CHARACTERIZING VEGETATION STRUCTURE AND BIOMASS USING LIDAR REMOTE SENSING

by

SHIHYAN LEE

A dissertation submitted to the Graduate Faculty in Earth and Environmental Sciences in partial fulfillment of the requirements for the detree of Doctor of Philosophy, The City University of New York

2010

This manuscript has been read and accepted for the
Graduate Faculty in Earth and Environmental Sciences in satisfaction of the
Dissertation requirement for the degree of Doctor of Philosophy

Wenge Ni-Meister, Ph.D.

_____	_____
Date	Chair of Examining Committee

Yehuda Klein, Ph.D.

_____	_____
Date	Executive Officer

Allan Frei, Ph.D.

Chuixiang Yi, Ph.D.

Supervisory Committee

THE CITY UNIVERSITY OF NEW YORK

Abstract

CHARACTERIZING VEGETATION STRUCTURE AND BIOMASS USING LIDAR REMOTE SENSING

by

Shihyan Lee

Adviser: Professor Wenge Ni-Meister

Precise characterization of vegetation structure and biomass is significant due to current high uncertainty in estimating global terrestrial carbon sink, ranging from 10 to 60% of total fossil fuel emission. LIght Detection And Ranging (LIDAR) remote sensing is an advanced tool developed for this purpose, and recent identified problems in this area include the need of interpreting lidar height on slopes and estimating forest above ground biomass. This research focuses on these two aspects by assessing the feasibility of analytical solutions, as well as investigating alternative physical interpretation of lidar data.

A recently developed slope correction scheme based on a Geometric Optical and Radiative Transfer (GORT) model was used to quantify the topographic impact on lidar measured vegetation height. By using this scheme, data from spaceborne Geoscience

Laser Altimeter System (GLAS) were compared to airborne Laser Vegetation Imaging Sensor (LVIS) and small-footprint lidar data, where LVIS data is regarded as less affected by slope, and small-footprint lidar data is regarded as ground truth. Analyses show slope-corrected GLAS vegetation heights match well with both small-footprint lidar (R^2 = 0.77, RMSE = 2.2 m) and slope-corrected LVIS heights (R^2 = 0.64, RMSE = 3.7 m). Both slope-corrected GLAS and LVIS height biases are independent on slope.

The investigation of the relationship between lidar data and in-situ measured vegetation structure parameters showed that it is scale- and vegetation type- dependant. For dense forest stands, vegetation biomass is more related to lidar height; while for sparse stands, lidar estimated canopy cover can be more important parameters in approximating tree density variation. To better link lidar data with vegetation structure, a lidar biomass index was developed based on height and estimated canopy cover profile to approximate vertical tree density distribution. Analyses in three different types of forests showed high correction (R^2=0.75-0.83) and near stable relationships between this index and in-situ measured biomass. This index also helps to explain why some height metrics are optimal based on the vegetation structure and topography.

The results presented in this dissertation suggest that the theoretical development can improve the accuracy and interpretation of lidar data, which, in turn, provide unique remote sensing datasets for studies of vegetation structure and biomass, and ultimately decrease the uncertainty in estimating terrestrial carbon sink.

Acknowledgements

I am deeply grateful to my major advisor, Dr. Wenge Ni-Meister, who gave me the opportunity to pursue higher education in the Earth and Environmental Program, the Graduate Center, City University of New York. Her scientific knowledge and vision have enriched my scientific pursuits. Her persistent, selfless and scientific supervision have fostered my enthusiasm for the research presented here. Many thanks to Dr. Allan Frei, as his insightful thoughts and teachings, expert knowledge in climate and statistics set a good example for me. Many thanks also to Dr. Chuixiang Yi, whose expert knowledge in the physics of vegetation dynamics broadened my vision in the field of land-atmosphere interaction. I am honored to have them as members of my dissertation committee.

I would like to express my special thanks to my office mate, Dr. Wenze Yang, for his encouragement and our excellent teamwork led me to succeed in my Ph.D. program. I would also like to thank all the faculty and staff of the Geography Department at Hunter College, whose teachings enriched my knowledge. Further, a special thanks to our system administrator, Nguyen Ngoc, his excellent support guaranteed a suitable and pleasant environment for study and research.

Finally, I would like to dedicate this dissertation to my parents, who offered me their endless love and supported my pursue of higher education in the United States; to my wife, Julie Gutin, who has given her love, support, and encouragement through the

good and bad times of my study; and to my daughter, Victoria Yi-Ling Lee, 9 months, whose birth has given me plenty of joy, happiness and new perspective of life, and an additional motivation to complete my degree.

Table of Contents

List of Tables

List of Figures

List of Abbreviations

AGB	Above-Ground Biomass
BA	Basal Area
BEF	Bartlett Experimental Forest, NH
CHP	Canopy Height Profile
CSIRO	Commonwealth Scientific and Industrial Research Organisation
DBH	Diameter Breast Height
DESDynI	Deformation, Ecosystem Structure, and Dynamics of Ice
DN	Digital Number
EVI	Echidna® Validation Instrument
FWHM	Full Width at Half Maximum
$FWHM_{GLAS}$	FWHM for GLAS
$FWHM_{LVIS}$	FWHM for LVIS
GLAS	Geoscience Laser Altimeter System
GORT	Geometric Optical and Radiative Transfer
H	Maximum vegetation Height
H_{GLAS}	Slope corrected GLAS vegetation height
H_{LVIS}	Slope-corrected LVIS vegetation height
H_{LVIS3}	Mean height of top 3 H_{LVIS} within each GLAS footprint
H_{SF}	98 percentile of small-footprint discrete-return lidar height
H_{field}	Field measured maximum vegetation height
H_g	Ground elevation

H_{w_end}	Waveform ending height (last detectable return)
HF	Harvard Forest, MA
HOME	Height of Medium Energy Returns
ICESat	Ice Cloud and Elevation Satellite
L_{ext}	GLAS waveform extent
Lidar	Light detection and ranging
LPA	Laser Profile Array
LVIS	Laser Vegetation Imaging Sensor
MODIS	Moderate Resolution Imaging Spectroradiometer
NASA	National Aeronautics and Space Administration
NED	National Elevation Data
QMCH	Quadratic Mean Canopy Height
RHx	Lidar height relative to ground at x% of accumulated waveform energy
RH50/75/100	Lidar height relative to ground at 50/75/100% of accumulated waveform energy
SLICER	Scanning Lidar Imager of Canopies by Echo Recovery
SRTM	Shuttle Radar Topography Mission
UNFCCC	United Nations Framework Convention on Climate Change
USGS	U.S. Geological Survey
WWI	Waveform Weighted Index

Chapter 1 Introduction

In December, 2009, at the United Nations Climate Change Conference in Copenhagen, an agreement had been made by countries to cap the global temperature rise to 2 degrees Celsius by committing to significant emission reductions among the developed countries and raise finance, provide technologies to kick start action in the developing world to deal with climate change. The goal is recommended by United Nations Framework Convention on Climate Change (UNFCCC), which was adopted by 189 nations in 2007 and has its objective as "stabilization of greenhouse gas concentrations in the atmosphere at a level that would prevent dangerous anthropogenic interference with the climate system". The challenge now is to turn what is agreed into something that is legally binding in the subsequent meetings. Understanding the global carbon cycle is critical to this objective because its processes defined how emissions of carbon dioxide (CO_2) from anthropogenic activity translate into concentration of CO_2 in the atmosphere. To facilitate fair international climate change policy, it is crucial to establish scientific understanding on how much of carbon emitted to the atmosphere remains; specifically, where and what are the carbon sources and sinks?

The current understanding of global carbon budget has indicated terrestrial ecosystem as one of the key uncertain components (Houghton 2007). Global terrestrial ecosystems absorbed carbon at a rate of 1–4 Pg/yr during the 1980s and 1990s, offsetting 10–60 per cent of the fossil-fuel emissions (Solomon et al., 2007; Houghton 2007). More importantly, the regional patterns and causes of terrestrial carbon sources and sinks remain highly uncertain (Solomon et al., 2007; Houghton 2007; Schimel, et al., 2001). Forest has thought to be the major source of the uncertainties in terrestrial carbon sinks

(Goodale et al. 2002; Luyssaert et al., 2009) due to lack of precise vegetation structure information to adequately monitor and model the carbon cycle and forest ecosystem dynamics (Rosenqvist et al., 2003; Hese et al., 2005; Houghton et al., 2009). In addition, it is necessary to know the three dimensional vegetation structure when calculating boundary layer energy balances in global climate, hydrological and numerical weather prediction models (Sud et al., 1988). Further, vegetation structure characteristics also affect forests' role in land surface stabilization processes, animal habitats, and the production of food, materials and pharmaceuticals (Bergen et al., 2009; Buermann, 2002). Therefore, there is a need to map and monitor vegetation structure and forest biomass globally with routine updates.

1.1 Monitoring forest structure and biomass

Traditionally, forest structure and biomass are measured in-situ, which is expensive and time consuming. In large-scale forest management practice, only a very small fraction of ground is inventoried. The inventory tree metrics are used to estimate their biomass and other structures using allometric equations and then extrapolate these estimates to forest stands and to entire forest ecosystems (Botkin and Simpson, 1990; Botkin et al., 1993; Fournier et al., 2003). Also, most ground measurements focus on what is commonly referred to as "stand structure", the size and number of woody stem per unit area (Miles, 1979; Oliver and Larson, 1996). Such measurements lack the information on crowns and heights distribution, which are more difficult to measure but critical components to fully describe forest canopy structure (Maser, 1989). Because the extremely low sample aerial percentage, extrapolating current ground forest measurements to global scale is doomed highly un-reliable.

For the last thirty-five years, remote sensing has greatly improved our understanding of the earth's surface. Observations from seven Landsat satellites to NASA's Earth Observing System (EOS) satellites, and numerous airborne and space-based sensors using multi-angle, multispectral and radar techniques, had created several global land cover data bases. However, the current best efforts have only led to the ability to classify our land surface which includes forest's spatial distribution but had limited success in characterizing vegetation's vertical structures, such as canopy height, density, age and biomass (Dubayah et al., 2000). Significant improvements from future advance in EOS-like satellites and inversion algorithms are not likely due to the inherent limitation on extracting three-dimensional information from two-dimensional measurements.

In recent decades, the development of lidar remote sensing has been shown to be a breakthrough technology for deriving forest canopy structure characteristics (Lim et al., 2003; Frolking et al., 2009). The advantage of lidar over other remote sensing technologies lies on the use of high power laser, which can penetrate forest canopy and interact with canopy components vertically. Many studies had demonstrated the strong relationships between field-measured vegetation structure and lidar derived parameters (Simarc et al, 2008; Lefsky et al., 1999, 2002, 2005; 2007; Harding et al., 2001, 2005; Drake et al., 2002a, 2002b, 2003; Nelson et al., 1988, 1997, 2009; Patenaude et al., 2004); however, the theoretical understanding of the relationships is still incomplete. For lidar remote sensing to aid the understanding of global carbon cycle, it is necessary to improve our understanding of the physical meaning of lidar data so that vegetation structures can be adequately quantified at broad scale.

1.2 lidar remote sensing on vegetation structure

Lidar (**li**ght **d**etection **a**nd **r**anging) is an active remote sensing technique, analogous to radar, but using laser light (Young, 1986). Lidar measures the roundtrip time for a pulse of laser energy travelling between the sensor and a target. There are two types of lidar systems, "discrete return" and "full waveform" systems. The later, full waveform systems, is consider better for capturing true vertical vegetation structure (Lim et al., 2003), and is the focus of this dissertation. In full waveform lidar system, the laser beam is diverged to cover a small area, or footprint, usually is a circle and measured by diameter (usually 8-70m, Means, 1999; Harding et al., 2000). The incident pulse is centered at 1064 nm, a near-infrared wavelength which reflectance from branches, leaves and barks are similar (Wehr and Lorh, 1999). The entire time history of the outgoing and return laser pulses is digitized as "waveform" using a single detector, digitizer and timing clock, and unambiguously describes the range to the surface as well as the vertical distribution of surfaces within each laser footprint (Blair et al., 1994, 1999). Because the lidar waveform is the direct measurement of the vertical distribution of the intercepted canopy elements, lidar data are particularly effective in estimating vegetation canopy height, its vertical profile and forest above ground biomass.

However, large uncertainties still exists when lidar data alone is used to retrieve vegetation structure parameters, such as height and above ground biomass. The current state of vegetation lidar remote sensing applications summarized by Nelson et al. (2009) show that the relationships between forest structure parameters and lidar data need to be empirically defined by ground measurements. This is because the lidar-vegetation structure relationships can be influenced by vegetation types, ages, canopy openness, soil

condition and even local climate. In addition, lidar waveform measurements are further complicated by topographical relief, surface roughness and spatial distribution of dominate trees (Harding et al., 2005; Yang et al., 2010; Hyde et al., 2005). Consequently, significant gaps remain in our knowledge of developing a standardized lidar method that can be used for global operational vegetation structure and biomass retrieval.

1.2.1 Topography Effect

Topography within lidar footprint can alter the relationships between lidar waveform and canopy structure. Although the basic waveform shape is determined by canopy structure, lidar waveform could be further modified by the factors like surface topography, footprint size, off-nadir pointing, laser pulse width, surface roughness and the ratio of vegetation and background reflectivity (Yang et al., 2010). For the typical vegetation lidar systems, topography effect has been identified as one of the biggest uncertainties for vegetation height and biomass retrieval on high slopped terrains (Harding et al., 2005; Lefsky et al., 2007). Recent observation and modeling studies had demonstrated that other factors impacting lidar waveform can be amplified by topography (Harding et al., 2005; Yang et al., 2010; Hyde et al., 2005); further strengthen the need to address topography effect when lidar remote sensing is used in non-flat surfaces.

1.2.2 Vegetation Structure Characteristics

The relationships between vegetation's vertical and horizontal structure and lidar data is intrinsically complicated (Neuenschwander et al., 2008). Lidar laser energy is

much stronger in the center than the edge of the footprint; therefore, trees located near the center will have stronger influence on recording waveform. In addition, lower canopy can be obscured by upper canopy which makes occlusion another issue in linking canopy structure with lidar waveform.

The recent advanced geometric and radiative transfer models had provided a mean to investigate the relationships between vegetation structure and waveforms (Sun and Ranson, 2000; Ni-Meister et al., 2001). Given the vegetation structure parameters, lidar waveforms can be accurately constructed using geometric and/or radiative transfer models. The inversion from lidar data, however, will require ancillary information such crown shape, foliage density, tree density, size and distribution, and the ratio of foliage and background reflectance (Koetz et al., 2006).

Forest structure characteristics are influenced by species composition, age, climate, topography, soil and other geographical/biological factors (Pacala et al., 1996). Investigating relationships between lidar data and vegetation structure parameters regarding to each specific condition will contribute to our understanding of how lidar data can be used retrieve vegetation structure without in-situ data.

1.2.3 Above Ground Biomass

Forest above ground biomass is the assemblage of stem, bark, branches and foliage biomass. Although above ground biomass is mostly contained in the woody components (e.g. stem), it was the lowest biomass component, foliage, has the highest impact on lidar waveform. Therefore, for operational lidar biomass estimates, the relationships of lidar waveform response to canopy foliage and forest's woody components need to be properly defined.

The relationships between foliage and woody components is, in principle, governed by the plant dynamic (vary by species composition) and boundary condition (e.g. age, climate, competition) (Hurtt et al., 1998). The studies in allometry of tree and forest stand provide evidence of the link between canopy foliage profile and forest's woody components (Pacala et al., 1996, Albini et al., 2006). The results suggest lidar waveform should respond to change in above ground biomass. While strong relationship had been demonstrated between above-ground biomass and lidar derived parameters, their relationship to canopy properties of tree cover, size and shape are unknown. The consequence is the need for in-situ data to aid the selection the optimal lidar parameters and determination of lidar-biomass relationship.

1.3 Objective and dissertation structure

The main objectives for this dissertation are: (1) to better understand the relationships of forest structure parameters and lidar data; (2) to examine the effect of vegetation structure on lidar response to forest biomass; (3) to develop a lidar biomass parameter that is more universal to vegetation structure and compare its performance with the other commonly used lidar biomass parameters in different biomes; and, (4) to evaluate topography effect on lidar waveform and develop a method to correct such effect on lidar derived vegetation parameter.

The first aim of this study is to investigate the potential of removing topography effect on spaceborne lidar (GLAS) data to retrieve vegetation height over slopes. Airborne lidar data collected in White Mountain in New England region were used to evaluate our correction scheme. I first analyzed the lidar waveform and height change due to slope effects and proposed a correction method based on a simple geometric

model. Thereafter, the un-corrected and corrected lidar vegetation height derived by GLAS and LVIS were compared to ground truth to verify the earlier analogy. This study addressed the topography impact on lidar vegetation height retrieval which is vital in improving lidar retrieved vegetation structure accuracy in the mountainous terrain (Chapter 2).

Thereafter, I investigated the relationships between measured vegetation structures and lidar parameters to better understand the lidar-biomass associations. For this, in-situ collected stem map and airborne/ground lidar data sets are used. A total of 50 forest plots data collected in 2003 and 2007 from three experimental forests in New England, USA were used in this study. The lidar data were collected using airborne lidar, the Laser Vegetation Imaging Sensor (LVIS, Blair et al., 1999) and ground scanning lidar, the Echidna® validation instrument (EVI) (Jupp et al. 2005). I attempted to assess the correlation between lidar and ground measurements of canopy height, cover, DBH, woody volume and above ground biomass with discussion of the role of forest type (chapter 3).

Next, I proposed a lidar biomass parameter and examined its potential in large scale application. Towards this goal, I performed a comprehensive analysis between all lidar height metrics and above-ground biomass and compared the results with those derive using my parameter. Forest data from three biomes, including coniferous and deciduous forests in New England, and Sierra Neveda Forest in California's Cascade Mountain region, were used. This study addressed why different lidar parameter was used to derive forest biomass in different biome in previous studies and provided a method to stabilize lidar-biomass relationship to improve large-scale biomass retrieval

accuracy (chapter 4). Finally, I present concluding remarks about my dissertation research in Chapter 5.

Chapter 2 Physically Based Vertical Vegetation Structure Retrieval from ICESat Data

2.1 Introduction

Remote estimation of vegetation structure characteristics and forest biomass has become essential to ecosystem modeling studies and advancement in understanding of many ecosystem processes. Unlike many other remote sensing measurements, vegetation lidars provide direct and indirect measurements of vegetation structure (Dubayah and Drake, 2000). Recently vegetation lidar data have become more widely available to study the link between vegetation lidar signals and vegetation structure characteristics. The spaceborne Geoscience Laser Altimeter System (GLAS), part of the ICESat mission, provides global lidar data with a variable diameter ~70 m footprint spaced at ~170 m (Zwally et al., 2002; Harding et al., 2005; Lefsky et al., 2005b). Airborne data have also been collected using a Scanning Lidar Imager of Canopies by Echo Recovery (SLICER) with a 15 m footprint and the Laser Vegetation Imaging Sensor (LVIS) with a footprint of 10 – 25 m over several large areas for improved vegetation structure characterization (Blair et al., 1999). Small-footprint lidar data are also accessible in many small regions of the globe and automatic methods were developed to extract vegetation structure characteristics from them (Chen, 2007; Popescu and Wynne, 2004). These global, regional, and local lidar data can provide detailed vegetation structure data necessary for carbon models and studies of ecosystem processes.

Many studies have demonstrated the potential use of spaceborne and airborne vegetation lidar data to estimate vegetation height, above-ground biomass characteristics, and other vegetation structure parameters (Lefsky et al., 1999b; 2002a; 2002b; 2005b; Harding et al., 2001; 2005; Drake et al., 2002a; 2003; Nilson et al., 2009; Patenaude et al., 2004; Anderson et al., 2006; 2008; Popescu, 2007; Popescu and Zhao, 2008; Chen et al., 2007; and Chen, 2010b). On flat surfaces, studies have shown GLAS data can be used to estimate important forest parameters, such as vegetation canopy height and above-ground biomass, over different vegetation types (Lefsky et al., 2005b; Simard et al., 2008). In recent studies, the accuracy of laser altimetry data from GLAS over flat terrains has been validated using small-footprint waveform lidar (Neuenschwander et al., 2008) and airborne vegetation lidar – LVIS (Sun et al., 2008). These studies found that the vertical structures in GLAS altimetry data (or waveforms) are similar to those synthesized airborne lidar data collected within GLAS footprints. The results confirmed this spaceborne sensor is capable of studying forest structure at low- to no-relief surfaces.

However, the ability of lidar to retrieve vegetation structure over mountainous regions is limited due to the complex interactions of lidar signals from vegetation and ground surface (Harding et al., 2005; Lefsky et al., 2007). Recently many studies have focused on estimating vegetation height from GLAS over mountainous regions. Two approaches are commonly used to derive vegetation height: One approach is to decompose lidar waveforms into multiple Gaussian distributions. The last or the second last peak is identified as the ground, and vegetation height is estimated as the distance between first and ground peak returns (Hofton et al., 1999; Rosette et al., 2009; 2008; Sun et al., 2008; Neuenschwander et al., 2008; Chen, 2010b). The advantage of this

approach is to detect slope from the last/second to last Gaussian peak. However, this method can lead to large uncertainties in height estimate when ground returns are well mixed with vegetation returns and it is impossible to find ground peak returns. The second approach uses statistical analysis to separate vegetation and ground returns from full waveform extents. Vegetation height is calculated from waveform extents and trailing and leading (Lefsky et al., 2007; Pang et al., 2008; Chen 2010b) or other waveform shape parameters (Duncanson et al., 2010). The advantage of this approach is that there is no need to detect ground returns, instead, slope is indirectly calculated from waveform shape parameters and removed from waveform extent. Chen (2010b) indicates the second approach achieves better height estimates than the first approach. However, the relationships between these waveform shape parameters and slope varies with vegetation structure and topographical conditions. Therefore, these types of statistical models often need to recalibrate the fitting coefficients for different vegetation types and/or different topographic configurations (Lefsky et al., 2007; Pang et al., 2008). A physical approach is required to develop a consistent vegetation height product from global ICESat/GLAS data.

Our recent modeling work used the vegetation canopy Geometric Optical and Radiative Transfer (GORT) model to quantify the impacts of surface topography, footprint size, off-nadir pointing, surface roughness and laser pulse distribution on lidar vegetation height and lidar waveforms (Yang et al., 2010). We developed an analytical approach to quantify the impact of surface topography and footprint size on vegetation height with the assumption that vegetation canopy is uniformly distributed within each footprint. Vegetation height can be estimated from either GLAS full waveform extent or

LVIS height metrics using this approach. The main objective of this study is to apply this approach to retrieve vegetation height from ICESat/GLAS and to evaluate our approach by comparing retrieved ICESat/GLAS vegetation heights with LVIS, small-footprint discrete-return lidar and field data collected in southern White Mountains, NH. National Elevation Data (NED) was used to estimate slope. However, with the availability of global elevation productions from Shuttle Radar Topography Mission (STRM) and Advanced Spaceborne Thermal Emission and Reflection (ASTER) data, our approach may have the potential to derive global vegetation heights from ICESat/GLAS data.

Another objective of this study is to assess the quality of GLAS measured vegetation structure. ICESat/GLAS data can be noisy due to strong backscattering from low clouds (Abshire et al., 2005) and waveforms can be saturated at canopy layer for dense vegetation or at ground for sparse vegetation. ICESat samplings are sparse, therefore, it would be valuable to assess if these noisy and saturated waveforms have any valuable vegetation structure information. We compared GLAS with aggregated LVIS waveforms and assessed impact of background noise and saturation on GLAS vegetation structure measurements.

The chapter is organized as follows: Section 2.2 describes the site and datasets. Section 2.3 presents LVIS waveform aggregation and height estimate algorithms. Section 2.4 first discusses the evaluation results of our slope correction method by inter-comparing heights from GLAS, LVIS, small-footprint discrete-return lidar and field data; then assesses GLAS waveform data quality under different conditions by comparing GLAS and aggregated LVIS waveforms. Section 2.5 discusses the uncertainties of our algorithm and differences from other methods. Section 2.6 is the conclusion.

2.2 Study Site

The study site is the southern portion of White Mountain National Forest in New Hampshire, USA. The region is featured with old-growth northern hardwoods. The dominant species include beech, yellow birch, sugar maple and eastern hemlock. In lower elevations, most of the forest is covered by tall canopies where sugar maple/beech/yellow birch dominates the upper canopy layer with maximum height around 30 m. At higher elevation before reaching the forest line, spruce, fir, and hemlock are commonly mixed with hardwoods, and can be dominant on cool steep slopes (Filip et al., 1971). The regional terrain varies from flat in the lower valley to slopes of over 30 degrees on higher mountain passes. Higher slopes usually occur at higher elevation (Figure 2.1).

2.3 Data sets

2.3.1 ICESat/GLAS data

Global ICESat/GLAS data is available from 2003 to 2008. ICESat/GLAS is the first spaceborne lidar system designed to observe surface structures, including ice sheet mass balance and vegetation characteristics in three dimensions at a global scale. Three onboard lasers were designed to sample Earth's surface at a 172 m interval along the track with a maximum cross track separation of 15 km at the equator (Zwally et al., 2002). Due to the unexpectedly short lifetime of the laser system, the GLAS mission started to operate with a 91-day repeat orbit (with a 33 day sub-cycle) (Sun et al., 2008). In each of the operating years, GLAS acquired data in winter (Feb–March), summer (May–June) and fall (October–November) periods. The three sub-cycles were designed to capture seasonal and interannual change on land surface, e.g. ice and forest canopy.

Studies had shown the typical geolocation accuracy of GLAS altimetry data is between 2.4 and 5.8 m (Abshire et al., 2005; Carabajal et al., 2005; Neuenschwander et al., 2008), although it can be as high as 60 m (Sun et al., 2008).

To measure the vertical structure of land and ice surfaces, GLAS uses the 1064-nm laser pulses and records the returned laser energy from an ellipsoidal footprint. The original GLAS footprint size was designed to be ~65 m; however, actual footprint diameter was about 110 m, 90 m and 55 m for Lasers 1, 2 and 3, respectively (Zwally et al., 2002; Abshire et al., 2005; Schutz et al., 2005). GLAS footprint size, orientation and ellipticity can vary shot by shot and precise measurements can only be obtained from laser far field patterns, or Laser Profile Array (LPA) images (Bae and Schutz, 2002). In general, GLAS footprint ellipticity is the smallest for Laser 3 and largest for Laser 1, and only Laser 3 footprint is close to a circle (http://www.nsidc.org/). Since most LPA images are not available in our study region, we chose to use only Laser 3 data, collected in June of 2005 (L3C) and 2006 (L3F), to reduce GLAS footprint coverage uncertainty due to lack of precise ellipticity information. We then obtained GLAS footprint sizes and ellipticities from the laser operation average, and assumed the footprint is a circle and the diameter equals the mean of major and minor axes.

The background noise level of GLAS data can increase under cloudy condition due to the backscattering from clouds (Abshire et al., 2005). Two rules were used to select good quality GLAS data: 1) The minimum signal-to-noise ratio was set to 15, and 2) The elevation differences from the built-in SRTM and GLAS were used to detect potential low cloud data. Because SRTM penetrates through clouds and GLAS does not, elevation difference from these two datasets indicates the lidar returns are from the cloud or

ground. The threshold of the elevation difference was set slightly larger than the maximum vegetation height in the region (50 m) to remove potential low cloud data.

2.3.2 LVIS data

LVIS data in southern White Mountain region of NH were acquired during July 18-20, 2003 with 20 m footprint size (Figure 2.1). The LVIS is an airborne lidar system designed to collect data on surface topography and vegetation structure (Blair et al., 1999). LVIS standard data products include the fully digitized waveforms and the height metrics (relative to surface) where 25%, 50%, 75% and 100% of the waveform energy occur, named as RH25, RH50, RH75 and RH100, respectively (Blair et al., 2006). Often all LVIS height metrics are released for public access, but the fully digitized waveforms are only publicly available for some intensive study sites.

For this study, we used relative heights at 100% waveforms energy, or RH100, to estimate forest canopy heights in the whole region and the available full waveform data within the 2 km radius area of the main flux tower in Bartlett Experimental Forest, NH for waveform comparisons with the GLAS data.

2.3.3 NED data

National Elevation Dataset (NED) is a seamless and nationally consistent set maintained by the U.S. Geological Survey (USGS) (http://egsc.usgs.gov). The standard 1 arc-second (~20x30 m resolution in our study region) NED digital elevation model was used to derive slope. The slope for each NED grid cell was determined as the maximum rate of change between each cell and its immediate neighbors. At 1 arc-second scale,

previous research had shown NED's slope accuracy is 3 degree or less (Holmes et al. 2000).

The slope for each ICESat or LVIS footprint was calculated as the mean of slopes for all grid cells within the footprint. Since LVIS footprint is similar in size to NED's 1 arc-second grid, we expect the slope within each LVIS footprint has similar accuracy to the NED grid scale i.e. 3 degree or less. For GLAS, better slope accuracy might be achieved due to the larger footprint size.

2.3.4 Small-footprint discrete-return lidar data

First- and last-return discrete lidar data were collected at Bartlett Experimental Forest (see Figure 2.1) using an Optech ALTM 2025/2050 sensor in August of 2005. The reported accuracies of lidar data were 15 cm vertically and 50 cm horizontally. To derive canopy height, a "bare earth" digital elevation model was first constructed at 1 m spatial resolution using the Toolbox for Lidar Data Filtering and Forest Studies (TIFFS) program (Chen, 2007). TIFFS uses morphological approaches with variable window sizes to resolve pixels without any lidar shot penetrating canopy to construct the best "bare earth" digital elevation model (Chen et al., 2007). Any lidar points above the bare earth surface are considered to be vegetation canopy returns. In this study, a 1 m lidar canopy height model was created using the highest lidar canopy returns within each 1 m grid cell.

2.3.5 Field data

Field data were collected during the summer 2009 Deformation, Ecosystem Structure and Dynamics of Ice (DESDnyI) field campaign in NH, within the LVIS

sampling range. Field data collection consists of eight 1-ha plots (50x200 m grid size), each plot was divided into 16 subplots (25x25 m). The heights of the tallest trees (varying from one to three) were measured and their average was used to compare with LVIS and small-footprint maximum height. Field plot geolocation uncertainties range from 4-8 m and 12-14 m for two subplots. 4 or 6 years differences between LVIS/small foot print and field data was taken into account in our analysis. No field data is available within GLAS footprints. Figure 2.1 shows the spatial coverage of GLAS, LVIS, small-footprint discrete-return lidar and field sites.

2.4 Method

2.4.1 Vegetation height retrieval from lidar over slope terrains

The two most commonly used height metrics to infer vegetation height from lidar are waveform extent for GLAS and RH100 for LVIS (Figure 2.2). Waveform extent is defined as the height difference between the first and last elevations at which the waveform energy exceeds a threshold. The threshold was set as 4.5 times of background noise in GLAS (Lefsky et al., 2005b and 2007) and three times of background noise in LVIS (Blair et al., 2006). RH100 is defined as the distance from the top of vegetation returns to the peak of last Gaussian pulse or ground returns (Blair et al., 1999).

Both waveform extent and RH100 are influenced by surface topography, footprint size and laser pulse energy distribution/surface roughness. Yang et al. (2010) extended the Geometric Optical and Radiative Transfer (GORT) vegetation lidar model to take into account the impacts of surface topography and off-nadir pointing on vegetation lidar waveforms and vegetation height retrieval. This study presented an analytical formula to

quantify the first-order influence of surface topography, footprint size, laser pulse energy distribution/surface roughness and off-nadir pointing effect on lidar vegetation height. For nadir-viewing lidar application, the relationship between vegetation height (H) (i.e. lidar vegetation height at a flat terrain), RH100, waveform extent (L_{ext}) and surface topography (with slope θ), footprint size (d) and the pulse width/surface roughness can be described by the following equations (see Figure 2.3):

$$\begin{aligned} L_{ext} &= H + d\tan\theta + c \cdot FWHM / 2 \\ RH100 &= H + d\tan\theta / 2 \end{aligned} \tag{2.1}$$

where *c* is the speed of light and FWHM (Full Width at Half Maximum) describes the laser pulse width, which is 6 ns (equivalent to 1.8 m) for GLAS and 10 ns (equivalent to 3 m) for LVIS. Surface roughness also plays the similar role of broadening waveforms as the Gaussian laser pulse energy distribution, and c*FWHM is the combined effect of laser pulse width, energy distribution and surface roughness.

Eq. (2.1) suggests that vegetation height (H) can be retrieved from RH100 or waveform extent (L_{ext}), knowing footprint size, d, slope (θ), and c*FWHM, and are described as follow:

$$\begin{aligned} H &= L_{ext} - d\tan\theta - c \cdot FWHM / 2 \\ H &= RH100 - d\tan\theta / 2 \end{aligned} \tag{2.2}$$

However, this relationship was derived based on very smooth waveforms, whose extent and canopy top is defined with zero threshold. In reality, waveforms have background noise and their extents or RH100 values are defined using a non-zero threshold, e.g. 4.5 times of background noise in GLAS, three times in LVIS. Therefore, in Eq. (2.2) waveform extent/RH100 needs to be corrected (extended).

We applied a simple correction scheme to extend waveforms with a zero-threshold: waveforms were smoothed first to reduce uncertainty caused by noises; then extended linearly with a zero-threshold using locations of the original zero and 10 percentile of waveforms as two reference points. The same procedure was applied to the lower end of waveforms. Then waveform extents were recalculated using the extended waveforms. In this study we were not able to correct RH100 due to unavailability of LVIS full waveforms within most GLAS footprints. We expect that the impact on RH100 would be much smaller (~1/3) comparing to GLAS waveform extent due to its smaller threshold (see discussion before) and for one side of waveform correction only.

To evaluate our approach, we compared LVIS RH100 with and without slope correction with small-footprint discrete-return lidar and field measured vegetation height. We also compared slope-corrected GLAS with slope-corrected LVIS and small-footprint discrete-return lidar heights (see Section 2.4.5).

2.4.2 GLAS and LVIS waveform comparison

To evaluate GLAS waveforms, GLAS waveforms were compared to aggregated LVIS waveforms, which were calculated as a summation of all LVIS waveforms within each GLAS footprint weighted by GLAS laser energy Gaussian distribution (see Figure 2.4):

$$
\begin{aligned}
F(z) &= \sum_i f_{Li}(z) * w_i \\
w_i &= g_i / \sum_i g_i \\
g_i &= e^{(-2r_i / d_{GLAS})}
\end{aligned}
\qquad (2.3)
$$

where $F(z)$ is the aggregated LVIS waveform intensity at height z, $f_{Li}(z)$ is the i^{th} LVIS waveform intensity at height z, z is the relative height to the mean LVIS ground

elevation, this indicates that during the aggregation, heights for each LVIS waveform on a slope were shifted vertically relative to mean ground elevation in the GLAS footprint. w_i is the normalized weight for i^{th} LVIS waveform, g_i is the weight for i^{th} LVIS waveform, r_i is the distance from GLAS centroid to i^{th} LVIS footprint center, and d_{GLAS} is the GLAS footprint diameter.

For comparison, LVIS waveforms were aggregated vertically from its original resolution, 0.3 m, to GLAS resolution, 0.6 m. GLAS and the aggregated LVIS waveforms are not directly comparable because the total waveform energy is different. We also compared normalized waveforms, instead of empirically re-scaling them by peak energy returns (Neuenschwander et al., 2008).

The elevations reported in GLAS and LVIS or other airborne data do not have exact agreement (Sun et al., 2008; Neuenschwander et al. 2008). To overlay GLAS and LVIS aggregated waveforms, this study used an automatic approach by first adjusting the height of GLAS and LVIS waveforms relative to their corresponding ground elevation; then the waveforms are aligned by assuming the same ground elevation. This approach allows us to overlay LVIS and GLAS waveforms automatically for comparison as opposite to manual shifting waveforms used in Neuenschwander et al., (2008).

Ground elevation is a standard LVIS product, and LVIS mean ground elevation at the GLAS footprint was calculated as the mean of all LVIS data within each GLAS footprint. We chose not to use ground elevation estimated from the Gaussian peaks in the GLAS product to avoid ground peak identification uncertainties, which could be significant under dense canopy conditions and on slope terrains due to its large footprint size (Lefsky et al., 2007; Yang et al., 2010; Chen, 2010a). Original heights in GLAS

waveforms were relative to the last waveform return (H_{w_end}). In this study, heights in each GLAS waveform were shifted relative to GLAS ground elevation (H_g) using the elevation difference between H_g and H_{w_end}, and the GLAS ground elevation was treated as the same as the LVIS mean elevation. The height shift is described as,

$$H_g - H_{w\_end} = (d_{GLAS} \tan\theta + c \cdot FWHM_{GLAS}) / 2 \quad (2.4)$$

where $FWHM_{GLAS}$ is FWHM for GLAS. Since we don't have surface roughness information, the impact of $c \cdot FWHM_{GLAS}$ was ignored in Eq. (2.4).

2.5 Results

2.5.1 Vegetation height comparisons

2.5.1.1 *LVIS, field and small-footprint discrete-return lidar data comparison*

Although LVIS footprint size (20 m) is much smaller than GLAS', LVIS height could also be affected by slope as discussed before. To evaluate the quality of LVIS height data and the impact of slope on LVIS heights, we first compared LVIS height before and after slope correction with field measured maximum heights and small-footprint discrete-return lidar data. Figure 2.5 compares LVIS height, $H_{LVIS,}$ (with slope correction), RH100 (without slope correction) with field measured maximum tree height (H_{field}) and the dependences of their differences, (RH100-H_{field}) and (H_{LVIS}- H_{field}) on slope. The slope-corrected LVIS heights match better to field measurements than RH100 (RMSE 4.8 m and R^2 0.29 vs. RMSE 5.2 m and R^2 0.19). Overall, field measured heights tend to be taller than LVIS', which is likely due to the tree growth between 2003 (LVIS data collected) and 2009 (field data collected).

RH100 and measured maximum tree height difference, (RH100-H_{field}) increases with slope indicating the impact of slope on RH100 (Figure 2.5c). However after removing slope from RH100 using our scheme, LVIS and field measured height difference, (H_{LVIS}-H_{field}), indicates no relationship with slope (Figure 2.5d). RH100 without slope correction could overestimate height by 3 m over a terrain with a 20 degree slope. The result indicates our slope correction successfully removes slope-effect in RH100. The relative large RMSE could be caused by large geolocation uncertainties of field data (approximately 4-8m).

Figure 2.6 shows the same comparison as in Figure 2.5 except with small-footprint discrete-return lidar height (H_{SF}), calculated as 98 percentile of canopy height to reduce canopy top height uncertainty. This comparison shows similar results as shown in Figure 2.5. Our scheme improves the overall height retrieval accuracy by 0.7m (RMSE 3.7m in Figure 2.6a vs. 3.0m in Figure 2.6b). Similar slope dependency of (RH100-H_{SF}) as in Figure 2.5c was shown in Figure 2.6c. However after removing slope impact, (H_{LVIS}-H_{SF}) is independent on slope (Figure 2.6d) indicating the slope effect was removed from RH100. The height retrieval bias can reach 4-5 m at 15 degree slope terrain (Figure 2.6c). The overall uncertainty is smaller in Figure 2.6 (RMSE = 3.0 m) than the one in LVIS and field-measured height comparison as shown in Figure 2.5 (RMSE = 4.8 m) implying LVIS is better geolocated with small-footprint discrete-return lidar data than with the field data. Opposite to what is shown in Figure 2.5, LVIS slightly overestimates small-footprint lidar height. This is somewhat expected because LVIS height is very close to top of the canopy height, but small-footprint lidar sometimes misses the tallest tree tops.

We also compared the small-footprint discrete-return lidar and field measured height differences, (H_{field} - H_{SF}), which tend to show a slightly decreasing trend with slope (Figure 2.7). The trend might be caused by the mismatched geolocation problem or by some other random factors.

2.5.1.2 GLAS and small-footprint discrete-return lidar height comparison

We compared slope-corrected GLAS height (H_{GLAS}) with small-footprint discrete-return lidar (Figure 2.8). GLAS heights match well with 98 percentile of small-footprint lidar heights (H_{SF}) (RMSE = 2.2 m and R^2 = 0.77) (Figure 2.8a). Comparing to the LVIS and small-footprint discrete-returns lidar height comparison with a larger RMSE=3.0m, it might indicate that our scheme works better for a larger footprint lidar. However we can not rule out the possibility that the larger uncertainty in LVIS height estimates comes from the uncertainty of RH100 estimates, particularly for dense vegetation on steep slopes. RH100 was estimated using the Gaussian decomposition approach and particularly at steep slopes, the ground return is well mixed with vegetation returns and identifying ground peaks might be difficult at forest stands with dense canopies.

We also evaluated the waveform extent correction scheme (see Section 2.3.1) (Figure 2.8b). Comparison of small-footprint (H_{SF}) and slope-corrected GLAS heights with (Figure 2.8a) and without (Figure 2.8b) waveform extent correction indicates that correction of waveform extent improves height accuracy (R^2 = 0.77 and RMSE = 2.2 m in Figure 2.8a vs. R^2 = 0.74 and RMSE = 2.9 m in Figure 2.8b).

We further investigated the effectiveness of slope correction scheme by comparing GLAS and small-footprint lidar height differences (H_{GLAS} - H_{SF}) with slope (Figure 2.8c). To examine the impact of correction of waveform extent on height retrieval, we overlay

(H_{GLAS} - H_{SF}) with/without correction of waveform extent (Figure 2.8c). Retrieved heights can increase from 0 to 5 m using our correction scheme. Height differences, (H_{GLAS} - H_{SF}), show slightly decreasing trends with slope for both schemes. The trend is roughly the same as the slope dependency of small-footprint lidar and field measured height differences (Figure 2.7c). In summary, extending original GLAS waveform extent to a zero-threshold is necessary for our slope correction scheme. However, a more systematic study is required in the future. H_{GLAS} with the correction of waveform extent is used for the rest of the paper.

2.5.1.3 GLAS and LVIS height comparison

To further evaluate GLAS height at a large scale with an increased slope range, we compared GLAS with LVIS heights (Figure 2.9). LVIS height was calculated as the mean of three tallest LVIS heights (slope-corrected) within each GLAS footprint. Direct comparison of GLAS and LVIS height (Figure 2.9a) shows RMSE=3.7m and $R^2 = 0.64$. The RMSE here is larger than the value in GLAS and small-footprint lidar comparison (RMSE = 2.2 m). Two possible factors contribute to this: a). larger height range in LVIS data (13.6 – 35.5 m with standard deviation, std = 4.6 m) than in small-footprint lidar (19.6 – 29.9 m, std = 3.1 m), the possible uncertainty in RH100 estimate (see the discussion before).

Figure 2.9b and Figure 2.9c compare the slope dependencies of the uncorrected lidar height differences (L_{ext} - RH100) and slope-corrected lidar heights (H_{GLAS} - H_{LVIS3}). The slope dependency is strong between uncorrected lidar height differences (Figure 2.9b) but minimal after slope correction (Figure 2.9c). The dependence of these differences on slope can be derived from Eq. (2.1) and described as follows:

$$L_{ext} - RH100 = (d_{GLAS} - d_{LVIS} / 2)\tan\theta + c \cdot FWHM_{GLAS} / 2$$
$$H_{GLAS} - H_{LVIS} = c \cdot FWHM_{GLAS} / 2 \quad (2.5)$$

where d_{GLAS} and d_{LVIS} are GLAS and LVIS footprint diameters, $FWHM_{GLAS}$ is FWHM for GLAS. The positive correlation between (L_{ext}-RH100) and slope in Figure 2.9b is consistent with Eq. (2.5). Once slope effect is removed from GLAS waveform extent and LVIS RH100, the difference between GLAS and LVIS heights shows very little dependence on slope (Figure 2.9c). This result demonstrates the effectiveness of the slope correction used in this study.

2.5.2 GLAS and LVIS Waveform Comparison in BEF

GLAS and aggregated LVIS waveforms were analyzed under three different categories: typical, noisy and saturated waveforms. Noisy waveforms refer to those with background noise greater than 2 Digital Number (DN). Background noise is calculated as the standard deviation of DN values from non-vegetation and non-ground waveform returns. 2DN was selected as the threshold to separate lower and higher noise groups shown in the GLAS data. Saturated waveforms refer to those with sensor saturation, occurring in either vegetation returns for very dense canopies or ground returns for very sparse canopy partly due to a wrong sensor gain mode. Saturation index in GLAS products was used to determine if each GLAS waveform is saturated or not. We call the rest of the waveforms, which are acquired under more common atmospheric conditions with a proper sensor gain mode, typical waveforms. By our definition, for 31 leaf-on GLAS waveforms collected in BEF, 12 are noisy, 8 saturated and 13 typical.

2.5.2.1 Typical Waveforms

9 out of 13 GLAS waveforms matched well with aggregated LVIS waveforms over various slopes (0.2 to 14.4 degrees). Figure 2.10 shows comparison for four typical waveforms with two good matches and two mismatches. The matched waveforms occur in both relatively dense (Figure 2.10a) and multilayered (Figure 2.10b) vegetation canopies. Two types of mismatches were found in four mismatched waveforms: One type (Figure 2.10c) has two similar waveform shapes but with mismatched height, probably due to either GLAS or LVIS ground elevation error. The other type (Figure 2.10d) has two similar canopy or ground peaks but with mismatches in magnitudes. Several factors could contribute to this mismatch, such as geolocation uncertainty between datasets, LVIS' and GLAS' small off-nadir pointing effects and LVIS non-uniform sampling effects within each GLAS footprint. In forest stands with non-random spatial distribution of tree heights, the small LVIS off-nadir pointing angle can cause some waveform change, compared to the near-nadir viewing GLAS data (Neuenschwander et al., 2008). We found that the mismatches caused by the above sources are usually not significant, and overall most typical GLAS waveforms match well with the aggregated LVIS waveforms.

2.5.2.2 Saturated Waveforms

Figure 2.11 shows the same comparison as Figure 2.10 except for saturated GLAS waveforms at ground (Figure 2.11a and Figure 2.11b) and at canopy (Figure 2.11c and Figure 2.11d). For canopy-saturated waveforms, GLAS and LVIS waveform shapes are generally in good agreement, except for slight differences in peak return magnitudes.

This indicates that saturation in canopy might not be strong and these canopy-saturated waveforms are still good for extracting vegetation structure.

Ground-saturated waveforms (Figure 2.11a and Figure 2.11b) show very strong ground returns and weak canopy returns from both GLAS and LVIS data, indicating these sites are sparsely vegetated regions. Ground-saturated GLAS waveforms are consistently located several meters above LVIS data, which could be caused by underestimation in GLAS ground elevation or ignoring surface roughness effects in our analytical estimation of ground peak in GLAS waveforms. Comparisons for canopy-saturated waveforms (Figure 2.11c and Figure 2.11d) show little height differences between GLAS and LVIS waveforms. Ground-saturated waveforms show weaker ground peak returns in GLAS and some mismatch in canopy returns (Figure 2.11a and Figure 2.11b). Saturation in GLAS waveforms results in weaker ground peaks when comparing to LVIS waveforms. The mismatches in canopy returns could be the result of mismatches in LVIS and GLAS geolocation and uneven LVIS samplings in each GLAS waveforms. In summary, GLAS waveforms can saturate at canopy in dense canopy or at ground in sparse canopy, and ground saturation have slightly higher impact on GLAS data quality.

2.5.2.3 Noisy Waveforms

Figure 2.12 shows the same kind of comparison as in Figure 2.10 except for noisy GLAS waveforms. In general, GLAS and LVIS waveform shapes match well; however, GLAS waveform extent can be reduced due to larger threshold value used to determine top and bottom of the waveform extent. By examining the relationship between GLAS waveforms and their extent (shaded area) shown in Figure 2.12, we found that the

waveform extent is usually a few meters shorter than what the visually interpreted waveform range should be.

The degree of waveform extent underestimation is depending on waveform shape, or vegetation structure, slope and sensor view angle. When both canopy and ground returns are strong (Figure 2.12a), waveform extent underestimation is very small. However, significant underestimation can occurred in ground (Figure 2.12b and Figure 2.12d) or canopy (Figure 2.12c) when ground or canopy returns are weak. This indicates that waveform extent determined by the simple threshold method can be problematic in noisy waveforms especially when either canopy or ground returns are weak; therefore, our vegetation height method should be used with caution for noisy waveforms.

2.5.2.4 Evaluation of GLAS Vegetation Height

Figure 2.13 compares GLAS and mean of top three LVIS vegetation heights for all BEF data. The comparison was categorized into two groups: typical and saturated waveforms and noisy waveforms. Figure 2.13 shows that LVIS mean-top-three-heights match well with GLAS heights for typical and saturated waveforms (R^2 = 0.8 and RMSE = 2.4 m). The result is similar to the earlier GLAS and small-footprint lidar height comparison (Figure 2.8a), which covers a similar area. However, noisy waveforms often result in underestimating waveform extent and vegetation height. Height underestimation is usually about a few meters, but can be up to 15 m for very sparse and dense forests, in which identifying first and last return is particularly sensitive to the threshold.

2.6 Discussion

The most unique part of our height retrieval scheme is that slope is corrected using an analytical formula with auxiliary slope data. The same correction will be applied for the same slopes no matter what the vegetation condition is. Other GLAS height estimation schemes, for example statistical approaches, have their own advantages as slope information is estimated from waveform shape parameters and calibration coefficients and could be more accurate for specific sites. Nonetheless, recalibration is likely a necessary process as topography effects on waveform shape are affected by local vegetation conditions and sensor view angle (Chen, 2010b; Neuenschwander et al.; 2008; and Yang et al., 2010). Our physical approach provides a systematical analysis to quantify the specific error sources. Such a physical approach to retrieve vegetation height from lidar over slope terrains offers an important compliment to current retrieval methods and a significant step to map global vegetation structure from future lidar missions.

The slope-corrected vegetation height scheme presented in this study is different from other height retrieval schemes such as Lorey's height (Lefsky et al., 2007) or crown-weighted height (Pang et al., 2008), which consider the canopy distribution on height estimate. Our slope-based scheme is to retrieve maximum vegetation height, which is similar to RH100 in LVIS data. Maximum canopy height might be less related to biomass than other height metrics such as Lorey's height, crown-weighted height or RH75, RH50 derived from LVIS data. However, maximum canopy height is a critical structure input for surface roughness estimate in land surface biophysical models.

Retrieving other vegetation height metrics is more complicated and will be further investigated in our future work.

Our results have demonstrated that first-order topography effects on lidar heights can be removed analytically using the slope-correction scheme. However, its height retrieval accuracy is affected by several factors. One of the largest uncertainties might come from unevenly distributed vegetation. Our scheme was developed with the assumption that similar vegetation structure is evenly distributed within each footprint. This assumption may not be valid when vegetation distribution has a large variation in height and unevenly distributed over a slope.

GLAS, LVIS and small-footprint lidar height comparison results presented in this study show that slope-corrected LVIS height has larger RSME (3.0 m) than slope-corrected GLAS height (with RMSE = 2.2m) when compared to small-footprint lidar height. This result might indicate that our algorithm works better for large footprint lidar (ICESat/GLAS) than medium-footprint lidar (LVIS). However, uncertainty in *RH100* estimates over sloped terrains might cause larger uncertainty in LVIS height retrieval presented in this study (see discussion before). Further work is necessary to use both the full waveform extents from LVIS and GLAS data for our slope correction scheme in order to fully validate our conclusion.

Non-uniform spatial vegetation distribution might also contribute to the larger uncertainty in the two large footprint lidar height comparison (GLAS and LVIS, RMSE = 3.0 m) than in the GLAS and small-footprint lidar height comparison (RMSE = 2.2m), particularly when large variation of tree heights are on slopes. For example, lidar return from top canopy near the edge of GLAS footprint might not be detectable because laser

energy is much lower than it is in the center. But the same canopy can be detected by small-footprint lidar or LVIS footprint. We expect the sampling bias to be larger, thus larger RSME when GLAS is compared with LVIS than with small-footprint lidar because small-footprint lidar has higher spatial samplings density.

In addition, accurately quantifying surface roughness effects on waveforms can also reduce uncertainties on lidar heights retrieval when ground peaks are questionable. It is not trivial to quantify surface roughness effects because of the lack of knowledge of detailed surface elevation characteristics.

To fully quantify lidar height distortion on slope, both topography information (slope and aspect) and sensor zenith and azimuth angle are needed (Yang et al., 2010). We chose to use the simpler nadir correction scheme because the off-nadir angles are small and lack of accurate sensor zenith and azimuth angles and slope orientation. In our study region, the estimated off-nadir pointing angles are about 3-5 degrees for LVIS and 0-2 degrees for GLAS, resulting in height error ranging from 0.1 to 1 m for 2-degree off-nadir angle in GLAS and from 0.2 to 1.5 m for 5-degree off-nadir pointing angle in LVIS (estimated using equations in Yang et al., 2010). Furthermore, the uncertainty caused by the 3 degree accuracy of NED slope dataset is about 0.55 m for LVIS RH100 and 2.7 m for GLAS waveform extent (at typical GLAS footprint size (50 m) in our study) at 10 degree slope. For a global application, SRTM or ASTER DEM will be used to derive slope. Our future work will assess the overall slope-corrected height accuracy when SRTM or ASTER in stead of NED slope datasets are used.

2.7 Conclusions

Large uncertainties in vegetation structure retrieved from spaceborne ICESat/GLAS laser altimetry data are caused by surface topography, background noise, and sensor saturation. In this study we used an analytical approach to retrieve vegetation height by removing slope effect from lidar returns. We evaluated slope-corrected GLAS and LVIS heights using small-footprint lidar data and field measurements in southern White Mountain National Forest, NH. We also evaluated GLAS vertical vegetation structure content by comparing GLAS and corresponding LVIS waveforms and investigated the impact of sensor noise and saturation on vegetation structure retrieval in Bartlett Experimental Forest, NH.

Yang et al. (2010) presents an analytical scheme to quantify the impacts of surface topography, footprint size, laser pulse width and surface roughness on lidar vegetation height metrics over a forest stand. The scheme was used to remove surface topography from GLAS waveform extents and LVIS RH100 using slopes derived from the National Elevation Dataset (NED). The difference between original lidar height metrics, GLAS waveform extent and LVIS RH100 indicates a strong dependence on slope. However, after slope correction using our method, GLAS and LVIS heights match reasonably well with small-footprint data and field data with RMSE = 2.2 m for GLAS and 3 m for LVIS. Our analysis shows that the combined effect of surface topography and footprint size can be accurately removed from GLAS waveform extent and LVIS RH100. Slope-corrected GLAS heights match well with slope-corrected LVIS heights with RMSE = 3.7 m. The corrected GLAS and LVIS vegetation height differences are independent of slope, indicating the topography correction is effective.

GLAS and aggregated LVIS waveforms and vegetation height comparisons in Bartlett Experimental Forest, NH show that GLAS waveforms and heights matched well with LVIS for most waveforms without sensor saturation and with low background noise. Saturated waveforms show reasonably good match between GLAS and LVIS waveforms, indicating GLAS saturated waveforms should still capture vertical vegetation structure well. GLAS waveforms with large noise often mismatch LVIS waveforms resulting in underestimation of waveform extent and vegetation height

Applying our slope correction scheme outside of U.S. or at global scale requires surface slope information from alternative sources. It is possible to characterize large-scale slope information from the global canopy-adjusted surface elevation data generated from SRTM, owing the elevation bias is systematic and statistically insignificant. Our initial study in New England region shows that SRTM does not provide good quality elevation data, however the slope from SRTM matches reasonably well with NED slope (Mean error ~0.9 degree, RMSE ~3 degree). With the acceptable slopes from SRTM, our correction scheme can be applied to global scale after further evaluation in other regions and biomes.

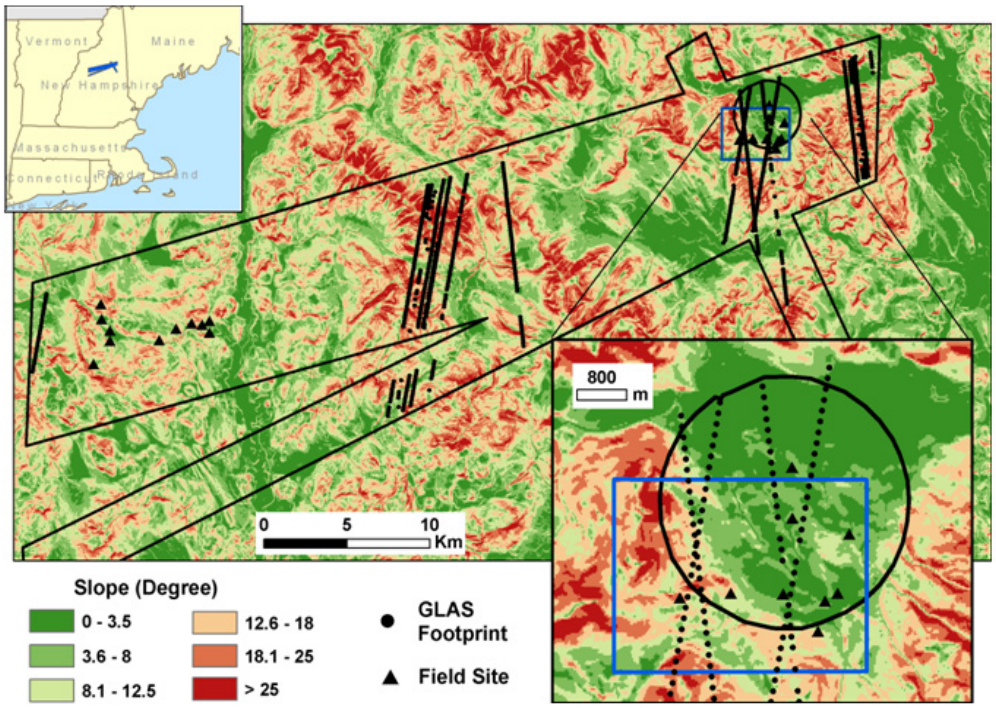

Figure 2.1: Slope map derived from National Elevation Data (NED) overlaid with GLAS footprints (thick black dots), field sample sites (triangulars), small-footprint discrete-return lidar sampling area (blue rectangle) and LVIS data extent (parallelograms) in southern White Mountain National Forests, NH. Circle indicates the location of Bartlett Experimental Forest where full LVIS waveform data are available.

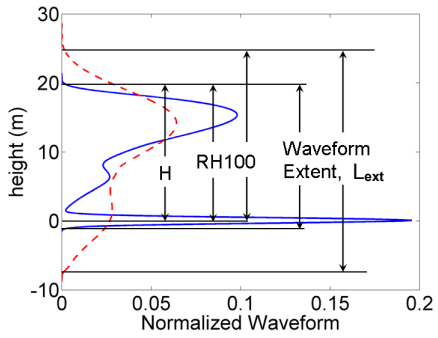

Figure 2.2: Schematic definitions of lidar-derived vegetation height (H), RH100, and waveform extent (L_{ext}) for one waveform on a flat terrain (solid line) and another over a slope (dash line).

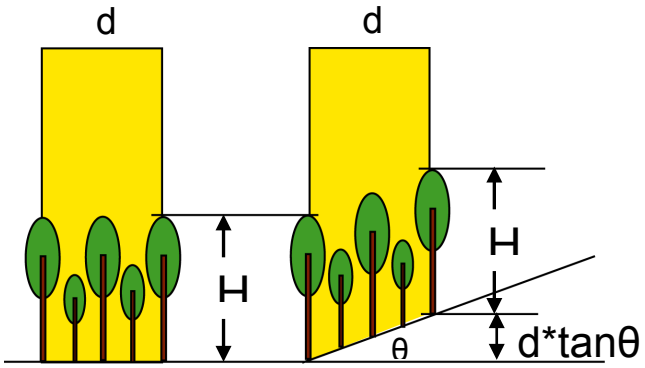

Figure 2.3: Impact of slope on lidar height retrieval. H is lidar height over a flat surface, d is lidar footprint diameter, and θ is the slope. The distance from canopy top to bottom on a slope is H + d*tan θ.

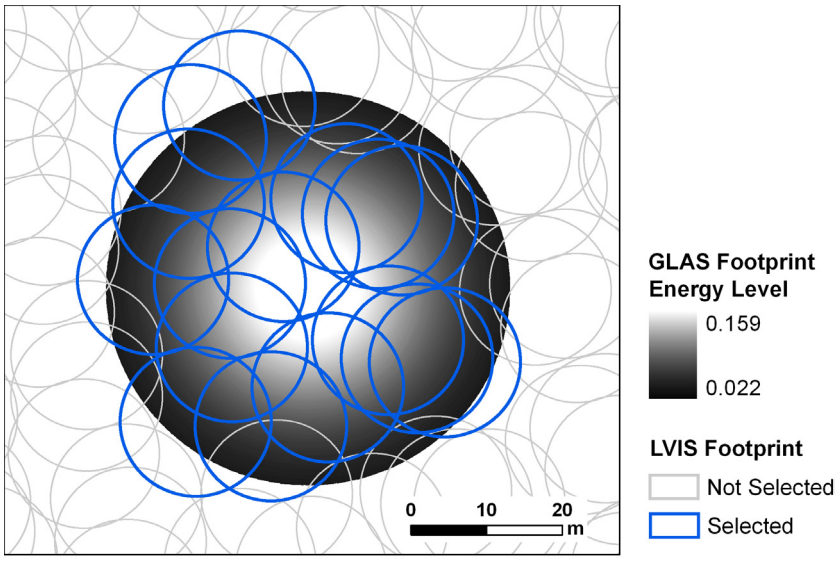

Figure 2.4: Scheme of to aggregate LVIS waveforms (thicker circles) within one GLAS footprint as a summation of all LVIS waveforms weighted by the normalized GLAS lidar energy intensity (shaded) ranging from 0.159 from the center to 0.022 at the edge).

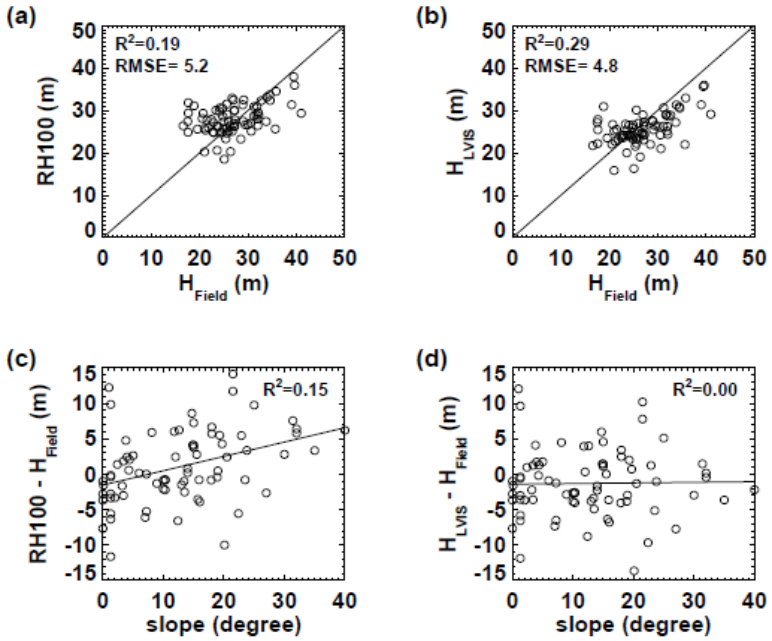

Figure 2.5: Comparison of a) LVIS RH100 with field measured maximum height (H_{field}) ; b) H_{LVIS} (with slope correction) with field measured maximum height (H_{field}) ; the dependence of c) LVIS height residuals, (RH100 - H_{field}) on slope and; d) (H_{LVIS} - H_{field}) on slope.

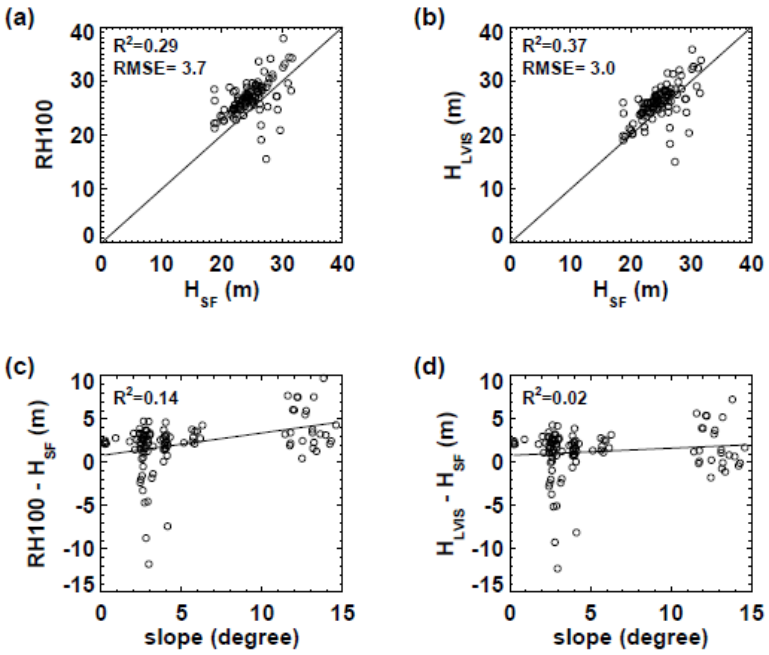

Figure 2.6: The same as Figure 2.5 except for comparison with 98 percentile of small-footprint lidar height (H_{SF}).

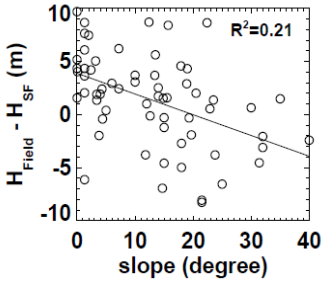

Figure 2.7: The relationship of height differences between the 98 percentile of small-footprint lidar height (H_{SF}) and field measured maximum height (H_{field}) with slope

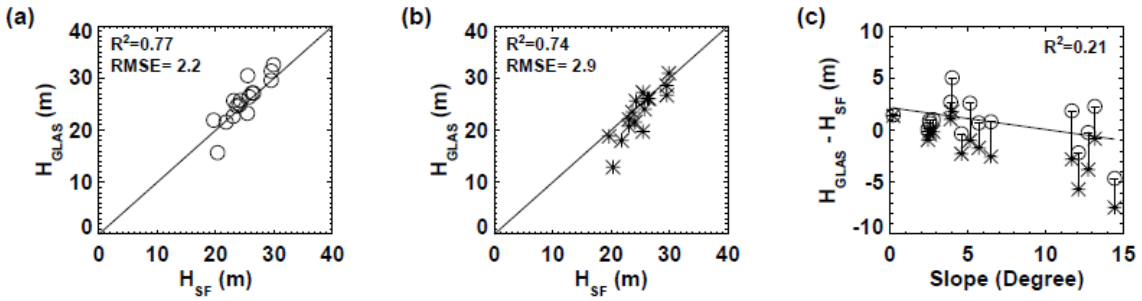

Figure 2.8: Comparison of slope-corrected GLAS heights (H_{GLAS}) with 98 percentile of small-footprint lidar height (H_{SF}) a) with the correction of waveform extension and b) without the correction) and c) the dependence of residuals of slope-corrected GLAS heights, (H_{GLAS}-H_{SF}) on slope with trend. Two GLAS heights in c): circles with the correction of waveform extent and stars without the correction.

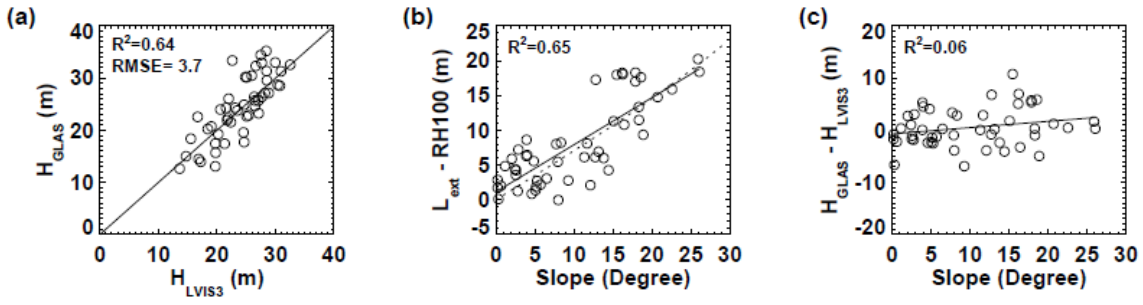

Figure 2.9: a) Slope-corrected GLAS and LVIS height comparison (a); b) the dependence of (L_{ext} – RH100) on slope and; c) (H_{GLAS}-H_{LVIS3}) on slope.

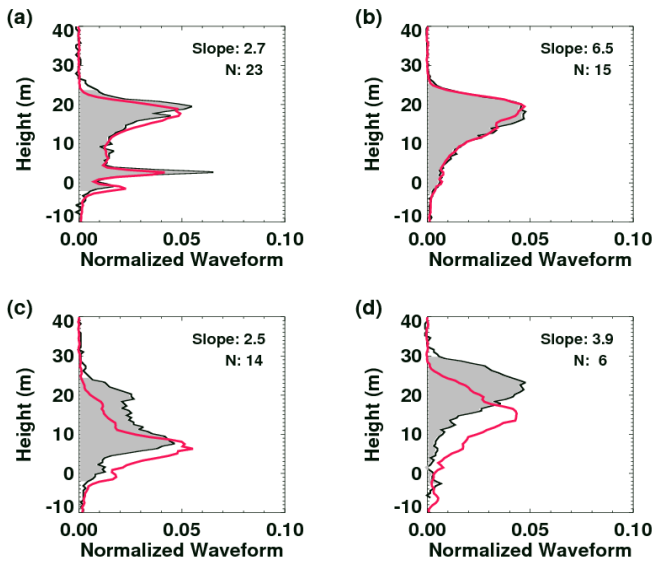

Figure 2.10: Comparison of aggregated LVIS (thick curve) and typical GLAS (shaded) waveforms with two well-matched (a and b), one matched in height but in magnitude (c), and one with similar waveform shape but mismatched height (d). Slope is the mean slope in GLAS footprint, N is the number of LVIS waveform used for aggregation.

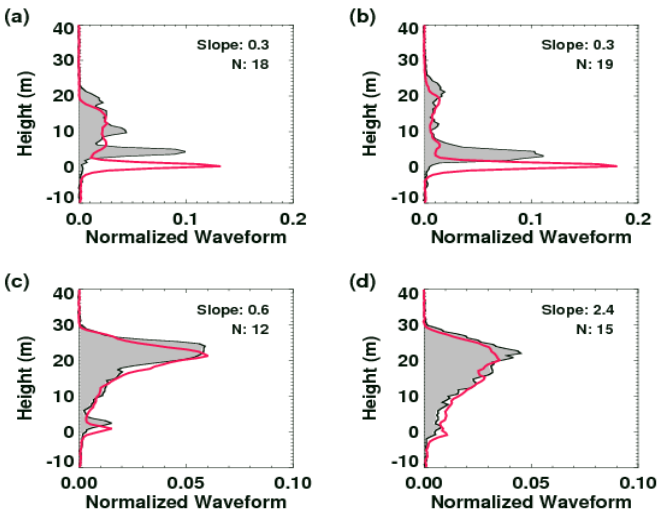

Figure 2.11: Comparison of LVIS and saturated GLAS waveforms with ground saturated (a and b) and canopy saturated (c and d).

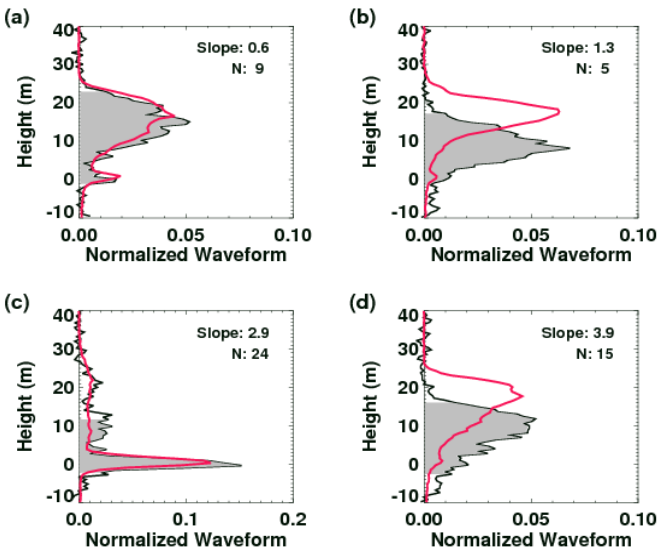

Figure 2.12: Comparison of LVIS and noisy GLAS waveforms.

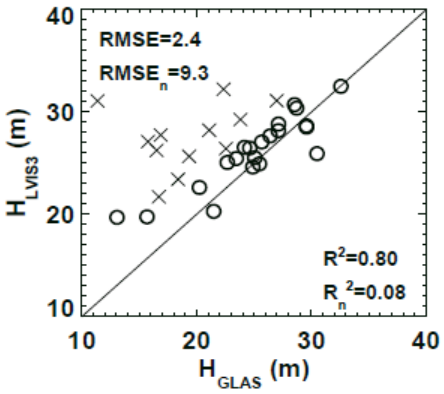

Figure 2.13: GLAS and LVIS height comparison in Bartlett, NH for typical and saturated (circle) and noisy (X) GLAS waveforms. R^2 and RMSE for typical and saturated waveforms and R_n^2 and $RMSE_n$ for noisy waveforms.

Chapter 3 Assessing General Relationships between Above-Ground Biomass and Vegetation Structure Parameters for Improved Carbon Estimate from Lidar Remote Sensing

3.1 Introduction

Lidar remote sensing provides measurements of horizontal and vertical vegetation structure of ecosystems. This information will be critical for estimating global carbon storage and assessing ecosystem response to climate change and natural and anthropogenic disturbances. Unlike many other remote sensing measurements, lidars provide direct and indirect measurements of vegetation structure which can be used to estimate global carbon storage (Dubayah and Drake, 2000). Recent advances in lidar technology have made lidar data widely available to study vegetation structure characteristics and forest biomass. The spaceborne Geoscience Laser Altimeter System (GLAS), part of the ICESat mission, provides global lidar data with a variable diameter of ~70m footprint spaced at ~170 m (Zwally et al. 2002; Harding et al. 2005; Lefsky et al. 2005). Airborne data have also been collected using a Scanning Lidar Imager of Canopies by Echo Recovery (SLICER) with a 15m footprint and the Laser Vegetation Imaging Sensor (LVIS) with a 20m/25m footprint over several large areas for improved vegetation structure characterization (Blair et al. 1999). Small-footprint multiple return

lidar data have also been collected in many regions of the globe (Jupp et al. 2005) and more recently small footprint scanning waveform systems have become operational (Gutierrez et al. 2005; Neuenschwander et al., 2008). Many studies have demonstrated the potential use of spaceborne and airborne lidar data to estimate vegetation height, above-ground biomass characteristics, and other vegetation structure parameters (Simarc et al, 2008; Lefsky et al., 1999, 2002, 2005; 2007; Harding et al., 2001, 2005; Drake et al., 2002a, 2002b, 2003; Nelson et al., 1984, 1988, 1997, 2009; Patenaude et al., 2004). The National Decadal Survey report (NRC, 2007) also recommended two new lidar missions - the Ice Cloud and Land Elevation Satellite-II (ICESat-II) and the Deformation, Ecosystem Structure, and Dynamics of Ice (DESDynI) to measure the horizontal and vertical structure of ecosystems for estimating global carbon storage and ecosystem response to climate change and human land use. Both the ICESat-II and DESDynI missions expect to provide measurements of vegetation structure to estimate above-ground biomass and carbon stocks with greatly reduced uncertainties.

Carbon stocks or above-ground biomass are not directly measured by lidar. But many studies have demonstrated the strong relationship between above-ground biomass and lidar measured tree height, ranging from boreal conifers to equatorial rainforests (Lefsky et al. 2005, Drake et al. 2002a; 200b). These relationships are used to derive above-ground biomass from lidar-measured vegetation height (Lefsky et al., 2005) or accumulated vegetation lidar returns (Drake et al. 2002a; 2002b; 2003) at large scales.

However, large uncertainties exist in the large-scale above-ground biomass estimates from lidar. Forest above-ground biomass is related to several woody structure parameters, including trunk diameter at breast height, height of canopy, stem density, and

branch distribution, but height is the only woody structure parameter directly measured by lidar. Above-ground biomass was indirectly derived based on empirical relationships with lidar measured vegetation height or accumulated lidar returns from vegetation. Often these relationships are site-dependent and lead to large uncertainties when applied over large regions, indicating that height may not be the only structure parameter relating to biomass. Drake et al. (2002a and 2002b) found that the height of medium energy returns (RH50) is better related to above-ground biomass than the height at full energy returns (RH100) due to RH50 being more sensitive to changes in both the vertical arrangement of canopy elements and the degree of canopy openness (including tree density) than height itself. Further investigation is required to have a clear understanding of the links between above-ground biomass, vegetation structure parameters measured from field and from lidar in order to develop a more physically based approach in order to improve above-ground biomass estimates from lidar. This will also help us to better understand the above-ground biomass retrieval accuracy from lidar for the ICESat-II and DESDynI missions. This study is a first step into this direction, using field measurements of vegetation structure and the airborne LVIS and hemispherical scanning ground lidar EVI data collected in the New England region in 2003 and 2007.

3.2 Review of Above-Ground Biomass Estimate

Two approaches are commonly used to estimate above ground biomass. One is allometric approach to estimate biomass based on a given tree diameter at breast height (DBH) and the other one is using woody volume to calculate biomass (see Brown, 2002 for an extensive review).

The most common approach to estimate above-ground biomass is using allometric equations to calculate above-ground tree biomass (AGB) to a given tree diameter value (D), usually measured at 1.3 m,

$$AGB = aD^b \tag{3.1}$$

where *a* and *b* are scaling coefficient and scaling exponent, respectively, which vary with species, sites and age. Usually *a* and *b* are obtained through empirical regression of log-transformed data of biomass and diameter pairs measured from destructively sampled trees. For example, Jenkin et al. (2004) and Ter-Mikaelian and Korzukhin (1997) list *a* and *b* values from species in the US. This approach is laborious and time-consuming.

Many studies have investigated how *a* and *b* are related to stand structure and age. West et al. (1999) used the fractal properties of tree branching networks and developed a geometric model of tree structure which predicts above-ground biomass from tree diameter and estimates the exponential factor *b=8/3=2.67* independently of species, site and age. Enquist et al. (1999) suggested that *a* is related to wood density, however Chambers et al., (2001) found *a=0.1*.

Many other studies found that *a* and *b* varies by species, sites, and even stand ages. Ketterings et al (2001) found that *a* and *b* vary between sites and suggested that *b* can be estimated from the site-specific relationship between height and diameter, $H = kD^{b-2}$ and $a = r\rho$, where r is expected to be relatively stable across sites and ρ is the wood density. Zianis and Mencuccini (2004) used fractal geometry and found that the scaling exponent *b* is between *2* and *3* and should be obtained based on tree height/size (or stand age). Their result is consistent with the finding by Ketterings et al (2001) that *b* is related to tree height. Their study also found a statistically significant difference between

theoretical and empirical values of the allometric exponent (*b=2.3679 vs b=2.67*). Pilli et al. (2006) analyzed 49 datasets of different species and found that *b* is also related to tree stand age, species and site and *a* is correlated to wood density. All these studies indicates that tree height/age based allometric equations are possible, This result may have a great implication for deriving tree height and stand age based allometric equations for biomass estimate from lidar.

The allometric approach to estimate above-ground biomass requires measurement of the diameter for each individual trees and the allometric equation for each individual tree species. Developing allometric equations for each individual species can be very difficult particularly in tropical forests. Chave et al. (2005) used a single pan-tropical woody volume-based approach to estimate above-ground biomass at plot level. The basic hypothesis is that above-ground biomass **for each tree** is proportional to woody volume (product of height and basal area) (Chave et al., 2005):

$$AGB = F \cdot (\rho \cdot (\pi D^2/4) \cdot H)^{\beta} \qquad (3.2)$$

where H is tree height, ρ is the wood density (dry weight per unit volume in g/cm^3) and $\beta<1$. F is a constant factor depending on tree taper. For broadleaf species, F=0.06 (Chave et al., 2005). Their study demonstrates that including wood density and woody volume improves biomass estimates.

However the woody volume-based approach is not often used due the difficulty to obtain height measurements for each individual tree. One way to overcome this problem is to use allometric equation, $H \approx D^B$, to calculate height from trunk diameter. However, in reality a power-law is not the best relationship for predicting height from diameter, and often lead to biased estimate of above-ground biomass equations (Chave et al., 2005).

Woody volume-based approach has been proved an effective approach to study the parametric values of diameter-based allometric equations (Pilli et al. 2006; Zianis and Mencuccini, 2004) and to estimate above-ground biomass at plot level (Chave et al., 2005). This study uses this approach to revisit the relationship between above ground biomass and vegetation structure parameters at different scales and use these relationships to further understand the relations between biomass and lidar structure parameters.

3.3 Study Sites

This study used data collected in temperate forests in New England. Our study sites include many plots in Harvard Forest (HF), MA, Bartlett Experimental Forest (BEF), NH, and the Forest Ecosystem Research site in Howland, ME (see Figure 3.1). HF is a 60-70 yr old mixed deciduous forest. The stand is in the transition hardwoods-white pine-hemlock zone (Spurr, 1956), and is comprised mainly of red oak, red maple, yellow birch, white birch, beech, white pine, and hemlock. BEF includes old-growth northern hardwoods with beech, yellow birch, sugar maple, and eastern hemlock. The natural stands in Howland, ME, is the boreal-northern hardwood transitional forest and consists of spruce-hemlock-fir, aspen-birch, and hemlock-hardwood mixtures. Our study sites cover large forested regions with various climate conditions and should be a good representation of the forests in the northeastern USA.

3.4 Data Sets

Data sources used in this study include: vegetation structure field data collected in 2003 and 2007, LVIS data collected in 2003 and ground lidar data collected in 2007.

There is a four year difference between these datasets. We observed that growth, for example, in sparse shelterwood site in Howland, Maine, plays a role in LVIS and modeled waveform comparison (Yang et al., 2009). However DBH comparison in Howland, MN from 1989 to 2003 does not show much growth. LVIS and EVI height comparison does not show impact of growth. Growth was ignored in this study.

Two sets of ground tree structure field data were used in this study. The first includes 6 stands (28 plots with a 20m/25m radius) of tree data collected in summer, 2007 in three forest sites (Harvard Forest, MA, Bartlett Experimental Forest, NH, and the Forest Ecosystem Research site in Howland, ME). The second is a complete stand map collected in 2003 in the Forest Ecosystem Research site in Howland, ME (see Figure 3.1 for the locations of our study plots and the stem map area). These two datasets were collected over relatively flat terrains.

3.4.1 2007 Field Data

The 2007 ground data include vegetation structure data collected in six stands with two stands selected in each forest (Harvard Forest, MA, Bartlett Experimental Forest, NH, and the Forest Ecosystem Research site in Howland, ME) and each stand with unique canopy cover composition commonly seen in this region, e.g. deciduous forest dominated, conifer forest dominated, or partially harvested forest. Our sites are purposely selected to encompass most forest cover types in New England. Within each stand, tree data were collected in 5 plots (except only 3 plots near the tower site in Howland). Each is a circle area with a 20m radius (except for 25m in the hardwood site in Harvard Forest due to lower tree density). One plot is at the stand center, the rest at 30-50m away from the center at each of the 90 degree angles (see the bottom right square in Figure 3.1).

Diameter at Breast Height (DBH), tree species, tree status, and crown status were sampled for trees with DBH >=10cm, smaller trees (DBH between 3-10cm) were measured within 10m of plot center. In addition, 10 trees were selected in each plot for additional tree height and crown size measurements.

3.4.2 2003 Stem Map Data

The 2003 stem map in the Forest Ecosystem Research site in Howland, ME is a 200mx150m rectangular area (except for a 30mx30m plot) near the experimental forest's flux tower (see Figure 3.1). Tree species, DBH, tree status (alive, dead), and crown status (dominance of the tree crown) were sampled for a total amount of ~7800 trees in 2003. Similar data collected in 1989 were used to verify our tree height allometric equations. Instead of using the originally designed rectangular plot for 2003 data, a circular plot (20m in radius) was used to be consistent with the other dataset (see the top right square in Figure 3.1).

3.5 Lidar Data

Two sets of lidar data collected in our study sites were used in our analysis. One is the LVIS, or the "Laser Vegetation Imaging Sensor" and the other is a hemispherical scanning ground lidar, the Echidna® validation instrument (EVI) (Jupp et al. 2005; 2009; Strahler et al. 2008).

3.5.1 Airborne Lidar Data

In summer 2003, LVIS was flown over several regions in New England (Blair et al., 2006). The LVIS data used in this study were acquired on July 18-20, 2003 in Bartlett, NH, July 26, 2003 in Howland, ME, and July 20, 2003 in Harvard, MA. The flights

altitude are ~10 km, producing footprint size of ~ 20 m, separated by roughly the same distance both along and across track.

LVIS is an airborne laser altimeter system, or lidar, designed, developed and operated by the Laser Remote Sensing Laboratory at NASA's Goddard Space Flight Center. The onboard laser generates Gaussian shaped optical pulses at a wavelength of 1064nm (Blair et al., 1999). The vertical sampling resolution of LVIS is 30cm (1ns). LVIS footprint sizes (diameter) typically vary between 10 to 25 m depending on the mission flight altitude. To be consistent with other datasets, the vegetation structure parameters derived from LVIS were also binned in each plot where ground structure data were collected. LVIS has a published geolocation accuracy of 1 m or less for this dataset (Blair et al., 2006). Our analysis indicates that LVIS geolocation data match very well with IKONOS data.

3.5.2 Ground Lidar Data

Concurrently with the ground structure data collection, a hemispherical scanning ground lidar, the Echidna® validation instrument (EVI) (Jupp et al. 2005; 2009; Strahler et al. 2008) was deployed in the same sites as our ground data collection. One set of EVI scan data were collected for each plot in each stand in 2007. EVI, developed by CSIRO Australia as part of its canopy lidar initiative, is a ground-based, upward hemispherical-scanning, full waveform digitized, terrestrial lidar instrument and allows acquisition of vegetation canopy structure data, including height, basal area, and stem counts, as well as accurate information on standing woody and green biomass for carbon balance inventory and mapping (Jupp et al. 2007; 2005; Strahler et al. 2008). EVI utilizes a horizontally positioned laser that emits pulses of near-infrared light at a wavelength of 1064nm.

Pulses are emitted at a rate of 2 kHz. Plot and stand level DBH, basal area, tree height in addition to foliage profiles were derived from each plot in our study sites using the EVI data (Tian et al., 2009; Zhao et al., 2009).

3.6 Field Data Analysis

Based on the vegetation structure parameters collected in 2003 and 2007, we examined the relationship between above-ground biomass for each individual tree and at plot level with height, DBH, basal area and woody volume. Above-ground biomass for each individual tree was calculated based on DBH-based allometric equations and plot level biomass was calculated as the sum of individual tree biomass within each plot. Vegetation height was not measured for each single tree in the field; however height was estimated from DBH using DBH-height allometric equations and was evaluated using height and DBH measurements from the same site.

The allometric equations for biomass were chosen from the literature where ground data collection best approximates the conditions in our study sites (Tritton et. al., 1982):

$$\ln(AGB) = a + b \ln D \tag{3.3}$$

where the coefficients a and b vary with species. The allometric equations for height were from Albani, et. al.(2001):

$$H = 1.3 + b_{1k} \cdot (1 - e^{b_{2h} D}) \tag{3.4}$$

where coefficients b_{1h} and b_{2h} vary with plant functional types (PFTs), which are differentiated by their leaf physiology, allometry, mortality and dispersal. We verified that the above allometric equation provides good estimates for tree height for our study

area by comparing tree height estimates with the tree height estimates in 1989 (R=0.87, RMSE=2.3m, for nearly 7800 trees).

3.6.1 Results for Individual Tree-Species-Based Analysis

Figure 3.2 compares the relationship between above-ground biomass with woody volume approximated as the product of basal area and tree height for major tree species in New England. The analysis was separated by two groups: conifer/softwood tree species and deciduous /hardwood tree species. Figure 3.2 shows that for trees within 10 to 68 cm DBH range, above-ground biomass is almost linearly related to the woody volume except for birch species. Previous studies on allometric equations have shown that the coefficients are dependent on species, possibly on stand age and structure (Pilli et al. 2006; Zianis and Mencuccini, 2004). As indicated in Eq. (2.2), the coefficients (slopes) of the linear relationship between above-ground biomass and woody volume are related to the wood density. Softwood and hardwood show a large difference in wood density, resulting in two distinguished groups as shown in Figure 3.2. For all hardwood tree species in our study area, except for birch, the biomass does not vary much between individual species. For all softwood tree species, except for northern white-cedar tree, biomass estimates all group together. We concluded that the above-ground biomass and woody volume relationships are vegetation type dependent. This result is consistent with the conclusion from Chave et al. (2005) that knowing wood density improves above-ground biomass estimates.

3.6.2 Results for Plot Level Analysis

Plot level analysis was based on two datasets collected in 2003 and 2007. These two datasets have similar geolocation accuracy, topography and sampling strategy. Data collected in 2003 were binned into 4 x 6 circular plots with 20m radius, the same plot size as for the 2007 data. Both datasets have about 5m geolocation accuracy.

Figure 3.3 compares the relationships among plot level above-ground biomass with various tree structure parameters including quadratic mean DBH (DBH_qm), basal area, quadratic mean tree height (H_qm), top 10% tree height, and woody volume (the product of basal area and tree height) for data collected in 2003 and 2007. It shows that plot level above-ground biomass is related to height metrics and to mean DBH structure parameters with root mean square error (RMSE) ranging around 1.1-4kg/m^2 and coefficient of determination R^2 ranging from 0.4 to 0.7. However there are five outliers in biomass and height/DBH relationships (Figures 3.3a, 3.3c and 3.3d). These outliers are from the shelterwood stand in Howland, MN, where vegetation is very sparse. Quadratic mean of DBH (DBH_qm), quadratic mean tree height (H_qm) and top 10% tree height were large, however total above-ground biomass was low in these sites.

Over all structure parameters, woody volume (calculated as the product of basal area and top 10% or quadratic mean of tree height) is the best biomass predictor when the analysis was based on vegetation type with RMSE from 1.1kg/m^2 to 1.6kg/m^2 and R^2 from 0.74 to 0.95. Conifer plots show a better relationship (R^2=0.95) between biomass and woody volume than deciduous (R^2= 0.74). Overall woody volume is a good predictor of plot level above-ground biomass when the forest type is known. The method used in this study is different from Chave et al. (2005) in that their study developed a

regression model to relate above-ground biomass with diameter and height for each individual tree at a plot level. Our study used basal area and top 10% tree height at the plot level which is more closely related to lidar measured tree height. Our conclusion is consistent with theirs in that woody density and woody volume a good estimate of biomass and our study also indicates that distinguishing hardwood and softwood vegetation type and woody volume are critical for biomass estimate.

3.7 LVIS Data Analysis

Lidar data analysis focuses on examining the physical meaning of vegetation height metrics and the link between plot level above-ground biomass with height metrics, vegetation cover and structure parameters combing height and gap fraction. This analysis will aid us to understand the error sources in lidar estimated biomass and investigate how much improvement in biomass estimate we can expect by adding additional vegetation structure parameter compared to using height metrics alone.

3.7.1 The LVIS Data Processing

To extract vegetation structure parameters for each plot from LVIS data, LVIS footprint-level structure parameters within each plot were binned to obtain plot level LVIS structure parameters. Figure 3.4 shows the spatial relationship between LVIS footprint and the plot size. To extract LVIS data for the corresponding plot, we averaged all LVIS data occupying an area slightly larger than plot range to reduce geolocation uncertainty. LVIS data (height and coverage) with footprint centers located within a 5m radius circle centered at each plot center are averaged (using quadratic mean for height) for all the 2003 and 2007 plots (see Figure 3.4).

Three structure parameters derived from LVIS were used in this study: The first, RH100, calculated as the distance of the top of vegetation returns and the peak of last Gaussian pulse (the ground returns); the second, RH50 as the height of medium energy occurs relative to the last ground returns (Drake et al., 2002a). Both height metrics are often used to estimate above-ground biomass (Drake et al., 2002a; 2003); and the third, vegetation cover, calculated from the full lidar waveforms based on the method described below:

Vegetation cover can be estimated from accumulated lidar returns from canopy and ground. However it is well known that the lidar returns are affected by the canopy and background reflectivity ratio. Based on the basic lidar equations (Ni-Meister et al., 2001), vegetation cover can be directly estimated from accumulated vegetation and background returns, R_v and R_g (see Figure 3.5 on how to obtain R_v and R_g from waveforms), and the ratio of canopy volume backscattering coefficient (ρ_V) and the background reflectivity (ρ_g), ρ_v / ρ_g ,:

$$C = 1 \Big/ (1 + \frac{\rho_v}{\rho_g} \frac{R_g}{R_v}) \tag{3.5}$$

Accumulated vegetation and background returns, R_v and R_g for each footprint can be easily retrieved from original lidar energy returns. For reflectivity ratio (ρ_v / ρ_g), we used values derived from previous study (Yang et al., 2010), which demonstrated that using the reflectivity adjusted waveforms can improve accuracy in lidar derived canopy gap profiles.

3.7.2 Results

Figure 3.6 compares RH100 and in-situ measured vegetation heights calculated as mean top 10%, maximum and quadratic mean tree height at LVIS footprint, plot and stand scales. At the LVIS footprint level (20m), RH100 is best represented as the maximum tree height; however there are a few outliers showing significant underestimation of maximum tree height. They are likely the results that taller tress are located near the edge of the footprint, where laser energy is weak; therefore there is not enough back scattering from the tree tops to record distinguishable returns (Hyde et al., 2005). Our result is comparable to other results by Lefsky et al. (2001) and Hyde et al. (2005).

RH100 overestimates the quadratic mean of tree height, but compares reasonably well with top 10% tree height at the footprint level. However both top 10% and the quadratic mean of tree height has much smaller variations than RH100, partly due to averaging often leading to reduced variations. The other reason could be that in-situ tree heights were not direct ground measurements, but were calculated based on allometric equations and such calculation loses its nature variation.

At the plot and stand levels, RH100 and top 10% height and quadratic mean height were all averaged values. RH100 slightly overestimates top 10% tree height, particularly for tall trees, however both variables show similar spatial variations. RH100 underestimates maximum tree height in most cases particularly for conifer sites. The underestimate is the result that RH100s at both plot and stand levels, are averaged value and maximum tree heights are not. It appears that crown shape influences RH100 estimates and in conifer forests, skinny crowns often lead to too little signal to detect the

tree tops, leading to underestimation of tree height for conifer forests. RH100 overestimates quadratic mean of tree height. The overestimation comes from that RH100 is an averaged value of taller trees and quadratic mean includes also small trees.

To better understand RH50, Figure 3.7 compares RH50 with top 10% and quadratic mean canopy height, canopy cover and canopy volume estimated as the product of vegetation height (RH100) and vegetation cover. It shows that RH50 is related well to vegetation cover, quadratic mean and top 10% averaged vegetation height with R^2 greater than 0.7 and RMSE ranging from 1.7m to 2.0m. However RH50 has the closest relationship with canopy volume (RH100* cover) with a coefficient of determination R^2 =0.92 and RMSE=1m, indicating that RH50 is a better representation of canopy volume. RH50, the height where 50% of waveform energy occurred, is sensitive to changes in canopy foliage structure, for which canopy height, density, cover and vertical foliage profiles are all determining elements. The value of RH50 is lowered in forest stand with lower canopy cover, when more energy can reach the ground (Drake et. al., 2002a; 200b). Conversely, RH50 is increased in high canopy covered forest. Figure 3.6 indicates that RH50 is about 60%-70% of canopy volume value. Our additional analysis (not showing plots here) found that the relationship of RH50 with crown volume (calculated based on crown size and stem density) and total foliage area (foliage area volume density*crown volume), appear quite complicated and crown shape may play a role in these relationships. Our understanding of why RH50 is more closely related to height*cover than to crown volume or foliage area is that RH50 is a combination of height and canopy gap fraction. A physical model like the Geometric Optical and Radiative Transfer

(GORT) (Ni-Meister et al., 2001; and Ni-Meister et al., 2009 and Yang et al. 2009b) is required for further analysis and the results will be presented in a separate study.

To investigate the relationships of plot level above-ground biomass with RH100, RH50, vegetation cover and canopy volume, Figure 3.8 compares above-ground biomass with RH100, RH50, vegetation cover and the combination of two. RH100, RH50 and vegetation cover are all good predictors of biomass. Previous study shows RH50 is a good predictor of above-ground biomass (Drake et al., 2002a and 2002b). Our LVIS data analysis shows that RH100*cover (R^2 =0.85 and RMSE=2.4 m) and RH50*cover (R^2 =0.87 and RMSE=2.2 m) perform similarly or even better than RH50 (R^2 =0.84 and RMSE=2.5m). However Figure 3.8 also indicates that in deciduous plots, lidar was not successful to identify biomass variations.

To understand why RH100*cover and RH50*cover are good predictors of above-ground biomass, Figure 3.9 compares RH100*cover and RH50*cover with woody volume (basal area * top 10% height) and it shows that both metrics are closely related to woody volume, particularly for conifer plots. For deciduous plots, RH100*cover and RH50*cover seem saturated and do not change much with woody volume, indicating that lidar might have difficult to identify biomass variation in deciduous forests, consistent with what is shown Figure 3.8. This result concurs with Nelson et al. (2007).

High correlation of RH50 with RH100*cover suggests that RH50 is directly related to canopy height and gap fraction, thus woody volume. This finding indicates that biomass should be better described by lidar measured height and canopy gap fraction than height alone. Even nadir-looking above canopy lidar does not provide woody volume information, above-ground biomass can still be estimated with good accuracy

from lidar data with structure parameters that combine height and gap fraction, such as RH50, RH100*cover and RH50*cover. However lidar may not be successful in to identify biomass variation in deciduous forests.

3.8 Analysis of Hemispherical Scanning Ground-based Lidar Data

Our ground data analysis demonstrates that woody volume (estimated as the product of vegetation height and basal area) is an excellent biomass predictor with knowing hardwood/deciduous or softwood/conifer forests. However nadir pointing lidar does not provide a direct estimate of woody volume except for tree height. One approach to estimate woody volume is using a full- digitizing hemispherical-scanning below-canopy lidar like the Echidna® validation instrument (EVI). As demonstrated in Jupp et al. (2009), Strahler et al. (2008), Tian et al. (2009) and Zhao et al. (2009), EVI provides plot level mean DBH, basal area, stem density, foliage profile and tree height structure information. Woody volume can be calculated from EVI basal area and height measurements. The purpose of this study is not repeating what exactly structure information EVI can provide as did in Jupp et al. (2009), Strahler et al. (2008), Tian et al. (2009) and Zhao et al. (2009), but to explore using EVI-basal area and height measurements in New England to estimate woody volume and then above-ground biomass at plot and stand levels. This approach will avoid using allometric equations to estimate above-ground at plot and stand levels.

Our first analysis is to compare EVI height with different field measured height metrics. Similar analysis was conducted as Figure 3.6. EVI height was compared to mean top 10%, maximum and quadratic mean tree heights at plot and stand levels (Figure

3.10). Similar to RH100, EVI height at both plot and stand levels matches reasonably well with top 10% tree height with slight overestimation. EVI height slightly underestimates maximum tree height at both plot and stand levels. However the difference from LVIS RH100 is that EVI height accuracy is less crown shape dependent as it is upward-looking and crown shape may not play a strong role on height estimate. Similar to RH100, EVI height overestimates quadratic means of tree height.

Further analysis includes comparison of EVI and LVIS canopy height, and EVI DBH and basal area with field data (see top two panels in Figure 3.11). EVI-height agrees reasonably well with LVIS height with slightly larger LVIS height. EVI basal area matches well with the field measurements.

Finally, above-ground biomass was compared to woody volume estimated by EVI or a fusion of EVI and LVIS at the plot level (middle two panels in Figure 3.11) and the stand level (bottom two panels in Figure 3.11). Figure 3.11 demonstrates that plot level above-ground biomass is well related to woody volume particularly for conifer plots. Although with low samplings, above-ground biomass was better estimated at the stand level for both conifer and deciduous forests (R^2 = 0.97 and RMSE = 0.3 kg/m^2 for conifer stand and with R^2 = 0.92 and RMSE = 1.6 kg/m^2 for deciduous stand) than at the plot level. Our analysis demonstrates that hemispherical-scanning ground lidar like EVI provides high quality estimates of woody volume, and thus accurate estimates of above-ground biomass. This analysis demonstrates the potential strength in biomass estimate at large scale using EVI.

3.9 Conclusions and Discussion

This study analyzes both field and lidar-based vegetation structure measurements collected in New England at different scales to investigate if additional vegetation structure parameters besides height are highly related to above-ground biomass and will improve above-ground biomass estimate. Field data analysis shows that above-ground biomass for each individual tree and at the plot level is closely related to woody volume and wood density, depending on vegetation type (conifer/softwood or deciduous/hardwood forests). Woody volume at plot level is best approximated by the product of basal area and top 10% tree height. Our study indicates that using woody volume and vegetation type rather than tree height alone has the potential to improve biomass estimates.

This result concurs with previous allometric equation-based biomass studies by Pilli et al. (2006); Zianis and Mencuccini, (2004) and Chave et al. (2005). Their studies found that above-ground biomass is closely related to woody volume and wood density. However our analysis method is different from theirs. Previous analysis calculated the statistics based on individual tree. Our analysis calculates the statistical tree structure parameters first and relates above ground biomass with tree structure statistical parameters at large scales. For example, our woody volume at plot level was approximated as the product of mean basal area and top 10% tree height. A similar study at plot level (Chave et al. 2005) did regression analysis to relate biomass with DBH and height for each individual tree.

Previous airborne lidar analysis found that RH50 is a good predictor of above-ground biomass (Drake et al., 2002a; 2003). Our LVIS data analysis shows that structure

parameters combining canopy height and gap fraction such as RH100*cover and RH50*cover perform similarly or even better than RH50. High correlations between these structure parameters with woody volume, particularly for conifer forests explain why they are good predictors of above-ground biomass. High correlation between RH50 RH100*cover suggests RH50 is a better predictor of biomass than RH100. Above-canopy lidar does not provide a direct measure of woody volume, however structure parameters that combine height and gap fraction, such as RH50, RH100*cover and RH50*cover are highly related to woody volume, above-ground biomass can therefore still be estimated with good accuracy from lidar data. Finaly our analysis also indicates that lidar might not be so successful to identify biomass variations in deciduous forests, which is consistent with what was discussed in Nelson et al. (2007). Future research is necessary to confirm this finding.

Our field data analysis indicates that woody volume (estimated as the product of vegetation height and basal area) is an excellent biomass predictor especially with known dominating forest type to be either deciduous or coniferous forests. Forest type information can be easily extracted from the existing land cover maps derived from optical remote sensing data. Our study implies that fusion of lidar and optical remote sensing will provide better above-ground biomass estimates than lidar alone.

Our result also demonstrates that a fully digitized hemispherical-scanning below-canopy lidar like the Echidna® validation instrument (EVI) provides excellent woody volume measurements, i.e. an accurate measure of above-ground biomass, particularly at the stand level.

Further, allometric equations to derive above-ground biomass in this study were derived in New England over flat areas. These equations can vary from site to site due to difference in soil, climate and topography. The best approach would be using allometric equations derived specifically for each site. However these equations are often not available. The error associated with using one set of allometric equations for all stands might lead to some bias in our relationship between biomass with structure parameters. Our findings need to be further tested in other regions.

This study uses ground and lidar data collected over flat terrains. The relationships between above-ground biomass and vegetation structure parameters are complicated by adding topography (Lefsky et al., 2005; 2007; Hyde et al., 2005 and Harding et al. 2005). Surface topography has been a known factor in above-ground biomass production (Whittaker et. al. 1974). Uncertainties in LVIS height metrics estimates from lidar also increase with terrain slope (Blair et. al. 2006). Correction of slope effect on height metrics is required to derive accurate vegetation height structure parameters (Lefsky et al., 2005; 2007; Harding et al., 2005, Hyde et al., 2005; Yang et al., 2009b and Lee et al., 2009). Recent study in montane ecosystems indicates that above-ground biomass is more closely related to RH75 than RH50 (Swatantran et al., 2009). This conclusion is different from what was found in Drake et al. (2002a, 2002b and 2003). Topography might be a factor for such a relationship. In montane regions, woody volume might be better related to RH75 than RH50 as found in our study. Our future work includes assessing the relationship between biomass and vegetation structure parameters over slope terrains by adding the surface topography effect in above-ground biomass estimates from lidar.

Finally, this study and previous work all indicate that deriving height based allometric equations for above ground biomass estimate is possible. These types of allometric equations will also be dependent on stand age (Ketterings et al., 2001; Pilli et al. 2006; and Zianis and Mencuccini, 2004). Optical remote sensing on stand age (Song et al., 2007 and Liu et al., 2008) may have to be fused with lidar remote sensing.

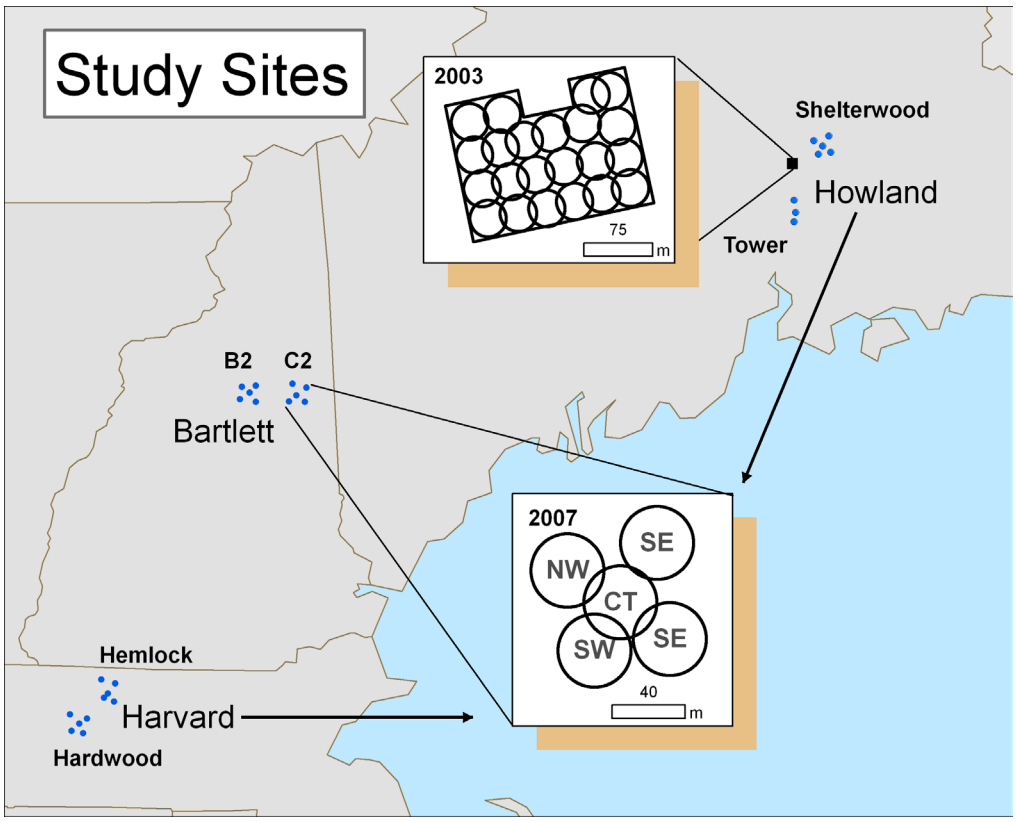

Figure 3.1: Locations of our study sites (blue dots and black square) in New England and sampling strategies, including 2003 stem map boundary and plot range (in circles) (top square) and the spatial arrangement of 2007 field data (bottom square). The plot sizes are 20m/25m radius circles with 5m geolocation accuracy.

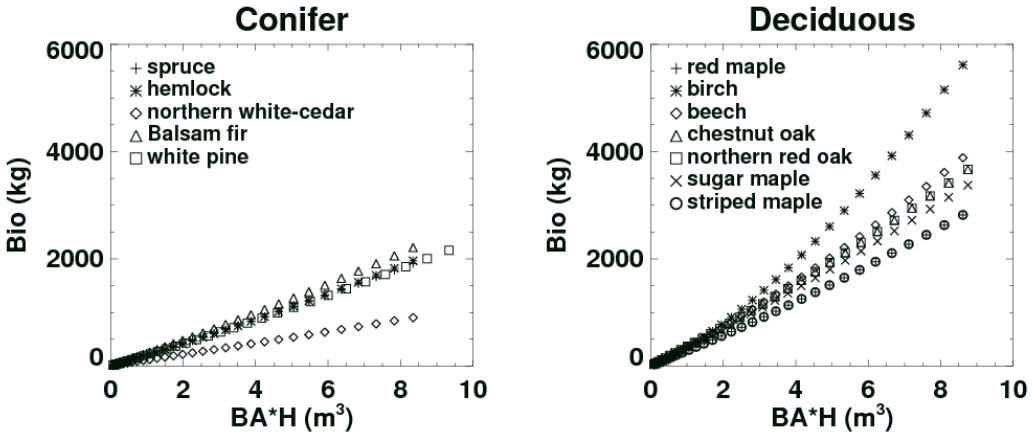

Figure 3.2: Relationship between single tree species above-ground biomass with woody volume (product of basal area (BA) and tree height (H)) for major tree species in the northeastern US. BA and H are calculated with DBH ranging from 10 to 68cm with a 2 cm increment.

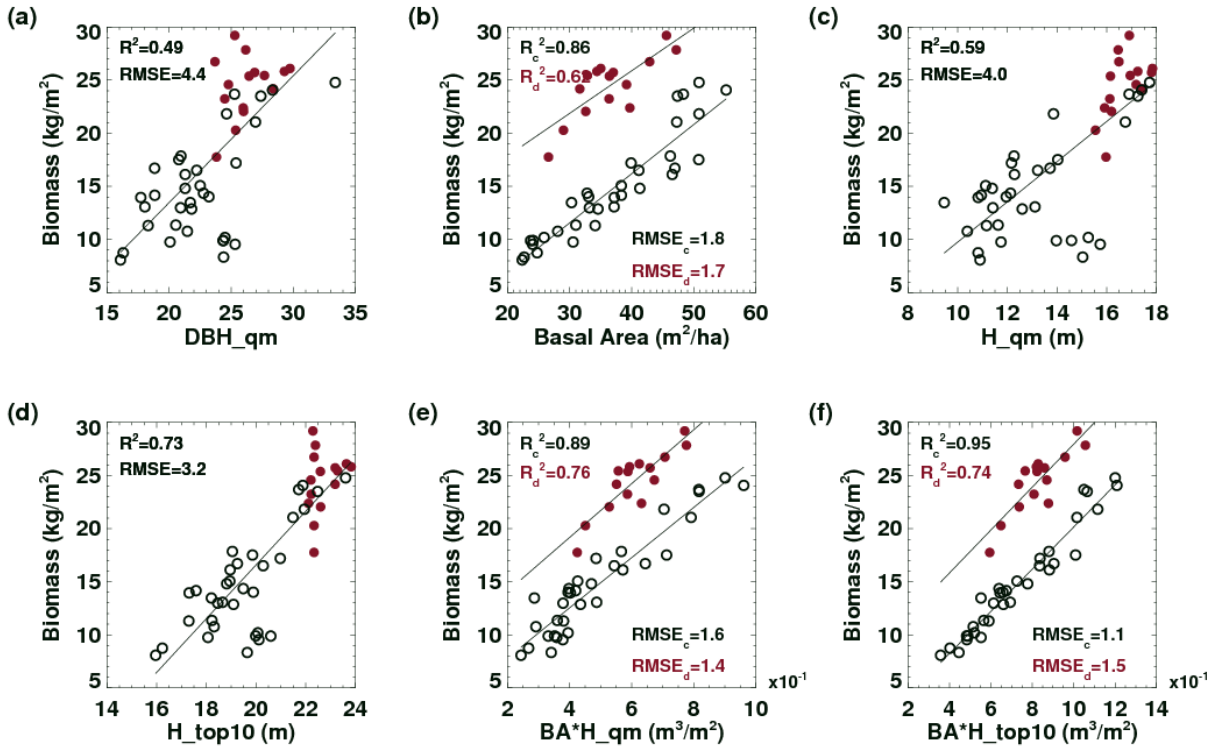

Figure 3.3: Relationships between plot level above-ground biomass and different tree structure parameters for all study plots in New England. Solid circles are deciduous plots and hollow circles for conifer plots.

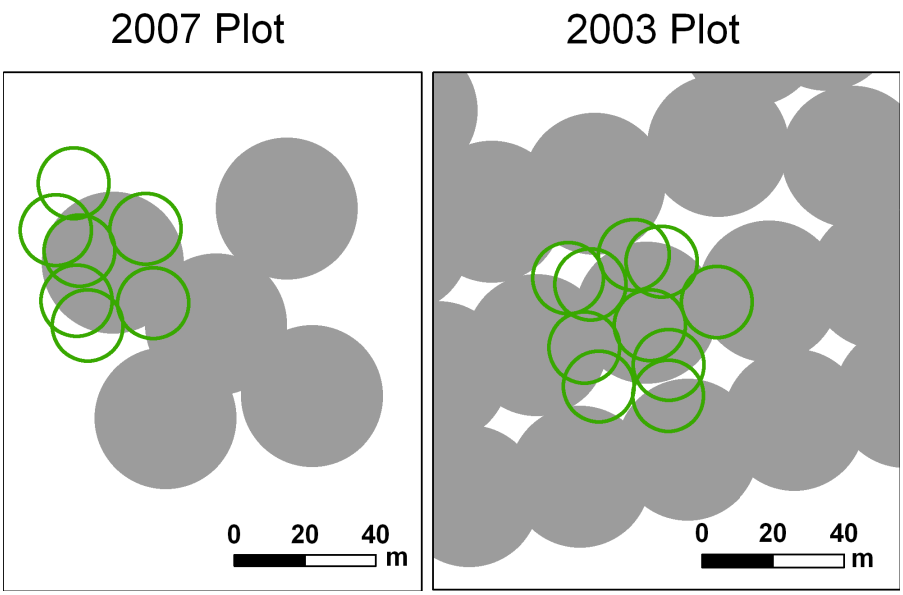

Figure 3.4: Spatial relationship between LVIS footprint (hollow circles) and plot range (solid circles). Plot level LVIS products were obtained by averaging LVIS data for footprint centers located within a 5m radius circular centered at each plot center.

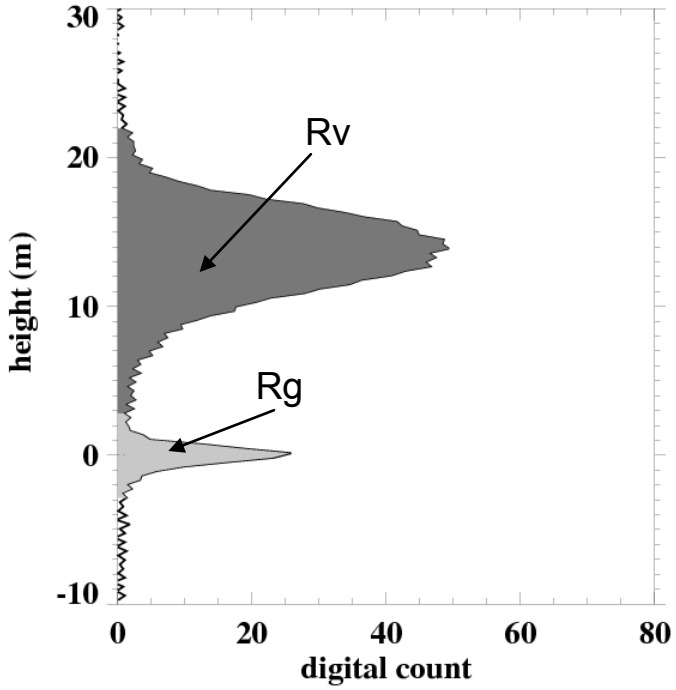

Figure 3.5: Illustration of accumulated vegetation returns (Rv) and ground returns (Rg) from a typical lidar waveform.

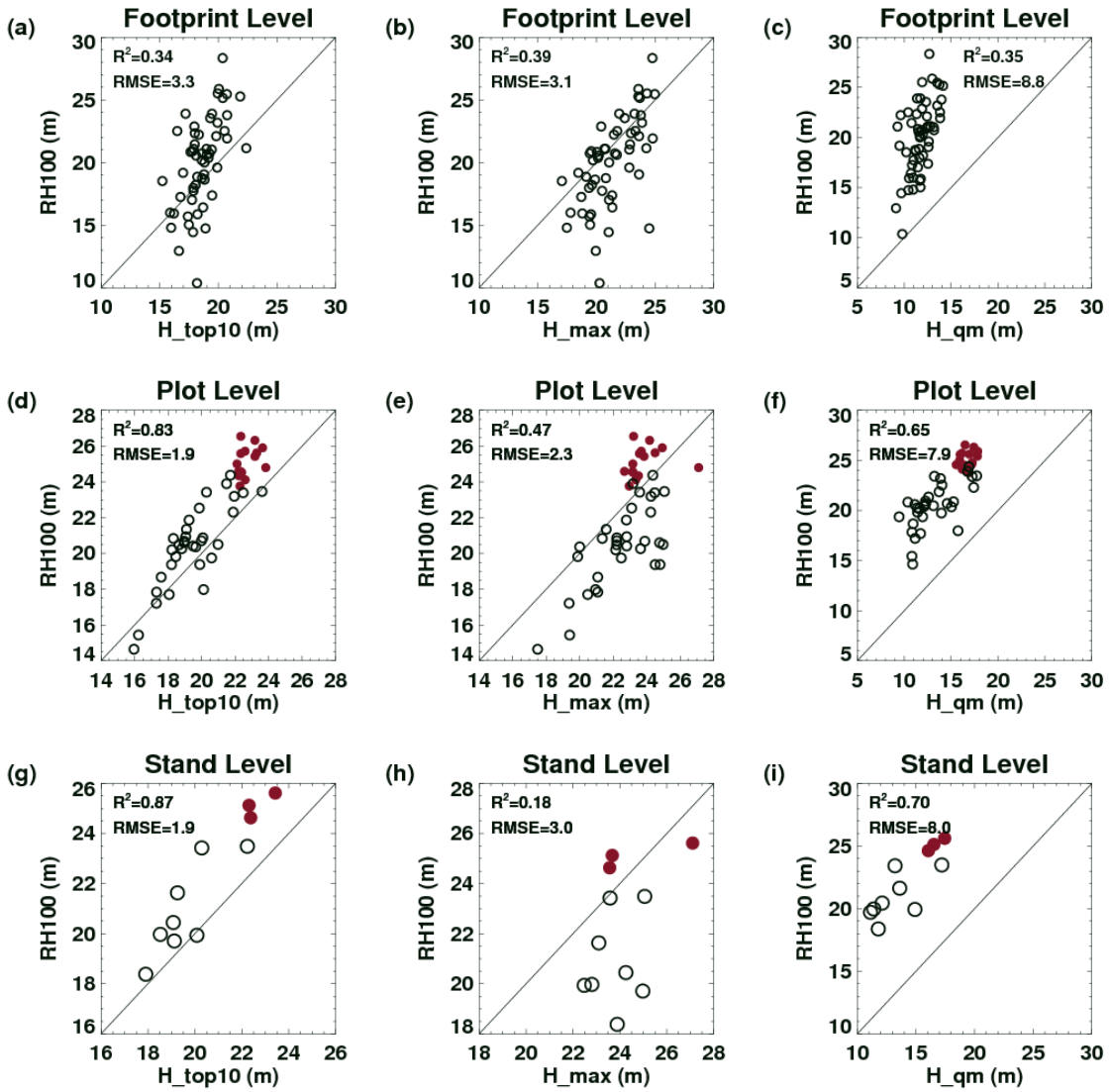

Figure 3.6: Comparison of RH100 with ground measured tree height metrics (left panel: top 10% mean tree height; middle panel: maximum tree height; right panel: quadratic mean tree height) at three scales (Top panel: footprint level (0.03ha), middle panel: plot level (0.12ha) and bottom panel: stand level (~1 ha)). Solid circles are deciduous plots and hollow circles for conifer plots.

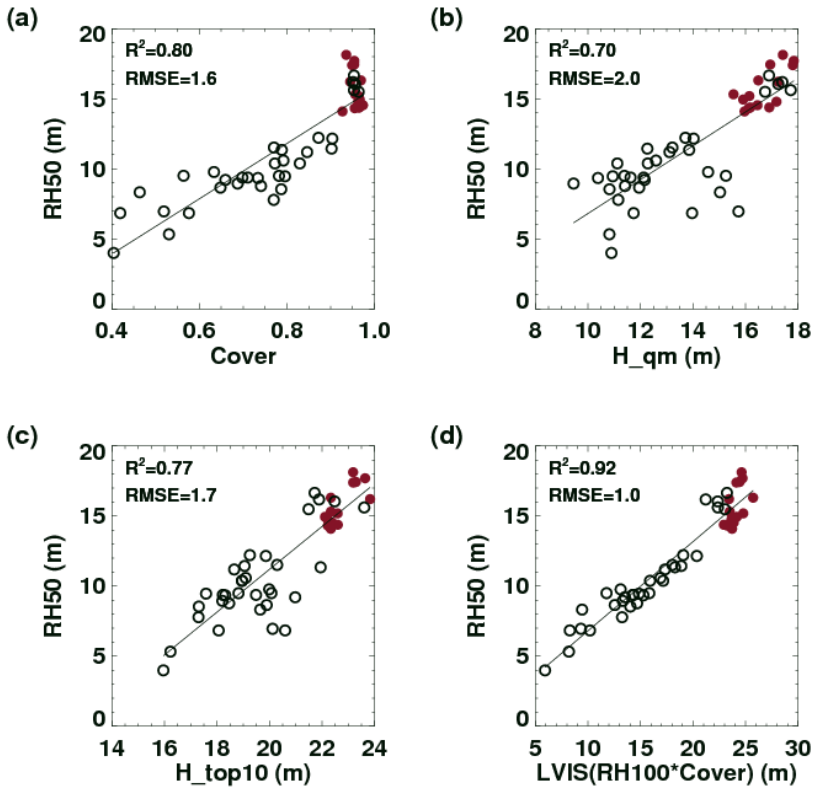

Figure 3.7: Comparison of plot level RH50 with vegetation cover, ground quadratic mean and top 10% height and relationship between RH50 and the product of RH100 and LVIS vegetation cover for all study plots (solid circles for deciduous plots and hollow circles for conifer plots) in New England.

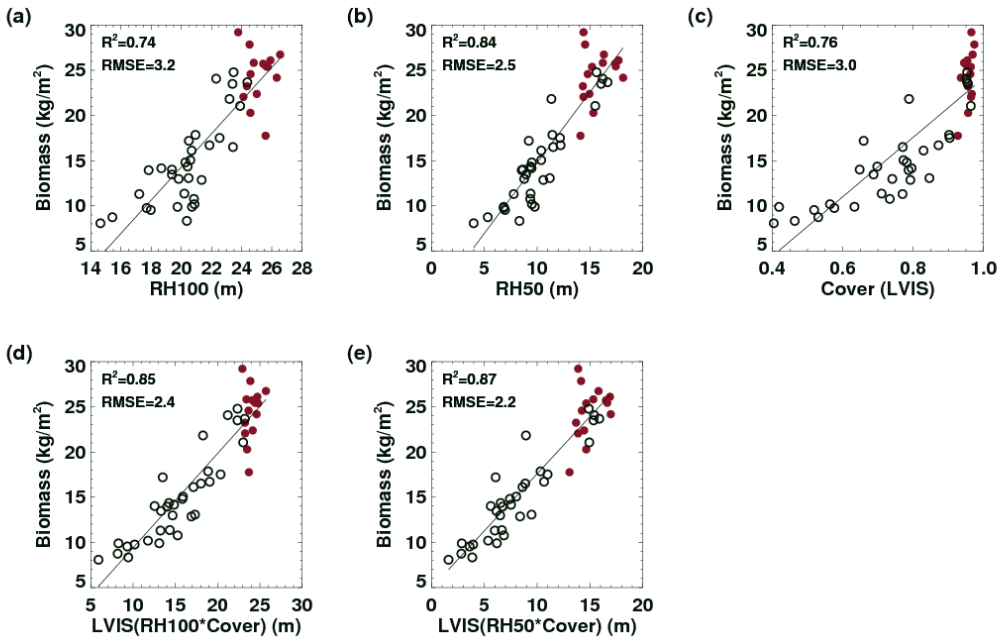

Figure 3.8: Relationships between plot level above-ground biomass with LVIS height metrics (RH100, RH50), vegetation cover and their combinations for all study plots (solid circles for deciduous plots and hollow circles for conifer plots) in New England.

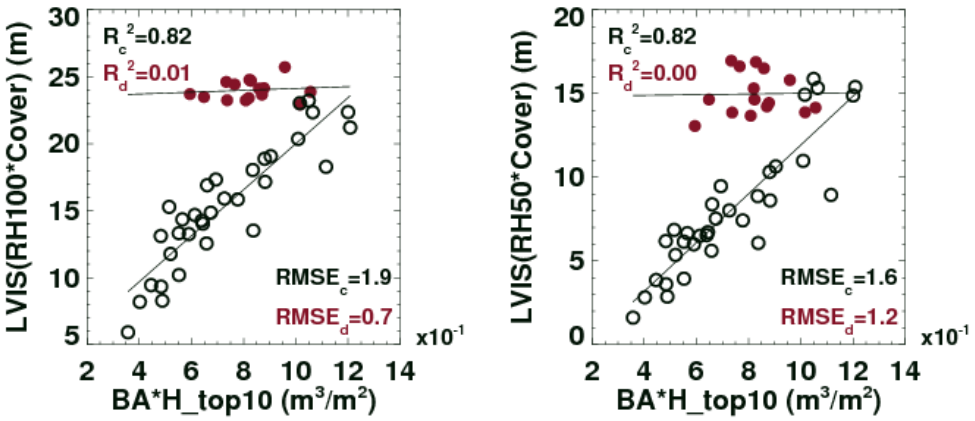

Figure 3.9: Relationships between LVIS RH100*cover, RH50*cover with woody volume for all study plots (solid circles for deciduous plots and hollow circles for conifer plots) in New England.

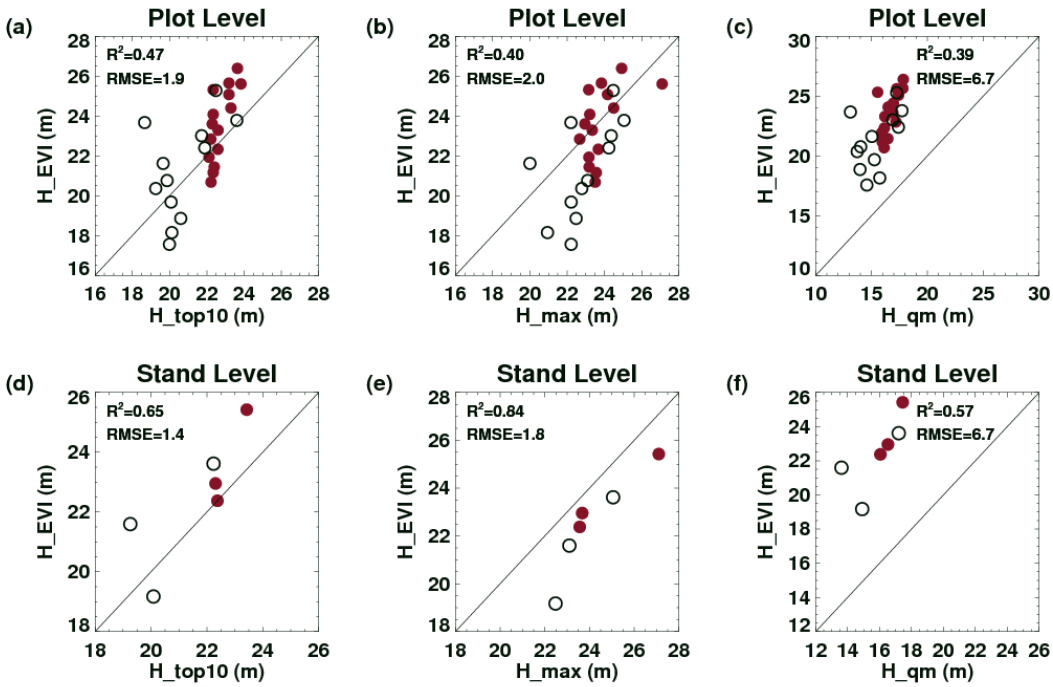

Figure 3.10: Comparison of EVI height with ground measured tree height metrics (left panel: top 10% mean tree height; middle panel: maximum tree height; right panel: quadratic mean tree height) at plot (0.12 ha) and stand (~ 1 ha) scales (solid circles for deciduous plots/stands and hollow circles for conifer plots/stands) in New England.

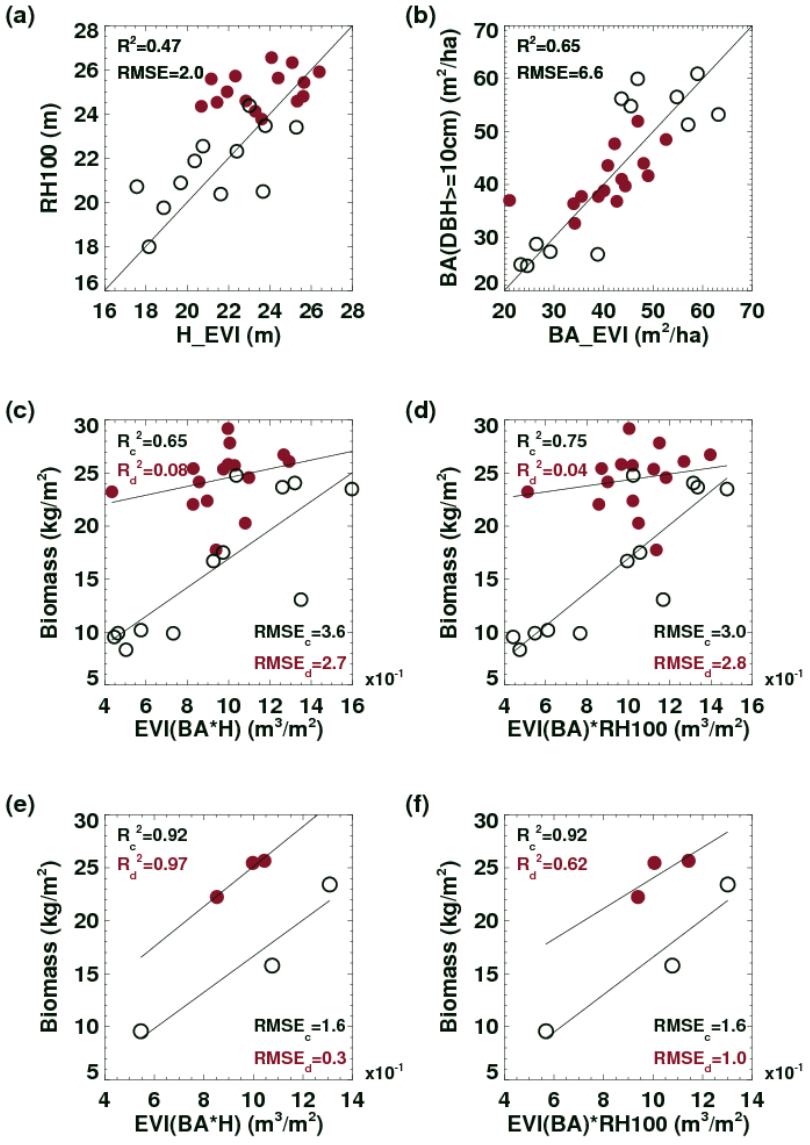

Figure 3.11: Comparison of LVIS and EVI vegetation heights, EVI measured with ground based basal area (top two panels) and above-ground biomass with woody volume measured by EVI or fusion of EVI and LVIS at the plot level (middle two panels) and the stand level (bottom two plots) (solid circles for deciduous plots/stands and hollow circles for conifer plots/stands) in New England.

Chapter 4 Toward Operational Biomass Estimation from Vegetation Canopy Lidar

4.1 Introduction

Many studies have shown the capability of vegetation canopy lidar in estimating above ground biomass (AGB). Several airborne and spaceborne lidar systems have demonstrated the ability to retrieve AGB over various biomes, including studies using Scanning Lidar Imager of Canopies by Echo Recovery (SLICER, Blair et al, 1994) in Cascade Mountain Range in Oregon and Washington States, U.S. (Lefsky et al., 2005a), in Annapolis, MD, USA (Lefsky et al., 1999a); studies using the airborne Laser Vegetation Imaging System (LVIS, Blair and Hofton, 1999) in La Selva (Drake et al., 2002), White Mountain, NH (Anderson et al., 2006, Ni-Meister et al., 2009), and Sierra Nevada, California (Hyde et al., 2005; Swatantran et al., 2010); and studies using the spaceborne Geosciences Laser Altimeter System (GLAS) in Ciénaga Grande de Santa Marta (CGSM) in Colombia (Simard et al., 2008), Quebec, Canada (Boudreau et al., 2008), conifer forest in Oregon, USA, deciduous in Tennessee, USA, and Santarem, Para State, Brazil (Lefsky et al., 2005b). On the global scale carbon storage, lidar remote sensing will provide critical information from the new generation spaceborne lidar systems, such as the second generation Ice, Cloud, and land Elevation Satellite (ICESat II) (http://nasascience.nasa.gov/missions/icesat-ii), and Deformation, Ecosystem Structure and Dynamics of Ice (DESDynI) missions.

However, current biomass estimation methods depend on empirical regression analysis to determine the site specific relationships between lidar data and biomass. There are more than 50 different lidar parameters have been used to estimate AGB (Nelson et al., 2009). The general lidar-biomass retrieval approach including select optimal parameter sets by step-wise regression analysis, neural network, or simply try-and-error depending on the size of the parameter pool used. The choice of parameter(s) are found to be site dependant, or more precisely, species, tree density, tree height and vertical distribution, topography, etc are all factors. Since the optimal lidar parameter(s) to estimate AGB changes by sites, the coefficients used in the biomass estimates are also different. At global scale, a dense ground validation network will be required if current methods were used to estimate AGB. A dense global ground calibration network is likely too expensive to be practical. Future global lidar biomass mapping requires an operational method that can provide good biomass estimates over a wide range of vegetation conditions with limited available ground calibration data.

The most common way of estimating biomass from lidar is using the lidar height metrics, such as RH100 and RH50. These lidar height metrics (RH100 or RH50) are defined as the heights relative to ground (denoted as "RH") at which accumulated waveform energy reaches X% (i.e. 100% for RH100 and 50% for RH50. see Blair et al., 2006 for detail). However, using a single height to refer biomass is somehow limited in dealing with forest with different structures, because AGB is affected by several forest structure parameters (as discussed in Ni-Meister et al., 2010). Drake et al. (2002a, b) found that the height of medium energy returns (HOME or RH50) is better related to AGB than the height at full energy returns (RH100). This is because RH100 is

determined solely by the height of the canopy top, but HOME/RH50 is affected by both canopy height, canopy structure, and the degree of canopy openness (Ni-Meister et al., 2009). Using parameters that can respond to both canopy height and openness (e.g. RH50) is particular important in the forests having highly variable canopy cover. For example, our previous investigation in New England forests (Ni-Meister et al., 2010) found that the AGB estimates from RH100 are similar to RH50 in forest stand with near complete canopy cover. RH100, however, fail to predict AGB in the shelterwood plots, where trees have been harvested in a series of narrow tracks, while RH50 still show reasonable AGB predictability.

Being able to respond to both canopy height and openness, the HOME/RH50 has been widely used to estimate AGB. RH50, however, fail to estimate biomass in the Sierra National Forest, CA, where both vegetation structure and topography are highly variable (Swatantran et al., 2010). Among standard LVIS height metrics (Blair et al., 2006), RH75 was found to be the strongest biomass predictor (Swatantran et al., 2010). The results indicate that neither RH50 nor RH75 is a universal lidar vegetation structure parameter that can be applied for various vegetation and climate conditions for above ground biomass estimate at large scales. Thus many studies use multiple regression analysis to estimates the optimal lidar AGB parameter(s) from different lidar metrics, e.g. RH100, RH75, RH50, and RH25 (Lefsky et al., 2005a, b; Drake et al., 2002). Variation in optimal lidar-AGB parameter is likely due to each lidar parameter (height metrics) only reflect some aspects of vegetation characteristics that is related to AGB. Past studies have shown different lidar parameter(s) is/are sometimes used to fit the unique

site configuration to improve AGB estimation accuracy (Drake et al., 2002b; Hyde et al., 2005; Lefsky et al., 2005b; Nelson et al., 2009).

Lidar returns are the collective results of forest structures. Therefore, a lidar parameter that better approximates forest structure characters should better relate to the AGB. Lidar waveform records back-scattering energy intensity at each height interval as waveform, which is affected by vegetation's vertical and horizontal structure (Drake et al., 2002b). For each height interval, the higher tree density should results in denser canopy and lead to stronger lidar return. Thus the whole lidar waveform could be used to approximate tree density at each measured height interval. Since AGB is highly related to tree height, both height and the magnitude of lidar return are important in estimating ABG.

The purpose of this study is to develop a lidar parameter better representing the vegetation structure information that can be applied to estimate AGB at various biomes. We will demonstrate and evaluate this using ground and LVIS data collected in three different biomes, deciduous and coniferous forest in New England, USA, and Sierra National Forest in California, USA. The performance of developed lidar biomass parameter will be compared with lidar height metrics, which are commonly used to estimate forest AGB. The paper is organized as follows: Section 4.2 and 4.3 describe the sites and datasets. Section 4.4 presents ground and LVIS data processing and a lidar biomass parameter that uses both height and gap fraction information in lidar waveform. Section 4.5 discusses the evaluation results of our lidar AGB retrieval method by comparing the relationships between standard lidar height metrics and our biomass

parameter at various site configurations. Section 4.6 discusses the uncertainties of our algorithm and with conclusions in Section 4.7.

4.2 Study site

Our study sites including conifer and deciduous forest stands in New England, and Sierra National Forest in California USA (Figure 4.1).

New England data are collected from three intensive ecological study sites: Harvard Forest (42°32′N, 72°11′W) in Massachusetts, Bartlett Experimental Forest (44°06′N, 71°17′W) in White Mountain region, New Hampshire, and Ecosystem Research Forest (45°12′N, 68°44′W) in Howland, Maine. This region is featured with old-growth northern hardwoods with beech, yellow birch, sugar maple, and eastern hemlock being the dominant species. Most of the forest is covered by high canopy where sugar maple/beech/yellow birch dominates its upper canopy with maximum height around 30 m (Filip et al., 1971).

The study site in Sierra National Forest (37°00′N, 119°10′W) lies along the western slopes of the central Sierra Nevada in California, USA. The region has a Mediterranean climate with elevations ranging from 853 to 2743 m (see Hunsaker et al., 2001 for detail). Forests in this region are dominant by red fir, white fir, Sierra mixed-conifer, ponderosa pine, montane hardwood-conifer and California black oak. Unlike the New England forest, which is usually dense with lower variation in canopy heights, forest structure is complex in Sierra sites and both canopy cover and tree height can varies significantly. Tree height can reach 70-80 m and multi-layers canopy cover is common. Forest biomass also vary significantly in Sierra Forest and it can reach more than 100

kg/m^2 in Giant Sequoia stands; comparing to 10-30 kg/m^2 typically observed in New England forest.

4.3 Data

4.3.1 Field data

Field data in New England were collected in 2003 and 2007 using 20-25 m radius cycle plots. All live stems with a diameter at breast height (DBH) >= 10 cm were inventoried with species information. The biomass is estimated using DBH based, species-specific allometric equation (Jenkins et al., 2004), since tree height is only measured for few selected trees. Detailed description of sites, dataset and biomass estimation method can be found in Ni-Meister et al. (2010).

Sierra data were collected in 2008 using 33 by 33 m rectangular plots. Similar to New England sites, all live stems with a DBH >= 10 cm were inventoried with species information. Additionally, top and bottom crown heights were also measured. Since tree height is measured, we estimated biomass using species-specific allometric equations based on both DBH and tree height from the USDA Forest Service (Waddell and Hiserote, 2005) for better accuracy.

4.3.2 LVIS data

The Laser Vegetation Imaging Sensor (LVIS) is a large footprint waveform digitizing, scanning laser altimeter, designed, and developed at NASA's Goddard Space Flight Center (See http://lvis.gsfc.nasa.gov/; Blair et al., 1999). The LVIS's vertical sampling resolution is 30cm (1ns). The sensor view angle and footprint sizes

(diameter) can typically vary between 0 to 12 degree and 10 to 25 m respectively, depending on the relative location of ground target and flight path and altitude.

LVIS data over the New England sites were collected in the summer of 2003 with an estimated 3-5 degree field of view. Data over the Sierra National Forest were collected in the summer of 2008. Four flight paths were collected with various degree of field of view. We used the data collected from the near-nadir viewing data (collected in day 1), which has estimated 0-5 degree field of view over the sample sites. Both data collected in New England (2003, 2007) and Sierra National Forest (2008) has footprint size of 20 m.

4.4 Method

4.4.1 Field data analysis

Plots data were selected to cover wide variety of canopy and topography conditions to investigate the impact of vegetation structure and topographical relief on lidar-AGB relationships. A total of 69 plots were collected, 48 in New England study sites, and 25 in Sierra Forest. In Sierra dataset, 4 plots data were excluded because no quality lidar data were found within the plot, leaving 21 suitable for analysis.

Table 4.1 shows the summary of plot structural parameters. It shows AGB, canopy cover and slope varied significantly among New England's deciduous and conifer dominated stands and Sierra Forest. In New England's deciduous dominated plots, AGB and canopy cover are generally higher and have lower variation comparing to conifer dominated plots. Part of the reasons that AGB and canopy cover varies more in conifer plots is due to shelterwood plots (5 of them), which their AGB and canopy cover is

roughly halved from recent harvest. Terrains in New England plots are mostly flat with the highest within plot slope reaching only 5 degree.

Plot structural parameters in Sierra plots are highly variable. AGB, slopes and canopy cover all show larger variation than those in New England plots. Slopes are significant for most plots. Canopy covers are not as high as New England plots. In fact, few Sierra plots have very low canopy cover (less than 25%) as opposite to New England, where most plots have high canopy cover. Mean AGB (28.94 kg/m^2) in Sierra is only slightly more than the mean of deciduous plots in New England; however, the range (7.04 – 59.4 kg/m^2) is much larger than New England plots.

In summary, New England plots are flat, have high canopy cover and low biomass variation especially in deciduous dominated plots. Sierra plots have high variability in AGB, canopy cover, and slope, and canopy cover is not very high.

4.4.2 LVIS data processing

To extract lidar vegetation structure parameters for each plot, LVIS waveforms within the plot were first aggregated to create a single normalized waveform at plot-level. This method is different than what we used in previous investigation (see Chapter 3.7 for detail), which LVIS parameters are calculated at footprint-level then averaged within each plot. Aggregating waveforms is more complicated, but can reduce uncertainty introduced by averaging method. In our New England sites, where vegetation structure is more homogeneous within the plot, the expected bias is small between the values produced by these two methods because the LVIS waveforms (metrics) within each plot have similar shape (values). However, in Sierra site where forest structure can be highly heterogeneously distributed within the plot, the mean value of lidar metrics can be

different, depending on which averaging method is used. Calculation using aggregated waveform can reduce uncertainty since lidar metrics is estimated from a single waveform representing the overall vegetation structure in the plot.

To get the plot-level LVIS waveform, we first selected the corresponding LVIS data for each plot using an area slightly larger than plot range to compensate some geolocation uncertainty. LVIS footprint centers within 5m of plot range were selected. For each waveform, the vertical gap fraction distribution (*dp/dz*) was derived by normalized waveform with the assumption of equal ground-crown reflectance ratio (Ni-Meister et al., 2001). The vertical gap fraction distribution (*dp/dz*) for each waveform is aggregated at each height and normalized to approximate the plot-level vertical gap fraction distribution, or,

$$dp(z)/dz = \frac{1}{N}\sum_{i=1}^{N} dp(z)_i / dz \qquad (4.1)$$

where $dp(z)/dz$ is the normalized (plot-level) gap fraction at height z, $dp(z)_i$ is the gap fraction for i^{th} LVIS waveform at height z, and N is the number of LVIS data within plot range.

Lastly, we used the aggregated plot-level LVIS gap fraction to develop lidar biomass index and calculate lidar height metrics (RH0 to RH100) for the analysis.

4.4.3 Lidar Biomass Index

The lidar returns, or waveform, can be used to indicate crown coverage at each height (Figure 3.1). Since the crown have few general shapes, lidar return at height (z) can be use to associate tree density at that particular height. For individual tree, above ground biomass is proportional to tree height (Agee, 1981; Chave et al., 2005), and the

biomass within a lidar footprint is the sum of all trees. Therefore, a summation of height adjusted lidar return intensity should produce a parameter better approximate AGB.

It is worth to note that above ground lidar waveform is corresponding to foliage profile, which itself contains two pieces of information, the amount and height of foliage. Integrating lidar waveform without height can not produce a good biomass parameter because if such relationship exists, tall and short tree will have the same biomass if their crown size is the same. In the study by Lefsky et al. (1999b), he defined a lidar parameter, quadratic mean canopy height (QMCH), as the square root of the summation of the product of the canopy height profile attributes and each element's squared height. He explains the weight (squared height) on canopy height profile as a factor that is proportional to the size of the tree required to support canopy at such height (Niklas 1996). The results confirm a weight, proportionally to height, is needed if multiple lidar canopy height returns were to use for AGB prediction.

We propose a lidar parameter, Waveform Weighted Index (WWI), defined as the summation of the product of the canopy gap distribution and each element's squared height. This parameter is similar to QMCH in Lefsky et al. (1999), but it uses the vertical gap profile instead of canopy height profile. The reason for such modification is because vertical gap information is readily in lidar waveform, but the process in deriving canopy height profile from lidar could subject to uncertainties, such as clumping factors. WWI can be described as,

$$WWI = \int_{w\_end}^{w\_first} (dp/dz) * z^2 dz \qquad (4.2)$$

where dp/dz is the gap fraction vertical distribution derived from lidar waveform, z is the lidar return height relative to ground, *w_end* is the last detectable return (the bottom of

ground return), *w_first* is the first detectable return (top of the canopy return). Since z is the height relative to ground, bellow ground waveform returns are integrated as negative values to negate the above ground portion of ground returns.

In Eq. 4.2, height (z) is squared to approximate its relation with AGB, which is usually a power function (Agee, 1981; Chave et al., 2005). Our analysis found power of ~2 produced the best site independent WWI-AGB correlations. However, during our analysis (not show here), the optimal power function is slightly deviated from 2 for site specific data. This suggests the power function is likely associated with site configuration, i.e. species composition, canopy openness, and topography. In this study, we did not attempt to fine tune the power function, instead, the focus is on finding the base and the more general form of lidar biomass parameter. Therefore, the square term is used for all study sites.

4.5 Results

To evaluate the performance for estimating above-ground biomass, we compares the above-ground biomass estimate using this index with the field measured above-ground biomass and using other commonly used lidar height metrics. In the end, we analyze the relationships between WWI, ABG and height profile (RH0 to RH100) to understand the physics of our index.

4.5.1 WWI vs. AGB

The regression analyses between WWI and AGB (Figure 4.3) shows that the correlations are good and consistent over sites (R^2=0.75-0.83, RMSE=2.5-7.5). More importantly, the WWI-AGB relationships described in regression equations coefficients

(Table 4.2) are very similar among New England, Sierra, and the combination of both. The consistent WWI-AGB relationships derived from field plots in different forest types indicates there are some general relationships exist between WWI and AGB for the forest types investigated in this study. The scatter plot of WWI and AGB using all plots (CA + NE in Figure 4.3 right panel) shows that there is some degree of separation between Sierra and New England data. This is partly caused by the site specific configuration which indicates slightly different relationships between WWI and AGB.

The correlations strength between AGB and WWI is similar to those achieved by RH75 (Table 4.2), however, regression coefficients in WWI-AGB are much more stable than those in RH75-AGB. The stability in regression coefficients for WWI based biomass models indicates WWI parameter likely capture the core element in vegetation structure relating to AGB.

Table 4.2 also shows the regression statistics of AGB and standard lidar height metrics, RH50, RH75 and RH100. For New England sites, both RH50/RH75 shows good correlations with AGB (R^2 = 0.83, 0.79, RMSE = 2.6 m, 2.9 m), and the correlation for RH100 is slightly lower (R^2 = 0.68, RMSE = 3.5 m). This result is similar to our previous investigation (Chapter 3); however the correlation strength is slightly different because of different approach in deriving plot-level LVIS parameters.

Analyses using data from Sierra site and all sites show very different results than those in New England (Table 4.2). AGB in Sierra site correlated poorly with RH50 (R^2 = 0.38, RMSE = 11.8 m), and RH100 (R^2 = 0.33, RMSE = 12.2 m). However, RH75 shows good correlation with AGB (R^2 = 0.8, RMSE = 6.7 m). RH75 also shows the best correlation with AGB when all data were used for analysis. RH100 is not strongly

correlated with AGB (R^2 = 0.46, RMSE = 8.1 m) and no correlation was found between RH50 and AGB (R^2 = 0.07, RMSE = 10.6 m).

Comparing the lidar-AGB regression equations derived from different site (Table 4.2), regression coefficients varied significantly for RH50 and RH100. For RH75, the regression coefficients are closer among sites; however, they are still not very stable. The results indicate the relationships between AGB and lidar height metrics are site dependant especially for RH50 and RH100. Therefore, the standard lidar height metrics is not suitable for large scale study as calibration remain an important part of process to ensure accuracy.

4.5.2 WWI vs. RH50/75/100

Figure 4.4 shows that a linear relationship exists between WWI and lidar height metrics (RH50/75/100) having high correlation with AGB. In New England sites, linear relationships were found between WWI and RH50/75 but RH100. The comparison between WWI and RH100 (Figure 4.4 left panel) shows there is high RH100 variation between WWI values of 50 and 150. On the contrary, there is lack of variation in RH100 when WWI is between 200 and 350. The results indicate RH100 is somewhat saturated in high biomass plots, and could be over sensitive in the mid-to-low biomass New England plots.

In Sierra site, WWI only shows correlation with RH75, which is the only height metrics having high correlation with AGB (Table 4.2). Similar to New England plots, RH100 seems to saturate when WWI is larger than 200. In addition, Figure 4.4 shows the values of RH50 are very small or even negative due to the low canopy cover and high slopes. The failure for RH50 to capture WWI is likely due to the very low values, which

could be over sensitive small errors. Using combined dataset, the scatter plot (Figure 4.4 right panel) shows RH75 is the only parameters having clear relationships with WWI. In addition, the plot shows two different groups for WWI vs. RH50/100, indicating two different relationships between WWI and RH50/100 for New England and Sierra sites.

The analysis indicates RH100 and RH50 can be saturated or over sensitive to errors in estimating AGB depending on the site configurations. And, a better fit between WWI and lidar height metrics seems to imply both parameters will be better correlated with AGB.

4.5.3 RHs vs. WWI and AGB

Figure 4.5 shows the correlations of determination (R^2) of all lidar height metrics (RH0 to RH100) with WWI and AGB. The RHx is the Lidar height relative to ground at x% of accumulated waveform energy. The R^2 values for each RH number were connected to show the correlation variation from RH0 to RH100.

In New England sites, the correlation between RH and WWI are very high between RH20 to RH98 ($R^2 > 0.9$). The correlations between RH numbers and AGB show a very similar curve as high correlations can be obtained for any RH number between 20 and 98. This is likely the results of simpler canopy structure and flatter terrain in New England sites. Vegetation in New England sites are usually very dense, has high canopy cover and low variation in vertical and horizontal structure. Therefore, there is high degree of autocorrelation for heights at each accumulated lidar energy return percentile (RH) above ground. The slight dip in correlation strength for the top lidar height metric (RH100) is partly due to the uncertainty in determining first return (see Section 3.7.2 for more discussion).

In Sierra site, the correlation curve of WWI and RH numbers shows a sinusoidal like curve. There are two RH number ranges with high correlations, RH5 to RH25 (R^2 = 0.6 – 0.75) and RH60 to RH80 (R^2 ~ 0.9). The correlations between RH numbers and AGB shows similar curve but a more distinct peak around RH73 instead of a plateau shows in WWI–RHs correlation curve. The results are likely caused by the complex canopy structure in Sierra site, where each plot also varies in density, and slopes.

Similar analysis using combined New England and Sierra dataset shows correlation curves of RH number vs. WWI or AGB that are very different than those derived from individual site. Comparing to either New England or Sierra site, the correlation curves derived using combined dataset shows a sharper single peak for both WWI and AGB, indicating fewer lidar height metrics are good AGB indicators when vegetation structure variation increases. The peak correlation for the combined sites (R^2 = 0.78 at RH80) between RH numbers and AGB is very similar to that of New England (R^2 = 0.83 at RH50) and Sierra (R^2 = 0.81 at RH73) (Table 4.2). The results indicate the best single lidar height metrics can explain about 80% of AGB variation, and the optimal lidar height metrics for estimating AGB is depending on vegetation structure.

4.6 *Discussion*

4.6.1 Lidar height metrics

Analysis in Section 4.5.1 shows both RH100 and RH50 fail to capture AGB variation in Sierra site. RH100 is not an ideal AGB predictor because it only records top canopy height which is not affected by tree density and can subject to large height uncertainty on steep slope (Yang et al., 2010). Although RH50 is response to change in

both vertical and horizontal structure change, large uncertainty can still occurred in low density forest, where RH50 values are small. Since lidar waveform is known to have tailing effects, making the bellow ground peak lidar energy return from ground slightly larger than that energy return from above ground peak. The artificially inflated bellow ground peak energy return will lower the height of RH50. The tailing effect on RH50 is small in dense canopy, where ground portion of lidar return is small, i.e. in New England sites. However, RH50 is close to 0 (ground height) in sparse canopy and the relative change (%) caused by tailing effect will be significant. Similarly, the ground height determination error will also have an amplified effect on the small RH50 values. Our Sierra study site shows such an example, where many plots having RH50 values close to zero and few are negative (Figure 4.4). Therefore, RH50 is not an ideal AGB predictor in sparse canopy over sloped terrain, as shown in the Sierra site analysis.

It is worth to note that some degree of correlation was observed between AGB and RH5 to 20 in Sierra site. Since canopy cover is not very dense in Sierra site, accumulated lidar energy returns under 20% are mostly coming from ground, and the RH are influenced by topography. In addition, because RH is defined relative to mean ground height, most values for the low RH numbers are negative, and are negatively correlated with AGB (in Figure 4.5, we show R^2). The result indicates AGB in Sierra plots is also related by topography. The topography effect on AGB is results of the variation in the biological, hydrological, and climatic condition on tree growth (Whittaker et al., 1974; Tajchman and Wiant, 1983). Therefore, in an undisturbed forest with similar tree species composition and age, topography can affect AGB because the tree allometry can be different on the slopes than on the flats.

4.6.2 Lidar AGB parameter

It is worth to note that above ground lidar waveform is corresponding to foliage profile, which itself contains two pieces of information, the amount and height of foliage. Integrating lidar waveform without height can not produce a good biomass parameter because if such relationship exists, tall and short trees will have the same biomass if their crowns having the same size. In the study by Lefsky et al. (1999b), he defined a lidar parameter, quadratic mean canopy height (QMCH), as the square root of the summation of the product of the canopy height profile attributes and each element's squared height. He explains the weight (squared height) on canopy height profile as a factor that is proportional to the size of the tree required to support canopy at such height (Niklas 1996). This is consistent with the WWI parameter where lidar derived gap fraction (*dp/dz*) is multiplied by its height square to approximate the size of the tree.

Some uncertainties could be introduced by including ground return in WWI calculation (Eq. 4.2). Lidar waveforms are known to have tailing effect (Harding and Carabajal, 2005), which result in over-estimation of bellow ground portion of ground returns. This error should be small in typical forests since the values of $(dp/dz)\cdot z^2$ are much smaller comparing to those from higher canopy returns. However, in forest stands with very sparse vegetation covers and/or on high slopes, the uncertainties could be significant due to stronger ground return and/or waveform tailing effect.

4.6.3 Lidar AGB Estimate at Large Scale

In Figure 4.6, the shape of RHs-AGB correlation curve changes over sites indicating the correlation strength between AGB and lidar height metrics are site dependent. Although the optimal lidar height metrics can explain roughly the same

amount of AGB variation (similar R^2), both the optimal lidar height metrics and regression coefficients are site specific (Table 4.2). The results indicate in-situ data remain an integral part of lidar-biomass application since lidar-biomass model needs to be developed and calibrated for each site. The need for ground data will limit lidar AGB retrieval within intensely sampled forests, which make it difficult for large scale applications.

Table 4.2 shows Canopy Weighted Height (WWI) can explain about 75-83% of AGB variation, which is close to the best lidar height metrics. This implies WWI likely contains most AGB information in lidar height metrics. Since AGB has different degree of association with most lidar height metrics, indicating a lidar AGB parameter incorporating all lidar height metrics could be more robust. Furthermore, how to weight each lidar height metrics based on its physical relation with AGB will be important in formulating parameter having more universal response to AGB. In WWI calculation, the height square is used to approximate the relationships between lidar height metrics and AGB.

Comparing WWI to commonly defined lidar height metrics, RH100 can be seen as a version of WWI parameter having weight = 1 at 100 percentile waveform height and weight = 0 everywhere else; and RH50 is another special case of WWI parameter having weight = 1 at 50 percentile waveform height and weight = 0 everywhere else. This analogy offers one explanation why the single lidar height metrics is likely not universally applicable to different biomes as each lidar height metrics only contain partial vegetation structure information. For example, RH100 only depends on the detectable

tallest tree height, which is not affected by other vegetation structures, i.e. canopy cover, or tree density.

Although the heights of accumulated lidar energy returns in RH50 or RH75 do reflect some aspects of vegetation structure, it is still not very robust. Take RH50 for example, the same value can have different upper or lower canopy as long as the height of 50% accumulated energy return is the same. The same RH50 height generated by a few tall trees or many medium size trees could have different biomass. This is, however, not a critical issue for sites having similar vertical vegetation structures (e.g. New England sites), since the tree height, size distribution is similar. In this kind of forests, lidar waveform return shapes will be similar, and the magnitude and height will likely correspond to density and tree size thus AGB. This is the reason that parameters like RH50 (or HOME) has been shown to be a choice of lidar height metrics in estimating AGB (Drake et al., 2002a) because in most studies, the field plots used for lidar-AGB calibration are usually within similar structured forests. However, RH50 could fail to capture AGB variation when the forest structure is highly heterogeneously distributed, such as the Sierra site in this study. In large scale studies, the high variation in plot vegetation structure is common as shown in this study when New England data are combined with Sierra data.

In our previous investigation (Ni-Meister et al., 2010), RH100 * cover was shown to be a good approximation of woody volume. The same principle is taken further in WWI as gap fraction * height is calculated at each height interval before integration. However, some uncertainties remain as the relationships between the crown shape/size and tree size/AGB are species dependent and could also depend on other environmental

factors. The variation in the characteristics of crown will affect the relationship between lidar estimated tree density and height distribution. The woody density for different species also varies slightly, and is not detectable by lidar. These are some of the reasons to cause the small variations in the WWI-AGB regression equations (Table 4.2) among different sites. To reduce such uncertainties in large scale studies, it is possible to categorize vegetation into several primary types, each having some general allometric relationships among the tree crown, height, woody density, and AGB (Albini et al., 2006; Jenkins et al., 2004; Waddell and Hiserote, 2005). The vegetation types can be obtained from other forms of remote sensing data and WWI-AGB model can be calibrated separately to improve accuracy.

4.7 Conclusion

Although past research has demonstrated the ability of lidar remote sensing to retrieve above ground biomass (AGB), the relationships between lidar metrics and AGB has been site dependent, both for optimal metrics and their relationships with AGB. In this study, we found that by integrating height adjusted energy intensity from lidar returns, a single parameter can be use to estimate AGB over a wide range of conditions (i.e. deciduous and coniferous forest in New England, Sierra National Forest's complex structured forest and terrain, and the combination). The WWI parameter consistently predicted AGB in high accuracy ($R^2 \sim 0.8$). More importantly, the WWI-AGB model coefficients are very stable over sites. Comparing to commonly used lidar biomass metrics (RH50/HOME, RH75 and RH100) which showed less stable relationships with AGB among different sites.

The analysis of all lidar height metrics and AGB showed the shape of lidar height metrics-AGB correlation curves are significantly different by site(s). The shape of correlation curve is likely caused by vegetation and/or topography structure. The highest lidar-AGB correlations are similar (~0.8) among sites, although the peak correlations are achieved using different lidar height metrics. The number of lidar height metrics having good correlation with AGB reduced when ground data contains larger vegetation and topography variation. The results indicate the amount of AGB variation can be explained by lidar height metrics is similar regardless of forest type, but the optimal lidar height metrics varies and depends on the composition of vegetation structure.

The correlation curves of all lidar height metrics and WWI are similar to those with AGB indicating close relationship between WWI and AGB. The amount of AGB variation that can be explained by WWI is very similar to those achieved by the best lidar height metrics. This indicates the WWI parameter likely contains most of the AGB information available in lidar height metrics. As the optimal lidar height metrics changes by site, the coefficients used to derive AGB are also different. Analysis of the AGB-WWI regression equations derived using different dataset showed the variation in coefficients are very small, much smaller than the AGB equations derived from RH50/75/100. The small deviation in regression coefficients are partly due to the variation in vegetation types, which the relationships among tree crown, height, woody density and AGB are different.

Because ground data required to develop site specific lidar-AGB relationships are limited, the near stable relationship between WWI and AGB presents a step toward large scale AGB retrieval using lidar remote sensing. More comprehensive studies in other

forest types, e.g. tropical forests, savannas, boreal forest, are required to test the WWI parameter and further develop operational lidar AGB retrieval method. As the new generations of spaceborne, global coverage lidar systems, i.e. ICSat II and Dynesty, are scheduled to be in orbit the next few years, an operational lidar AGB retrieval method can potentially bring the global AGB mapping to a new horizon.

Table 4.1: Sample plot structural summaries for all field data used in this study.

	plots	AGB (kg/m^2)	Slope[a] (degree)	Canopy cover[b]
New England	15 deciduous dominate	17.76-29.21 (24.45)	2.6-5.2 (3.7)	0.93-0.97 (0.956)
	33 conifer dominate	8.07-24.78 (14.77)	0.4-2.8 (1.2)	0.4-0.96 (0.74)
Sierra	21	7.04-59.4 (28.94)	3.4-24.4 (9.6)	0.05-0.71 (0.40)

For AGB, slope and canopy cover, range and mean values (in the parenthesis) are listed.

a: slope is derived from Nation Elevation Dataset

b: canopy cover is estimated from LVIS (Ni-Meister et al., 2010)

Table 4.2: Regression equations for field-estimated AGB vs. lidar metrics.

X	a	b	R^2	RMSE
New England (NE)				
RH100	1.79	-22.0	0.68	3.5
RH75	1.53	-6.2	0.79	2.9
RH50*	1.30	1.60	0.83	2.6
WWI	0.068	6.15	0.83	2.5
Sierra (CA)				
RH100	0.88	-10.0	0.33	12.2
RH75	1.25	4.96	0.80	6.7
RH50	1.23	23.1	0.38	11.8
RH73*	1.24	6.88	0.81	6.6
WWI	0.065	10.3	0.75	7.5
New England + Sierra (NE + CA)				
RH100	0.62	3.03	0.46	8.1
RH75	1.41	-2.5	0.76	5.4
RH50	0.46	16.5	0.07	10.6
RH80*	1.37	-3.8	0.78	5.1
WWI	0.07	6.77	0.80	4.9

* Best correlated lidar height metrics

Regression equation:

$$AGB = aX + b$$

Where

AGB = total above ground biomass (kg/m2)

X = lidar metrics

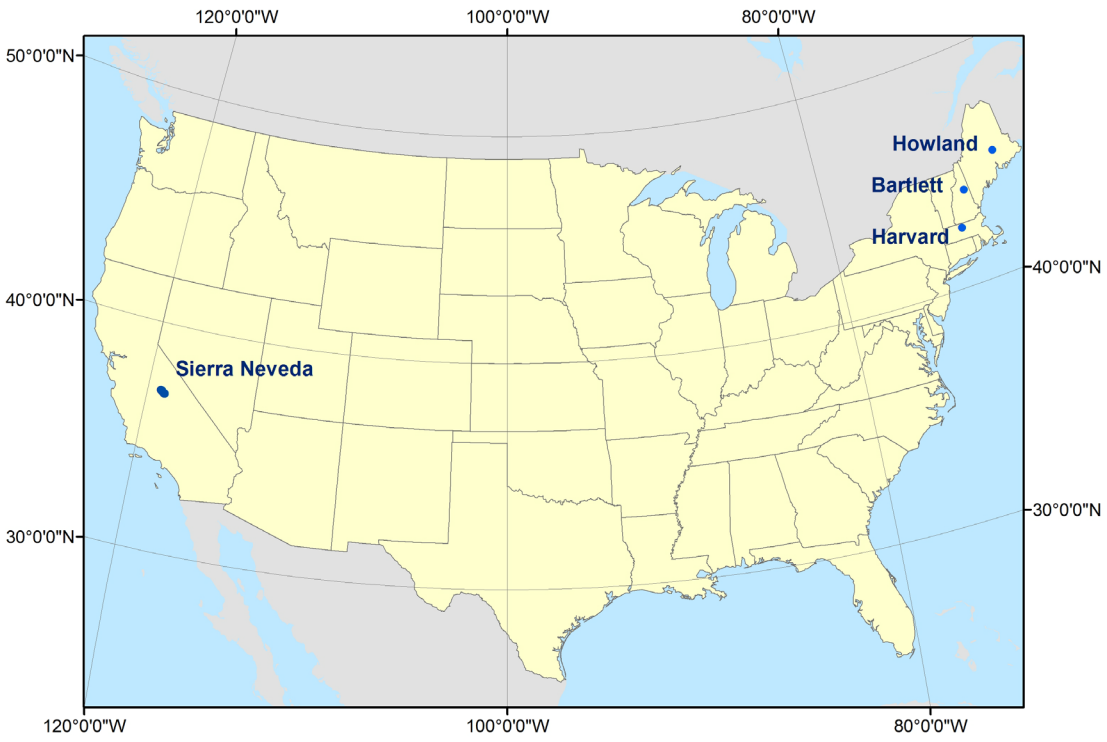

Figure 4.1: Locations of study sites.

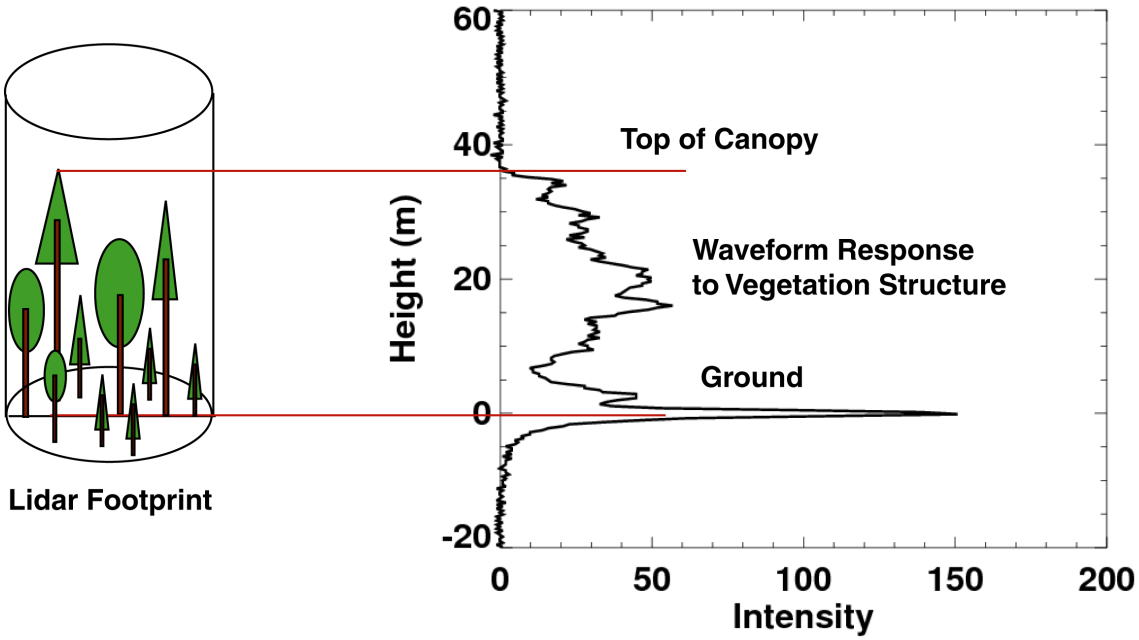

Figure 4.2: Conceptual basis of large-footprint lidar remote sensing. Pulses of laser energy reflect off canopy (e.g., leaves and branches) and ground surfaces, resulting in a waveform, shown in the upper right.

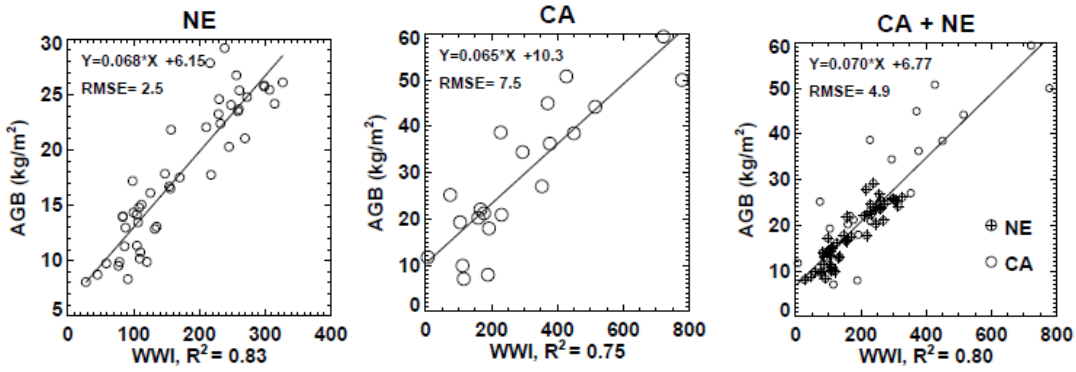

Figure 4.3: Scatter plot of Above Ground Biomass (AGB) with Waveform Weighted Index (WWI) for New England (NE), Sierra National Forest (CA) and the combined dataset (CA + NE).

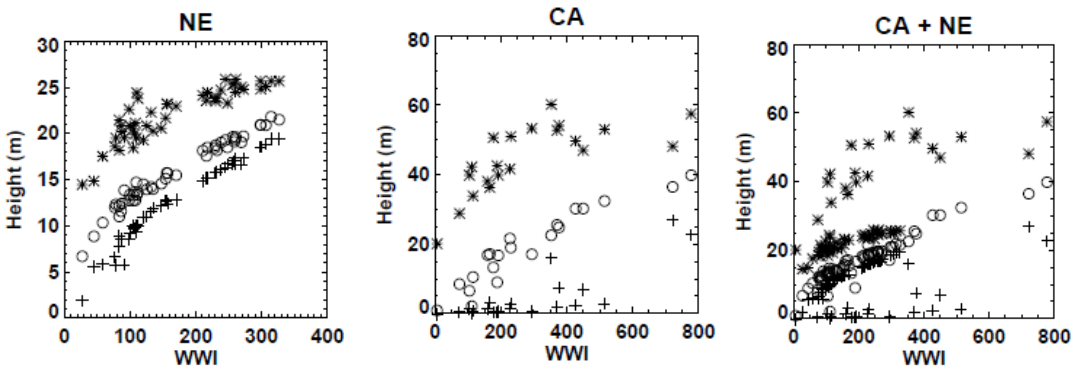

Figure 4.4: Scatter plot of Waveform Weighted Index (WWI) with lidar height metrics (plus sign for RH50, circle for RH75 and star for RH100), for New England (NE), Sierra National Forest (CA) and the combined dataset (CA + NE).

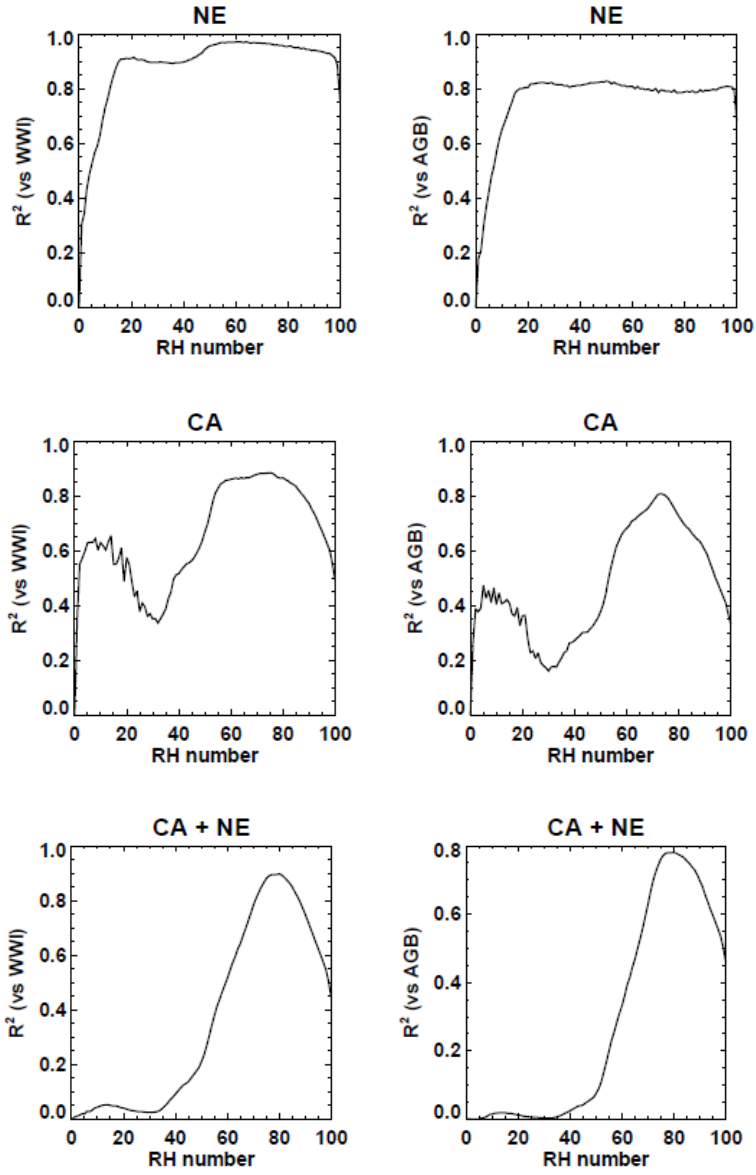

Figure 4.5: Correlations between Waveform Weighted Index (WWI) and RH0 to RH100, and AGB and R0 to RH100. Plots in the top panel show analysis using New England (NE) dataset. Plots in the middle panel show Sierra Forest (CA) dataset. Plots in the bottom panel show all dataset (CA + NE).

Chapter 5 Concluding Remarks

Globally, forests store vast pools of carbon and even small shifts in the balance between photosynthesis and ecosystem respiration can result in a large change in the uptake or emission of carbon dioxide (CO2) from forests to the atmosphere (Pregitzer and Euskirchen 2004). Knowing the spatial-temporal distribution of vegetation structure and biomass can provide a link between disturbances and changes in biomass, which is important in understanding global carbon cycle (Houghton et al., 2005, 2009; Magnani et al., 2007). Direct measurement of forest structure and biomass on the ground is time and financially un-realistic at global scale. The possibility that vegetation structure and above-ground forest biomass might be determined from space is a promising alternative to ground-based methods (Hese et al., 2005, Nelson et al., 2009, Piao et al., 2009). Among remote sensing methods, lidar is most promising in determining three-dimensional vegetation structure and accurately estimate aboveground biomass (Frolking et al., 2009; Hawbaker et al., 2009). Therefore, studies of how to improve lidar remote sensing methods to estimate global forest structure and biomass are important at this juncture.

The first aim of this dissertation was to assess the potential of using spaceborne lidar (GLAS) data for large scale vegetation structure retrieval. Specifically, how to correct the lidar height bias caused by topographical effect (Harding et al., 2005) and whether the vegetation structure content from spaceborne lidar system (GLAS) is as accurate as the LVIS, a proven airborne lidar system in characterizing vegetation structure. Overall, GLAS data was found to contain vegetation structure similar to those in LVIS; however, its accuracy can be affected by large noise and ground saturation.

A simple geometric model was used to quantify the lidar height change with respect to slope and footprint size and a correction scheme was developed. Then, I validated the correction scheme by comparing GLAS and LVIS data collected in White Mountain, NH, USA, with slope reaching 30 degree. The comparison of original GLAS and LVIS shows a height bias that is strongly dependent on slope. Such relationships concur with what is described in the geometric model. After applying the correction, the difference between GLAS and LVIS height is independent of slope. However, further analysis shows large disagreement between GLAS and LVIS height can be caused by tree height distribution. The results indicate topography effect on lidar height can be effectively removed given the slope and footprint size information, but the interpretation of GLAS canopy height should be cautious especially when it were used to estimate above-ground biomass.

In the second part of this dissertation, the relationships of how vegetation structure and biomass changes with respect to lidar derived metrics at various scales were examined. For this, stem inventories and ground lidar data collected from three intensive study forests in New England during summer of 2007 field campaign were utilized in conjunction with airborne lidar data collected in summer of 2003. In order to understand the relationships between vegetation structure and lidar metrics, vegetation structure were estimated using several different methods and the comparison were done in three different spatial scales (lidar footprint, plot and stand level). Lidar metrics associated with vegetation structure parameters found to be linked to aboveground biomass was then used to test the potential of improving the above-ground biomass estimate. The results presented in this dissertation suggest reasonable lidar response to vegetation structure

variations but lidar height metrics representing a slightly different meaning at different scale.

Analysis of above-ground biomass and vegetation structure at plot scale showed knowing woody volume and vegetation type has the potential to improve biomass estimates from height alone. Although such allometric relationship is known at individual tree level, the main contribution is to extend such relationship at scale that vegetation lidar normally operate. Both woody volume and vegetation type can not be directly measured by above-canopy lidar. However, it is possible to approximate woody volume from the combination of height (RH100) and gap fraction (canopy cover) parameters or a single structure parameter (e.g. RH50) that contains both information. In New England sites, RH50 has almost linear relationship with RH100*cover indicating RH50 is a good approximation of both height and gap fraction composition in the forest stand. Comparing to woody volume and biomass, both RH100*cover and RH50 showed high correlation. The results indicate the forest above-ground biomass can be more accurately estimated from lidar metrics resembles woody volume. In New England's forest, RH50 was shown to be the best parameter, however, the optimal lidar metrics could change depending on forest type and topography.

The third part of the study was to develop a more general method so that biomass can be retrieved by lidar at large scale with limited ground calibration. Particularly, how to derive a lidar parameter that has more universal responses to biomass variation in different biomes was discussed. In earlier part of this dissertation, I had presented results showing that forest stand woody volume can be estimated using the product of lidar

canopy height and cover. Taking this concept further, I formulated a lidar biomass index (WWI) by integrating height adjusted energy intensity of the full lidar waveform.

I analyzed the lidar biomass models based on WWI and other lidar height metrics in three different biomes. The analysis showed ~80% of above-ground biomass variation can be explained by lidar height metrics based model, but the optimal model depends on site's vegetation and topography characteristic. WWI based models are about as accurate as models based on the best lidar height metrics. More importantly, the regression coefficients in WWI based models only varied slightly and are partly caused by vegetation type and topography. The results indicate the WWI parameter likely contains most biomass information available in lidar height metrics, and forest type and topography only has minimal impact on it.

With the next generation global lidar missions, ICESat-II and DESDynI, are scheduled to launch within the next few years, a new era will soon begin in the three-dimensional remote sensing of the Earth system (http://nasascience.nasa.gov/missions/icesat-ii). Until then, the continuing efforts of understanding lidar sensed vegetation structure information will provide knowledge necessary for future global vegetation structure and biomass mapping missions. The new generation sensors are expected to provide improved quality and dramatically increase sampling density, which can reveal the Earth surface's three-dimensional structure in the unprecedented detail. Such information will reduce uncertainties in terrestrial carbon estimates and improve our understanding in global carbon cycle. The results can then be turn to better, more effective, coherent efforts among countries to mediate anthropogenic effects on global climate change.

List of Journal Abbreviations

Agr. Forest. Meterorol.	Agricultural and Forest Meteorology
Annu. Rev. Earth Pl. Sc.	Annual Review of Earth and Planetary Sciences
Can. J. Forest Res.	Canadian Journal of Forest Research
Can. J. Remote Sens.	Canadian Journal of Remote Sensing
Ecol. Monogr.	Ecological Monographs
Environ. Pollut.	Environmental Pollution
Forest Ecol. Manag.	Forest Ecology and Management
Geophys. Res. Lett.	Geophysical Research Letters
Global Change Biol.	Global Change Biology
Global Ecol. Biogeogr.	Global Ecology and Bio-geography
IEEE Trans. Geosci. Remote Sens.	IEEE Transactions on Geoscience and Remote Sensing
IEEE Geosci. Remote Sens. Lett.	IEEE Geoscience and Remote sensing Letters
Int. J. Remote Sens.	International Journal of Remote Sensing
ISPRS J. Photogramm.	ISPRS Journal of Photogrammetry and Remote Sensing
J. Appl. Meteor..	Journal of Applied Meteorology
J. Appl. Remote Sens.	Journal of Applied Remote Sensing
J. Geophys. Res.	Journal of Geophysical Research
J. Hydrology	Journal of Hydrology
J. Forestry	Journal of Forestry

Prog. Phys. Geog.	Progress in Physical Geography
Remote Sens. Environ.	Remote Sensing of Environment
Tree Physiol.	Tree Physiology

Bibliography

Abshire, J. B., X. Sun, H. Riris, J. M. Sirota, J. F. McGarry, S. Palm, D. Yi, and P. Liiva, (2005). Geoscience Laser Altimeter System (GLAS) on the ICESat mission: On-orbit measurement performance. *Geophys. Res. Lett.*, 32, L21S02, doi:10.1029/2005GL024028

Agee, J.K. (1981) Biomass of coniferous understory trees in Crater Lake National Park, Oregon. National Park Service Cooperative Park studies unit, college of forest resources university of Washington Seattle, Washington 98195.

Albani, M., D. Medvigy, G. C. Hurtt, and P. R. Moorcroft, (2006). The contributions of land-use change, CO2 fertilization, and climate variability to the Eastern US carbon sink. *Global Change Biol., 12(12)*, 2370--2390.

Anderson, J., M. E. Martin, M-L. Smith, R. O. Dubayah, M. A. Hofton, P. Hyde, B.E. Peterson, J. B. Blair, R.G. Knox, (2006). The use of waveform lidar to measure northern temperate mixed conifer and deciduous forest structure in New Hampshire. *Remote Sens. Environ., 105*, 248-261.

Anderson, J. E., Plourde, L. C., Martin, M. E., Braswell, B. H., Smith, M. L., Dubayah, R. O., Hofton, M. A. and Blair, J. B. (2008). Integrating waveform lidar with hyperspectral imagery for inventory for a northern temperate forest. Remote Sensing of Environment, 112:1856-1870.

Bae, S., and B. E. Schutz (2002), Precision attitude determination (PAD), Geoscience Laser Altimeter System (GLAS) Algorithm Theoretical Basis Document Version 2.2,

http://www.csr.utexas.edu/glas/pdf/atbd_pad_10_02.pdf, last accessed 10 January 2010.

Bergen, K. M., S. J. Goetz, R. O. Dubayah, G. M. Henebry, C. T. Hunsaker, M. L. Imhoff, R. F. Nelson, G. G. Parker, and V. C. Radeloff (2009). Remote sensing of vegetation 3-D structure for biodiversity and habitat: Review and implications for lidar and radar spaceborne missions, *J. Geophys. Res., 114*, G00E06, doi:10.1029/2008JG000883.

Blair, J. B., D. B. Coyle, J. L. Bufton, and D. J. Harding, (1994). Optimization of an airborne laser altimeter for remote sensing of vegetation and tree canopies. *Proceedings of the International Geosciences Remote Sensing Symposium* (pp. 939–941). Pasadena, CA7 California Institute of Technology.

Blair, J. B., D. L. Rabine and M. A. Hofton, (1999). The Laser Vegetation Imaging Sensor: a medium-altitude, digitisation-only, airborne laser altimeter for mapping vegetation and to-pography. *ISPRS J. Photogramm*., 54(2-3), 115-122.

Blair, J. B., Hofton, M. A., and D. L. Rabine, (2006). Processing of NASA LVIS elevation and canopy (LGE, LCE and LGW) data products, version 1.01. https://lvis.gsfc.nasa.gov

Botkin, D. B., and L. G. Simpson, (1990). Biomass of the North America Boreal Forest: A step toward accurate global measures. *Biogeochemistry, 9*, 161−174.

Botkin, D.B., L.G. Simpson, and R.A. Nisbet, (1993). Biomass and carbon storage of the North American deciduous forest. *Biogeochemistry, 20*, 1−17.

Boudreau, J., R. F., Nelson, A. Beaudoin, L. Guindon, D. S. Kimes, and H. A. Margolis, (2008). Regional forest biomass using airborne and spaceborne LiDAR in Québec. *Remote Sens. Environ., 112*, 3876−3890.

Brown, S. (2002). Measuring carbon in forests: current status and future challenges, *Environ. Pollut., 116*, 363-372.

Buermann, W. (2002). The Impact and Response of Vegetation to Climate at Inter-annual Timescales. PhD thesis, 160pp.

Carabajal, C. C., and D. J. Harding, (2005). ICESat validation of SRTM C-band digital elevation models. *Geophys. Res. Lett., 33*, L22S01, doi:10.1029/2005GL023957.

Chave, J., C., Andalo, S. Brown, M. A. Cairns, J. Q. Chambers, D. Eamus, H. Folster, F. Fromard, N. Higuchi, T. Kira, J. P. Lescure, B. W. Nelson, H. Ogawa, H. Puig, B. Riera, and T. Yamakura, (2005). Tree allometry and improved estimation of carbon stocks and balance in tropical forests. *Oecologia, 145*, 87-99.

Chen, Q. (2010a). Assessment of terrain elevation derived from satellite laser altimetry over mountainous forest areas using airborne lidar data. ISPRS Journal of Photogrammetry and Remote Sensing 65(1), 111-122.

Chen, Q. (2010b). Retrieving vegetation height of forests and woodlands over mountainous areas in the Pacific coast region using satellite laser altimetry, Remote Sensing of Environment, , 114(7), 1610-1627.

Chen., Q., Gong, P., Baldocchi, D., and Tian, Y. Q. (2007). Estimating basal area and stem volume for individual trees from lidar data. Photogrammetric Engineering and Remote Sensing, 73(12), 1355-1365.

Chen, Q. (2007). Airborne lidar data processing and information extraction. Photogrammetric Engineering and Remote Sensing, 62:109-112.Curtis, R. O., and D. J. Marshall, (2000). Technical Note: Why quadratic mean diameter? *Western Journal of Applied Forestry, 15(3)*, 137-139.

Curtis, R. O., and Marshall, D. J. (2000). Technical Note: Why quadratic mean diameter? Western Journal of Applied Forestry, 15(3): 137-139.

Drake, J. B., R. O. Dubayah, D. B. Clark, R. G. Knox, J. B. Blair, M. A. Hofton, R. L. Chazdon, J. F. Weishampel, and S. D. Prince (2002a). Estimation of tropical forest structural charac-teristics using large-footprint lidar. *Remote Sens. Environ., 79*, 305-319.

Drake, J. B., R. O. Dubayah, R. G. Knox, D. B. Clark and J. B. Blair (2002b), Sensitivity of large-footprint lidar to canopy structure and biomass in a neotropical rainforest. *Remote Sens. Environ., 81*, 378-392.

Drake, J. B., R. G. Knox, R. O. Dubayah, D. B. Clark, R. Condit, J. B. Blair and M. Hofton (2003). Above-ground biomass estimation in closed canopy neotropical forests using lidar remote sensing: factors affecting the generality of relationships. *Global Ecol. Biogeogr., 12(2)*, 147-159.

Dubayah, R. O. and J. B. Drake (2000). Lidar remote sensing for forestry. *J. Forestry, 98(6)*, 44-46.

Duncanson, L. I., Niemann, K. O., and Wulder, M. A. (2010). Estimating forest canopy height and terrain relief from GLAS waveform metrics. Remote Sensing of Environment, 114:138-154.

Filip, S.M., and E. L. Little, (1971). Trees and shrubs of the Bartlett Experimental Forest, Carroll County, New Hampshire. Res. Paper NE-211. Upper Darby, PA: U.S. Department of Agriculture, Forest Service, Northeastern Forest Experiment Station, 20 p.

Fournier, R. A., Luther, J. E., Guindon, L., Lambert, M. -C., Piercy, D., Hall, R. J., et al. (2003). Mapping aboveground tree biomass at the stand level from inventory information: Test cases in Newfoundland and Québec. *Can. J. Forest Res., 33*, 1846−1863.

Frolking, S., M. W. Palace, D. B. Clark, J. Q. Chambers, H. H. Shugart, and G. C. Hurtt (2009), Forest disturbance and recovery: A general review in the context of spaceborne remote sensing of impacts on aboveground biomass and canopy structure, *J. Geophys. Res., 114*, G00E02, doi:10.1029/2008JG000911

Gutierrez, R., A. Neuenschwander, and M.M. Crawford (2005). Development of laser waveform digitization for airborne LIDAR topographic mapping instrumentation. Proc. 2005 *International Geoscience and Remote Sensing Symposium*, Seoul, South Korea, July 25-29, 1154-1157.

Harding, D.J., Blair, J., Rabine, D. and Still, K., (2000). SLICER airborne laser altimeter characterization of canopy structure and sub-canopy topography for the BOREAS Nothern and Southern Study Regions: Instrument and Data Product Description. In Technical Report Series on the Boreal Ecosystem-Atmosphere Study (BOREAS), Volume 93, 45pp.

Harding, D. J., M. A. Lefsky, G. G. Parker and J. B. Blair (2001). Laser altimeter canopy height profiles - Methods and validation for closed-canopy, broadleaf forests. Remote *Sens. Environ., 76(3)*, 283-297.

Harding, D. J., and C. C. Carabajal (2005). ICESat waveform measurements of within-footprint topographic relief and vegetation vertical structure, *Geophys. Res. Lett., 32*, L21S10, doi:10.1029/2005GL023471.

Hawbaker, T. J., N. S. Keuler, A. A. Lesak, T. Gobakken, K. Contrucci, and V. C. Radeloff (2009), Improved estimates of forest vegetation structure and biomass with a LiDAR-optimized sampling design, *J. Geophys. Res., 114*, G00E04, doi:10.1029/2008JG000870.

Hese S., W. Lucht, C. Schmullius et al. (2005). Global biomass mapping for an improved understanding of the CO2 balance – the Earth observation mission carbon-3D. *Remote Sens. Environ., 94*, 94–104.

Holmes, K.W., O.A. Chadwick, and P.C. Kyriakidis (2000). Error in a USGS 30m digital elevation model and its impact on terrain modeling, *J. Hydrology, 233*, 154–173.

Houghton, R.A. (2007). Balancing the Global Carbon Budget, *Annu. Rev. Earth Pl. Sc., 35*, 313-347.

Houghton, R.A. (2005). Aboveground Forest Biomass and the Global Carbon Balance, *Global Change Biol., 11*, 945-958.

Houghton, R. A., F. Hall, and S. J. Goetz (2009), Importance of biomass in the global carbon cycle, *J. Geophys. Res., 114*, G00E03, doi:10.1029/2009JG000935

Hunsaker, C., B. Boroski, and G. Steger (2001). Relations between canopy cover and occurrence and productivity of California spotted owls. Predicting Species Occurences, Issues of Accuracy and Scale. Covelo, CA' Island Press.

Hurtt G. C., P. R. Moorcroft, S. W. Pacala et al. (1998) Terrestrial models and global change: challenges for the future. *Global Change Biol., 4*, 581–590.

Hurtt, G. C., R. Dubayah, J. Drake, P. R. Moorcroft, S. W. Pacala J. B. Blair (2004). Beyond potential vegetation: Combining lidar data and a height-structured model for carbon studies. *Ecol. Appl., 14(3)*, 873−883.

Hyde, P., Dubayah, B. Peterson, J. B. Blair, M. Hofton, C. Hunsaker, R. Knox, and W. Walker, (2005). Mapping forest structure for wildlife habitat analysis using waveform lidar: Validation of montane ecosystems. *Remote Sens. Environ., 102*, 63−73.

Jenkins J. C., D. C., Chojnacky , L. S. Heath, and R. A. Birdsey (2004). Comprehensive database of diameter-based biomass regressions for North American tree species. Gen. Tech. Rep. NE-319. Newtown Square, PA: US Department of Agriculture, Forest Service, Northeast Research Station. 45 p.

Jupp, D. L. B., D. S. Culvenor, J. Lovell, G. Newnham, A.H. Strahler and C.E. Woodcock (2009). Estimating forest LAI profiles and structural parameters using a ground based laser called "Echidna®. *Tree Physiol., 29*, 171–181, doi:10.1093/treephys/tpn022

Jupp, D. L. B., D. S. Culvenor, J. Lovell and G. Newnham, (2005). Evaluation and Validation of Canopy Laser Radar (LIDAR) Systems for Native and Plantation Forest Inventory. Canberra, ACT, Australia, CSIRO Earth Observation Centre and CSIRO Forestry and Forest Products Division: 150 pp.

Ketterings Q. M., R. Coe, M. van Noordwijk, Y. Ambagau C. A. Palm, (2001), Reducing uncertainty in the use of allometric biomass equations for predicting above-ground tree biomass in mixed secondary forests, *Forest Ecol. Manag., 146*, 199-209.

Koetz, B., F. Morsdorf, G. Sun, K.J. Ranson, K. Itten,, and B. Allgower, (2006). Inversion of a lidar waveform model for forest biophysical parameter estimation. *IEEE Geosci. Remote Sens. Letts.*, 1545-598.

Lee, S., W. Ni-Meister and W. Yang (2010), Physically based vertical vegetation structure retrieval from ICESat data: Validation using LVIS in White Mountain National Forest, New Hampshire, USA, Remote Sens. Environ., (accepted).

Lefsky, M.A, D. Harding, W. Cohen, G. Parker, and H. Shugart, (1999a), Surface lidar remote sensing of basal area and biomass in deciduous forests of easterm Maryland, USA. *Remote Sens. Environ., 67*, 83-96.

Lefsky, M. A., W. B. Cohen, S. A. Acker, G. G. Parker, T. A. Spies and D. Harding, (1999b). Lidar remote sensing of the canopy structure and biophysical properties of Douglas-fir western hemlock forests. *Remote Sens. Environ., 70(3)*, 339-361.

Lefsky, M. A., W. B. Cohen, and R. A. Spies (2001). An evaluation of alternate remote sensing products for forest inventory, monitoring, and mapping of Douglas-fir forests in western Oregon. *Can. J. Forest Res, 31*, 78-87.

Lefsky, M. A., W. B. Cohen, D. J. Harding, G. G. Parker, S. A. Acker and S. T. Gower (2002a). Lidar remote sensing of above-ground biomass in three biomes. *Global Ecol. Biogeogr., 11(5)*, 393-399.

Lefsky, M. A., W. B. Cohen, G. G. Parker and D. J. Harding, (2002b). Lidar remote sensing for ecosystem. *Bioscience, 52(1)*, 19-30.

Lefsky, M.A., Turner, D.P., Guzy, M., Cohenc, W.B. (2005a) Combining lidar estimates of aboveground biomass and Landsat estimates of stand age for spatially extensive validation of modeled forest productivity. *Remote Sens. Environ., 70(3)*, 339-361.

Lefsky, M. A., D. J. Harding, M. Keller, W. B. Cohen, C. C. Carabajal, F. D. B. EspiritoSanto, M. O. Hunter, and R. de Oliveira Jr. (2005b). Estimates of Forest Canopy Height and Aboveground Biomass using ICESat, *Geophys. Res. Lett., 32*, L22S02, doi:10.1029/2005GL023971.

Lefsky, M. A., M. Keller, Y. Pang, P. de Camargo, and M. O. Hunter, (2007). Revised method for forest canopy height estimation from the Geoscience Laser Altimeter System waveforms. *J. Appl. Remote Sens., 1*, 013537, doi:10.1117/1.2795724.

Lim, K., Treitz, P., Wulder, M., St-Onge, B., and Flood, M, (2003). LiDAR Remote Sensing of Forest Structure. *Prog. Phys. Geog., 27*, 88-106.

Liu, W., Song, C., Schroder, T. A. and Cohen, W. B., (2008). Predicting forest successional stages with multitemporal Landsat imagery and forest inventory and analysis data. *Int. J. Remote Sens., 29(13)*, 3855-3872.

Luyssaert, S., Schulze, E.D., Borner, A., Knohl, A., Hessenmoller, D., Law, B.E., Ciais, P., and G., J. (2009). Old-growth forests as global carbon sinks. *Nature, 455*, 213–215.

Magnani, F., Mencuccini M, and Borghetti M et al. (2007). The human footprint in the carbon cycle of temperate and boreal forests. *Nature, 447*, 848–850.

Maser, C. (1989), Forest Primeval—The Natural History of an Ancient Forest, Sierra Club Books, San Francisco, 282 pp.

Means, J. 1999. Design, capability and uses of large-footprint and small-footprint lidar systems. In Csatho, B.M., editor, *International Archieves of the Phtogrammetry and Remote Sensing*. Volume 32, Part 3-W14, La Jolla, California, 9-11 November, 201-207.

Miles. J. (1979), Vegetation Dynamics, Chapman & Hall, London.

National Research Council (NRC) (2007). Earth Science and Applications from Space: National Imperatives for the Next Decade and Beyond, The National Academies Press.

Nelson, R., W. Krabill and G. MacLean (1984). Determining forest canopy characteristics using airborne laser data. *Remote Sens. Environ., 15(3)*, 201-212.

Nelson, R., W. Krabill, and J. Tonelli, (1988). Estimating forest biomass and volume using airborne laser data. *Remote Sens. Environ., 24*, 247-267.

Nelson, R., R. Oderwald and T. G. Gregoire, (1997). Separating the ground and airborne laser sampling phases to estimate tropical forest basal area, volume, and biomass, *Remote Sens. Environ., 60*, 311-326.

Nelson R. F., R. Hyde, P. Johnson, B. Emessiene M. L. Imhoff, R. Campbell, W. Edwards (2007). Investigating Radar-Lidar synergy in a North Carolina pine forest, *Remote Sens. Environ., 11*, 98-108.

Nelson, R. K. J. Ranson, G. Sun, D. S. Kimes, V. Kharuk, P. Montesano, (2009). Estimating Siberian timber volume using MODIS and ICEsat/GLAS, *Remote Sens. Environ., 113*, 691-701.

Neuenschwander, A. L., T. J. Urban, R. Gutierrez, and B. E. Schutz, (2008). Characterization of ICESat/GLAS waveforms over terrestrial ecosystems:

Implications for vegetation mapping, *J. Geophys. Res., 113*, G02S03, doi:10.1029/2007JG000557.

Niklas, K. J. (1996), Plant Allometry: The Scaling of Form and Process, University of Chicago Press, Chicago, 395 pp.

Ni-Meister, W., D. L. B. Jupp and R. Dubayah, (2001). Modeling lidar waveforms in heterogeneous and discrete canopies. *IEEE Trans. Geosci. Remote Sens., 39(9)*, 1943-1958.

Ni-Meister, W., Yang, W., & Kiang, N. Y. (2009). A clumped-foliage canopy radiative transfer model for a global dynamic terrestrial ecosystem model I: Theory. Agr. *Forest. Meterorol.* (accepted).

Ni-Meister, W., S. Lee, A. Strahler, C. E. Woodcock, C. Schaaf, T. Yao, G. Sun, K. J. Ranson, and J. B. Blair (2009). Assessing general relationship between above-ground biomass and vegetation structure parameters for improved carbon estimate from vegetation lidar. *J. Geophys. Res.* (in press).

Oliver, C. D., and Larson, B. C. (1996), Forest Stand Dynamics, Wiley, New York, 435 pp.

Pacala, S., Canham, C., Saponara, J., Silander, J., Kobe, R., and Ribbens, E. (1996). Forest models defined by field measurements: II. Estimation, error analysis and dynamics, *Ecol. Monogr., 66*, 1-44.

Pang, Y., Lefsky, M., Andersen, H., Miller, M. E., Sherill, K. (2008). Validation of the ICESat vegetation product using crown-area-weighted mean height derived using crown delineation with discrete return lidar data. *Can. J. Remote Sens., 34(Suppl.2)*:S471-S484.

Patenaude, G., R. A. Hill, R. Milne, D. L. A. Gaveau, B. B. J. Briggs and T. P. Dawson (2004). Quantifying forest above ground carbon content using LiDAR remote sensing. *Remote Sens. Environ., 93(3)*, 368-380.

Piao S., Fang, J., Ciais, P., Peylin, P, Huang, Y, Sitch, S., and Wang, T. (2009). The carbon balance of terrestrial ecosystems in China. *Nature, 458*, 1009-1015.

Pilli, R., R. Anfodillo and M. Carrer, (2006). Toward a functional and simplified allometry for estimating forest biomass, *Forest Ecol. Manag., 237*, 583-593.

Popescu, S. C. and Zhao, K. (2008). A voxel-based lidar method for assessing crown base height. Remote Sensing of Environment 112(3): 767-781.

Popescu, S. C. (2007). Estimating biomass of individual pine trees using airborne lidar. Biomass & Bioenergy 31(9): 646-655.

Popescu , S. C., and Wynne, R. H. (2004). Seeing the trees in the forest: using lidar and multispectral data fusion with local filtering and variable window size for estimating tree height. Photogrammetric Engineering & Remote Sensing, 70(5): 589-604.

Pregitzer, K. S. and E. S. Euskirchen, (2004). Carbon cycling and storage in world forests: biome patterns related to forest age. *Globle Change Biol., 10*, 2052–2077.

Rosette, J. A., North, P. R. J., Suarez, J. C., and Armston, J. D. (2009). A comparison of biophysical retrieval for forestry using airborne and satellite lidar. International Journal of Remote Sensing, 30(19):5229-5237. DOI: 10.1080/01431160903022944.

Rosette, J. A. B., North, P. R. J., Suarez, J. C. (2008). Vegetation height estimates for a mixed temperate forest using satellite laser altimetry. International Journal of Remote Sensing, 29(5):1475-1493. . DOI: /10.1080/01431160701736380.

Sarabandi, K., and Y.C. Lin, (2000). Simulation of Interferometric SAR Response for Characterization of Scattering Phase Center Statistics of Forest Canopies", *IEEE Trans. Geosci. Remote Sens., 38 (1)*, 115-125.

Schutz, B. E., H. J. Zwally, C.A. Shuman, D. Hancock, and J. P. DiMarzio, (2005). Overview of the ICESat Mission. *Geophys. Res. Lett., 32*, L21S01, doi:10.1029/2005GL024009,

Simard, M, V. Rivera-Monroy, J. Mancera-Pineda, E. Castañeda-Moya, and R. R. Twilley, (2008). A systematic method for 3D mapping of mangrove forests based on Shuttle Radar Topography Mission elevation data, ICEsat/GLAS waveforms and field data: Application to Cienaga Grande de Santa Marta, Columbia, *Remote Sens. Environ., 112*, 2131–2144

Song, C., Schroeder, T. A. Cohen, W. B., (2007). Predicting temperate conifer forest successional stage distributions with multitemporal Landsat Thematic Mapper imagery. *Remote Sens. Environ., 106*, 228-237.

Strahler, A. H, D. L. B. Jupp, C. E. Woodcock, C. B. Schaaf, T. Yao, F. Zhao, X. Yang, J. Lovell, D. Culvenor, G. Newnham, W. Ni-Meister, and W. Boykin-Morris, (2008). Retrieval of Forest Structural Parameters Using a Ground-Based Lidar Instrument (Echidna®), *Can. J. Remote Sens., 34(Suppl. 2)*, S426-S440.

Spurr, S. H. (1956). Forest associations of the Harvard Forest. *Ecol. Monogr., 26*, 245-262.

Sud, Y.C., J. Shukla, and Y. Mintz, (1988). Influence of land-surface roughness on atmospheric circulation and precipitation: A sensitivity study with a general circulation model. *J. Appl. Meteor., 27*, 1036-1054.

Sun G., K.J. Ranson, D.S. Kimes, J.B. Balir, and K. Kovacs, (2008). Forest vertical structure from GLAS: An evaluation using LVIS and SRTM data. *Remote Sens. Environ., 112*, 107-117.

Swatantran, A., R. Dubayah, D. Roberts, M. Hofton and J. B. Blair, (2009). Mapping biomass and stress in Sierra Nevada using lidar and hyperspectral data fusion, *Remote Sens. Environ.* (accepted).

Tajchman, S.J., and H.V. Jr. Wiant, (1983), Topography and biomass characteristics of a forested catchment in the Northern Appalachians, *Forest Ecol. Manag., 5(1)*, 55-69.

Ter-Mikaelian, M.T. and M. D. Korzukhin, (1997). Biomass equations for sixty-five North American tree species. *Forest Ecol. Manag., 97*, 1–24.

Tritton, L. M. and J.W. Hornbeck, (1982). Biomass equations for major tree species of the Northeast. USDA Forest Service, Northeastern Forest Experiment Station General Technical Report NE-69.

U.S. Forest Inventory Analysis, 2009. FIA Database Description and Users Manual for Phase 2, version 4.0, revision 1.

Waddell, K., & Hiserote, B. (2005). Technical documentation for the integrated database, version 2.0. USDA Forest Service Pacific Northwest Research Station.

Wehr, A. and Lohr, U. 1999: Airborne laser scanning – an introduction and overview. *ISPRS J. Photogramm., 54*, 68-82.

West, G.B., Brown, J.H., Enquist, B.J., (1999). A general model for the structure and allometry of plant vascular systems. *Nature, 400*, 664–667.

Whittaker, R. H., F. H. Bormann, G. E. Likens and T. G. Siccama (1974). The Hubbard Brook Ecosystem Study: Forest biomass and production, *Ecological Monogr., 44(2)*, 233-254.

Yang, W., W. Ni-Meister, S. Lee, A. H. Strahler and F. Zhao (2010), Evaluation of a canopy geometric optical and radiative transfer lidar model, *Remote Sens. Environ.* (in revision).

Yang W., W. Ni-Meister, and S. Lee, (2010). Assessment of the impacts of surface topography, off-nadir pointing and vegetation structure on vegetation lidar waveforms using an extended Geometric Optical and Radiative Transfer model. *Remote Sens. Environ.* (in press).

Yao, T. , X. Yang, F. Gao, Z. Wang, Q. Zhang, D. Jupp, D. Culvenor, G. Newnham, W. Ni-Meister, C. Schaaf, C. Woodcock, and A. Strahler, (2009). Estimation of forest structure parameters using the Echidna® ground-based lidar at New England forest sites. *Remote Sens. Environ.* (accepted).

Young, M. 1986. Optics and lasers: including fibers and optical waveguides. Berlin: Springer Verlag.

Zhao, F., M. Schull, M. Roman-Colon, T. Yao, X. Yang, Z. Wang, Q. Zhang, D. Jupp, D. Culvenor, G. Newnham, W. Ni-Meister, C. Schaaf, C. Woodcock, and A. Strahler, (2009). Comparison of leaf area index and foliage profile retrievals from the Echidna® ground-based lidar, digital hemispherical photography, and LAI-2000 for New England forest sites. *Remote Sens. Environ.* (accepted).

Zianis, D. and M. Mencuccini, M., (2004). On simplifying allometric analyses of forest biomass. Forest Ecol. Manag., 187, 311–332.

Zwally, H. J., B. Schutz, W. Abdalati, J. Abshire, C. Bentley, A. Brenner, J. Bufton, J. Dezio, D. Hancock, D. Harding, T. Herring, B. Minster, K. Quinn, S. Palm, J. Spinhirne and R. Thomas, (2002). ICESat's laser measurements of polar ice, atmosphere, ocean, and land, *J. Geodyn., 34(3-4)*, 405-445, . doi:10.1016/S0264-3707(02)00042