2015

Advances in measuring forest structure by terrestrial laser scanning with the Dual Wavelength ECHIDNA® LIDAR (DWEL)

https://hdl.handle.net/2144/14063

Boston University

BOSTON UNIVERSITY

GRADUATE SCHOOL OF ARTS AND SCIENCES

Dissertation

ADVANCES IN MEASURING FOREST STRUCTURE BY TERRESTRIAL LASER SCANNING WITH THE DUAL WAVELENGTH ECHIDNA® LIDAR (DWEL)

by

ZHAN LI

B.S., Nanjing University, 2008
M.S., Graduate School of Chinese Academy of Sciences, 2011

Submitted in partial fulfillment of the
requirements for the degree of
Doctor of Philosophy

2015

Approved by

First Reader

Alan H. Strahler, Ph.D.
Professor of Earth and Environment

Second Reader

Crystal L. Schaaf, Ph.D.
Adjunct Research Professor of Earth and Environment

Third Reader

Curtis E. Woodcock, Ph.D.
Professor of Earth and Environment

DEDICATION

I would like to dedicate this work to my patient, caring and supportive parents.

献给永远耐心，关心和支持我的父母

ACKNOWLEDGMENTS

I would like to sincerely thank all the people who have helped and accompanied me through the journey to this thesis finish. First I have to thank my advisor, Dr. Alan H. Strahler for his tremendous time and efforts in advising me extensively from fostering my scientific attitude and vision to inspiring my ideas of solving data puzzles and writing scientific articles. I also have to thank Dr. Crystal L. Schaaf for her continuous and countless support in every field trip, every science meeting and every stage of the research project. I have been lucky to work with Alan and Crystal. This thesis would not have been possible without their guidance and support.

Special thanks go to Dr. David L. B. Jupp for his patient and meticulous explanations on the theories and practices about lidar applications to forests. I am also grateful to Dr. Curtis E. Woodcock and Dr. Lucy. R. Hutyra for their insightful comments and advices that have improved my work.

I would not have been able to use the DWEL instrument in reality without the engineering efforts from the instrument building team, particularly Dr. Supriya Chakrabarti, Dr. Timothy A. Cook, Ewan S. Douglas, Jason Martel, Glenn Howe, and Kuravi Hewawasam. I have learned a great deal from all the discussions and troubleshooting over the instrument with them.

The TLSIIG group has been a collaborative and resourceful team for my work. In particular, I would like to thank Dr. John Armston for his timely help on coding and data processing tools, Dr. Michael Schaefer for his cooperative efforts on the DWEL instruments, Jasmine Muir for her genuine assistance of field trips, Dr. Steven Hancock for his valuable discussions on radiative transfer theories, and Dr. Pasi Raumonen for his generous sharing of the great modeling tools.

I deeply appreciate the enormous support and assistance in every walk of my Ph.D. study and life from my current and former colleagues, Dr. Zhuosen Wang, Dr. Qingsong Sun, Dr. Xiaoyuan Yang, Dr. Tian Yao, Ian Paynter, Edward J. Saenz, Yan Liu, Angela Erb, JiHyun Kim, Shabnam Rouhani, Francesco Peri, and Yiming Chen.

I would like to thank Dr. Yuri Knyazikhin, Dr. Michael Dietz, Dr. Nathan Phillips and Dr. Pamela Templer for sharing their expertise with me. I appreciate the memorable time with the current and former graduate students in the department at BU, particularly Dr. Xiaoman Huang, Dr. Zhe Zhu, Ha Nguyen, Xiaojing Tang, Shixiong Wang, Sungho Choi, Dr. Jian Bi, Margaret Hendrick, Christopher Holden, Dan Gianotti, Jonathan Wang, and Taejin Park. I would also like to thank Emily Johnson, Frederick George, Nora J. Watson, Sayaka Yamaki, Josh Parsons and Christopher DeVits who have been very helpful to solve all the administrative hassles. I also acknowledge National Science Foundation for their funding (MRI-0923389) that has supported the DWEL instrument building and my research work.

Last but not least, I give my earnest and deepest thanks to my always patient, caring and supportive parents. I dedicate this thesis to them.

ADVANCES IN MEASURING FOREST STRUCTURE BY TERRESTRIAL LASER SCANNING WITH THE DUAL WAVELENGTH ECHIDNA® LIDAR (DWEL)

(Order No.)

LI, ZHAN

Boston University Graduate School of Arts and Sciences, 2015

Major Professor: Alan H. Strahler, Professor of Earth and Environment

ABSTRACT

Leaves in forests assimilate carbon from the atmosphere and woody components store the net production of that assimilation. Separate structure measurements of leaves and woody components advance the monitoring and modeling of forest ecosystem functions. This dissertation provides a method to determine, for the first time, the 3-D spatial arrangement and the amount of leafy and woody materials separately in a forest by classification of lidar returns from a new, innovative, lidar scanner, the Dual-Wavelength Echidna® Lidar (DWEL). The DWEL uses two lasers pulsing simultaneously and coaxially at near-infrared (1064 nm) and shortwave-infrared (1548 nm) wavelengths to locate scattering targets in 3-D space, associated with their reflectance at the two wavelengths. The instrument produces 3-D bispectral "clouds" of scattering points that reveal new details of forest structure and open doors to three-dimensional mapping of biophysical and biochemical properties of forests.

The three parts of this dissertation concern calibration of bispectral lidar returns; retrieval of height profiles of leafy and woody materials within a forest canopy; and virtual reconstruction of forest trees from multiple scans to estimate their aboveground woody biomass. The test area was a midlatitude forest stand within the Harvard Forest, Petersham, Massachusetts, scanned at five locations in a 1-ha site in leaf-off and leaf-on conditions in 2014. The model for radiometric calibration assigned accurate values of spectral apparent reflectance, a range-independent and instrument-independent property, to scattering points derived from the scans. The classification of leafy and woody points, using both spectral and spatial context information, achieved an overall accuracy of 79±1% and 75±2% for leaf-off and leaf-on scans, respectively. Between-scan variation in leaf profiles was larger than wood profiles in leaf-off seasons but relatively similar to wood profiles in leaf-on seasons, reflecting the changing spatial heterogeneity within the stand over seasons. A 3-D structure-fitting algorithm estimated wood volume by modeling stems and branches from point clouds of five individual trees with cylinders. The algorithm showed the least variance for leaf-off, woody-points-only data, validating the value of separating leafy and woody points to the direct biomass estimates through the structure modeling of individual trees.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

LIST OF ABBREVIATIONS

AGB	Aboveground Biomass
ALS	Airborne Laser Scanner/Scanning
CSIRO	Commonwealth Scientific and Industrial Research Organization
DBH	Diameter at Breast Height
DHP	Digital Hemispherical Photography
DN	Digital Number or Digital Counts
DWEL	Dual Wavelength Echidna Lidar
EVI	Echidna Validation Instrument
EW	Equivalent Width
FWHM	Full Width at Half Maximum
LAI	Leaf Area Index
LAVD	Leaf Area Volume Density
NDI	Normalized Difference Index
NEP	Net Ecosystem Productivity
NIR	Near Infrared
PAI	Plant Area Index
PAVD	Plant Area Volume Density
PRF	Pulse Repetition Frequency
QSM	Quantitative Structure Model
SWIR	Shortwave Infrared
TLS	Terrestrial Laser Scanner/Scanning

TRAC	Tracing Radiation and Architecture of Canopies
VAI	Vegetation Area Index
VAVD	Vegetation Area Volume Density
WAI	Woody Area Index
WAVD	Woody Area Volume Density

CHAPTER 1. INTRODUCTION

Understanding the carbon cycle and energy balance of the Earth is critical to project the climate change impacts and devise proper mitigation policies. The role of terrestrial vegetation is especially important in the carbon cycle because of its active, large and continuous exchange of carbon, water and energy between the land and the atmosphere (Sellers et al., 1997). Forests, about one third of the Earth's land surface cover, sequester and store more carbon than any other terrestrial ecosystem (Bonan, 2008; Gibbs et al., 2007). Forest structure interrelates with forest functioning and significantly affects carbon assimilation and storage of forest ecosystems (Baldocchi et al., 2002). The primary objective of the studies presented in this dissertation is to evaluate the quantification of forest structure using new dual-wavelength terrestrial lidar technology. This chapter (1) overviews the roles of two important forest structural measures, leaf area and aboveground biomass that will be retrieved by lidar; (2) discusses approaches to forest structure measurements and specifically those using terrestrial lidar for ground-based measurements; (3) introduces the Dual-Wavelength Echidna Lidar, a novel terrestrial lidar instrument we used in this study for forest structure measurements; and (4) presents the objectives and structure of the dissertation.

1.1 Importance of Leaf Area and Aboveground Biomass

Forest structure refers to the size (length, area, volume and mass) and spatial arrangement of vegetation elements (leaves, shoots, branches, stems and roots) in forests. Forest structures in aboveground (mainly leaves and stems/branches) and belowground (mainly roots) systems are often approached separately, due to the clear separation of their medium (atmosphere and soil). For this work, we are only concerned with aboveground structure. Leaf area index (LAI) and aboveground biomass (AGB) are probably the two most important forest structure measures because leaf area determines the potential of carbon assimilation and evapotranspiration, and biomass measures the amount of carbon in storage and, over time, its change.

Leaf area index (LAI), defined as half of leaf surface area (simply referred as ‘leaf area’) per unit ground area (Chen & Black, 1992), is one of the primary forest structural parameters that governs the radiation interception in forest canopies and the capacity for photosynthesis. The LAI is used in the ecophysiological models to predict forest carbon uptake and also in the remote sensing approach to observe forest carbon uptake over large areas (Baldocchi et al., 2002; Bonan, 1993; Hanson et al., 2004; Medvigy et al., 2009; Running & Coughlan, 1988). Besides LAI, estimates of net ecosystem productivity (NEP) in ecophysiological modeling have also shown the importance of including the detailed three-dimensional distribution of leaf areas, e.g. the vertical foliage profile, especially for open canopies and multi-layered stands (Law et al., 2001). Measurements of the vertical foliage profile are strongly related to (Parker et al., 2004) and significantly correlated with (Stark et al., 2012) forest carbon dynamics, further demonstrating the importance of the three-dimensional distribution of leaf area.

Vegetation biomass, particularly aboveground biomass (AGB), is exposed to natural and human-induced disturbance and thus controls the amount of carbon emission caused by disturbance to ecosystems (Houghton et al., 2009; Houghton, 2007). Moreover, biomass also affects ecosystem functioning through its control over the magnitude and rate of autotrophic respiration (Houghton et al., 2009; Turner et al., 2005). About 70% - 90% of terrestrial vegetation biomass is in forests, of which about 70% - 90% is in aboveground forest biomass (Cairns et al., 1997; Houghton et al., 2009). Therefore, measurement of forest aboveground biomass is critical information for monitoring and predicting carbon emission and uptake from forest ecosystems in the global carbon balance.

1.2 Measurements of Forest Structure

1.2.1 Measurement at coarser scales

Over large areas from regional to global scales, LAI has been mainly derived from the inversion of radiative transfer models from optical remotely sensed imagery (Koetz et al., 2005; Myneni et al., 2002), and by using empirical relationships between field-measured LAI and various metrics from remote sensing data. The metrics have included vegetation indexes from optical imagery (Cohen et al., 2003); backscatter cross section and polarization ratio from synthetic aperture radar (Manninen et al., 2005); laser return energy and canopy heights from airborne and spaceborne lidar (Korhonen et al., 2011; Lefsky et al., 1999; Luo et al., 2013; Solberg et al., 2009); and combinations of the above (Ma et al., 2014). Airborne and spaceborne full-waveform lidar data have demonstrated the capability of deriving vertical foliage profiles over large areas, as either relative (Lefsky et al., 1999) or absolute quantities (Harding et al., 2001; Ni-Meister et al., 2001; Tang et al., 2014, 2012).

Current approaches to large-scale AGB estimation and mapping mainly extrapolate plot-level AGB measurements on the ground over landscapes based on other spatially explicit datasets, such as existing maps of environmental variables (e.g. geolocation, climate, soils, topography, land use, tree cover, etc.) (Brown et al., 1999; Houghton et al., 2001; Sales et al., 2007), and/or remotely sensed data and metrics (e.g. spectral reflectance, vegetation indexes, backscatter cross section, lidar waveform metrics and etc.) (Baccini et al., 2012; Saatchi et al., 2011, 2007).

1.2.2 Measurements on the ground

Most approaches to large-area measurements of both LAI and AGB require ground-based measurements for calibration and all approaches need ground-based measurement to validate their LAI and AGB estimates. Ground-based measurements of LAI and AGB are thus critically important.

Ground-based LAI measurement techniques generally fall into two categories: direct, which involves destructive sampling or litter-fall collection, and indirect, which involves tree allometry, or gap probability measurements (Asner et al., 2003; Bréda, 2003; Chen et al., 1997; Jonckheere et al., 2004). Direct destructive sampling and litter-fall collection are time-consuming, laborious and sometimes impractical, e.g. in areas of preservation or scientific interest where destructive sampling is prohibited (Jonckheere et al., 2004). Also litter-fall collection only works for deciduous trees. The indirect allometry-based method requires site-specific allometric equations from detailed destructive sampling. Indirect gap-based methods invert gap probability in the canopy to LAI based on Beer’s law of light transmission through an absorbing medium (Chen et al., 1997; Lang, 1987; Welles & Cohen, 1996). The gap probability can be obtained by a point quadrat method using a thin probe that passes through the canopy (Wilson, 1960) or noncontact optical techniques such as hemispherical photos (HP), radiation measurement devices (e.g. DEMON, ceptometers, LAI-2000 and TRAC), and ground-based laser technology (Bréda, 2003; Jonckheere et al., 2004; Jupp et al., 2009; Zhao et al., 2012). The point quadrat method is impractical in forest stands because of the tall stature of trees and the high density of leaves (Chen et al., 1997). Indirect noncontact optical methods are quick, low-cost and more commonly used. Direct inversion by Beer's law typically underestimates LAI due to clumping of leaves or overestimates LAI due to lack of differentiation between leaves and woody materials. Clumping effects can be corrected with an index derived from gap size theory with the TRAC instrument (Tracing Radiation and Architecture of Canopies) (Chen et al., 1997). The contribution of woody material to LAI measurements is usually removed with an empirical woody-to-total ratio. Kucharik et al. (1998) found the nonrandom positioning of branches/stems with regard to leaves causes inaccurate LAI with this simple ratio correction, especially when branches/stems are not preferentially shaded by leaves. They removed the woody contribution directly with a Multiband Vegetation Imager. But this approach cannot correct vertical foliage profiles to remove the woody contribution. The separation of leaves from woody materials in 3-D space is needed to remove the woody contribution to vertical foliage profiles derived from gap probability measurements.

The most direct and accurate way to measure AGB is destructive sampling, which is prohibitively laborious, time-consuming, and sometimes impractical (Gibbs et al., 2007). Another commonly used indirect approach to ground-based biomass measurements relates AGB of a tree with more easily measured quantities such as diameter at breast height (DBH) and/or tree height, using allometric equations from destructive sampling. Although extensive compilations of allometric equations have covered many important species across several continents and biomes (Chave et al., 2005; Jenkins et al., 2004; Zianis & Seura, 2005), allometric equations are still large sources of error in AGB estimation (Basuki et al., 2009; Chave et al., 2004; Keller et al., 2001; Wang, 2006), due to inadequacies imposed by the time-consuming, laborious, and costly development and update of allometric equations through destructive sampling. The errors in AGB estimates using allometric equations may be particularly high for large trees, which have large amounts of biomass but are rare in destructive sampling (Basuki et al., 2009; Chave et al., 2004; Keller et al., 2001).

1.2.3 Terrestrial Lidar for forest structure measurements

Lidar, an active remote sensing technology, has been widely explored in forest structure measurements because of its accurate range measurements using lasers and 3-D imaging capability from laser scanning. Two technologies, pulsed and continuous-wave (CW), are used by lidar to measure ranges. Here we address only pulsed laser technology, which is more commonly used in forest applications (Lim et al., 2003).

At coarser scales, airborne and spaceborne lidars have shown the ability to directly measure canopy height, subcanopy topography, and intercepted surface profile (Dubayah & Drake, 2000) and empirically or semi-empirically estimate forest structural parameters such as aboveground biomass, basal area, mean stem diameter, foliage cover, gap probability, projected foliage profile, and leaf area index (Armston et al., 2013;

Baccini et al., 2012; Drake et al., 2002; Harding et al., 2001; Korhonen et al., 2011; Lefsky et al., 1999; Tang et al., 2012).

At tree or plot scales, terrestrial lidar (or terrestrial laser scanning, TLS) has shown great ability to provide fast, accurate and detailed forest structure measurements that can be used in precision forestry, novel ecological applications, and ground truth data collection for the calibration and validation of large-area forest structure measurements (Dassot et al., 2011). Compared with airborne and spaceborne lidar, terrestrial lidar has a much higher sampling resolution and provides a direct and 3-D view of the lower canopy and shrub layer through its scanning capability. Thus, those parameters that are most directly related to aboveground biomass, such as DBH and stem count density can be estimated more easily and more accurately with TLS. Additionally, scanning at various zenith angles by terrestrial lidar provides multi-angle gap probabilities that are essential for LAI and vertical foliage profile estimates, but are not provided by current airborne and spaceborne lidar instruments.

LAI and vertical foliage profile have been retrieved from terrestrial laser scanning data for plots and stands through gap probability inversion (Calders et al., 2015; Jupp et al., 2009; Zhao et al., 2011) or individual trees through voxel-based contact frequency methods (Béland et al., 2011; Hosoi & Omasa, 2006). Use of TLS data to estimate LAI from gap probability shares the same basic principle used with traditional passive optical method such as hemispherical photography and radiation-measuring devices (Chen et al., 1997; Lang, 1987; Welles & Cohen, 1996). But the ranging information in TLS data enables easy retrieval of gap probability as a function of distance, or over range, and thus provides a vertical foliage profile. There are currently two major techniques for retrieving gap probability over range from TLS scanning data. One technique uses the ratio of laser returns within a distance to total number of outgoing laser beams as gap probability (Calders et al., 2015; Danson et al., 2007). The other technique uses calibrated laser return signals to estimate gap probability from the basic physics of the lidar equation (Jupp et al., 2009; Zhao et al., 2011). For the correction of leaf area clumping, TLS data has also been used to estimate the clumping index at scales coarser than shoots or leaf clusters (Moorthy et al., 2008; Zhao et al., 2012). For the correction of the woody contribution to estimates of LAI and the vertical foliage profile, a few studies have tried to use TLS data to classify leaves and woody materials in 3-D space using calibrated laser return intensities at a single wavelength (Béland et al., 2014) or laser return pulse shapes (Yang et al., 2013; Zhao et al., 2011). However, the interacting effects of reflectance, size, and orientation of targets may generate similar return intensities or return pulse shapes from leaves and woody materials.

A number of studies have used 3-D terrestrial lidar scanning to extract dendrometric parameters of forest stands, i.e. stem diameter at breast height (DBH), stem profile (stem diameter at different heights), tree height, stem density, basal area, and commercial wood volume (Dassot et al., 2011; Huang et al., 2011; Lovell et al., 2011; Strahler et al., 2008; Yang et al., 2013; Yao et al., 2011). These studies mainly demonstrated the potential of TLS to provide faster and more accurate forest inventory data.

For plot-scale AGB estimates, Yao et al. (2011) used DBH and stem count density extracted from TLS data with site-averaged allometric equations of dominant species and showed the resultant AGB estimates were in good agreement with estimates using species-specific allometric equations. Such species-specific allometric equations could not be used directly because identification of individual tree species from TLS data is quite difficult, if not impossible. Reulke & Haala (2005) explored a tree species classification method using the combination of geometric information from high resolution TLS and bark texture information from high resolution panoramic images, but did not put the method into practice. Further studies and developments of instruments and algorithms are needed to provide tree species identification that can be used with TLS data to fully automate forest inventory data collection and improve plot-scale biomass estimates.

For individual-tree AGB estimates, some recent studies have developed new, nondestructive approaches using TLS data to estimate aboveground biomass, independent of allometric equations. These approaches combine *a priori* wood density information with wood volume that is directly calculated from tree models built from terrestrial laser scanning data (Calders, Newnham, et al., 2014; McHale et al., 2009; Vonderach et al., 2012). Current tree model building approaches designed for wood volume estimates generally fall into two categories: geometric modeling and voxel modeling. Raumonen et al. (2013) used a geometric modeling approach called Quantitative Structure Modeling (QSM) to fit TLS data of stems and branches with collections of connected cylinders from which branching structure was derived and wood volumes of a whole tree was calculated from the sum of cylinder volumes. QSM is quite automated with little manual editing, but needs a proper selection of input parameters. In an earlier study, Thies et al. (2004) modeled stems with cylinders and calculated only stem volumes, not including branches.

Voxel modeling represents a tree using voxels. The voxel model of Vonderach et al. (2012) represented stem and all branches together with connected voxels. The voxel model of Hosoi et al. (2013) represented a tree with two components, a solid model composed of connected voxels that fill the outer surface and interior of stem and large branches, and a cloud of voxels equivalent to branches discretely scattered in the canopy. Voxel modeling needs a careful selection of voxel size (Hosoi et al., 2013), and currently requires leaf-off scans or explicit removal of leaf points for wood volume calculation by summing voxels. The branching structure is not directly available from voxel-based tree models. Lefsky & McHale (2008) combined geometric and voxel modeling in tree model building by using cylinders to represent main stems and large branches and voxels to represent small branches, but required significant manual editing for tree model correction.

1.3 Dual-Wavelength Echidna Lidar

The novel terrestrial lidar used in this dissertation, the Dual-Wavelength Echidna Lidar (DWEL) is designed to separate leaves and woody materials in forests in three-dimensional space using their two spectral reflectance values. Based on the design of the Echidna® Validation Instrument (EVI), built by Australia’s Commonwealth Scientific and Industrial Research Organization (CSIRO) (Jupp et al., 2009; Parkin et al., 2001), DWEL uses two coaxial pulsed lasers and acquires full-waveform scans at both near-infrared (NIR, 1064 nm) and shortwave infrared (SWIR, 1548 nm) wavelengths with simultaneous laser pulses. At the SWIR wavelength, laser power returned from leaves is much lower than from woody materials, such as trunks and branches, due to absorption by liquid water in leaves. In contrast, returned power from leaves and woody materials is similar at the NIR wavelength.

DWEL's two infrared lasers emit unpolarized pulses with a full-width half-maximum (FWHM) of 5±0.1 ns; the two laser beams are aligned coaxially to within 1 mrad (Figure 1-1). Although the pulse repetition frequency (PRF) of the lasers is 20 kHz, only one of every ten pulses is actually recorded, providing an effective PRF of 2 kHz. The collimated beam diameters of the two lasers are 6 mm; one of three pairs of interchangeable beam divergence optics of 1.25, 2.5 and 5 mrad is then used to expand each beam. The DWEL scanning step is set slightly smaller than the beam divergence, which ensures continuous coverage at the horizon (Douglas et al., 2015). A third continuous-wave green marker laser is also aligned with the two infrared signal lasers; since it is readily visible, it is used to position the triple beam or mark the scan path in the laboratory (Figure 1-1).

Similar to EVI, DWEL uses a rapidly rotating zenithal scan mirror and a slowly rotating azimuth platform to provide full coverage of the angular scan space (Douglas et al., 2015). Each rotation of the scan mirror directs the beam through 360°; returns from the environment are acquired at zenith angles from –117° to +117° (234° extent), while returns from the instrument housing, used for calibration, are acquired as the beam passes through angles of +117° to –117° (126° extent). For azimuth, the instrument platform rotates through 180°, thus providing a complete spherical scan. Angular measurement precision is assured by a Gurley Precision Instruments 12-bit optical shaft encoder wheel installed on each zenith and azimuth rotation shaft, with readout electronics providing 19 bits of interpolated absolute angular resolution (Douglas et al., 2015). Measurement error is less than 2 mrad (one sigma), below the operational scanning resolution.

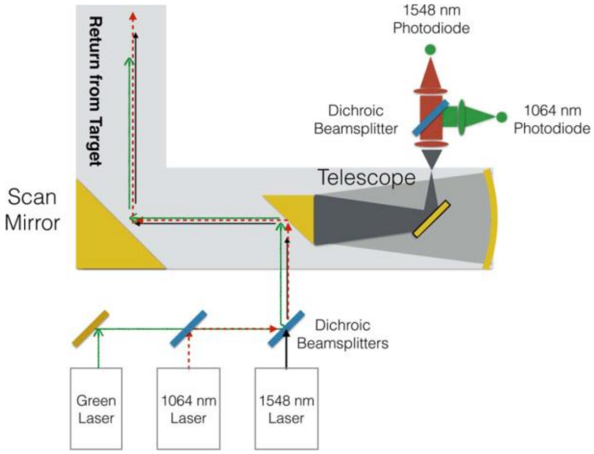


Figure 1-1 DWEL instrument schematic (Douglas et al., 2015). Thin lines represent the outgoing, coaxial beams. Gray lines represent the field of view of the instrument and the path of observed returns. The green laser serves as a marker for laser alignment and is not recorded.

1.4 Research Objective and Dissertation Structure

The primary objective of this dissertation is to evaluate the use of dual-wavelength spectral information, acquired by the DWEL, to separate leaves from woody and ground components of the forest in 3-D terrestrial lidar scans, and to explore how this new capability can improve the remote retrieval of forest structure. This evaluation has required three major steps: (1) calibration of lidar returns at the two wavelengths; (2) retrieval of the proportions of leafy and woody materials with height in a forest canopy;

and (3) estimation of woody aboveground biomass of forest trees through the virtual reconstruction of tree models from multiple DWEL scans.

In Chapter 2, I establish the basis of the lidar calibration for studies of canopy structure. I formulate the lidar equation to describe the interaction of the laser beam with vegetative elements and identify the objective variable of the calibration ("apparent reflectance") from the basic scattering lidar equation (Measures, 1991) using Ross's framework of radiation regime modeling of the vegetation canopy (Ross, 1981). I lay out the procedures to collect data and estimate parameters for radiometric calibration of dual-wavelength lidar. I also assess the accuracy of radiometric calibration and analyze the sensitivity of calibration to range errors and noise in lidar intensities.

In Chapter 3, I analyze DWEL scans of a midlatitude forest site (N 42° 31′ 51.48″, W 72° 10′ 55.56″) at Harvard Forest in Central Massachusetts, USA, in both leaf-off and leaf-on conditions. I develop a 3-D classification approach using both spectral and spatial context information from the dual-wavelength lidar data. I assess this classification in 3-D space with an indirect classification accuracy assessment procedure, and observe an average overall accuracy of 79±1% and 75±2% for leaf-off and leaf-on scans, respectively. Then I derive separate leaf and woody area profiles from both leaf-off and leaf-on scans.

In Chapter 4, I build virtual tree models from multiple scans of the study site at Harvard Forest with a state-of-the-art geometric fitting algorithm, Quantitative Structure Modeling (QSM) (Raumonen et al., 2013). I estimate the aboveground woody biomass of five individual trees by combining the wood volume from QSM tree models and *a priori* wood density. I use a least-variance approach to select the best input parameter to QSM for DWEL scanning data. Finally, I show the improvement of wood volume estimates and resultant biomass estimates by the separation of leaves and woody materials in the DWEL point clouds.

In Chapter 5, I present the overall conclusions and discussions.

CHAPTER 2. RADIOMETRICALLY CALIBRATED POINT CLOUDS FROM A DUAL-WAVELENGTH FULL-WAVEFORM TERRESTRIAL LIDAR

2.1 Introduction

A new and important application of lidar is the quantification of vegetation structure, principally measures of the physical dimensions of trees, amount and location of leaves, and gaps between and within tree canopies, through airborne and terrestrial laser scanning. Early lidar scanners used for this purpose were developed primarily to detect and locate scattering events in three-dimensional space, for example providing canopy height from airborne systems or 3-D models of forests from terrestrial systems. However, more complete inference of vegetation structure requires using the intensity of the scattered return detected by the lidar scanner, which in turn requires calibration. In its most useful form, this calibration provides a signal expressed as a property of the scattering event itself, such as the apparent reflectance of the scattering surface. This measure, defined as the reflectance of a perfectly diffuse target orthogonal to the lidar beam and completely filling the field of view that returns the same energy from the same range as an actual target, is directly useful in retrieving vegetation properties such as leaf area index, foliage density with height, and separation of returns of leaves from those of trunks, branches and ground. For terrestrial lidars, it also provides additional value for measuring mean tree diameter, canopy height, stem count density, and indirectly, above-ground biomass.

While calibration of airborne lidars has been well documented in the literature, calibration of terrestrial lidar scanners has not. It faces unique challenges, including (1) a very large variation in intensity with range that can induce saturation of the detector system by bright targets in the near field and reduced intensities that merge with the noise in the far range; and (2) strong telescopic effects, with defocusing producing weak signals at near range. This paper documents how these and other challenges have been addressed for a dual-wavelength, full-waveform terrestrial scanner, the Dual-Wavelength Echida® Lidar (DWEL). The simultaneous calibration of returns from DWEL's two lasers, pulsing at wavelengths of 1064 and 1548 nm, also demonstrates how calibration can ensure both radiometric and spectral fidelity in a unified process, thus providing a pathway for calibration of other dual and multiple wavelength terrestrial lidars now in various stages of development and application.

2.1.1 Lidar and its forest application

Light detection and ranging (lidar) is an active remote sensing technique using an instrument that emits coherent laser light and records the travel time and intensity of that light as scattered by targets along the laser transmission path. The range to a target is determined as the product of the speed of light and the one-way travel time of light between the instrument and the target. The travel time can be measured using either pulse ranging or continuous wave ranging techniques (Lim et al., 2003). For lidar remote sensing of forests, which is the application of concern in this paper, pulse ranging lidar is much more commonly used (Lim et al., 2003), and thus we will confine our discussion here to pulsing lidar unless otherwise noted. Moreover, we will focus our discussion below on terrestrial lidar scanning (TLS) from ground-based instruments, although much of the discussion will also apply to airborne lidar scanning.

The intensity[1] of light backscattered from a target is usually recorded in digital counts (DN), a measure provided by the optical and electronic system (detector-amplifier-digitizer) of the lidar instrument. Here we will assume a linear system response; *i.e.,* that the intensity measured for a given target at a specific distance is a linear function of the output energy of the laser pulse. For pulsing lidars, the intensity may be recorded either as a small number of discrete scattering events encountered by an outgoing pulse (discrete-return lidar), or as a continuous return waveform (full-waveform lidar) containing return pulses related to scattering events. Discrete-return lidar typically records only peak intensities of scattering events, detected on the fly by fast and simple algorithms, and saves them as points at particular locations in space. In contrast, full-waveform lidar retains a complete record of the scattering of the pulse with range as it travels away from the instrument in a particular direction. The advantage of full-waveform lidar is that the recorded waveform can be post-processed in different ways, particularly to yield more accurate representations of scattering events as points in space and to trace the continuous reduction in gap probability with range along the laser pulse path.

Through accurate measurements of ranges to scattering targets and the directions of emitted laser beams, lidar directly generates three-dimensional representations of targets as seen from the viewpoint of the lidar instrument. This acquisition of 3-D information by lidar technology has been widely explored on different remote sensing

---

[1] While the term “intensity,” as defined in optical physics, refers to the energy flow rate in W sr^{-1} from a point source of emission, we will use “intensity” here to refer to a measure, normally in digital counts, of the response of the detector-amplifier-digitizer system to the return of energy from a laser pulse or power of a continuous wave scattered by a target into the aperture of the telescope of the lidar instrument and reaching the detector system.

platforms (terrestrial, airborne, and spaceborne) and shown to facilitate research and management of forest ecosystems (Dubayah & Drake, 2000; Lefsky et al., 2002; Lim et al., 2003; Wulder et al., 2013) by describing the heterogeneity of forest structure in 3-D space and its relation to forest function (Baldocchi et al., 2002; Stark et al., 2012).

2.1.2 Motivation and status of lidar calibration

Of the two primary attributes of scattering events recorded by lidar, spatial location and intensity, location by range and angle (or as resolved into Cartesian coordinates) has found wide use in the retrieval of forest structural parameters (Strahler et al., 2008), such as diameter at breast height (DBH) (Lovell et al., 2011), tree and canopy height (Huang et al., 2011; Thies & Spiecker, 2004; Yao et al., 2011), timber volume (Murphy, 2008; Raumonen et al., 2013), leaf area index (LAI) (Béland et al., 2011; Hilker et al., 2010; Jupp et al., 2009; Korhonen et al., 2011; Zhao et al., 2012) and others (Dassot et al., 2011; Wulder et al., 2012).

However, intensity, the other attribute of the return signal, does not provide straightforward interpretation and has been underutilized. Intensities in digital counts output by lidar instruments neither give actual backscattered energy from targets nor relate directly to target physical properties. Accordingly, they are usually processed to remove electronic effects and normalize the decrease of observed intensity with range. The processed intensities thus provide the relative distribution of target return energy from which forest structural parameters can be inferred directly or through empirical regressions, such as gap fraction (Hancock et al., 2014), canopy height profile (Lefsky et al., 1999), basal area and above-ground biomass (Drake et al., 2002) and others (Wulder et al., 2012). Normalized intensity has also been used in target classification (Béland et al., 2011; Ramírez et al., 2013) and estimation of biochemical properties of vegetation (Eitel et al., 2014; Gaulton et al., 2013). Although these studies have documented the usefulness of lidar intensity values, their simple and empirical normalization limits inter-comparison of instruments and makes merging data from two or more instruments difficult. Even comparisons between datasets of the same instrument from different scanning campaigns can be problematic because of inconsistent normalization ranges or approaches and different interpretations of normalized values.

The need for a consistent definition of calibrated lidar intensity is also driven by recent design or fabrication of bi- or multispectral lidar instruments using lasers at different wavelengths or white lasers to exploit the spectral signatures of targets (Danson et al., 2014; Douglas et al., 2015, 2012; Eitel et al., 2010; Gong et al., 2012; Hakala et al., 2012; Tan & Narayanan, 2004; Woodhouse et al., 2011). For example, the Dual-Wavelength Echidna Lidar (DWEL) instrument, which is the focus of this paper, uses two coaxial lasers at 1064 nm and 1548 nm wavelengths to differentiate leaves from branches, trunks, and ground by taking advantage of the distinctive spectral response of leaves at the two wavelengths (Douglas et al., 2015).

To compare lidar intensities from different scanning campaigns, instruments or wavelengths, a consistent and unified radiometric calibration is needed. For airborne lidar scanning (ALS), recent studies have reviewed the physical concepts of return intensity and radiometric calibration (Höfle & Pfeifer, 2007; Roncat, Morsdorf, et al., 2014; Wagner, 2010). Various calibration targets and procedures for ALS data have been proposed and evaluated for several calibration scenarios, including different scanning campaigns with the same instrument (Kaasalainen, Hyyppä, et al., 2009; Roncat, Briese, et al., 2014; Wagner et al., 2006), different instruments at the same wavelength (Kaasalainen et al., 2011) or different wavelengths (Briese et al., 2012). Meanwhile, calibrated ALS intensities have been explored to improve key forest structure measurements, such as gap fraction (Armston et al., 2013), and even differentiate tree species (Reitberger et al., 2008).

In terrestrial lidar scanning (TLS), calibrated intensities have also been used to measure canopy structure. Examples include retrieval of multi-angle gap fraction and then LAI (Jupp et al., 2009); clumping index estimate (Zhao et al., 2012); and target classification with calibrated intensity alone (Béland et al., 2014) or along with pulse width from full-waveform data (Yang et al., 2013). However, with the exception of a few recent studies on both pulse-ranging TLS (Hartzell et al., 2015; Pfeifer et al., 2008; Pfennigbauer & Ullrich, 2010) and continuous-wave ranging TLS (Kaasalainen, Krooks, et al., 2009), radiometric calibrations of TLS data are currently scattered among various application studies and are poorly documented or rely on undocumented proprietary calibration algorithms from instrument manufacturers. The DWEL instrument we use in this study is a pulse-ranging TLS with two coaxial lasers pulsing at different wavelengths. No evaluation of radiometric calibration of similar TLS data has yet been documented for different wavelengths.

As a prerequisite to advance forest structure measurements by TLS with dual wavelengths, this study presents the procedure and evaluation of radiometric calibration of dual-wavelength point clouds from the DWEL. We begin with the theoretical lidar equation for canopy structure study using TLS data, and then describe the preprocessing of the DWEL waveform data and point cloud generation from the waveforms. Our calibration model, based on a generalized logistic function for telescope efficiency and an inverse power fall-off with range, is fitted to stationary scans of panels with known reflectance values at different ranges. We conclude by evaluating the range uncertainty and the calibration accuracy of dual-wavelength point clouds from DWEL as well as the sensitivity of the calibration accuracy to errors in both range and intensity measurements.

2.2 Physical Background

2.2.1 Basic lidar equation for forest canopies

The goal of our radiometric calibration is to obtain range-independent, instrument-independent, and physically well-defined measurements for canopy structure modeling and estimation from returned power as detected and recorded by the lidar instrument's optical and electronic systems. Previous studies (Ni-Meister et al., 2001; Sun & Ranson, 2000) formulated lidar equations as a function of canopy structure parameters to model large-footprint lidar waveforms but did not identify a realizable quantity for lidar radiometric calibration. To establish the basis of the lidar calibration for canopy structure study, we formulate the lidar equation to describe the interaction of the laser beam with vegetative elements and identify the objective variable of the calibration ("apparent reflectance") from the basic scattering lidar equation (Measures, 1991) using Ross's framework of radiation regime modeling of the vegetation canopy (Ross, 1981).

Consider an angular voxel, an elemental volume enclosed by a laser beam between range r to $r + \Delta r$ from the lidar instrument. Vegetative elements inside one such angular voxel are modeled as a turbid medium composed of tiny thin facets of different orientations in space. We shall not specify the size and thickness of these facets nor their location inside the angular voxel (Ross, 1981). Laser radiation incident into this angular voxel can be absorbed, reflected back toward the lidar instrument, or transmitted through it without interaction with vegetative facets. Simulations of lidar waveforms with Monte Carlo ray tracing have shown that multiple scattering by vegetative elements in the canopy largely has no effect on return waveform shapes and contributes little to return energy, especially for the small laser beam divergence as we use here (Hancock, 2010). Thus, it is reasonable to assume only single scattering in the interaction between laser beams and vegetative facets.

The probability that a laser beam in a given direction reaches an angular voxel at range r without interaction with vegetative facets is given by the gap probability $P_{gap}(r)$

$$P_{gap}(r) = e^{-G\int_0^r u_L(r')dr'} \qquad \text{2-1}$$

where G is the Ross G-function, which describes the projection of a unit vegetative area in a given direction; $u_L(r)$ ($\mathrm{m}^2 \cdot \mathrm{m}^{-3}$) is the total upper side surface area of all tiny facets (no mutual-shading between facets, i.e. no clumping is considered) within a unit volume at range r along the laser beam (Ross, 1981). Note that in expressions below, we will consider only a single laser beam and omit the laser beam direction in $P_{gap}(r)$ and G.

Let J_0 be the total outgoing laser radiation energy (units: J) within an infinitesimal time, i.e. an impulse laser energy. The received energy J (units: J) from the angular voxel at range r by the telescope of the lidar instrument is

$$J = J_0 P_{gap}(r) \cdot \beta \cdot K(r) \cdot \eta_{sys}\eta_{atm}$$

$$\beta = \int_{\Omega_T} \frac{1}{\pi}\Gamma(r, \Omega_i \to \Omega_v) d\Omega_v \cdot \int_r^{r+\Delta r} u_L(r')dr' \qquad \text{2-2}$$

$$\Omega_T = \frac{A_T}{r^2}$$

In these equations, $J_0 P_{gap}(r)$ (J), is the laser radiant energy that reaches the angular voxel at range r; β (dimensionless) is the effective backscatter ratio of the angular voxel, i.e., the proportion of the incident laser radiation energy that is scattered back from the angular voxel into the solid angle subtended by the telescope receiving area; the expression $(1/\pi)\Gamma(r, \Omega_i \to \Omega_v)$ (sr^{-1}) is the area scattering phase function at range r, i.e. the portion of the radiant energy received in direction Ω_i that is scattered by a unit area of vegetative facets in the direction Ω_v within a unit solid angle, a term which contains both the radiative and structural characteristics of the vegetative facets (Ross, 1981); Ω_T (sr) is the solid angle subtended by the area of the telescope aperture (A_T) of the lidar instrument from range r; $K(r)$ (dimensionless) is telescope efficiency at range r (see Section 2.2.1.3); and η_{sys} and η_{atm} (dimensionless) are transmission factors that account for energy loss due to the sensor system and atmosphere, respectively.

2.2.1.1 Apparent reflectance

In the simplest case, (1) only one angular voxel along a laser beam is filled with vegetative elements at range r, that is $P_{gap}(r) = 1$, and all laser radiation energy falls onto this voxel; (2) all the vegetative facets in the angular voxel are Lambertian with the same diffuse reflectance ρ_d (dimensionless) and all faces are orthogonal to the laser beam, that is $(1/\pi)\Gamma(\Omega_i \rightarrow \Omega_v) = \rho_d/\pi$; and (3) all the vegetative facets together fill the whole laser beam, that is $\int_r^{r+\Delta r} u_L(r')dr' = 1$. The return energy from such a voxel is then

$$J = J_0 \cdot \rho_d \cdot \frac{A_T}{\pi r^2} \cdot K(r) \cdot \eta_{sys}\eta_{atm} \qquad \text{2-3}$$

This simplest case is equivalent to that an extended Lambertian panel with reflectance ρ_d that fills the whole laser beam orthogonally at range r. The expression $A_T/(\pi r^2)$ in equation 2-3 is the proportion of the total hemispherical backscattering that is intercepted by the telescope aperture given this simplest case. Assuming η_{sys} and η_{atm} are constant with the range of our instrument operation and the laser wavelength in consideration, we may simplify equation 2-3 by combining all constants into Φ_0 ($\mathrm{J} \cdot \mathrm{m}^2$), i.e.

$$\begin{gathered} J = \Phi_0 \cdot \frac{\rho_d \cdot K(r)}{r^2} \\ \Phi_0 = J_0 \frac{A_T}{\pi} \eta_{sys}\eta_{atm} \end{gathered} \qquad \text{2-4}$$

The reflectance ρ_d, describing the physical properties of targets in the simplest case, is computable from the received energy and also range-independent and instrument-independent. For the general case observed in reality, we may identify a variable called apparent reflectance, ρ_{app} (dimensionless), that is calculated in the same way as ρ_d for the simplest case, i.e.

$$J = \Phi_0 \cdot \frac{\rho_{app} \ \cdot K(r)}{r^2} \qquad \text{2-5}$$

Comparing equation 2-5 with equation 2-2,

$$\rho_{app} = P_{gap}(r) \frac{\beta}{A_T/(\pi r^2)} \qquad \text{2-6}$$

Apparent reflectance, our range-independent and instrument-independent quantity, provides the objective variable for our lidar calibration for canopy structure study for two reasons. First, ρ_{app} can be modeled to derive canopy radiation and structural information. According equation 2-6, ρ_{app} includes $P_{gap}(r)$ as determined by the structural characteristics of vegetative facets in the canopy, and β as determined by both the radiant and structural characteristics of the vegetative facets of the canopy. Second, ρ_{app} can be interpreted as the reflectance factor value for a diffusely-reflecting, partly-absorbing panel filling the laser beam orthogonally that would produce the same received laser return energy as the actual target at the same range. It can be realized as the ratio of (dark current corrected) lidar intensities from a target and a white Lambertian panel (orthogonal to the laser beam, reflectance of 1) at the same range, as shown in equation 2-7

$$J = \Phi_0 \cdot \frac{\rho_{app} \cdot K(r)}{r^2}, J_w = \Phi_0 \cdot \frac{1 \cdot K(r)}{r^2}$$

$$\rho_{app} = \frac{J}{J_w} \tag{2-7}$$

Thus, ρ_{app} is theoretically useful for canopy structure study as well as practically realizable for lidar calibration. It was introduced in Parkin et al. (2001) and used in data interpretation by Jupp et al. (Jupp et al., 2009).

2.2.1.2 Physical interpretation of apparent reflectance

When the thickness of the angular voxel $\Delta r \to 0$, we have the differential form of equation 2-2 as

$$J(r) = \frac{\partial J}{\partial r} = J_0 P_{gap}(r) \cdot \beta(r) \cdot K(r) \cdot \eta_{sys}\eta_{atm}$$

$$\beta(r) = \frac{\partial \beta}{\partial r} = \int_{\Omega_T} \frac{1}{\pi} \Gamma(r, \Omega_i \to \Omega_v) d\Omega_v \cdot u_L(r) \tag{2-8}$$

where $J(r)$ ($\mathrm{J \cdot m^{-1}}$) is the received laser energy from range r per unit length of laser beam travel; $\beta(r)$ ($\mathrm{m^{-1}}$) is the effective volume backscatter ratio at range r, i.e. the proportion of the incident laser radiation energy that is scattered back into the solid angle subtended by the telescope receiving area at range r per unit length of laser beam travel. The quantity $(1/\pi)\Gamma(r, \Omega_i \to \Omega_v) \cdot u_L(r)$ ($\mathrm{sr^{-1} \cdot m^{-1}}$) is the volume scattering phase function, which defines the part of the radiant energy coming in the direction Ω_i that is scattered by a unit volume of vegetative facets in the direction Ω_v within the unit solid angle. The quantity $\beta(r)$ is the integral of the volume scattering phase function over the solid angle subtended by the telescope aperture from range r. Accordingly, we have the differential apparent reflectance $\rho_{app}(r)$ (m^{-1}) from equation 2-5 and 2-6 respectively.

$$\rho_{app}(r) = J(r) \cdot \frac{r^2}{\Phi_0 \cdot K(r)} \tag{2-9}$$

$$\rho_{app}(r) = P_{gap}(r) \frac{\beta(r)}{A_T/(\pi r^2)} \tag{2-10}$$

The equation 2-9 shows how to practically calculate and interpret apparent reflectance from the received laser energy. The equation 2-10 shows how to theoretically model apparent reflectance from the radiation and structural characteristics of the canopy. To separate the radiation and structural information about canopy from $\rho_{app}(r)$, we assume Lambertian facets and the same diffuse reflectance ρ_d for all vegetative elements that contribute to the received laser energy from which $\rho_{app}(r)$ is calculated. Then $(1/\pi)\Gamma(r, \Omega_i \rightarrow \Omega_v)$ can be approximated by $G^2 \cdot \rho_d/\pi$ (see Appendix A.1 for derivation). From equation 2-8 and 2-9, we have,

$$\rho_{app}(r) = P_{gap}(r) \cdot G^2 \cdot u_L(r) \cdot \rho_d \tag{2-11}$$

Applying equation 2-1,

$$P_{hit}(r) = -\frac{\partial P_{gap}(r)}{\partial r} = P_{gap}(r) \cdot G \cdot u_L(r) \tag{2-12}$$

$$\rho_{app}(r) = -\frac{\partial P_{gap}(r)}{\partial r} \cdot G \cdot \rho_d \tag{2-13}$$

Here, $P_{hit}(r)$ (m^{-1}) is the laser beam interception density (Ni-Meister et al., 2001), i.e. the interception fraction by vegetative facets per unit length along the laser beam. Taking the integral over range on both sides of the equation 2-13,

$$\int_0^r \rho_{app}(r')dr' = \left(1 - P_{gap}(r)\right) \cdot G \cdot \rho_d \qquad 2\text{-}14$$

Thus we can estimate P_{gap} over range from the integral of differential apparent reflectance over range calculated with equation 2-9 from received laser energy if G-function and leaf diffuse reflectance are known. Gap probability with range is an important function for indirect measurement of canopy structure, such as leaf area index (LAI), clumping index, and foliage profile (Chen & Cihlar, 1995; Jupp et al., 2009; Welles & Cohen, 1996). Calibrating lidar return intensity to apparent reflectance enables better estimates of gap probability than just counting numbers of points returned from targets along a laser beam (Jupp et al., 2009). The relation between apparent reflectance and gap probability given by equation 2-14 has been used to estimate gap probability and foliage profile consistently with waveform data from the Echidna Validation Instrument, the predecessor of the DWEL (Jupp et al., 2009; Zhao et al., 2011).

Note equation 2-14 omits the laser beam direction and implies two assumptions: (1) the *G*-function is constant over range; and (2) vegetative facets are all Lambertian with the same diffuse reflectance ρ_d. This might not be true for vegetative elements traversed by a single laser beam. In practice, $\rho_{app}(r)$ from multiple laser shots are often averaged together (e.g. over all azimuth angles within zenith angle ranges), which will reduce the variance in estimating $P_{gap}(r)$ for the canopy as a whole.

2.2.1.3 Telescope efficiency

The telescope efficiency function $K(r)$ is needed by geometric laser systems using a telescope to focus the return power (Halldórsson & Langerholc, 1978; Measures, 1991). It is theoretically zero at zero range (the focal point of the telescope) and rises to unity at the range at which the focused return beam falls entirely within the detector. $K(r)$ is usually omitted in airborne lidar equation (Wagner, 2010) because almost all targets on the ground are far enough from the instrument for $K(r)$ to reach unity. For terrestrial lidar, many returns are from near-range targets, which requires including $K(r)$ (Douglas et al., 2015; Pfennigbauer & Ullrich, 2010).

2.2.2 Spatial and temporal description of return power

The basic lidar equation above only describes return power from a single target at an instant time. For a pulsing laser, the description of actual return power over time involves both the spatial and temporal shape of the outgoing pulse and the spatial distribution of multiple targets in clusters intercepting laser beams.

2.2.2.1 Beam cross section

First, the energy is not uniformly distributed across the laser beam cross section. The laser beam profile is often Gaussian-shaped (Harding et al., 2001; Jutzi et al., 2003), thus eliminating the effect of diffraction at the beam aperture and maintaining the Gaussian beam cross section with range (Douglas et al., 2015). For an extended Lambertian target, the Gaussian laser beam profile theoretically has no effect on return power. For general targets as the vegetative facets in our canopy structure study here, the modeling of the effects of a Gaussian beam profile on return power is not trivial (Jutzi et al., 2003). However, in processing terrestrial lidar data, returns from many laser shots are usually aggregated together, which greatly reduces any variance due to the beam cross section shape.

2.2.2.2 Pulse shape

Second, the setup of lidar equation in equation 2-2 uses an impulse laser energy as the total outgoing laser radiation energy J_0, i.e. laser energy within an infinitesimal time. In reality, outgoing laser energy is spread over a finite time which makes a pulse shape, $P_0(t)$, ($\mathrm{J} \cdot \mathrm{s}^{-1}$) i.e. laser energy per unit time. After converting the time to apparent range with $r = c(t - t_p)/2$ where t_p is the time at which the outgoing pulse peak occurs and c is the speed of light, the outgoing pulse, $P_0(t)$ as a function of time becomes $P_0(r)$ ($\mathrm{J} \cdot \mathrm{m}^{-1}$) a function of range. Correspondingly, the term Φ_0 in equation 2-4 becomes a function of range $\Phi_0(r)$ and the return signal $P(r)$ is the convolution of $J(r)$ (equation 2-9) and $\Phi_0(r)$:

$$\begin{aligned} P(r) &= \int_0^r \Phi_0(r - r') \cdot \frac{\rho_{app}(r') \cdot K(r')}{r'^2} dr' \\ &= \Phi_0(r) * \frac{\rho_{app}(r) \cdot K(r)}{r^2} \end{aligned} \qquad \text{2-15}$$

To get $\rho_{app}(r)$, we need to deconvolve $\Phi_0(r)$ from received laser signal $P(r)$. However deconvolution is very sensitive to noise. To reduce the effects of noise on data interpretation, we modeled $\rho_{app}(r)$ as a sequence of Dirac delta functions. In this so-called delta-sequence model, vegetative facets are conceptually distributed inside

N_h $(N_h \geq 1)$ very thin angular voxels (i.e. $\Delta r \rightarrow 0$) at discrete ranges R_j $(j = 1 \ldots N_h)$, that is,

$$\rho_{app}(r) = \sum_{j=1}^{N_h} \rho_{app}(R_j)\delta(r - R_j) \qquad \text{2-16}$$

The apparent reflectance of the $j\text{-}th$ voxel, ρ_{app_j} that is the integral of differential apparent reflectance over the infinitesimal interval at range R_j is given by the differential apparent reflectance at range R_j calculated from $P(r)$

$$\rho_{app_j} = \int \rho_{app}(R_j)\delta(r - R_j)dr = \rho_{app}(R_j) \qquad \text{2-17}$$

Using the delta-sequence model, $P(r)$ is given by

$$\begin{aligned} P(r) &= \sum_{j=1}^{N_h} \Phi_0(r - R_j) \cdot \frac{\rho_{app}(R_j) \cdot K(R_j)}{R_j^2} \\ &= \sum_{j=1}^{N_h} \Phi_0(r - R_j) \cdot \frac{\rho_{app_j} \cdot K(R_j)}{R_j^2} \end{aligned} \qquad \text{2-18}$$

Thus ρ_{app_j} the quantity that relates to gap probability (equation 2-14) can be calculated from received waveforms while avoiding deconvolution with this delta-sequence model.

2.2.2.3 System response

Moreover, the actual shape of the recorded return signal is also changed by the characteristics of the detector-amplifier system of the instrument. For a system with impulse response function $h(r)$, the system response is $k \cdot P_0(r) * h(r)$ where k is an optical-to-digital conversion factor, i.e., the gain of the detector-amplifier-digitizer system in units of digital counts (DN) per unit of power received by the detector. The factor k is constant for detectors on many lidar instruments, such as the DWEL instrument used here (Douglas et al., 2015; Pfennigbauer et al., 2013). If the detector gain is not constant, the varying k function needs to be applied before relating recorded return intensities to apparent reflectance. As $P_0(r)$ is altered by the detector-amplifier system, the $\Phi_0(r)$ is changed accordingly. The corresponding outcome term is denoted as $S_R(r)$ in the lidar equation for $I(r)$, the recorded return signal in digital counts.

$$I(r) = S_R(r) * \left[\frac{\rho_{app}(r) \cdot K(r)}{r^2}\right] \quad \text{2-19}$$

With vegetative facets modeled as a sequence of Dirac delta functions (equation 2-16),

$$I(r) = \sum_{j=1}^{N_h} S_R(r - R_j) \cdot \frac{\rho_{app}(R_j) \cdot K(R_j)}{R_j^2} \quad \text{2-20}$$

The pulse shape $S_R(r)$ is important for extracting the target signal from noise contaminating the recording of returns, since the actual target signal is the convolution of $S_R(r)$ and the target response $\rho_{app}(r) \cdot K(r)/r^2$. The pulse shape $S_R(r)$ can be measured from targets with spatial extent of zero along the laser beam direction (Pfennigbauer et al., 2013), such as a single extended Lambertian target perpendicular to the laser beam.

2.3 The Dual-Wavelength Echidna Lidar (DWEL)

2.3.1 Instrument description

The scientific objective of DWEL instrument design is to separate leaves and woody materials in forests readily in three dimensional space using their different spectral reflectance values. Based on the design of the Echidna® Validation Instrument (EVI), built by Australia's Commonwealth Scientific and Industrial Research Organization (CSIRO) (Jupp et al., 2009; Parkin et al., 2001), DWEL uses two coaxial pulsed lasers and acquires full-waveform scans at both near-infrared (NIR, 1064 nm) and shortwave infrared (SWIR, 1548 nm) wavelengths with simultaneous laser pulses. At the SWIR wavelength, laser power returned from leaves is much lower than from woody materials, such as trunks and branches, due to absorption by liquid water in leaves. In contrast, returned power from leaves and woody materials is similar at the NIR wavelength.

2.3.1.1 Lasers and scanning

DWEL's two infrared lasers emit unpolarized pulses with a full-width half-maximum (FWHM) of 5±0.1 ns; the two laser beams are aligned coaxially to less than 1 mrad. Although the pulse repetition rate (PRF) of the lasers is 20 kHz, only one of every ten pulses is actually recorded, providing an effective PRF of 2 kHz. The collimated beam diameters of the two lasers are 6 mm; one of three interchangeable beam divergence optics of 1.25, 2.5 and 5 mrad is then used to expand each beam. DWEL scans presented here were collected at the standard operational settings of 2.5 mrad beam divergence optics and a scanning step of 2 mrad. This step setting, slightly smaller than the beam divergence, ensures continuous coverage of the hemispheres (Douglas et al., 2015). A third continuous-wave green marker laser is also aligned with the two infrared signal lasers; since it is readily visible, it is used to position the triple beam or mark the scan path in the laboratory.

Similar to EVI, DWEL uses a rapidly rotating zenithal scan mirror and a slowly rotating azimuth platform to provide full coverage of the angular scan space (Douglas et al., 2015). Each rotation of the scan mirror directs the beam through 360°; returns from the environment are acquired at zenith angles from –117° to +117° (234° extent), while returns from the instrument housing, used for calibration, are acquired as the beam passes through angles of +117° to –117° (126° extent). For azimuth, the instrument platform rotates through 180°, thus providing a complete spherical scan. Angular measurement precision is assured by a Gurley Precision Instruments 12-bit optical shaft encoder wheel installed on each zenith and azimuth rotation shaft, with readout electronics providing 19 bits of interpolated absolute angular resolution (Douglas et al., 2015). Measurement error is less than 2 mrad (one sigma), below the operational scanning resolution.

2.3.1.2 Internal calibration objects

Two scattering objects are fixed to the instrument to calibrate range and outgoing laser intensity. First, a fine stainless steel (removable) wire crosses the edge of the outgoing beam before it hits the scan mirror, thus scattering a small fraction of each outgoing pulse into the telescope and detectors. This allows a small part of outgoing laser pulse be present in the recorded signal, which assures the temporal alignment of individual waveforms and gives range precision values of one-sigma error of 2.59 cm at 1064 nm and of 2.52 cm at 1548 nm. Second, a small circular Spectralon® panel with nominal reflectance of 0.99 is affixed to the case so that each mirror rotation will acquire samples of outgoing pulses acquired from a fixed target at a fixed range. These sampled waveforms are used primarily to track drifts in laser output power that occur through the scan, but can also establish the mean outgoing pulse times of the lasers for each mirror rotation in the absence of a wire signal or refine the temporal alignment of waveforms by wire signal.

2.3.1.3 Signal recording and system response function

DWEL detects and digitizes the return signal at 2 GHz, i.e. every half nanosecond, and records the returns as full waveforms returning from a range up to 70 m. Figure 2-1 shows the mean of multiple samples of system response $S_R(r)$ at each wavelength after background noise is removed and the pulses are normalized by peak intensity. The "ringing" response after the maximum is produced by the modulation transfer function of the combined detector-amplifier.

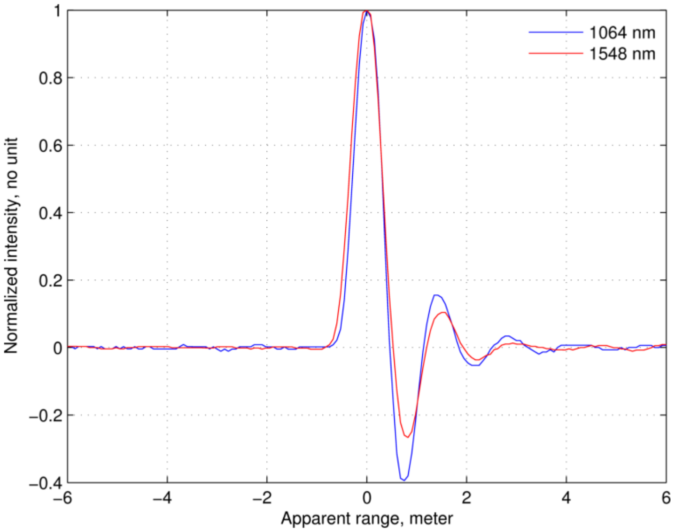


Figure 2-1 DWEL system response pulse. The two pulse peaks at the two wavelengths are aligned.

2.3.2 Basic processing

Before point cloud generation and radiometric calibration, the raw waveforms from DWEL are preprocessed to (1) remove background noise; (2) convert digitizer time to apparent range by aligning each waveform to the peak of the outgoing pulse using the signals from the scattering wire or internal Spectralon panel; (3) detect and correct saturated return pulses (see Section 2.3.2.1); (4) correct laser power drift, typically due to instrument temperature change, by scaling recorded intensities according to mean intensities observed from the internal Spectralon panel for each mirror rotation; and (5) calculate cross-covariance between waveforms and system response function. This cross-covariance function changes the original asymmetric return pulses $I(r)$ seen in each DWEL waveform to symmetric pulses. The new waveform of symmetric return pulses $I_p(r)$ is written as ($\star$ denotes cross correlation)

$$
\begin{aligned}
I_p(r) &= S_R(r) \star I(r) \\
&= \varphi(r) * \left[\frac{\rho_{app}(r) \cdot K(r)}{r^2}\right]
\end{aligned}
\tag{2-21}
$$

where $\varphi(r) = S_R(r) \star S_R(r)$ is a symmetric pulse after cross-covariance calculation. Multiple return pulses identified from a waveform are then treated as a sequence of multiple target returns, each of which is modeled as a Dirac delta function (equation 2-16). Comparing equations 2-19, 2-20 and 2-21, we have

$$
I_p(r) = \sum_{j=1}^{N_h} \varphi\left(r - R_j\right) \cdot \rho_{app}\left(R_j\right) \cdot \frac{K\left(R_j\right)}{R_j^2}
\tag{2-22}
$$

The preprocessed waveform $I_p(r)$ then provides the input to point cloud generation, thus avoiding pulse peak shifts due to the asymmetry of the original DWEL return pulse (see Figure 2-3 and Section 2.4.1 for more information). An additional significant benefit of this operation is that it reduces uncorrelated noise and increases the signal-to-noise ratio prior to extraction of the signal in later processing.

2.3.2.1 Saturation fix

Terrestrial laser scanners, in contrast to airborne scanners, will provide returns from close targets, sometimes within one meter for a placement in a forest with a dense or patchy understory, while also detecting targets at ranges of 100 m or more. This large relative variation in range provides a wide variation in return power that can exceed the limits of detector-amplifier systems (linear or nonlinear) and/or digitizers available for terrestrial scanners, and as a result, close targets can produce saturated pulse waveforms. Moreover, direct solar irradiance or specular reflectance may also produce saturated waveforms. The result may be either detector saturation, which produces an overloaded signal that persists through multiple digitizer bins or even multiple pulses, or digitizer saturation, which produces a flat-topped return pulse as the return signal exceeds the quantization range of the digitizer. In either case, the result is an unusual return pulse shape that cannot be used in calibration or to generate a scattering point with a correct apparent reflectance value.

In the DWEL instrument, detector saturation occurs in the rare case of a pulse striking an orthogonal specular target or corner reflector; normal target returns are well within the incoming power bounds of the DWEL’s detector-amplifier (Thorlabs PDA10CF) given the outgoing laser energy of DWEL. If the field of view of the telescope includes the sun or the sun’s aureole, the pulse may be lost completely as the detector and/or digitizer saturates or records high levels of continuous noise through the entire waveform. This situation is easy to detect and such waveforms are identified as solar-saturated.

Digitizer saturation, however, is commonly encountered in pulses returned from near objects. Here, a “saturation fix” is employed (Figure 2-2). Saturation creates a flat-topped pulse as the digitizer reaches its limit; however, the side-lobe trough and secondary peak are recorded correctly. By comparing saturated and unsaturated waveforms acquired from targets with high and low reflectance at the same range, we determined empirical ratios between the magnitudes of the side lobes and the unsaturated peak. These ratios are used to generate a pulse peak that is located at the mean range of the saturated bins. This pulse is identified as a “desaturated” pulse for further processing.

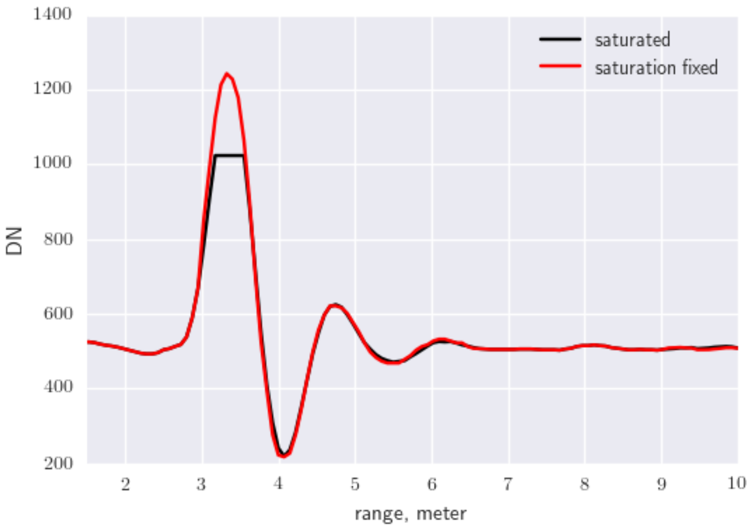


Figure 2-2. Example of saturated pulse and saturation fix.

2.4 Methods

2.4.1 Point cloud generation

The goal of basic processing of DWEL waveform data is to identify, quantify, and locate the scattering events encountered by each outgoing laser pulse from the recording of the return value by the instrument. In this way, the scattering events may be recorded as points in a cloud with attributes of intensity, or after radiometric calibration, of apparent reflectance.

2.4.1.1 Normalized pulse shape

For a theoretical return intensity pulse $\psi(r)$ (DN) of which peak intensity α is at range zero, i.e. $\psi(0) = \alpha$, the integral of the digital counts recorded by the instrument's digitizer is denoted as E^*,

$$E^* = \int_{-\infty}^{+\infty} \psi(r)dr \tag{2-23}$$

The equivalent width (EW) of the pulse $\psi(r)$ can be approximated by (Yang, 2012)

$$EW = \frac{E^*}{\max[\psi(r)]} = \frac{\int_{-\infty}^{+\infty} \psi(r)dr}{\max[\psi(r)]} \tag{2-24}$$

Equivalent width of a pulse is the width of a rectangular band with the same integral as the pulse and the same max value centered at the pulse peak. If $\psi_N(r)$ is the standard pulse from normalizing the pulse peak $\psi(0)$ to one, i.e. $\psi(r) = \alpha \cdot \psi_N(r)$ and $\psi_N(0) = 1$,

$$EW = \frac{\alpha \int_{-\infty}^{+\infty} \psi_N(r)dr}{\alpha} = \int_{-\infty}^{+\infty} \psi_N(r)dr \tag{2-25}$$

A preprocessed DWEL return waveform of symmetric pulses (see Section 2.3.2) can be written as a sum of several return pulses from several targets,

$$I_p(r) = \sum_{j=1}^{N_h} \alpha_j \cdot \psi_{N_j}(r - R_j) \tag{2-26}$$

where j represents the j-th return pulse resolvable and decomposed from the return waveform, N_h is the number of pulses, α_j is the peak intensity of the symmetric DWEL return pulse, R_j is the range to the return pulse peak, and $\psi_{N_j}(r)$ is the j-th normalized pulse. The α_j and R_j will be assigned as two attributes, intensity and range to points extracted from return waveform decomposition.

2.4.1.2 Normalized system response and apparent reflectance

The standard pulse of a single return from an extended target orthogonal to the laser beam, i. e. $\psi_{N_j}(r)$, will be a copy of the system response ($\varphi(r) = S_R \star S_R$, see Section 2.3.2) according to equation 2-21, that is normalized to a peak of 1; we refer to this function as the normalized system response,

$$\psi_{N_j}(r) = \varphi_N(r) = \frac{\varphi(r)}{\max[\varphi(r)]} \qquad \text{2-27}$$

and

$$\alpha_j = C_0 \cdot \rho_{app}(R_j) \cdot \frac{K(R_j)}{R_j^2} \qquad \text{2-28}$$

where $C_0 = \max[\varphi(r)]$. With targets modeled as a sequence of Dirac delta functions (equation 2-16), $\rho_{app_j} = \rho_{app}(R_j)$ (equation 2-17). Thus, the peak intensity of a return pulse, α_j, is directly related to our calibration objective, apparent reflectance ρ_{app_j} of an angular voxel filled with vegetative facets. If the spatial extent of a target is not zero (violation of Dirac delta model), $\psi_{N_j}(r)$ will be an elongated version of $\varphi_N(r)$ and have a larger FWHM than $\varphi_N(r)$ according to equation 2-21. Now the shape of $\psi_{N_j}(r)$ becomes unique to the scattering event. The solution to the unknown shape of $\psi_{N_j}(r)$ is addressed in the following section 2.4.1.3.

2.4.1.3 Waveform decomposition strategy

Many lidar studies model $\psi_{N_j}(r)$, the shape of the return pulses, with a mathematical function, such as Gaussian, lognormal, or generalized Gaussian (Chauve et al., 2008). Individual pulses can then be fitted to the model with one or two parameters. However, theoretical mathematical functions can never reconstruct the actual system response and may introduce extra uncertainties in waveform decomposition (Hartzell et al., 2015). Moreover, the system response of DWEL is not easily modeled with a simple function, which makes waveform decomposition by fitting pulses to mathematical functions unstable and potentially very sensitive to noise.

A workable model for decomposition that we exploit here is to assume that all target returns have the shape $\varphi_N(r)$, but differ in peak intensity α. Thus, we ignore variation in FWHM pulse width at this stage of processing. We apply this model to the preprocessed waveform $I_p(r)$ (see Section 2.3.2) and write the preprocessed waveform as

$$I_p(r) = \sum_{j=1}^{N_h} \alpha_j \cdot \varphi_N\left(r - R_j\right) \qquad \text{2-29}$$

where $\varphi_N(0) = 1$.

Suppose a waveform has M_b digitized bins and we denote the apparent range of the *k*-th bin as r_k. Equation 2-29 for all the waveform bins then provides a $M_b \times N_h$ linear system if we know $R_1, R_2, \cdots, R_{N_h}$. However, to avoid uncertainty and instability in solving this linear system caused by bins containing only noise, we do not include all M_b bins. Instead, we use the N_h peak bins of return pulses. That is, a $N_h \times N_h$ linear system as follows,

$$\underset{\sim}{I_p} = \Phi\underset{\sim}{\alpha}$$

$$\underset{\sim}{I_p} = \begin{bmatrix} I_p(R_1) \\ I_p(R_2) \\ \vdots \\ I_p\left(R_{N_h}\right) \end{bmatrix}, \underset{\sim}{\alpha} = \begin{bmatrix} \alpha_1 \\ \alpha_2 \\ \vdots \\ \alpha_{N_h} \end{bmatrix} \tag{2-30}$$

$$\Phi = \begin{bmatrix} 1 & \varphi_N(R_1 - R_2) & \cdots & \varphi_N\left(R_1 - R_{N_h}\right) \\ \varphi_N(R_2 - R_1) & 1 & \cdots & \varphi_N\left(R_2 - R_{N_h}\right) \\ \vdots & \vdots & \ddots & \vdots \\ \varphi_N\left(R_{N_h} - R_1\right) & \varphi_N\left(R_{N_h} - R_2\right) & \cdots & 1 \end{bmatrix}$$

In addition, we add a second $N_h \times N_h$ linear system using convolution to stabilize the estimate of $\underset{\sim}{\alpha}$ with a linear least squares solution:

$$I_{p2}(r) = \varphi_F(r) * I_p(r) = \sum_{j=1}^{N_h} \alpha_j \cdot (\varphi_F * \varphi_N)(r - R_j)$$

$$I_{p2}(R_k) = \sum_{j=1}^{N_h} \alpha_j \cdot (\varphi_F * \varphi_N)(R_k - R_j) \tag{2-31}$$

$$\underset{\sim}{I_{p2}} = \Phi_2\underset{\sim}{\alpha}$$

where φ_F is a normalized symmetric pulse, i.e. $\varphi_F(0) = 1$. Usually we choose $\varphi_F = \varphi_N$. In matrix notation,

$$\begin{bmatrix} \Phi \\ \Phi_2 \end{bmatrix} \underset{\sim}{\alpha} = \begin{bmatrix} \underset{\sim}{I_p} \\ \underset{\sim}{I_{p2}} \end{bmatrix} \tag{2-32}$$

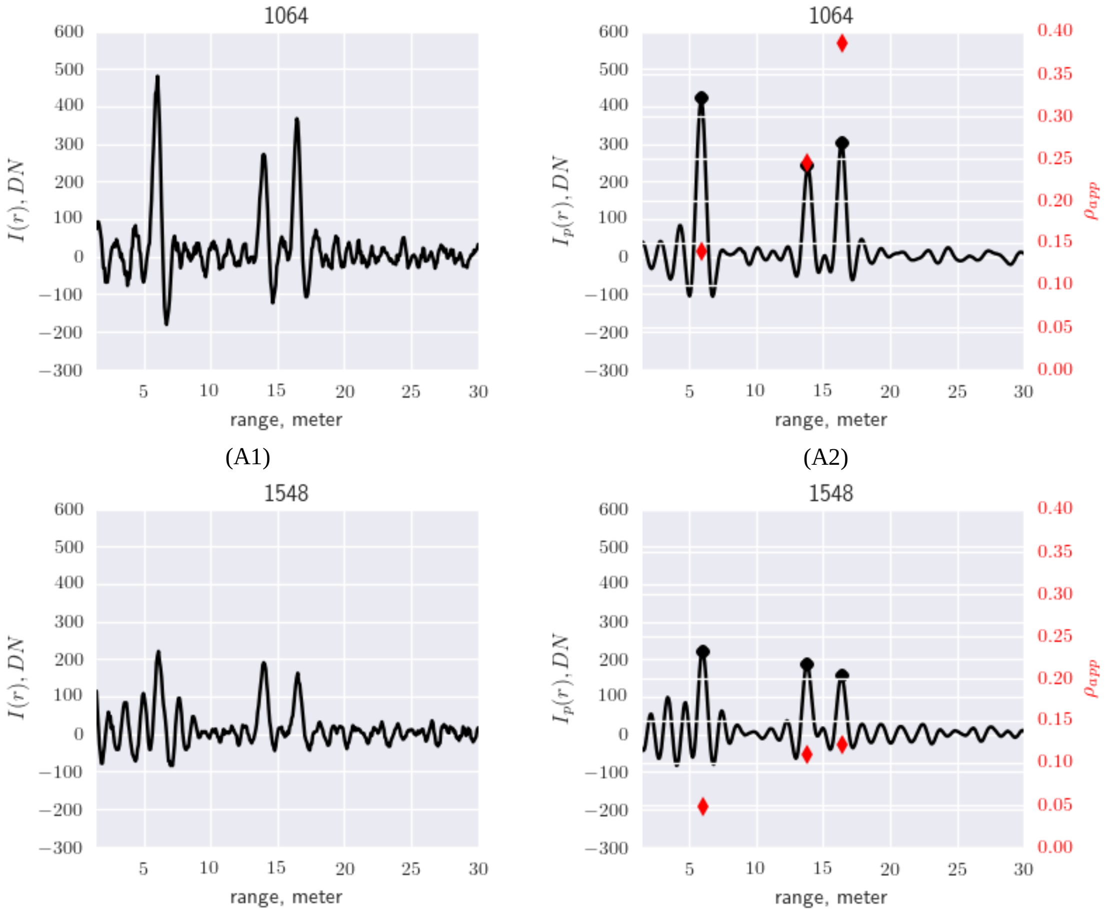
We obtain R_j, locations of return pulse peaks, by searching waveform sections with shapes similar to φ_N using cross-correlation. To be specific, we identify the peak bins of return pulses with zeros of derivatives and associated second derivatives from

$I_{p2}(r)$ and obtain refined peak locations $R_1, R_2, \cdots, R_{N_h}$ by quadratic interpolation with three bins around the identified peak bin. The peaks identified from $I_{p2}(r)$ are at the same locations as those of $I_p(r)$ because both φ_N and φ_F are symmetric. Finally, we solve the $2N_h \times N_h$ linear system for $\underset{\sim}{\alpha}$ given by equation 2-32 using least squares:

$$\underset{\sim}{\hat{\alpha}} = [\Phi^T\Phi + \Phi_2^T\Phi_2]^{-1}[\Phi^T \underset{\sim}{I_p} + \Phi_2^T \underset{\sim}{I_{p2}}] \qquad \text{2-33}$$

Note that in the above linear system we have assumed that each pulse return in $I_p(r)$ follows the shape of the symmetric system response, $\varphi_N(r)$. If a target j has spatial extent greater than zero along the laser beam direction, its return pulse shape will have a larger EW than φ_N and thus $I_{p2}(R_j)$ will be larger than it would be if the return pulse had the shape of φ_N. Then $\hat{\alpha}_j$ will be larger, which will compensate for some but not all of the underestimate of return pulse energy due to the assumption of the return pulse shape as the system response shape. A more accurate model might vary the return pulse shape as $\varphi_N(r/s_j)$, where s_j is a scale factor to accommodate the EW. However, the estimate of this extra scale parameter s_j requires the solution of a nonlinear system and will be pursued at a later time.

Figure 2-3 (left column) gives an example of extracted points from waveform decomposition in DN (dark dots) and in ρ_{app} (red diamonds) after radiometric calibration (see Section 2.4.2 in the following).



(A1) (A2)

(B1) (B2)

Figure 2-3. First row, 1064 nm. Second row, 1548 nm. First column, a waveform sample before filtering by cross covariance. Second column, the waveform sample after filtering as the input to point cloud generation. Black dots are extracted points from waveform decomposition and show the digital count before radiometric calibration. Red diamonds are points of apparent reflectance after radiometric calibration.

2.4.2 Radiometric calibration

2.4.2.1 Calibration model

The purpose of calibration as it is pursued here is to convert the intensity (DN) associated with each pulse peak to apparent reflectance, a measure that removes the effects of telescope efficiency and fall-off with range. For the calibration model, we may choose between two alternatives: a physical model based on the optical design of the instrument and an empirical model designed to best fit the data. A physical model describing the returned power P_r of the DWEL instrument, derived from first principles, is

$$P_r = \frac{C_0 \rho}{r^2}\left(1 - \exp\left[-\frac{(C_2 r)^2}{(C_1 + r)^2}\right]\right)$$

$$C_0 = \frac{P_t A_T}{\pi}, C_1 = \frac{\sigma_0}{\theta}, and\ C_2 = \frac{R}{\theta\, f_{sys}} \qquad \text{2-34}$$

where r is the range, ρ is the reflectance of a diffuse target, P_t is the outgoing pulse power, A_T is the telescope receiving area, σ_0 is the initial beam half width (at e^{-1}), θ is the beam divergence (half angle), R is the radius of the detector, and f_{sys} is the focal length. However, initial tests of this model with the calibration data (see Section 2.4.2.2) were not satisfactory. While the general shape of the response function fit the observations, the model showed significant departures from the behavior actually observed for the instrument, especially at near range. We believe that second-order effects, such as imperfect alignment interacting with the Gaussian beam cross section, departure of divergence from nominal specifications, or un-modeled electronic effects, were responsible for this variance.

Accordingly, we used a semi-empirical model to fit the data. According to equation 2-28, for each extracted point with range R and digital count intensity α, we need the constant C_0 and the function $K(r)$ to calculate apparent reflectance ρ_{app}. To find these unknowns, we require a collection of data points of $\left(R_j, \alpha_j, \rho_{app_j}\right)$ from targets of different reflectance values at multiple ranges. The quantities R_j and α_j are derived from point cloud generation; apparent reflectance ρ_{app_j} may be taken as the diffuse reflectance ρ_d of an extended Lambertian target held perpendicular to the laser beam. The near-range effect, $K(r)$, is a function theoretically rising from zero at the focal point of the lidar optical system to unity at a range at which divergence of the scattered beam entering the telescope aperture is negligible; it will be modeled with a generalized logistic function (Richards, 1959). The empirical calibration function for the DWEL is thus

$$\rho_{app}(R_j) = \alpha_j \cdot \frac{R_j^b}{C_0 \cdot K(R_j)}$$

$$K(r) = \frac{1}{(1 + C_1 \cdot e^{-C_2 \cdot r})^{C_3}} \qquad 2\text{-}35$$

where five parameters need to be estimated for the DWEL calibration, (C_0, C_1, C_2, C_3, b). Of the three parameters for $K(r)$, C_1 and C_3 together determine the range at which the function approaches its asymptote of one; C_2 controls the rate at which telescope efficiency rises from 0 to 1 in the near range.

Note that the exponent of range is taken as a variable, b, for two reasons. First, a calibration target surface, for example a manufactured Spectralon Lambertian panel, may not provide perfectly isotropic diffuse reflectance (Bhandari et al., 2011). The farther a calibration target is from the instrument, the smaller is the solid angle Ω_T subtended by the telescope area over which the integral is taken to give the effective backscatter ratio β (equation 2-2). If the anisotropic target preferentially reflects radiation into a small solid angle in the direction of instrument's observation, the smaller Ω_T causes the integral over directions where the area scattering function has larger values and thus larger β and larger return energy J than it would be if the target were perfectly isotropic. However, the reflectance value of a calibration panel is typically taken as a constant, for example measured by an integrating sphere, and assumed to be isotropic in the calibration model. From equation 2-5, if ρ_{app} is kept constant but J becomes larger, the exponent of range r will be smaller to compensate. Second, previous studies have also suggested that the exponent may need to accommodate electronic effects (Yang et al., 2013). The exponent has therefore been treated as a variable in the calibration. Although the number of model parameters is large and they are not independent of each other, the calibration function can be fitted across its full range of application, thus avoiding issues of extrapolation beyond the limits of the fitting.

2.4.2.2 Calibration experiment

To acquire the calibration data, three panels of different reflectance values were scanned by the DWEL from a nearly perpendicular direction at 33 range locations from 0.5 m to 70 m (Table 2-1). The range sampling intervals were based on a provisional calibration, made at the time of commissioning, that established the general shape of the $K(r)$ curve. The instrument was set in stationary mode, i.e. without scan mirror or azimuth platform rotation, and the green marker laser was used to manually point the lasers to the center of each panel at each placement. The panel sizes are large enough to intercept the whole laser beam at the 70 m. For each panel at each range, we collected around 150,000 waveform samples as candidates for calibration model fitting and evaluation.

Table 2-1 Range sample design

Range (m)	Range Interval (m)	Measurement Positions
[0.5, 10]	0.5	20
(10, 15]	1	5
(15, 40]	5	5
(40, 70]	10	3

The three panels included a white Spectralon panel and two foam boards painted with flat interior wall paint in light and dark gray tones derived by mixing black and white paints together. Panel reflectance values are shown in Table 2-2. The Spectralon panel reflectance is from the manufacturer's specification. The gray panel reflectance values were first measured using a FieldSpec Pro spectrometer (Analytical Spectral Devices) fitted with a self-illuminating spectral probe. However, this device measures reflectance at 0° incidence angle and 10° view angle rather than by retroreflection, and the values appeared to be underestimated slightly due to reflectance anisotropy (BRDF effect). Moreover, the reflectance of flat wall paint may have changed with time as the paint slowly cured between the spectrometer measurement and the acquisition of calibration data (about 20 weeks). We thus calculated the gray panel reflectance values using the manufacturer's Spectralon panel reflectance measurement and the ratio of return intensities between the Spectralon panel and gray panels at each range. We then used the mean of these calculated reflectance values as anisotropy-corrected gray panel reflectance values. The adjusted values were 28 – 34 percent higher than originally measured.

Table 2-2 Reflectance values of panels used in calibration.
White panel: Manufacturer's calibration value. Measured value: Spectrometer with illuminated probe. Adjusted value: Adjusted to provide consistent ratio with Spectralon panel.

Target	NIR Reflectance		SWIR Reflectance	
	Measured Value	Adjusted Value	Measured Value	Adjusted Value
White Spectralon panel	0.99		0.98	
Gray painted panel 1	0.436	0.574	0.349	0.447
Gray painted panel 2	0.320	0.431	0.245	0.329

2.4.2.3 Calibration model fitting

A primary objective of the DWEL instrument is to separate returns of leaves from those of other targets, such as trunks, branches, and ground, by taking advantage of the difference in spectral reflectance values of leaves and those of other targets at NIR and SWIR bands. As we try to minimize errors in ρ_{app} at individual wavelengths in calibration model fitting, we also try to ensure the two wavelengths have the same or similar relative errors in ρ_{app} in order to minimize artificial variations in spectral difference due to different errors in ρ_{app} at the two wavelengths. Thus, we estimate the calibration parameters of NIR and SWIR bands together in a joint calibration model that is fitted using an objective error function including relative errors in ρ_{app} from individual wavelengths and spectral constraints from both wavelengths as follows:

$$
\begin{gathered}
f\left(\underset{\sim}{C}\right) = f_1\left(\underset{\sim}{C}\right) + f_2\left(\underset{\sim}{C}\right) \\
f_1\left(\underset{\sim}{C}\right) = \sum_{i=1}^{N_f}\left(\frac{\widehat{\rho_{Ni}} - \rho_{Ni}}{\rho_{Ni}}\right)^2 + \sum_{i=1}^{N_f}\left(\frac{\widehat{\rho_{Si}} - \rho_{Si}}{\rho_{Si}}\right)^2 \\
f_2\left(\underset{\sim}{C}\right) = var\left(\widehat{NDI}\right) + \sum_{i=1}^{N_f}\left(\frac{\widehat{\rho_{Ni}} + \widehat{\rho_{Si}} - \rho_{Ni} - \rho_{Si}}{\rho_i + \rho_{Si}}\right)^2 \\
\widehat{NDI} = \frac{\widehat{\rho_N} - \widehat{\rho_S}}{\widehat{\rho_N} + \widehat{\rho_S}}
\end{gathered}
\tag{2-36}
$$

where $\underset{\sim}{C}$ is the vector of 5 calibration parameters for NIR and 5 parameters for SWIR; $f\left(\underset{\sim}{C}\right)$ is the objective error function as a sum of two components, the error $f_1\left(\underset{\sim}{C}\right)$ from individual wavelengths, and the spectral constraints $f_2\left(\underset{\sim}{C}\right)$ from both wavelengths; subscript i represents the i-th data point and subscript N and S represent NIR and SWIR; N_f is the total number of data points used in calibration fitting; $\hat{\rho}$ is the apparent reflectance of panels estimated from calibration model while ρ is the adjusted diffuse reflectance of panels; $\widehat{NDI}$ is normalized difference index to identify the spectral difference of target reflectance between NIR and SWIR; $var(\widehat{NDI})$ is the variance of NDI of data points in calibration fitting. In addition, we had the same calibration parameter C_1 and C_3 for NIR and SWIR to make $K(r)$ at the two wavelengths asymptote at the same range. The objective error function $f\left(\underset{\sim}{C}\right)$ has many local minima due to the high nonlinearity of the DWEL calibration model brought about by the $K(r)$ function. We used the Genetic Algorithm implemented in MATLAB (MATLAB, 2015b) to search for initial parameter values that will approach the global minimum before using the Nelder-Mead method (MATLAB, 2015a) to reach the global minimum.

All the waveforms from stationary panel scans were processed to points with intensities and ranges. Saturated waveforms were excluded from calibration model fitting to avoid uncertainty from the saturation fix procedure. We randomly divided the remaining returns (about 24,000 samples for each range) into a training set (80 percent) and a validation set (20 percent). In the training set, return intensities were normalized by the corresponding panel reflectance to provide equivalent target reflectance values of 1.0 and then averaged together for each range to reduce noise in the data. Mean normalized intensities and ranges at 1064 nm and 1548 nm were paired according to panel range locations. Thirty pairs of data points from 1064 nm and 1548 nm were used to estimate the calibration parameters of 1064 nm and 1548 nm jointly by minimizing the error function $f\left(\underset{\sim}{C}\right)$.

2.5 Results and Discussion

2.5.1 Dual-wavelength point cloud

We produced single point clouds with dual-wavelength apparent reflectance values by merging two point clouds generated separately from waveforms at the two wavelengths to examine and reduce the range bias between point clouds at the two wavelengths. To assess the variation in range between point clouds at the two wavelengths, we calculated the difference in ranges (1064-nm range – 1548-nm range) of unsaturated (see Section 2.5.1) single returns at the two wavelengths from stationary scans of the Spectralon panel at different ranges (see Section 2.4.2.2 for more details). The histogram of range differences (Figure 2-4) presents a near-Gaussian shape, suggesting random variation. The mean difference of –6.861 cm indicates a bias, in which the range measured by the 1064-nm laser is shorter. This difference is within the nominal waveform sampling interval of 0.5 ns, i.e. 7.5 cm, and is thus within one digitizer bin.

To reduce the effect of this bias in registering the two point clouds, the ranges of common points in the two point clouds are averaged, thus providing a single point cloud of dual calibrated apparent reflectance values. Common points from the two point clouds are identified if they share the same laser shot sequence number and their range difference is within three standard deviations (±12.24 cm) of the histogram value in Figure 2-4.

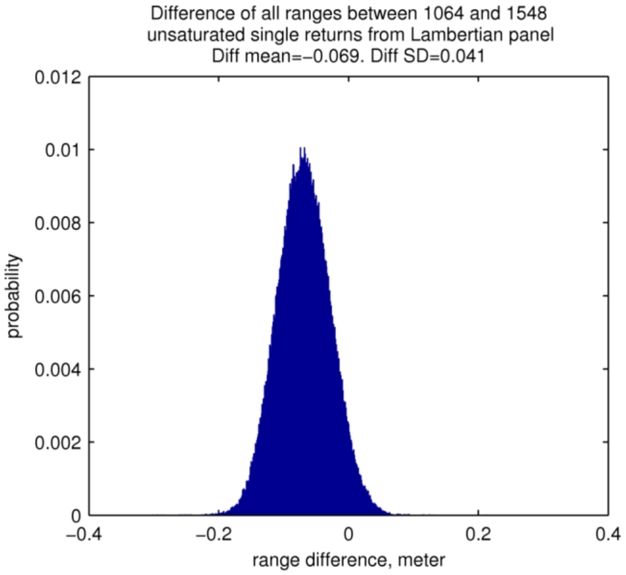


Figure 2-4 Histogram of range difference between the two wavelengths. Range from 1064 nm minus that from 1548 nm. Mean: -6.861 cm. Standard deviation: 4.081 cm.

2.5.2 Radiometric calibration

2.5.2.1 Fitting of the empirical model

Table 2-3 provides values for the model coefficients derived from the model and procedure described in Section 2.4.2; they may be taken as examples, since recalibration will be necessary during the lifetime of the instrument.

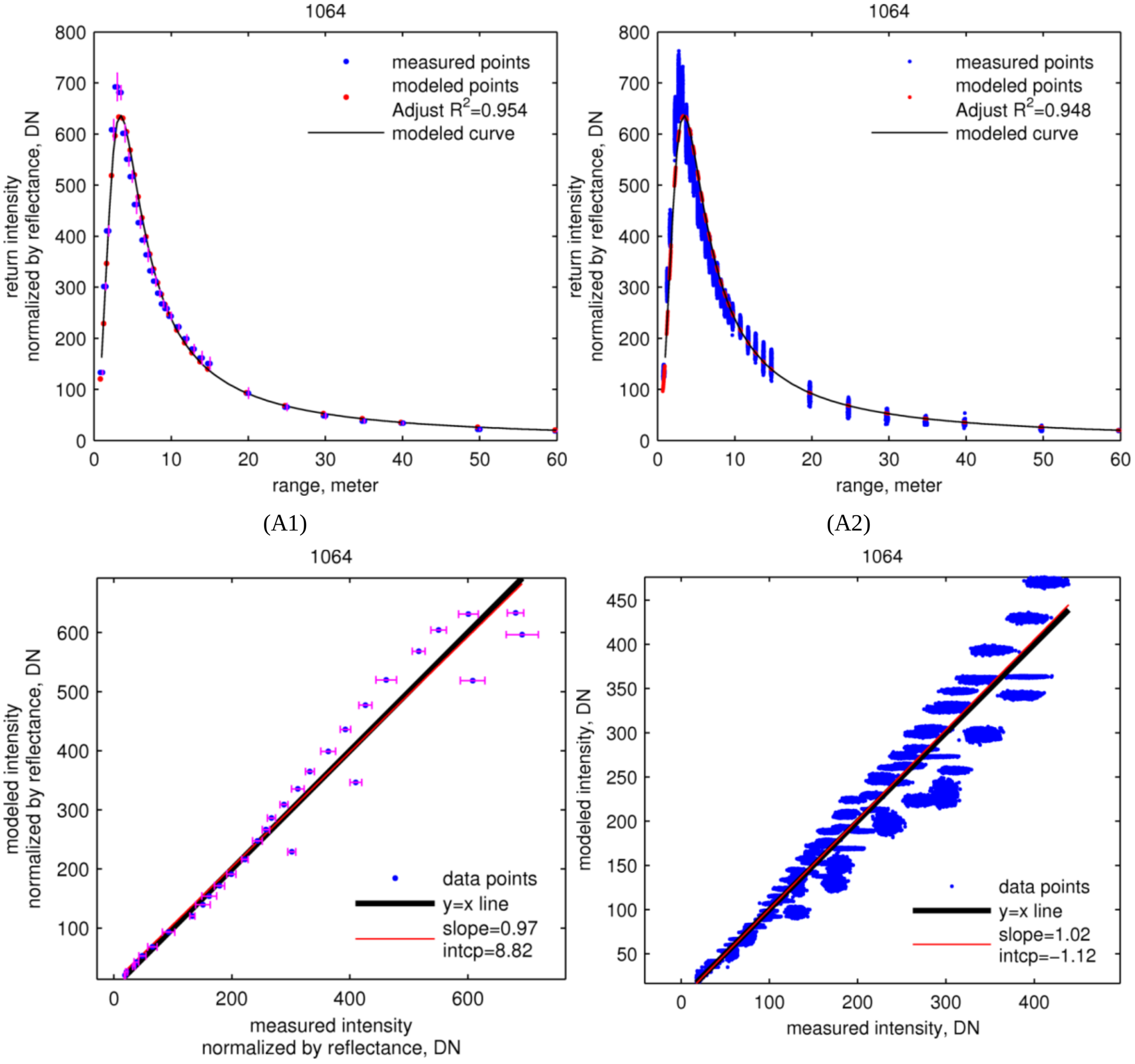
Table 2-3 An example set of DWEL calibration parameters

Parameter	1064	1548
C0	5788.265818	22054.218342
C1	0.000319	0.000319
C2	0.808880	0.540762
C3	25176.835032	25176.835032
b	1.384297	1.585985

Figure 2-5 and Figure 2-6, rows A and B, show the fits of the calibration functions (equation 2-35) for the two wavelengths to the training and validation data. The adjusted coefficient of determination (R^2) of modeled intensity at both wavelengths for both training and validation data (Table 2-4) indicates the proposed calibration function and estimated parameters predict the return intensity well. The linear regressions between measured and modeled intensity for both training and validation data (Figure 2-5 and Figure 2-6, row B) yield values of slope very close to unity as well as very small intercepts, indicating very good fits.

The calibration functions (Figure 2-5 and Figure 2-6, A1), which provide the return intensity of a target with unit reflectance, increase sharply and then fall exponentially. The normalized 1064-nm return intensity peaks at ~3.5 m (Figure 2-5, row A) and the 1548-nm peaks at ~5 m (Figure 2-6, row A). The curves of $K(r)$, shown in

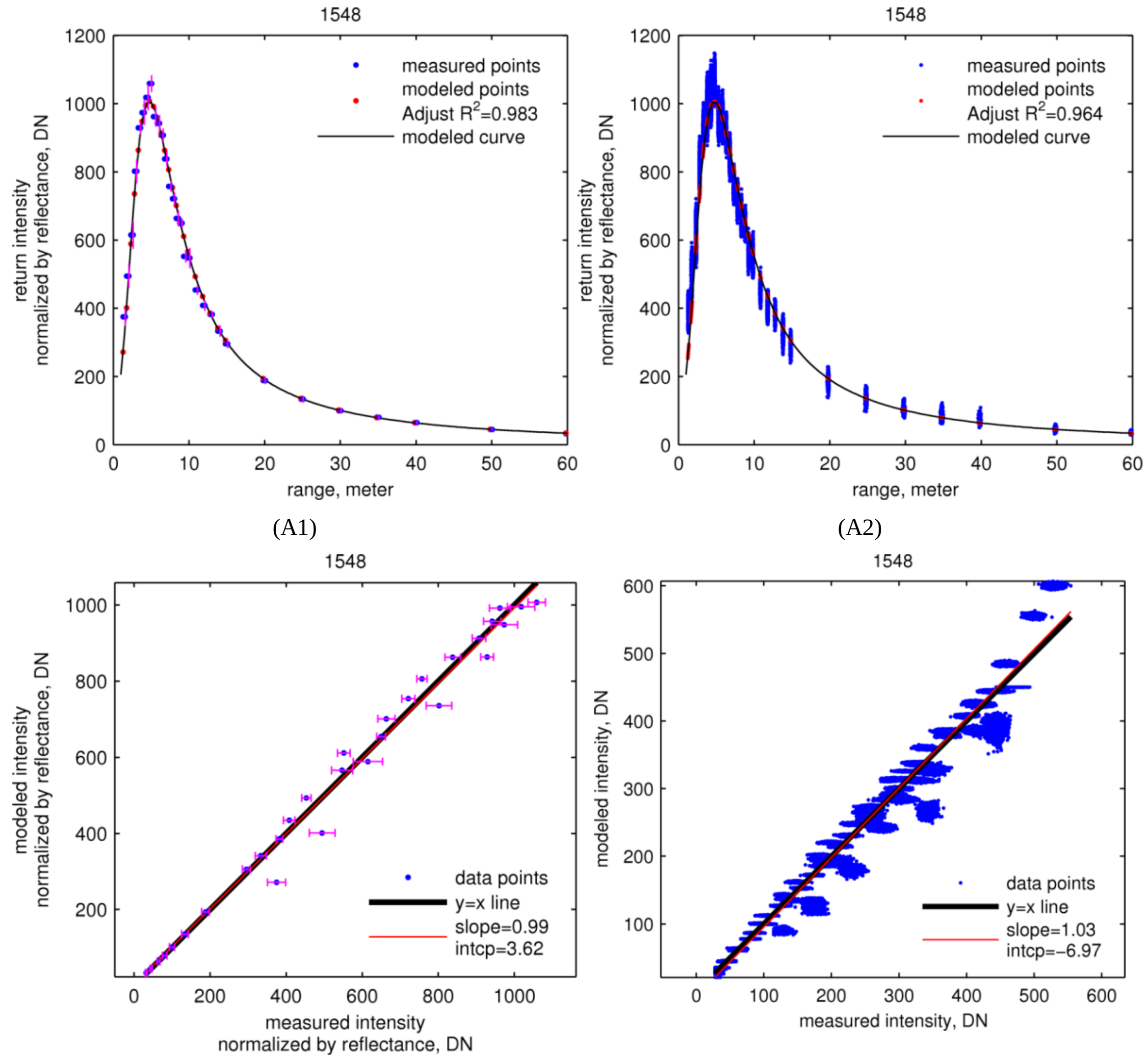
Figure 2-7, rise from zero and plateau at about unity at ~10 m for 1064-nm laser and ~15 m for 1548-nm. The functions differ because of small differences in alignment, and because each laser beam uses individual wavelength-dependent optics in the divergence lens assembly and the detector focusing lens (see Section 2.5.3).



(A1) (A2)

(B1) (B2)

Figure 2-5 Estimation and validation of calibration of 1064 nm data. In rows A and B, the left column shows the calibration function as fitted to training data, and the right column shows the fit to the validation data. Row A: Measured and modeled intensity normalized by reflectance. Row B: scatter plots of measured against modeled intensity. The vertical error bars in (A1) and horizontal error bars in (B1) are one standard deviation of measured intensities normalized by reflectance.



(A1) (A2)

(B1) (B2)

Figure 2-6 Estimation and validation of calibration of 1548 nm data. In rows A and B, the left column shows the calibration function as fitted to training data, and the right column shows the fit to the validation data. Row A: Measured and modeled intensity normalized by reflectance. Row B: scatter plots of measured against modeled intensity. The vertical error bars in (A1) and horizontal error bars in (B1) are one standard deviation of measured intensities normalized by reflectance.

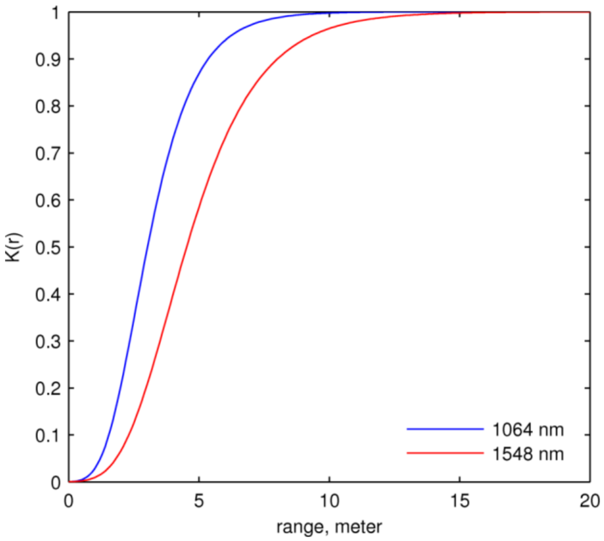


Figure 2-7. Telescope efficiency $K(r)$ of the two wavelengths

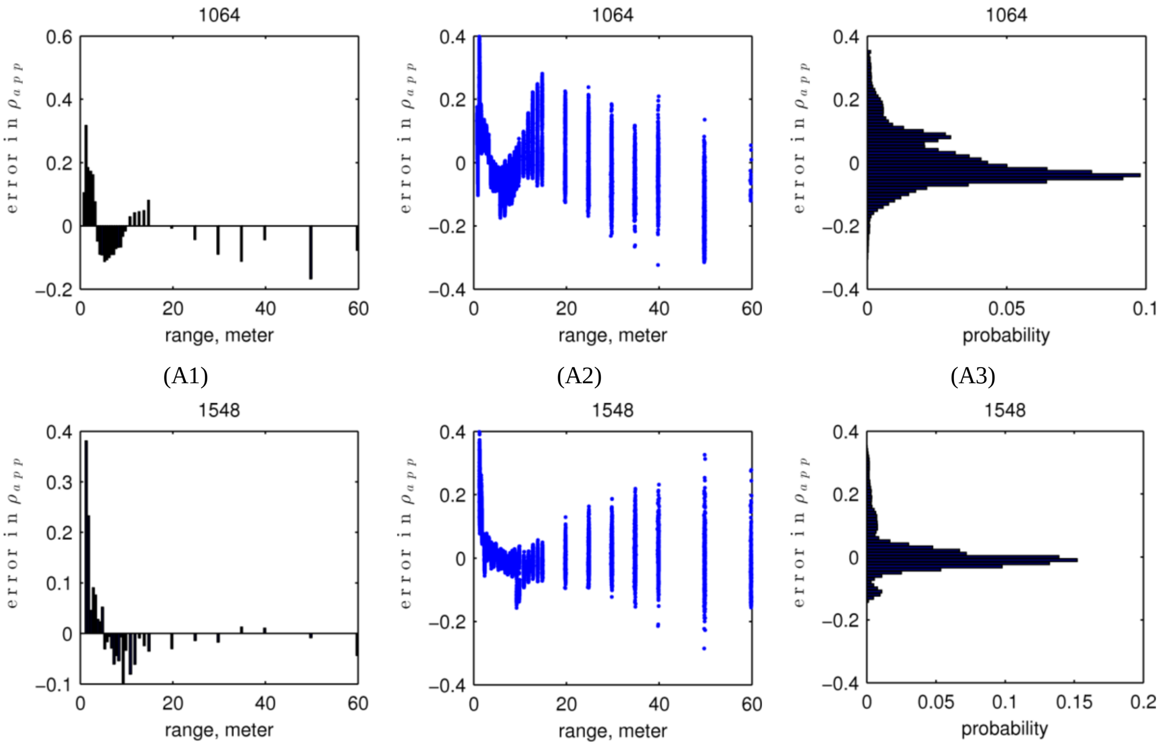
2.5.2.2 Apparent reflectance error

Because all intensities in calibration fitting and validation were normalized by corresponding panel reflectance, the error in ρ_{app} hereinafter means relative error unless otherwise noted. The estimates of apparent reflectance ρ_{app} by the calibration model from validation data show root mean squared errors (RMSE) of 0.092 at 1064 nm and 0.108 at 1548 nm (Table 2-4) (Figure 2-8). The histograms of errors (Figure 2-8, A3, B3) are centered around zero, which indicates no systematic offset in the apparent reflectance estimate.

The plots of errors in estimated ρ_{app} against range for the validation dataset (Figure 2-8, A2, B2) show larger and dispersed errors at very near range (< ~3.5 m for 1064 nm and < ~2 m for 1548 nm) and farther range (> ~10 m for 1064 nm and > ~20 m for 1548 nm), in contrast to smaller and less dispersed errors in between. This pattern of errors in ρ_{app} over range is a combination of errors from range uncertainty (Δr) and return intensity uncertainty ($\Delta\alpha$). We observed how Δr and $\Delta\alpha$ contribute to errors in ρ_{app} separately with our calibration model over the range of our calibration data, 0.5 m to 70 m. We simulated $\delta\rho_{\alpha}$, the relative errors in ρ_{app} only due to different return intensity uncertainty levels (±15 DN) at different ranges by keeping range error at zero (Figure 2-9, A1, B1). Then we simulated $\delta\rho_{r}$, the relative errors in ρ_{app} only due to different range uncertainty levels (±15 cm) at different ranges by keeping return intensity error at zero (Figure 2-9, A2, B2). These relative error ranges in simulation of Δr and $\Delta\alpha$ are more than three times larger than the standard deviation of range measurements and root mean squared noise.

Table 2-4 Assessment of calibration fitting and validation

Wavelength		1064 nm	1548 nm
Measured vs. Modeled Intensity, Adjusted R^2	Training	0.954	0.983
	Validation	0.948	0.964
RMSE of Apparent Reflectance	Training	0.108	0.092
	Validation	0.081	0.064



(A1) (A2) (A3)

(B1) (B2) (B3)

Figure 2-8. Errors in apparent reflectance. Row A shows 1064 nm and Row B shows 1548 nm. The left column is the deviation from calibration fitting with range. The middle column is the deviation of validation points with range. The right column is the histogram of deviations.

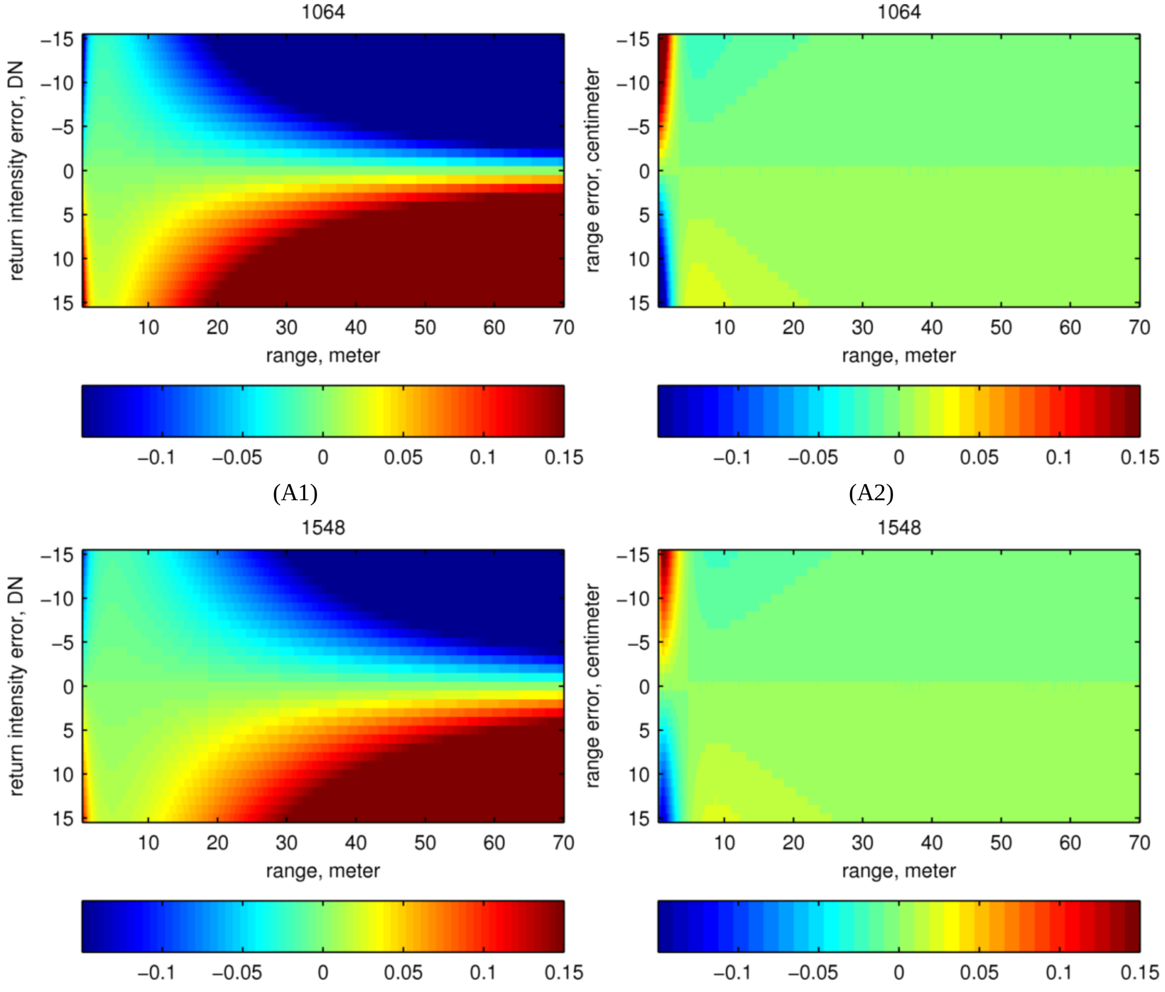
2.5.2.3 Sensitivity of estimated apparent reflectance to intensity and range

Figure 2-9 presents a graphical display of the estimated relative errors calculated using the above procedure. The total relative error in $\rho_{app,}$ $\delta\rho$, can be approximated by $\delta\rho_\alpha + \delta\rho_r$ (see Appendix A.2 for this derivation). Graphs A1 and B1 in Figure 2-9 show the relative error in apparent reflectance ($\delta\rho_\alpha$) produced by changes in return intensity ($\Delta\alpha$) of ±15 digital counts (DN) (y-axis), as the error varies with range. At very near range (<~3 m), small changes in DN produce large errors in apparent reflectance; this effect arises because the telescope efficiency $K(r)$ is very low and the return signal is weak. At near range between about 2 and 10 m, the signal is much stronger and thus the errors produced are fairly small (green colors). Between near and far range (10–70 m), the exponential decrease in signal provides a smooth transition from low sensitivity of apparent reflectance with DN error to high sensitivity. At far range (70 m), the signal is sufficiently diminished by fall-off with range that errors in apparent reflectance are large given a deviation of just a few counts from true values.

In contrast, large relative error in apparent reflectance ($\delta\rho_r$) produced by changes in range (Δr) of ±15 cm (Figure 2-9, A2, B2) (y-axis) is limited to the very near range (<~5 m). The weak signal in this range provides large errors, which decrease rapidly as the telescope efficiency function $K(r)$ increases the signal strength. Beyond this range, the relative error in apparent reflectance remains low.

From this analysis, we see that the error in apparent reflectance due to error in range ($\delta\rho_r$) dominates at near ranges while the error due to return intensity ($\delta\rho_\alpha$) dominates at far ranges. The exact range at which $\delta\rho_\alpha$ surpasses $\delta\rho_r$ and becomes dominant depends on the uncertainty level of return intensity given the calibration model. Thus, we see larger and more dispersed errors at very near ranges in validation data (Figure 2-9, A2, B2) mainly due to range uncertainty, and at far ranges mainly due to return intensity uncertainty. The range accuracy is therefore more critical in the near range target calibration while the return intensity accuracy becomes more critical in the far range target calibration.

The problem for calibration here is that returns from far ranges have a lower signal-to-noise ratio, but their calibration is highly sensitive to return intensity uncertainty. Thus, the noise level of lidar return intensity needs to be characterized to find the range at which the reflectance uncertainty $\delta\rho$ exceeds a desirable level given the calibration model. For returns from far ranges, the apparent reflectance should be used carefully. For returns from near ranges, $\delta\rho$ could be very high if the range uncertainty is not low enough. However, lidar instruments generally give range measurements of high accuracy.



(A1) (A2)

(B1) (B2)

Figure 2-9 Sensitivity of ρ_{app} estimate on errors in return intensity and range. The image color shows relative error in ρ_{app} estimate (estimate - measurement). The color map scale is unified for all images for comparison purpose but the actual error ranges of the four images are different and given here: (A1), $\delta\rho_{\alpha}$ at 1064 nm, [-0.928, 0.928]; (A2), $\delta\rho_{r}$ at 1064 nm, [-0.226, 0.290]; (B1), $\delta\rho_{\alpha}$ at 1548 nm, [-0.574, 0.574]; (B2), $\delta\rho_{r}$ at 1548, [-0.133, 0.154].

2.5.3 Calibration comparison of the two wavelengths

The two telescope efficiency functions $K(r)$ at the two wavelengths (Figure 2-7) suggest different optical characteristics of the two beam pathways through the DWEL instrument. As noted earlier, each pathway uses a separate wavelength-dependent divergence optic and focusing lens in the detector assembly, which can produce small differences in beam width and detector field of view. Moreover, the two laser beams may not be exactly coincident due to small errors in alignment. As a result, the two functions show different shapes.

In addition, the range exponent values b for the two wavelengths are different, but both are smaller than the theoretical value of 2 that applies to scattering from a perfectly diffuse surface. The observed value may depart from 2 for a number of reasons, including slight misalignment of the optical path as it interacts with the Gaussian beam cross section and electronic effects in the detector-amplifier-digitizer systems. In addition, target surfaces may depart from diffuse reflectance and show partial specularity (i.e., bidirectional reflectance effects). In this case, the telescope aperture will see a narrower solid angle of scattering from the target with range, thus including a greater proportion of scattering from a retroreflection peak. Moreover, as a free parameter in the model inversion, the range exponent may be adjusted by the nonlinear fitting procedure to better shape the telescope efficiency function. While a more physical model grounded in instrument optics and first principles of scattering might be desirable, an empirical model will capture the data trend more accurately in the face of physical and electronic unknowns. As shown above, our calibration functions clearly fit the data well, predicting observed intensities from calibration targets and retrieving reflectance from observed intensities with low errors.

2.6 Conclusion

Calibration of our full-waveform, dual-wavelength, terrestrial laser scanner presents a number of challenges relevant to the next generation of terrestrial laser scanners. The need to scan from near to far range requires characterizing both telescopic effects, which reduce the near-range signal with increasing proximity due to defocusing, and saturation effects, which alter the return pulse shape of near-range scattering events. Fortunately, digitizing the full return waveform allows the ready detection of saturation and accurate description of telescope efficiency in the near range, leading to mitigation of both effects.

In addition, dual- or multiple-wavelength data must be consistent in spectral performance. By using an empirical calibration model fitted to data, it is not difficult to add a constraint that optimizes spectral "flatness" with range using a Spectralon target. This step is particularly useful for the DWEL, since the laser wavelengths are chosen specifically for their ability to separate hits of water-bearing leaves from hits of the dry bark of trunks and branches and dry ground surfaces.

Our calibration procedure provides apparent reflectance, a physically-defined value related to the size, orientation, and reflectance of a target, independent of range and instrument optics and electronics. Because apparent reflectance provides gap fraction with range and height, it is used to retrieve plant and leaf area index as well as plant and foliage profiles. Apparent reflectance values also allow algorithms to find and count trunks and measure their diameters at near and far ranges (Lovell et al., 2011; Strahler et al., 2008) without reference to range. These data can then be used in estimating above-ground biomass with appropriate allometric equations (Yao et al., 2011). And, as an instrument-free measure, apparent reflectance facilitates merging point data from multiple instruments, allowing assessment of the effects of angular resolution and beam divergence on structure retrieval and even providing multiple wavelength information for scanners using different laser wavelengths.

To obtain point clouds of apparent reflectance, we solve a linear system of equations established from an empirical system response model, rather than use waveform decomposition with analytical mathematical functions, such as the widely-used Gaussian distribution. This linear system approach to point cloud generation from lidar full waveforms overcomes the noise in the waveforms and accommodates the unique DWEL pulse shape, which is not easy to model with analytical mathematical functions.

The DWEL calibration model explicitly includes the telescope efficiency function $K(r)$, modeled by a generalized logistic function, to correct defocusing of return signals on the detector from near range targets. The calibration model is parameterized by nonlinear fitting to a calibration dataset of intensities returned by targets of known reflectance at known ranges. A sensitivity analysis shows that the apparent reflectance error from radiometric calibration is dominated by range errors at near ranges but by return intensity errors at far ranges. In spite of such errors, the RMSE values (relative errors) of apparent reflectance, 0.081 for 1064 nm and 0.064 for 1548 nm, obtained from separate training and test measurement datasets, show that the parameterized model quite accurately converts lidar return intensities in digital counts to apparent reflectance. This calibration model can apply to almost any terrestrial lidar instrument using a telescope to focus the return power.

While the calibration of a terrestrial lidar is by nature more difficult than that of an airborne lidar, one advantage is that controlled laboratory or field calibration measurements are readily designed and executed. If stationary operation is possible to fix the beam on the target panel, it is easy to acquire pulses at measured ranges in a long corridor or outdoor environment. If stationary operation is not possible, only short scan segments crossing the target need to be acquired. Calibration is also aided by having targets of different reflectance. While our painted panels functioned well, a set of Spectralon panels with well-characterized diffuse reflectance properties ranging from light to dark would be desirable.

In summary, we show how to overcome the challenges posed in calibrating a dual-wavelength terrestrial laser scanner by formulating a flexible calibration model, acquiring appropriate calibration data, fitting the model with a constraint providing spectral consistency, and testing the results and the sensitivity of errors to uncertainties in range and intensity. We also provide solutions to the problems of saturated returns, slow change in laser output pulse energy, and variance in the timing of laser pulse emissions. The next step is to use calibrated data to retrieve forest structural parameters with the new dual-wavelength data, following the pathways pioneered with the heritage Echidna Validation Instrument, but extending them to new information from the Dual-Wavelength Echidna Lidar. This is the subject of the following two chapters.

CHAPTER 3. SEASONAL CHANGE OF LEAF AND WOODY AREA PROFILES IN A MIDLATITUDE DECIDUOUS FOREST CANOPY FROM CLASSIFIED DUAL-WAVELENGTH POINT CLOUDS

3.1 Introduction

Forest canopy structure regulates radiation interception through the canopy, affects the canopy microclimate, and consequently influences the energy, water, and carbon fluxes between soil, vegetation and atmosphere through its interactions with leaf physiological functioning (Baldocchi et al., 2002). Leaf area index (LAI), defined as half of total leaf surface area (simply referred as ‘leaf area’) per unit ground area (Chen & Black, 1992), is one of the primary canopy structural measures used in both ecophysiological and remote sensing models to govern the radiation interception through forest canopy and the capacity of canopy photosynthesis (Baldocchi et al., 2002; Bonan, 1993; Hanson et al., 2004; Medvigy et al., 2009; Running & Coughlan, 1988). In addition to LAI, simulations of net ecosystem productivity (NEP) have also shown the necessity of including detailed two- and three-dimensional distribution of leaf areas, e.g. vertical foliage profile, in ecophysiological modeling especially for open canopies and multi-layered stands (Law et al., 2001). Measurements of vertical foliage profile are closely related to (Parker et al., 2004) and significantly correlated with (Stark et al., 2012) forest functioning measures, which also demonstrates the importance of the three-dimensional distribution of leaf area.

LAI and vertical foliage profile are typically measured across different spatial scales. LAI over large areas has been derived from the inversion of radiative transfer modeling from optical remotely sensed imagery (Koetz et al., 2005; Myneni et al., 2002) and by using empirical relationships between field-measured LAI and various metrics from remotely sensing data. Examples include vegetation indexes from optical imagery (Cohen et al., 2003); backscatter cross section and polarization ratio from synthetic aperture radar (Manninen et al., 2005); laser return energy and canopy heights from airborne and space-borne lidar (Korhonen et al., 2011; Lefsky et al., 1999; Luo et al., 2013; Solberg et al., 2009); and combinations of above (Ma et al., 2014). Airborne and spaceborne full-waveform lidar data have demonstrated the capability to derive vertical foliage profiles over large areas, either relative (Lefsky et al., 1999) or absolute (Harding et al., 2001; Ni-Meister et al., 2001; Tang et al., 2014, 2012).

All the above methods to derive LAI and vertical foliage profile over large areas need ground truth data to calibrate and validate their empirical and physical retrieval models. Thus, the quality and detail of the ground truth data are crucially important. Various ground-based LAI measurement techniques have been developed. The major methodologies for ground-based LAI measurements generally fall into two categories: direct, which involves destructive sampling or litter-fall collection, and indirect, which involves tree allometry, or gap probability measurements (Asner et al., 2003; Bréda, 2003; Chen et al., 1997; Jonckheere et al., 2004). Direct destructive sampling and litter-fall collection are time-consuming, laborious and sometimes impractical, e.g. in areas of preservation or scientific interest where destructive sampling is prohibited (Jonckheere et al., 2004). Also litter-fall collection only works for deciduous trees. The indirect allometry-based method requires site-specific allometric equations from detailed destructive sampling. Indirect gap-based methods invert gap probability in the canopy to LAI based on Beer's law of light transmission through an absorbing medium (Chen et al., 1997; Lang, 1987; Welles & Cohen, 1996). The gap probability can be obtained by a point quadrat method using a thin probe that passes through canopy (Wilson, 1960) or noncontact optical techniques such as hemispherical photos (HP), radiation measurement devices (e.g. DEMON, ceptometers, LAI-2000 and TRAC), and ground-based laser technology (Bréda, 2003; Jonckheere et al., 2004; Jupp et al., 2009; Zhao et al., 2012). The point quadrat method is impractical in forest stands because of the tall stature of trees and the high density of leaves (Chen et al., 1997). Indirect noncontact optical methods are quick, low-cost and more commonly used.

Ground-based measurements of vertical foliage profiles date back to early work, including stratified clipping and inversion of leaf contact frequency measured by point quadrats or a camera with telephoto lens (Aber, 1979; MacArthur & Horn, 1969). Vertical foliage profile has also been obtained by taking LAI measurements using hemispherical photography acquired from a crane gondola with increasing canopy height (S. B. Weiss, 2000). All these early methods are time-consuming, inconvenient and sometimes impractical. Recent ground-based active optical methods with terrestrial laser scanners (TLS) have demonstrated great potential to measure gap probability with canopy height from lidar range in 3-D space, and thus can retrieve vertical foliage profile readily and accurately (Béland et al., 2011; Calders, Armston, et al., 2014; Jupp et al., 2009; Zhao et al., 2011).

However, gap probability measurements by these optical methods to measure LAI or vertical foliage profile typically include both leaves and woody materials, and thus actually measure plant area index (PAI, leaves and woody materials) and its vertical profile. The contribution of woody material to LAI measurements is usually removed with an empirical woody-to-total ratio. Kucharik et al. (1998) found the nonrandom positioning of branches/stems with regard to leaves causes inaccurate LAI with this simple ratio correction, especially when branches/stems are not preferentially shaded by leaves. They removed the woody contribution directly with a Multiband Vegetation Imager. But this approach cannot correct vertical foliage profiles to remove the woody contribution. The separation of leaves from woody materials in 3-D space is needed to remove the woody contribution to vertical foliage profiles derived from gap probability measurements.

Moreover, the separation between leaves and woody materials in 3-D will also improve the simulation and inversion of ecophysiological and 3-D radiative transfer models. Kobayashi et al. (2012) found the effect of woody elements on energy balance simulation in ecophysiological modeling is not negligible for a heterogeneous landscape because of the radiation absorption and heat storage by woody elements. Some studies have also shown that the explicit inclusion of woody elements in the 3-D radiative transfer models improves canopy reflectance modeling and thus model inversion to estimate both biophysical (Malenovský et al., 2008) and biochemical variables (Verrelst et al., 2010) at high resolution.

Three-dimensional scans of forests by terrestrial lidar capture provide the potential of separating leaves from woody materials. However, currently there are only a few studies on the classification of leaves and woody materials in 3-D space. Béland et al. (2014) separated leaves and woody materials using lidar return intensities from a discrete TLS operating in the shortwave infrared (1535 nm – 1550 nm), where leaf and woody reflectance are significantly different. Yang et al. (2013) used lidar return pulse shapes from a full-waveform TLS for classification. However, the compound effects of reflectance, size and orientation of targets may generate similar return intensities or return pulse shapes from different target classes. For example, a large leaf fully intercepting the laser beam orthogonally would give the same return intensity as a woody surface that is half the size of the leaf but twice as reflective, if intercepting the laser beam orthogonally at the same range. Also, a trunk edge intercepting a laser beam at an oblique angle can give a similar return pulse shape as a leaf cluster (Yang et al., 2013).

To address the challenge to the classification of leaves and woody materials from 3-D lidar scanning data posed by the compound effects of reflectance, size and, orientation of targets, we built the Dual-Wavelength Echidna Lidar (DWEL), a terrestrial lidar that acquires full-waveform scans at both near-infrared (NIR, 1064 nm) and shortwave infrared (SWIR, 1548 nm) wavelengths with simultaneous laser pulses (Douglas et al., 2015). The difference in return signals at the two wavelengths from a target is theoretically only affected by the difference in its reflectance at the two wavelengths and thus provides a way to resolve the classification challenge posed by optical and geometric properties of targets.

This study tested the ability to separate leaves from woody materials in 3-D space and derive separate leaf and woody area profiles using novel dual-wavelength scans of a midlatitude forest site by DWEL in both leaf-off and leaf-on conditions. To carry out this objective, we first developed a 3-D classification approach using both spectral and geometrical information from the dual-wavelength lidar data, and assessed our classification in 3-D space with an indirect classification accuracy assessment procedure. Next we derived separate leaf and woody area profiles from both leaf-off and leaf-on scans. Lastly, we assessed the error sources of the 3-D classification and the implications of the separation of leaf and woody profiles.

3.2 The Dual-Wavelength Echidna Lidar

The Dual-Wavelength Echidna Lidar (DWEL), based on the design of the Echidna Validation Instrument (EVI) built by Australia's Commonwealth Scientific and Industrial Research Organization (CSIRO) (Jupp et al., 2005; Parkin et al., 2001), uses two pulsed lasers and acquires full-waveform scans at both near-infrared (NIR, 1064 nm) and shortwave infrared (SWIR, 1548 nm) wavelengths with simultaneous laser pulses (Douglas et al., 2015). DWEL uses a rapidly rotating zenithal scan mirror and a slowly rotating azimuth platform to provide full coverage of the angular scan space. Each rotation of the scan mirror directs the laser beam through 360°; returns from the environment are acquired at zenith angles from –117° to +117° (234° extent), while returns from the instrument housing, used for calibration, are acquired as the laser beam passes through angles of +117° to –117° (126° extent). For azimuth, the instrument platform rotates through 180°, thus providing a complete spherical scan. The scanning resolution was set 2 mrad with slightly larger beam divergence, 2.5 mrad to ensure continuous coverage of the hemispheres for the scans used in this study. DWEL detects and digitizes the return signal at 2 GHz, i.e. every half nanosecond, and records the returns as full waveforms from a range up to 70 m. It samples return waveforms at the pulse repetition rate (PRF) of 2 KHz.

3.3 Study Area and Data

For this study, we established a 100-m-by-100-m deciduous forest site (N 42° 31′ 51.48″, W 72° 10′ 55.56″) at Harvard Forest in central Massachusetts, USA. This generally flat 1-ha site is dominated by red maple (*Acer rubrum*), red oak (*Quercus rubra*) and white birch (*Betula papyrifera*), with an understory of these species accompanied by American beech (*Fagus grandifolia*), American chestnut (*Castanea dentata*) and others. A few large white pines (*Pinus strobus*) and several hemlocks (*Tsuga canadensis*) are also present within the plot. At five circular plots of 20-m radius (Figure 3-1), we collected biometric data including diameter at breast height (DBH), species, location (range and compass azimuth from the circular plot center), and crown position. For a subsample of trees, we also acquired tree heights, crown diameters at two orthogonal dimensions, and crown heights. In data of September 2014, the average stem density at this site was 769 trees ha^{-1} and the basal area was around 38.5 m^{2} ha^{-1}. The average tree height of sampled trees at this site was 20.3 m and the average crown diameter was 8.7 m.

We scanned the forest site with the Dual-Wavelength Echidna Lidar (DWEL) at five scan locations (Figure 3-1) with the deciduous trees in both leaf-off (May 3rd 2014) and leaf-on (Sept 19th 2014) conditions. The wind speed on both scanning days was low with some light leaf motion at the canopy top. DWEL waveform data of each collected scan were processed and calibrated to two point clouds of apparent reflectance (ρ_{app}) at the two wavelengths (see Chapter 2).

We measured effective plant area index and the clumping index with a TRAC instrument (Chen & Cihlar, 1995) along two 100-m transects through the site center in the leaf-off season, but on different days from the DWEL scanning dates due to the availability of the TRAC instrument. On the same dates as DWEL scanning, we also took hemispherical photos at the plot centers and scan locations, as well as at 10-m intervals along the two TRAC transects. The hemispherical photos provided plant area index measurements (including both leaves and woody materials) and indirect comparison with DWEL scanning images for classification accuracy assessments. In June 2014, between leaf-off and leaf-on scanning, we collected spectral measurements of green-leaf and bark samples of dominant tree species at the site with an ASD FieldSpec spectrometer and plant probe.

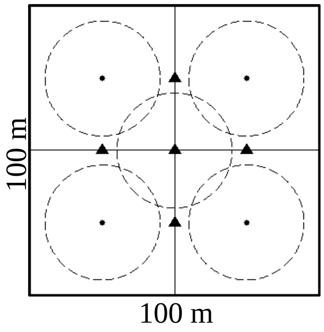


Figure 3-1. DWEL scans were collected at the five solid triangles. Tree measurements were collected at circular plots centered at the four dots and the center triangle. The dashed circles represent the coverage of each tree measurement plot.

3.4 Methodology

3.4.1 Classification of dual-wavelength point clouds

3.4.1.1 Generation of dual-wavelength point clouds of forests

Two point clouds of apparent reflectance (ρ_{app}) at the two wavelengths from a single DWEL scan were merged together to provide dual-wavelength ρ_{app} values for each point. This procedure first identified two types of points: matched points between NIR and SWIR point clouds and unmatched points. Here, matched points have return signals at both NIR and SWIR wavelengths, while unmatched points only have return signals at one wavelength but not the other. Unmatched points may arise in four situations:

1. A target has much lower reflectance at one wavelength than the other, and it generates a return signal that is not detected by the instrument at one wavelength but detected at the other. This is the most common cause of unmatched points.
2. The return signal is lost in random transient electronic noise at one wavelength, but not at the other.
3. The laser beam at one wavelength sees the edge of a target at far range but the other laser misses the edge due to a slight misalignment between the two laser beams.
4. A noise signal at one wavelength is erroneously taken as a return signal, while the noise at the other wavelength is correctly removed or does not occur. Our

waveform preprocessing uses aggressive criteria to eliminate noise points in the generation of point clouds, and this cause is the least common.

The reflectance information at a single wavelength from unmatched points is still helpful for the classification of leaves and woody materials and the sounding of the forest canopy, even though unmatched hits lack information from the second wavelength. To add this information and provide a point cloud with bispectral information for each point, we synthesize ρ_{app} in the missing wavelength based on the reflectance information from matched points in the same laser shot or neighboring laser shots. This synthesis of ρ_{app} at the signal-missing wavelength uses two assumptions: (1) all targets sensed by an individual laser beam are of the same material and therefore share similar reflectance values in the two wavelengths; and (2) laser shots close to each other see a similar composition of targets.

The procedure for merging the two point clouds into a single, dual-wavelength point cloud utilizes the following steps.

S1. Search for matching points in NIR and SWIR point clouds by comparing the laser shot numbers and ranges of points at the two wavelengths. If two points from the two wavelengths have the same laser shot number and their range difference is less than a given threshold, they are matched. Their coordinates are averaged and output as a single point with dual-wavelength ρ_{app} values. Otherwise, points are identified as unmatched.

S2. For unmatched points, synthesize the ρ_{app} value at the missing wavelength using the normalized difference index (NDI) value of the laser shot containing the point. The NDI of the laser shot is defined by equation 3-1, where ρ_{app}^{nir} and ρ_{app}^{swir} are average apparent reflectance values at NIR and SWIR wavelength of points in the laser shot. If the shot already contains one or more matched points between the two wavelengths, the NDI of the laser shot is calculated from average apparent reflectance values of those matched points. Otherwise, the NDI of the laser shot is interpolated from five closest neighboring shots. Given the NDI, equation 3-1 is solved for the missing apparent reflectance.

$$NDI = \frac{\rho_{app}^{nir} - \rho_{app}^{swir}}{\rho_{app}^{nir} + \rho_{app}^{swir}} \qquad 3\text{-}1$$

3.4.1.2 Synergistic use of spectral information and spatial context for classification

We used both spectral information from NIR and SWIR apparent reflectance and spatial context information given by the 3-D spatial distribution pattern of each point and its neighbors to classify points into leaves and woody materials. Our unsupervised spectral classifier, the K-means clustering algorithm implemented in Scikit-learn package (Pedregosa et al., 2011), first divided all points into 100 clusters using three variables: apparent reflectance values at the two wavelengths and the normalized difference index value. Each cluster was viewed interactively with a 3-D display program and labeled as composed of leaf points, woody points, or mixed points including both leafy and woody hits. Points in mixed point clusters were further classified using the spatial context information of the point and surrounding points according to multiscale dimensionality criteria, described further below (Brodu & Lague, 2012).

While the spectral information has the capacity to classify many points of leaves and woody materials using the contrast in spectral reflectance between NIR and SWIR returns, leaves and woody materials may sometimes be confused for a number of reasons, of which four are identified below.

1. Some leaves and woody materials may show an atypical contrast between NIR and SWIR apparent reflectance, depending on water content. For example, some young stems or newly-grown branches may have thin or green bark with a higher water content; or unhealthy or senescent leaves may have a lower water content.
2. Many tree trunks at our study site exhibit lichens growing on bark surfaces. Arboreal lichens are typically moist, and their SWIR absorption can produce a leaf-like signal for a trunk point.
3. Error in the calibration of raw lidar return intensities to apparent reflectance may reduce the contrast in NIR and SWIR apparent reflectance, leading to classification of a point into a mixed cluster.
4. Laser beam misalignment can cause the two laser beams to have slightly different paths into the canopy and therefore interact with different mixtures of leafy and woody surfaces. Since leaves and wood have similar reflectance values in the NIR, this error will primarily affect SWIR apparent reflectance values and may lead to classification into a mixed cluster. As previously noted, misalignment at the edge of a leaf or branch may also produce a hit in one wavelength but not in the other. While the

alignment error is less than the scanning resolution (2 mrad for scans in this study), this type of confusion could happen for small targets at longer range.

When the spectral information is not sufficient to correctly classify some points in the cases mentioned above, spatial context information can augment the classification. The spatial context of a point, given by the multi-scale dimensionality of the local point cloud, characterizes the local 3-D organization of the points within spheres of different diameters (scales) centered on the point of concern; the spatial organization varies from 1-D (points set along a line), to 2-D (points set on the surface of a plane), and 3-D (points located freely in a 3-D volume) (Brodu & Lague, 2012). Trunk, branch, and ground surfaces produce 1- and 2-D organizations, while leaves in the canopy show a 3-D organization.

While this technique works well for true woody points with equivocal spectral values located on or near larger branches or trunks, it may also increase the number of leaf hits within the canopy. For example, a fine branch point high in the canopy with a weak woody signal may be surrounded by a 3-D pattern of leaf points and thus classified as a leaf point based on spatial context, and we may anticipate some commission errors in leaf hit counts.

3.4.1.3 Classification accuracy assessment

Classification accuracy assessment generally calls for a reference data source of higher quality than data used to create the classification (Olofsson et al., 2014). However, currently there is no reference classification of points in 3-D of higher quality for our study site. Accordingly, we developed an indirect approach to accuracy assessment of our 3-D point classification using color hemispherical photos and manual photointerpretation selected according to a random stratified sample design.

We first projected 3-D point clouds of dual-wavelength apparent reflectance and classification into a 2-D hemispherical projection of 2043 by 2043 pixels. The class label of a projection pixel (leaf or wood) is assigned as the mode of class labels of points projected into the pixel. The apparent reflectance value of a projection pixel is assigned as the average apparent reflectance value of points projected into the pixel. We set the view point of each hemispherical projection at a standard height so that the projection images from different scans collected at slightly different heights are directly comparable[2].

As our reference data, we used the true-color hemispherical photos taken at the same scan locations and the false-color projection images from DWEL scans as the reference data sources. We registered the hemispherical photos to the hemispherical projection images of the points with the ENVI registration module using identifiable branch forks and crossings in images as tie points. However, due to slight distortions in the photographic hemispherical projections, the registered hemispherical photos and projection images of points could not be aligned exactly together. Thus, visual interpretation using image context was required to give a reference label to each selected validation pixel. To reduce any possible bias in the labeling, we chose primary and

---

[2] The side effect is the hemispherical projection image produced in this way could have some artificial gap pixels if the height of scanning is different from the chosen height of the view point of the hemispherical projection.

secondary reference labels for each validation pixel. The primary validation label generated the best possible accuracy, while the secondary label generated the worst possible accuracy. The primary reference labeling always tried to follow the mapped class label unless we were very confident in a change of label to the other class (our classification is two-class). The secondary reference labeling always tried to change the mapped class label unless we were very confident in retaining the class label.

We selected validation sample pixels for each scan using stratified random sampling with the DWEL-labeled classes as strata, and then merged validation samples from the five scans of the site together to generate the classification error matrix. The total sample size for five scans together was determined according to Olofsson et al. (2014, 2013) with expected user's accuracy of the two classes and targeted standard error of overall accuracy estimates (Table 3-1). The allocation of samples to each strata of classes followed the recommendation by Olofsson et al. (2014) to balance the standard errors of user's accuracy estimates for rare classes and overall accuracy estimates. In this procedure, 75 samples of the total were allocated to the rare class in two-class classification. Here, leaf was considered the rare class in the leaf-off classification while wood was considered the rare class in the leaf-on classification.

Table 3-1. Design of validation sample (projection pixel) size

Scanning season	Expected $\hat{U}_i$[1]		Targeted $S(\hat{O})$[2]	Total sample size	Sample allocation	
	Leaf	Wood			Leaf	Wood
Leaf-off	0.6	0.8	0.02	416	75	341
Leaf-on	0.8	0.6	0.02	454	379	75

[1]: user's accuracy for the *i*-th class

[2]: standard error of overall accuracy estimate

Our goal was to establish the accuracy of the label of each point; however, our validation label applied to each projection pixel, not to each point. To overcome this problem, we first obtained a classification error matrix by comparing the validation label to the modal value of the classified points within the projected pixel. We then converted it to an error matrix in number of points based on a simple enumeration approach considering all possible distributions of individual point labels that would provide a mode matching the observed label (see Appendix A.3 for a more detailed description). Finally, to establish the range of accuracies, we compared outcomes using primary and secondary validation labels.

3.4.2 Estimation of leaf and woody area profiles

Most indirect measurements of leaf area index (LAI) and leaf area volume density (LAVD) rely on gap probability with height or range, the probability that a light beam or a thin probe penetrating the canopy will have no contact with a vegetative element as height or range increases (M. Weiss et al., 2004). The LAI and LAVD are related to gap probability as (Jupp et al., 2009; Miller, 1967):

$$P_{gap}(\theta, z) = e^{-G(\theta)\Omega(\theta)L(z)/\cos\theta}$$

$$L(z) = \int_{z_0}^{z} f(z')dz' \qquad \text{3-2}$$

$$L(z) = 2\int_0^{\pi/2} \frac{-\log\left(P_{gap}(\theta, z)\right)}{\Omega(\theta)} \cos\theta \sin\theta \, d\theta$$

In these equations, $P_{gap}(\theta, z)$ is the gap probability in the direction of zenith angle θ from height z_0 where the measurements are taken (usually ground) up to canopy height z. The function $G(\theta)$ is the Ross G-function (Ross, 1981), which describes the mean projection of a unit area of vegetative elements in the direction θ. The variable $\Omega(\theta)$ is the clumping index, which is determined by the nonrandomness of leaf spatial distribution (Chen & Cihlar, 1995) and describes the extent to which LAI retrieved using a random model of leaf placement (usually termed effective LAI, LAIe) differs from the true LAI (Zhao et al., 2012). The variable $L(z)$ is the LAI up to the canopy height z. The function $f(z)$ is the vertical profile of LAVD, or the horizontally averaged LAVD as a function of canopy height z.

Although the above equations are formulated in terms of leaf area index and leaf area volume density, most optical methods measure gap probability of all vegetative elements without differentiating leaves and woody materials; that is, plant area index (PAI), including both leaves and woody materials, is actually measured. LAI is then calculated from PAI using an empirical woody-to-total ratio usually obtained by destructive sampling, either for a given site or more generically for a particular vegetation type.

Classification of leaves and woody materials in 3-D with dual-wavelength point clouds from DWEL enables measurements of the gap probabilities of leaves and woody materials separately, thus providing leaf area index and woody area index (WAI) directly without a woody-to-total ratio. Similar to the definition of LAI for cylindrical shaped leaves such as needles (Chen & Black, 1992), WAI is defined as one half the total surface area (hemisurface area) of woody components per unit ground area (Kucharik et al., 1998). Hereafter, we will use plant area index (PAI) to represent the total hemisurface area index of all vegetative elements including both leaf area index (LAI) and woody area index (WAI), i.e. PAI=LAI+WAI. We will use vegetation area index (VAI) to mean PAI, LAI, or WAI generally.

3.4.2.1 Apparent reflectance and gap probability

We calibrated DWEL lidar return intensities in digital counts at both wavelengths to apparent reflectance (ρ_{app}) following the procedures in Chapter 2:

$$\rho_{app}(r) = \frac{I(r) \cdot r^b}{C_0 \cdot K(r)} \tag{3-3}$$

In this equation, $I(r)$ is the lidar return intensity in digital counts from targets at range r and C_0 is the instrument calibration constant. The telescope efficiency factor, $K(r)$, is needed by geometric laser systems using a telescope to focus the return power (Halldórsson & Langerholc, 1978; Measures, 1991). It is theoretically zero at zero range (the focal point of the telescope) and rises to unity at the range at which the focused return beam falls entirely within the detector. The exponent of range, b, describes the decrease of return intensity over range. While b is theoretically 2 for a diffuse target, the exponent of range can deviate from 2 due to anisotropy of actual targets and electronic effects, and has been taken as a parameter that is estimated by our calibration procedure.

The apparent reflectance, ρ_{app}, can be interpreted as the equivalent reflectance of a diffuse and partially-absorbing panel filling the laser beam perpendicularly that would return the same intensity as the actual target. For ρ_{app} averaged over all azimuths for a small range of zenith angles (a zenith ring), we identified the following computable quantity $I_a(\theta, r)$ (Jupp et al., 2009):

$$I_a(\theta, r) = 1 - \frac{1}{\rho_d}\int_0^r \rho_{app}(\theta, r')dr' \qquad 3\text{-}4$$

where ρ_d is the normal reflectance of a vegetative facet and I_a provides a range-based measure of the stepwise reduction in the power of outgoing laser beam (Jupp et al., 2009). The measure I_a will equal the P_{gap} if all illuminated facets are diffuse and perpendicular to the laser beam. If not, which is the usual case, I_a needs to be scaled to remove the effect represented by the so-called canopy "phase function" (Jupp et al., 2009) to give P_{gap}. The canopy phase function describes the interaction between bidirectional reflectance and the angular distribution of vegetative facets. We scaled I_a so that $P_{gap} = 0$ for targets that fully intercept a laser beam perpendicularly and $P_{gap} = 1$ for pure sky laser shots.

Given two lasers and two types of hits, four different $P_{gap}(r)$ functions were derivable from the point clouds (Figure 3-2).The four $P_{gap}(r)$ functions estimated were $P_{gap}^{L,NIR}$ and $P_{gap}^{L,SWIR}$ for leaves and $P_{gap}^{W,NIR}$ and $P_{gap}^{W,SWIR}$ for woody materials. For the normal reflectance ρ_d at NIR and SWIR wavelengths, we used reflectance values measured by an ASD FieldSpec VNIR spectrometer with plant probe and averaged for the two dominant tree species at our site, red maple and red oak (Table 3-2).

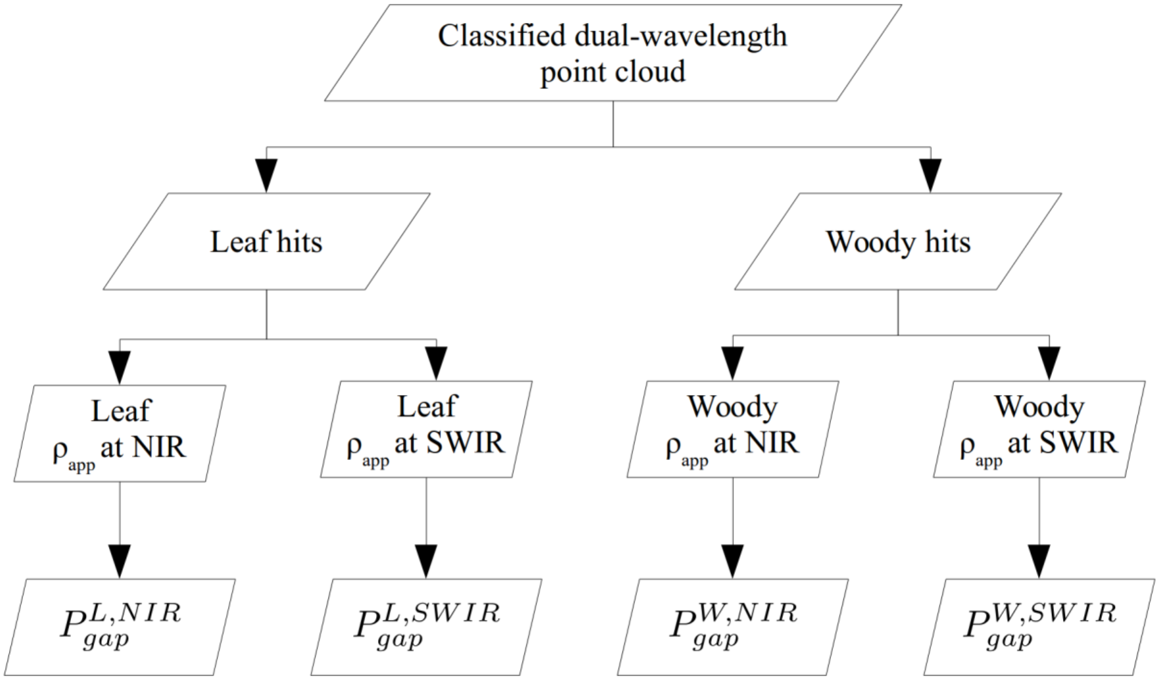


Figure 3-2. Scheme of four estimates of gap probability brought by leaves and woody materials from apparent reflectance at the two wavelengths.

Table 3-2. Normal reflectance of vegetative elements in P_{gap} estimation.

Vegetative element	NIR reflectance		SWIR reflectance	
	Mean	SD	Mean	SD
Green leaf	0.413	0.036	0.284	0.023
Trunk surface	0.541	0.126	0.540	0.033

3.4.2.2 Leaf and woody area profiles

The four $P_{gap}(r)$ functions, estimated as above, provided four vegetation area profiles. However, we found that the profiles for the NIR and SWIR wavelengths were virtually identical (see Section 3.5.2) and thus averaged them to derive LAI and WAI profiles independent of wavelength. To estimate effective LAI (LAIe) and WAIe, we used a linear regression method from gap probability out of the canopy top (height H) in multiple view directions (zenith rings) (Jupp et al., 2009; Zhao et al., 2011). This method is derived using a simple model for the Ross G-function, and thus the LAI can be regarded as largely corrected for leaf angle effects. For simplicity, we use L for both leaf area and woody area in the equations below:

$$\begin{gathered} L_h + L_v X(\theta) \approx -\log P_{gap}(\theta, H) \\ X(\theta) = \frac{2}{\pi} \tan\theta \\ LAI_e = L_h + L_v \end{gathered} \tag{3-5}$$

A solid-angle-weighted averaging approach was suggested by Jupp et al. (2009) to estimate the vertical profile of LAVD:

$$\begin{gathered} \frac{L(z)}{LAI} = \frac{\log P_{gap}(\bar{\theta}, z)}{\log P_{gap}(\bar{\theta}, H)} \\ f(z) = LAI \frac{\partial}{\partial z}\left(\frac{\log P_{gap}(\bar{\theta}, z)}{\log P_{gap}(\bar{\theta}, H)}\right) \end{gathered} \tag{3-6}$$

In these equations, the notation $\bar{\theta}$ means the ratio of the logarithms has been averaged over zenith ranges weighted by solid angle size, rather than a mean angle. Because the ratio provides a means to obtain a profile largely independent of clumping (Jupp et al., 2009), a true LAI value applied to equation 3-6 gives true leaf area profile, while an effective LAI value gives an effective leaf area profile. Here we used an effective vegetation area index (LAIe or WAIe), so that our following investigation of mutual occlusion between leaves and woody materials avoids the complication brought by potentially different clumping indexes for leaves, woody materials, and all vegetative elements considered together.

Estimating LAIe and WAIe independently assumes no mutual occlusion between leaves and woody materials. However, a simple analysis can determine the possible extent of such occlusion, given leaf-off and leaf-on scans. For leaf-off scans, no mutual occlusion is a reasonable assumption; leaf hits are a small proportion of the total and most are hits of evergreen white pine branchlets that are clustered at twig tips. For leaf-on scans, leaves will occlude many branches and stems, while few stems and branches will occlude leaves. A reasonable assumption is thus that occlusion of leaves by woody materials is zero, but the woody materials will be significantly occluded. If we assume that the true woody plant area remains the same in leaf-on condition as in leaf-off condition, the difference in two observed woody areas will estimate the proportion of woody materials occluded by leaves in the leaf-on scan.

For a robust linear regression between $X(\theta)$ and $-\log P_{gap}(\theta, H)$ to estimate L_h and L_v (equation 3-5), gap probability to the canopy top, $P_{gap}(\theta, H)$, from a sufficient number of zenith rings is required. However, our current scans were not able to provide zenith rings of P_{gap} to the canopy top beyond about a 35° zenith angle because of limited measurement range due to insufficient laser power and high signal noise. To overcome this shortage of current DWEL data while taking advantage of range-based leaf-woody separation, we extracted PAIe for the whole study site from the hemispherical photos using CAN-EYE software (M. Weiss & Baret, 2013) and split PAIe to LAIe and WAIe according to a wood-to-leaf ratio λ estimated from each DWEL scan in the following way:

$$\lambda = \frac{\log P_{gap}^{W}(\bar{\theta}, H)}{\log P_{gap}^{L}(\bar{\theta}, H)}$$

$$LAI_e = PAI_e \cdot \frac{1}{1+\lambda} \qquad \text{3-7}$$

$$WAI_e = PAI_e \cdot \frac{\lambda}{1+\lambda}$$

The notation $\bar{\theta}$ in the first equation here has the same meaning as that in the equation 3-6. Note that the calculation of wood-to-leaf ratio λ assumes leaves and woody materials have the same *G*-function; while somewhat restrictive, this assumption is probably better than accepting a value from the literature in the absence of destructive sampling. We then substituted LAIe and WAIe in equation 3-6 to derive the vertical profiles of effective LAVD and WAVD from the corresponding gap probability functions, P_{gap}^{L} and P_{gap}^{W}, to obtain $f_e^L(z)$ and $f_e^W(z)$. The sum of $f_e^L(z)$ and $f_e^W(z)$ gave $f_e^P(z)$, the vertical profile of effective plant area volume density (PAVD).

3.4.2.3 Variance of leaf, woody and plant area profiles

We examined the variance of leaf, woody and plant area profiles, $f_e^L(z)$, $f_e^W(z)$ and $f_e^P(z)$ from two sources: classification error and heterogeneity of canopy structure at the study site.

For variance from the classification error, we generated 1000 resamples of the classification of dual-wavelength point clouds from five scans in each season (leaf-off, leaf-on) based on the classification user’s accuracy (see Appendix A.4 about the resampling procedure of a point cloud classification). These resamples of classification then resulted in 1000 profiles of $f_e^L(z)$, $f_e^W(z)$ and $f_e^P(z)$. We calculated the variance of these profiles as the variance in the vegetation area profiles due to classification error.

For variance from the heterogeneity of canopy structure at the study site, we calculated the variance of the five sets of vegetation area profiles from the five scan locations at the site for each season.

3.5 Results

3.5.1 Classification of dual-wavelength point clouds

Figure 3-3 shows the color-composite images of apparent reflectance and classifications in hemispherical projections along with the registered hemispherical photos for both leaf-off (left column) and leaf-on (right column) seasons from the center plot. In the color composite images (first row of Figure 3-3), leaves show green colors, while trunks and big branches show a spectrum of greenish-yellow, yellow, and brown colors. Hemispherical photos are in true color; in the classified images, woody points are red, while leaf points are green.

Five zoom-in areas (Figure 3-4) of the color composite, classification, and hemispherical photos demonstrate several classification challenges, as well as the strengths of our classification approach. Figure 3-4(A) demonstrates the ability of the spectral information to pick out the evergreen canopies of two white pine trees, which are clearly visible in the classification image.

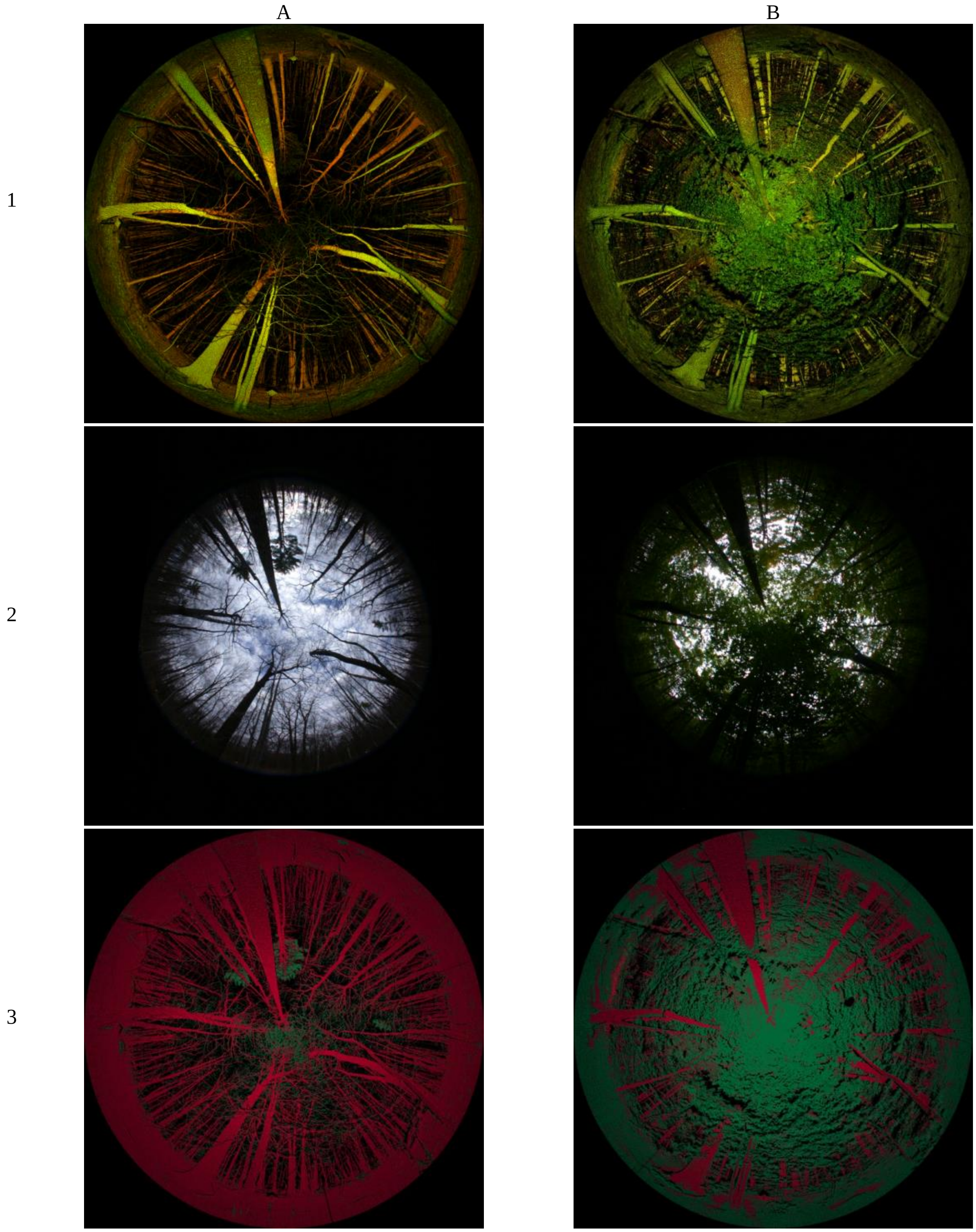
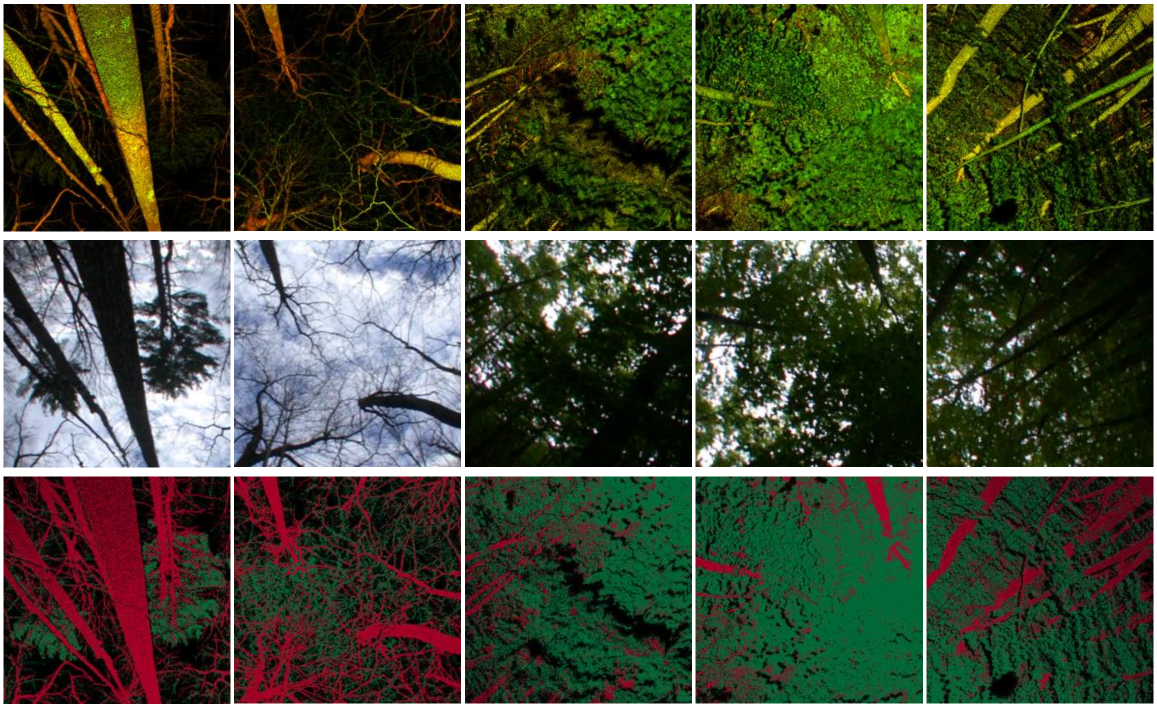


Figure 3-3. Center plot in leaf-off (left column) and leaf-on (right column) seasons. First row: hemispherical projection of color-composite apparent reflectance from dual-wavelength point clouds, red by NIR ρ_{app}, green by SWIR ρ_{app} and blue by dark constant. Second row: hemispherical photos registered to DWEL hemispherical projection. Third row: hemispherical projection of classification of point clouds.

Figure 3-4(B) shows fine branches with dark green colors in the leaf-off color-composite image. These dark-green branches appear at the far-range canopy top; they are most likely caused by laser beam misalignment, which reduces the SWIR return from their hits. The synergistic use of spectral and shape information classified some fine branches correctly, but left some fine branches labeled as leaves, especially in the area around the zenith point.

In the leaf-on image, Figure 3-4(C) shows a dark yellowish leafy branch quite close to the lidar arcing across the image from upper left to lower right. The yellow color of the leafy branch, which probably arises from near-field calibration variance, produced leaf points that were first mixed in clusters with woody materials by K-means spectral clustering. These points were then separated from woody materials by shape information and nearly all were labeled correctly in the classification. Above and to the right of the branch is a dark virtual shadow, caused by the reprojection of the lidar image to match the height of the camera. In this dark region, the hemispherical photo may record leaves or branches which are unseen by the lidar. However, these pixels of virtual shadow were not included in the accuracy assessment.

Figure 3-4(D) shows some branches cut into sections inside the canopy due to occlusion by leaves. They were identified by spectral information, but it was difficult to tell if they were all true branches in our accuracy assessment. Figure 3-4(E) shows some greenish small stems and branches, probably due higher water content than normal woody materials. These were mixed in clusters with leaves by K-means spectral clustering, but were then separated out by shape information. However, a few fragments and edges of these stems and branches were misclassified as leaves.



(A) leaf-off (B) leaf-off (C) leaf-on (D) leaf-on (E) leaf-on

Figure 3-4. Five zoom-in areas of the color-composite, classification and hemispherical photos of center plot shown in Figure 3-3.

Table 3-3 and Table 3-4 show the error matrices and accuracies (user's, producer's and overall) of the classifications of five scans together for leaf-off and leaf-on seasons respectively. The first two error matrices in each table are from the primary and secondary reference labels, giving the best and worst accuracy estimates respectively. The third error matrix in each table gives the average accuracies. The leaf-off overall classification accuracy, ranging from 0.70 ± 0.02 to 0.88 ± 0.01 with the average 0.79 ± 0.01, is slightly better than leaf-on, which ranges from 0.70 ± 0.02 to 0.81 ± 0.01 with the average 0.75 ± 0.02.

The leaf-off season classification shows both high user's and producer's accuracy for woody materials, but low values for leaves. The low user's accuracy suggests a large commission error in the leaf classification of leaf-off scans caused by misclassification of fine branches as leaves. The smaller omission error than commission error of leaf classification in leaf-off scans, suggested by the producer's accuracy being higher than the user's accuracy, implies that the classification of leaf-off scans overestimates the number of leaf points, or conversely underestimates the number of woody points. The main reason for misclassification of fine branches, particularly at far ranges at the canopy top, appears to be the misalignment of the two laser beams, even though the misalignment is smaller than scanning resolution.

Table 3-3. Error matrix in number of points, classification of five leaf-off point clouds.

Class	Wood	Leaf	Total	User's[1]	Producer's[1]	Overall[1]
Primary reference label: best possible accuracy						
Wood	428	20	448	0.96 ± 0.01	0.88 ± 0.01	0.88 ± 0.01
Leaf	57	109	166	0.66 ± 0.04	0.85 ± 0.06	
Total	485	129	614			
Secondary reference label: worst possible accuracy						
Wood	385	63	448	0.86 ± 0.02	0.76 ± 0.01	0.70 ± 0.02
Leaf	120	46	166	0.28 ± 0.04	0.42 ± 0.05	
Total	506	108	614			
Average error matrix from the primary and secondary						
Wood	407	41	448	0.91 ± 0.01	0.82 ± 0.01	0.79 ± 0.01
Leaf	89	77	166	0.47 ± 0.04	0.65 ± 0.06	
Total	495	119	614			

[1] Standard error of classification accuracy was calculated according to (Olofsson et al., 2014) with the proportions of number of points in mapped leaves (0.120) and woody materials (0.880).

The leaf-on season classification shows similar user's accuracy for woody materials and leaves (around 0.75). But woody materials show much lower producer's accuracy than user's accuracy, i.e. larger omission error and smaller commission error, which implies the classification of leaf-on scans underestimates the number of woody points, or conversely overestimates the number of leaf points.

The fine branches misclassified as leaves in leaf-off scans (Figure 3-4(B)) were rare in leaf-on scans because they were mostly occluded by leaves. However, the classification of leaf-off scans did not see as many parts of green stems and branches misclassified into leaves as seen by leaf-on scans (Figure 3-4(E)). There are two possible reasons for the change of classification of those stems and branches. First, higher water content in small stems and branches and moister lichens on trunk surfaces in leaf-on season than leaf-off season may lead to more misclassification of these woody materials by spectral information. Measurements of spectra of the same trunk/stem spots in leaf-off and leaf-on seasons are needed to confirm such changes in water content. Second, stems and branches are less obscured in the leaf-off scans and spatial context information may correct more of misclassified stems and branches in that point cloud.

Table 3-4. Error matrix in number of points, classification of five leaf-on point clouds.

Class	Wood	Leaf	Total	User's[1]	Producer's[1]	Overall[1]
Primary reference label: best possible accuracy						
Wood	91	29	120	0.76 ± 0.04	0.38 ± 0.02	0.81 ± 0.01
Leaf	147	656	803	0.82 ± 0.01	0.96 ± 0.01	
Total	238	685	923			
Secondary reference label: worst possible accuracy						
Wood	85	35	120	0.71 ± 0.04	0.26 ± 0.02	0.70 ± 0.02
Leaf	245	558	803	0.70 ± 0.02	0.94 ± 0.02	
Total	330	593	923			
Average error matrix from the primary and secondary						
Wood	88	32	120	0.73 ± 0.04	0.31 ± 0.02	0.75 ± 0.02
Leaf	196	607	803	0.76 ± 0.02	0.95 ± 0.01	
Total	284	639	923			

[1] Standard error of classification accuracy was calculated according to (Olofsson et al., 2014) with the proportions of number of points in mapped leaves (0.770) and woody materials (0.230).

3.5.2 Estimation of leaf, woody and plant area profiles

Figure 3-5 shows the equal-angle Andrieu projection images of the gap probability estimates at far range brought by leaves and woody materials separately in the center plot in leaf-off and leaf-on seasons. The separation between leaves and woody materials (also including ground and any non-leaf targets) is good by visual inspection. The change of gap probability brought about by leaves and woody materials from leaf-off season to leaf-on season is clear.

Table 3-5 provides values for LAIe, WAIe, and PAIe. PAIe values are 1.11 for the leaf-off season and 3.42 for the leaf-on season. As noted previously, PAIe values are from hemispherical photos and applied to LAIe and WAIe proportions. Leaf-on PAIe is close to the estimates by a previous study at the same site using hemispherical photos (3.46±0.46) (Zhao et al., 2011), There are no prior estimates of leaf-off PAIe, but estimates from both hemispherical photos and the TRAC instrument are consistent with each other.

The table also shows that band-averaged WAIe changed from 0.863 in the leaf-off season to 0.307 in the leaf-on season. As previously noted, leaf-on WAIe is an apparent value because branches and stems are occluded by leaves, while our estimate does not account for mutual occlusion. The difference in WAIe from leaf-off season cited above suggests that about 65 percent of the WAIe is occluded by leaves at our study site.

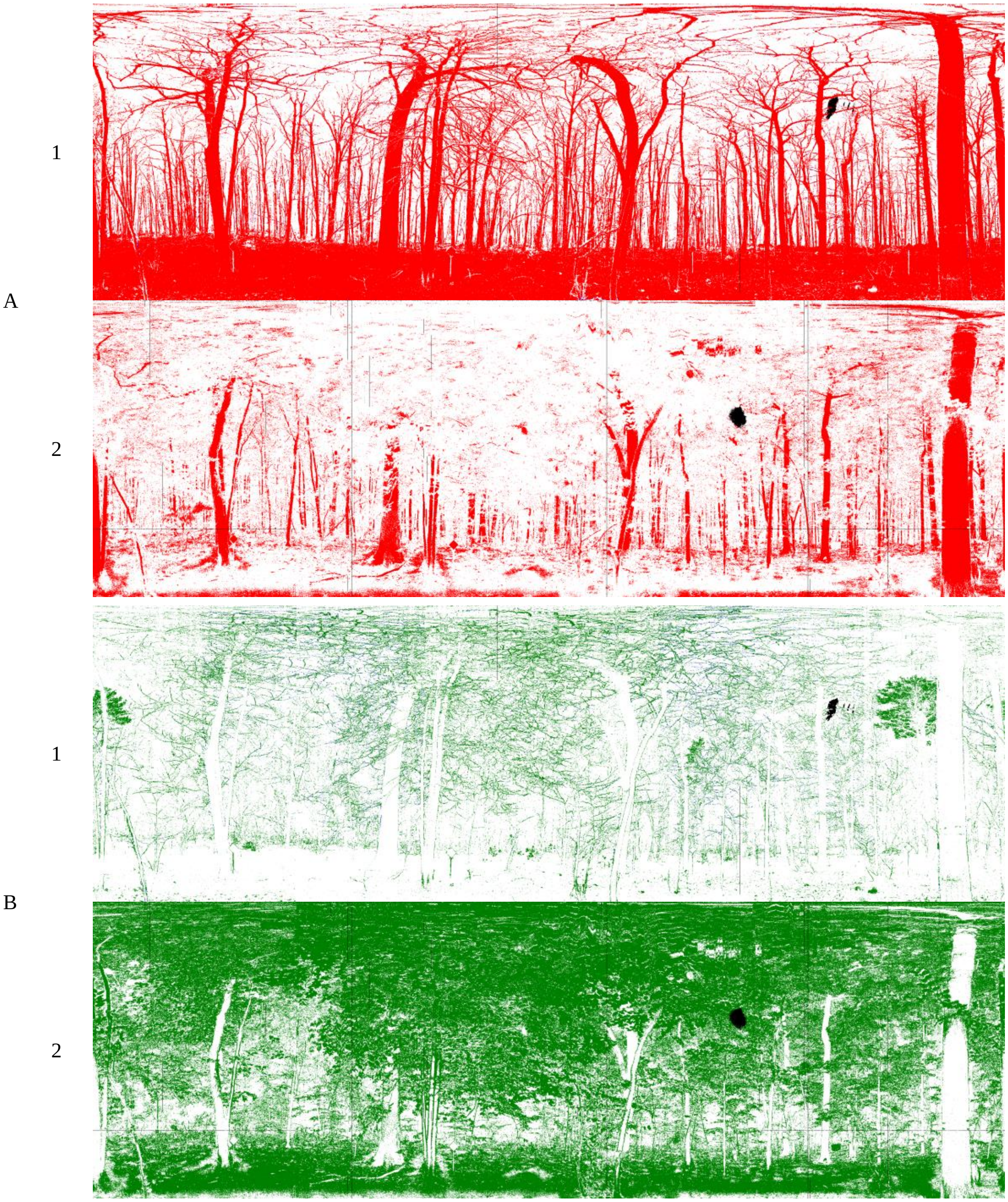


Figure 3-5. Gap probability at far range for woody materials (Part A) and leaves (Part B) separately from the scans at the center plot, in equal-angle Andrieu projection with X dimension as azimuth angles from 0° to 360° and Y dimension as zenith angles from 0° to 117°. Gap probability was calculated from NIR data. Row 1: Leaf-off season. Row 2: Leaf-on season. Red: $P_{gap}^{W} = 0$, totally attenuated by woody materials. Green: $P_{gap}^{L} = 0$, totally attenuated by leaves. White: $P_{gap}^{W} = 1$ or $P_{gap}^{L} = 1$, no attenuation by woody materials in pair A or leaves in pair B. Blue (scattered, sparse): partial gap by woody materials in pair A or leaves in pair B. Black: No data (largely solar aureole).

Table 3-5. LAIe, WAIe and PAIe from NIR and SWIR data in leaf-off and leaf-on seasons.

Vegetation component	NIR	SWIR	Band average
Leaf-off season			
LAIe	0.249	0.245	0.247
WAIe	0.861	0.865	0.863
PAIe	1.11	1.11	1.11
Leaf-on season			
LAIe	3.111	3.115	3.113
WAIe	0.309	0.305	0.307
PAIe	3.42	3.42	3.42

The leaf, woody, and plant area profiles from NIR and SWIR data for leaf-off and leaf-on seasons are shown in Figure 3-6. For the leaf-off season, cumulative effective vegetation index area curves increase quite smoothly to about 20 m, where the canopy begins to thin as dominants emerge at the canopy top. The vegetation area volume density curves are also more-or-less uniform with height; the peaks between 5 and 10 m are largely due to some close large trunks at the north and south scan points. In the leaf-on season graphs, the cumulative VAIe profiles show more variable structure and gently curve, indicating more variation with height and a slow decrease in the progressive accumulation of vegetation area. The woody area volume density curve is relatively constant with height, while the leaf area volume density peaks at about 5–7 m due to close leafy vegetation at the north site. With some variance, leaf and plant area volume density remain high through about 15 m, where LAVD and PAVD decreases steadily to the top of the canopy, at about 25 m. Some of this decrease may be due to weak or lacking returns at higher zenith angles produced by insufficient laser power.

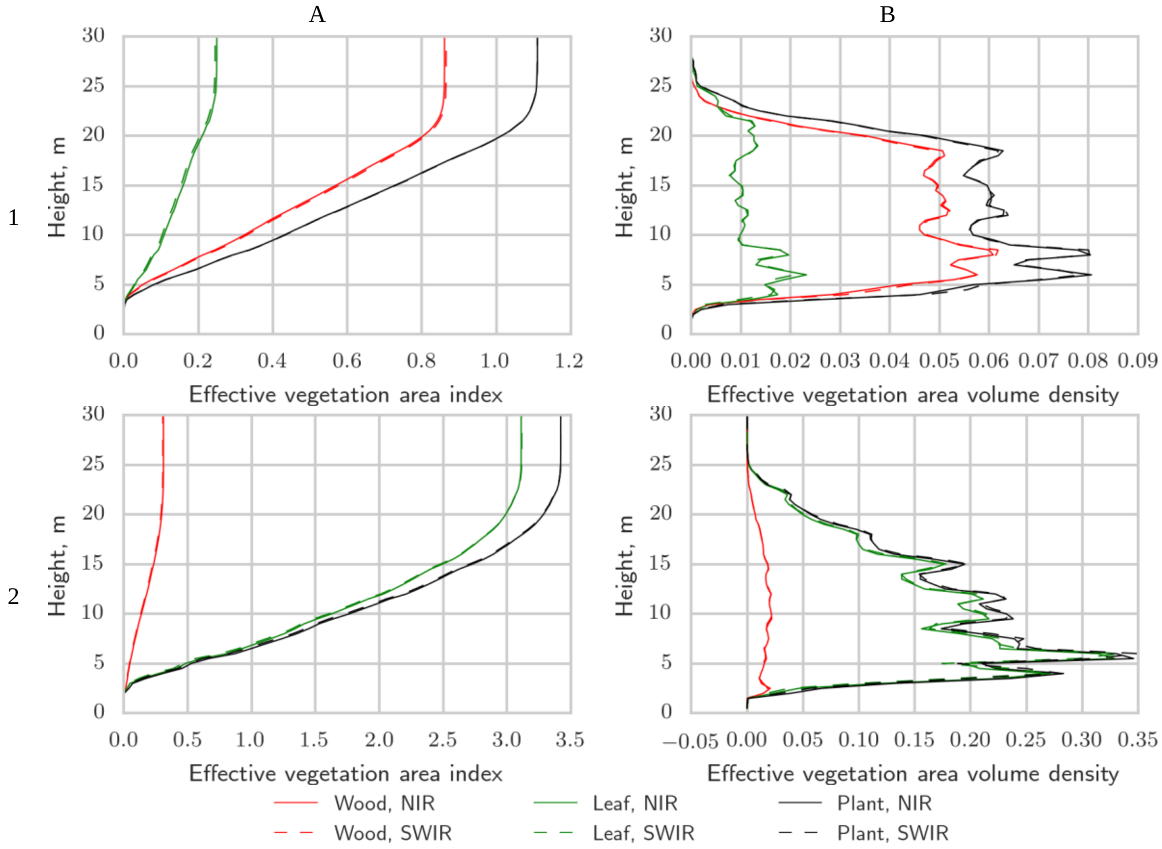
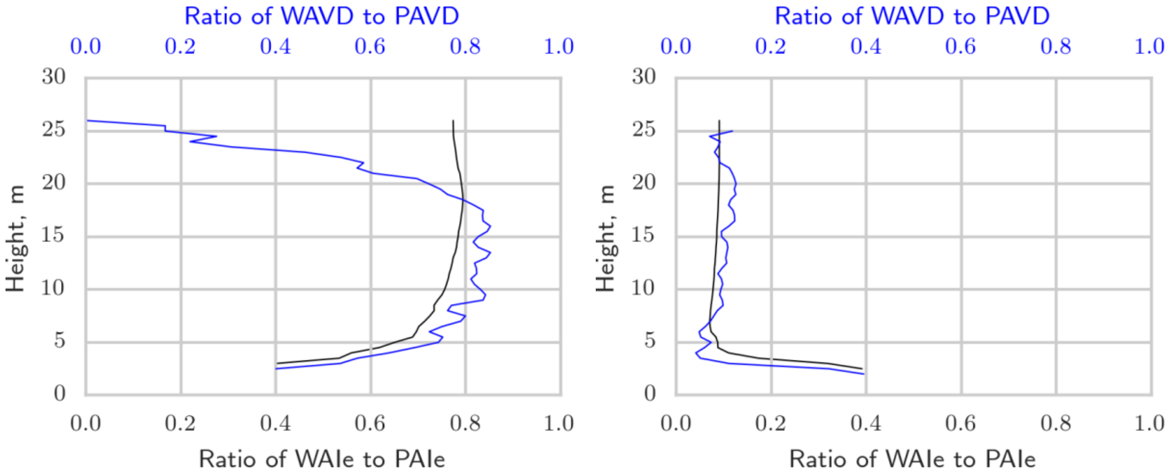


Figure 3-6. Vertical profiles of vegetation area index (cumulative VAVD, Column A) and vegetation area volume density (VAVD, Column B) averaged from five scans. Row 1: Leaf-off season. Row 2: Leaf-on season. NIR and SWIR curves are very close and are not visibly separated in many lines.

Table 3-5 and Figure 3-6 also show that NIR and SWIR curves are very close. Accordingly, we will present and interpret in the following only NIR data for simplicity; Appendix A.5 presents corresponding SWIR graphs and tables.

The woody-to-total ratios along canopy heights from VAVD and VAI show variation in both leaf-off (Figure 3-7(A)) and leaf-on conditions (Figure 3-7(B)). In the leaf-off condition, woody-to-total ratios (WAVD to PAVD) are stable around 0.8 from 8 m to 18 m with some variation. Above 18 m, the woody-to-total ratio of WAVD to PAVD decreases with canopy height. This is likely to be an artificial decrease due to misclassification of fine branches at far ranges in the high canopy either by laser misalignment or by their volume-like spatial pattern similar to leaf clusters. In the leaf-on condition, the woody-to-total ratio stays stable around 0.1 above 5 m with some variation.



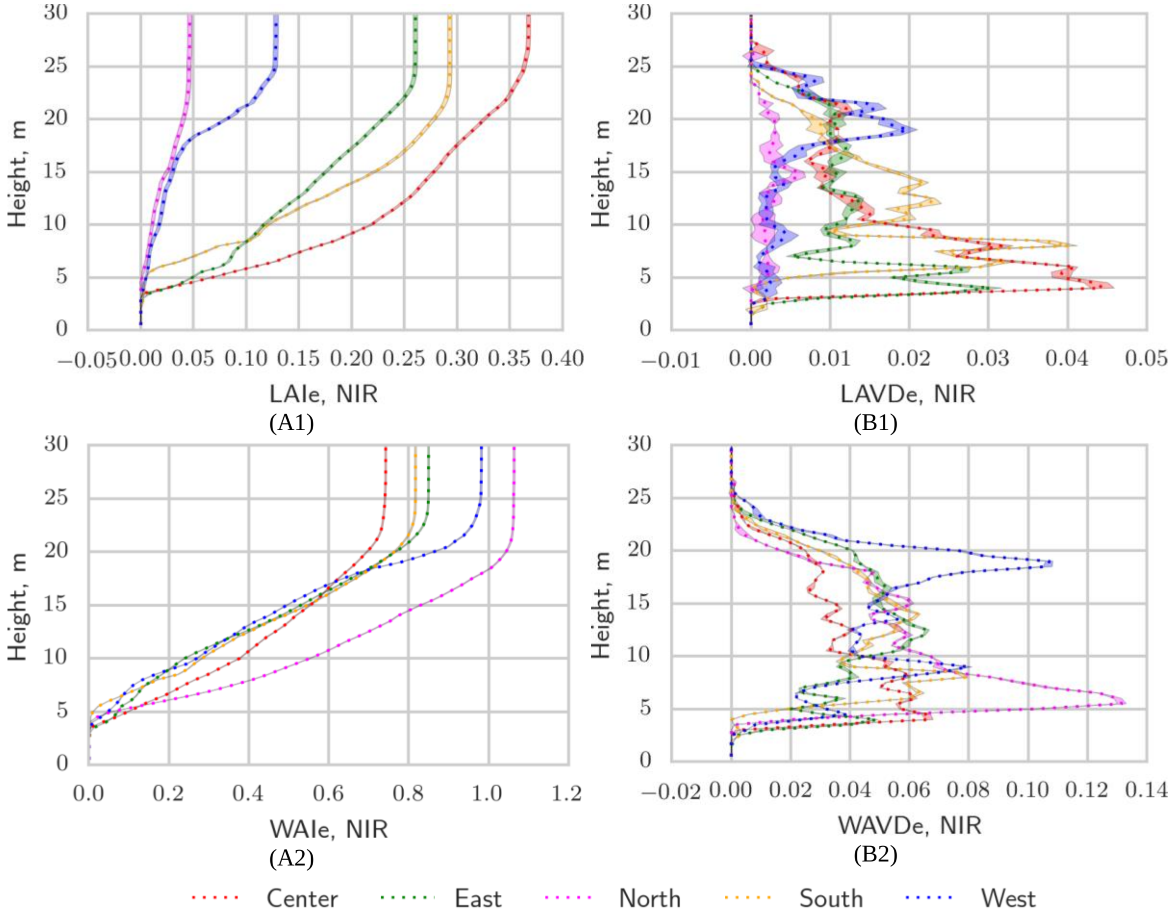
(A) (B)

Figure 3-7. Woody-to-total ratio profile along canopy heights. (A) ratios from leaf-off scans; (B) ratios from leaf-on scans. Ratios from PAIe or PAVD smaller than 0.001 are removed to avoid extremely large values due to numerical errors in the ratio calculation.

At heights below 5 m, the initial increase in woody-to-total ratios in the leaf-off scans and decrease in the leaf-on scans are effects of the field procedure for locating scan points. To avoid excessive occlusion from large trunks or shrub branches, the exact instrument position is moved away from such obstacles before scanning. This creates a bias in observations in which leaf points are reduced around and above the instrument. The profiles are thus obtained from points within an upward-opening cone of 35° half-angle from which the 0–10° zenith ring is omitted. Just above the instrument, the small volume of this cone changes little with season and the few initial points there remain in the same proportion of wood, about 0.4. In the leaf-off season, the cone volume rapidly acquires additional woody hits with height, while in the leaf-on season, additional leaf hits are acquired. By about 5 m height, this initial bias disappears.

Resampling of classifications (see Section 3.4.2.3) shows that the variance in vegetation profiles at each scan location (within-scan variances) due to classification error is very small for both leaf-off scans (Figure 3-8) and leaf-on scans (Figure 3-9). In contrast, the variance in vegetation profiles between scan locations (between-scan variances) is much larger (Figure 3-10), which suggests that the heterogeneity of canopy structure across the site dominates the variance in estimating vegetation profiles of the site. Because we used a single averaged PAIe that was split into LAIe and WAIe for the five scans, the variance in vegetation profiles does not include the heterogeneity in the total leaf, woody and plant area at different locations across the site. Instead, the variance describes the heterogeneity in the proportions of leaves and woody materials and the distributions of leaf, woody, and plant area with canopy height across the site. More scan locations can help improve the estimates of separate leaf and woody area profiles and better characterize the heterogeneity of leaf-versus-wood proportions and distribution of leaves and woody materials along canopy heights across the site.

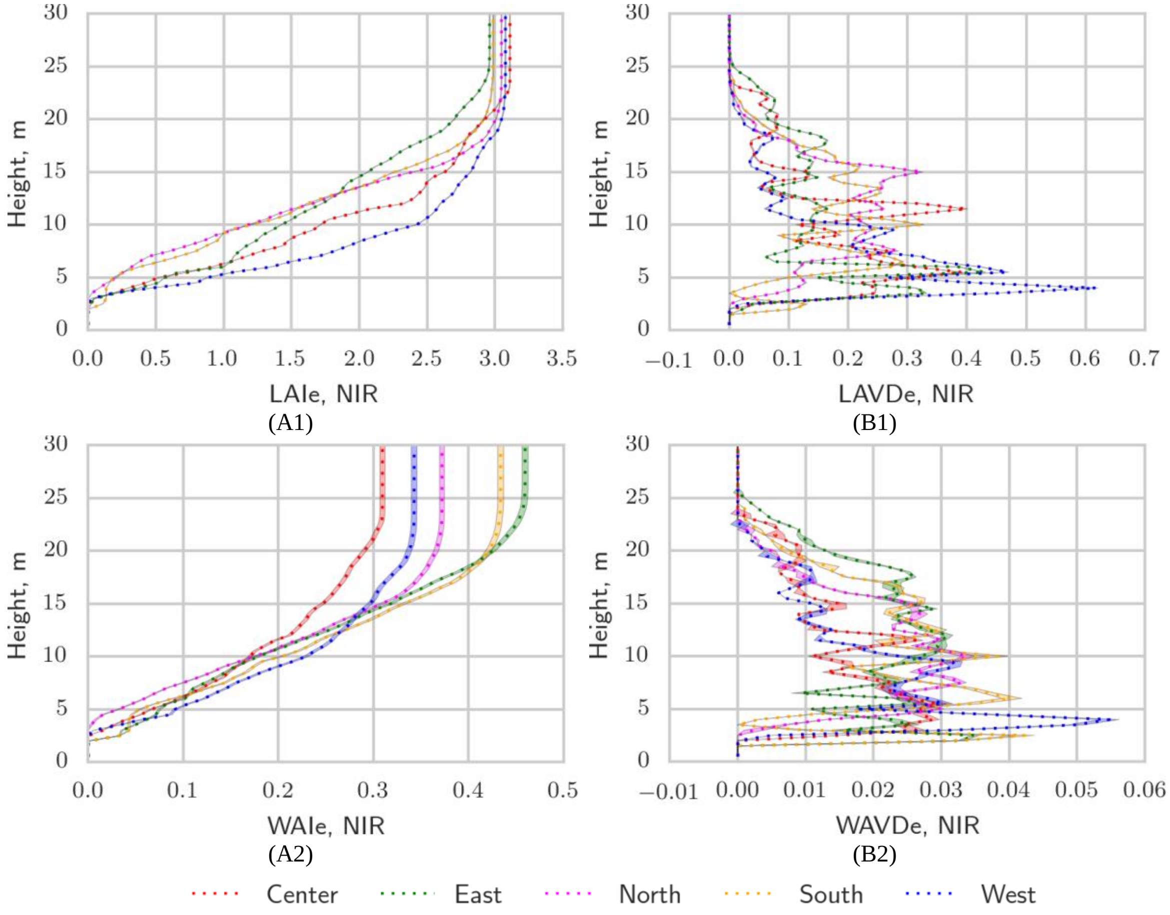
To further describe the variance in vegetation profiles, we normalized the between-scan variances in the profiles by the corresponding average vegetation area indexes (LAIe, WAIe and PAIe) (Figure 3-11). The quartiles and estimated probability density (by kernel density estimation) of the relative between-scan variances in vegetation profiles in the leaf-off season (Figure 3-11 (A1) & (B1)) show larger relative variances in leaf area profiles than woody area profiles and plant area profiles. In contrast, in the leaf-on season (Figure 3-11 (A2) & (B2)), the relative variances in leaf, woody and plant area profiles have similar shapes.



(A1) (B1)

(A2) (B2)

Figure 3-8. Variance in profiles due to classification error from each of the five leaf-off scans. Shaded area shows three standard deviations; for LAIe and WAIe and some of the LAVDe and WAVDe curves, shaded area is indistinguishable from graphic line width.



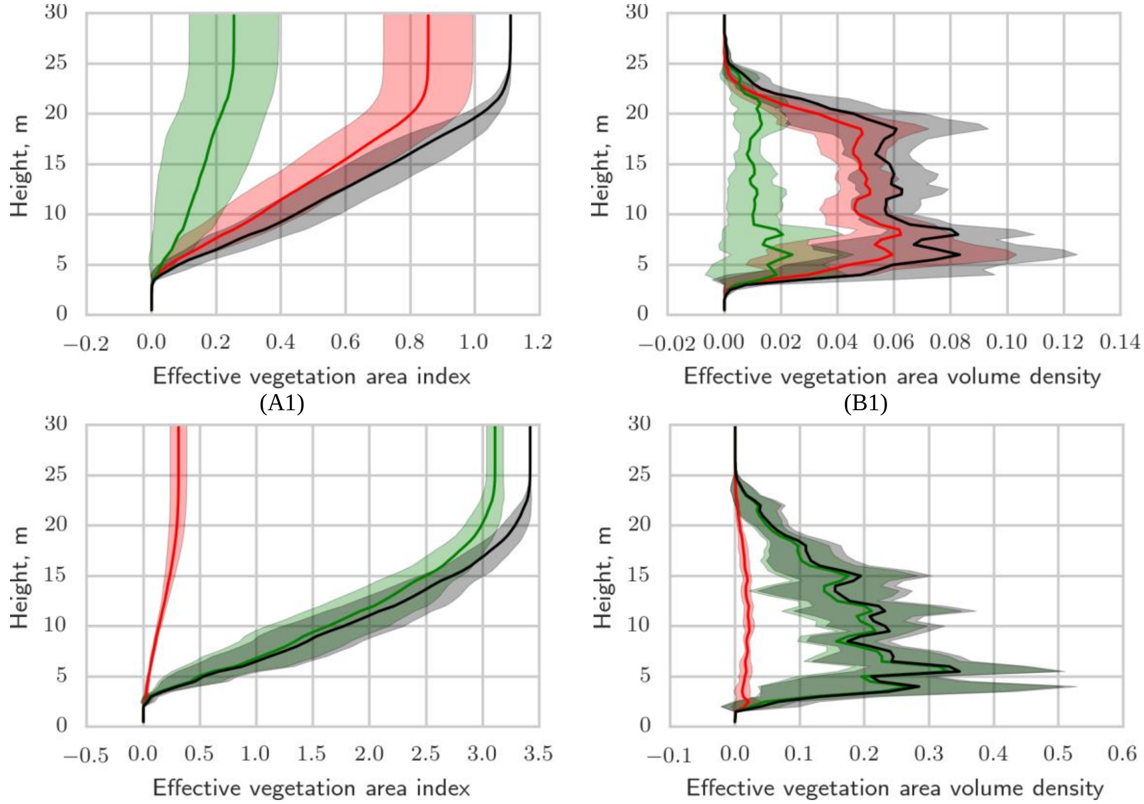
(A1)

(B1)

(A2)

(B2)

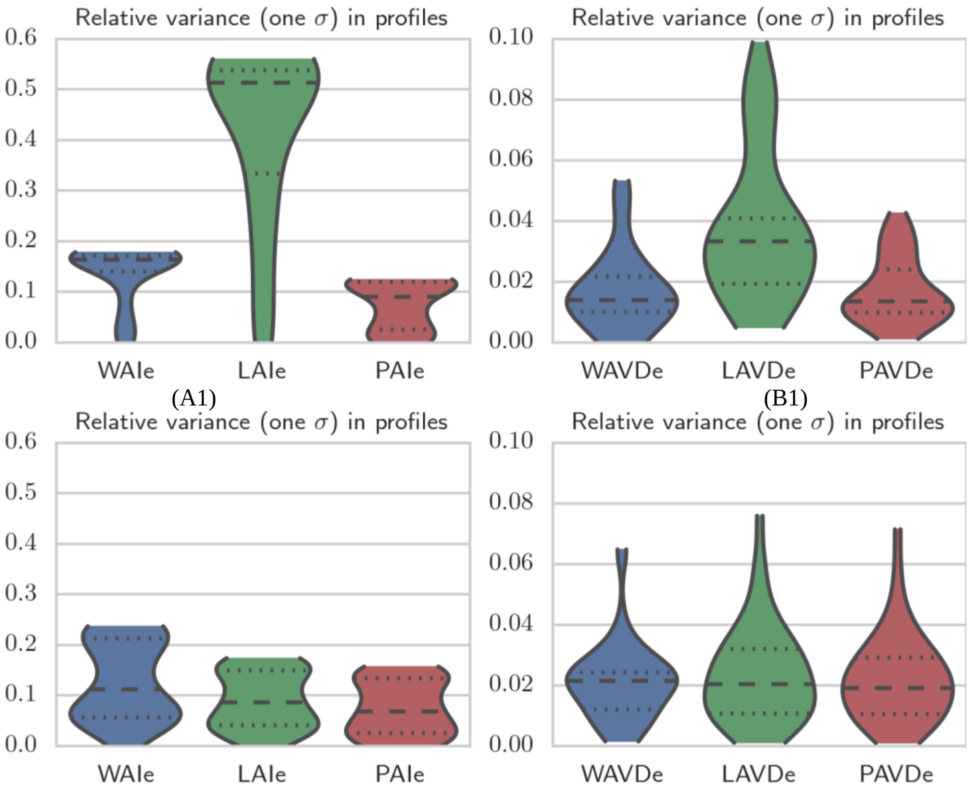
Figure 3-9. Variance in profiles due to classification error from each of the five leaf-on scans. Shaded area shows three standard deviations; for LAIe and WAIe and some of the LAVDe and WAVDe curves, shaded area is indistinguishable from graphic line width.



(A1) (B1)

(A2) (B2)

Figure 3-10. Variance in profiles due to heterogeneity of canopy structure across the study site, from NIR data. First row: leaf-off season. Second row: leaf-on season. Shaded area shows one standard deviation.



(A1) (B1)

(A2) (B2)

Figure 3-11. Violin box plots of variance (one standard deviation) in vegetation profiles due to heterogeneity of the study site, normalized by average vegetation area indexes. First row: leaf-off season. Second row: leaf-on season. Left column: cumulative VAVD (VAI up to a canopy height). Right column: VAVD. The dash line is the median. The dotted lines are 25 and 75 percentiles. The ends of the violin boxes show the minimum and maximum values. The widths of violin boxes show the probability density of the data from kernel density estimate.

3.6 Discussion

3.6.1 Point classification

The synergistic use of spectral and spatial information yields good overall accuracy of point classification of leaves and woody materials in 3-D space (Table 3-3 and Table 3-4), demonstrating the value of integrating spectral information with 3-D spatial locations through spectral lidar instruments like the DWEL. While the spectral contrasts at the NIR and SWIR wavelengths given by NDI for leaves and woody materials are theoretically distinctive, using NDI alone is not sufficient to discriminate leaves and woody materials in point clouds in practice. The spectral reflectance of individual leaves and woody materials at the two wavelengths may show less distinct difference or even similarity due to complex variation in the properties of leaves and woody materials. For example, higher water content in small stems or lichen spots on stems could result in atypical woody materials with the dual-wavelength spectral reflectance values similar to typical leaves (Figure 3-4(E), first row). Conversely, lower water content in unhealthy or senescent leaves could results in the spectral reflectance values similar to typical woody materials. In addition, the difference in the NDI between leaves and woody materials may be further reduced by uncertainties in terrestrial laser scanning data, such as radiometric calibration error and laser alignment error (Figure 3-4, (B) & (C), first row).

Given such variance, we turned to the unsupervised K-Means classifier to cluster the points in spectral measurement space and thus separate those points that were clearly the result of leaf hits and woody hits from points that were less spectrally definitive. Clusters were labeled by manual interpretation using color-composite images and the geometric shapes of point clusters, with clusters of mixed leaf and woody points then further processed using the multiscale dimensionality algorithm discussed in Section 3.4.1.2. A disadvantage to unsupervised classification is that it must be repeated for each scan, and thus is not practical for large numbers of scans. Supervised classification provides an alternative that could be explored for future work, but still requires identification of training data that could vary significantly from scan to scan. An intermediate method might be to label unsupervised clusters automatically based on spectral means and variances to identify homogeneous clusters of leaf and woody points and then pool the points in remaining clusters for input to the spatial classification step.

The spatial context information given by the 3-D spatial distribution pattern of each point and its neighbors provided a needed second step to classify the points of mixed clusters as identified by the K-Means classifier. Points of stems and large branches have distinct spatial distribution pattern from points of leaf clusters, which resulted in good classification using spatial context information. However, points of fine branches surrounded by leaves inside canopies have a spatial distribution pattern similar to leaf points (Figure 3-4(B)). Using spatial context information can thus label fine branch points as leaves. In future work, these fine branches need to be identified through their spectral information rather than the spatial context information after the laser beam alignment is improved. It should also be noted that the computational cost of calculating multiscale dimensionality for spatial context information is quite high for the whole point cloud; reducing the number of points to be examined by prior spectral classification significantly reduces the computational time and cost.

3.6.2 Accuracy assessment

The accuracy assessment of point classification in 3-D space is challenging without high accuracy reference data. Our indirect accuracy assessment approach using registered hemispherical photos suggests a range of reasonably good classification accuracy using spectral and spatial information. The misclassification of many fine branches, particularly at far ranges at the canopy top (Figure 3-4(B)) highlights the importance of laser alignment accuracy for spectral lidar instrument built with multiple lasers. Our indirect approach can be limited by the registration error between hemispherical photos and the projection images of lidar point clouds. In future work, artificial targets on the trees, e.g. reflective crosses, visible in both hemispherical photos and the DWEL scans, may help the registration and thus reduce error in the classification accuracy assessment.

Collecting “real truth” data of classification in 3-D for individual points is quite difficult, if not impossible. An alternative is simulation of lidar scanning data and point clouds from tree models to test classification algorithms, as the “truth” is known from the tree models (Burt et al., 2013; Disney et al., 2006). However, the simulation needs to incorporate the complexity and the error sources of the actual lidar scanning to assess the classification accuracy more realistically. For example, wet lichen spots could be placed on stems of tree models to change the woody surface reflectance. Some small stems could be assigned lower reflectance values at the SWIR wavelength. Leaf and woody reflectance could be varied randomly according to their reflectance variance from spectral measurements, and effects of leaf angle distribution models could be assessed. Two-laser alignment error could also be included in the simulated lidar data, allowing the sensitivity of classification accuracy to laser alignment error to be estimated.

3.6.3 Foliage profiles

The separate profiles of leaf and woody area from leaf-off and leaf-on scans show the seasonal change of the amount of leaves and woody materials visible from under canopy. In the leaf-off season, the distribution of leaf and woody area volume density with height in the middle canopy between 10 m to 20 m (Figure 3-6 (B1)) was relatively constant. In the leaf-on season, the leaf area profile shows a peak around 6 m and a more or less steadily decreasing trend, while woody area remains nearly constant (Figure 3-6 A2).

Our leaf-on leaf area profile is different from that derived by Zhao et al. (2011) using the Echidna Validation Instrument (EVI) at the same site in 2007; they observed a smoother profile that peaked around 21 m and had a slightly but steadily increasing trend of leaf area between 10 m to 20 m, rather than a steady decrease. This difference occurred even though our profile estimation method was similar to that used in the EVI scanning. The main reason for the different leaf profiles from the current DWEL data in 2014 and the heritage EVI data in 2007 is the limited measurement range of the DWEL. Only pulses at zenith angles of 35° or less were able to exit the 25-m canopy to provide a proper estimate of gap probability with height. As a result, we used only zenith rings between 10° to 35°, a much smaller zenith angle range than that used by Zhao et al. (2011) for EVI data. Even with this zenith range reduction, it appears that weak and partial hits, accentuated by misalignment errors, are being lost with distance, producing the gradual reduction in both leaf and woody area volume density noted in Figure 3-6 (B2). The smoother profile of Zhao et al. is probably due to scanning a larger volume of the canopy (5–60° zenith angle) and thus averaging more points to reduce variance.

This comparison shows the importance of an adequate measurement range and signal-to-noise ratio that is capable of measuring gap probability through the top of the canopy to zenith angles of 60° or greater. Considering the low reflectance of leaves at the

SWIR wavelength, this requirement is significantly more challenging for the DWEL than for an NIR instrument like the EVI.

In retrieving foliage profiles, variation within the same stand from one location to the next can cause significant variance in profiles (Figure 3-8, Figure 3-9, and Figure 3-10). The variance is probably higher than expected for more powerful scanners like the EVI, in which the volume of canopy effectively sampled by each scan is much larger. Given the present capabilities of the DWEL, more scans within a stand are likely to be needed to add precision to estimates of leaf and woody areas and foliage profiles.

The change from the leaf-off season to the leaf-on season in the spatial heterogeneity of the distribution of leaves and woody materials with height is captured by the different variances in the leaf and woody profiles (Figure 3-11). Higher variance in the leaf-off leaf area profile than in the woody area profile (B1) reflects the spatial variance of leaves due to the chance occurrence of evergreen needleleaf canopies across the forest site. Relatively similar variance in the leaf and woody area profiles in the leaf-on season (B2) suggests a similar distribution pattern of leaves and woody materials with canopy height across the site.

While the classification errors cause small variances in leaf and woody area profiles, these variances can be further removed using the error-adjusting method in Olofsson et al. (2013). However the error-adjusting method was designed to eliminate bias attributable to map classification error on a 2-D surface. Further changes are needed to extend this error-adjusting method from 2-D surface to 3-D space in a future study.

3.6.4 Woody-to-total ratios and occlusion

The height profiles of woody-to-total ratios suggest a generally stable value of 0.7 – 0.8 for the leaf-off season and about 0.1 for the leaf-on season above about 5 – 8 m (Figure 3-7), indicating a constant ratio could be used to remove the woody contribution to vegetation profiles above a certain canopy height (5 – 8 m at this site). At lower canopy heights, woody-to-total ratio is more variable, and the removal of the woody contribution may depend on the scan location. For the profiles of understory leaf area at low canopy heights, it would be better to estimate the ratio from the direct discrimination of leaf and woody areas from the DWEL scanning data.

The comparison of WAIe from leaf-off and leaf-on scans found that about 65 percent of the WAIe is occluded by leaves in the leaf-on season at our study site. This occlusion ratio could be used as an empirical value to correct the apparent WAIe value for occlusion by leaves to give an actual WAIe for similar forest stands in leaf-on condition.

3.7 Conclusions

Explicit separation of leaves and woody materials directly in 3-D space of a forest improves the retrieval of LAI and vertical foliage profiles from gap probability measurements by removing the woody contribution without using a site-specific empirical woody-to-total ratio from destructive sampling or a typical value from the literature. The 3-D separation may also benefit ecophysiological and 3-D radiative transfer models by improving the simulation and inversion of biophysical and/or biochemical properties of a forest.

We obtained separate vertical profiles of leaf and woody areas through the 3-D classification of leaves and woody materials in dual-wavelength point clouds from the Dual-Wavelength Echidna Lidar (DWEL), a terrestrial laser scanner. The 3-D classification was created with an approach using both spectral information from NIR and SWIR apparent reflectance and spatial context information given by the 3-D spatial distribution pattern of each point and its near neighbors. The classification accuracy (in number of points) from an indirect assessment approach using registered hemispherical photos is 0.70 ± 0.02 to 0.88 ± 0.01 for leaf-off point clouds and 0.70 ± 0.02 to 0.81 ± 0.01 for leaf-on point clouds, demonstrating the value of combining spectral information with 3-D spatial information. The misclassification of fine branches, especially at far ranges, emphasizes the importance of high alignment accuracy of laser beams in multispectral lidar instruments built with multiple lasers, like the DWEL. This challenge to multispectral lidar instruments built with multiple lasers may be eased by using super-continuum (white) lasers but with extended laser measurement ranges (Hakala et al., 2012; Kaasalainen et al., 2007; Puttonen et al., 2015).

We obtained the gap probabilities by leaves and woody materials separately from classified point clouds and retrieved the separate vertical profiles of leaf and woody areas. The five-scan averaged woody-to-total ratios at different canopy heights are generally stable in the middle and upper canopy for the stand, but vary in the lower canopy as sampled by the instrument. The uncertainty in leaf and woody area profiles from classification errors was negligible based on a bootstrapping analysis. The variance in leaf and wood area profiles over this forest site was dominated by the heterogeneity of canopy vertical structure as sampled at the five scan locations. More scan locations over the site are needed for a more complete characterization of the leaf and woody area profiles. The variance in leaf area profiles with canopy height was relatively larger than that of the woody area profile in the leaf-off season because of the presence of a few large white pine evergreen canopies scattered throughout the site, while the variances were relatively similar in leaf and woody area profiles with canopy height in the leaf-on season.

Although the estimates of total leaf area index and woody area index were limited by the current restricted measurement range of the DWEL, due to insufficient laser power and signal-to-noise ratio, the relative proportions of leaf and woody areas were successfully retrieved with the dual-wavelength point clouds. An increase in laser power and reduction of electronic noise in the instrument should provide better dual-wavelength scans and better estimates of LAI, WAI and their vertical profiles.

CHAPTER 4. A NONDESTRUCTIVE APPROACH TO ABOVEGROUND BIOMASS ESTIMATION OF INDIVIDUAL FOREST TREES USING CLASSIFIED THREE-DIMENSIONAL DWEL POINT CLOUDS

4.1 Introduction

The amounts of carbon uptake, storage and release in terrestrial ecosystems are among the largest uncertainty sources in the global carbon balance under a changing climate (Denman et al., 2007). Large spatial and temporal variations in terrestrial carbon storage come from the biomass of terrestrial vegetation (biomass, as dry weight, is about 50% carbon) (Houghton et al., 2009). Vegetation biomass, particularly aboveground biomass, is exposed to natural and human-induced disturbances and determines the amount of carbon emission caused by disturbances to ecosystems (Houghton et al., 2009; Houghton, 2007). Moreover, biomass affects ecosystem functioning through its control over the magnitude and rate of autotrophic respiration (Houghton et al., 2009; Turner et al., 2005).

About 70% - 90% of terrestrial vegetation biomass is in forests, of which about 70% - 90% is in aboveground forest biomass (AGB) (Cairns et al., 1997; Houghton et al., 2009). But estimates of aboveground biomass density (AGB per unit area) over large areas are uncertain and estimates by different approaches are not consistent (Hill et al., 2013). Airborne and space-borne remote sensing data are useful in evaluating aboveground biomass because of their ability and potential to provide low cost, consistent, operational and wall-to-wall estimates at large scales (Gibbs et al., 2007). But no methods based on remote sensing data measure AGB across a landscape continuously.

Ground-based data are required to calibrate biomass-retrieval methods and validate AGB estimates at large scales. Improving the quality and detail of ground-based estimates of aboveground biomass is essential to reducing AGB estimate uncertainty and making consistent and comparable maps of aboveground biomass through time for AGB change detection.

The direct and probably the most accurate way to measure aboveground biomass is destructive sampling (Picard et al., 2012). However, this approach is extremely labor intensive and sometimes impractical. Another commonly used indirect approach to ground-based biomass measurement relates AGB of a tree with more easily measured quantities such as diameter at breast height (DBH), and/or tree height, using allometric equations derived by destructive sampling (Gibbs et al., 2007). Although extensive compilations of allometric equations have covered many important species across several continents and biomes (Chave et al., 2005; Jenkins et al., 2004; Zianis & Seura, 2005), allometric equations are still large sources of error in AGB estimation (Basuki et al., 2009; Chave et al., 2004; Keller et al., 2001; Wang, 2006), due to inadequacies imposed by the time-consuming, laborious, and costly development and update of allometric equations through destructive sampling. The errors in aboveground biomass estimates using allometric equations may be particularly large for large trees, as the destructive sampling of large trees is rare (Basuki et al., 2009; Chave et al., 2004; Keller et al., 2001) and large trees have large amounts of biomass.

A ground-based active remote sensing instrument, the terrestrial laser scanner (TLS) or terrestrial lidar, emits laser light, detects its return from scattering surfaces and records the time (or phase) and intensity of signals reflected back from objects at different distances along the laser transmission path. Terrestrial laser scans image forest trees in fine detail and scanners have shown great ability to estimate aboveground biomass in nondestructive way. However, most methods to estimate AGB by laser scanning still rely on allometric equations applied to stem diameters, tree heights, and stem count densities extracted from the laser scans (Strahler et al., 2008; Thies & Spiecker, 2004; Yao et al., 2011).

A recent TLS-based nondestructive approach to aboveground biomass estimation combines *a priori* wood density information with wood volume that is directly calculated from cylinder tree models built from TLS point clouds using Quantitative Structure Modeling (QSM) (Burt et al., 2013; Calders, Newnham, et al., 2014; Kaasalainen et al., 2014; Raumonen et al., 2013). This nondestructive approach is independent of allometric equations and has been validated against destructive sampling measurements in a eucalyptus forest, showing an overestimation error of 9.68% by QSM compared to an underestimation error of 36.57% - 29.85% by allometric equations (Calders, Newnham, et al., 2014). The key to accurate aboveground biomass estimates by this approach is correct tree models and thus accurate wood volume estimates, which depend on a proper selection of input parameters to the QSM algorithm. The selection of parameters for QSM is not trivial and requires tuning for different scanning resolutions and branch sizes to find a proper and objective criterion. Moreover, previous QSM trials on both simulated and real TLS point clouds have suggested that leaf points are an important error source in modeling trunk and branch structure for wood volume estimates (Burt et al., 2013) and called for the removal of leaf points to improve the accuracy of structure modeling.

In this study, we scanned a midlatitude forest site with deciduous trees in both leaf-off and leaf-on condition using the Dual-Wavelength Echidna Lidar (DWEL), a dual-wavelength terrestrial lidar that acquires full-waveform scans at both near-infrared (NIR, 1064 nm) and shortwave infrared (SWIR, 1548 nm) wavelengths with simultaneous laser pulses (Douglas et al., 2015). The dual-wavelength scanning was designed to distinguish leaves from woody materials and has shown its ability to classify leafy and woody points in 3-D space in the Chapter 3. This study applies quantitative structure modeling to the classified DWEL point clouds to answer the following questions pertaining to accurate aboveground biomass estimates using QSM:

(1) Is the standard deviation of wood volume estimates by multiple QSM runs a good criterion for parameter selection? Which parameter settings seem to work best with DWEL point clouds of 2-mrad angular resolution?

(2) Does the removal of leaf points from the point cloud of a tree improve the tree modeling and wood volume estimates? If so, how much will the estimate be improved? Are the improvements consistent between leaf-off and leaf-on scans?

4.2 Study Area and Data

For this study, we established a 100-m-by-100-m deciduous forest site (N 42° 31′ 51.48″, W 72° 10′ 55.56″) at Harvard Forest in central Massachusetts, USA. This generally flat 1-ha site is dominated by red maple (*Acer rubrum*), red oak (*Quercus* *rubra*) and white birch (*Betula papyrifera*), with an understory of these species accompanied by American beech (*Fagus grandifolia*), American chestnut (*Castanea dentata*) and others. A few conifers are also present, including several large white pines (*Pinus strubus*) and a few smaller eastern hemlocks (*Tsuga canadensis*). At five circular plots of 20-m radius (Figure 4-1), we collected biometric data including diameter at breast height (DBH), species, location (range and compass azimuth from the circular plot center), and crown position. For a subsample of trees, we also acquired tree heights, crown diameters at two orthogonal dimensions, and crown heights. In data of September 2014, the average stem density at this site was 769 trees ha^{-1} and the basal area was around 38.5 m^2 ha^{-1}. The average tree height of sampled trees at this site was 20.3 m and the average crown diameter was 8.7 m.

We scanned the forest site with the Dual-Wavelength Echidna Lidar (DWEL) at five scan locations (Figure 4-1) in both leaf-off (May 3rd 2014) and leaf-on (Sept 19th 2014) conditions. The wind speed on both scanning days was low with occasional light leaf and branch motion at the canopy top. At the time of the leaf-off scan, tree leaf and flower buds were breaking, but little leaf area was visible except for conifers. The DWEL acquires full-waveform scans with two coaxial lasers at NIR (1064 nm) and SWIR (1548 nm) wavelengths. It collects return waveforms at the pulse repetition rate (PRF) of 2 KHz. The scanning resolution was set 2 mrad with slightly larger beam divergence, 2.5 mrad to ensure continuous coverage of the hemispheres for the scans used in this study.

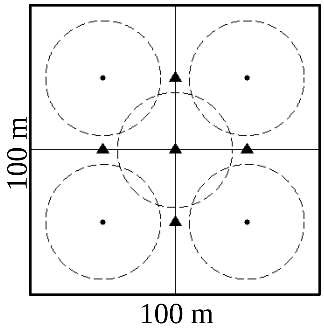


Figure 4-1. DWEL scans were collected at the five solid triangles. Tree measurements were collected at circular plots centered at the four dots and the center triangle. The dashed circles represent the coverage of each tree measurement plot.

4.3 Methodology

4.3.1 Preprocessing and classification of TLS scanning data

We generated radiometrically calibrated dual-wavelength point clouds from DWEL waveform data at the two wavelengths through a chain of preprocessing steps including (1) background noise removal; (2) conversion of digitizer time to apparent range by aligning each waveform to the peak of the outgoing pulse; (3) detection and correction of saturated return pulses; (4) correction of laser power drift; (5) detection of return pulse peaks; (6) calibration of peak intensity from digital counts to apparent reflectance (ρ_{app}); and (7) assignment of dual-wavelength intensity and apparent reflectance to each point by merging two point clouds at the two wavelengths.

We used both spectral information from NIR and SWIR apparent reflectance and spatial context information given by the 3-D spatial distribution pattern of each point and its neighbors to classify points into leaves and woody materials. First, an unsupervised classifier, the K-means clustering algorithm implemented in Scikit-learn package (Pedregosa et al., 2011), divided all points in one scan into 100 clusters using three variables: apparent reflectance values at the two wavelengths and a normalized difference index (NDI) value given by equation 4-1. Each cluster was inspected by visual inspection of the 3-D point cloud and labeled as being dominated by leaf or woody hits.

Where clusters included both leafy and woody hits, a supervised classification method based on multiscale dimensionality criteria (Brodu & Lague, 2012) labeled each cluster point into a canopy point (taken as a leaf hit) or a trunks/coarse branches/ground hit using shape information of points. The shape information of a point, given by the point cloud multiscale dimensionality, characterizes the local 3-D organization of the point cloud within spheres of different diameters (scales) centered on the measured point and varies from being 1-D (points set along a line), 2-D (points set along a plane) to the full 3-D volume (Brodu & Lague, 2012). Points of trunks/branches or ground shape lines or planes while points of canopies shape 3-D volumes.

$$NDI = \frac{\rho_{app}^{nir} - \rho_{app}^{swir}}{\rho_{app}^{nir} + \rho_{app}^{swir}} \tag{4-1}$$

4.3.2 Marker-free registration of multiple scans

The merging of the five leaf-off/leaf-on scans using a marker-free registration approach generated full 360° views of trees in point clouds, which are required by the quantitative structure models for AGB estimates from TLS scanning data (see next section 4.3.3). This semi-automatic registration approach does not require reflective marker targets, which are often deployed to provide common registration points within multiple scans (Béland et al., 2011; Blanchette et al., 2015; Pueschel, 2013). Although we deployed eight reflective spheres on the site between scan locations, not enough targets were captured by or clearly identifiable from multiple scans for registration, due to the high density of occluding trees and understory at our study site, coupled with low laser power and low signal-to-noise ratios. To solve this problem, we developed a marker-free, semi-automatic registration approach. It is a pair-wise registration, i.e. all the scans are aligned to the center scan to merge together, and has two steps: (1) coarse registration with trunk centers as homologous markers; and (2) fine registration using an iterative closest point (ICP) algorithm (Besl & McKay, 1992).

For coarse registration, trunk centers extracted from single-scan point clouds were manually matched and aligned on the XY plane for translation along X- and Y- axis and rotation around the Z-axis. An approximate translation along Z-axis was further manually performed by aligning ground points if necessary. The extraction of trunk centers at breast height above ground used an automatic detection method by Huang et al. (2011) with some changes to adapt to single-scan point clouds (Figure 4-2). To first locate ground points, the variable scale and threshold filtering algorithm in this method grids points on the XY plane and finds potential ground points with a threshold of distance between points and the lowest height in a grid cell. It repeats this procedure with decreasing grid size and decreasing distance thresholds to extract final ground points. These points are then connected in a triangulated irregular network to fashion a digital elevation model (DEM). To find trunk centers, a slice of points from a trunk at breast height is modeled as an arc from a circle. Circle detection uses a circular Hough transform (Ioannou et al., 1999) and removes noise points. Circle fitting uses a simple least squares algorithm.

However, in single-scan point clouds, occlusion will increase with range. As a result, both ground and trunk points were often sparse in the overlapping area between adjacent scans, where homologous trunks area sought. To increase the number of extracted homologous trunk centers for coarse registration, we added two steps. First, the digital elevation model was extrapolated to a larger area (bounding box of points on XY plane) with a plane fitted to the DEM by the random sample consensus (RANSAC) algorithm (Fischler & Bolles, 1981), allowing identification of more trunk points at breast height. Second, if circle detection and fitting failed in a point cluster, a simple average of point coordinates the point clusters was generated as a possible extra trunk center.

For a fine registration after each pair of point clouds were approximately aligned together, we first extracted points in the general overlapping area and then performed a fine registration using the ICP function in CloudCompare (Girardeau-Montaut, 2015).

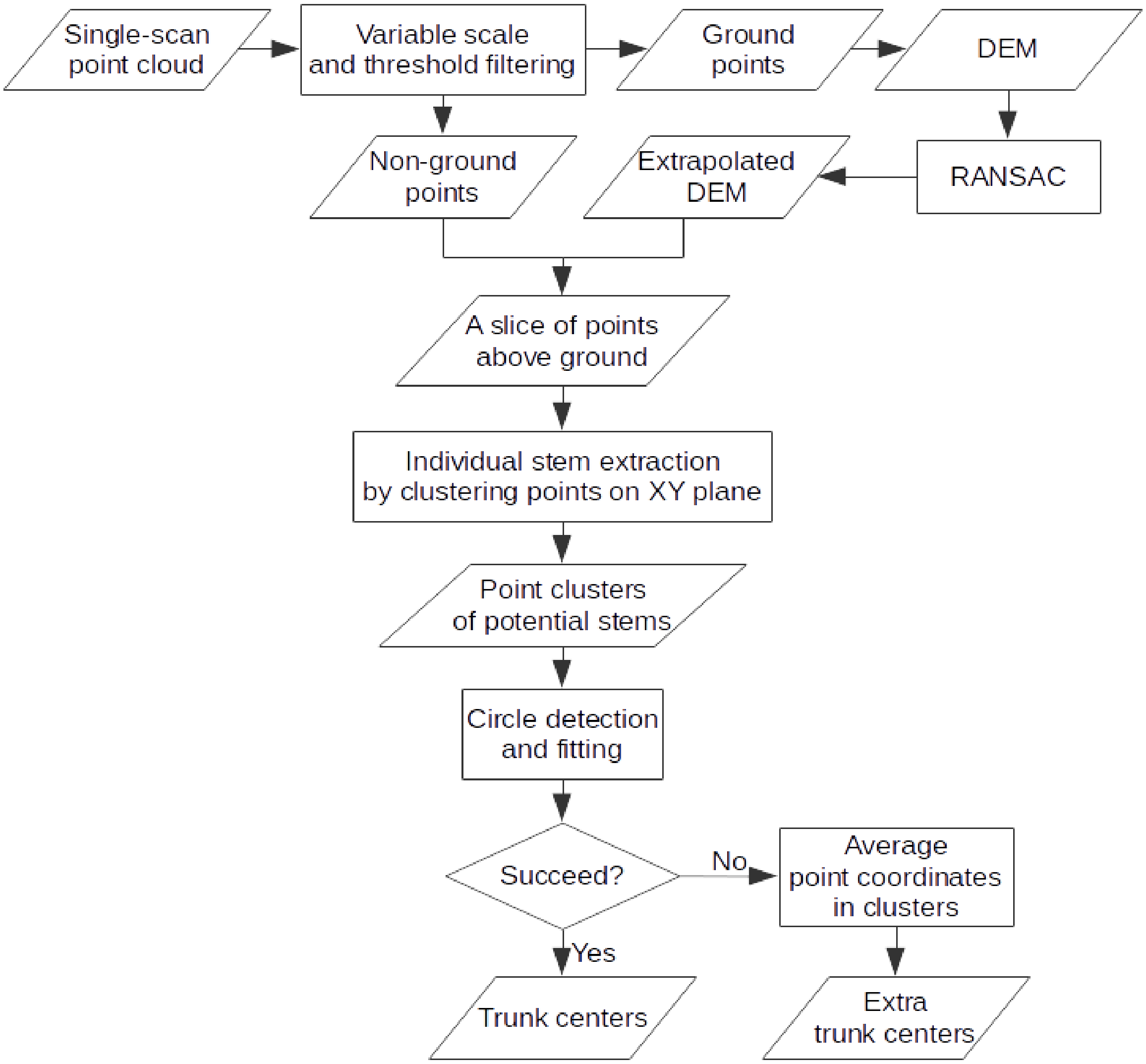


Figure 4-2. Flow chart of trunk center extraction from a single-scan point cloud.

4.3.3 AGB estimate by quantitative structure models

4.3.3.1 Description of the Quantitative Structure Model (QSM)

A newly developed technique (Raumonen et al., 2013), referred to as quantitative structure modeling (QSM), provides a method of estimating the woody volume of a single tree from a lidar point cloud by developing a 3-D model for the structure of the trunk and branches. With that estimate, and given the proper wood density, the aboveground biomass (AGB) of the tree can then be determined. Previous validations of QSM-based wood volume with simulated TLS point clouds from model trees show a relative error within 2% – 10.8% (Burt et al., 2013; Disney et al., 2012), although the result is quite sensitive to registration errors in the scans comprising the point cloud (Burt et al., 2013). A recent validation of QSM-based AGB from lidar scans of eucalyptus trees in Australia against destructive sampling shows an overestimation error of 9.68% compared to an underestimation error of 36.57% – 29.85% by allometric equations (Calders, Newnham, et al., 2014).

The QSM algorithm constructs a collection of cylinders of varying sizes for each single-tree point cloud to represent tree stems and branches for wood volume calculation. After filtering isolated noise points, QSM builds the cylindrical tree models using two steps to (1) segment a point cloud into parts of stems and branches, and (2) fit cylinders to the segments of points and fill the possible gaps between cylinders.

In the first step, point cloud segmentation, a point cloud is partitioned into small cover sets that correspond to connected surface patches of points on the tree surface. These cover sets are then grouped together to form segments of stems and branches according to their neighbor relations (connectivity between patches) and geometrical characterizations (direction of connected patches as stem and branches). One important parameter in this step, the patch size d, controls the partitioning and subsequent segmentation, and thus strongly affects the final cylindrical tree model. The larger the patch size, the coarser is the detail of stem and branches that can be modeled.

The above segmentation procedure is carried out twice in QSM first with a constant patch size, and then with varying patch sizes. The first segmentation uses a larger user-defined patch size to quickly but roughly construct segments of stem and branches. These segments generate information about branching structure and local sizes of stem and branches. The second segmentation then uses varying patch sizes so that the maximum size is at the base of the stem (trunk) and given by a smaller user-defined value than the first segmentation. The patch size then decreases linearly from the base of stem/branches to the tip of stem/branches. At the same time, the patch size at the base of branches also decreases linearly with branching order.

In the second step, cylinder fitting and gap filling, cylinders are first fit to subregions of segments. The approximate length of a subregion is controlled by a parameter l, the relative length of a cylinder defined as the ratio of cylinder length to its base radius. After fitting cylinders to segments, small gaps between cylinders due to occlusions are identified and filled with cylinders that are fitted to the previously fitted cylinders as data. The complete cylindrical tree model is reconstructed and wood volume can be calculated.

Because the segmentation step in QSM starts from random locations in the point cloud, cylindrical tree models and resultant wood volume estimates from multiple QSM will vary. Multiple runs of QSM are usually carried out for a point cloud using the same input parameters to output a mean and standard deviation of wood volume estimates.

4.3.3.2 Use of QSM on Harvard Forest scanning data for AGB estimates

We extracted five individual trees from both the leaf-off and leaf-on registered point clouds to estimate their AGB with QSM. The five trees were matched with the field measurement records to identify their species (Table 4-1). The species type is used to determine the wood densities for the AGB estimates of the five trees. Wood density here is defined as oven-dried weight divided by fresh volume including bark. Within-tree and within-species variations in wood density are not considered here as they are usually small or negligible for trees with similar climatic and geographic conditions (Baker et al., 2004; Chave et al., 2006; Gominho et al., 2001; Knapic et al., 2008). Two datasets of wood density (Chave et al., 2009; Jenkins et al., 2004) provide the same wood densities for the three species in this study (Table 4-1).

The standard deviation or variance of multiple QSM runs is an indicator of the robustness and accuracy of the reconstructed tree model given a set of model parameters. The patch size parameter d is the most critical parameter in the QSM. To find an optimal d for the DWEL point cloud and its 2-mrad angular resolution, we examined the change of standard deviations observed in 30 QSM runs with a series of d values from 4 cm to 20 cm in 2-cm increments for each tree point cloud.

Leaf points in a single-tree point cloud have been suggested as an error source for QSM (Burt et al., 2013). Leaf points thicken branches and increase the variance of segmentations, both of which can lead to larger standard deviations of multiple model runs. To examine the effects of separating leaf points from woody points by our point classification for QSM-based AGB estimates, we ran QSM with all points in each point cloud and then with only those classified as woody material, and repeated this procedure for the five leaf-off single-tree point clouds and the five leaf-on single-tree point clouds. Then we compared the standard deviation of QSM runs with all points and with those classified as woody material.

Table 4-1. List of five trees for QSM.

Tree ID	Species	Common name	DBH[1], cm	Wood density[2], g/cm^3
1	*Pinus strobus*	White pine	70.0	0.34
2	*Quercus rubra*	Red oak	60.4	0.56
3	*Quercus rubra*	Red oak	61.2	0.56
4	*Acer rubrum*	Red maple	26.1	0.49
5	*Quercus rubra*	Red oak	59.1	0.56

[1] from field measurements

[2] oven-dried weight / fresh volume, including barks, from published literature and data (Chave et al., 2009; Jenkins et al., 2004)

4.3.4 Aboveground biomass estimated by allometric equations

As a benchmark for comparison with our application of QSM, we selected species-specific allometric equations to estimate the aboveground woody biomass (stem and branches) of each tree from a widely-used database of equations for trees in the United Stated compiled by Jenkins et al. (2004). We selected equations that were developed from tree samples covering the DBHs of the studied trees here; however, the white pine was larger than the white pine sampled for all the allometric equation developments compiled by Jenkins et al. (2004), which adds some unknown uncertainty to its biomass estimate (Table 4-2). The table also shows a correction factor, which is applied to correct for potential underestimation resulting from a back-transformation of logarithmic predictions to arithmetic units (Jenkins et al., 2004). If multiple equations were available, their average was used and standard deviation was calculated.

Table 4-2. Allometric equations for the species in the study.

Species	Allometric Equation (B: biomass, D: DBH)	Correction factor	DBH of tree samples to develop the equation, cm	
			Minimum	Maximum
White pine	$\ln B = 4.19 + 2.43 \cdot \ln D$ D: cm, B: g	1	1.70	25.70
Red oak	$B = 3.28071 \cdot D^{2\times1.20936}$ D: in., B: lb.	1	28.19	62.74
Red maple	$B = 2.0772 \cdot D^{2.508}$ D: in., B: lb.	1	5.08	50.80
	$B = 1.81301 \cdot D^{2.56226}$ D: in., B: lb.	1	12.70	40.64
	$\log_{10} B = -1.096 + 2.591 \cdot \log_{10} D$ D: cm, B: kg	1.003	6.30	52.40

4.4 Results and Discussion

4.4.1 Sensitivity of QSM to the main input parameter for DWEL point clouds

Figure 4-3 and Figure 4-4 show two examples of original point clouds from tree #1 (white pine) and tree 3 (red oak) and fitted cylindrical tree models from one realization of QSM using four types of inputs: (1) all points from leaf-off scans, (2) woody points from leaf-off scans, (3) all points from leaf-on scans and (4) woody points from leaf-on scans. Here, “leaf-off” refers to the deciduous canopy; the evergreen white pine still possessed needles during the leaf-off scans. The first three columns in each figure show the point clouds for leaf-off and leaf-on acquisitions. Points in the three columns are displayed according to the scan position from which they were acquired; as a color-composite; and as classified into wood and leaf hits (see Chapter 3). There are clearly fewer points in the upper crowns of each tree in the leaf-on scans, which result from occlusion by leaves and loss of points by weak partial returns. The last two columns compare the results using all points and only woody points (see the discussion in the section 4.4.2). The different branching structures of the two species are clearly shown. The white pine (Figure 4-3) displays branches from near the base to the top; irregular branching at the canopy top could result from the old damage to the top of the tree or just the missing view of upper crowns due to occlusions by leaves and loss of points by weak partial returns. The red oak shows a much cleaner main stem and two prominent upper-canopy branches arching out to either side.

Figure 4-5 shows the mean and standard deviation of tree wood volume estimates of QSM structures for different patch size values using the four types of input point clouds for each of the five trees. The wood volume estimates increase with the QSM parameter d, point patch size for all the four types of input point clouds to QSM (Figure 4-5, left column). This steady increase was also observed in the QSM results in a Eucalyptus forest (Calders, Newnham, et al., 2014). Although some of the wood volume estimates seem to show a possible asymptote as the QSM point patch size goes even larger, the optimal patch size value is not expected to be at the asymptote. Larger patch size leads to coarser segmentation of point clouds. Points of several nearby branches are more likely to be aggregated together as one coarse segment under a large QSM point patch size, resulting in one large cylinder that gives a greater volume than the total volume of several smaller cylinders correctly fitted to those branches. Thus, the wood volume estimate by QSM will increase with the patch size until almost all the branching structure details are lost. Calders et al. (2014, fig. 8) also showed that the patch size that gave the closest AGB estimate by QSM to the destructive measurement is not a value that provides a possible asymptote volume but a value that models the stem well without over-fitting to the point cloud.

If a patch size captures stem and branches well with correct and sufficient details but without over-fitting with too many small cylinders, the cylindrical tree model will be stable and the variance of multiple wood volume estimates is expected to be small. Thus, the standard deviation of wood volume estimates is used to select the optimal QSM point patch size. Generally, the standard deviation in wood volume estimates decreases with increasing point patch size and then increases (Figure 4-5, right column). When the point patch size is too small, the points of a branch are segmented into multiple smaller parts and over-fitted with multiple smaller cylinders, especially when the branch is not fully covered by points due to some occlusion. As noted above, the variation of tree model building and wood volume estimates comes from the random start locations of segmentation. As the point patch size goes even smaller than the optimal value, more but smaller segments and resultant cylinders that are over-fitted to the points of branches will have higher chance to vary from different start locations of segmentation, which therefore results in more variation in tree models and larger standard deviation in wood volume estimates. On the other hand, when the QSM point patch is too large, points of multiple branches will be aggregated together as one coarse segment. As point patch size goes even larger than the optimal value, more different aggregations of multiple branches are possible and therefore different start locations of segmentation will also cause more variation in tree models and larger standard deviation in wood volume estimates.

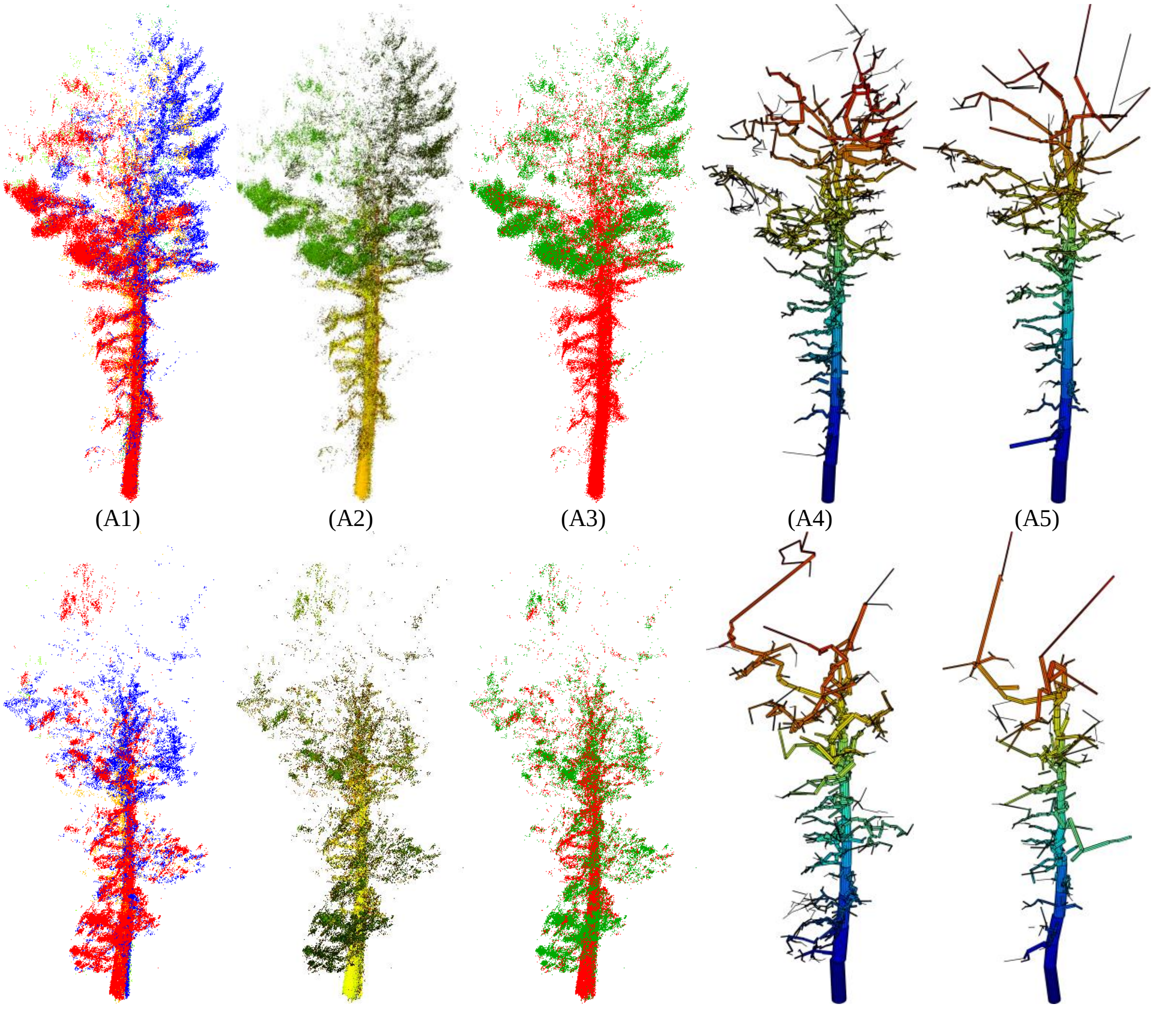
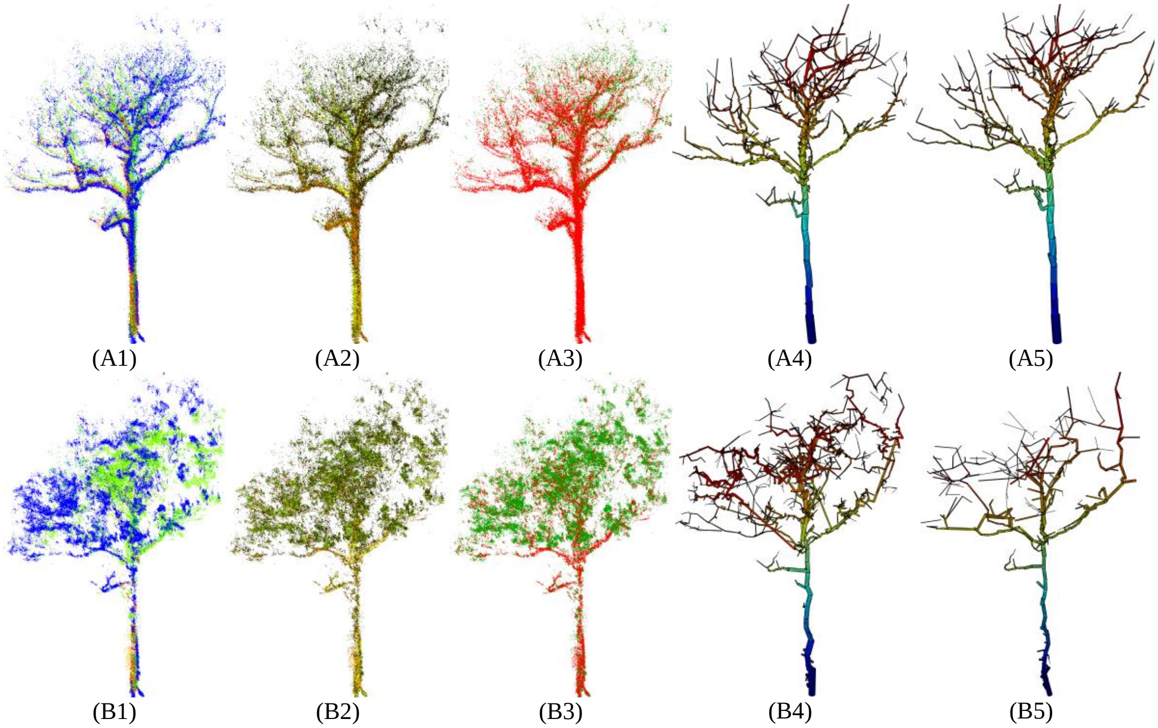


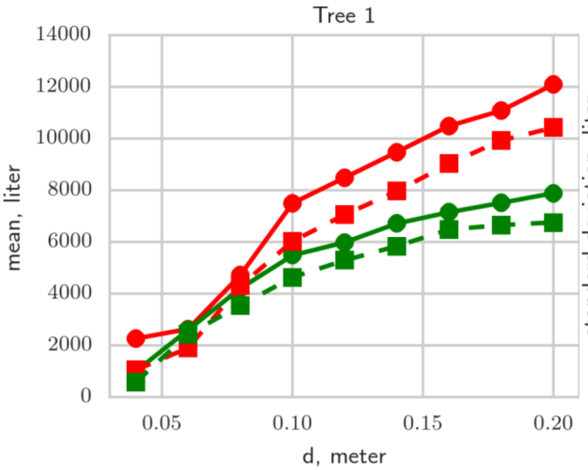
Figure 4-3. Point clouds of the white pine (Tree ID = 1) and cylindrical tree model from one sample QSM realization using a patch size of 0.11 m. First row: leaf-off; Second row: leaf-on. First column: point colors show different scan locations; Second column: color-composite points, red (1548 nm), green (1064 nm) and blue (dark); Third column: point colors show classification, red as woody materials and green as leaves; Fourth column: tree model with all points; Fifth column: tree model with only woody points.

The general trend of standard deviation against QSM point patch size for our five trees (Figure 4-5, right column) suggests the optimal patch size for the DWEL point clouds here was 0.10 – 0.12 m when the standard deviation is low. We chose the middle value 0.11 m as the optimal patch size for QSM with the 2-mrad DWEL point cloud here.

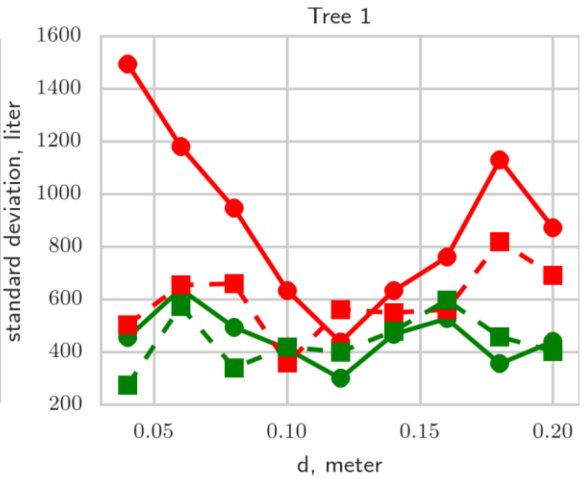

(A1) (A2) (A3) (A4) (A5)

(B1) (B2) (B3) (B4) (B5)

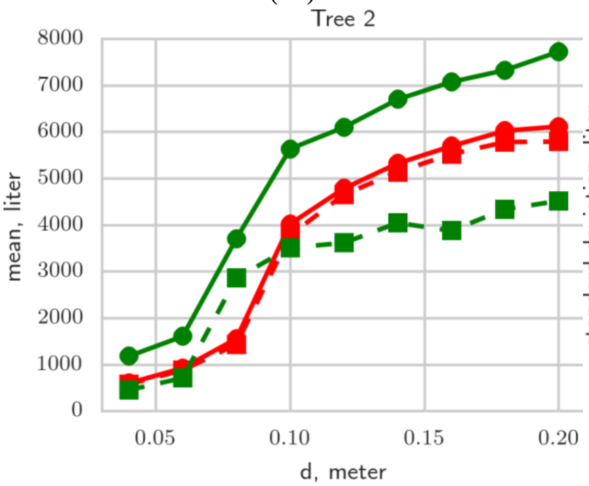
Figure 4-4. Point clouds of the red oak (Tree ID = 3) and cylindrical tree model from one sample QSM realization using a patch size of 0.11 m. First row: leaf-off; Second row: leaf-on. First column: point colors show different scan locations; Second column: color-composite points, red (1548 nm), green (1064 nm) and blue (dark); Third column: point colors show classification, red as woody materials and green as leaves; Fourth column: tree model with all points; Fifth column: tree model with only woody points.



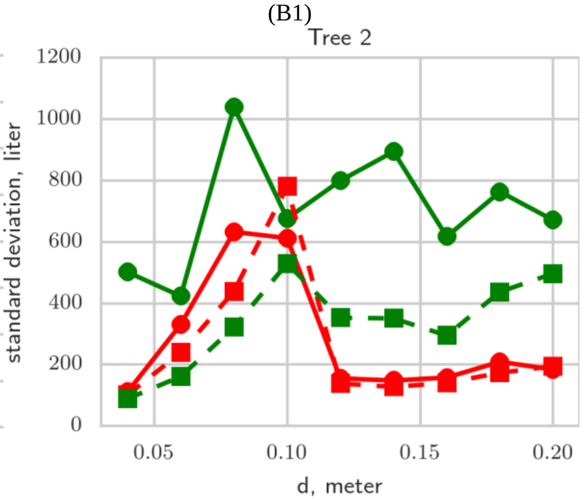
(A1)



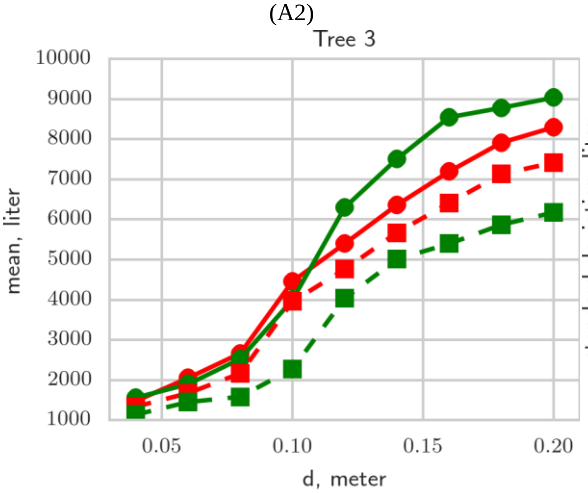
(B1)



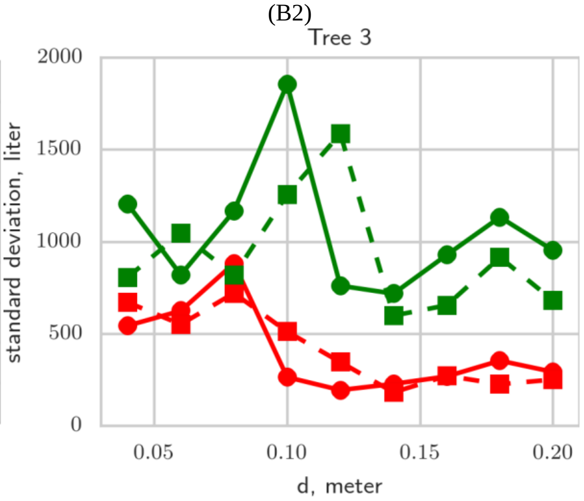
(A2)



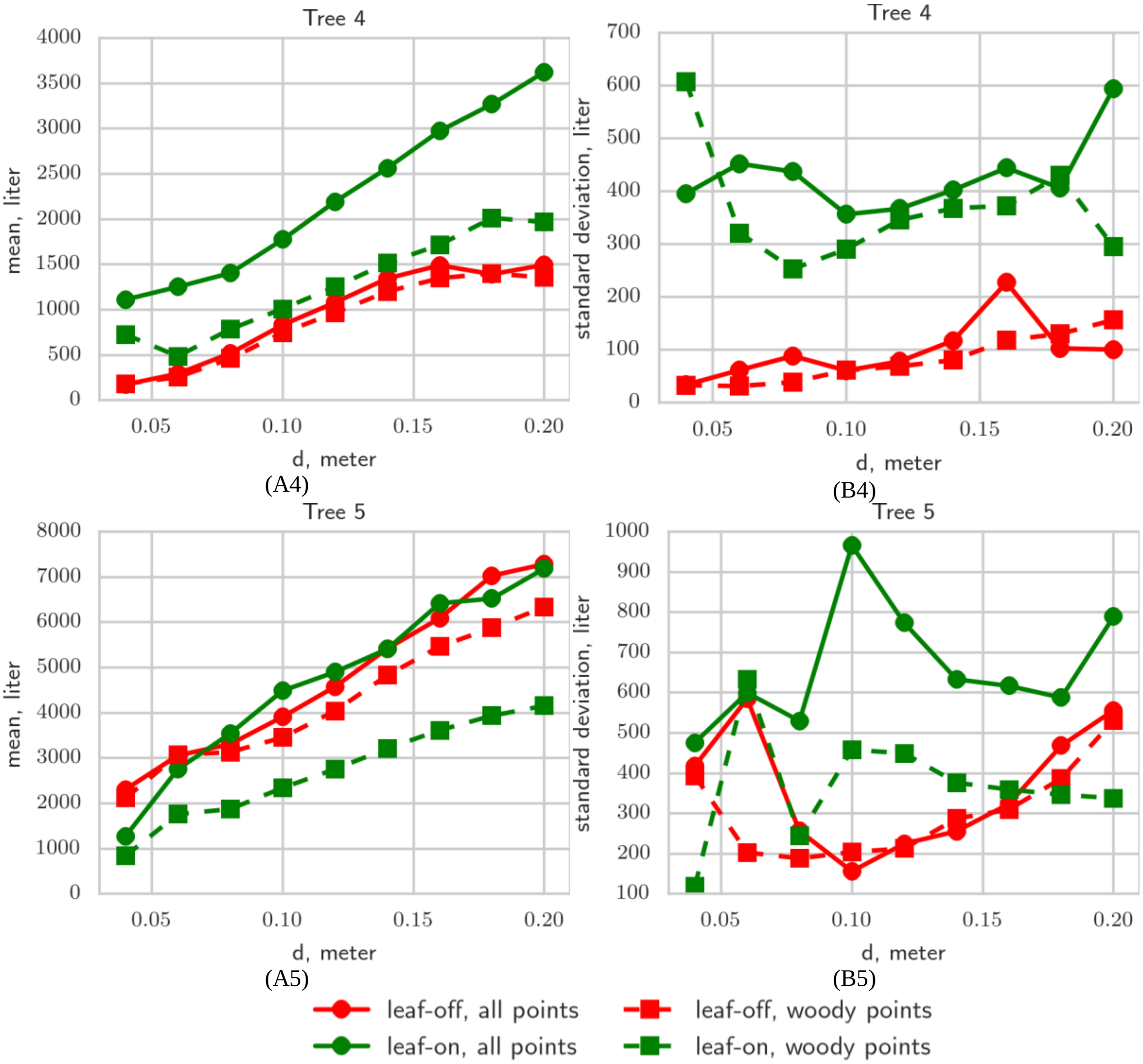
(B2)



(A3)



(B3)



(A4) (B4)

(A5) (B5)

Figure 4-5. Mean (left column) and standard deviation (right column) of tree wood volume from multiple QSM runs against patch size parameter. Each row is for one tree. Red dot: QSM with all points in leaf-off point clouds. Red square: QSM with woody points in leaf-off point clouds. Green dot: QSM with all points in leaf-on point clouds. Green square: QSM with woody points in leaf-on point clouds.

4.4.2 Effects on QSM results by the separation of leaf and woody points

More leaf hits in point clouds tend to increase the variance of cover set generation by QSM and thus the variance of wood volume estimates. For the deciduous trees (Figure 4-5, 2nd to 5th row), the difference in both the means and standard deviations between using all points and using only woody points from leaf-on scans is larger than those from the leaf-off scans because increased leaf hits in the leaf-on scans add more variance to the tree model reconstruction by QSM.

For the evergreen white pine (Figure 4-5, 1st row), the differences in the means and standard deviations between using all points and using only woody points from leaf-on scans are closer to those from leaf-off scans. This is because all leaf points exist in the point clouds of the white pine from both leaf-off and leaf-on scans. We also note that for the white pine, the difference between using all points and only woody points from leaf-off scans is slightly larger than that from leaf-on scans. This is probably because less occlusion by broadleaf deciduous trees and shrubs in leaf-off scans causes a larger proportion of leaf points in the point cloud derived from leaf-off scans than from leaf-on scans. That is, the decrease of leaf point numbers from leaf-off to leaf-on scans is larger than that of woody point numbers for the white pine.

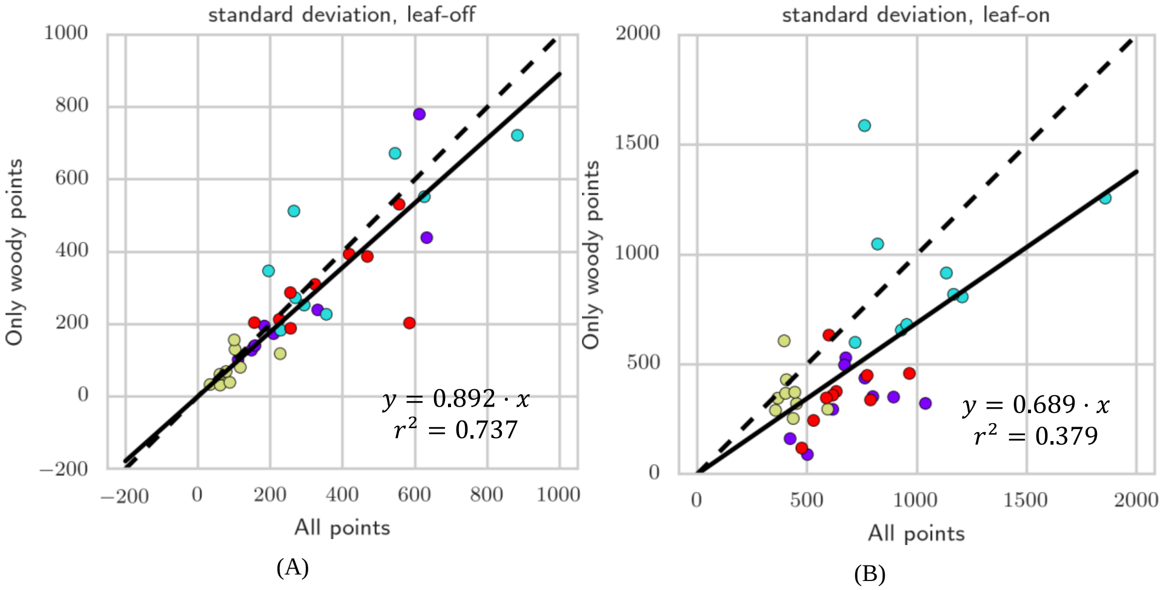
To quantitatively examine whether and how much the separation of woody points from leaf points in the point cloud can reduce the variance of wood volume estimates by QSM, we plotted standard deviations of the wood volume estimates of all deciduous trees for all the test QSM point patch size values using all points against those using only woody points. Figure 4-6 shows the scatter plot and the linear fitting with fixed intercept at zero for leaf-off scans (left panel) and leaf-on scans (right panel). The slope of the linear fitting is a ‘mean’ ratio (denoted as α) of the standard deviation of wood volume estimate using only woody points to those using all points. The lower the ratio, the more the separation of woody point from leaf points reduces the variance of wood volume estimates by QSM. Both leaf-off and leaf-on scans show α ratio less than 1, indicating the reduction of variance of wood volume estimates by separating woody point from leaf points. The scatter plot of leaf-on scans shows more data points under the one-to-one line and a lower α ratio, which suggests more improvement of wood volume estimates by only using the woody points in a leaf-on point cloud.

The stronger effect by leaf point removal, i.e. lower α ratio, for leaf-on wood volume estimates is expected because leaf-on scans have many more leaf points that will add variance to the fitted tree models and wood volume estimates. Indeed, visual inspection of the fitted tree models from all points versus woody points revealed more erroneous cylinders caused by leaf points. For example, comparing the tree models of tree 3 in Figure 4-4 (B4) and (B5), the model from the all-points situation (Figure 4-4 (B4)) has many tiny cylinders in the middle and right side of the canopy that cause fuzzy edges of large branches, while the model from only the woody points (Figure 4-4 (B4)) does not. The tree models from the leaf-off scans (Figure 4-4 (A4) & (A5)) can serve a relative benchmark of tree branching structure, as they have clearer views and less occlusion of branches inside the canopy. By comparing tree models from the leaf-on scans to the leaf-off scans and the point cloud (Figure 4-4 (B3)), it is evident that these tiny cylinders in Figure 4-4 (B4) are mostly errors caused by leaf points.

The extent of occlusion also affects the variance in the wood volume estimates. The forest was much more open in the leaf-off season than the leaf-on, and thus less occlusion occurs in the point clouds. For the deciduous trees (trees #2 – #5), using either all points in a point cloud or only the woody points, the standard deviations in wood volume estimates from the leaf-off scans are generally lower than those from leaf-on scans (Figure 4-5, right column, 2nd row to 5th row).

For the evergreen white pine tree (tree #1), the standard deviations in wood volume estimates from the leaf-off scans are not as different (Figure 4-5(B1)) because all leaf points exist in the point clouds of the white pine from both leaf-off and leaf-on scans. But the standard deviations using all points from leaf-off scans (red solid line in Figure 4-5(B1)) is larger than using the other types of point cloud inputs for the QSM. This is caused by the larger proportion of leaf points in the point cloud from the ‘leaf-off’ scans of the white pine than from the leaf-on scans. The increase of variation in tree models from the increase of leaf points in the ‘leaf-off’ point cloud of the white pine outranges the decrease of variation from the decrease of the occlusion of woody components in leaf-off scans.

The increase of standard deviations in wood volume estimates caused by the increase of occlusion is likely from two reasons. First, more occlusion reduces the number of points covering the surfaces of stem and branches. The cylinder fitting thus becomes more sensitive to the segmentation. Second, the more complete and clearer views of the woody components of the trees, especially of the branches in the canopy, in the leaf-off point clouds cause fewer gaps between fitted cylinders that QSM has to identify and fill. The identification and filling of gaps adds less accurate cylinders to tree models than those directly fitted to points and can result in incorrect branching structure if too many gaps exist. Figure 4-3 (B4 & B5) shows an example of incorrect branching structure resulting from incorrect gap filling by the QSM at the canopy top because of the large occlusion and insufficiency of points. Thus, more gaps in the cylinder fitting procedure for the leaf-on point clouds add more variance or errors to the tree models and wood volume estimates.



(A)

(B)

Figure 4-6. Scatter plots of standard deviations of tree wood volume (unit: liter) between using all points (Y axis) and using only woody points (X axis). Dot colors are for different trees, purple (tree 2, red oak), cyan (tree 3, red oak), yellow (tree 4, red maple), red (tree 5, red oak). Black solid line is linear fit with intercept fixed at zero. Black dash line is one-to-one line. (A): leaf-off scans; (B): leaf-on scans.

4.4.3 Comparison between AGB estimates by QSM and allometric equations

Figure 4-7 shows the AGB estimates of the five trees using allometric equations and field measured DBH alongside the AGB estimates with QSM when using the chosen optimal patch size (0.11 m) for the four types of input point clouds to QSM. Of the four types of QSM-based estimates, woody points only from leaf-off scans gives the lowest variance, which suggests it is the most credible estimate by QSM. Indeed, leaf-off scans have less occlusion and thus provide better views of stem and trunks, and removal of leaf points further reduces variances in cylinder fitting. For all the deciduous trees (tree #2 – #5), using all points from both leaf-off scans and leaf-on scans gives larger AGB estimates than woody points only from leaf-off scans. The larger estimates are probably due to the thickening of branches by leaf points and addition of erroneously fitted cylinders from leaf points. In contrast, the woody points only from leaf-on scans gives smaller AGB estimates than woody points only from leaf-off scans. Using woody points only from leaf-on scans removes positive errors caused by leaf points, but also increases negative errors caused by the greater number of gaps and their larger size, produced by strong leaf occlusion in the tree point clouds, that cannot be filled by QSM.

The comparison between the four types of QSM-based AGB estimates for the evergreen needleleaf white pine (tree 1) is somewhat different from the deciduous trees. Using all points from leaf-off scans still gives a larger AGB estimate than using woody points only from leaf-off scans. But all points from leaf-on scans gives a smaller AGB estimate than woody points only from leaf-off scans. Comparing the point clouds of the white pine from leaf-off and leaf-on scans (Figure 4-3, A-B), a large part of the white pine canopy top in the leaf-on scans is missing due to strong occlusion, which causes the lower wood volume.

As the comparison between QSM-based AGB estimates using the four types of input point clouds suggests, different degrees of occlusion and different amounts of leaf points in input point clouds not only result in different variances in AGB estimates, but may also lead to different AGB estimate values. Small gaps caused by small occlusions can be identified and filled by QSM, and probably will not cause significantly different wood volume estimates. However, large and numerous gaps caused by large occlusions cannot be filled by QSM and also can result in incorrect branching structure, as we have seen in Figure 4-3 (B4 & B5). Thus, strong occlusion will produce different wood volume estimates, which are usually underestimates because of lost stems and branches in point clouds. Leaf points will increase wood volume estimates by thickening branches and causing more erroneously fitted cylinders. Therefore, different amounts of leaf points in point clouds may result in different wood volume estimates as well.

In general, we may conclude that it is better to use leaf-off scanning data in QSM for AGB estimates to avoid errors caused by leaf points. But leaf-off scanning is not always an option, e.g. for evergreen conifers or broadleaf evergreen tropical trees. In such cases, separation of woody points from leaf points can produce a better estimate. However, more leaf points usually indicate stronger occlusion due to more leaves, which may cause underestimation of wood volume due to large and numerous gaps. To fill these gaps in leaf-on canopies, the solution lies in improving scanning data collection, including a better design of scan location layout, longer lidar measurement range, finer scanning resolution, and stronger capability of detecting multiple returns.

The effects of occlusion in the current DWEL point clouds, especially of branches in upper canopies, are worsened by the instrument’s limited measurement range due to low laser power and high noise levels, which causes distant and partial hits to be lost. Even leaf-off point clouds lose small branches and some parts of big branches in the upper canopy at far range. As a result, AGB estimates here by QSM are likely to be underestimated. This underestimate seems to partly explain the generally lower AGB values by QSM as compared to the results from allometric equations (Figure 4-7). But it is not possible to say which estimate is more accurate without direct destructive sampling.

Calders et al. (2014) compared both AGB values obtained by QSM and values obtained from allometry with biomass measured by destructive sampling in Australian eucalypt woodlands and reported a slight overestimate of AGB by QSM and significant underestimate by allometric equations. However, no destructive sampling has yet been reported for direct comparisons of QSM-based estimates against allometric equations of North American species. Moreover, allometric equations are best applied to large sizes of tree samples rather than individuals. Only five well-reconstructed trees are selected here due to limited lidar measurement range, which is a very small sample size. Clearly, we need destructive sampling and a larger sample size to directly validate the accuracy of AGB estimates by QSM and allometric equations.

Point cloud registration error is another source that contributes to the AGB estimate uncertainty and to the difference in AGB estimates between leaf-off and leaf-on scans. Trunks and branches are the main targets for our registration procedure. The more open view of trunks and branches in the leaf-off scans likely gives higher registration accuracy than leaf-on scans, which may lead to more accurate wood volume estimates from leaf-off scans. The AGB estimates can be further improved by taking into account within-tree variations in wood density. Wood densities of many species show radial variations from pith to bark (Espinoza, 2004; Henry et al., 2010; Woodcock & Shier, 2002). Stems and branches of different sizes will thus have different wood densities. The detailed branch sizes from cylinder models can be utilized to include the variation in wood density into the AGB biomass to reduce estimate errors.

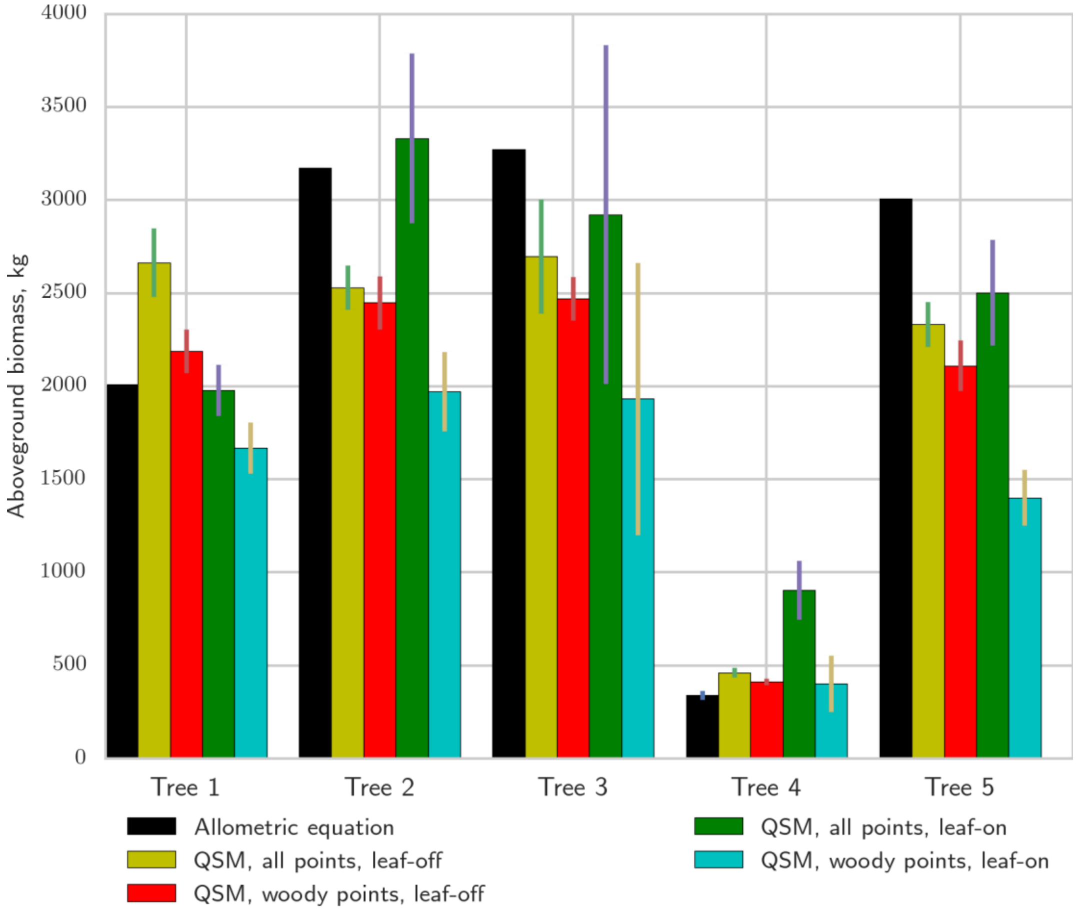


Figure 4-7. AGB estimates (unit: kg) of five trees from allometric equations and QSM with all points or woody points from leaf-off and leaf-on point clouds. Error bars are one standard deviation.

4.5 Conclusions

This study has shown that the DWEL's dual-wavelength data has the capability to improve nondestructive estimates of aboveground biomass of individual trees using Quantitative Structure Modeling (QSM). This new technique builds tree models from virtual cylinders using registered multiscan terrestrial lidar point clouds, determines the volume of each tree structure, applies a wood density factor, and then estimates the biomass of the tree. We used the standard deviation of wood volume estimates from multiple QSM runs as a criterion for the selection of point patch size, a critical input parameter to QSM. We reduced the variance in QSM-based estimates through the separation of leaves and woody materials with DWEL scanner data, exploring the use of all points and only woody points as applied to a sample of forest trees scanned in both leaf-off and leaf-on seasons.

In our data analysis, we found that leaf points in point clouds of individual trees increase the variance in the wood volume estimate by QSM; leafless point clouds are better for QSM to build tree models and estimate wood volumes and resultant AGB. Scanning leaf-off trees is a way to obtain leafless point clouds. But this is not always an option, e.g. for conifer or tropical trees or when the field campaign is carried out in leaf-on seasons. In that case, DWEL's classification of leaf and woody points in three-dimensional space data can be used to remove leaf points, leaving only woody points for the QSM procedure.

Occlusions in point clouds also increase the variance in QSM-based estimates. First, occlusions increase the variance in cylinder fitting by reducing the points covering stem and branches. Occlusions also increase the gaps that QSM has to identify and fill and can result in incorrect branching structure in tree cylinder models, which increases the variance in wood volume estimates.

The point cloud registration error is another source that contributes to the variances and errors in AGB estimates, and more work is needed to assess the registration accuracy quantitatively and check differences in AGB estimates caused by registration errors. For example, brighter marker targets need to be placed in scan areas where they are visible from different scan views to assess the registration accuracy. Alignment between scans could then be varied to create different registration accuracies and observe the sensitivity of AGB estimates to registration accuracy change. This procedure would thus estimate the difference in AGB estimates that is caused by the difference in the registration accuracy between leaf-off and leaf-on scans.

Proper parameter selection criterion and removal of leaf points based on point classification improve the QSM-based AGB estimates by reducing the estimate variance. This facilitates the application of the QSM-based approach, which helps the calibration and validation of large-scale AGB mapping by providing fast and accurate ground-based AGB estimates directly. Moreover, the QSM approach is independent of allometric equations, which are known as a source of biases in biomass estimation, and can help establish and update allometric equations fast and accurately for more forest sites, more species, and more trees, especially large ones. However, destructive sampling remains the best method to further validate AGB estimates by QSM and quantitatively check the sensitivity of QSM to point patch size, leaf point labeling and registration error.

CHAPTER 5. CONCLUSION

Forest structure is critical information for monitoring and modeling carbon sequestration, emission, and storage of forests to balance the global carbon budget and manage the forest carbon cycle for climate change mitigation. Almost all current approaches to forest structure retrieval at coarse scales (landscape, regional and global) require ground measurements for calibration; moreover, all methods require ground measurements for validation. This dissertation uses bispectral information from a novel terrestrial laser scanner, the Dual-Wavelength Echidna Lidar (DWEL), to evaluate how dual-wavelength lidar data can improve ground-based measurements of forest structure. For this task, I addressed the technical challenges and scientific implications that relate to the extraction and interpretation of separate structure measurements of leaves and woody materials from the scans of a midlatitude forest site at Harvard Forest, Petersham, Massachusetts, USA, using this full-waveform dual-wavelength terrestrial lidar. The integration of spectral and 3-D spatial information by the DWEL and the resultant differentiation of leaves and woody materials, for the first time in three-dimensional space reveal new details of forest structure, improve forest structure measurements, and open possibilities of collecting more detailed biophysical and biochemical properties of forest ecosystems in three-dimensional space to broaden and deepen our understanding of forest ecosystems.

5.1 Radiometric Calibration of the Dual-wavelength Terrestrial Lidar

Measurements of intensities returned to a lidar instrument by scattering of lidar pulses, coupled with much more commonly used 3-D spatial locations of scattering events from terrestrial laser scanners, have demonstrated great value in forest structure measurements, including leaf area index (LAI), vertical foliage profile, diameter at breast height (DBH), and stem count density, (Jupp et al., 2009; Lovell et al., 2011), as well as in target classification (Béland et al., 2014) and estimation of plant physiological properties (Eitel et al., 2010; Gaulton et al., 2013; Magney et al., 2013). Radiometric calibration allows for generalized and deterministic uses of lidar intensities in space and time. It is a prerequisite to use lidar intensities synergistically from different instruments, different scanning campaigns, and different wavelengths from emerging spectral lidar instruments such as the DWEL.

Calibration of our full-waveform, dual-wavelength, terrestrial laser scanner presents a number of challenges relevant to the next generation of terrestrial laser scanners. The need to scan from near to far range requires characterizing both telescopic effects, which reduce the near-range signal with increasing proximity due to defocusing, and saturation effects, which alter the return pulse shape of near-range scattering events. In addition, dual- or multiple-wavelength lidar data must be consistent in spectral performance to avoid instrument optics and electronics from causing artifacts in derived spectral contrasts for further data processing and information extraction, such as classification of vegetation components and estimation of plant physiological properties (Eitel et al., 2014; Gaulton et al., 2013; Morsdorf et al., 2009).

In Chapter 2, the formulation of the lidar equation using Ross’s framework of radiation regime modeling of vegetation canopies (Ross, 1981) allows us to connect the lidar return signal explicitly to forest canopy structural parameters. This lidar equation identifies the objective variable of lidar calibration, i.e. apparent reflectance, a physically-defined value related to the size, orientation, and reflectance of a target, independent of range and instrument optics and electronics. Apparent reflectance is theoretically useful because of its connection with forest structural parameters and practically realizable for interpreting it as a reflectance factor.

Apparent reflectance is a function of size, orientation, and reflectance characteristics of a target. Further extraction of biophysical and biochemical information about forests usually requires disentangling these three components of apparent reflectance, which is not easy. Spectral lidar instruments, such as the DWEL, have the potential to solve the problem by canceling out size and orientation by comparing lidar return signals at different wavelengths from the same target.

However, when comparing these apparent reflectance values, errors in the calibration at individual wavelengths can be offsetting or amplified, causing a larger range of uncertainty in spectral contrast measurements from spectral lidar instruments, thus requiring high calibration accuracy. The RMSE values (relative errors) of apparent reflectance achieved here, 0.081 for 1064 nm and 0.064 for 1548 nm from the DWEL calibration, show that the calibration model and parameter fitting procedure accurately convert lidar return intensities in digital counts to apparent reflectance.

The methodology for DWEL calibration includes three unique features: (1) explicit inclusion of telescopic effects in a semi-empirical calibration model; (2) a joint calibration model fitting that estimates calibration parameters for both wavelengths together; and (3) a spectral constraint using the normalized difference index in this joint calibration model fitting. This calibration methodology can apply to almost any terrestrial lidar instrument using a telescope to focus the return power, and thus it is of potentially wider use for calibration of other terrestrial lidar systems, particularly spectral lidar ones.

5.2 Separate Height Profiles of Leaf and Woody Areas

Leaf area index and vertical foliage profile regulate the radiation interception of the forest canopy and thus the canopy photosynthetic capacity and carbon assimilation of a forest stand. Direct removal of nonphotosynthetic elements, mainly woody materials in the 3-D space of a forest canopy, improves the measurement of LAI and the vertical foliage profile. Separate height profiles of leaf and woody areas, explicitly in 3-D space, may also improve the simulation and inversion of ecophysiological and 3-D radiative transfer models of forest ecosystems.

In Chapter 3, the differentiation of leaves and woody materials in three-dimensional space was improved by the synergistic use of spectral information from dual-wavelength apparent reflectance and spatial context information given by the 3-D spatial distribution pattern of each point and its neighbors in the bispectral lidar point cloud. Given the lack of reference leaf-versus-wood classification of higher quality in three-dimensional space, I proposed an indirect approach to accuracy assessment of our 3-D point classification using color hemispherical photos and manual photointerpretation selected according to a random stratified sample design. The classification accuracy observed (in number of points), 0.70 ± 0.02 to 0.88 ± 0.01 for leaf-off point clouds and 0.70 ± 0.02 to 0.81 ± 0.01 for leaf-on point clouds, proved the value of combining the spectral and 3-D spatial information from the two scanning lasers. The examination of classification error sources highlighted the importance of radiometric calibration accuracy and laser beam alignment accuracy in using the added spectral information of multiple lasers at different wavelengths.

Separate height profiles of leaf and woody areas were retrieved from separate gap probabilities over range brought by leaves and woody materials through their classification in 3-D space. The variance in leaf and wood area profiles over the test forest site from different scan locations revealed the large and changing spatial heterogeneity of vertical structures of leaves and woody materials at this 1-ha forest site over the seasons. In the leaf-off season, the variance in the leaf area profile at different canopy heights was relatively larger than that of the woody area profile, reflecting the presence of needleleaf crowns of several white pines at different locations in this site. In the leaf-on season, the variances were relatively similar in leaf and woody area profiles at different canopy heights, reflecting the similar spatial distribution of leaves and woody materials with canopy height over the forest site.

The separate height profiles of leaf and woody areas also revealed the vertical pattern of woody-to-total ratios (woody area to total vegetation area including both leaf and woody areas), a commonly-used empirical parameter for the correction of woody contributions to LAI retrievals using optical methods. The woody-to-total ratios were generally stable in the middle and upper canopy but varied in the lower canopy with height due to the varying distributions of leaf area from the understory trees and woody areas from stems of large trees in the scanned volume above each scan location.

5.3 Improvement of a Nondestructive Biomass Measurement Approach

Allometric equations are commonly used in ground measurements of forest aboveground biomass for the calibration and validation of large-area measurements. However, inaccuracies in, and inappropriate applications of, allometric equations are still large sources of error in AGB estimation (Basuki et al., 2009; Chave et al., 2004; Keller et al., 2001; Wang, 2006).

Chapter 4 introduced a state-of-the-art, nondestructive approach to AGB estimation that is independent of allometric equations. It combines *a priori* wood density information with wood volume that is directly calculated from cylinder tree models, built from TLS point clouds using Quantitative Structure Modeling (QSM) (Burt et al., 2013; Calders, Newnham, et al., 2014; Kaasalainen et al., 2014; Raumonen et al., 2013). The removal of leaf points in a single-tree point cloud through the classification of leaves and woody materials in 3-D space reduced the variance in the QSM-based wood volume estimate and thus improved the biomass estimate. This improvement shows another advantage of differentiating leaves and woody materials in 3-D space and further supports the addition of spectral information to the 3-D spatial domain with dual-wavelength scanning. The reduction of variance in QSM estimates on the removal of leaf points calls for more accurate 3-D classifications of points in lidar point clouds.

The examination of variance sources of QSM-based AGB estimates and the comparison with AGB estimates by allometric equations also highlighted the importance of reducing occlusion in DWEL multiscan point clouds for more complete 3-D reconstruction of forests. The improvement of 3-D reconstructions with the DWEL will require a longer measurement range and a better ability to detect the weaker signals of partial hits in the leaf-on canopy, as well as more scan locations or a more adaptive design of scan location layout.

5.4 Future Research

To map and monitor the spatial and temporal variations of forest structure in support of understanding, modeling and managing the carbon cycle under a changing climate, ground measurements of forest structure need to be collected extensively, accurately, and routinely for the calibration and validation of large-area measurements. This need is quite challenging to traditional ground measurement approaches, but can be achieved through emerging and innovative technologies, such as terrestrial laser scanning, as many studies and this dissertation have demonstrated. However, terrestrial laser scanning cannot fully fulfill its promise to provide rapid and accurate forest structure measurements without further efforts in instrument building and a collaborative framework of data collection and processing.

For the DWEL instrument, we need to increase measurement range through higher laser power and a higher signal-to-noise ratio to overcome current limitations on LAI retrieval from gap probability inversion and wood volume estimates from quantitative structure modeling. The DWEL scanning reported here demonstrated the value of the integration of spectral and 3-D spatial information from spectral lidar instruments. However, to obtain a stable spectral contrast from multispectral lidar return signals in further data processing, a rigorous requirement of laser alignment accuracy must be met for spectral lidar instruments using multiple lasers. This laser alignment is not trivial but quite challenging and needs improvement in engineering techniques. A different approach is to build a spectral lidar using a single supercontinuum laser, i.e. a white laser, without having to align multiple lasers (Kaasalainen et al., 2007). To bring this type of spectral lidar instrument into safe and operational use in forests, however the laser power needs to be reduced while the signal detection ability of each wavelength will need to be enhanced. Also, the cost of white lasers will need to fall through technology development.

Occlusion of upper canopy is one challenge to forest structure measurements by terrestrial laser scanning, especially for tall and/or dense forests, such as old-growth forests, tropical forests and etc. Combining under-canopy scans from terrestrial lidars and above-canopy scans from airborne lidars is a way to reduce the occlusion. Deploying terrestrial lidar instruments at different heights is also a viable approach, but requires solving the technical problem of moving lidar instruments to different canopy heights. The vertical movement of lidar instruments can be simply achieved through a permanent or movable (quick and temporary setup) tower or mast inside forests, if available. Flying a lidar instrument with a multicopter unmanned aerial vehicle (UAV) along the canopy height is also possible, but needs further technology development to navigate UAVs through complex forest canopy environments.

In addition, both tower and UAV approaches prefer and/or require light lidar instruments. Further technology development is required to reduce the payload of lidar instruments while maintaining good scanning quality and as much information as regular lidar instruments provide for forest structure measurements. Moreover, lighter terrestrial lidars are easier to carry into rugged forest areas at more geolocations, which reduces the bias in the collection of ground measurements towards more easily-accessed forest areas.

Terrestrial lidar data processing for forest structure measurement is advancing rapidly and extensively for different forest types and geographic conditions through various studies. These studies focus on their own unique challenges and problems and therefore may use different instruments, different scanning setups, different scan location layouts, and/or different processing approaches even for the estimation of the same structural parameter. However, in support of operational monitoring of forest structure, a collaborative framework of TLS data collection and processing needs to be set up after intercomparison of instruments, scanning protocols, and processing algorithms. A guideline for data collection at different forest environments is needed, including minimum and recommended requirements of TLS instruments, scan location layouts, scanning condition (wind, temperature, and etc.), and scanning data quality (measurement range, scanning resolution, intensity noise level and etc.).

In addition, a top-level processing scheme is needed to specify several essential and basic modules in the retrieval of forest structure from low-level lidar scanning data. Data processing modules can be implemented according to different forest environments and specific objectives. An open source library of data processing tools needs to be collaboratively developed and maintained.

Moreover, we need to set up a collaborative framework to facilitate the conversion of the many TLS scans collected by forest research and management communities around the world to consistent and comparable ground measurements of forest structure across different forest types and geographic conditions. These ground measurements from terrestrial laser scanning, acquired through this collaborative framework, will result in progress in understanding and projecting the carbon cycle forward and greatly benefit human management of the carbon cycle and resulting climate change.

APPENDIX

A.1 Area scattering phase function for Lambertian facets of the same diffuse reflectance

Let $g_L(r, \Omega_L)$ (sr^{-1}) be the probability function of leaf angle distribution at location designated by range r along a laser beam of interest, the Ross *G*-function is (Ross, 1981),

$$G(r, \Omega) = \frac{1}{2\pi} \int_{\Omega_1} g_L(r, \Omega_L) |\cos\langle \Omega, \Omega_L \rangle| d\Omega_L \qquad \text{A 1}$$

where the integral interval Ω_1 defines all the possible leaf normal directions within an elemental volume at range r. Let $(1/\pi)\gamma_L(\Omega_L, \Omega_i \to \Omega_v)$ be the leaf scattering phase function that defines the part of irradiance in the direction Ω_i which is scattered from the leaf unit area perpendicular to the direction Ω_L to the unit solid angle around the direction Ω_v (Ross, 1981). The area scattering phase function is derived as,

$$\begin{aligned} &\frac{1}{\pi}\Gamma(r, \Omega_i \to \Omega_v) \\ &= \int_{\Omega_1} \frac{g_L(r, \Omega_L)}{2\pi} \frac{\gamma_L(\Omega_L, \Omega_i \to \Omega_v)}{\pi} |\cos\langle \Omega_i, \Omega_L \rangle| |\cos\langle \Omega_v, \Omega_L \rangle| d\Omega_L \end{aligned} \qquad \text{A 2}$$

If all vegetative facets are Lambertian and have the same diffuse reflectance ρ_d, since lidar instruments make observations at backscattering direction, i.e. $\Omega_v = \Omega_i - \pi = \Omega$, then we have,

$$\begin{aligned}\frac{1}{\pi}\Gamma(r,\Omega_i\to\Omega_v) &= \frac{\rho_d}{\pi}\int_{\Omega_1}\frac{g_L(r,\Omega_L)}{2\pi}\ |\cos\langle\Omega,\Omega_L\rangle|^2 d\Omega_L \\ &\approx \frac{\rho_d}{\pi}G(r,\Omega)^2\end{aligned} \qquad \text{A 3}$$

The two *G*-functions in the area scattering phase function correspond to the two cosine terms in the integral. One cosine term ($|cos\langle\Omega_i,\Omega_L\rangle|$) accounts for the projected area that intercepts incoming irradiance per unit vegetative facet area. The other cosine term ($|\cos\langle\Omega_v,\Omega_L\rangle|$) accounts for the decrease of radiant intensity ($\mathrm{w}\cdot\mathrm{sr}^{-1}$) with the cosine, i.e. projected area for a Lambertian surface. Also we assume the leaf angle distribution is constant within the canopy and so is *G*-function and area scattering function.

$$\frac{1}{\pi}\Gamma(\Omega_i\to\Omega_v)\approx\frac{\rho_d}{\pi}G(\Omega)^2 \qquad \text{A 4}$$

A.2 Error in apparent reflectance from two sources: range and return intensity

According to equation 2-35, at a given range r,

$$\begin{aligned}\alpha &= \frac{C_0K(r)}{r^b}\cdot\rho_{app} \\ \rho_{app} &= \frac{r^b}{C_0K(r)}\cdot\alpha\end{aligned} \qquad \text{A 5}$$

We omit the subscript j here for clarity. If the absolute errors in range r and return intensity α are Δr and $\Delta\alpha$ respectively

$$\Delta\rho_r = \frac{(r+\Delta r)^b}{C_0K(r+\Delta r)}\cdot\alpha-\rho_{app} \qquad \text{A 6}$$

$$\begin{aligned}\Delta\rho_\alpha &= \frac{r^b}{C_0 K(r)} \cdot (\alpha + \Delta\alpha) - \rho_{app} \\ &= \frac{r^b}{C_0 K(r)} \cdot \Delta\alpha = \frac{\rho_{app}}{\alpha} \cdot \Delta\alpha\end{aligned} \qquad \text{A 7}$$

where $\Delta\rho_r$ and $\Delta\rho_\alpha$ are errors in apparent reflectance estimate due to Δr and $\Delta\alpha$ separately. Notice $\Delta\rho_r$ and $\Delta\rho_\alpha$ both change with range. Now the total absolute error in apparent reflectance, $\Delta\rho$ due to $\Delta\rho_r$ and $\Delta\rho_\alpha$ together is

$$\Delta\rho = \frac{(r + \Delta r)^b}{C_0 K(r + \Delta r)} \cdot (\alpha + \Delta\alpha) - \rho_{app} \qquad \text{A 8}$$

Also according to equation A 6 and A 7, we have

$$\begin{aligned}\frac{(r + \Delta r)^b}{C_0 K(r + \Delta r)} &= \frac{\Delta\rho_r + \rho_{app}}{\alpha} \\ \Delta\alpha &= \alpha \cdot \frac{\Delta\rho_\alpha}{\rho_{app}}\end{aligned} \qquad \text{A 9}$$

Combining equation A 8 and A 9,

$$\begin{aligned}\Delta\rho &= \frac{\Delta\rho_r + \rho_{app}}{\alpha} \cdot \left(\alpha + \Delta\rho_\alpha \cdot \frac{\alpha}{\rho_{app}}\right) - \rho_{app} \\ &= \Delta\rho_r + \Delta\rho_\alpha + \frac{\Delta\rho_r \cdot \Delta\rho_\alpha}{\rho_{app}}\end{aligned} \qquad \text{A 10}$$

The relative errors in estimated apparent reflectance due to Δr and $\Delta\alpha$ separately are denoted as $\delta\rho_r = \Delta\rho_r / \rho_{app}$ and $\delta\rho_\alpha = \Delta\rho_\alpha / \rho_{app}$. Then the total relative error in apparent reflectance estimate due to Δr and $\Delta\alpha$ together is,

$$\delta\rho = \frac{\Delta\rho}{\rho_{app}}$$

$$= \frac{1}{\rho_{app}}\left(\delta\rho_r \cdot \rho_{app} + \delta\rho_\alpha \cdot \rho_{app} + \frac{\delta\rho_r \cdot \rho_{app} \cdot \delta\rho_\alpha \cdot \rho_{app}}{\rho_{app}}\right) \quad \text{A 11}$$

$$= \delta\rho_r + \delta\rho_\alpha + \delta\rho_r \cdot \delta\rho_\alpha$$

According to Figure 2-9, at near range $\delta\rho_r$ is large while $\delta\rho_\alpha$ is very small. At far range, it is the opposite. Thus $\delta\rho_r \cdot \delta\rho_\alpha$ is small for any given range and we can have,

$$\delta\rho \approx \delta\rho_r + \delta\rho_\alpha \quad \text{A 12}$$

A.3 Estimation of error matrix in number of points for a projection pixel

The reference labels of points in a projection pixel are unknown but the mode of the point reference labels is assumed to be same as the pixel reference label because the class of a pixel is determined from the mode of all the point classes in a pixel. For example, suppose one projection pixel has 11 points, 3 of which were classified as leaves and 8 were classified as woody materials. The reference label for this pixel is woody materials. First we enumerate all possible combinations of leaf and woody point counts in the reference labeling that result in the mode of point reference labels to be the pixel reference label (woody materials in the example here as shown by Figure A. 1(1)). Next, for one combination of leaf and woody point counts in the reference labeling, we enumerate all possible error matrices in number of points that result in the leaf and woody point counts in both reference and classification labeling. Figure A. 1(2) shows this enumeration of error matrices for one combination, 9 woody points and 2 leaf points in reference labeling, along with 8 woody points and 3 leaf points in classification. In this way, we enumerate all error matrices in number of points for all the combinations of leaf and woody point counts in reference labeling. The probability of each error matrix is the same. The average of these error matrices gives an estimate of error matrix in number of points for this projection pixel, given the pixel reference label and classification of points.

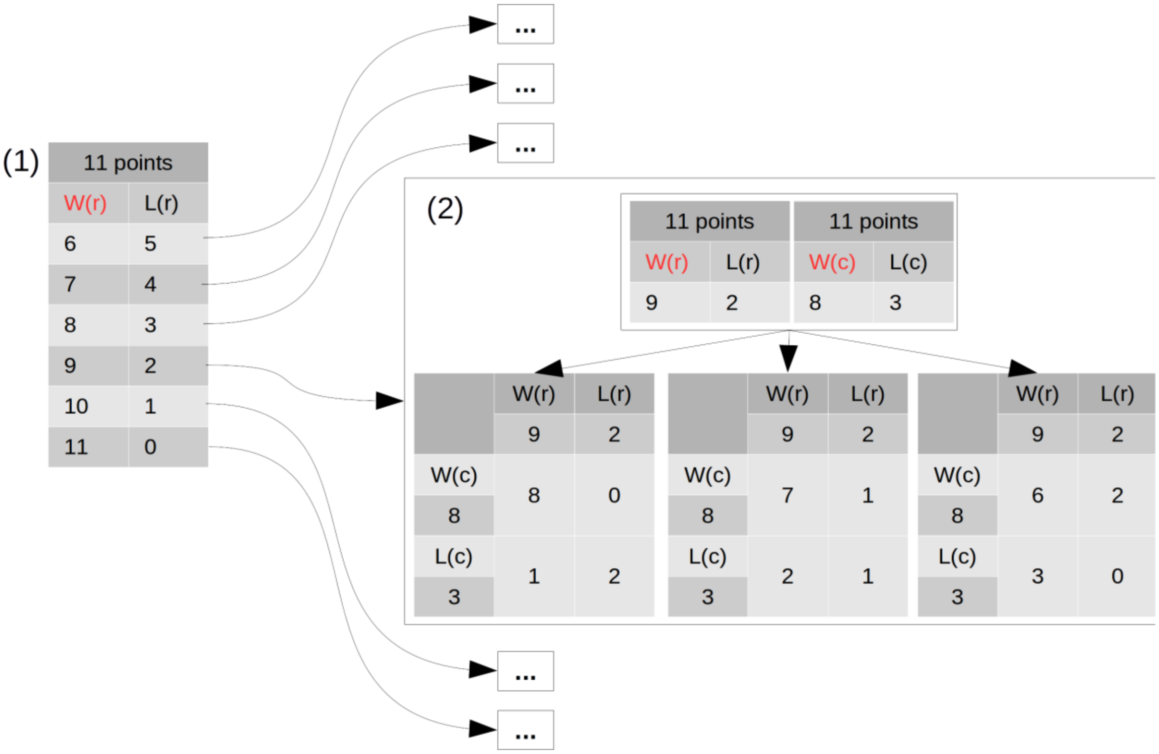


Figure A. 1. Diagram of error matrix estimation in number of points for a projection pixel. W(r): woody materials for reference label; L(r): leaves for reference label; W(c): woody materials for classification; L(c): leaves for classification. Numbers in the tables are point counts.

A.4 Resample of a point cloud classification based on classification accuracy

In a point cloud classification, the number of leaf points is N_c^L and the number of woody points is N_c^W. The classification accuracy assessment in number of points gives the user's accuracies for leaves and woody materials as $\widehat{U_L}$ and $\widehat{U_W}$. To generate a resample of a point cloud classification, we first randomly select $N_c^L \cdot \widehat{U_L}$ leaf points and $N_c^W \cdot \widehat{U_W}$ woody points as confidently classified points and remain the class labels of these point unchanged in the resample. For all the remaining unconfident points including both leaf and woody points, we randomly shuffle their classification labels but remain the proportion of points in each class. To be specific, we randomly select $N_c^L \cdot \left(1 - \widehat{U_L}\right)$ points from the unconfident points and assign them as leaf points while assigning the rest as woody points.

A.5 Results from SWIR data for estimation of leaf, woody and plant profiles

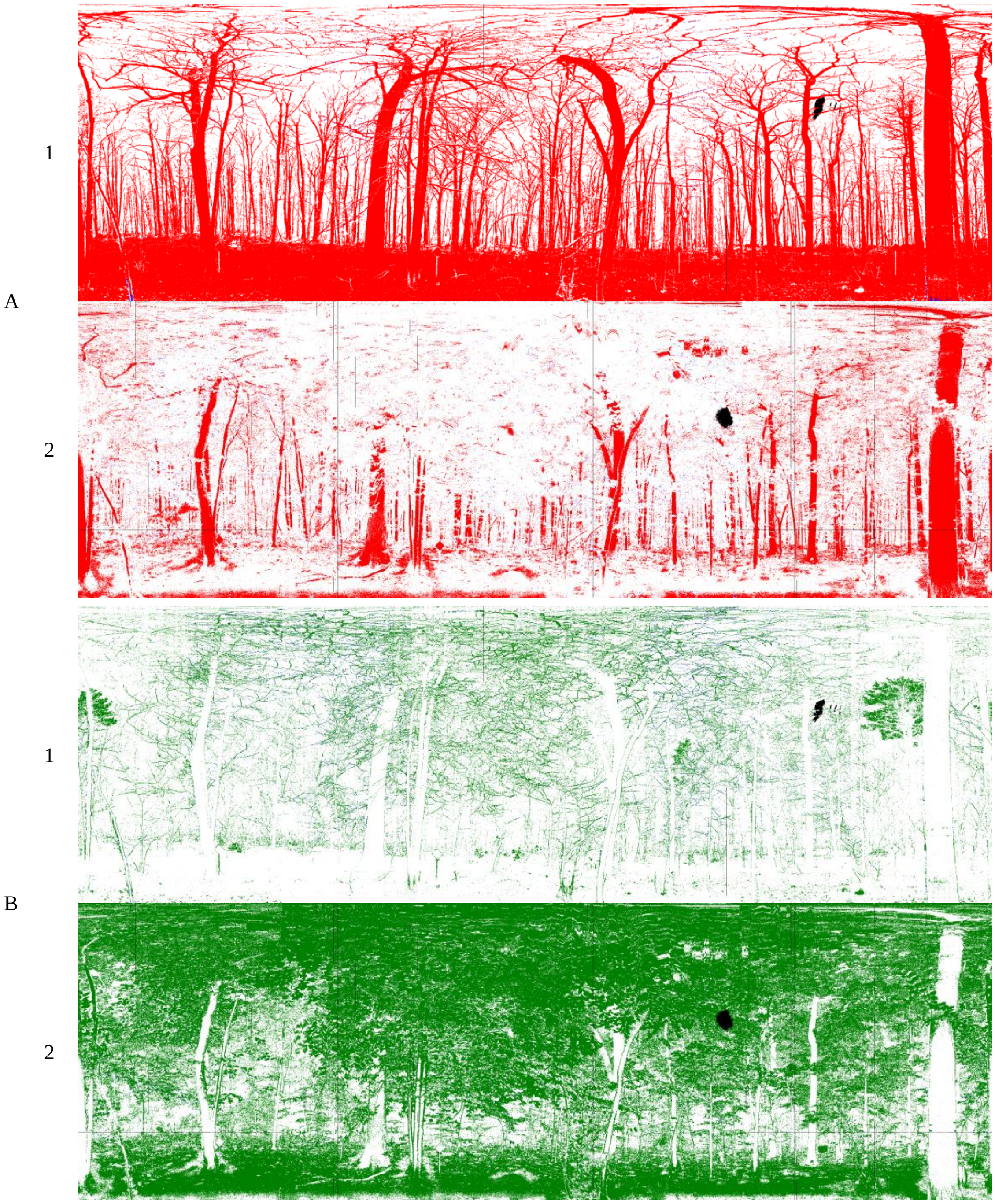
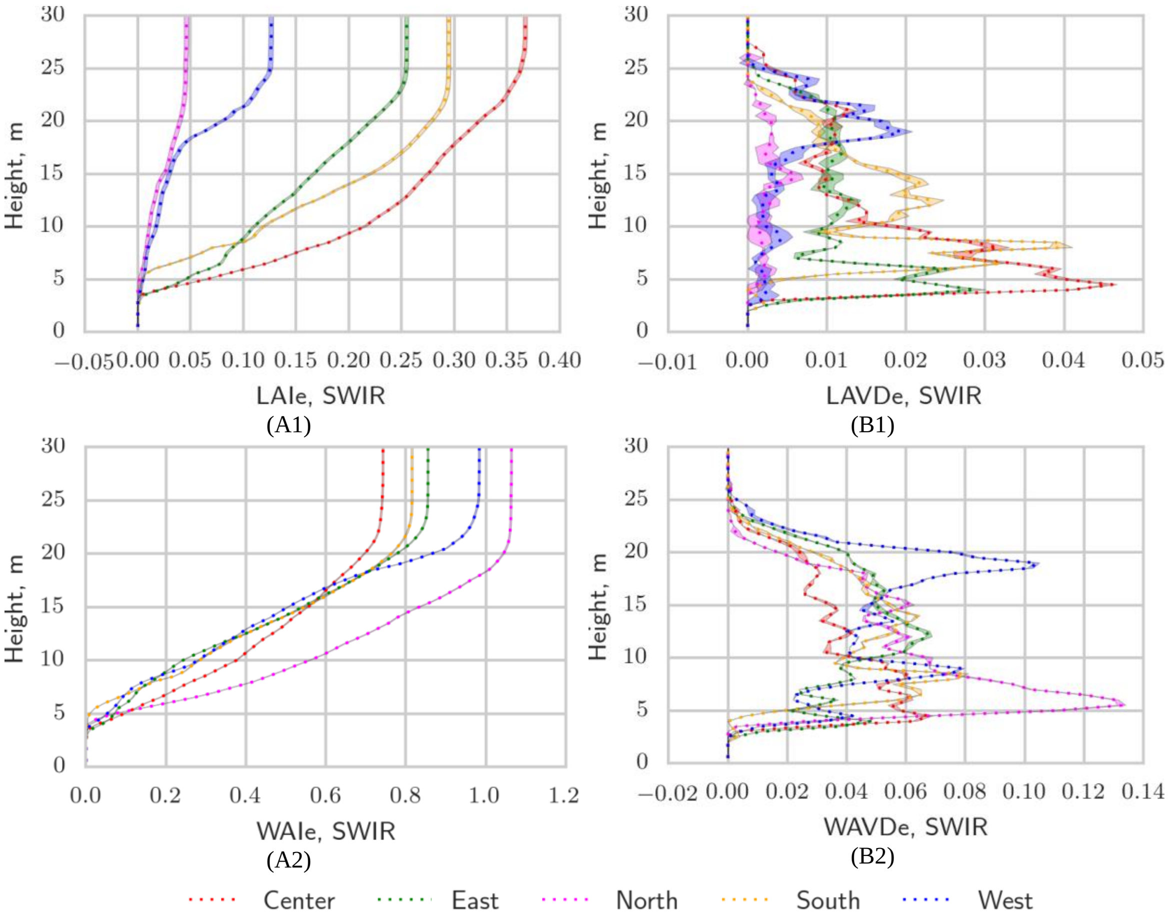


Figure A. 2. Gap probability at far range for woody materials (Part A) and leaves (Part B) separately from the scans at the center plot, in equal-angle Andrieu projection with X dimension as azimuth angles from 0° to 360° and Y dimension as zenith angles from 0° to 117°. Gap probability was calculated from SWIR data. Row

1: Leaf-off season. Row 2: Leaf-on season. Red: $P_{gap}^{W} = 0$, totally attenuated by woody materials. Green: $P_{gap}^{L} = 0$, totally attenuated by leaves. White: $P_{gap}^{W} = 1$ or $P_{gap}^{L} = 1$, no attenuation by woody materials in pair A or leaves in pair B. Blue (scattered, sparse): partial gap by woody materials in pair A or leaves in pair B. Black: No data (largely solar aureole).



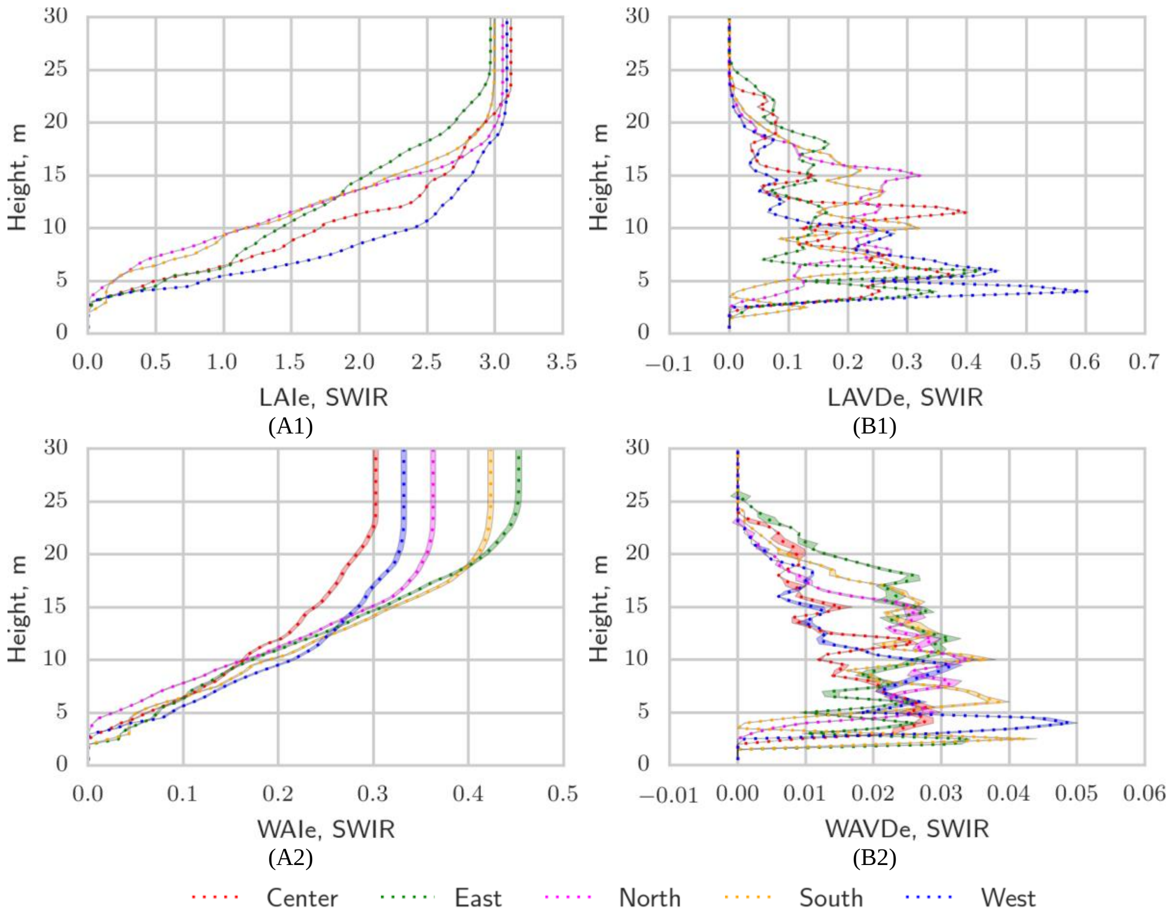
(A1)

(B1)

(A2)

(B2)

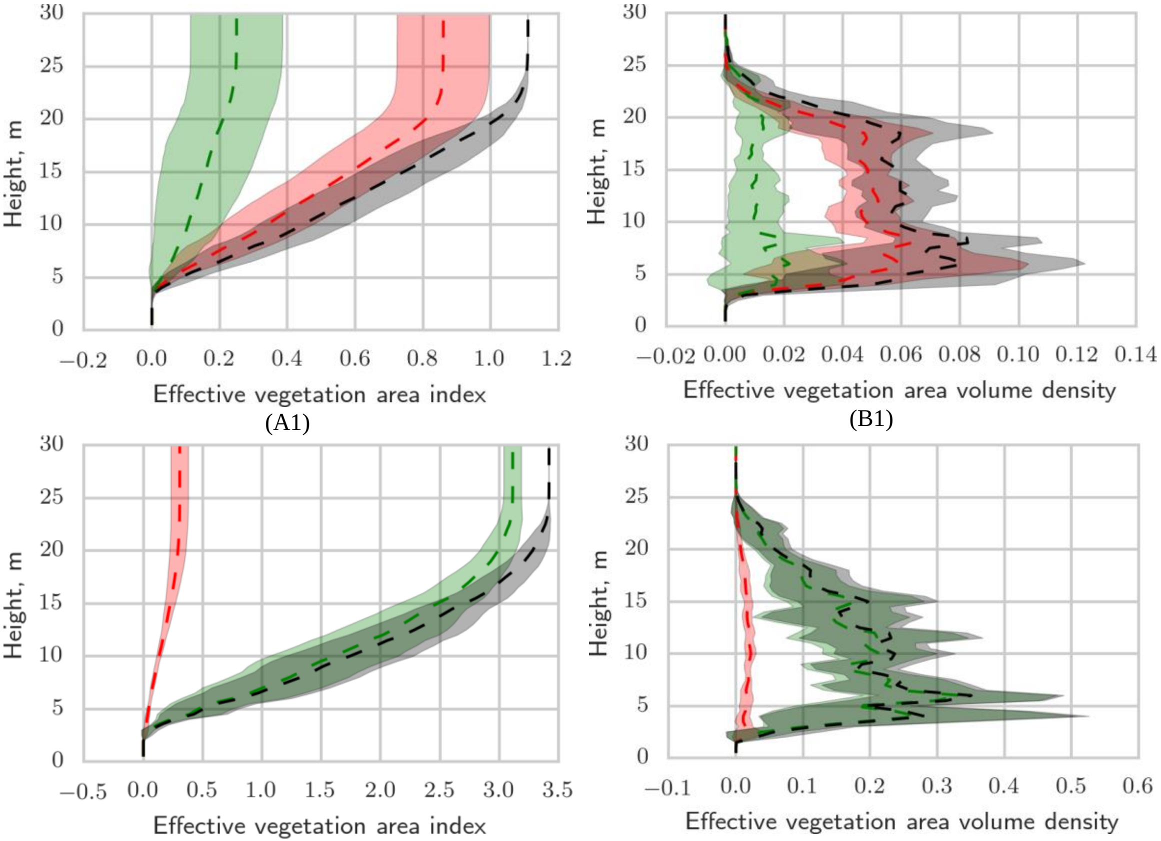
Figure A. 3. Variance in profiles due to classification error from each of the five leaf-off scans. Shaded area shows three standard deviations; for LAIe and WAIe and some of the LAVDe and WAVDe curves, shaded area is indistinguishable from graphic line width.



(A1) (B1)

(A2) (B2)

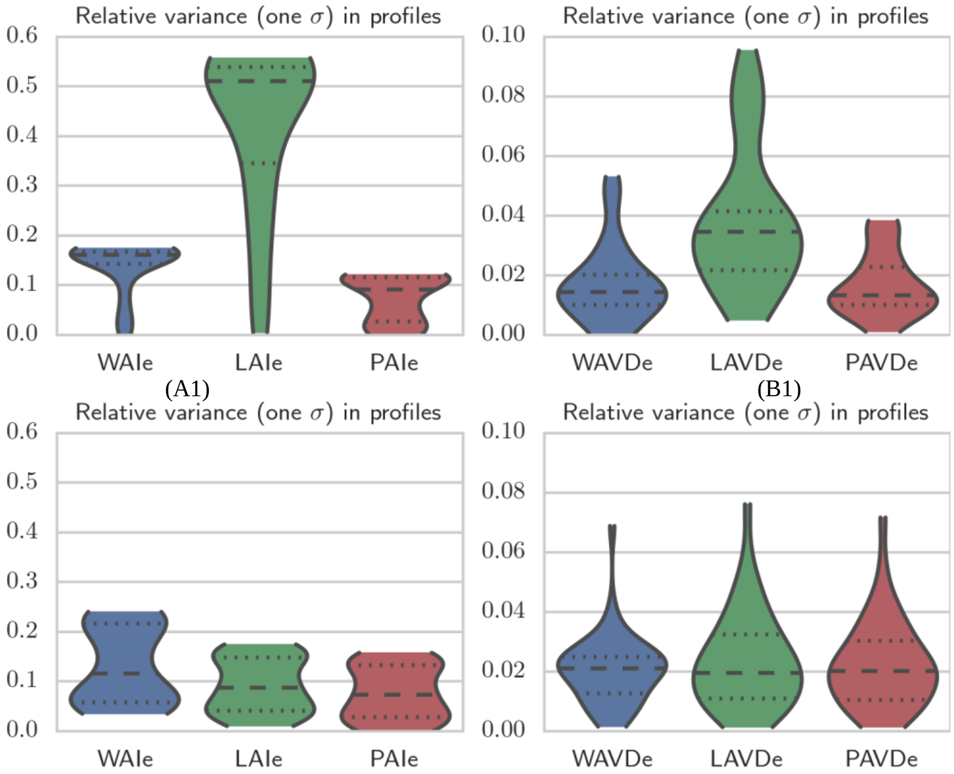
Figure A. 4. Variance in profiles due to classification error from each of the five leaf-on scans. Shaded area shows three standard deviations; for LAIe and WAIe and some of the LAVDe and WAVDe curves, shaded area is indistinguishable from graphic line width.



(A1) (B1)

(A2) (B2)

Figure A. 5. Variance in profiles due to heterogeneity of canopy structure across the study site, from SWIR data. First row: leaf-off season. Second row: leaf-on season. Shaded area shows one standard deviation.



(A1) (B1)

(A2) (B2)

Figure A. 6. Violin box plots of variance (one standard deviation) in vegetation profiles due to heterogeneity of the study site, normalized by average vegetation area indexes. First row: leaf-off season. Second row: leaf-on season. Left column: cumulative VAVD (VAI up to a canopy height). Right column: VAVD. The dash line is the median. The dotted lines are 25 and 75 percentiles. The ends of the violin boxes show the minimum and maximum values. The widths of violin boxes show the probability density of the data from kernel density estimate.

BIBLIOGRAPHY

Aber, J. D. (1979). Foliage-Height Profiles and Succession in Northern Hardwood Forests. *Ecology*, *60*(1), 18–23.

Armston, J., Disney, M. I., Lewis, P., Scarth, P., Phinn, S. R., Lucas, R. M., Bunting, P., & Goodwin, N. (2013). Direct Retrieval of Canopy Gap Probability Using Airborne Waveform Lidar. *Remote Sensing of Environment*, *134*(0), 24–38.

Asner, G. P., Scurlock, J. M. O., & A. Hicke, J. (2003). Global Synthesis of Leaf Area Index Observations: Implications for Ecological and Remote Sensing Studies. *Global Ecology and Biogeography*, *12*(3), 191–205.

Baccini, A., Goetz, S. J., Walker, W., Laporte, N., Sun, M., Sulla-Menashe, D., Hackler, J. L., Beck, P. S. A., Dubayah, R. O., Friedl, M., Samanta, S., & Houghton, R. A. (2012). Estimated Carbon Dioxide Emissions from Tropical Deforestation Improved by Carbon-Density Maps. *Nature Climate Change*, *2*(3), 182–185.

Baker, T. R., Phillips, O. L., Malhi, Y., Almeida, S., Arroyo, L., Di Fiore, A., Erwin, T. L., Killeen, T. J., Laurance, S. G., Laurance, W. F., Lewis, S. L., Lloyd, J., Monteagudo, A., Neill, D. A., Patino, S., Pitman, N. C. A., M. Silva, J. N., Vasquez Martinez, R., Patiño, S., Silva, J. N. M., & Martínez, R. V. (2004). Variation in Wood Density Determines Spatial Patterns in Amazonian Forest Biomass. *Global Change Biology*, *10*(5), 545–562.

Baldocchi, D. D., Wilson, K. B., & Gu, L. (2002). How the Environment, Canopy Structure and Canopy Physiological Functioning Influence Carbon, Water and Energy Fluxes of a Temperate Broad-Leaved Deciduous Forest--an Assessment with the Biophysical Model CANOAK. *Tree Physiology*, *22*(15-16), 1065–1077.

Basuki, T. M., van Laake, P. E., Skidmore, A., & Hussin, Y. A. (2009). Allometric Equations for Estimating the above-Ground Biomass in Tropical Lowland Dipterocarp Forests. *Forest Ecology and Management*, *257*(8), 1684–1694.

Béland, M., Baldocchi, D. D., Widlowski, J.-L., Fournier, R. A., & Verstraete, M. M. (2014). On Seeing the Wood from the Leaves and the Role of Voxel Size in Determining Leaf Area Distribution of Forests with Terrestrial LiDAR. *Agricultural and Forest Meteorology*, *184*(0), 82–97.

Béland, M., Widlowski, J.-L., Fournier, R. A., Côté, J.-F., & Verstraete, M. M. (2011). Estimating Leaf Area Distribution in Savanna Trees from Terrestrial LiDAR Measurements. *Agricultural and Forest Meteorology*, *151*(9), 1252–1266.

Besl, P. J., & McKay, N. D. (1992). A Method for Registration of 3-D Shapes. *Pattern Analysis and Machine Intelligence, IEEE Transactions on*, *14*(2), 239–256.

Bhandari, A., Hamre, B., Frette, Ø., Zhao, L., Stamnes, J. J., & Kildemo, M. (2011). Bidirectional Reflectance Distribution Function of Spectralon White Reflectance Standard Illuminated by Incoherent Unpolarized and Plane-Polarized Light. *Applied Optics*, *50*(16), 2431–42.

Blanchette, D., Fournier, R. A., Luther, J. E., & Côté, J.-F. (2015). Predicting Wood Fiber Attributes Using Local-Scale Metrics from Terrestrial LiDAR Data: A Case Study of Newfoundland Conifer Species. *Forest Ecology and Management*, *347*(0), 116–129.

Bonan, G. B. (2008). Forests and Climate Change: Forcings, Feedbacks, and the Climate Benefits of Forests. *Science*, *320*(5882), 1444–1449.

Bonan, G. B. (1993). Importance of Leaf Area Index and Forest Type When Estimating Photosynthesis in Boreal Forests. *Remote Sensing of Environment*, *43*(3), 303–314.

Bréda, N. J. (2003). Ground-based Measurements of Leaf Area Index: A Review of Methods, Instruments and Current Controversies. *Journal of Experimental Botany*, *54*(392), 2403–2417.

Briese, C., Pfennigbauer, M., Lehner, H., Ullrich, A., Wagner, W., & Pfeifer, N. (2012). Radiometric Calibration of Multi-Wavelength Airborne Laser Scanning Data. *ISPRS Annals of Photogrammetry, Remote Sensing and Spatial Information Sciences*, *I-7*, 335–340.

Brodu, N., & Lague, D. (2012). 3D Terrestrial Lidar Data Classification of Complex Natural Scenes Using a Multi-Scale Dimensionality Criterion: Applications in Geomorphology. *ISPRS Journal of Photogrammetry and Remote Sensing*, *68*(0), 121–134.

Brown, S. L., Schroeder, P., & Kern, J. S. (1999). Spatial Distribution of Biomass in Forests of the Eastern USA. *Forest Ecology and Management*, *123*(1), 81–90.

Burt, A., Disney, M. I., Raumonen, P., Armston, J., Calders, K., & Lewis, P. (2013). Rapid Characterisation of Forest Structure from TLS and 3D Modelling. In *Geoscience and Remote Sensing Symposium (IGARSS), 2013 IEEE International* (pp. 3387–3390). Melbourne.

Cairns, M. A., Brown, S. L., Helmer, E. H., & Baumgardner, G. A. (1997). Root Biomass Allocation in the World's Upland Forests. *Oecologia*, *111*(1), 1–11.

Calders, K., Armston, J., Newnham, G. J., Herold, M., & Goodwin, N. (2014). Implications of Sensor Configuration and Topography on Vertical Plant Profiles Derived from Terrestrial LiDAR. *Agricultural and Forest Meteorology*, *194*(0), 104–117.

Calders, K., Newnham, G. J., Burt, A., Murphy, S., Raumonen, P., Herold, M., Culvenor, D. S., Avitabile, V., Disney, M. I., Armston, J., & Kaasalainen, M. (2014). Nondestructive Estimates of above-Ground Biomass Using Terrestrial Laser Scanning. *Methods in Ecology and Evolution*, *6*(2), 198–208.

Calders, K., Schenkels, T., Bartholomeus, H., Armston, J., Verbesselt, J., & Herold, M. (2015). Monitoring Spring Phenology with High Temporal Resolution Terrestrial LiDAR Measurements. *Agricultural and Forest Meteorology*, *203*, 158–168.

Chauve, A., Mallet, C., Bretar, F., Durrieu, S., Pierrot Deseilligny, M., & Puech, W. (2008). Processing Full-Waveform Lidar Data: Modelling Raw Signals. In *International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 2007* (pp. 102–107). Espoo, Finland.

Chave, J., Andalo, C., Brown, S. L., Cairns, M. A., Chambers, J. Q., Eamus, D., Fölster, H., Fromard, F., Higuchi, N., & Kira, T. (2005). Tree Allometry and Improved Estimation of Carbon Stocks and Balance in Tropical Forests. *Oecologia*, *145*(1), 87–99.

Chave, J., Condit, R., Aguilar, S., Hernandez, A., Lao, S., & Perez, R. (2004). Error Propagation and Scaling for Tropical Forest Biomass Estimates. *Philosophical Transactions of the Royal Society of London. Series B: Biological Sciences*, *359*(1443), 409–420.

Chave, J., Coomes, D., Jansen, S., Lewis, S. L., Swenson, N. G., & Zanne, A. E. (2009). Towards a Worldwide Wood Economics Spectrum. *Ecology Letters*, *12*(4), 351–366.

Chave, J., Muller-Landau, H. C., Baker, T. R., Easdale, T. A., Steege, H. ter, & Webb, C. O. (2006). Regional and Phylogenetic Variation of Wood Density across 2456 Neotropical Tree Species. *Ecological Applications*, *16*(6), 2356–2367.

Chen, J. M., & Black, T. A. (1992). Defining Leaf Area Index for Non-Flat Leaves. *Plant, Cell and Environment*, *15*(4), 421–429.

Chen, J. M., & Cihlar, J. (1995). Plant Canopy Gap-Size Analysis Theory for Improving Optical Measurements of Leaf-Area Index. *Applied Optics*, *34*(27), 6211–6222.

Chen, J. M., Rich, P. M., Gower, S. T., Norman, J. M., & Plummer, S. (1997). Leaf Area Index of Boreal Forests: Theory, Techniques, and Measurements. *Journal of Geophysical Research: Atmospheres*, *102*(D24), 29429–29443.

Cohen, W. B., Maiersperger, T. K., Gower, S. T., & Turner, D. P. (2003). An Improved Strategy for Regression of Biophysical Variables and Landsat ETM+ Data. *Remote Sensing of Environment*, *84*(4), 561–571.

Danson, F. M., Gaulton, R., Armitage, R. P., Disney, M. I., Gunawan, O., Lewis, P., Pearson, G., & Ramirez, A. F. (2014). Developing a Dual-Wavelength Full-Waveform Terrestrial Laser Scanner to Characterize Forest Canopy Structure. *Agricultural and Forest Meteorology*, *198-199*, 7–14.

Danson, F. M., Hetherington, D., Morsdorf, F., Koetz, B., & Allgöwer, B. (2007). Forest Canopy Gap Fraction From Terrestrial Laser Scanning. *Geoscience and Remote Sensing Letters, IEEE*, *4*(1), 157–160.

Dassot, M., Constant, T., & Fournier, M. (2011). The Use of Terrestrial LiDAR Technology in Forest Science: Application Fields, Benefits and Challenges. *Annals of Forest Science*, *68*(5), 959–974.

Denman, K. L., Brasseur, G., Chidthaisong, A., Ciais, P., Cox, P. M., Dickinson, R. E., Hauglustaine, D., Heinze, C., Holland, E., Jacob, D., Lohmann, U., Ramachandran, S., Dias, P. L. da S., Wofsy, S. C., & Zhang, X. (2007). The Contemporary Carbon Budget. In S. Solomon et al. (Eds.), *Climate Change 2007: The Physical Science Basis. Contribution of Working Group I to the Fourth Assessment Report of the Intergovernmental Panel on Climate Change* (pp. 517–525). Cambridge, United Kingdom and New York, NY, USA.: Cambridge University Press.

Disney, M. I., Lewis, P., & Saich, P. (2006). 3D Modelling of Forest Canopy Structure for Remote Sensing Simulations in the Optical and Microwave Domains. *Remote Sensing of Environment*, *100*(1), 114–132.

Disney, M. I., Raumonen, P., & Lewis, P. (2012). Testing a New Vegetation Structure Retrieval Algorithm from Terrestrial Lidar Scanner Data Using 3D Models. In *SilviLaser 2012*. Vancouver.

Douglas, E. S., Martel, J., Li, Z., Howe, G., Hewawasam, K., Marshall, R., Schaaf, C. L., Cook, T., Newnham, G. J., Strahler, A. H., & Chakrabarti, S. (2015). Finding Leaves in the Forest: The Dual-Wavelength Echidna Lidar. *Geoscience and Remote Sensing Letters, IEEE*, *12*(4), 776–780.

Douglas, E. S., Strahler, A. H., Martel, J., Cook, T., Mendillo, C., Marshall, R., Chakrabarti, S., Schaaf, C. L., Woodcock, C. E., Li, Z., Yang, X., Culvenor, D. S., Jupp, D. L. B., Newnham, G. J., & Lovell, J. L. (2012). DWEL: A Dual-Wavelength Echidna Lidar for Ground-Based Forest Scanning. In *Geoscience and Remote Sensing Symposium (IGARSS), 2012 IEEE International* (pp. 4998–5001). Munich, Germany.

Drake, J. B., Dubayah, R. O., Clark, D. B., Knox, R. G., Blair, J. B., Hofton, M. a., Chazdon, R. L., Weishampel, J. F., & Prince, S. (2002). Estimation of Tropical Forest Structural Characteristics Using Large-Footprint Lidar. *Remote Sensing of Environment*, *79*(2-3), 305–319.

Dubayah, R. O., & Drake, J. B. (2000). Lidar Remote Sensing for Forestry. *Journal of Forestry*, *98*(6), 44–46.

Eitel, J. U. H., Magney, T. S., Vierling, L. A., & Dittmar, G. (2014). Assessment of Crop Foliar Nitrogen Using a Novel Dual-Wavelength Laser System and Implications for Conducting Laser-Based Plant Physiology. *ISPRS Journal of Photogrammetry and Remote Sensing*, *97*(0), 229–240.

Eitel, J. U. H., Vierling, L. A., & Long, D. S. (2010). Simultaneous Measurements of Plant Structure and Chlorophyll Content in Broadleaf Saplings with a Terrestrial Laser Scanner. *Remote Sensing of Environment*, *114*(10), 2229–2237.

Espinoza, J. (2004). Within-Tree Density Gradients in Gmelina Arborea in Venezuela. *New Forests*, *28*(2-3), 309–317.

Fischler, M. A., & Bolles, R. C. (1981). Random Sample Consensus: A Paradigm for Model Fitting with Applications to Image Analysis and Automated Cartography. *Communications of the ACM*, *24*(6), 381–395.

Gaulton, R., Danson, F. M., Ramírez, F., & Gunawan, O. (2013). The Potential of Dual-Wavelength Laser Scanning for Estimating Vegetation Moisture Content. *Remote Sensing of Environment*, *132*(0), 32–39.

Gibbs, H. K., Brown, S. L., Niles, J. O., & Foley, J. A. (2007). Monitoring and Estimating Tropical Forest Carbon Stocks: Making REDD a Reality. *Environmental Research Letters*, *2*(4), 045023.

Girardeau-Montaut, D. (2015). CloudCompare. Retrieved January 1, 2015, from http://www.danielgm.net/cc/

Gominho, J., Figueira, J., Rodrigues, J., & Pereira, H. (2001). Within-Tree Variation of Heartwood, Extractives and Wood Density in the Eucalypt Hybrid Urograndis (*Eucalyptus Grandis* × *E. Urophylla*). *Wood and Fiber Science*, *33*(1), 3–8.

Gong, W., Song, S., Zhu, B., Shi, S., Li, F., & Cheng, X. (2012). Multi-Wavelength Canopy LiDAR for Remote Sensing of Vegetation: Design and System Performance. *ISPRS Journal of Photogrammetry and Remote Sensing*, *69*, 1–9.

Hakala, T., Suomalainen, J., Kaasalainen, S., & Chen, Y. (2012). Full Waveform Hyperspectral LiDAR for Terrestrial Laser Scanning. *Optics Express*, *20*(7), 7119.

Halldórsson, T., & Langerholc, J. (1978). Geometrical Form Factors for the Lidar Function. *Applied Optics*, *17*(2), 240–244.

Hancock, S. (2010). *Understanding the Measurement of Forests with Waveform Lidar*. UCL (University College London).

Hancock, S., Essery, R., Reid, T., Carle, J., Baxter, R., Rutter, N., & Huntley, B. (2014). Characterising Forest Gap Fraction with Terrestrial Lidar and Photography: An Examination of Relative Limitations. *Agricultural and Forest Meteorology*, *189-190*, 105–114.

Hanson, P. J., Amthor, J. S., Wullschleger, S. D., Wilson, K. B., Grant, R. F., Hartley, A., Hui, D., Hunt E. R., J., Johnson, D. W., Kimball, J. S., King, A. W., Luo, Y., McNulty, S. G., Sun, G., Thornton, P. E., Wang, S., Williams, M., Baldocchi, D. D., & Cushman, R. M. (2004). Oak Forest Carbon and Water Simulations: Model Intercomparisons and Evaluations against Independent Data. *Ecological Monographs*, *74*(3), 443–489.

Harding, D. J., Lefsky, M. A., Parker, G. G., & Blair, J. B. (2001). Laser Altimeter Canopy Height Profiles: Methods and Validation for Closed-Canopy, Broadleaf Forests. *Remote Sensing of Environment*, *76*(3), 283–297.

Hartzell, P., Glennie, C., & Finnegan, D. C. (2015). Empirical Waveform Decomposition and Radiometric Calibration of a Terrestrial Full-Waveform Laser Scanner. *Geoscience and Remote Sensing, IEEE Transactions on*, *53*(1), 162–172.

Henry, M., Besnard, A., Asante, W. A., Eshun, J., Adu-Bredu, S., Valentini, R., Bernoux, M., & Saint-André, L. (2010). Wood Density, Phytomass Variations within and among Trees, and Allometric Equations in a Tropical Rainforest of Africa. *Forest Ecology and Management*, *260*(8), 1375–1388.

Hilker, T., Leeuwen, M., Coops, N. C., Wulder, M. A., Newnham, G. J., Jupp, D. L. B., & Culvenor, D. S. (2010). Comparing Canopy Metrics Derived from Terrestrial and Airborne Laser Scanning in a Douglas-Fir Dominated Forest Stand. *Trees*, *24*(5), 819–832.

Hill, T. C., Williams, M., Bloom, A. A., Mitchard, E. T. A., & Ryan, C. M. (2013). Are Inventory Based and Remotely Sensed Above-Ground Biomass Estimates Consistent? *PLoS ONE*, *8*(9), e74170.

Höfle, B., & Pfeifer, N. (2007). Correction of Laser Scanning Intensity Data: Data and Model-Driven Approaches. *ISPRS Journal of Photogrammetry and Remote Sensing*, *62*(6), 415–433.

Hosoi, F., Nakai, Y., & Omasa, K. (2013). 3-D Voxel-Based Solid Modeling of a Broad-Leaved Tree for Accurate Volume Estimation Using Portable Scanning Lidar. *ISPRS Journal of Photogrammetry and Remote Sensing*, *82*(0), 41–48.

Hosoi, F., & Omasa, K. (2006). Voxel-Based 3-D Modeling of Individual Trees for Estimating Leaf Area Density Using High-Resolution Portable Scanning Lidar. *Geoscience and Remote Sensing, IEEE Transactions on*, *44*(12), 3610–3618.

Houghton, R. A. (2007). Balancing the Global Carbon Budget. *Annual Review of Earth and Planetary Sciences*, *35*(1), 313–347.

Houghton, R. A., Hall, F., & Goetz, S. J. (2009). Importance of Biomass in the Global Carbon Cycle. *Journal of Geophysical Research: Biogeosciences*, *114*(G2), n/a–n/a.

Houghton, R. A., Lawrence, K. T., Hackler, J. L., & Brown, S. L. (2001). The Spatial Distribution of Forest Biomass in the Brazilian Amazon: A Comparison of Estimates. *Global Change Biology*, *7*(7), 731–746.

Huang, H., Li, Z., Gong, P., Cheng, X., Clinton, N., Cao, C., Ni, W., & Wang, L. (2011). Automated Methods for Measuring DBH and Tree Heights with a Commercial Scanning Lidar. *Photogrammetric Engineering & Remote Sensing*, *77*(3), 219–227.

Ioannou, D., Huda, W., & Laine, A. F. (1999). Circle Recognition through a 2D Hough Transform and Radius Histogramming. *Image and Vision Computing*, *17*(1), 15–26.

Jenkins, J. C., Chojnacky, D. C., Heath, L. S., & Birdsey, R. A. (2004). *Comprehensive Database of Diameter-Based Biomass Regressions for North American Tree Species*. USDA Forest Service.

Jonckheere, I., Fleck, S., Nackaerts, K., Muys, B., Coppin, P., Weiss, M., & Baret, F. (2004). Review of Methods for in Situ Leaf Area Index Determination: Part I. Theories, Sensors and Hemispherical Photography. *Agricultural and Forest Meteorology*, *121*(1–2), 19–35.

Jupp, D. L. B., Culvenor, D., Lovell, J., & Newnham, G. (2005). *Evaluation and Validation Laser Radar Systems for Native and Plantation Forest Inventory*. Canberra: CSIRO.

Jupp, D. L. B., Culvenor, D. S., Lovell, J. L., Newnham, G. J., Strahler, A. H., & Woodcock, C. E. (2009). Estimating Forest LAI Profiles and Structural Parameters Using a Ground-Based Laser Called “Echidna®.” *Tree Physiology*, *29*(2), 171–181.

Jutzi, B., Eberle, B., & Stilla, U. (2003). Estimation and Measurement of Backscattered Signals from Pulsed Laser Radar. In *Proceedings of SPIE* (Vol. 4885, pp. 256–267).

Kaasalainen, S., Hyyppä, H., Kukko, A., Litkey, P., Ahokas, E., Hyyppä, J., Lehner, H., Jaakkola, A., Suomalainen, J., Akujarvi, A., Kaasalainen, M., & Pyysalo, U. (2009). Radiometric Calibration of LIDAR Intensity With Commercially Available Reference Targets. *Geoscience and Remote Sensing, IEEE Transactions on*, *47*(2), 588–598.

Kaasalainen, S., Krooks, A., Kukko, A., & Kaartinen, H. (2009). Radiometric Calibration of Terrestrial Laser Scanners with External Reference Targets. *Remote Sensing*, *1*(3), 144–158.

Kaasalainen, S., Krooks, A., Liski, J., Raumonen, P., Kaartinen, H., Kaasalainen, M., Puttonen, E., Anttila, K., & Mäkipää, R. (2014). Change Detection of Tree Biomass with Terrestrial Laser Scanning and Quantitative Structure Modelling. *Remote Sensing*, *6*(5), 3906–3922.

Kaasalainen, S., Lindroos, T., & Hyyppä, J. (2007). Toward Hyperspectral Lidar: Measurement of Spectral Backscatter Intensity With a Supercontinuum Laser Source. *Geoscience and Remote Sensing Letters, IEEE*, *4*(2), 211–215.

Kaasalainen, S., Pyysalo, U., Krooks, A., Vain, A., Kukko, A., Hyyppä, J., & Kaasalainen, M. (2011). Absolute Radiometric Calibration of ALS Intensity Data: Effects on Accuracy and Target Classification. *Sensors*, *11*(11), 10586–10602.

Keller, M., Palace, M., & Hurtt, G. C. (2001). Biomass Estimation in the Tapajos National Forest, Brazil: Examination of Sampling and Allometric Uncertainties. *Forest Ecology and Management*, *154*(3), 371–382.

Knapic, S., Louzada, J. L., Leal, S., & Pereira, H. (2008). Within-Tree and between-Tree Variation of Wood Density Components in Cork Oak Trees in Two Sites in Portugal. *Forestry*, *81*(4), 465–473.

Kobayashi, H., Baldocchi, D. D., Ryu, Y., Chen, Q., Ma, S., Osuna, J. L., & Ustin, S. L. (2012). Modeling Energy and Carbon Fluxes in a Heterogeneous Oak Woodland: A

Three-Dimensional Approach. *Agricultural and Forest Meteorology*, *152*(0), 83–100.

Koetz, B., Baret, F., Poilvé, H., & Hill, J. (2005). Use of Coupled Canopy Structure Dynamic and Radiative Transfer Models to Estimate Biophysical Canopy Characteristics. *Remote Sensing of Environment*, *95*(1), 115–124.

Korhonen, L., Korpela, I., Heiskanen, J., & Maltamo, M. (2011). Airborne Discrete-Return LIDAR Data in the Estimation of Vertical Canopy Cover, Angular Canopy Closure and Leaf Area Index. *Remote Sensing of Environment*, *115*(4), 1065–1080.

Kucharik, C. J., Norman, J. M., & Gower, S. T. (1998). Measurements of Branch Area and Adjusting Leaf Area Index Indirect Measurements. *Agricultural and Forest Meteorology*, *91*(1-2), 69–88.

Lang, A. R. G. (1987). Simplified Estimate of Leaf Area Index from Transmittance of the Sun's Beam. *Agricultural and Forest Meteorology*, *41*(3-4), 179–186.

Law, B. E., Cescatti, A., & Baldocchi, D. D. (2001, August). Leaf Area Distribution and Radiative Transfer in Open-Canopy Forests: Implications for Mass and Energy Exchange. *Tree Physiology*.

Lefsky, M. A., Cohen, W. B., Acker, S. A., Parker, G. G., Spies, T. A., & Harding, D. J. (1999). Lidar Remote Sensing of the Canopy Structure and Biophysical Properties of Douglas-Fir Western Hemlock Forests. *Remote Sensing of Environment*, *70*(3), 339–361.

Lefsky, M. A., Cohen, W. B., Parker, G. G., & Harding, D. J. (2002). Lidar Remote Sensing for Ecosystem Studies. *Bioscience*, *52*(1), 19–30.

Lefsky, M. A., & McHale, M. R. (2008). Volume Estimates of Trees with Complex Architecture from Terrestrial Laser Scanning. *Journal of Applied Remote Sensing*, *2*(1), 23519–23521.

Lim, K., Treitz, P., Wulder, M. A., St-Onge, B., & Flood, M. (2003). LiDAR Remote Sensing of Forest Structure. *Progress in Physical Geography*, *27*(1), 88–106.

Lovell, J. L., Jupp, D. L. B., Newnham, G. J., & Culvenor, D. S. (2011). Measuring Tree Stem Diameters Using Intensity Profiles from Ground-Based Scanning Lidar from a Fixed Viewpoint. *ISPRS Journal of Photogrammetry and Remote Sensing*, *66*(1), 46–55.

Luo, S., Wang, C., Li, G., & Xi, X. (2013). Retrieving Leaf Area Index Using ICESat/GLAS Full-Waveform Data. *Remote Sensing Letters*, *4*(8), 745–753.

Ma, H., Song, J., Wang, J., Xiao, Z., & Fu, Z. (2014). Improvement of Spatially Continuous Forest LAI Retrieval by Integration of Discrete Airborne LiDAR and Remote Sensing Multi-Angle Optical Data. *Agricultural and Forest Meteorology*, *189–190*(0), 60–70.

MacArthur, R. H., & Horn, H. S. (1969). Foliage Profile by Vertical Measurements. *Ecology*, *50*(5), 802–804.

Magney, T. S., Eusden, S. A., Eitel, J. U. H., Logan, B. A., Jiang, J., & Vierling, L. A. (2013). Assessing Leaf Photoprotective Mechanisms Using Terrestrial LiDAR: Towards Mapping Canopy Photosynthetic Performance in Three Dimensions. *New Phytologist*, *201*(1), 344–356.

Malenovský, Z., Martin, E., Homolová, L., Gastellu-Etchegorry, J. P., Zurita-Milla, R., Schaepman, M. E., Pokorný, R., Clevers, J. G. P. W., & Cudlín, P. (2008). Influence of Woody Elements of a Norway Spruce Canopy on Nadir Reflectance Simulated by the DART Model at Very High Spatial Resolution. *Remote Sensing of Environment*, *112*(1), 1–18.

Manninen, T., Stenberg, P., Rautiainen, M., Voipio, P., & Smolander, H. (2005). Leaf Area Index Estimation of Boreal Forest Using ENVISAT ASAR. *Geoscience and Remote Sensing, IEEE Transactions on*, *43*(11), 2627–2635.

MATLAB. (2015a). Fminsearch Algorithm. Retrieved from http://www.mathworks.com/help/matlab/math/optimizing-nonlinear-functions.html#bsgpq6p-11

MATLAB. (2015b). Genetic Algorithm. Retrieved from http://www.mathworks.com/discovery/genetic-algorithm.html

McHale, M. R., Burke, I. C., Lefsky, M. A., Peper, P. J., & McPherson, E. G. (2009). Urban Forest Biomass Estimates: Is It Important to Use Allometric Relationships Developed Specifically for Urban Trees? *Urban Ecosystems*, *12*(1), 95–113.

Measures, R. M. (1991). Laser-Remote-Sensor Equations. In *Laser Remote Sensing: Fundamentals and Applications* (pp. 237–243). Krieger Publishing Company.

Medvigy, D., Wofsy, S. C., Munger, J. W., Hollinger, D. Y., & Moorcroft, P. R. (2009). Mechanistic Scaling of Ecosystem Function and Dynamics in Space and Time: Ecosystem Demography Model Version 2. *Journal of Geophysical Research: Biogeosciences*, *114*(G1), n/a–n/a.

Miller, J. R. (1967). A Formula for Average Foliage Density. *Australian Journal of Botany*, *15*(1), 141–144.

Moorthy, I., Miller, J. R., Hu, B., Chen, J. M., & Li, Q. (2008). Retrieving Crown Leaf Area Index from an Individual Tree Using Ground-Based Lidar Data. *Canadian Journal of Remote Sensing*, *34*(3), 320–332.

Morsdorf, F., Nichol, C., Malthus, T., & Woodhouse, I. H. (2009). Assessing Forest Structural and Physiological Information Content of Multi-Spectral LiDAR Waveforms by Radiative Transfer Modelling. *Remote Sensing of Environment*, *113*(10), 2152–2163.

Murphy, G. (2008). Determining Stand Value and Log Product Yields Using Terrestrial Lidar and Optimal Bucking: A Case Study. *Journal of Forestry*, *106*(6), 317–324.

Myneni, R. B., Hoffman, S., Knyazikhin, Y., Privette, J. L., Glassy, J., Tian, Y., Wang, Y., Song, X., Zhang, Y., Smith, G. ., Lotsch, A., Friedl, M., Morisette, J. T., Votava, P., Nemani, R. R., & Running, S. W. (2002). Global Products of Vegetation Leaf Area and Fraction Absorbed PAR from Year One of MODIS Data. *Remote Sensing of Environment*, *83*(1-2), 214–231.

Ni-Meister, W., Jupp, D. L. B., & Dubayah, R. O. (2001). Modeling Lidar Waveforms in Heterogeneous and Discrete Canopies. *Geoscience and Remote Sensing, IEEE Transactions on*, *39*(9), 1943–1958.

Olofsson, P., Foody, G. M., Herold, M., Stehman, S. V., Woodcock, C. E., & Wulder, M. A. (2014). Good Practices for Estimating Area and Assessing Accuracy of Land Change. *Remote Sensing of Environment*, *148*, 42–57.

Olofsson, P., Foody, G. M., Stehman, S. V., & Woodcock, C. E. (2013). Making Better Use of Accuracy Data in Land Change Studies: Estimating Accuracy and Area and Quantifying Uncertainty Using Stratified Estimation. *Remote Sensing of Environment*, *129*(0), 122–131.

Parker, G. G., Harmon, M. E., Lefsky, M. A., Chen, J., Pelt, R. Van, Weiss, S. B., Thomas, S. C., Winner, W. E., Shaw, D. C., & Franklin, J. F. (2004). Three-Dimensional Structure of an Old-Growth Pseudotsuga-Tsuga Canopy and Its Implications for Radiation Balance, Microclimate, and Gas Exchange. *Ecosystems*, *7*(5), 440–453.

Parkin, D. A., Jupp, D. L. B., Poropat, G. V, & Lovell, J. L. (2001). Lidar System and Method. US 7,187,452; Australia 2002227768; New Zealand 527547; China ZL 02 8 07968.X; Japan 4108478; Hong Kong 04108479.9; Canada 2,437,897 and EU 1358508. EU patent validated in UK, France, Germany, Austria, Sweden and Finland.

Pedregosa, F., Varoquaux, G., Gramfort, A., Michel, V., Thirion, B., Grisel, O., Blondel, M., Prettenhofer, P., Weiss, R., Dubourg, V., Vanderplas, J., Passos, A., Cournapeau, D., Brucher, M., Perrot, M., & Duchesnay, É. (2011). Scikit-Learn: Machine Learning in Python. *Journal of Machine Learning Research*, *12*, 2825–2830.

Pfeifer, N., Höfle, B., Briese, C., Rutzinger, M., & Haring, A. (2008). Analysis of the Backscattered Energy in Terrestrial Laser Scanning Data. In *Proc. in the XXIth ISPRS Congress, Silk Road for Information from Imagery* (Vol. 37, p. B5).

Pfennigbauer, M., & Ullrich, A. (2010). Improving Quality of Laser Scanning Data Acquisition through Calibrated Amplitude and Pulse Deviation Measurement. In *Proceedings of SPIE* (Vol. 7684, p. 76841F–76841F–10).

Pfennigbauer, M., Wolf, C., & Ullrich, A. (2013). Enhancing Online Waveform Processing by Adding New Point Attributes. In *Proceedings of SPIE* (Vol. 8731, pp. 873104–873109).

Picard, N., Saint-André, L., & Henry, M. (2012). *Manual for Building Tree Volume and Biomass Allometric Equations*. FAO Forestry Department.

Pueschel, P. (2013). The Influence of Scanner Parameters on the Extraction of Tree Metrics from FARO Photon 120 Terrestrial Laser Scans. *ISPRS Journal of Photogrammetry and Remote Sensing*, *78*, 58–68.

Puttonen, E., Hakala, T., Nevalainen, O., Kaasalainen, S., Krooks, A., Karjalainen, M., & Anttila, K. (2015). Artificial Target Detection with a Hyperspectral LiDAR over 26-H Measurement. *Optical Engineering*, *54*(1), 13105.

Ramírez, F., Armitage, R. P., & Danson, F. M. (2013). Testing the Application of Terrestrial Laser Scanning to Measure Forest Canopy Gap Fraction. *Remote Sensing*, *5*(6), 3037–3056.

Raumonen, P., Kaasalainen, M., Åkerblom, M., Kaasalainen, S., Kaartinen, H., Vastaranta, M., Holopainen, M., Disney, M. I., & Lewis, P. (2013). Fast Automatic Precision Tree Models from Terrestrial Laser Scanner Data. *Remote Sensing*, *5*(2), 491–520.

Reitberger, J., Krzystek, P., & Stilla, U. (2008). Analysis of Full Waveform LIDAR Data for the Classification of Deciduous and Coniferous Trees. *International Journal of Remote Sensing*, *29*(5), 1407–1431.

Reulke, R., & Haala, N. (2005). Tree Species Recognition with Fuzzy Texture Parameters. In R. Klette & J. Žunić (Eds.), *Combinatorial Image Analysis SE - 45* (Vol. 3322, pp. 607–620). Springer Berlin Heidelberg.

Richards, F. J. (1959). A Flexible Growth Function for Empirical Use. *Journal of Experimental Botany*, *10*(2), 290–301.

Roncat, A., Briese, C., Jansa, J., & Pfeifer, N. (2014). Radiometrically Calibrated Features of Full-Waveform Lidar Point Clouds Based on Statistical Moments. *Ieee Geoscience and Remote Sensing Letters*, *11*(2), 549–553.

Roncat, A., Morsdorf, F., Briese, C., Wagner, W., & Pfeifer, N. (2014). Laser Pulse Interaction with Forest Canopy: Geometric and Radiometric Issues. *Forestry Applications of Airborne Laser Scanning: Concepts and Case Studies*, *27*, 19–41.

Ross, J. (1981). *The Radiation Regime and Architecture of Plant Stands.* (H. Lieth, Ed.)*The radiation regime and architecture of plant stands.* The Hague: Dr W. Junk Publishers.

Running, S. W., & Coughlan, J. C. (1988). A General Model of Forest Ecosystem Processes for Regional Applications I. Hydrologic Balance, Canopy Gas Exchange and Primary Production Processes. *Ecological Modelling*, *42*(2), 125–154.

Saatchi, S. S., Harris, N. L., Brown, S. L., Lefsky, M. A., Mitchard, E. T. A., Salas, W., Zutta, B. R., Buermann, W., Lewis, S. L., Hagen, S., Petrova, S., White, L., Silman, M. R., & Morel, A. (2011). Benchmark Map of Forest Carbon Stocks in Tropical Regions across Three Continents. *Proceedings of the National Academy of Sciences*, *108*(24), 9899–9904.

Saatchi, S. S., Houghton, R. A., DOS SANTOS ALVALÁ, R. C., SOARES, J. V, & YU, Y. (2007). Distribution of Aboveground Live Biomass in the Amazon Basin. *Global Change Biology*, *13*(4), 816–837.

Sales, M. H., Souza Jr., C. M., Kyriakidis, P. C., Roberts, D. A., & Vidal, E. (2007). Improving Spatial Distribution Estimation of Forest Biomass with Geostatistics: A Case Study for Rondônia, Brazil. *Ecological Modelling*, *205*(1–2), 221–230.

Sellers, P. J., Dickinson, R. E., Randall, D. A., Betts, A. K., Hall, F., Berry, J. A., Collatz, G. J., Denning, A. S., Mooney, H. A., Nobre, C. A., Sato, N., Field, C. B., & Henderson-Sellers, A. (1997). Modeling the Exchanges of Energy, Water, and Carbon Between Continents and the Atmosphere. *Science*, *275*(5299), 502–509.

Solberg, S., Brunner, A., Hanssen, K. H., Lange, H., Næsset, E., Rautiainen, M., & Stenberg, P. (2009). Mapping LAI in a Norway Spruce Forest Using Airborne Laser Scanning. *Remote Sensing of Environment*, *113*(11), 2317–2327.

Stark, S. C., Leitold, V., Wu, J. L., Hunter, M. O., de Castilho, C. V, Costa, F. R. C., McMahon, S. M., Parker, G. G., Shimabukuro, M. T., Lefsky, M. A., Keller, M., Alves, L. F., Schietti, J., Shimabukuro, Y. E., Brandão, D. O., Woodcock, T. K., Higuchi, N., de Camargo, P. B., de Oliveira, R. C., & Saleska, S. R. (2012). Amazon Forest Carbon Dynamics Predicted by Profiles of Canopy Leaf Area and Light Environment. *Ecology Letters*, *15*(12), 1406–1414.

Strahler, A. H., Jupp, D. L. B., Woodcock, C. E., Schaaf, C. L., Yao, T., Zhao, F., Yang, X., Lovell, J. L., Culvenor, D. S., Newnham, G. J., Ni-Meister, W., & Boykin-Morris, W. (2008). Retrieval of Forest Structural Parameters Using a Ground-Based Lidar Instrument (EchidnaÂ®). *Canadian Journal of Remote Sensing*, *34*, S426–S440.

Sun, G., & Ranson, K. J. (2000). Modeling Lidar Returns from Forest Canopies. *Geoscience and Remote Sensing, IEEE Transactions on*, *38*(6), 2617–2626.

Tan, S., & Narayanan, R. M. (2004). Design and Performance of a Multiwavelength Airborne Polarimetric Lidar for Vegetation Remote Sensing. *Applied Optics*, *43*(11), 2360–2368.

Tang, H., Dubayah, R. O., Brolly, M., Ganguly, S., & Zhang, G. (2014). Large-Scale Retrieval of Leaf Area Index and Vertical Foliage Profile from the Spaceborne Waveform Lidar (GLAS/ICESat). *Remote Sensing of Environment*, *154*(0), 8–18.

Tang, H., Dubayah, R. O., Swatantran, A., Hofton, M. a., Sheldon, S., Clark, D. B., & Blair, B. (2012). Retrieval of Vertical LAI Profiles over Tropical Rain Forests Using Waveform Lidar at La Selva, Costa Rica. *Remote Sensing of Environment*, *124*(0), 242–250.

Thies, M., Pfeifer, N., Winterhalder, D., & Gorte, B. G. H. (2004). Three-Dimensional Reconstruction of Stems for Assessment of Taper, Sweep and Lean Based on Laser Scanning of Standing Trees. *Scandinavian Journal of Forest Research*, *19*(6), 571–581.

Thies, M., & Spiecker, H. (2004). Evaluation and Future Prospects of Terrestrial Laser Scanning for Standardized Forest Inventories. *Forest*, *2*(2.2), 192–197.

Turner, D. P., Ritts, W. D., Cohen, W. B., Maeirsperger, T. K., Gower, S. T., Kirschbaum, A. A., Running, S. W., Zhao, M., Wofsy, S. C., Dunn, A. L., Law, B. E., Campbell, J. L., Oechel, W. C., Kwon, H. J., Meyers, T. P., Small, E. E., Kurc, S. A., & Gamon, J. A. (2005). Site-Level Evaluation of Satellite-Based Global Terrestrial Gross Primary Production and Net Primary Production Monitoring. *Global Change Biology*, *11*(4), 666–684.

Verrelst, J., Schaepman, M. E., Malenovský, Z., & Clevers, J. G. P. W. (2010). Effects of Woody Elements on Simulated Canopy Reflectance: Implications for Forest Chlorophyll Content Retrieval. *Remote Sensing of Environment*, *114*(3), 647–656.

Vonderach, C., Vögtle, T., Adler, P., & Norra, S. (2012). Terrestrial Laser Scanning for Estimating Urban Tree Volume and Carbon Content. *International Journal of Remote Sensing*, *33*(21), 6652–6667.

Wagner, W. (2010). Radiometric Calibration of Small-Footprint Full-Waveform Airborne Laser Scanner Measurements: Basic Physical Concepts. *ISPRS Journal of Photogrammetry and Remote Sensing*, *65*(6), 505–513.

Wagner, W., Ullrich, A., Ducic, V., Melzer, T., & Studnicka, N. (2006). Gaussian Decomposition and Calibration of a Novel Small-Footprint Full-Waveform Digitising Airborne Laser Scanner. *ISPRS Journal of Photogrammetry and Remote Sensing*, *60*(2), 100–112.

Wang, C. (2006). Biomass Allometric Equations for 10 Co-Occurring Tree Species in Chinese Temperate Forests. *Forest Ecology and Management*, *222*(1-3), 9–16.

Weiss, M., & Baret, F. (2013). CAN-EYE. Retrieved June 15, 2015, from http://www6.paca.inra.fr/can-eye

Weiss, M., Baret, F., Smith, G. J., Jonckheere, I., & Coppin, P. (2004). Review of Methods for in Situ Leaf Area Index (LAI) Determination: Part II. Estimation of LAI, Errors and Sampling. *Agricultural and Forest Meteorology*, *121*(1–2), 37–53.

Weiss, S. B. (2000). Vertical and Temporal Distribution of Insolation in Gaps in an Old-Growth Coniferous Forest. *Canadian Journal of Forest Research*, *30*(12), 1953–1964.

Welles, J. M., & Cohen, S. (1996). Canopy Structure Measurement by Gap Fraction Analysis Using Commercial Instrumentation. *Journal of Experimental Botany*, *47*(9), 1335–1342.

Wilson, J. W. (1960). Inclined Point Quadrats. *New Phytologist*, *59*(1), 1–7.

Woodcock, D., & Shier, A. (2002). Wood Specific Gravity and Its Radial Variations: The Many Ways to Make a Tree. *Trees*, *16*(6), 437–443.

Woodhouse, I. H., Nichol, C., Sinclair, P., Jack, J., Morsdorf, F., Malthus, T., & Patenaude, G. (2011). A Multispectral Canopy LiDAR Demonstrator Project. *Geoscience and Remote Sensing Letters, IEEE*, *8*(5), 839–843.

Wulder, M. A., Coops, N. C., Hudak, A. T., Morsdorf, F., Nelson, R. F., Newnham, G. J., & Vastaranta, M. (2013). Status and Prospects for LiDAR Remote Sensing of Forested Ecosystems. *Canadian Journal of Remote Sensing*, *39*(s1), S1–S5.

Wulder, M. A., White, J. C., Nelson, R. F., Næsset, E., Ørka, H. O., Coops, N. C., Hilker, T., Bater, C. W., & Gobakken, T. (2012). Lidar Sampling for Large-Area Forest Characterization: A Review. *Remote Sensing of Environment*, *121*(0), 196–209.

Yang, X. (2012). *Using a Ground-Based Lidar Instrument (echidna®) to Reconstruct Three-Dimensional Forest Structure for Biophysical and Ecological Studies*. Boston University.

Yang, X., Strahler, A. H., Schaaf, C. L., Jupp, D. L. B., Yao, T., Zhao, F., Wang, Z., Culvenor, D. S., Newnham, G. J., Lovell, J. L., Dubayah, R. O., Woodcock, C. E., & Ni-Meister, W. (2013). Three-Dimensional Forest Reconstruction and Structural Parameter Retrievals Using a Terrestrial Full-Waveform Lidar Instrument (Echidna®). *Remote Sensing of Environment*, *135*(0), 36–51.

Yao, T., Yang, X., Zhao, F., Wang, Z., Zhang, Q., Jupp, D. L. B., Lovell, J. L., Culvenor, D. S., Newnham, G. J., Ni-Meister, W., Schaaf, C. L., Woodcock, C. E., Wang, J., Li, X., & Strahler, A. H. (2011). Measuring Forest Structure and Biomass in New England Forest Stands Using Echidna Ground-Based Lidar. *Remote Sensing of Environment*, *115*(11), 2965–2974.

Zhao, F., Strahler, A. H., Schaaf, C. L., Yao, T., Yang, X., Wang, Z., Schull, M. A., Román, M. O., Woodcock, C. E., Olofsson, P., Ni-Meister, W., Jupp, D. L. B., Lovell, J. L., Culvenor, D. S., & Newnham, G. J. (2012). Measuring Gap Fraction, Element Clumping Index and LAI in Sierra Forest Stands Using a Full-Waveform Ground-Based Lidar. *Remote Sensing of Environment*, *125*(0), 73–79.

Zhao, F., Yang, X., Schull, M. A., Román-Colón, M. O., Yao, T., Wang, Z., Zhang, Q., Jupp, D. L. B., Lovell, J. L., Culvenor, D. S., Newnham, G. J., Richardson, A. D., Ni-Meister, W., Schaaf, C. L., Woodcock, C. E., & Strahler, A. H. (2011). Measuring Effective Leaf Area Index, Foliage Profile, and Stand Height in New England Forest Stands Using a Full-Waveform Ground-Based Lidar. *Remote Sensing of Environment*, *115*(11), 2954–2964.

Zianis, D., & Seura, S. (2005). Biomass and Stem Volume Equations for Tree Species in Europe. *Silva Fennica Monographs*, *4*, 1–63.

CURRICULUM VITAE

[redacted]