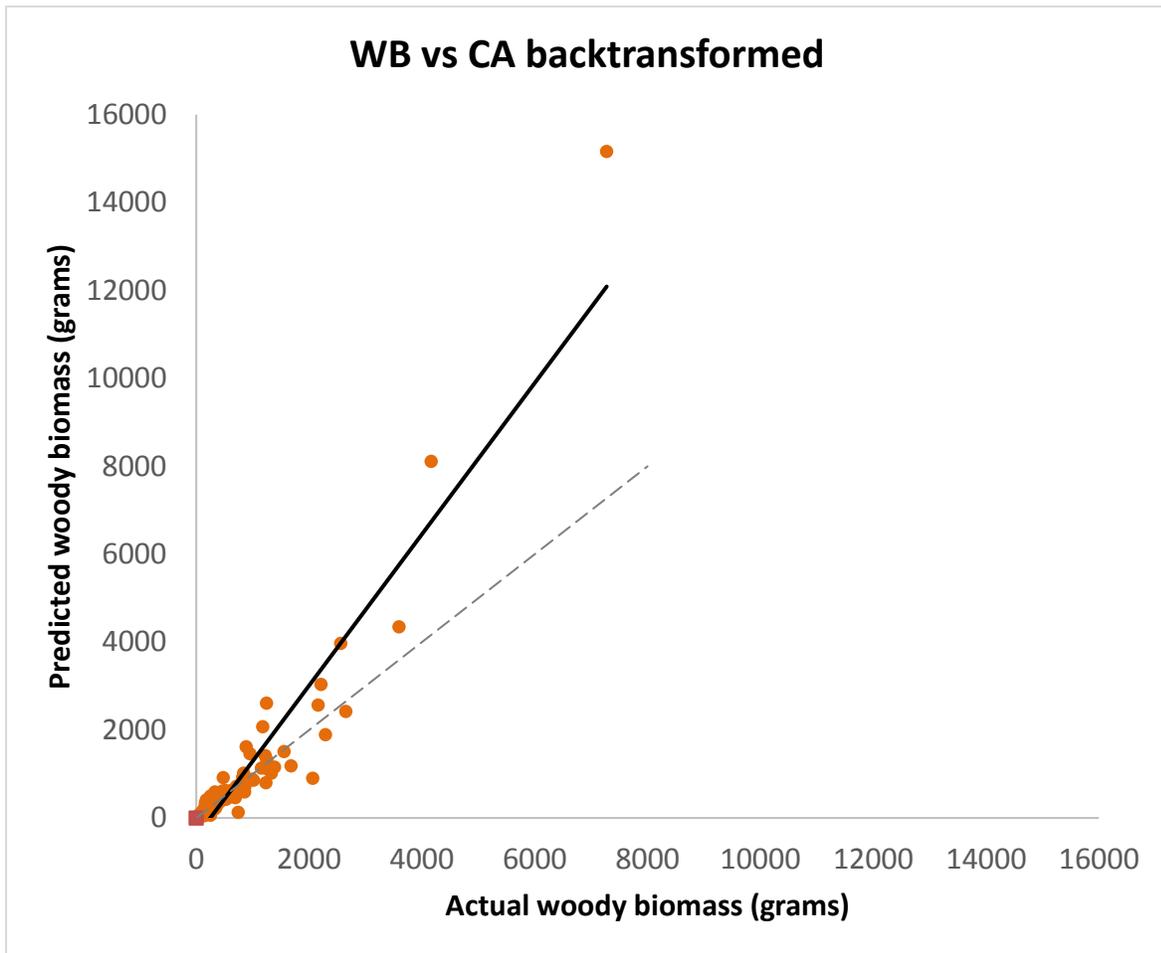


Figure 8. Woody biomass results. The woody biomass (WB) prediction (y) using the crown area (CA) (x) yields the equation:  $WB = e^{[1.2543 * \ln(CA) + 7.6312]}$  ( $r^2=0.88$ ,  $P \leq 0.01$ ,  $n=91$ ).

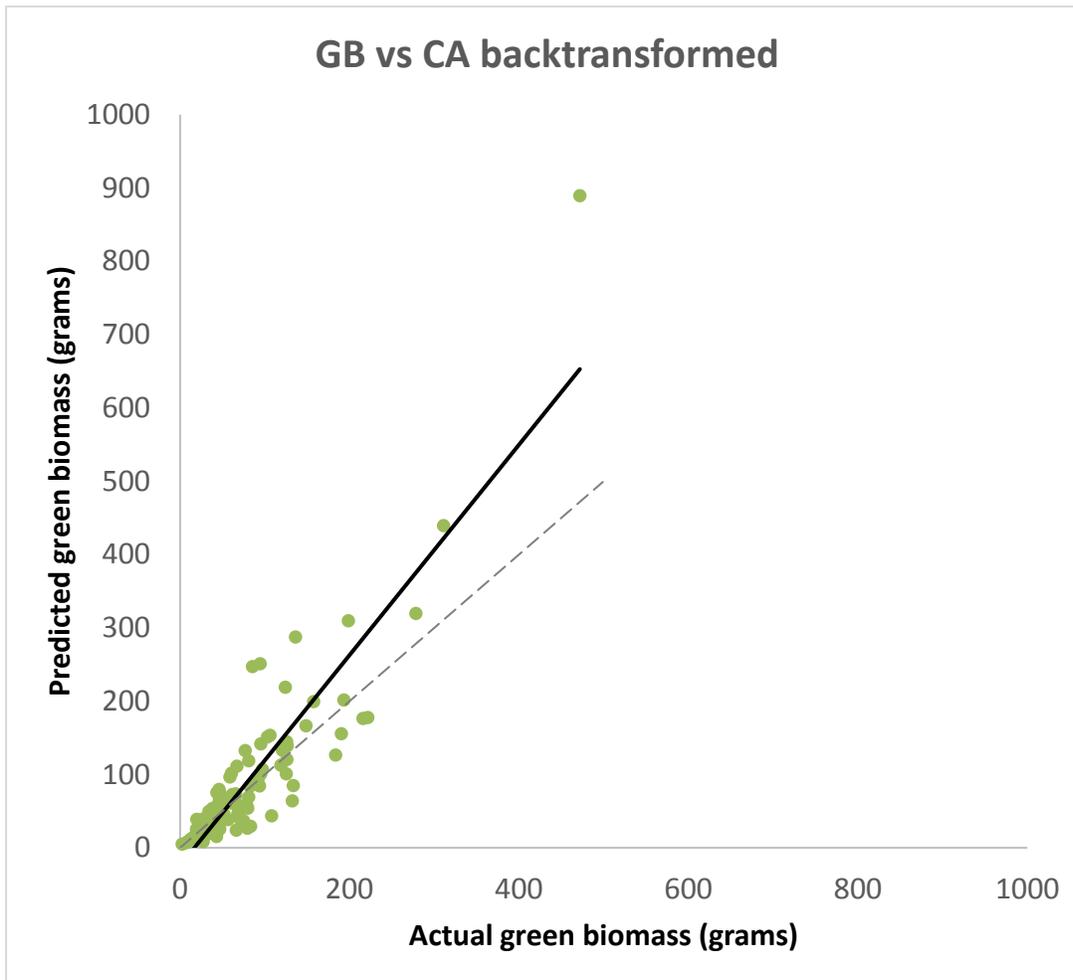


Figure 9. Green biomass results. The green biomass (GB) prediction (y) using the crown area (CA) (x) yields the equation:  $GB = e^{[0.9387 * \ln(CA) + 5.2129]}$  ( $r^2=0.80$ ,  $P \leq 0.01$ ,  $n=91$ ).

### 3.2 Partial least squares regression

Variations within scaling methods were tested. Averaging shrub level field sampling before calculation improved the Cover/LMA models compared to averaging after calculation, despite expectations that averaging later preserved accuracy and would improve models. Use of all live shrubs in density counts, rather than all shrubs or only live sagebrush shrubs, improved models for Biomass/Density. However, using a density count of only live sagebrush shrubs

improved performance over all live shrubs for the LAI/Density/LMA method. Improvements were generally slight. Removing all scaling adjustments to regress against average percent of foliar N for the plot improved over two methods in this case but not Cover/LMA or Biomass/Density (live). Restricting plot cover was also tested, though there were not enough plots with cover greater than 20% to successfully run the model. Using only plots with cover greater than 15% restricted the sample to 14 plots. The resulting models had lower adjusted  $R^2$  values but inconsistently improved RMSE values.

Biomass/Density, using all live shrubs for density, was used going forward with the absorbance and SG (1,3,1) smoothing transformations. This is one of the methods which showed RMSE improvement and only small change to adjusted  $R^2$  values after applying restricting samples to only those with cover greater than 15%. However, due to concern about the already small sample size, subsetting for cover was not used further. Further variations in model settings were tested. Model repetitions were run to compare variation in performance metrics caused by the random selection elements of the bootstrapping components of PLSR. Repetitions produced variation in adjusted  $R^2$  in the hundredths and in RMSE values in the tenths. Exclusion of wavelength values (leaving out 0, 1 or 3 values on either side) near water absorbance regions and weighting were applied out of concern that smoothing over large ranges of removed bands would affect wavelength values and degrade analysis. However, after averaging five repetitions, exclusion of values produced similar RMSE and adjusted  $R^2$  values. No exclusions and multiple model runs were used going forward.

Preliminary testing on the model prediction on imaging spectroscopy rasters showed low values were not produced for bare ground or other areas that were likely low in foliar N due to the model's beta coefficients including a high y-intercept value. While a homogenous sample set

is useful to identify consistently important wavelength predictors, the regression is limited to the range of its samples. Lacking a range in landscape types degrades the models ability to predict for types not represented in the sample. To test this, imaging spectroscopy data for a bare ground plot was visually selected, extracted, and transformed, and then assigned a value of zero for foliar N. This artificial bare ground plot was added to the plot samples in replicates of 1, 10, and 21. The mean was used to compare results rather than sum as the pixels overlapped the plot boundaries, selecting areas larger than 10x10 m, resulting in the model overestimating the foliar N sum compared to the plot scaling method. Figure 13 compares the scaling mean to the means including 1, 12, and 21 (equal to sample size) ground samples set at zero foliar N. These were selected to prevent the model from assuming a high base average N for all pixels.

Inclusions of bare ground samples in increasing amounts lowered the y intercept, moved predictions closer to scaling estimations, and appeared to more accurately represent a lack of foliage in the bare ground plot. However, while this experiment illustrates a need for further sampling, and representative sampling in number and type to the landscape being predicted, the bare ground samples were not used for the final model as they were not based on verified ground sampling.

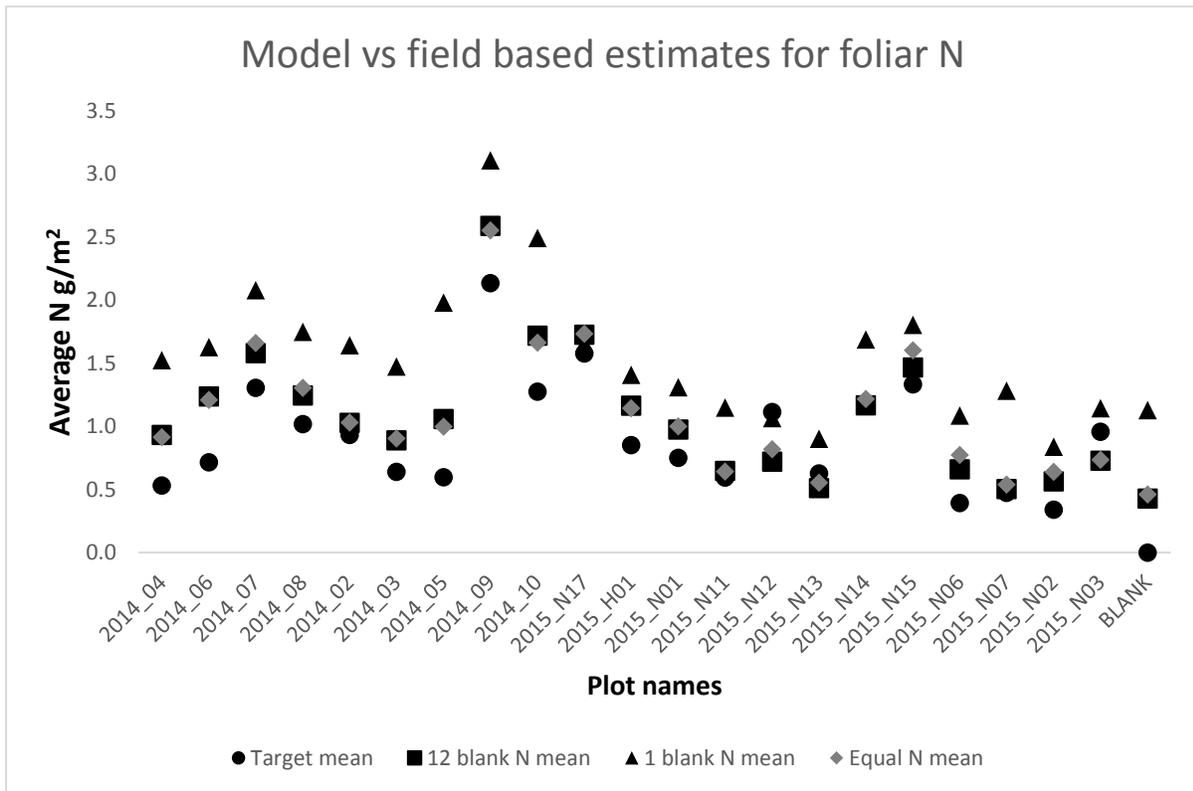


Figure 13. Higher amounts improved prediction for all sites and the bare ground sample, though higher amounts did not differ from each other.

Ten repetitions of the Biomass/Density (live) SG (1,3,1) models were run using the uncertainty test provided in the Unscrambler validation options. The uncertainty test works with the submodels created during cross validation to measure stability and identify important predictor variables. Predictors that were significant at a .05 alpha were marked, and the models were each rerun. The unmarked variables that showed unstructured variation and added noise to the model were excluded using downweighting, which multiplies insignificant variables by a very small value to reduce their influence. This inconsistently improved the model RMSE and adjusted  $R^2$  values.

When the predicted plot values of the two best performing models were graphed against the field estimates, they showed an association to either 2014 or 2015 samples. Samples were subset by year and rerun.

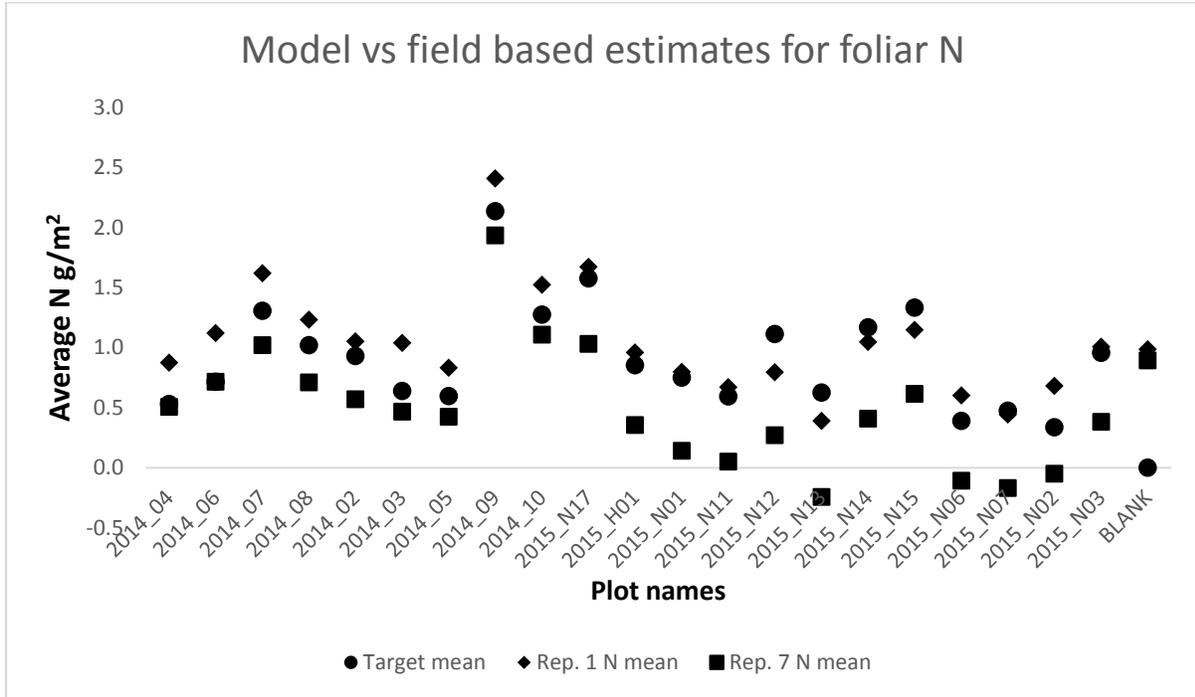


Figure 14. The first repetition of the model better predicted the 2015 plots, while the seventh repetition was a better predictor of 2014 plots. The five samples below zero are spurious, a function of the PLSR model allowing negative predictions.

Scaling method	Spectral transformation	RMSE (grams per m <sup>2</sup> )	Adjusted R <sup>2</sup>	Factors	Sample size	Description
Biomass/Density (live)	SG (1,3,1) downweighted	.18	.77	5	21	Rep. 1
Biomass/Density (live)	SG (1,3,1) downweighted	.18	.78	5	21	Rep. 7
Biomass/Density (live)	SG (1,3,1) downweighted	.19	.76	3	9	2014 field samples
Biomass/Density (live)	SG (1,3,1) downweighted	.09	.85	7	12	2015 field samples
Biomass/Density (live)	SG (1,3,1) downweighted	.14	.88	3	21	Combined repetitions

Table 3. The two best performing final model repetitions (1 and 7) and repetitions using subsets of 2014 and 2015 samples only, as well as the final multi-modal combination of repetitions.

While the 2014 showed similar RMSE and adjusted  $R^2$  values, the 2015 run showed improvement. The significant coefficients used for downweighting the 2014 and 2015 models match poorly and contradict. For example, in the 1100 nm range 2015 negatively correlates with the results of the repetitions which used both sets of samples, while the 2014 has a strong positive correlation at 1128 nm that isn't found in the other models. From the 600 to 900 nm range, the 2015 model showed a number of significant wavelengths that were not identified by the 2014 model or models using both sets of samples.

The selected significant wavelengths of the best performing models are shown in Figure 15 (Appendix F), which includes two repetitions using all samples and two subsets, each using either only 2014 or 2015 samples. They are annotated with absorption features from Curran (1989). The results are notably different between 2014 and 2015, and though the repetitions and 2014 samples often grouped together, there are contradictions such as 2014 selecting no VIS wavelengths, or Repetition 7 contradicting values around 640 nm. Excluding the C-H stretch at

1690 for the 2014 model, most wavelengths associated with N features were not identified as significant or were minimal contributors; in the 2015 model, starch overtones were selected while N features were not.

### 3.3 Plot level scaling of foliar N content

The Biomass/Density based scaling method produced the best N estimation results with PLSR adjusted  $R^2 = .63$ , and the Cover/LMA produced similar prediction results ( $R^2 = .58$ ). Different scaling methods favored different SG settings. Due to the disparate units between methods, models using the same method were compared to each other using RMSE and adjusted  $R^2$  first. The best model of each group was selected and then compared using adjusted  $R^2$ . Using 3 points on either side for smoothing worked best for three methods.

Scaling method	Spectral transformation	RMSE (grams per plot)	RMSE coefficient of variance	Adjusted $R^2$	Factors
Biomass/Density (live)	SG (1,3,1)	22.69 (grams per plot)	25%	0.63	5
LAI/Density/LMA	SG (1,3,2)	407.8 (grams per m <sup>2</sup> )	163%	-0.03	5
Cover/LMA	SG (1,2,2)	.257 (grams per cm <sup>2</sup> * fractional cover)	40%	0.58	5
Height/Cover	SG (1,3,4)	10.48 (absolute bulk cover)	53%	.21	2
Average percent foliar N	SG (1,3,2)	0.22 (average percent of foliar N at plot level)	12%	.47	1

Table 4. SG transformation settings were varied, annotated above as: (derivative, number of symmetrical points, and polynomial order used). Models within each scaling method were compared using RMSE, adjusted  $R^2$  and the number of ideal components selected.

As Biomass/Density and Cover/LMA had similar performance, all beta coefficients for their SG models were graphed in order to determine 1) if there were notable differences between SG transformation settings, and 2) if the spectral features being identified by the models were correlating with known N absorption features. The graphs were annotated with vegetative features that may be relevant to the coefficients shown. (See Figures 10, 11, and 12 in Appendix F.)

The general shape of features produced through different SG settings was consistent for both models, particularly the better performing Biomass/Density models. More noise and disagreement appears in the Cover/LMA model. Similar patterns appear in both. There is a possibly shifted spike of chlorophyll a near 670, a large cellulose/lignin/water drop at 1200, and a spread increase around both the 1690 and 2130 nitrogen features. The imaging spectroscopy data were sampled in intervals and then smoothed further, so some shift is expected. However, many N features were not notably detected or readings were inconsistent.

Average nitrogen was additionally examined with two settings. No scaling was applied. Results appeared noisy. The chlorophyll a spike was reduced; N features at 1020, 1980, and 2350 N feature were identified which were absent in the scaling methods' coefficients; the 1200 feature was notably inversely related.

### **3.4 Spectral transformation**

The spectral wavelength dataset was separated into subsets to compare performance of portions of spectrum known to be associated with N. The Biomass/Density (live) scaling method performed well previously and was used in subsequent experimentation to optimize the model. Breaking data into intervals when using reflectance data showed improvements when using the 1433-1773 nm range rather than the entire wavelength set; the 451-1338 range nm gave similar results to using the entire range, with worse performance using the 1959-2464 range nm. Models using log transformed absorbance data reacted similarly. Transforming the data using SG improved performance, but eliminated benefits of subsetting. These results guided the decision to use SG without subsetting in this work.

Spectral transformation method	RMSE (grams per plot)	$R^2$ Pearson	Adjusted $R^2$	Factors
Reflectance	34.57	0.41	0.16	6
Reflectance (451-1338 nm)	34.76	0.42	0.23	5
Reflectance (1433-1773 nm)	31.68	0.53	0.28	7
Reflectance (1959-2464 nm)	42.29	0.16	-0.05	4
Log Transformed Reflectance (451-1338 nm)	34.68	0.46	0.17	7
Log Transformed Reflectance (1433-1773 nm)	32.76	0.48	0.31	5
Log Transformed Reflectance (1959-2464 nm)	30.40	0.56	0.32	7
Log Transformed Reflectance (1959-2464 nm)	40.96	0.24	-0.01	5
First Derivative Reflectance SG (1,1,2)	27.98	0.58	0.44	5
First Derivative Reflectance SG (1,1,2) (451-1338 nm)	28.08	0.6	0.50	4
First Derivative Reflectance SG (1,1,2) (1433-1773 nm)	35.46	0.33	0.21	3
First Derivative Reflectance SG (1,1,2) (1959-2464 nm)	38.80	0.29	-0.01	6

Table 5. Within the Biomass/Density scaling method, using all live shrubs for the density count, transforming the data improved measurements. In raw data, subsetting data shows improvements over using the full dataset; however, as data were smoothed, gains lessened. The subset above 1959 nm consistently performed poorly.

Additional applications of gap derivative smoothing did not perform as well as SG smoothing. Detrending transformation (using polynomial settings of 1,2, and 4) followed by a Standard Normal Variate (SNV) transformation on SG (1,3,1) data for models using the

Biomass/Density scaling method also performed poorly, with low adjusted  $R^2$  values in comparison to models without the additional transformations. Given those results, additional transformations were not investigated in this study.

### **3.5 Wavelength selection using model repetitions**

Random elements of the PLSR model produced different wavelength selections and beta coefficients when repeatedly run using the same model parameters and input data (Table 6). Multimodal analysis was performed in this study as it has been shown to improve consistency in selection (Feilhauer et al., 2015). To improve consistency here, the beta coefficients of the downweighted reruns of ten repetitions of the Biomass/Density (live) SG (1,3,1) models were examined and combined. Three models with poor results were excluded. If a wavelength was selected as a positive or negative predictor by four of the remaining seven, and was not contradicted by any model (i.e. if a wavelength was positively indicated by at least four models it would be used unless another model selected it as a negative correlation) it was included in the subset.

A PLSR model was created using this combined subset. The model was rerun using downweighting. This selected the following bands: 677 (near the 660 absorption feature for chlorophyll a), 992 (near the 990 absorption feature for starch), 1133, 1208 (near the 1200 absorption feature for water, cellulose, starch, lignin), 1213, 1218, 1223, and 1263 nm. The model produced a 0.14 g/m<sup>2</sup> RMSE and 0.88 adjusted  $R^2$ , performing better than the 2014 subset and uncombined model results, and with an improved adjusted  $R^2$  over all models while using the entire set of plots.

Method	Wavelength selection
Biomass/Density (live) – Rep. 1	592, 597, 602, 607, 612, 622, 627, 672, 677, 987, 992, 997, 1133, 1138, 1143, 1208, 1213, 1218, 1223, 1233, 1263, 1318, 1528, 1583, 1588, 1714, 1749, 1974, 1989, 2044, 2049, 2099, 2104, 2109
Biomass/Density (live) – Rep. 3	672, 917, 922, 927, 952, 957, 1133, 1208, 1213, 1218, 1223, 1233, 1248, 1679, 1729, 1734, 1744, 2149, 2340
Biomass/Density (live) – Rep. 5	917, 922, 927, 952, 1133, 1208, 1213, 1218, 1223, 1233, 1248, 1588, 1669, 1679, 1729, 1734, 1744, 2149, 2340
Biomass/Density (live) – Rep. 7	672, 677, 992, 1133, 1138, 1143, 1208, 1213, 1218, 1223, 1233, 1588, 1744, 1749, 2104, 2109
Biomass/Density (live) – Rep. 8	592, 597, 602, 607, 617, 622, 627, 672, 677, 987, 992, 997, 1138, 1143, 1208, 1213, 1218, 1223, 1233, 1263, 1528, 1583, 1588, 1714, 1749, 1974, 1989, 2049, 2099, 2104, 2109
Biomass/Density (live) – Rep. 9	597, 602, 617, 622, 627, 677, 1138, 1143, 1208, 1213, 1218, 1223, 1233, 1238, 1248, 1263, 1528, 1583, 1588, 1739, 1744, 1749, 2099, 2104, 2109
Biomass/Density (live) – Rep. 10	592, 597, 602, 607, 612, 617, 622, 627, 672, 677, 987, 992, 997, 1138, 1143, 1208, 1213, 1218, 1223, 1233, 1263, 1528, 1583, 1588, 1714, 1749, 1974, 1989, 2049, 2099, 2104, 2109
Biomass/Density (live) – 2014 subset	917, 922, 932, 987, 992, 1128, 1213, 1679
Biomass/Density (live) – 2015 subset	622, 627, 642, 687, 697, 702, 712, 717, 727, 732, 837, 937, 942, 952, 962, 1118, 1133, 1138, 1143, 1148, 1158, 1163, 1323, 1593

Table 6. Wavelengths selected as significant by multiple repetitions using the full sample set as well as year subsets.

### 3.6 SLA and Cover results

Specific leaf area (the inverse of leaf mass area) was modeled against spectral predictors using PLSR and SG smoothing with settings of a first derivative transformation and then symmetrical points of 1, 2, or 3 and polynomial orders of 1, 2, 3, or 4. Models were then rerun using a subset of plots with cover of  $\geq 15\%$ . The best models of each group are listed below (Table 7). Results for both groups of models produced similar results, all with RMSE ranging 10-13  $\text{cm}^2/\text{g}$  and adjusted  $R^2$  ranging .82-.85. Models were rerun, using downweighting to

include only variables significant at a .05 alpha. The downweighted models selected similar wavelengths as significant predictors, both selecting: 502-547, 697-717 (near chlorophyll a absorption feature), 942 (near an oil feature), 1048-1053 (near an oil feature), 1128-1178, 1248-1283, 1298, 1333, 1443 (near a starch and sugar feature), 1458-1468 (near a cellulose and water feature), 1478-1483 (near cellulose, sugar, and water features), 1553-1563 (near a starch and sugar feature), 1749-1764, 2375, 2435, and 2455 nm (Curran, 1989). The model using all samples selected more narrow regions than the model using only samples with plot cover  $\geq 15\%$ . Similar to other models, downweighting insignificant indicators appeared to reduce noise and improve the relationship, indicated by the RMSE and adjusted  $R^2$ , while reducing the number of factors needed, though the improvements made by excluding low cover plots are small. LMA (using the inverse, SLA) was additionally graphed against foliar N content, but had no correlation.

Predicted value	Spectral transformation	RMSE (cm <sup>2</sup> /g)	Adjusted $R^2$	Factors	Sample size
SLA (cm <sup>2</sup> /g)	SG (1,1,2)	10.7	.85	3	21
SLA (cm <sup>2</sup> /g)	SG (1,1,2) unmarked downweighted	9.5	.90	2	21
SLA (cm <sup>2</sup> /g)	SG (1,2,2)	11.4	.84	3	14
SLA (cm <sup>2</sup> /g)	SG (1,2,2) unmarked downweighted	9.0	.92	2	14

Table 7. Results for PLSR models comparing SLA (the inverse of LMA) to smoothed spectral wavelengths.

### **3.7 Map of predicted foliar N values**

The raster file of the foliar N map created using the final models (Figure 16) is included in the supplementary materials.

# Reynolds Creek Experimental Watershed

 RCEW boundary

Foliar N estimates

(g/m<sup>2</sup>)

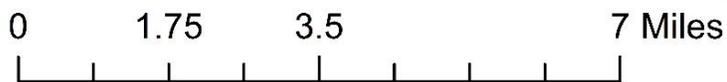
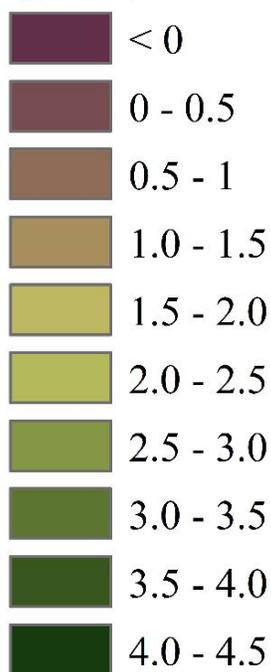


Figure 16. A wall to wall map depicting the distribution of foliar N for the Reynolds Creek Experimental Watershed (RCEW) at 1m<sup>2</sup> pixel resolution.

## **4. Discussion**

### **4.1 Partial least squares regression**

PLSR results were limited by constraints of method development, as time did not allow for full exploration of parameter options that could have influenced model performance. For example, models that performed poorly in this study may improve with different subsetting of wavelengths, exclusion of outlier samples, or increased repetitions. The high variability between repetitions when using the PLSR model suggests more experimentation is needed to determine how many repetitions are needed for a given sample size in order for results to stabilize. The ranking of the scaling methods is very sensitive to what bands were dismissed from further analysis. For example, some bands appear “noisy” in the regression coefficient graphs for some scaling methods and not others. The result is that if a band that is only noisy for one method is included as input in the model, that method may appear to underperform due to an arbitrary choice made by the modeler. One of the challenges to comparing methods was the large number of variations to test in PLSR. Later improvements provided by the "unmarked downweighted" refinement may have improved some models but could also have made discarded sampling methods more competitive. Models based on different sampling methods responded to different bands. Overall, data exploration and model performance was limited due to the small number of samples. For example, a larger dataset would allow experimentation with thresholds or gradients of variables, such as subsets of plots with greater cover or density. Increasing the sample size would give more options for investigating other variables and subsets and excluding potential

outliers without decreasing predictive power. Most importantly, the ability of the model to predict foliar N content for new locations is limited by a low number of sample plots with high foliar N content.

Variation in PLSR results produced by the random elements of the model enforces the need for repetition and is exacerbated by the small sample size of the dataset. Early data exploration using untransformed reflectance data and only subsets of selected wavelength intervals suggest irregular amounts of noise along the spectrum, particularly beyond second water absorption feature (1773-1959 nm). However, many of the best performing models had significant, though often contradicting, predictors in this region.

The final models were consistent with earlier work in showing shifted and unexpected peaks that did not correlate well to known N features, though other features were selected, particularly starch features (Figure 10-12, 15; Table 6; Curran, 1989). Some known nitrogen features were absent (Curran, 1989). Other unknown features were strongly and consistently present, such as peaks in the 1100-1160 nm range. A subset of 2014 samples did select the 1690 C-H stretch which includes N, and the mixed sample repetitions identified some N features in the 2000-2100 nm range. Outputs aligned with some expected features but had no response to others. Selected wavelengths did not align with N features identified by Curran (1989) (few within 20 nm of 2060 and 2130 and none at 1020, 2180, and 2300 and only slight negative correlation in one final model with 1510 and 1980); however, the 2015 model identified numerous positive correlations with known wavelength features in the VIS, including 640 and 660 chlorophyll features and one water absorbance feature around 970.

All models found strong (though often negative) correlations in numerous NIR and SWIR bands in agreement with literature findings (Ollinger et al., 2008; Ollinger, 2011; Martin et al.,

2008). The majority of model repetitions which used the full sample set and the 2015 subset also selected VIS wavelengths as predictors, diverging from previous findings that the inclusion of VIS wavelengths in analysis added noise and degraded correlation (Lepine et al., 2016; Martin et al., 2008; Ollinger et al., 2008; Ollinger, 2011) and in agreement with Mitchell et al. (2012). The majority of models using all samples and the 2014 subset model selected bands with strong negative correlations near the 1200 nm water absorption features in the NIR region, similar to Smith et al. (2003) and overlapping with results from Mitchell et al. (2012), which identified similar or identical bands at 1200, 1232, and 1263 nm. Ollinger (2011) states bands near water absorption features are influenced by structural effects.

While the models support the idea that the number of needed wavelengths could be reduced for prediction, though in model creation a large range of input bands is helpful, the disparity in the 2014 and 2015 model selections of important bands suggest that within dryland sagebrush ecosystems there is additional influence either from site or seasonal structural characteristics. The two subset models had little overlap in selection. The disparate selection in significant wavelengths may be a function of the discussed limitation introduced by field collection dates. Samples collected in the early wet May, 2015, season have a higher LAI (more dense canopy) and double the leaf area of the September, 2014, samples which were taken after summer droughts (Table 1). However, the 2014 season field samples on average had higher foliar N content than in 2015 and roughly half the SLA (Table 1). Thus the 2014 subset might be expected to produce a stronger spectral signature, having more and more concentrated foliar N; however, the 2015 subset produced a stronger model relationship. This may indicate a structural influence of the larger leaves and denser canopy or associated traits in 2015 allowed the model to

rely on VIS wavelengths that increased prediction, as well as relying on NIR and SWIR bands selected by the 2014 model.

As leaf absorbance and scattering within the canopy are influenced by leaf orientation and size, which differs between the ephemeral and perennial phenological states in sagebrush, the 2014 and 2015 differences may be due to structural differences in the canopy of these shrubs, particularly in the NIR response at 1213 nm. This may be useful to characterize signal differences between ephemeral and perennial foliage and identify the timing and extent of summer drought conditions on ephemeral foliage. Subsetting the 2014 samples produced a model fitting the spectral features of a canopy dominated by perennial foliage, while the 2015 samples identified spectral features from a canopy dominated by ephemeral foliage, and using the full sample set may identify generalized wavelengths predictors that can infer foliar N for both canopy types. While the full dataset may give a good general detection, subset models appear to generate more specialized datasets for ephemeral or perennial leaves, though the validity of that inference is unclear until it is verified by field tests. Additionally, this difference in selected wavelength may be attributable to site choice, since the 2015 field collection did not return to 2014 sites.

It is unusual that the larger sample size of 2015 was not more influential in dampening characteristic signals of the 2014 plots in the two repetitions using mixed samples. This may be related to the 2014 beta coefficients having greater magnitude than those produced in the 2015 model or higher and more concentrated foliar N in sampling. Future researchers may want to note that samples which mix phenological states selected different predictors than either of the more homogenous subsets in this work. For landscape modeling, it may be preferable to restrict sampling and aerial collection dates to within dates capturing a single phenological state. The

choice of which state should be informed by the goals of the project. In projects seeking to look at ecological questions, certain ecological states may be better suited, or both may be collected to investigate generalized ecological dynamics.

The final model, which used only wavelengths consistently selected in the same direction, evidencing benefits of using multiple repetitions to stabilize the random variation produced within the model development. Multimodal analysis or repetitive combinations may stabilize predictive wavelengths in future work, as suggested in literature (Feilhauer et al., 2015). This may improve transferability between sites within the Great Basin. The model repetitions notably relied on different subsets of wavelengths to generate predictions with similar success, suggesting combinations of wavelengths may substitute for one another as indicators. This supports the suggestion that use of large VIS, NIR, and SWIR ranges may be stronger than attempts to use limited regions to reduce noise or to assume model transferability.

#### **4.2 Field collection discussion**

Several field data variables occur in both successful and underperforming methods: LMA, density counts, and fractional cover (Table 8). Shrub height and LAI appear only in methods which performed poorly. Canopy width was successfully used to establish allometric relationships that reduced need for destructive sampling, and scaling using allometric estimation of green biomass performed well. Future data collection may exclude low performing variables such as LAI and height, while including cover, canopy allometry, LMA, and density counts. Future scaling work may examine additional combinations of variables.

	Shrub height	Canopy width	Leaf area	Leaf weight	Cover	Density	LAI
LAI/Density/LMA			yes	yes		yes	yes
Height/Cover	yes				yes		
Cover/LMA			yes	yes	yes		
Biomass/Density		yes				yes	

Table 8. A comparison of variables present in each model.

The Cover/LMA model ( $R^2 = .58$ ) supports previous work (Mitchell et al., 2012b) that suggest LMA is a useful metric for scaling sagebrush foliar N in airborne imaging spectroscopy studies. In contrast, alternative scaling metrics explored in this thesis such as LAI/Density/LMA and Height/Cover scaling methods performed poorly in this study, with  $R^2$  values consistently below 0.60. Underperformance of the LAI/Density/LMA model may be attributable to the open canopy structure of sagebrush, which correlates poorly with LAI (Hurcom and Harrison, 1998; Lepine et al., 2016; Ollinger et al., 2008). Ollinger (2011) criticized LAI for aggregating the canopy variables which relate to foliage distribution patterns and branching and thus obscuring differences in scattering. LAI may be unsuited to evaluating dryland ecosystems with low density, cover, and canopy closure despite its use in other ecosystems (Hurcom and Harrison, 1998).

In the development of the allometry values, use of a height variable in a volume formula reduced correlations, as compared to excluding height in area formula. This suggests height may also interfere with scaling methods for estimating sagebrush foliar N, as evidenced by the poor performance of the Height/Cover scaling method ( $R^2 = .21$ ). The distance between shrubs as a function of water competition and abundant light availability may reduce the advantages of using height in sagebrush scaling methods.

The Biomass/Density (live only density) scaling method performed well ( $R^2 = .63$ ), as did the Cover/LMA model ( $R^2 = .58$ ). Canopy width measurements may be a better representation of sagebrush foliage than height or LAI. The Biomass/Density model also did not rely on leaf area directly, but through allometric data based on a much larger dataset with mixed seasons (fall and summer). The greater sample size may have stabilized the relationships or its mixed seasonality may have better represented the mixed seasonality of RCEW samples. However, as the model based on the Cover/LMA scaling method used the 2014/2015 RCEW leaf area values and performed similarly, this would suggest that seasonal differences in LMA may not be the variable negatively influencing the model based on the LAI/Density/LMA scaling method.

### **4.3 Transformations and subsetting**

In this work, using raw data, first derivative, or gap derivative processing alone was insufficient. The combination of first derivative and SG smoothing prepared the data well for PLSR. However, while transforming the data using SG improved performance, the benefits of subsetting data declined, suggesting irregular levels of noise between subsets was smoothed out by the transformation. More noise and disagreement appears in the Cover/LMA model results than the biomass as well, which may indicate a greater sensitivity for this method to transformation settings.

### **4.4 Mapping predicted foliar N**

Mapping produced anomalous results over land cover types which were not included in the creation of the model. For example, water showed unusual and varying patterns in N distribution and such results are unlikely to be valid, while irrigated fields associated with relatively large patches of high biomass did not display high N values as expected. In particular,

water and irrigated fields could be excluded with masking in future mapping. Trees appear to be represented more accurately, as well as low stature shrub vegetation that was well represented in the sampled plots. Additionally, despite physical impossibility, the model allows for the generation of negative predictions.

Depending on the goals of future studies, the sample set may benefit from including representation of multiple landscape types. While a homogenous sample is useful for identifying consistent spectral features, heterogeneous plots such as those with bare ground should be included to represent instances that will occur in high resolution mapping. This work would indicate that models which do not include the landscape types in the sample will likely overestimate foliar N for bare ground cover and incorrectly represent water and agricultural fields.

Full geotiff files for this model are included in supplementary materials.

## **5. Conclusions**

Findings from this work can limit the time and resources necessary to collect vegetation variables in the field. In addition, this study identifies transformations and regression methods that are useful in identifying wavelengths for continuous N mapping at the watershed scale. The best performing scaling method included density counts and canopy width measurements of live shrubs, as well as leaf sampling and laboratory processing for N. Development of allometric biomass relationships based on canopy width inputs can replace destructive sampling. If allometry is not available, use of fractional cover and leaf sampling performs similarly. Height and LAI appear unnecessary to collect. Estimating vegetative area through cover or density is important. Species identification in density counts was not important when used with the canopy

allometry, but was used for fractional cover; species composition and abundance may influence whether identification is necessary. Canopy widths, density counts without species identification, and leaf nitrogen analysis would be sufficient to apply this model.

The lack of alignment to known N spectral features and between models suggests that regional field sampling is still necessary, as wavelength selections may not be accurately transferred between sites, seasons, or years. Transformations can be used to reduce noise. PLSR and subsetting can be used to further exclude noise and reduce wavelengths needed for prediction. However, model results are sensitive to band removal, transformations, sampling biases, model settings, and random variation within model creation. Few band selections from the subsets corresponded to N features, while the best performing model used wavelengths associated with starches, suggesting that correlations should be investigated for physical causes to better interpret model predictions. This may support the assertion that indirect measurements of N are being selected as predictors.

Large scale mapping of foliar N can be combined with maps of biomass to assess nutritional quality and quantity of foliage for rangeland management and to infer patterns of wildlife habitat selection, movement, and abundance. Foliar N maps can also be used to assess photosynthetic activity and N-cycling to improve climate models. Over a large area, these maps can be combined with climate variables to investigate vegetative growth patterns for invasive and fire management. Land covers not represented in sampling were poorly predicted and should either be excluded from mapping or included in sampling and model development. Future work would benefit from increased sampling which represented seasonal and regional variations as well as different land cover types. The results of this study can be used to evaluate N prediction from satellite platforms such as Landsat, Sentinel-2, and MODIS.

## References

- ACCP Project Page [WWW Document], n.d. URL <http://daac.ornl.gov/ACCP/accp.shtml> (accessed 7.30.17).
- Adeel, Z., World Resources Institute (Eds.), 2005. Ecosystems and human well-being: desertification synthesis: a report of the Millennium Ecosystem Assessment. World Resources Institute, Washington, D.C.
- Ahlstrom, A., Raupach, M.R., Schurgers, G., Smith, B., Arneeth, A., Jung, M., Reichstein, M., Canadell, J.G., Friedlingstein, P., Jain, A.K., Kato, E., Poulter, B., Sitch, S., Stocker, B.D., Viovy, N., Wang, Y.P., Wiltshire, A., Zaehle, S., Zeng, N., 2015. The dominant role of semi-arid ecosystems in the trend and variability of the land CO<sub>2</sub> sink. *Science* 348, 895–899. doi:10.1126/science.aaa1668
- Asner, G.P., 1998. Biophysical and biochemical sources of variability in canopy reflectance. *Remote Sensing of Environment* 64, 234–253.
- Asner, G.P., Knapp, D.E., Kennedy-Bowdoin, T., Jones, M.O., Martin, R.E., Boardman, J., Field, C.B., 2007. Carnegie airborne observatory: in-flight fusion of hyperspectral imaging and waveform light detection and ranging for three-dimensional studies of ecosystems. *Journal of Applied Remote Sensing* 1, 013536–013536.
- Asner, G.P., Lobell, D.B., 2000. A biogeophysical approach for automated SWIR unmixing of soils and vegetation. *Remote Sensing of Environment* 74, 99–112. doi:10.1016/S0034-4257(00)00126-7
- Asner, G.P., Martin, R.E., 2008. Spectral and chemical analysis of tropical forests: Scaling from leaf to canopy levels. *Remote Sensing of Environment* 112, 3958–3970. doi:10.1016/j.rse.2008.07.003
- Asner, G.P., Martin, R.E., Tupayachi, R., Emerson, R., Martinez, P., Sinca, F., Powell, G.V.N., Wright, S.J., Lugo, A.E., 2011. Taxonomy and remote sensing of leaf mass per area (LMA) in humid tropical forests. *Ecological Applications* 21, 85–98. doi:10.1890/09-1999.1
- Barrett, S.W., Adams, T., Kapus, J., 2016. Secretarial Order 3336 Science Priorities: The role of science past, present, and future. *Fire Science Digest* 23, 1–12.
- Bobbink, R., Hicks, K., Galloway, J., Spranger, T., Alkemade, R., Ashmore, M., Bustamante, M., Cinderby, S., Davidson, E., Dentener, F., Emmett, B., Erisman, J.-W., Fenn, M., Gilliam, F., Nordin, A., Pardo, L., De Vries, W., 2010. Global assessment of nitrogen deposition effects on terrestrial plant diversity: a synthesis. *Ecological Applications* 20, 30–59.

- Bolster, K.L., Martin, M.E., Aber, J.D., 1996. Determination of carbon fraction and nitrogen concentration in tree foliage by near infrared reflectances: a comparison of statistical methods. *Can. J. For. Res.* 26, 590–600. doi:10.1139/x26-068
- Borel, C.C., Gerstl, S.A., 1994. Nonlinear spectral mixing models for vegetative and soil surfaces. *Remote Sensing of Environment* 47, 403–416.
- Breshears, D.D., 2006. The grassland–forest continuum: trends in ecosystem properties for woody plant mosaics? *Frontiers in Ecology and the Environment* 4, 96–104. doi:10.1890/1540-9295(2006)004[0096:TGCTIE]2.0.CO;2
- Breshears, D.D., Nyhan, J.W., 1998. Effects of woody plants on microclimate in a semiarid woodland: Soil temperature and evaporation in canopy and intercanopy patches. *International Journal of Plant Sciences* 159, 1010.
- Campbell, J. B., Wynne, R. H., 2011. *Introduction to Remote Sensing (Fifth Edition)*. Photogrammetric Record. Guilford Press, New York, USA, 2011. ISBN 978.160918.176.5.
- Cao, M., Prince, S.D., Small, J., Goetz, S.J., 2004. Remotely sensed interannual variations and trends in terrestrial net primary productivity 1981-2000. *Ecosystems* 7. doi:10.1007/s10021-003-0189-x
- Chrien, T., Green, R., Eastwood, M., 1990. Accuracy of the spectral and radiometric laboratory calibration of the airborne visible/infrared imaging spectrometer (AVIRIS) [WWW Document]. URL [https://aviris.jpl.nasa.gov/proceedings/workshops/90\\_docs/1.PDF](https://aviris.jpl.nasa.gov/proceedings/workshops/90_docs/1.PDF) (accessed 11.4.16).
- Cleary, M.B., Pendall, E., Ewers, B.E., 2008. Testing sagebrush allometric relationships across three fire chronosequences in Wyoming, USA. *Journal of Arid Environments* 72, 285–301. doi:10.1016/j.jaridenv.2007.07.013
- Curran, P.J., 1989. Remote sensing of foliar chemistry. *Remote Sensing of Environment* 30 (3), 271-278.
- Dashti, H., Glenn, N.F., Ilangakoon, N., Mitchell, J.J., Spaete, L., Flores, L., Qi, Y., De Graaff, M.A., Ustin, S., in review. Remote sensing of shrublands.
- Donohue, R.J., Roderick, M.L., McVicar, T.R., Farquhar, G.D., 2013. Impact of CO<sub>2</sub> fertilization on maximum foliage cover across the globe's warm, arid environments. *Geophys. Res. Lett.* 40, 3031–3035. doi:10.1002/grl.50563
- Evans, J.R., 1989. Photosynthesis and nitrogen relationships in leaves of C<sub>3</sub> plants. *Oecologia* 78, 9–19. doi:10.1007/BF00377192

- Feilhauer, H., Asner, G.P., Martin, R.E., 2015. Multi-method ensemble selection of spectral bands related to leaf biochemistry. *Remote Sensing of Environment* 164, 57–65. doi:10.1016/j.rse.2015.03.033
- Field, C., Mooney, H.A., 1986. The photosynthesis-nitrogen relationship in wild plants, in: T.J. Givnish (Ed.), *Proceedings of the sixth Maria Moors Cabot Symposium, Evolutionary Constraints on Primary Productivity: Adaptive Patterns of Energy Capture in Plants*, Harvard Forest, August 1983. Cambridge University Press, Cambridge, UK, 25-55.
- Field, C. B., Barros, V. R., & Intergovernmental Panel on Climate Change (Eds.). (2014). *Climate Change 2014: Impacts, Adaptation, and Vulnerability: Working Group II Contribution to the Fifth Assessment Report of the Intergovernmental Panel on Climate Change*. Cambridge University Press, New York.
- Forbey, J., Wiggins, N., Frye, G., Connelly, J., 2013. Hungry grouse in a warming world: emerging risks from plant chemical defenses and climate change. *Wildlife Biology* 19, 374–381.
- Frye, G.G., Connelly, J.W., Musil, D.D., Forbey, J.S., 2013. Phytochemistry predicts habitat selection by an avian herbivore at multiple spatial scales. *Ecology* 94, 308–314. doi:10.1890/12-1313.1
- Gilbert, N., 2011. Science enters desert debate. *Nature News* 477, 262–262. doi:10.1038/477262a
- Hasselquist, N.J., Germino, M.J., Sankey, J.B., Ingram, L.J., Glenn, N.F., Bahn, M., 2011. Aeolian nutrient fluxes following wildfire in sagebrush steppe: implications for soil carbon storage. *Biogeosciences* 8, 3649–3659. doi:10.5194/bg-8-3649-2011
- Hooper, D.U., Johnson, L., 1999. Nitrogen limitation in dryland ecosystems: Responses to geographical and temporal variation in precipitation. *Biogeochemistry* 46, 247–293.
- Hu, H.-W., Macdonald, C.A., Trivedi, P., Anderson, I.C., Zheng, Y., Holmes, B., Bodrossy, L., Wang, J.-T., He, J.-Z., Singh, B.K., 2016. Effects of climate warming and elevated CO<sub>2</sub> on autotrophic nitrification and nitrifiers in dryland ecosystems. *Soil Biology and Biochemistry* 92, 1–15. doi:10.1016/j.soilbio.2015.09.008
- Hunt Jr, E.R., Everitt, J.H., Ritchie, J.C., Moran, M.S., Booth, D.T., Anderson, G.L., Clark, P.E., Seyfried, M.S., 2003. Applications and research using remote sensing for rangeland management. *Photogrammetric Engineering & Remote Sensing* 69, 675–693.
- Hurcom, S.J., Harrison, A.R., 1998. The NDVI and spectral decomposition for semi-arid vegetation abundance estimation. *International Journal of Remote Sensing* 19, 3109–3125. doi:10.1080/014311698214217

- Jarou, Z., 2009. Measuring Leaf Area with Adobe Photoshop 3. [WWW Document]. URL <https://www.youtube.com/watch?v=E3O-V6WLw0g> (accessed 11.4.16).
- Johnson, L., Hlavka, C., Peterson, D., 1994. Multivariate analysis of AVIRIS data for canopy biochemical estimation along the Oregon transect. *Remote Sensing of Environment* 47, 216–223.
- Knyazikhin, Y., Schull, M.A., Stenberg, P., Möttus, M., Rautiainen, M., Yang, Y., Marshak, A., Carmona, P.L., Kaufmann, R.K., Lewis, P., Disney, M.I., Vanderbilt, V., Davis, A.B., Baret, F., Jacquemoud, S., Lyapustin, A., Myneni, R.B., 2013. Hyperspectral remote sensing of foliar nitrogen content. *Proceedings of the National Academy of Sciences of the United States of America* 110, 811–812.
- Kokaly, R.F., Clark, R.N., 1999. Spectroscopic determination of leaf biochemistry using band-depth analysis of absorption features and stepwise multiple linear regression. *Remote Sensing of Environment* 67, 267–287. doi:10.1016/S0034-4257(98)00084-4
- Kokaly, R.F., Despain, D.G., Clark, R.N., Livo, K.E., 2003. Mapping vegetation in Yellowstone National Park using spectral feature analysis of AVIRIS data. *Remote Sensing of Environment* 84, 437–456. doi:10.1016/S0034-4257(02)00133-5
- Kokaly, R.F., Asner, G.P., Ollinger, S.V., Martin, M.E., Wessman, C.A., 2009. Characterizing canopy biochemistry from imaging spectroscopy and its application to ecosystem studies. *Remote Sensing of Environment, Imaging Spectroscopy Special Issue* 113, Supplement 1, S78–S91. doi:10.1016/j.rse.2008.10.018
- Latorre-Carmona, P., Knyazikhin, Y., Alonso, L., Moreno, J.F., Pla, F., Yan, Y., 2014. On hyperspectral remote sensing of leaf biophysical constituents: decoupling vegetation structure and leaf optics using CHRIS PROBA data over crops in Barrax. *IEEE Geoscience and Remote Sensing Letters* 11, 1579–1583. doi:10.1109/LGRS.2014.2305168
- Lepine, L.C., Ollinger, S.V., Ouimette, A.P., Martin, M.E., 2016. Examining spectral reflectance features related to foliar nitrogen in forests: Implications for broad-scale nitrogen mapping. *Remote Sensing of Environment* 173, 174–186. doi:10.1016/j.rse.2015.11.028
- Li, A., Glenn, N.F., Olsoy, P.J., Mitchell, J.J., Shrestha, R., 2015. Aboveground biomass estimates of sagebrush using terrestrial and airborne LiDAR data in a dryland ecosystem. *Agricultural and Forest Meteorology* 213, 138–147. doi:10.1016/j.agrformet.2015.06.005
- Marks, D., Seyfried, M., Flerchinger, G., Winstral, A., 2007. Research data collection at the Reynolds creek experimental watershed. *Journal of Service Climatology* 1, 1–18.
- Martin, M.E., Aber, J.D., 1997. High spectral resolution remote sensing of forest canopy lignin, nitrogen, and ecosystem processes. *Ecological Applications* 7, 431–443. doi:10.2307/2269510

- Martin, M.E., Plourde, L.C., Ollinger, S.V., Smith, M.-L., McNeil, B.E., 2008. A generalizable method for remote sensing of canopy nitrogen across a wide range of forest ecosystems. *Remote Sensing of Environment* 112, 3511–3519. doi:10.1016/j.rse.2008.04.008
- Mattson, 1980. Herbivory in relation to plant nitrogen content. *Annual Review of Ecology and Systematics* 11, 119-161.
- Miller, R.F., Shultz, L.M., 1987. Development and longevity of ephemeral and perennial leaves on *Artemisia tridentata* Nutt. ssp. *wyomingensis*. *The Great Basin Naturalist* 227–230.
- Mitchell, J.J., Glenn, N.F., Sankey, T.T., Derryberry, D.R., Anderson, M.O., Hruska, R.C., 2012a. Spectroscopic detection of nitrogen concentrations in sagebrush. *Remote Sensing Letters* 3, 285–294. doi:10.1080/01431161.2011.580017
- Mitchell, J.J., Glenn, N.F., Sankey, T.T., Derryberry, D.R., Germino, M.J., 2012b. Remote sensing of sagebrush canopy nitrogen. *Remote Sensing of Environment* 124, 217–223. doi:10.1016/j.rse.2012.05.002
- Mitchell, J.J., Shrestha, R., Spaete, L.P., Glenn, N.F., 2015. Combining airborne hyperspectral and LiDAR data across local sites for upscaling shrubland structural information: Lessons for HypsIRI. *Remote Sensing of Environment, Special Issue on the Hyperspectral Infrared Imager (HypsIRI)* 167, 98–110. doi:10.1016/j.rse.2015.04.015
- Moorcroft, P.R., 2006. How close are we to a predictive science of the biosphere? *Trends in Ecology & Evolution, Twenty years of TREE - part 2* 21, 400–407. doi:10.1016/j.tree.2006.04.009
- Mueller, E.N., Wainwright, J., Parsons, A.J., 2007. The stability of vegetation boundaries and the propagation of desertification in the American Southwest: A modelling approach. *Ecological Modelling* 208, 91–101. doi:10.1016/j.ecolmodel.2007.04.010
- Noojipady, P., Prince, S.D., Rishmawi, K., 2015. Reductions in productivity due to land degradation in the drylands of the southwestern United States. *Ecosystem Health and Sustainability* 1, 1–15. doi:10.1890/EHS15-0020.1
- Okin, G.S., 2008. A new model of wind erosion in the presence of vegetation. *J. Geophys. Res.* 113, F02S10. doi:10.1029/2007JF000758
- Okin, G.S., Roberts, D.A., Murray, B., Okin, W.J., 2001. Practical limits on hyperspectral vegetation discrimination in arid and semiarid environments. *Remote Sensing of Environment* 77, 212–225. doi:10.1016/S0034-4257(01)00207-3
- Ollinger, S.V., 2011. Sources of variability in canopy reflectance and the convergent properties of plants. *The New Phytologist* 189, 375–394.

- Ollinger, S.V., Richardson, A.D., Martin, M.E., Hollinger, D.Y., Frohling, S.E., Reich, P.B., Plourde, L.C., Katul, G.G., Munger, J.W., Oren, R., Smith, M.-L., U, K.T.P., Bolstad, P.V., Cook, B.D., Day, M.C., Martin, T.A., Monson, R.K., Schmid, H.P., 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 of the United States of America* 105, 19336–19341.
- Ollinger, S.V., Smith, M.L., 2005. Net primary production and canopy nitrogen in a temperate forest landscape: an analysis using imaging spectroscopy, modeling and field data. *Ecosystems* 8, 760–778.
- Olsoy, P.J., Forbey, J.S., Rachlow, J.L., Nobler, J.D., Glenn, N.F., Shipley, L.A., 2015. Fearscape: mapping functional properties of cover for prey with terrestrial LiDAR. *BioScience* 65, 74–80.
- Olsoy, P.J., Glenn, N.F., Clark, P.E., Derryberry, D.R., 2014. Aboveground total and green biomass of dryland shrub derived from terrestrial laser scanning. *ISPRS Journal of Photogrammetry and Remote Sensing* 88, 166–173. doi:10.1016/j.isprsjprs.2013.12.006
- Olsoy, P.J., Griggs, T.C., Ulappa, A.C., Gehlken, K., Shipley, L.A., Shewmaker, G.E., Forbey, J.S., 2016. Nutritional analysis of sagebrush by near-infrared reflectance spectroscopy. *Journal of Arid Environments* 134, 125–131. doi:10.1016/j.jaridenv.2016.07.003
- Peterson, D., Aber, J.D., Matson, P., Card, D., Swanberg, N., Wessman, C.A., Spanner, M., 1988. Remote sensing of forest canopy and leaf biochemical contents. *Remote Sensing of Environment* 24, 85–108.
- Polley, H.W., Briske, D.D., Morgan, J.A., Wolter, K., Bailey, D.W., Brown, J.R., 2013. Invited synthesis: Climate change and North American rangelands: Trends, projections, and implications. *Rangeland Ecology & Management* 66, 493–511. doi:10.2111/REM-D-12-00068.1
- Poulter, B., Frank, D., Ciais, P., Myneni, R.B., Andela, N., Bi, J., Broquet, G., Canadell, J.G., Chevallier, F., Liu, Y.Y., Running, S., Sitch, S., van der Werf, G.R., 2014. Contribution of semi-arid ecosystems to interannual variability of the global carbon cycle. *Nature* 509, 600–604. doi:10.1038/nature13379
- Prater, M.R., DeLucia, E.H., 2006. Non-native grasses alter evapotranspiration and energy balance in Great Basin sagebrush communities. *Agricultural and Forest Meteorology* 139, 154–163. doi:10.1016/j.agrformet.2006.08.014
- Qi, Y., Ustin, S., Glenn, N.F., in review. Imaging spectroscopic analysis of biochemical traits for shrub species in the Great Basin, USA.

- Ravi, S., D'Odorico, P., Wang, L., White, C.S., Okin, G.S., Macko, S.A., Collins, S.L., 2009. Post-fire resource redistribution in desert grasslands: A possible negative feedback on land degradation. *Ecosystems* 12, 434–444. doi:10.1007/s10021-009-9233-9
- Ray, T.W., Murray, B.C., 1996. Nonlinear spectral mixing in desert vegetation. *Remote Sensing of Environment* 55, 59–64. doi:10.1016/0034-4257(95)00171-9
- Roberts, D.A., Smith, M.O., Adams, J.B., 1993. Green vegetation, nonphotosynthetic vegetation, and soils in AVIRIS data. *Remote Sensing of Environment* 44, 255–269.
- Safriel, U., Adeel, Z., Niemeijer, D., Puigdefabregas, J., White, R., Lal, R., Winsolow, M., Ziedler, J., Prince, S., Archer, E. and King, C., 2006. Dryland systems, in: *Ecosystems and Human Well-being. Current State and Trends*, Vol. 1 (pp. 625-656). Island Press.
- Sankey, J.B., Germino, M.J., Benner, S.G., Glenn, N.F., Hoover, A.N., 2012. Transport of biologically important nutrients by wind in an eroding cold desert. *Aeolian Research* 7, 17–27.
- Savitzky, A., Golay, M.J., 1964. Smoothing and differentiation of data by simplified least squares procedures. *Analytical Chemistry* 36, 1627–1639.
- Serbin, S.P., Singh, A., McNeil, B.E., Kingdon, C.C., Townsend, P.A., 2014. Spectroscopic determination of leaf morphological and biochemical traits for northern temperate and boreal tree species. *Ecological Applications* 24, 1651–1669. doi:10.1890/13-2110.1
- Serrano, L., Peñuelas, J., Ustin, S.L., 2002. Remote sensing of nitrogen and lignin in Mediterranean vegetation from AVIRIS data: Decomposing biochemical from structural signals. *Remote Sensing of Environment* 81, 355–364. doi:10.1016/S0034-4257(02)00011-1
- Sinsabaugh, R.L., Belnap, J., Rudgers, J., Kuske, C.R., Martinez, N., Sandquist, D., 2015. Soil microbial responses to nitrogen addition in arid ecosystems. *Front Microbiol* 6. doi:10.3389/fmicb.2015.00819
- Skidmore, A.K., Ferwerda, J.G., Mutanga, O., Van Wieren, S.E., Peel, M., Grant, R.C., Prins, H.H.T., Balcik, F.B., Venus, V., 2010. Forage quality of savannas — Simultaneously mapping foliar protein and polyphenols for trees and grass using hyperspectral imagery. *Remote Sensing of Environment* 114, 64–72. doi:10.1016/j.rse.2009.08.010
- Smith, M.L., Martin, M.E., Plourde, L., Ollinger, S.V., 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, 1332–1337. doi:10.1109/TGRS.2003.813128

- Smith, M.L., Martin, M.E., 2001. A plot-based method for rapid estimation of forest canopy chemistry. *Canadian Journal of Forest Research* 31, 549–555. doi:10.1139/cjfr-31-3-549
- Svejcar, T., Angell, R., Bradford, J.A., Dugas, W., Emmerich, W., Frank, A.B., Gilmanov, T., Haferkamp, M., Johnson, D.A., Mayeux, H., Mielnick, P., Morgan, J., Saliendra, N.Z., Schuman, G.E., Sims, P.L., Snyder, K., 2008. Carbon fluxes on North American rangelands. *Rangeland Ecology & Management*, 61, 465-474.
- Townsend, P.A., Serbin, S.P., Kruger, E.L., Gamon, J.A., 2013. Disentangling the contribution of biological and physical properties of leaves and canopies in imaging spectroscopy data. *PNAS* 110, E1074–E1074. doi:10.1073/pnas.1300952110
- United Nations Environment Management Group, 2011. Global drylands: A UN system-wide response. [WWW Document]. URL [http://www.unccd.int/Lists/SiteDocumentLibrary/Publications/Global\\_Drylands\\_Full\\_Report.pdf](http://www.unccd.int/Lists/SiteDocumentLibrary/Publications/Global_Drylands_Full_Report.pdf) (accessed 10.30.16).
- Ustin, S.L., 2013. Remote sensing of canopy chemistry. *PNAS* 110, 804–805. doi:10.1073/pnas.1219393110
- Ustin, S.L., Roberts, D.A., Gamon, J.A., Asner, G.P., Green, R.O., 2004. Using imaging spectroscopy to study ecosystem processes and properties. *AIBS Bulletin*, 54(6) 523-534.
- Ustin, S.L., Roberts, D.A., Pinzón, J., Jacquemoud, S., Gardner, M., Scheer, G., Castañeda, C.M., Palacios-Orueta, A., 1998. Estimating canopy water content of chaparral shrubs using optical methods. *Remote Sensing of Environment* 65, 280–291. doi:10.1016/S0034-4257(98)00038-8
- Ustin, S.L., Smith, M.O., Roberts, D., Gamon, J.A., Field, C.B., 1992. Using AVIRIS images to measure temporal trends in abundance of photosynthetic and nonphotosynthetic canopy components, in: *JPL, Summaries of the Third Annual JPL Airborne Geoscience Workshop. Volume 1: AVIRIS Workshop*; 5-7.
- Wang, Z., Wang, T., Darvishzadeh, R., Skidmore, A.K., Jones, S., Suarez, L., Woodgate, W., Heiden, U., Heurich, M., Hearne, J., 2016. Vegetation indices for mapping canopy foliar nitrogen in a mixed temperate forest. *Remote Sensing* 8, 1–20. doi:10.3390/rs8060491
- Wertin, T.M., Belnap, J., Reed, S.C., 2017. Experimental warming in a dryland community reduced plant photosynthesis and soil CO<sub>2</sub> efflux although the relationship between the fluxes remained unchanged. *Funct Ecol* 31, 297–305. doi:10.1111/1365-2435.12708
- Wessman, C.A., Aber, J.D., Peterson, D.L., Melillo, J.M., 1988. Foliar analysis using near infrared reflectance spectroscopy. *Can. J. For. Res.* 18, 6–11. doi:10.1139/x88-002

- West, N. E., Young, J.A., 2000. Intermountain valleys and lower mountain slopes, in: Barbour M.G., Billings, W.D.. North American Terrestrial Vegetation. Cambridge University Press, New York, 255-284.
- WRCC 2008 LCD for Boise, Idaho, Western Regional Climate Center, 2017. [WWW Document]. URL [https://wrcc.dri.edu/Climate/west\\_lcd\\_show.php?iyear=2008&sstate=ID&stag=boise&sloc=Boise](https://wrcc.dri.edu/Climate/west_lcd_show.php?iyear=2008&sstate=ID&stag=boise&sloc=Boise) (accessed 6.6.17).
- Wright, I.J., Reich, P.B., Westoby, M., Ackerly, D.D., Baruch, Z., Bongers, F., Cavender-Bares, J., Chapin, T., Cornelissen, J.H.C., Diemer, M., Flexas, J., Garnier, E., Groom, P.K., Gulias, J., Hikosaka, K., Lamont, B.B., Lee, T., Lee, W., Lusk, C., Midgley, J.J., 2004. The worldwide leaf economics spectrum. *Nature* 428, 821–827. doi:10.1038/nature02403
- Yang, J., Weisberg, P.J., Bristow, N.A., 2012. Landsat remote sensing approaches for monitoring long-term tree cover dynamics in semi-arid woodlands: Comparison of vegetation indices and spectral mixture analysis. *Remote Sensing of Environment* 119, 62–71. doi:10.1016/j.rse.2011.12.004

## Appendices

### Appendix A. Leaf area estimation

Specific leaf area (SLA) is a ratio measure of foliar weight to area. This is useful in estimating proportional foliar nitrogen (Smith et al., 2001) and was used in three of the scaling methods tested. Leaf samples from six randomly selected shrubs of representative size were collected and weighed twice, first fresh to determine wet weight, and then after drying. Prior to drying, the leaf samples were scanned in groups, as seen in the image below. These images were processed for area estimates as subsequently described.



Figure 17. Leaf subset, single shrub, grouped with label.

The data were received as photographs taken at a measured distance, with one to three groups of 10 - 15 leaves per photo, labeled and on a white background. These were processed in Adobe Photoshop according to a process described in an instructional video by Jarou (2009), in which a collection of sample pixels are manually selected and extrapolated to a larger selection to encompass the selected group of leaves.

Multiplying the selected leaf pixels by dots per inch (DPI) provided leaf area that could be compared to other leaf sample measurements to form metrics used in models. This process has the benefit of precision and human judgement and attention, but is time intensive, requires

training, and is based on subjective judgement which produced results that can vary between researchers. Because of this, the process was conducted twice on each group of leaves and averaged to reduce error, although it did increase the time and labor costs.

To process photos more quickly and reduce subjective judgment, the photos were read into Matlab software using an image processing routine. To segment the multigroup photos, pixels were compared over a three dimensional scaling of hue, saturation, and value.

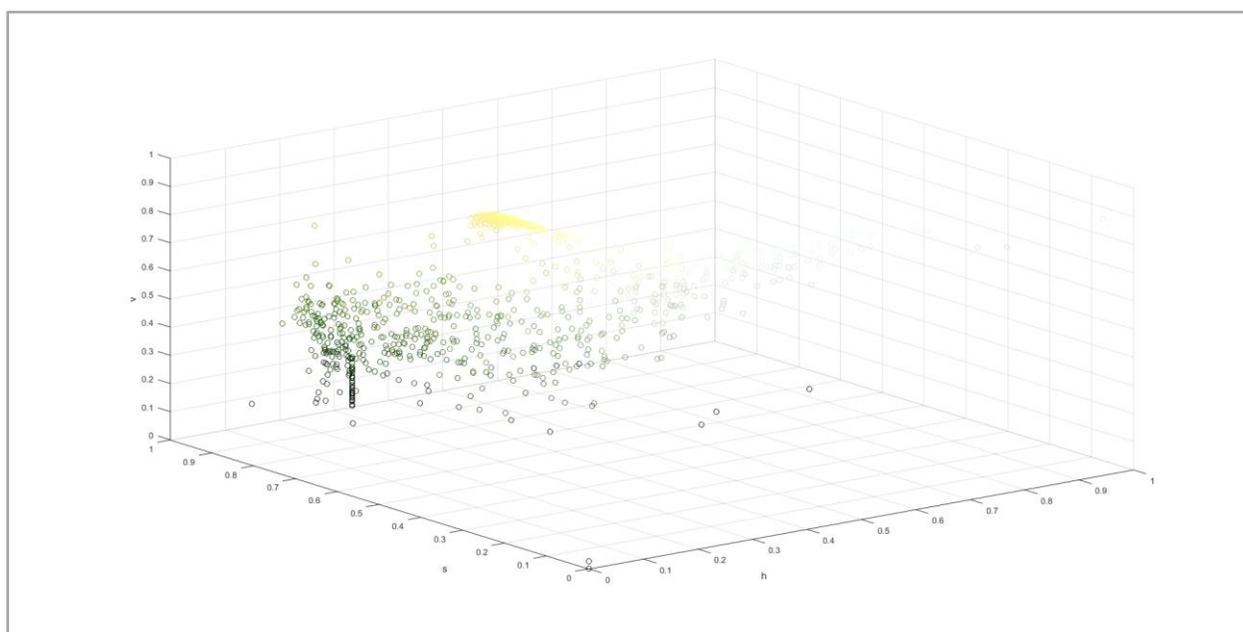
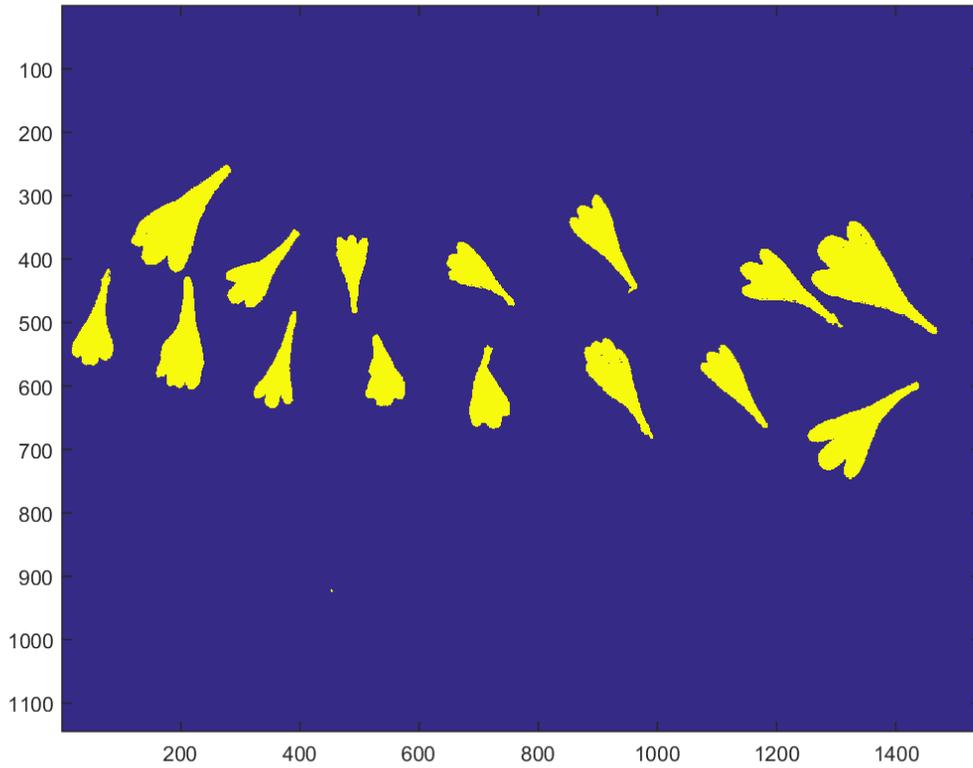


Figure 18. Pixels from leaf scans were projected into two types of multidimensional space.

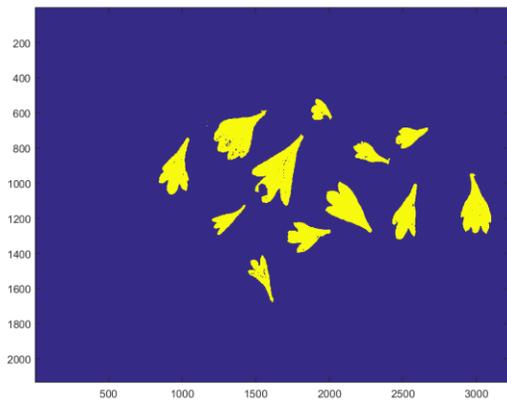
Green pixels were selected and used to mark rows in which groups of leaves were present. Green was defined as a hue value between 0.2 and 0.3, and a saturation between 0.5 and 0.8. Periods of vacant rows were used to mark the boundaries of these groups. Segmented subsets of images were then cropped horizontally along these boundaries. The process was repeated vertically using the same method to exclude the labels.

The leaves were then processed in a three dimensional space of lightness, range from red to green, and range from blue to yellow. Limits were set on these to categorize which pixels belonged to leaves and which to background or shadow. Selection included lightness less than a value of 25 and a red/green ranking of less than 120. Pixels were counted for the leaf category of the binary image produced, seen below, and area was calculated from the pixels.

A.



B.



C.

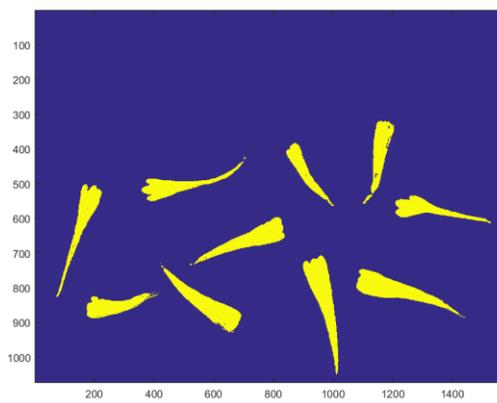


Figure 19. Three images depict the final pixel selections for segmented leaf groups. These are used for pixel counts and saved for manual quality control. Image A, B, and C show the variation

in spacing, shape, and number of leaves. In B, a leaf scar is visible, which was excluded due to the yellow and grey coloring, which was similar to shadow.

The choice of range for the selections was done through visual inspection and trial and error, and introduced human subjectivity and reliance on training data that promoted overfitting, as discussed in more detail below. The program also had a number of segmenting limitations relating to the irregular arrangement and colors of labels used, which require oversight or if possible more regular placement during photography. The program can be set to save the binary images it uses in pixel estimations to allow for quick visual QA/QC.

The automatic estimation was compared to manual estimations through a series of 2-tailed paired sample *t*-tests. The two manual counts that I performed were compared to each other and then to the automatic count. The error between my counts was similar to the error between the program count and manual count. This suggests that using the program is as accurate as having the same researcher take a count at different times, and may offer an efficient, practical alternative.

However, the program did not perform as well on a separate group of leaves. This suggests that the color thresholds should be reset and the performance closely monitored. Developing a GUI for this may be a venue for future work. Code is included below.

## **Appendix B. Leaf area code**

The Matlab code for automatic leaf analysis is included below.

```
% Specific Leaf Analysis  
% Developed by Amanda Smith, Megan Maloney, Dr. Jessica Mitchell, Dr. Mitch Parry of  
Appalachian State University  
% Conceptual development acknowledgement: Zach Jarou, Biva Shresta
```

```

%closes anything open that might be running from other projects
close all;
%clear any variables in the workspace leftover from other projects
clear all;

%sets "cform" variable equal to a colorspace transformation from sRGB into L*a*b, a more
inclusive colorspace whose components are:
%L = lightness
%a = positive for redness and negative for greenness
%b = positive for yellow colors and negative for blue colors
%setting this here as it only needs to be made once, not each iteration
cform = makecform('srgb2lab'); %for leaf selections

%Create a cell array to store counts and names. This is later saved as an Excel file with xlswrite
%We also set headings for the three columns of data we'll be copying over
output = {'Filename', 'Group', 'Pixel Count'};

%set directory name to d, so we don't have to keep retyping it
d = 'C:/Users/MaloneyMC/Downloads/Final_SLA/Final_SLA/';
%The line below reads in the image file that will be processed
files = dir(strcat(d, '*.jpg'));

for i = 1:length(files)
    %defines file as whatever file we're on, based on i
    file = files(i);
    disp(file.name);
    I = imread(strcat(d, file.name));

    [m,n,p]=size(I);
    %figure(12); imagesc(I) %sets figure number to prevent overwriting, shows image
    %set(gcf, 'name', 'Image to be processed');

    %convert to hue
    I_hsv = rgb2hsv(I);

    %Plots 3d scatterplot of pixels on h s v dimensions
    I_hsv = reshape(I_hsv,m*n,p); %reshapes to 3 columns
    % index = randperm(size(I_hsv,1), 10000);
    % I = reshape(I, m*n, p);
    % csB = double(I(index,:));
    % figure(13); scatter3(I_hsv(index,1), I_hsv(index,2), I_hsv(index,3), [], csB/255); %csB sec
uses original colors on points
    % xlabel('h')
    % ylabel('s')
    % zlabel('v')

```

```

I = reshape(I,[m,n,p]); % reshaping from pixel matrix to row/col/rgb

%first, select only green pixels
green = I_hsv(:,1)>.2 & I_hsv(:,1)<.3 & I_hsv(:,3)<.8 & I_hsv(:,3)>.5; %thresholds define
green color
green = reshape(green,m,n); %makes into matrix
figure(14); imagesc(green);
set(gcf, 'name', 'Binary image of groups with break lines');
S = sum(green,2);
S(S < 10) = 0;
S([1,end]) = 0;
%figure(15); plot(S)

w = hann(201); %creates window to fit data into
w = w/sum(w); % fits data in and normalizes, so takes mean rather than sum
filtered =conv(S,w,'same'); %filtered S by window, same size as S....sums 101 values to
smooth curve (convolution)
%figure(15); plot(1:length(S),[S,filtered]) %plots green sum against
%set(gcf, 'name', 'Binary, green pixels by rows');

filtered = [0;filtered]; filtered = [filtered;0];
S = [0;S];
S = [S;0];
binary = logical(filtered); %T/f if value
%figure(15); plot(1:length(S),[S,filtered,binary*100]);

breaks = diff(binary);
breaks(end+1)=0;
%figure(15); plot(1:length(S),[S,filtered,binary*100,(breaks+1)*100]);
%saveas(gcf,strcat(row{1},'_',row{2},'_fig15.png'))

%first start
starts = find(breaks == 1);
stops = find(breaks == -1);

half = round((starts(2:end) - stops(1:end-1))/2 + stops(1:end-1));
figure(14); hold on;
if ~isempty(half)
    plot([0,n],[half(1), half(1)],'w-')
end
if length(half) > 1
    plot([0,n],[half(2), half(2)],'w-')
end
hold off;

```

```

%%crop out post its - move greens
vertcol = sum(green,1); %sum green pixels in each column, 1
vertcol(vertcol < 10) = 0;
vertcol([1,end]) = 0;

%figure(16); plot(vertcol); set(gcf, 'name', 'Binary, green pixels by columns');
w = hann(201); %creates window to fit data into, 101 pixels
w = w/sum(w); % fits data in and normalizes, so takes mean rather than sum
filtered =conv(vertcol',w,'same'); %filtered S by window, same size as S....sums 101 values to
smooth curve (convolution)
%figure(16); plot(1:size(vertcol'),[vertcol',filtered]) %plots green sum against
binary = logical(filtered);
%figure(16); plot(1:size(vertcol'),[vertcol',filtered,binary*100]);
breaks2 = diff(binary);
breaks2(end+1)=0;
%figure(16); plot(1:size(vertcol'),[vertcol',filtered,binary*100,(breaks2+1)*100]);
stops = find(breaks2 == -1);
if length(stops) ~= 1
    stops = stops(1); %this tells it that if there are two (vertical) stops, take the first one. This
MAY clip off a leaf if it's separated from the group. It is designed to deal with "green" pixels
being selected in the post its
end

[pathstr,name,ext] = fileparts(file.name) %breaks input image file name into part to separate
name from extension etc

edge = stops + n*.005;
figure(14); hold on;
plot([edge,edge], [0,m], '-w');
hold off;
saveas(gcf,strcat(d,name,'_fig14_breaklines.png'))

groups = [1, half', m];
for j = 1:length(groups)-1
    %defines a row with three things to be added to output (filename, group, pixel count)
    row = cell(1,3);
    row{1} = name
    row{2} = j;
    J = I(groups(j):groups(j+1), 1:ceil(edge), :); %create subimage
    lab_he = applycform(J, cform); %applies the cform colorspace to image
    [m, n, p] = size(lab_he); %creates an array of transformed image, assigning the size of
    %each dimension to a separate variable

%resets to three columns L a b
lab_he = reshape(lab_he, m*n, p);

```

```

%%Begin classifications:

%This section looks at the second col a, which is red vs green and checks that red-green is >
120.
%(120 was selected manually from a visual examination of the scatterplot)
%The below lines select more reddish pixels (warm yellow tones in shadow, white
background) for exclusion.
%shadow = lab_he(:,2) > 120;
%figure(7); imagesc(reshape(shadow,m,n));

%The below lines select greenish pixels for inclusion
leaf = lab_he(:,2) < 120; %Reminder: this is in LAB now, not HSV - make new scatterplot
to view.
count = sum(leaf);
%figure(11); imagesc(reshape(leaf,m,n));
%saveas(gcf,strcat(row{1},'_',num2str(row{2}),'_fig11_leafbinary.png'))

%looks at first col checks if lightness is >25
%This catches black values that are actually leaf.
black = lab_he(:,1) < 25;
count2 = sum(black);
%figure(9); imagesc(reshape(black,m,n));

%adds black and green selections for the leaf pixel counts
totalcount = count + count2;
%Store with name
row{3} = totalcount;

%append output cell to array that will be written to the output file
output = [output;row];

%%final figure, related to totalcount
figure(19); imagesc(reshape(black | leaf,m,n));
set(gcf, 'name', 'Final pixel selection for count');
saveas(gcf,strcat(d,num2str(row{1}),'_group',num2str(row{2}),'_fig11_leafbinary.png'))
end

%% Turn on below line to show what's being selected as shadow.
%figure(4); imshow(reshape(white & ~shadow, m, n));

%%Turn on below line to see what's selected as
%figure(5); imagesc(reshape(new_image,m,n));

end

```

```
% Writes the collected rows in 'output' to an Excel file, located in same
% folder
xlswrite(strcat(d, 'PixelCount.xlsx'), output);
```

## **Appendix C. Lidar fusion**

Lidar datasets be used to derive metrics such as canopy relief ratio, foliage height diversity, and texture of vegetation heights which can then be used for classification (Glenn et al, 2010) and characterization of canopy structural measurements such as height (Mitchell et al, 2011), crown shape, vertical layer, and position in terrain (Asner et al, 2007). These metrics can be related to estimate age of vegetation, cover area, use as wildlife habitat, fuel loading, erosion, infiltration, evapotranspiration, disturbance history, and biomass, which are related to ecosystems' resilience, vulnerability, and ability to provide services. Lidar measures of vegetative structure may directly relate to ecosystem dynamics, such as nitrogen cycling (Moorcraft et al, 2001). Lidar has additional value in improving the analysis of hyperspectral data, such as estimations of chlorophyll (Gokkaya et al, 2014; Simic et al, 2009; Thomas et al, 2006), photosynthetic pigments (Blackburn, 2002) and absorbed photosynthetically active radiation (fPAR) (Thomas et al, 2006) by masking noise (Blackburn, 2002) or providing relevant structural information to stratify spectrally similar groups (Koetz et al, 2007). Common synergies of lidar and hyperspectral include adjusting for the effects of vegetative structure on reflectance and the expression of biochemical and physical traits within the canopy, and using hyperspectral data to characterize lidar estimates of biomass through functional type and for water content (Asner et al, 2007).

In sagebrush specifically, integrating lidar and hyperspectral imagery improved classification accuracy of sagebrush from 74% to 89% (Mundt et al, 2006). As foliar N is closely

related to growth and structure, additional lidar metrics may also improve hyperspectral detection. When lidar is combined with hyperspectral imaging spectroscopy in other ecosystems, Luo et al. reported overall classification accuracies improved by 9.1% over multispectral data alone and 19.6% over lidar data alone (Luo et al, 2016a). In a separate forest biomass estimation project, compared to lidar alone the “fused lidar and hyperspectral data improved  $R$ -squared ( $R^2$ ) by 5.8%, 2.2% and 2.6%...and reduced RMSE by 8.6%, 7.9% and 8.3%” for below ground, above ground, and total biomass respectively (Luo et al, 2016b). In a shrub ecosystem, Riano et al. were able to use orthoimagery to improve shrub height predictions from lidar significantly (2007). Fusion of lidar and imagery offers complementary strengths that improve results.

### C.1 Lidar collection

Lidar (Leica ALS50II) data were collected from an airborne platform for the entire RCEW site in 2014. This was during the same phenological period in which the hyperspectral data were collected (Glenn, et al - proposal). Lidar data were pre-processed by Nayani Ilangakoon at Boise State University; vegetation products include height intervals, canopy relief ratio, foliage height diversity, and texture of vegetation heights.

### C.2 Lidar data preprocessing

Pre-processed lidar data from 2015, containing 35 vegetation metrics at 1 and then 5 meter resolution, were extracted for 2014 sites and 2015 sites. These metrics include height intervals, canopy relief ratio, foliage height diversity, and texture of vegetation heights.

However, addition of lidar variables did not improve performance consistently in early trials. Lidar metrics were compared to average foliar percent N content in simple linear

regression (Table 9). Height related variables showed some correlation and potential use in the Height/Cover method, which number to total and ground returns correlated to Cover/LMA.

Method	Lidar variable	P value	Corr. Coef. (R)	R <sup>2</sup> (corr coef <sup>2</sup> )
Biomass/ Density (live)	MAD height	.0686	.405	0.16
	25 <sup>th</sup> percentile height	.0840	.386	0.15
	50 <sup>th</sup> percentile height	.0616	.415	0.17
Height/Cover	Max height	.0235	.492	0.24
	Range height	.0238	.491	0.24
	Mean height	.0321	.469	0.22
	MAD height	.0398	.452	0.20
	AAD height	.0302	.473	0.22
	Variance height	.0267	.483	0.23
	Standard deviation height	.0249	.488	0.24
	Interquartile range height	.0392	.453	0.21
	25 <sup>th</sup> per height	.0624	.414	0.17
	50 <sup>th</sup> per height	.0400	.451	0.20
	75 <sup>th</sup> per height	.0394	.453	0.21
	90 <sup>th</sup> per height	.0399	.452	0.20
	95 <sup>th</sup> per height	.0187	.508	0.26
	Per ground returns	.0337	.465	0.22
	Veg % 0-1m in height	.0771	-.394	0.16
Cover/LMA	Number of returns	.0071	.568	0.32
	Number of ground returns	.0099	.549	0.30

Table 9. Correlation of lidar variables to scaled foliar N estimates.

#### Appendix D. Code for foliar N prediction using PLSR coefficients

Below is the R code for processing preprocessed, transformed, and smoothed spectral input into foliar N maps using beta coefficients of a model corresponding to the input imagery.

```
setwd("L:/Research/maloneyMC/N") #Replace with address of desired workspace folder.  
getwd() #Check workspace.
```

```
rm(list=ls(all=TRUE)) #Clear workspace  
graphics.off() #Clear graphics
```

```
#Install packages and libraries as needed  
library(rgdal)  
#install.packages("raster")  
library(raster)  
#install.packages("signal")  
library(signal)  
library(prospectr)  
install.packages("pracma")  
library(pracma)  
install.packages("pls")  
library(pls)  
install.packages("xlsx")  
library(xlsx)
```

```
#get weighted beta coefficients into formats for use  
beta_c<-read.xlsx("C:/Users/MaloneyMC/Documents/ASU Winter  
2017/Thesis/Final_model_wout1208.xlsx", 1) #Replace with desired file address.  
yint<- beta_c[1,2] #Assign y integer  
beta_c<- beta_c[2:nrow(beta_c),] #Truncate y integer from coefficients  
waveschar <- as.character(beta_c[,1]) #Convert to character for string editing  
typeof(waveschar) #check type  
wavescharshort <- gsub("\\.+", "", waveschar) #Edit string  
beta_t <- t(beta_c) #Transpose  
beta_t <- rbind(beta_t, wavescharshort) #Reattach edited names  
colnames(beta_t) = beta_t[3, ] # Assign third row as column headers  
beta_t = beta_t[-1, ] #Delete redundant first row  
beta_t = beta_t[-2, ] #Delete redundant second row  
beta_t <- t(as.matrix(beta_t)) #Change to matrix to view easily in RStudio to check data  
colnames(beta_t) <- paste0("X", colnames(beta_t)) #May be required  
test <- colnames(beta_t) #Make list for selection
```

```

sagebrick=brick(choose.files()) #Choose a hyperspectral brick
names(sagebrick) <- gsub(".*img...", "", names(sagebrick)) #Edit names for selection
names(sagebrick) <- gsub("\\.*", "", names(sagebrick)) #Edit names
sagedrop <- subset(sagebrick, test, drop=TRUE) #Drop layers lacking a matching coefficient

#Generate a rasterlayer obj to store output values
Nitro <-raster(sagedrop)

#Apply coefficients
for (i in 1:nlayers(sagedrop))
{
  sagelayer <- raster(sagedrop, layer=i)
  beta_coef <- as.numeric(beta_t[i])
  newlayer <- sagelayer*beta_coef
  Nitro <- addLayer(Nitro, newlayer)
}

Nsum <- calc(Nitro, sum) #Sum layers which were multiplied by coefficients
NsumY <- Nsum+yint #Add y integers

#Write in chosen format
writeRaster(NsumY, filename = "Nitrogen_203933ENVI", format="ENVI", overwrite=TRUE)
writeRaster(NsumY, filename = "Nitrogen_200350_no1208dGTiff", format="GTiff",
overwrite=TRUE)
writeRaster(Nitro, filename = "NitroStack_200350_no1208ENVI", format="ENVI",
overwrite=TRUE)
#####

```

## Appendix E. Bands from Curran (1989)

Table 1. Forty-two Absorption Features in Visible and Near-Infrared Wavebands That Have Been Related to Particular Foliar Chemical Concentrations

Wave length ( $\mu\text{m}$ )	Electron Transition or Bond Vibration	Chemical(s)	Remote Sensing Considerations
0.43	Electron transition	<b>Chlorophyll a</b> <sup>a</sup> <b>Chlorophyll b</b> <b>Chlorophyll b</b> <b>Chlorophyll a</b>	Atmospheric scattering
0.46	Electron transition		
0.64	Electron transition		
0.66	Electron transition		
0.91	C—H stretch, 3rd overtone	Protein	
0.93	C—H stretch, 3rd overtone	Oil	
0.97	O—H bend, 1st overtone	<b>Water</b> , starch	
0.99	O—H stretch, 2nd overtone	Starch	
1.02	N—H stretch	Protein	
1.04	C—H stretch, C—H deformation	Oil	
1.12	C—H stretch, 2nd overtone	Lignin	
1.20	O—H bend, 1st overtone	<b>Water</b> , cellulose, starch, lignin	
1.40	O—H bend, 1st overtone	<b>Water</b>	
1.42	C—H stretch, C—H deformation	Lignin	
1.45	O—H stretch, 1st overtone, C—H stretch, C—H deformation	Starch, sugar, lignin, water	Atmospheric absorption
1.49	O—H stretch, 1st overtone	Cellulose, sugar	
1.51	N—H stretch, 1st overtone	<b>Protein, nitrogen</b>	
1.53	O—H stretch, 1st overtone	Starch	
1.54	O—H stretch, 1st overtone	Starch, cellulose	
1.58	O—H stretch, 1st overtone	Starch, sugar	
1.69	C—H stretch, 1st overtone	<b>Lignin</b> , starch, protein, nitrogen	
1.78	C—H stretch, 1st overtone/ O—H stretch/H—O—H deformation	<b>Cellulose, sugar, starch</b>	
1.82	O—H stretch/C—O stretch, 2nd overtone	Cellulose	
1.90	O—H stretch, C—O stretch	Starch	
1.94	O—H stretch, O—H deformation	<b>Water</b> , lignin, protein, nitrogen, starch, cellulose	Atmospheric absorption
1.96	O—H stretch/O—H bend	Sugar, starch	
1.98	N—H asymmetry	Protein	
2.00	O—H deformation, C—O deformation	Starch	
2.06	N=H bend, 2nd overtone/ N=H bend/N—H stretch	Protein, nitrogen	
2.08	O—H stretch/O—H deformation	Sugar, starch	
2.10	O=H bend/C—O stretch/ C—O—C stretch, 3rd overtone	<b>Starch</b> , cellulose	
2.13	N—H stretch	Protein	
2.18	N—H bend, 2nd overtone/ C—H stretch/C—O stretch/ C=O stretch/C—N stretch	<b>Protein, nitrogen</b>	
2.24	C—H stretch	Protein	
2.25	O—H stretch, O—H deformation	Starch	Rapid decrease in signal-to-noise ratio of sensors
2.27	C—H stretch/O—H stretch CH <sub>2</sub> bend/CH <sub>2</sub> stretch	Cellulose, sugar, starch	
2.28	C—H stretch/CH <sub>2</sub> deformation	Starch, cellulose	
2.30	N—H stretch, C=O stretch, C—H bend, 2nd overtone	Protein, nitrogen	
2.31	C—H bend, 2nd overtone	<b>Oil</b>	
2.32	C—H stretch/CH <sub>2</sub> deformation	Starch	
2.34	C—H stretch/O—H deformation/ C—H deformation/O—H stretch	Cellulose	
2.35	CH <sub>2</sub> bend, 2nd overtone, C—H deformation, 2nd overtone	Cellulose, protein, nitrogen	

Sources: Rotolo (1979), Pacific Scientific (1984), Peterson et al. (1985), Osborne and Fearn (1986), Williams and Norris (1987), Card et al. (1988), Elvidge (1990), and unpublished regional reports from nine USDA near-infrared spectroscopy workshops.

<sup>a</sup>Chemicals in bold have a wavelength of stronger absorption.

Appendix F. Annotated graphs of model coefficients along the spectrum



Figure 10. Biomass/Density model coefficients displayed with annotation of vegetative spectral features.

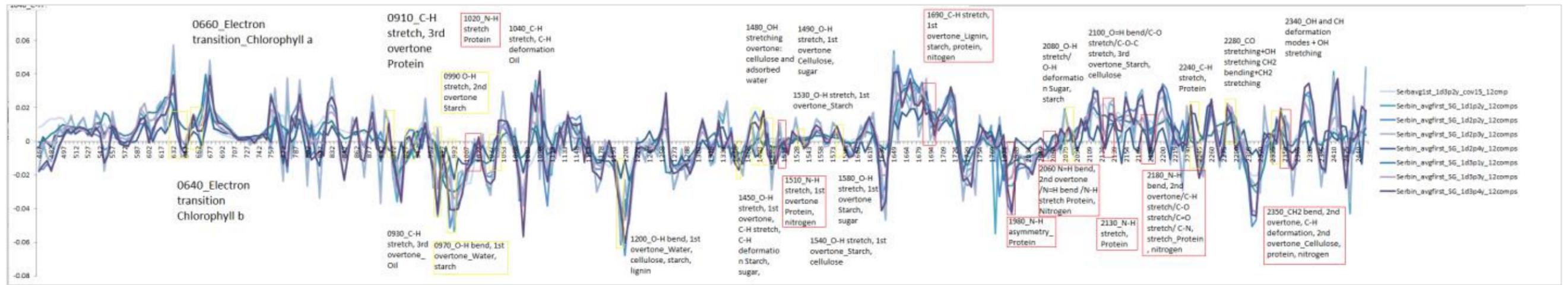


Figure 11. Cover/LMA model coefficients displayed with annotation of vegetative spectral features.

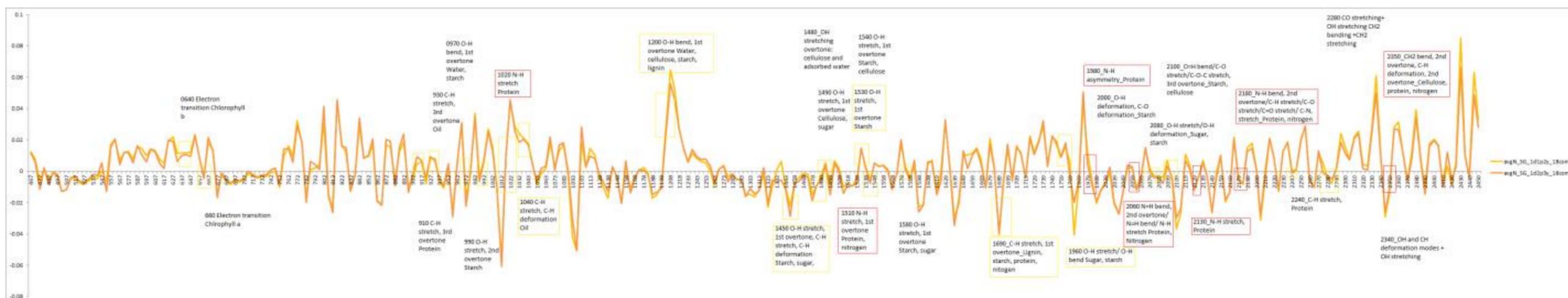


Figure 12. Average N values from lab sampling, which were used in multiple scaling methods, were applied without scaling in PLSR to two SG transformations.

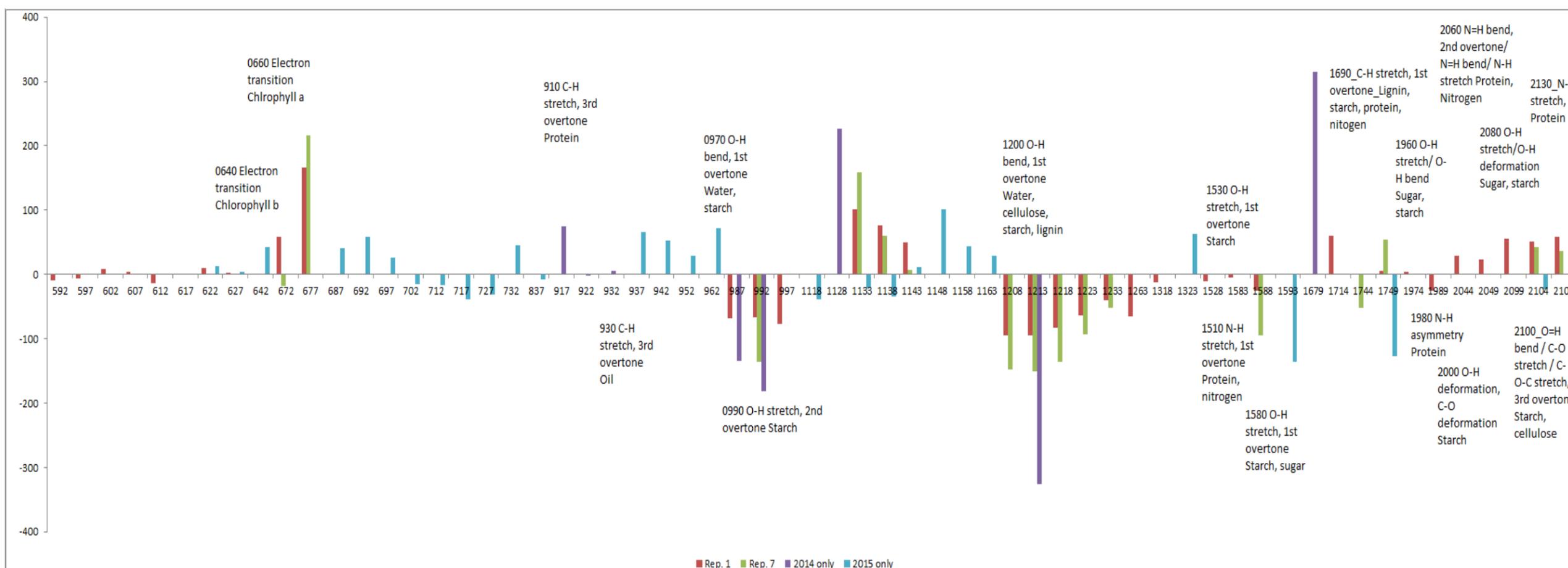


Figure 15. Selected significant wavelengths for two repetitions using all samples and two subsets using either only 2014 or 2015 samples, annotated with absorbance features from Curran (1989).

## **Vita**

Megan Maloney grew up exploring the Appalachian Mountains of North Carolina with her adventurous family and friends, which sparked her interest in ecology, mapping, and exploration. She graduated from Asheville High School in 2004 and attended Sweet Briar College in Virginia. In 2008, Megan earned her Bachelor of Arts degree in Environmental Studies with a minor in Philosophy. She immediately went to work with conservation and sustainable community planning nonprofits in Asheville, NC, and Manomet, MA. Megan became a research associate with the Climate Change Science Institute and Geographic Information Science & Technology group at Oak Ridge National Laboratory. For four years, Megan supported projects investigating coastal risk assessment and population dynamics. She then pursued a research assistantship at Appalachian State University in Boone, NC, where she used fieldwork, GIS, and remote sensing with lidar and imaging spectroscopy to investigate ecological questions relating to sustainable land management. Megan received a Master of Arts degree in 2017 from the Department of Geography and Planning.