EVALUATION OF AIRBORNE LIDAR AS A TOOL FOR OBTAINING SUSTAINABLE FOREST MANAGEMENT OF MAINE'S FORESTS

By

Rei Hayashi

B.S. Oregon State University, 2003

M.S. University of Georgia, 2006

A THESIS

Submitted in Partial Fulfillment of the

Requirements for the Degree of

Master of Science

(in Forest Resources)

The Graduate School

The University of Maine

December 2014

Advisory Committee:

Steven A. Sader, Professor of Forest Resources, Co-Advisor

Aaron Weiskittel, Associate Professor of Forest Biometrics and Modeling, Co-Advisor

Cynthia Loftin, Associate Professor of Wildlife Ecology and Unit Leader, U.S. Geological Survey, Maine Cooperative Fish and Wildlife Research Unit

LIBRARY RIGHTS STATEMENT

EVALUATION OF AIRBORNE LIDAR AS A TOOL FOR OBTAINING SUSTAINABLE FOREST MANAGEMENT OF MAINE'S FORESTS

By Rei Hayashi

Thesis Co-Advisors: Dr. Steven Sader
Dr. Aaron Weiskittel

An Abstract of the Thesis Presented
in Partial Fulfillment of the Requirements for the
Degree of Master of Science
(in Forest Resources)
December 2014

The first objective of this study was to evaluate the applicability of using a low density (*ca.* 1 point m^{-2}) discrete-return LiDAR for predicting maximum tree height, stem density, basal area, quadratic mean diameter, and stem volume using an area-based approach. The research was conducted at the Penobscot Experimental Forest in central Maine, where a range of stand structures and species composition is present and generally representative of northern Maines forests. Using a variety of high dimensional LiDAR metrics, a prediction model was developed using random forest, a nonparametric approach, based on reference data collected in fixed radius circular plots. For comparison, the volume model used two sets of reference data with one being fixed radius circular plots and the other were variable sampling plots. Prediction biases were evaluated with respect to five silvicultural treatments and softwood species composition based on the coefficient of determination (R^2), root mean square error, and mean bias as well as residual scatter plots. LiDAR tended to underestimate forest inventory attributes, regardless of silvicultural treatments and species composition. However, the unmanaged units had particularly larger prediction biases, while the prediction biases also tended to be larger when softwood species composition was greatest. The maximum tree height model had the largest R^2 (86.9%) followed by the volume model (72.1%), while the stem density had the smallest (R^2) (28.7%). Reference data collected in the 0.08-ha fixed radius circular plots resulted in a volume prediction model with a larger R^2. While it was difficult to develop models with a large (R^2) owing to

complexities of Maines forest structures and species composition, low density LiDAR with the area-based approach can be used as a supporting tool in forest management for this region.

The second objective of this thesis was to investigate the applicability of low density (*ca.* 3 pulses m^{-2}) LiDAR data to deploy an individual tree-based approach. Specifically, this study focused on species classifications as well as total height and volume predictions for stem mapped trees. The research was conducted at the Penobscot Experimental Forests in central Maine, where a range of stand structures and species composition is present and generally representative of northern Maine's forests. First, a random forest technique classified species type and softwood species based on LiDAR metrics. Second, the random forest technique was employed to calibrated individual tree height and volume prediction models. Classification errors for species were evaluated with a confusion matrix, while height and volume prediction biases were evaluated based on the coefficient of determination R^2, root mean square error, and mean bias, as well as residual scatter plots with respect to three silvicultural treatments and softwood species composition. Overall, both species type and softwood species classifications had poor classification accuracy, inferring that calibration of LiDAR pulse intensity is necessary. Also, the height and volume models had small R^2 values of 0.38 and 0.30, respectively. This limited accuracy of both models likely is caused by low LiDAR pulse density, which prevents an accurate representation of trees in subcanopy positions.

THESIS ACCEPTANCE STATEMENT

On behalf of the Graduate Committee for Rei Hayashi, I affirm that this manuscript is the final and accepted thesis. Signatures of all committee members are on file with the Graduate School at the University of Maine, 42 Stodder Hall, Orono, Maine.

____ Date: _____________

Dr. Steven Sader
Professor of Forest Resources

____ Date: _____________

Dr. Aaron Weiskittel
Associate Professor of Forest Biometrics and Modeling

ACKNOWLEDGMENTS

First and foremost, I would like to thank my chair advisors, Dr. Steven Sader and Dr. Aaron Weiskittel, for guiding me through this thesis but more so for inspiring me to perform my best throughout my graduate education. Also, I was glad that I could successfully publish a journal article based on research conducted for Chapter 1 of my thesis. Without your patience through multiple manuscript revisions, I could not have accomplished this. Additionally, I have learned extensively from a number of researchers at several conferences (NESAF and NEMO), a workshop (ISPRS Summer LiDAR School) and an internship (NASA) as well as other seminars thanks to your funding. I am also grateful to Dr. Weiskittel for supporting my graduate assistantship for an extra two years. Due to your extended guidance and patience, I had the opportunity to publish another article, currently under review, in addition to my thesis topics.

Of course, I would also like to thank Dr. Cynthia Loftin for graciously serving on my advisory committee. I appreciate that you provided me with valuable advice and suggestions for my thesis and at my defense. Furthermore, I have learned key concepts about landscape ecology from several of the articles you suggested for my graduate education.

I would like to thank Dr. Andrew Finley at Michigan State University and Dr. Jeremy Wilson at the Harris Center for Conservation Education for providing me with valuable advice.

I would like to send a special thanks to Dr. John Kershaw at the University of New Brunswick for providing me with priceless datasets and research topics. My hope is to publish at least one paper, currently under review, in the near future.

I would like to thank Dave Sandilands and Sam Grimm at the School of Forest Resources for helping me revise my resume and cover letters. While I was working on research, your support enabled me to complete numerous job applications.

Last but not least, I would like to thank many former and current SFR graduate students and research associates, Dr. Andrew Nelson, Ben Rice, Brianne Du Clos, Jereme Frank, Josh Puhlick, Kasey Legaard, Dr. Matt Russell, Pat Clune, and Patrick Hiesl.

TABLE OF CONTENTS

LIST OF TABLES

LIST OF FIGURES

CHAPTER 1

ASSESSING THE FEASIBILITY OF LOW DENSITY LIDAR FOR STAND INVENTORY ATTRIBUTE PREDICTIONS IN COMPLEX AND MANAGED FORESTS OF NORTHERN MAINE, USA

1.1 Introduction

Data on forest structure, such as stem density, basal area, and timber volume are used in both strategic and tactical forest management plans. To achieve goals of sustainable forest management, managers need to acquire accurate forest structural and conditional information for a variety of spatial scales including the stand, landscape, or regional levels, depending on management objectives. Conventionally, information acquired at the ground plot level is collected and expanded for estimates of total tree volume per stand, per county or even larger areas. However, conventional field measurements generally consist of a limited number of sampling plots that are established in stands where the forest structural variability within and between stands would not be accounted for Stone et al. (2011).

In contrast, airborne discrete-return LiDAR (Light Detection and Ranging), a type of remote sensing, has been widely accepted as an appropriate technology and supporting tool in ecosystem studies and sustainable forest management Akay et al. (2009); Hudak et al. (2009); Lefsky et al. (2002); Woods et al. (2011). Using an airborne discrete-return LiDAR system, forest managers can deploy a robust and reliable data sampling approach to complement conventional field measurements for estimating volume or other forest inventory attributes from the plot to landscape level Evans et al. (2009); Woods et al. (2011). The ability of LiDAR to inform forest management decisions has been demonstrated in a range of forest types, including boreal forest Woods et al. (2011), mixed softwood Hummel et al. (2011), mixed hardwood Anderson and Bolstad (2013), and single-species softwood plantations Goerndt et al. (2010).

However, there are three major issues associated with airborne discrete-return LiDAR system based forest inventory estimations. First, LiDAR tends to underestimate tree heights because pulse hits directly on treetops are generally insufficient Næsset (1997); Clark et al. (2004). Also, pulse returns are difficult to discriminate either from other nearby treetops or objects other than treetops (*e.g.* from bare-ground, understory vegetation, sides of crowns), a problem which has been avoided in some studies by arbitrarily defining a fixed threshold height García et al. (2010); Goerndt et al. (2010); Jaskierniak et al. (2011); Næsset (1997). For instance, bare-ground is presumably 1 m below the lowest pulse return to account for height of understory vegetation García et al. (2010); Næsset (1997), which can differ among forest ecosystems and silvicultural regimes. Thus, certain preliminary information is necessary to define threshold heights, particularly in northern Maine where the forests have extensive advance regeneration due to past and present silvicultural treatments McWilliams et al. (2005). Finally, extracting height information accurately at the individual tree level may not be possible from LiDAR data, despite a number of studies that have pursued such a goal Falkowski et al. (2006); Popescu (2007). Thus, LiDAR based predictions need to be carried out with a different approach because conventional volume or biomass equations often require both individual tree diameter and height information.

Regarding the second issue, LiDAR pulse footprint sizes and pulse densities may strongly affect prediction accuracy levels in forest inventory estimations. Nilsson (1996) reported that three different footprint sizes (0.75, 1.50 and 3.00 m in diameter) did not affect mean tree height estimations. However, Thomas et al. (2006) suggested that smaller pulse footprint sizes might be suitable for acquiring subdominant canopy information, while Zimble et al. (2003) reported that a low pulse density LiDAR (0.5 pulses m^{-2}) resulted in insufficient pulse direct hits on treetops; thus, height estimations at the stand-level were significantly underestimated compared to field measured tree heights. Popescu and Wynne (2004) as well as Falkowski et al. (2006) suggested that individual tree based estimation need a rather high LiDAR pulse density (6-8 pulses m^{-2}) to provide a sufficient number of pulse hits at treetops. However, some prior studies established strong correlations between LiDAR metrics and forest inventory attributes on plot-level based on low pulse density LiDAR (< 2 pulse m^{-2}) Hawbaker et al. (2010); Jensen et al. (2006); Means et al. (2000); Næsset (2004); Thomas et al. (2006). For example, Treitz et al. (2012) reported that a low pulse density, such as 0.5 pulses m^{-2} was sufficient for forest inventory attribute prediction regarding tactical forest management. Thus, if the objective is to predict forest attributes at the plot and stand level (instead of individual tree level), relatively low pulse density LiDAR should be sufficient.

Regarding the third issue, few LiDAR studies have been reported for relatively complex forest structures, such as those that are predominant in northern Maine. While a number of studies have reported that low pulse density LiDAR metrics and field measured forest inventory attributes, particularly volume, height, and biomass related attributes, on plot- and stand-levels showed relatively high coefficient of determination (R^2) in other forest ecosystems, limited work has been in regions dominated by mixed species and multi-canopy stands. Recently, Anderson and Bolstad (2013) evaluated the use of LiDAR in various forest types in the Great Lakes which are quite similar to those found in Maine, and found a strong relationship between ground-based measurements, regardless if the LiDAR data was collected with hardwood species leaf-on or leaf-off.

We assessed the feasibility of predicting various plot and stand level forest inventory attributes based on airborne low-density discrete-return LiDAR in a range of stand structures and species composition that are representative of northern Maine's forest. The primary objectives of this analysis were to: (1) establish empirical relationships between LiDAR data and forest inventory attributes such as maximum tree height, stem density, quadratic mean diameter (QMD), basal area and stem volume; (2) assess prediction accuracy across a range silvicultural treatments and species compositions; and (3) evaluate the influence of reference data acquired from research- and operational-grade sampling protocols on attribute predictions.

1.2 Experimental Section

1.2.1 Study Area

The study was conducted on the Penobscot Experimental Forest (PEF) near Orono, Maine, USA (N44°49'30", W68°39'00") (Figure 1.1). The PEF was established in 1952 by U.S. Forest Service, and a number of studies regarding timber management, stand dynamics, productivity, biological diversity and more have been conducted Sendak et al. (2003). The total area of the PEF is 1,619 ha, and various silvicultural treatments (*e.g.* natural area, clearcut,

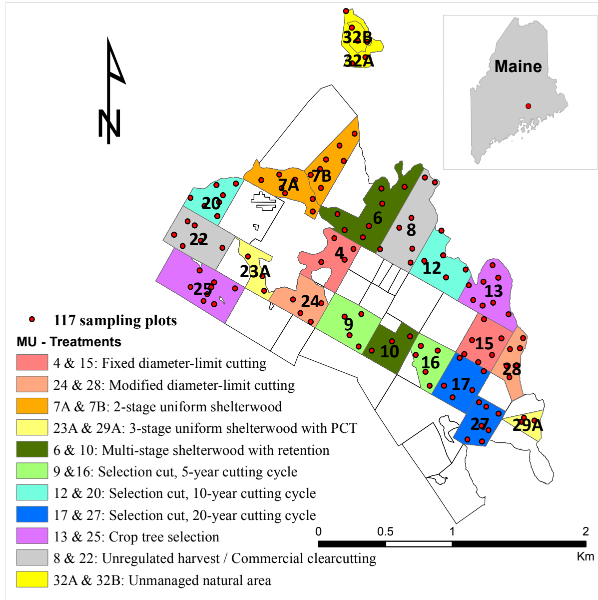

Figure 1.1. The Penobscot Experimental Forest near Orono, Maine, USA (N44°49'30", W68°39'00"). Eleven different silvicultural treatments have been monitored with long term permanent sampling plots.

shelterwood, diameter-limit cutting) have been twice replicated for long term observations. The treatments generally range in size of 0.5 to 22.4 ha and are representative of typical northern Maine's silvicultural practices (Table 1.1). With a few exceptions, most treatments are replicated in the PEF, and field data (*e.g.* DBH) for each of replicated treatments are collected at about 600 permanent sampling plots on a 10-year cycle.

Overall, the PEF is defined as a mixed northern conifer dominant forest as a part of Acadian ecosystem Sendak et al. (2003). The major hardwood species in the PEF are red maple (*Acer rubrum* L.), birches (*Betula* spp.) and aspens (*Populus* spp.), while the major softwood species are spruces (*Picea* spp.), balsam fir (*Abies balsamea* L. (Mill.)), nothern white cedar (*Thuja occidentalis* L.) and eastern white pine (*Pinus strobus* L.). The range of elevation above sea level is between 20 and 70 m.

1.2.2 Inventory Attributes Data

For this study, eleven replicated management units (total of 22 silvicultural treatment units) that varied from 2.86 to 19.58 ha in size were selected (Figure 1.1). Within these 22

Table 1.1. Description of silvicultural treatments in management units (MUs) in the study area.

MU	Area (ha)	Treat-ment year	Inven-tory year	Plot (n)	Description of Silvicultural Treatment	Treatment Group
4	10.1	1994	2009	4	Fixed diameter-limit cutting. Thresholds are 14.0 cm for balsam fir, 24.1 cm for spruce and hemlock, 26.7 cm for white pine, 19.1 cm for cedar and paper birch, and 14.0 cm for other hardwoods.	Diameter-limit (DL)
15	10.3	2001	2007	6		
24	9.4	1996	2005	4	Modified diameter-limit cutting. The third modified diameter-limit cut was applied in 1995. Portions of the stand are in the stem exclusion and understory reinitiation stages of development.	
28	7.3	1997	2007	6		
9	12.2	2003	2003	4	5-year cutting cycle. Structural goal is to retain 24.1 m^2 ha^{-1} (trees > 11.4 cm).	Selection (SEL)
16	8.6	2006	2006	5		
12	12.5	1994	2004	5	10-year cutting cycle. Structural goal is to retain 20.7 m^2 ha^{-1} (trees > 11.4 cm).	
20	8.8	1998	2008	7		
17	10.9	1994	2005	5	20-year cutting cycle. Structural goal is to retain 16.1 m^2 ha^{-1} (trees > 11.4 cm).	
27	8.4	1997	2007	7		
13	13.2	1995	2009	8	Crop tree selection	
25	18.0	2009	2009	8		
7A	10.6	1979	2003	7	2-stage uniform shelterwood. Overstory was removed in two harvests, unmerchantable trees > 5.08 cm in DBH felled after final overstory removal.	Shelterwood (SHE)
7B	10.9	1979	2003	7		
23A	5.3	2007	2007	3	3-stage uniform shelterwood with PCT. Manual PCT to a residual spacing of 2×3 m was applied in 1983. The canopy is not closed, and volunteer growth has occurred between crop trees.	
29A	3.6	2009	2010	3		
6	19.6	1995	2010	7	Multi-stage shelterwood with retention. Overstory will be removed in series of harvests at 10-year intervals, approximately 2 overstory trees $acre^{-1}$ will be retained through the next rotation.	
10	9.2	1995	2010	3		
8	17.6	1983	2008	7	Unregulated harvest/commercial clearcutting. This compartment was initially cut with unregulated (”loggers choice”) harvests. The second harvest was a commercial clearcut in 1982. The stands are in the stand initiation and stem exclusion phases of development.	Clearcut (CC)
22	13.6	1988	2004	6		
32A	5.2	-	2009	3	Unmanaged natural area (partial cutting had been practiced prior to 1900).	Unmanaged (NAT)
32B	2.9	-	2009	3		

management units, a total of 117 permanent sampling plots were established with a range of 3-7 fixed, nested circular permanent sampling plots established in each management unit. On each 0.02-ha (1/20th-acre) permanent sampling plot, diameter at breast height (DBH) were collected from all trees with a DBH greater than 6.35 cm (2.5 inches) between 2003 and 2010, depending on the management unit. On each 0.08-ha (1/5th-acre) permanent sampling plot, DBH was collected from all trees with DBH greater than 11.25 cm (4.5 inches). On a subsample of permanent sampling plots (n = 117), the total height (HT) and height to crown base were measured on all trees within the 0.08-ha plot.

Based on DBH and HT, total tree volume was calculated using a species-specific taper equation Li et al. (2012); Weiskittel and Li (2012). Given the differences between plot measurement and acquisition of the LiDAR data in the fall of 2010, the Acadian Variant of the Forest Vegetation Simulator was used Weiskittel et al. (2012) to project DBH and HT to a common year with the number of projections ranging from 1 to 7 annual cycles. Preliminary results indicated that projected inventory data improved the prediction models in comparison to using data that was not projected. Here after, this sampling method and data are called "research-grade" in this paper. All inventory attributes (maximum tree height, stem density, QMD, basal area, volume) were set in the metric unit at the plot-level, and a total volume prediction was scaled to the management unit level (*e.g.* m^3 management unit^{-1}) from the mean of the plot level data and the total acreage of the unit. Thus, a total of 117 plot-level and 22 management-unit data were available for analysis (Table 1.2).

In addition to the research-grade plots, a total of 44-20 basal area factor (BAF) variable sampling plots were established in a total of nine management units between 2010 and 2011. Locations of the plots were the same as the research-grade plots. At each plot, DBH was measured for all tallied trees, while a local height equation was derived using multi-level mixed effects to impute height values Robinson and Wykoff (2004), and volume was estimated using the same equations as described before. Here after, this sampling approach is called "operational-grade" plots and data in this paper.

Table 1.2. Examined attributes (mean ± standard deviation) by management unit (MU).

MU	Maximum Tree Height (m)		Stem Density (trees ha^{-1})		QMD (cm)		Basal Area (m^2 ha^{-1})		Stem Volume (m^3 ha^{-1})		%Softwood Basal Area	
	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD	mean	SD
4	11.51	2.46	7964	4284	6.6	1.1	24.6	6.9	82.15	25.66	66	5
15	11.02	2.67	6694	4601	6.9	1.3	21.2	7.6	56.74	17.37	81	4
6	12.70	4.18	12167	5224	5.9	1.7	28.9	5.0	110.63	69.06	86	5
10	15.20	3.94	7966	4054	7.3	1.5	30.0	3.6	139.95	46.14	76	5
7A	10.87	1.64	467	271	18.0	0.6	11.6	6.2	88.10	49.54	95	6
7B	10.67	1.79	321	118	17.4	0.7	7.6	2.8	57.02	19.91	84	6
8	11.14	2.17	8507	3170	7.0	1.0	30.4	4.4	58.63	31.30	56	4
22	10.10	2.32	8277	2892	6.3	0.9	24.2	4.1	32.02	22.81	51	3
9	15.62	4.25	3948	2223	11.2	3.2	31.7	5.9	288.53	44.46	91	6
16	15.10	3.79	2281	1691	15.0	5.6	28.6	5.5	253.21	55.88	84	5
12	15.00	3.73	11240	5177	9.0	2.0	63.0	2.8	248.56	27.05	80	7
20	14.13	4.39	15533	10945	7.8	1.9	59.5	7.8	169.17	91.60	71	7
13	12.43	4.35	12668	4663	7.8	1.7	55.2	10.0	133.98	39.81	84	7
25	13.49	4.15	3894	1824	8.7	3.0	18.6	5.0	162.36	33.66	69	5
17	15.73	4.33	7668	4549	8.2	3.2	30.2	6.4	195.10	62.79	87	6
27	13.56	4.13	12126	3450	6.5	1.0	37.9	4.7	144.66	49.83	79	5
24	15.20	3.94	3665	1953	11.2	2.2	32.4	4.0	249.83	76.84	83	4
28	14.33	3.90	4708	2431	10.5	3.7	32.6	4.5	206.60	64.12	77	4
23A	12.94	2.51	6971	2222	8.5	1.4	37.7	2.1	344.73	81.62	76	3
29A	10.60	1.34	1915	1811	10.9	1.7	15.3	10.1	353.03	84.37	98	2
32A	16.25	5.10	8479	2588	7.6	1.5	36.8	7.9	235.11	96.55	63	2
32B	21.48	6.24	864	267	28.1	5.0	50.8	7.8	662.45	88.81	90	1
Overall	13.59	3.51	6742	3200	10.3	2.1	32.2	5.7	194.21	53.60	78	5

1.2.3 LiDAR System Specifications

The LiDAR data were acquired along U.S. Geological Survey National Geospatial Program - Lidar Base Specification Version 1.0 Heidemann (2012). Airborne discrete-return laser scanner data were acquired using an Optech Gemini 246 instrument in the late October 2010, and the mean flying altitude above sea level was about 1,982 m. LiDAR data was intended to be collected under a leaf-off condition, but most deciduous trees in the PEF kept leaves at that time due to an abnormal prolonged summer period in 2010. The sensor generated the pulse repetition frequency of 50 KHz with, and the laser pulse intensity was 1064 nm with the scan angle of $< 20°$ from the nadiar. Mean laser point density was 1.1 pulses m^{-2} with footprint of 30 cm, and the sensor collected up to 4 pulse returns.

1.2.4 LiDAR Data Processing and Model Calibration Predictions

All LiDAR data processing including creation of a digital terrain model and LiDAR metrics were deployed in FUSION v2.90, developed by the U.S. Forest Service Pacific Northwest Research Station McGaughey (2013). The software sorted raw LiDAR data into various metrics containing a number of potential predictor variables of inventory attributes. In our case, 97 potential predictor variables were created. To calibrate prediction models, FUSION extracted raw LiDAR data from 117 0.08-ha circular plots coincidental to the research-grade plots in the management units. On the other hand, for prediction models based on operational-grade sampling, empirical relationships were established between raw LiDAR data extracted from 44 0.08-ha plots coincidental to the research-grade plots, because the size of plots varied.

Although understory vegetation heights varied largely depending on silvicultural treatments in each management unit, we disregarded pulse return within 2 m above ground as preliminary results indicated better model fit (greater R^2 values) during the LiDAR data extraction. A few example predictor variables in the LiDAR metrics were maximum height, the number of 1st return pulses in the 90^{th} percentile height, and standard deviation of 1^{st} return pulses. Consequently, two LiDAR metrics were generated based on research- and operational-grade samples.

However, predictor variables in the LiDAR metrics tend to be highly correlated with others, and some variables would not meet normal distribution criteria Stone et al. (2011); Hudak et al. (2008); Li et al. (2008). These issues violate the assumption inherent to linear regression models. In addition, variable selection with high dimensionality metrics is not a simple process and typical data transformations might not be effective for highly skewed or bimodal data. Although Akaikes Information Criteria (AIC) is a popular approach for variable selection in stepwise regression, the developed regression models tend to have model overfit issues, which are generally not stable when outside of the calibration data. Therefore, the development of inventory prediction models based on simple and multiple linear regressions would not be suitable for this type of dataset.

Alternatively, the random forest technique proposed by Breiman Breiman (2001), a nonparametric approach, may be a more effective technique. Random forest is developed based on the regression trees algorithm, where predictor variables are split to grow a number of nodes to select the best predictor variables. In the random forest approach, the regression tree process is continued to multiple times and compared against a bootstrapped validation dataset. A key advantage in random forest is that a greater number of predictor variables of various types (categorical, continuous, binary) can be handled and the relative importance of each predictor variable can be estimated during the model calibration process. In this analysis, the random forest algorithm was run iteratively in that the model initially included all covariates, the least influential covariate dropped, and the model reran until there were only five covariates, which preliminary analysis had suggested was most effective for prediction accuracy.

Stone et al. (2011) reported that inventory prediction models such as a volume prediction based on random forest had significantly lower R^2 values than prediction models based on other methods, such as regression trees. However, the developed models were based on small number of reference plot data, and some variables in this study required data transformations for meeting normal distribution which random forest might more effectively handle. The 'randomforest' package Liaw and Wiener (2002) available in R v2.15 R Development Core Team (2012), was used to calibrate the inventory attribute prediction models in this analysis. Each of calibrated models was evaluated using the coefficient of determination (R^2), mean bias, and root mean square error (RMSE) between field measured and LiDAR predicted inventory attributes on the plot and management unit levels. For each inventory attribute, prediction models were calibrated based on research- and operational-grade data in the random forest.

To examine the performance of the various models, the bias (observed − predicted) was examined graphically with the use of lowess regression splines. To simplify interpretation of the differences between the original eleven different silvicultural treatments, the treatments were narrowed down to five broad categories, which included diameter-limit, selection, shelterwood, clearcut, and unmanaged (Table 1.1). To examine the influence of species composition, the % of softwood vs. hardwood basal area was computed and the plots were typed as either softwood-dominant (% softwood species ≥ 70) or mixedwood (% softwood species < 70).

Finally, for producing a volume spatial distribution map, a wall-to-wall of 900 m^2 grid cells was overlaid on the PEF area. This size was chosen because it is similar to the size of the research-grade plots, and total volumes in each management unit (m^3 MU^{-1}) were derived as following equation (1.1):

$$Vol_j = \frac{1}{n_j}\left(\sum_{i=1}^{n} Vol_{ij}\right) \times A_j \tag{1.1}$$

Where Vol_j is the total volume (m^3 management unit^{-1}) for management unit j, Vol_{ij} is the volume (m^3 ha^{-1}) for 900 m^2 grid i in management unit j, n_j is the number of grids in management unit j, and A_j is the total area (ha) of management unit j.

1.3 Results

Overall, the random forest technique satisfactorily produced a volume prediction model, but the rest of inventory prediction models did not reach anticipated accuracy levels (Table 1.3). In general, the three most key variables were LiDAR measured height variables rather than pulse return counts.

1.3.1 Maximum Tree Height

Our preliminary analysis indicated that a variable of maximum height elevation was strongly correlated to field measured maximum height. Thus, we did not develop a maximum height prediction model through random forest.

In general, LiDAR underestimated the maximum tree height by 1.89 $\pm$ 2.06 m, regardless of silvicultural treatments and species composition, while an agreement between field and LiDAR measured maximum height was strong (Table 1.3). In particular, the diameter-limit and shelterwood units had a constant trend over the LiDAR measured maximum heights as both RMSEs were relatively small (Table 1.4 and Figure 1.2a). The unmanaged units had the largest mean bias and RMSE, and the largest variation between underestimation and overestimation. Also, LiDAR tended to greatly underestimate heights in softwood plots (Figures 1.2b) with greater mean bias and RMSE than hardwood plots.

1.3.2 Stem Density

In general, LiDAR underestimated the stem density by 9 $\pm$ 5013 trees ha^{-1} regardless of silvicultural treatments and species composition, while an agreement between field measured and LiDAR predicted stem density was weak (Table 1.3) as mean bias and RMSE were fairly large. Although there were no strong spatial trends in the stem density estimation bias based on the silvicultural treatments, the model generally underestimated stem density in the selection and clearcut units, while overestimating it in the diameter-limit, shelterwood, and unmanaged units (Table 1.4 and Figure 1.3a). In particular, the prediction in the diameter-limit units was increasingly overestimated with increasing predicted stem density. Also, the prediction in the clearcut units was fluctuated from underestimation to overestimation with increasing predicted stem density.

1.3.3 Quadratic Mean Diameter

In general, LiDAR overestimated the QMD by -0.05 $\pm$ 3.69 cm regardless of silvicultural treatments and species composition, while an agreement between field measured and

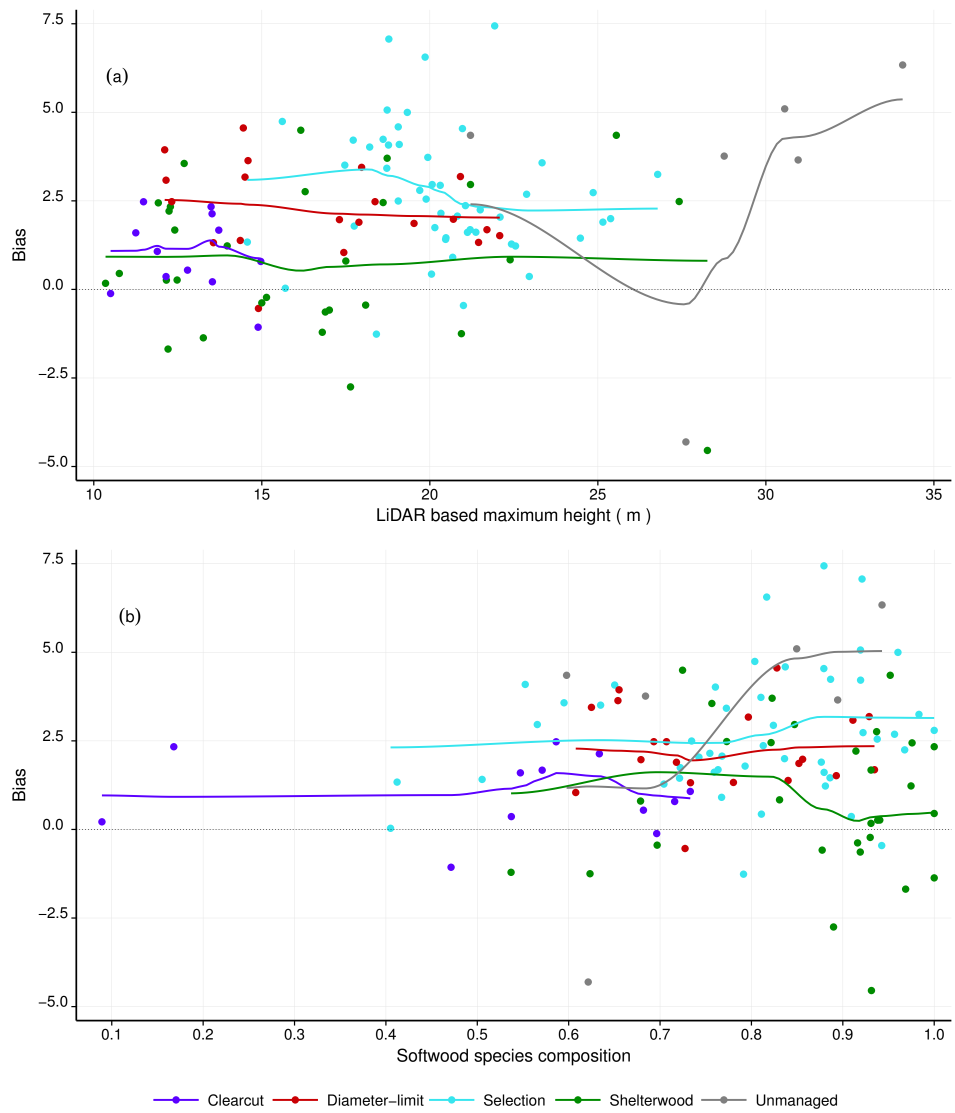

Figure 1.2. Scatterplot of maximum tree height prediction bias (observed - predicted; m) over LiDAR predicted values with lowess regression splines for the different silvicultural treatments **(a)**, and plot species composition based on basal area **(b)**.

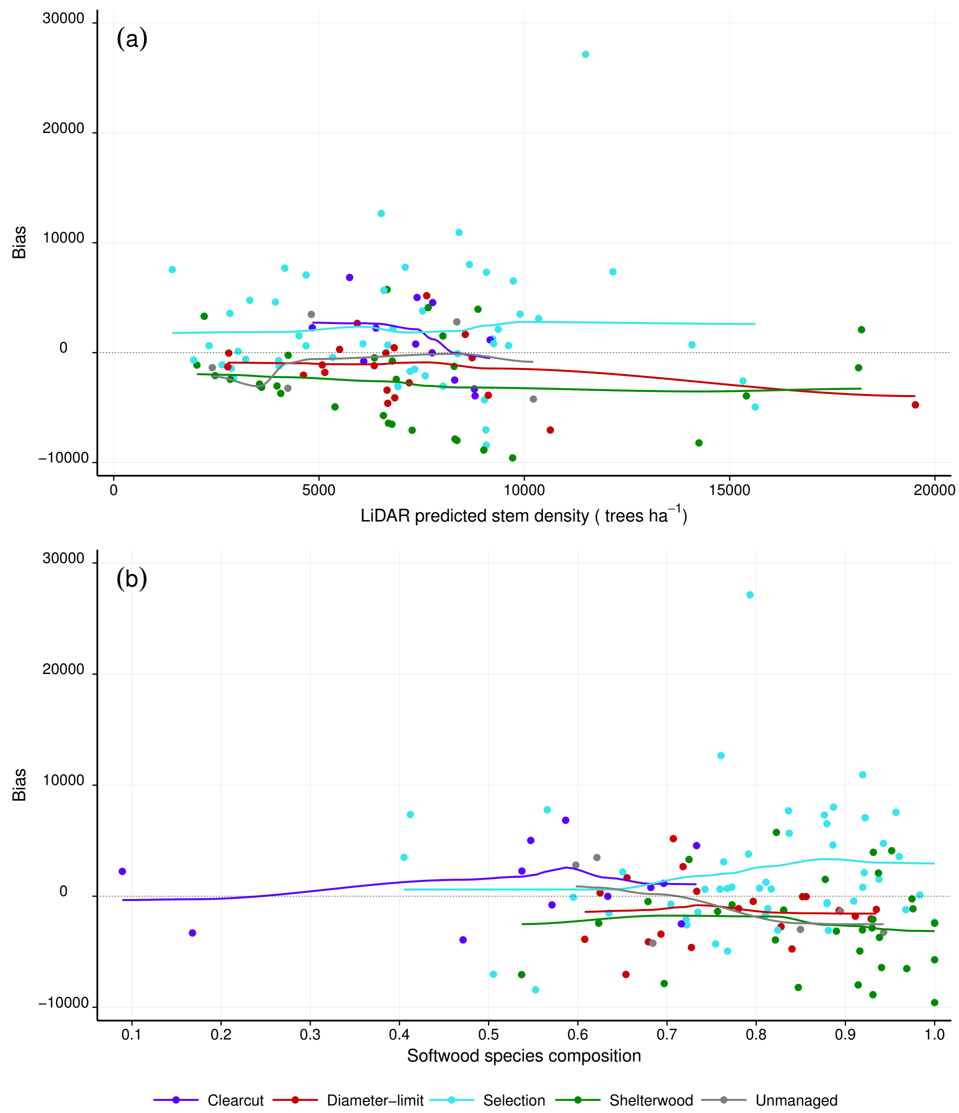

Figure 1.3. Scatterplot of stem density prediction bias (observed - predicted; trees ha^{-1}) over LiDAR predicted values with lowess regression splines for the different silvicultural treatments **(a)**, and plot species composition based on basal area **(b)**.

LiDAR predicted QMD was low (Table 1.3), while the diameter-limit units had the smallest mean bias and RMSE (Table 1.4). In particular, the model largely underestimated in the unmanaged and shelterwood units, and overestimated it in the clearcut and selection units with increasing predicted QMD (Figure 1.4a). Also, within those three management units except the clearcut units, the prediction biases were most prominent in plots with a greater softwood composition (Figure 1.4b).

1.3.4 Basal Area

In general, LiDAR underestimated the basal area by 0.03 $\pm$ 13.07 m^2 ha^{-1} regardless of silvicultural treatments and species composition, while an agreement between field measured and LiDAR predicted basal area was low (Table 1.3). Regarding RMSE, the shelterwood and selection units had the largest precision bias (Table 1.4). While the model constantly overestimated in the shelterwood and diameter-limit units, the prediction in the clearcut units tended to fluctuate from underestimation to overestimation with increasing the predicted basal area (Figure 1.5a). Also, plots with greater softwood species composition tended to have larger prediction biases (Figure 1.5b).

1.3.5 Stem Volume

In general, LiDAR underestimated the volume by 1.81 66.96 m^3 ha^{-1} across silvicultural treatments and species composition, while the plot-level volume prediction model based on the 117 research-grade plots achieved a relatively strong agreement between field measured and LiDAR predicted volume (Table 1.3). The prediction bias in the clearcut and diameter-limit units was fairly constant as those RMSEs were relatively small, while predictions particularly in the shelterwood and unmanaged units were varied over the predicted volume as those RMSEs were large (Table 1.4 and Figure 1.6a). In general, the model underestimated the volume in the selection and unmanaged units, while overestimated in the diameter-limit and clearcut units. The prediction in the shelterwood units varied between underestimation and overestimation with increasing predicted volume. Except for the selection units, prediction biases tended to increase with greater softwood species composition (Figure 1.6b).

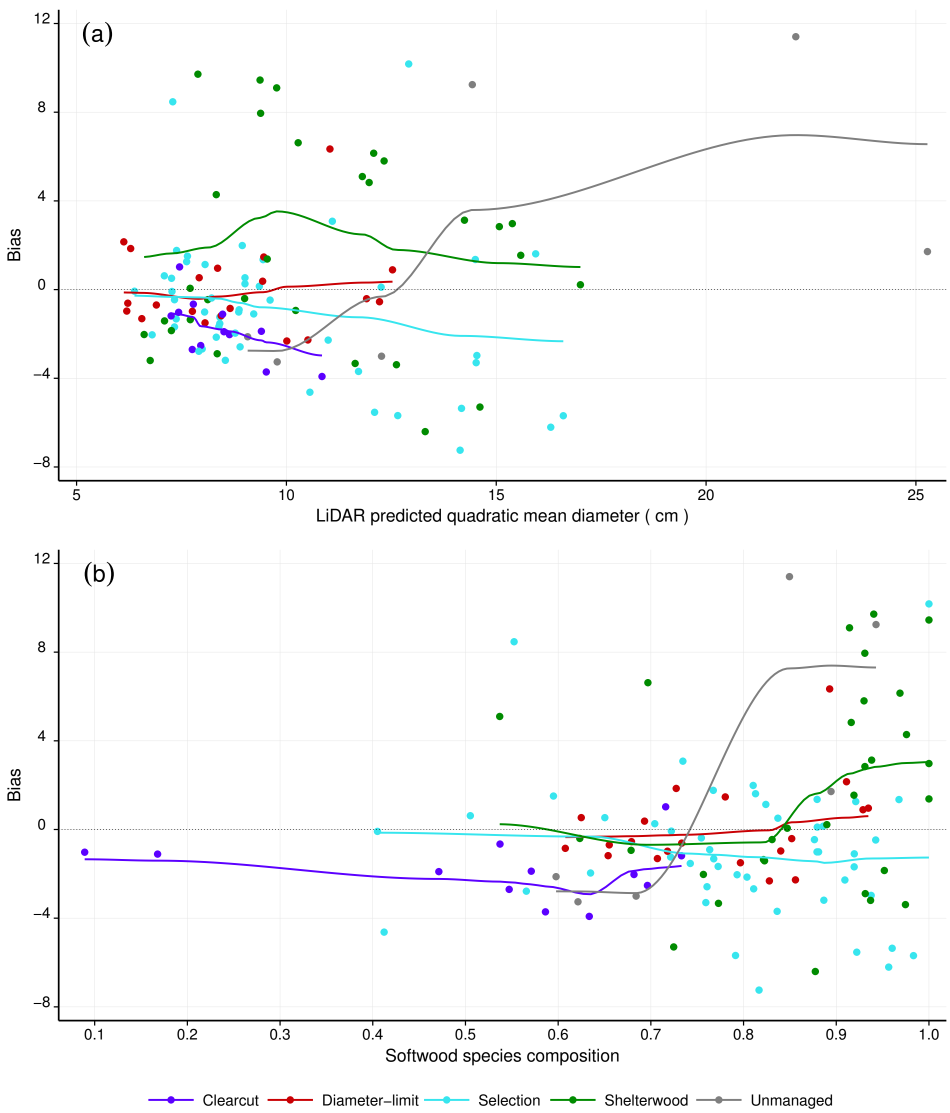

Figure 1.4. Scatterplot of quadratic mean diameter prediction bias (observed - predicted; cm) over LiDAR predicted values with lowess regression splines for the different silvicultural treatments **(a)**, and plot species composition based on basal area **(b)**.

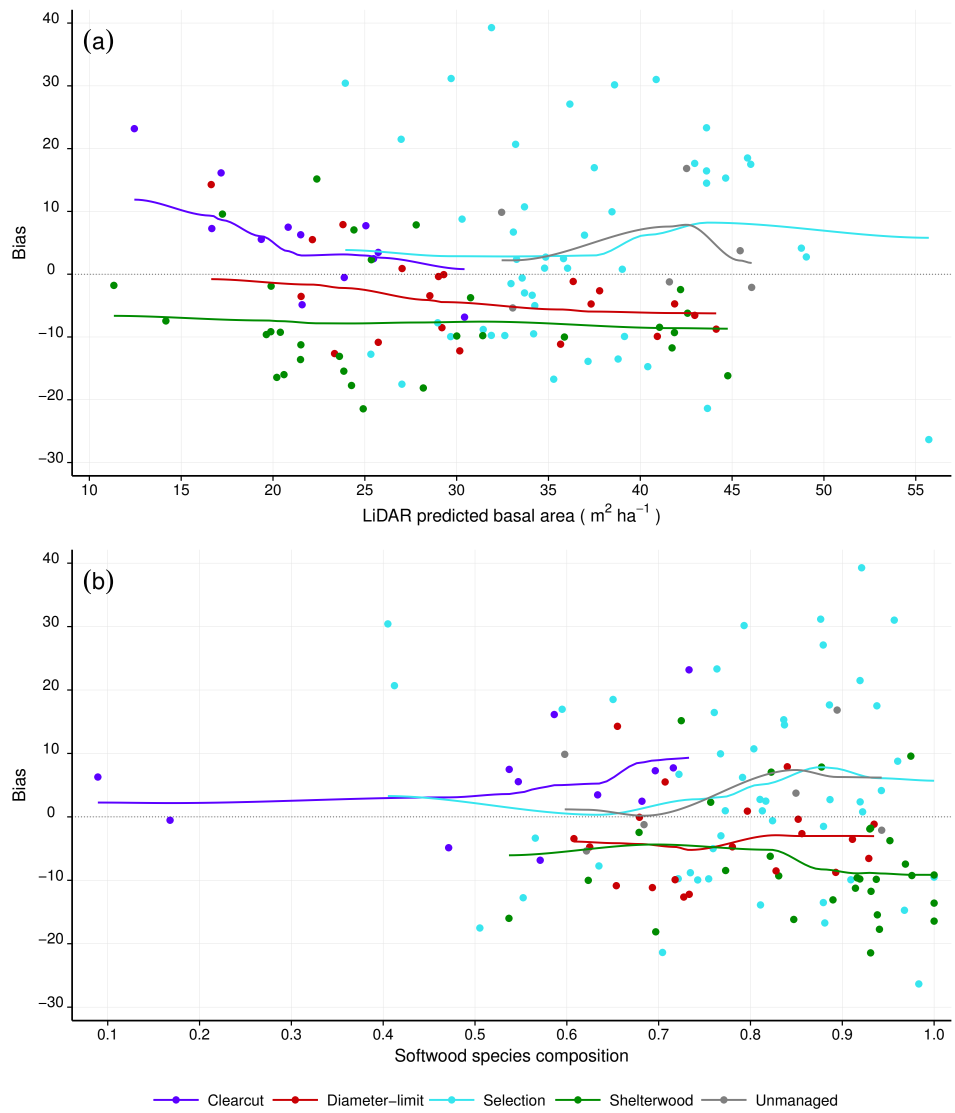

Figure 1.5. Scatterplot of basal area prediction bias (observed - predicted; m^2 ha^{-1}) over LiDAR predicted values with lowess regression splines for the different silvicultural treatments **(a)**, and plot species composition based on basal area **(b)**.

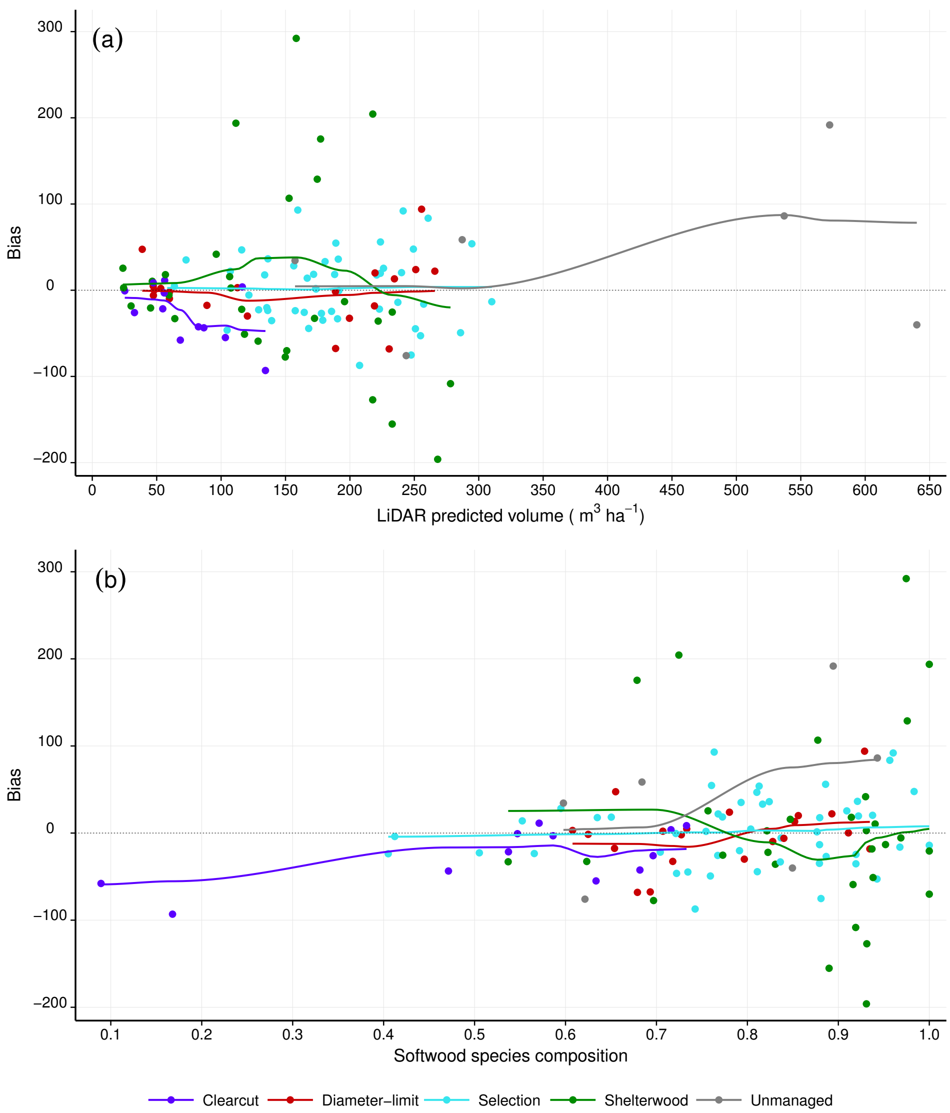

Figure 1.6. Scatterplot of volume prediction bias (observed - predicted; m^3 ha^{-1}) over LiDAR predicted values with lowess regression splines for the different silvicultural treatments **(a)**, and plot species composition based on basal area **(b)**.

An agreement between the LiDAR prediction model based on the 44 operational-grade sampling plots and the matched locations of the 44 research-grade sampling plots, in the nine management units, was relatively high (Table 1.5). The difference in those two R^2 values was about 0.07 with a RMSE difference of 14.81 m^3 ha^{-1}. The operational-grade model had prediction biases between overestimation and underestimation in the diameter-limit and selection units (Figure 1.7a). The research-grade model had prediction biases from underestimation to overestimation in the selection units, and from overestimation to underestimation in the diameter-limit units (Figure 1.7b). In general, the model based on the research-grade plots showed better accuracy and precision in the diameter-limit and selection units (Table 1.6). Also, the research-grade model had smaller mean bias in the mixedwood and softwood plots, although RMSE for mixedwood plots was larger than the operational-grade model.

At last, an agreement between field and model estimates of total volume in the management unit was strong (R^2=0.92). A volume distribution map based on the model with the research-grade plots is presented in Figure 1.8.

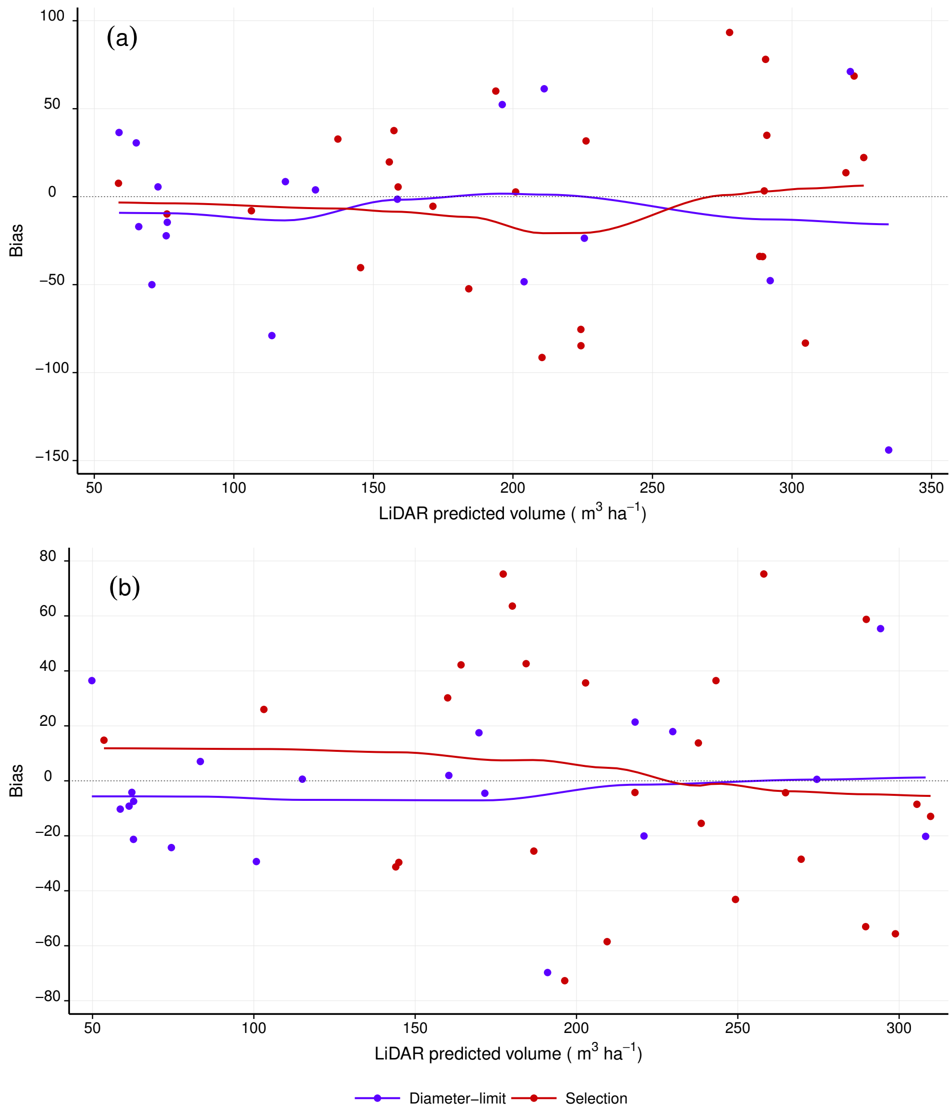

Figure 1.7. Scatterplot of volume prediction bias (observed predicted; m^3 ha^{-1}) over LiDAR predicted values with lowess regression splines for the two silvicultural treatments. The volume prediction model was calibrated based on the 44 operational-grade plot data **(a)**, and the 44 research-grade plot data **(b)**.

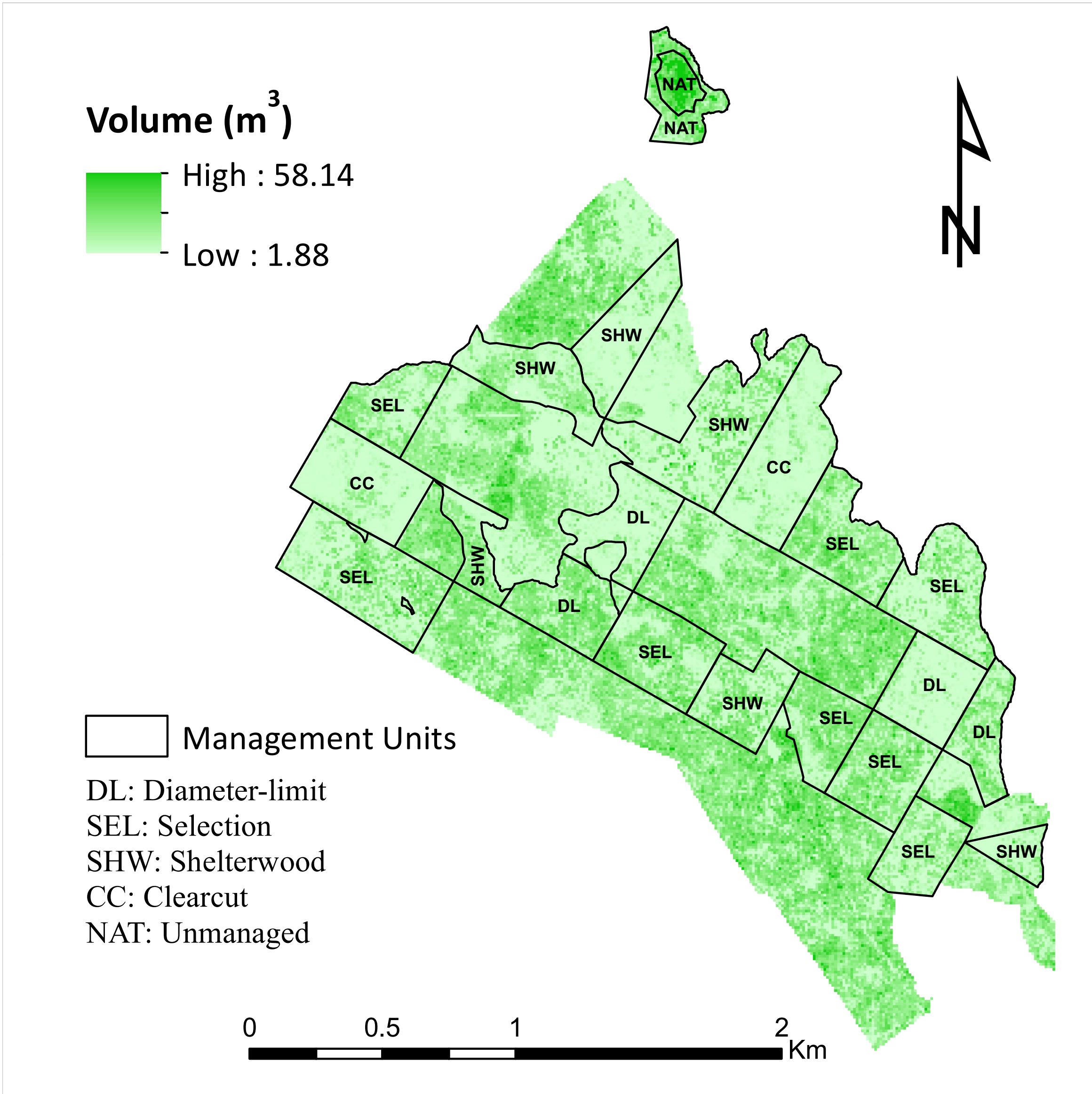

Figure 1.8. Volume (m^3 ha^{-1}) distribution map over the PEF. The prediction model was developed based on 117 research-grade plot data. Each grid represents 900 m^2.

Table 1.3. Developed prediction models with the three most key predictor variables with respect to mean square error in random forest with the coefficient of determination (R^2), mean bias (MB) with standard deviation (SD), and root mean square error (RMSE).

Attributes	Key Variables (mean square error)	R^2 (Adj R^2)	MB (SD)	RMSE
Maximum Tree Height (m)	Maximum height	0.869 (0.867)	1.89 (± 2.06)	2.80
Stem Density (trees ha^{-1})	5^{th} percentile height (3.302) Height kurtosis (5.982) Height L-skewness (6.198)	0.287 (0.280)	9 (± 5013)	4993
QMD (cm)	Percent 1^{st} return above mean (6.591) Percent 1^{st} return above 1 m (7.854) 25^{th} percentile height (8.362)	0.489 (0.434)	-0.05 (± 3.69)	3.68
Basal Area (m^2 ha^{-1})	Percent all returns above 1 m (7.262) Height L-kurtosis (7.564) 99^{th} percentile height (7.614)	0.344 (0.339)	0.03 (± 13.07)	13.01
Volume (m^3 ha^{-1})	90^{th} percentile height (7.795) 20^{th} percentile height (8.724) 75^{th} percentile height (9.757)	0.721 (0.719)	1.81 (± 66.96)	66.70

Table 1.4. Mean bias (MB) with standard deviation (SD), and root mean square error (RMSE) by silvicultural treatments, and species composition. Mixedwood plots had % basal area softwood < 70, and softwood-dominant plots had % basal area softwood ≥ 70.

Silvicultural treatment / Species composition	Plot (n)	Maximum Tree Height MB ± SD [RMSE] (m)	Stem Density MB ± SD [RMSE] (trees ha^{-1})	QMD MB ± SD [RMSE] (cm)	Basal Area MB ± SD [RMSE] (m^2 ha^{-1})	Volume MB ± SD [RMSE] (m^3 ha^{-1})
Diameter-limit	20	2.27±1.19 [2.55]	-1415±2843 [3111]	0.05±1.95 [1.90]	-3.67±7.08 [7.80]	-1.12±36.04 [35.14]
Selection	49	2.73±1.81 [3.26]	2119±5755 [6078]	-0.96±3.21 [3.32]	4.40±15.86 [16.30]	2.70±40.76 [40.43]
Shelterwood	30	0.81±2.15 [2.26]	-2712±4132 [4884]	1.61±4.51 [4.71]	-7.60±8.88 [11.58]	5.56±108.05 [106.37]
Clearcut	30	1.00±1.08 [1.44]	1028±3377 [3392]	-1.80±1.36 [2.22]	5.61±8.24 [9.68]	-26.63±32.46 [40.93]
Unmanaged	6	3.15±3.78 [4.67]	-925±3287 [3140]	2.33±6.48 [6.36]	3.62±8.38 [8.46]	42.47±95.22 [96.75]
Mixedwood	31	1.59±2.02 [2.55]	-589±4496 [4462]	-0.65±2.89 [2.91]	0.72±12.02 [11.85]	-12.34±51.53 [52.17]
Softwood	86	2.15±2.06 [2.97]	224±5196 [5171]	0.17±3.93 [3.91]	-0.22±13.49 [13.41]	6.91±71.30 [71.22]
All Plots	117	2.00±2.05 [2.87]	8.06±5013 [4993]	-0.05±3.69 [3.68]	0.03±13.07 [13.01]	1.81±66.96 [66.70]

Table 1.5. Developed volume prediction models based on research- and operational-grade plot data with the three most key predicator variables regarding the coefficient of determination (R^2), mean bias (MB) with standard deviation (SD), and root mean square error (RMSE).

Sampling Type	Key Variables (mean square error)	R^2 [Adj R^2]	MB [SD] (m^3 ha^{-1})	RMSE (m^3 ha^{-1})
Research-grade	Mean height (6.777) 75^{th} percentile height (6.784) 40^{th} percentile height (6.873)	0.828 [0.824]	0.20 [± 36.74]	36.33
Operational-grade	30^{th} percentile height (6.349) 25^{th} percentile height (6.397) 80^{th} percentile height (7.344)	0.755 [0.749]	-4.21 [± 51.22]	50.81

Table 1.6. Mean bias (MB) with standard deviation (SD), and root mean square error (RMSE) by silvicultural treatments, and species composition. The prediction models were calibrated based on 44 research- and 44 operational-grade plot data. Mixedwood plots had % basal area softwood $<$ 70, and softwood-dominant plots had % basal area softwood $\geq$ 70.

Silvicultural treatment Species composition	Plots (n)	Volume MB ± SD (RMSE) Research-grade (m^2 ha^{-1})	 Operational-grade (m^2 ha^{-1})
Diameter-limit	18	-3.09 ± 26.36 (25.88)	-9.90 ± 53.04 (52.49)
Selection	26	2.73 ± 43.42 (42.67)	-0.28 ± 50.60 (49.62)
Mixedwood	8	-3.96 ± 37.80 (35.58)	13.43 ± 26.10 (27.86)
Softwood	36	1.07 ± 36.97 (36.49)	-8.13 ± 54.77 (54.62)
All Plots	44	0.20 ± 36.74 (36.33)	-4.21 ± 51.22 (50.81)

1.4 Discussion

1.4.1 Predictor variables in LiDAR metrics

Overall, maximum tree height and volume prediction models showed relatively high correlation between field measured values and LiDAR metrics (Table 1.3). Although some previous studies only used the first return and the last return data García et al. (2010); Hawbaker et al. (2010); Kim et al. (2009); Means et al. (2000); Næsset and Økland (2002); Parker and Glass (2004) for inventory attribute predictions, random forest allowed for the use of all return information in this study. To explain the complex vertical structures observed at the PEF, we expected that the first and the last returns information would not be sufficient as the 2^{nd}, 3^{rd} and 4^{th} returns would sense variability under overstory canopy structures. For the volume prediction model, certain percentile heights were necessary to account for multiple canopy layers in plots, and it would be important to acquire not only overstory canopy height distribution, but also lower height (*e.g.* the 20^{th} percentile height) data to distinguish between ground and understory vegetation.

LiDAR intensity values were available in this study, but we did not have an appropriate tool and other auxiliary data to calibrate for flying altitudes, terrain conditions, and atmospheric conditions for the intensity values. While the intensity values have the potential to discriminate between hardwood and softwood species García et al. (2010) or live and dead standing trees Kim et al. (2009), they may not improve accuracy levels for the forest inventory attributes examined in this analysis Goerndt et al. (2010).

1.4.2 Silvicultural treatments and species composition

The unmanaged units tended to results in large prediction errors (Table 1.4). For instance, the unmanaged units had the highest bias in the maximum height and volume predictions. Although total area of unmanaged units is smaller than other four management units, it tends to have the highest variability and the six sampling plots might not have accounted for this variability. Also, management units with softwood species composition greater than 80% tended to result in large prediction errors. For example, the volume prediction tended to be greatly toward underestimation in the softwood species dominant plots. On the other hand, the prediction errors were fairly constant in mixedwood plots, when compared with softwood plots. However, the number of mixedwood plots was small in this study. The PEF is a relatively complex forest and descriptive statistics (*e.g.* mean and standard deviation, Table 1.2) indicated high variability between plots in each of management units. In general, the plots with the highest softwood composition had multiple layer canopy structures, which can be problematic for prediction using LiDAR metrics. In particular, balsam fir is a prolific species, and tends to establish a number of advance seedlings under a range of overstory conditions on the PEF Olson and Wagner (2010). Thus, this creates a rather complex vertical structure and can make it quite difficult to develop forest inventory prediction models based solely on remotely sensed attributes.

1.4.3 Maximum Tree Height

The maximum tree height in plots was generally underestimated, and such result was similar to most other studies Clark et al. (2004); Magnussen and Boudewyn (1998); Magnusson et al. (2007); Næsset (1997). A number of laser pulses likely returned from below treetops Magnussen and Boudewyn (1998), and prediction in the softwood dominant plot had a larger underestimation than the mixedwood plots. The RMSE of 2.75 m between field measured and the LiDAR measured maximum heights in this study was similar to those observed by Means et al. (2000) and Jensen et al. (2006) who also used a low pulse density LiDAR. In contrast, Persson et al. (2002) achieved a RMSE of 0.63 m when a relatively higher pulse density LiDAR was used. In general, higher pulse density LiDAR is necessary to achieve better accuracy levels for maximum height predictions Zimble et al. (2003). In contrast, Magnusson et al. (2007) pointed out that achievable accuracy levels in tree height predictions depends also on canopy structure. For example, uniformly distributed canopy height structural stands may not require the use of high pulse density LiDAR. In addition, the creation of digital terrain models is a difficult task where understory vegetation grows thick Clark et al. (2004) such as in the stands examined in this study. For example, Clark et al. (2004) had a high RMSE for tree height estimations in a tropical rainforest, despite using high pulse density LiDAR.

1.4.4 Stem Density

The stem density had the lowest R^2 and highest prediction bias. Likewise, a majority of other studies also reported that stem density predictions had the lowest accuracy when compared to other forest inventory attributes Hawbaker et al. (2010); Næsset and Økland (2002); Woods et al. (2011). Woods et al. (2011) indicated that a higher pulse density LiDAR in conjunction with sophisticated analytical methods such as individual crown segmentation may be needed to achieve better accuracy for stem density predictions. The prediction accuracy will probably always be rather low in the Acadian Forest due to very high densities of natural regeneration and the tendency for clumped spatial distributions.

1.4.5 Quadratic Mean Diameter

The accuracy of QMD model fit was similar to other previous studies. For example, Jensen et al. (2006) had a R^2 and RMSE of 0.61 and 6.31 cm, while Treitz et al. (2012) achieved a R^2 and RMSE of 0.8 and 0.8 cm, respectively. Similar to this study, both Jensen et al. (2006) and Treitz et al. (2012) used a low pulse density LiDAR, and they found that predicted QMD had lower accuracy levels when compared the maximum tree height, basal area and volume.

1.4.6 Basal Area

The developed plot-level basal area (m^2 ha^{-1}) had a lower R^2 (0.34) when compared to Jaskierniak et al. (2011), Næsset (2002), Jensen et al. (2006) and Gobakken and Næsset (2009), which achieved values of 0.67 to 0.89. Næsset (2002) found that R^2 varied depending on stand age and site quality for his study area in a Norway spruce (*Picea abies* (L.) Karst.) and Scots pine (*Pinus sylvestris* L.) dominated forest, which is a simpler stand structure when compared to those examined in this study. Not only do the stands examined in this analysis have a relatively complex species composition, they also have multiple layer canopy structures due to the range of management histories present. In particular, the shelterwood systems have created very high structural variability between plots in the same management unit.

1.4.7 Stem Volume

The developed plot-level volume (m^3 ha^{-1}) had the highest R^2 value of the various equations evaluated in this study (0.72), which was relatively similar to other studies such as van Aardt et al. (2006) and Hawbaker et al. (2010). Like this analysis, both of these studies were based on low pulse density LiDARs. Magnusson et al. (2007) indicated that relative RMSE in volume predictions increased as pulse density decreased. However, the accuracy of volume prediction models is likely influenced by not only pulse density, but also the stand types examined. For example, Jaskierniak et al. (2011) developed models with R^2 values of 0.59-0.80 based on 2 pulses m^{-2} in an eucalyptus forest in Australia, while Means et al. (2000) developed models with high R^2 values based on a low pulse density in a Douglas-fir (*Pseudotsuga menziesii* (M.) Franco var. *menziesii*) dominated forest in Oregon. In contrast, Magnusson et al. (2007) developed models with a R^2 greater than 0.90 in Norway spruce and Scots pine dominated forests in southern Sweden. When compared to the PEF, the stand structures in these aforementioned studies are relatively simple. Like this study, van Aardt et al. (2006) and Hawbaker et al. (2010) conducted the study in mixed softwood-hardwood forests in Virginia and Wisconsin, respectively, which would have stand structures similar to the PEF. Woods et al. (2011) also worked in a mixed softwood-hardwood forests in Ontario, Canada and were able to achieve a much lower RMSE than our study. Woods et al. (2011) likely did this by stratifying their study area into four stand types based on species composition rather than past silvicultural treatments. Likewise, Anderson and Bolstad (2013) found that stratification of models by forest type was necessary to improve prediction accuracy.

In this study, the volume prediction as well as other inventory attributes was particularly problematic in the shelterwood and unmanaged units. Despite twenty nine and six research-grade 0.08-ha plots being established in these management units, respectively, the high variability between plots suggests that this might be an inadequate sample. Shelterwood systems tend to leave a small number of large trees in the overstory with the intent of promoting a great number of young trees and seedlings in the understory. Likely, a greater number of field plots or larger size plots would be needed to account for this large variability Anderson and Bolstad (2013); Gobakken and Næsset (2009).

In general, the mixedwood plots had smaller prediction biases than the softwood plots for all inventory attributes. However, while Anderson and Bolstad (2013) predicted biomass in a mixed softwood-hardwood forest in Wisconsin, they reported an opposite result that they had less prediction bias in the softwood forests than mixedwood forests. Complexities of stand structures and species composition were somewhat similar to our study site, but the number of mixedwood plots was small in this study; thus further investigation is necessary to resolve such a disagreement.

When comparing the research- and operational-grade plots, overall prediction errors were smaller based on the research-grade sampling plots. Therefore, although such a comparison has not been reported previously to our knowledge, this study suggests that reference data for model calibrations be based on fixed radius plots with a subsample of measured tree heights rather than using variable radius plots with limited or no height measurements.

Although a comparison between the field and LiDAR based total volume prediction (the model calibrated by research-grade plot data) at the management unit-level showed general agreement, both methods were quite different (R^2=0.92). Given the ability to better account for within-stand variability, the LiDAR based volume estimates should be considered superior to the volume estimates based on conventional field measurements.

1.5 Conclusion

Development of inventory attribute prediction model based on a nonparametric regression technique allowed us to explore all potential LiDAR predictor variables and account for highly nonlinear relationships. In general, the low density LiDAR used in this study was able to capture the variability, despite a wide range of stand structure and species composition mixtures examined. However, there were certain stand structures and species composition mixtures where low density LiDAR was ineffective. Although costs of LiDAR data acquisition for large areas are still relatively high, this study highlights that use of LiDAR based inventory attribute predictions are a valuable option for achieving efficient and effective forest assessment from a variety of spatial scales, even in regions dominated by naturally-regenerated, mixed species stands.

CHAPTER 2

PREDICTION OF INDIVIDUAL TREE ATTRIBUTES USING LOW DENSITY LIDAR DATA ACROSS A RANGE OF SILVICULTURAL TREATMENTS IN CENTRAL MAINE, USA

2.1 Introduction

Inventory data are key information for the forest management planning. The ability to accurately project individual tree or stand level growth largely depends on inventory data quality. Based on these projections, foresters plan thinning and other necessary management actions for selected stands. However, it is difficult to conduct conventional field measurements for large or remote areas, and acquiring high resolution inventory data is costly and time limited. In addition, conventional field measurements tend to establish a limited number of sampling plots in each stand, and it is generally assumed that they are representative of the entire stand. In the case of Maine's forests, uniform conditions within a stand may not be met because widely used silvicultural treatments such as a shelterwood system tend to create greatly variable structures and species composition within each stand.

Predictions of forest inventory attributes, especially tree height or stem volume derived from light detection and ranging (LiDAR) metrics, have been demonstrated in various forest types including softwood forests Gobakken and Næsset (2009); Goerndt et al. (2010); Hudak et al. (2008); Means et al. (2000), hardwood forests Hawbaker et al. (2010); Lim and Treitz (2004), and mixedwood forests Anderson and Bolstad (2013); Treitz et al. (2012); Woods et al. (2011). LiDAR-based forest inventory predictions can be accomplished using either area-based methods Anderson and Bolstad (2013); Gobakken and Næsset (2009); Hawbaker et al. (2010); Lim and Treitz (2004); Næsset (2004); Woods et al. (2011) or individual tree-based methods Chen et al. (2006); Falkowski et al. (2008); Jing et al. (2012); Kwak et al. (2010); Lee et al. (2010); Popescu (2007).

In the area-based methods, forest inventory attributes are predicted for a plot-level attribute, such as m^3 ha^{-1} for stem volume, while model calibration data through certain field measurements are necessary. Subsequently, parametric (*e.g.,* stepwise regression) (*e.g.,* Næsset

(2002)) or non-parametric (*e.g.,* random forest) (*e.g.,* Stone et al. (2011); Yu et al. (2010)) statistical techniques are used to develop plot-level prediction models. Area-based methods accurately predicted inventory attributes, such as mean tree height, volume and biomass Clark et al. (2004); Hawbaker et al. (2010); Woods et al. (2011).

In contrast, inventory attribute predictions at high accuracy levels deploying the individual tree-based methods require a different approach. First, individual trees must be accurately discriminated, but previous results have shown that tree discrimination accuracy seemed to depend on a choice of segmentation techniques, and forest structures. Basically, the height of each treetop and its associated crown boundary need to be identified using LiDAR metrics. Employing the crown maxima model with variable window size, Chen et al. (2006) reported that about 64% of individual trees were successfully segmented in an open oak savanna woodland in California. Comparing between the spatial wavelet analysis method and the variable window filters, Falkowski et al. (2008) reported that about 80% and 85%, respectively, of individual trees were successfully segmented in an Idaho mixed softwood forest, and plot crown closure levels influenced the results. The watershed segmentation method with the extended maxima transformation was applied by Kwak et al. (2010). These authors reported that about 50-80% of individual trees were successfully segmented in Korean pine (*Pinus koraiensis* Sieb. et Zucc.) stands, and stem density greatly influenced accuracy of the results. Employing the multi-scale analysis and segmentation method, Jing et al. (2012) reported that approximately 61% of individual trees were successfully segmented in mixedwood forests in Ontario, Canada, and the crown size of each tree influenced accuracy of the results.

Despite the difficulties in identifying individual trees, the approach has some key advantages when compared to the area-based approach. For example, Yu et al. (2010) compared an area-based and an individual tree-based method for mean height, mean diameter and volume predictions in a boreal forest in Finland, and reported that the individual tree-based method resulted in slightly lower prediction errors than the area-based method. Also, silvicultural management plans such as single tree selections or shelterwoods can be accomplished without physically visiting forest stands, if we could accurately segment individual trees. Additionally, species specific volume and biomass equations can be developed if segmentation and individual tree species identification is successful.

Two issues to be considered when using either individual tree-based methods or the area-based methods are the LiDAR pulse density and the complexity of the data analysis methods. Regarding the first issue, Popescu and Wynne (2004), and Falkowski et al. (2006) reported that pulse density greater than 5 pulses m^{-2} was necessary for successfully segmenting individual trees. Low density LiDAR would be more likely to miss the highest treetop position, while multiple pulses may hit on lower crown positions, which could cause a commission error as those pulses might be identified as returned from different trees during the segmentation process. The issue of low pulse density is particularly problematic for softwood species as it can result in greater underestimations in a height prediction and overestimation in a stem density prediction Chen et al. (2006). Additionally, low density LiDAR pulses may not adequately reach and sense individual trees in subcanopy positions. In northern Maine's forests, prolific advanced regeneration positions a number of individual trees in intermediate crown positions that difficult to discriminate.

Regarding the second issue of data analysis, previously developed individual tree-based methods tend to be rather complex. Thus, without knowing advance image analysis theories with equivalent programming skills, novice LiDAR analysts would rarely be able to implement the methods. In addition, field measurements are necessary to support the area-based methods for collecting model calibration data, while the individual tree-based methods need certain calibration data depending on target prediction attributes such as stem volume. For example, conventional volume equations generally require diameter breast height (DBH) and tree height (*e.g.* Li et al. (2012)), while crown width or area measured in the individual tree-based methods would need to be used as a surrogate measure for DBH. In that case, a parametric or nonparametric statistical technique is needed to calibrate volume prediction models.

The goal of this research was to apply an individual tree-based method using relatively low density LiDAR data with model calibration data to predict individual tree attributes. Specific objectives were to: (1) classify tree species type, and softwood species; (2) predict individual tree height and volume; and (3) assess the influence of species, canopy position, and silvicultural regime on model accuracy. In particular, this research employed a relatively simple non-parametric approach, the random forest technique Breiman (2001), to predict key individual tree attributes in a managed mixed-species and multi-age forest.

2.2 Methods

2.2.1 Study Area

The study was conducted on the Penobscot Experimental Forest (PEF) near Orono, Maine, USA (N44°49’30”, W68°39’00”) (Figure 2.1). The PEF was established in 1952 by U.S. Forest Service, and a number of studies regarding timber management, stand dynamics, productivity, biological diversity and more have been conducted Sendak et al. (2003). The total area of the PEF is 1,619 ha, and various silvicultural treatments (*e.g.* natural area, clearcut, shelterwood, diameter-limit cutting) have been twice replicated for long term observations. The treatments generally range in size from 0.5 to 22.4 ha and are representative of typical northern Maine’s silvicultural practices (Table 2.1). With a few exceptions, most treatments are replicated in the PEF, and field data (*e.g.* DBH) for each of replicated treatments are collected at about 600 permanent sampling plots on a 10-year cycle.

The PEF is defined as a mixed northern conifer dominant forest as a part of Acadian ecosystem Sendak et al. (2003), with major hardwood species including red maple (*Acer rubrum* L.), birches (*Betula* spp.) and aspens (*Populus* spp.), while the major softwood species are spruces (*Picea rubens* Sarg., *Picea glauca* (Moench) Voss and *Picea mariana* (Mill.) BSP), balsam fir (*Abies balsamea* L. (Mill.)), northern white cedar (*Thuja occidentalis* L.) and eastern white pine (*Pinus strobus* L.). The range of elevation above sea level is 20 - 70 m.

2.2.2 Inventory Attributes Data

For this study, six replicated silvicultural treatment units that varied from 7.25 to 17.55 ha in size were selected (Figure 2.1). Within these six management units, a total of 57 0.08-ha (1/5th-acre) plots, with a range of 4-7 plots in each management unit, were selected for this study.

Table 2.1. Description of silvicultural treatments in management units (MUs) in the study area.

MU	Area (ha)	Treat-ment year	Inven-tory year	Plot (n)	Description of Silvicultural Treatment	Treatment Group
4	10.1	1994	2009	4	Fixed diameter-limit cutting. Thresholds are 14.0 cm for balsam fir, 24.1 cm for spruce and hemlock, 26.7 cm for white pine, 19.1 cm for cedar and paper birch, and 14.0 cm for other hardwoods.	Diameter-limit
15	10.3	2001	2007	6		
24	9.4	1996	2005	4	Modified diameter-limit cutting. The third modified diameter-limit cut was applied in 1995. Portions of the stand are in the stem exclusion and understory reinitiation stages of development.	
28	7.3	1997	2007	5		
9	12.2	2003	2003	4	5-year cutting cycle. Structural goal is to retain 24.1 m^2 ha^{-1} (trees > 11.4 cm).	Selection
16	8.6	2006	2011	6		
12	12.5	1994	2004	5	10-year cutting cycle. Structural goal is to retain 20.7 m^2 ha^{-1} (trees > 11.4 cm).	
17	10.9	1994	2005	5	20-year cutting cycle. Structural goal is to retain 16.1 m^2 ha^{-1} (trees > 11.4 cm).	
27	8.4	1997	2007	7		
8	17.6	1983	2008	7	Unregulated harvest /commercial clearcutting. This compartment was initially cut with unregulated ("loggers choice") harvests. The second harvest was a commercial clearcut in 1982. The stands are in the stand initiation and stem exclusion phases of development.	Clearcut
22	13.6	1988	2004	6		

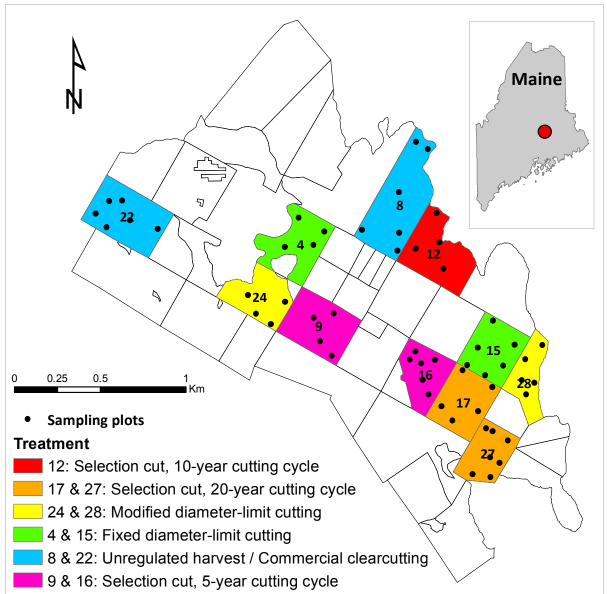

Figure 2.1. The Penobscot Experimental Forest near Orono, Maine, USA (N44°49'30", W68°39'00"). At each plot, 25% of all trees with a DBH greater than 11.25 cm are spatially mapped within a circular 0.08-ha area.

On each 0.08-ha plot, DBH was measured on all trees greater than 11.25 cm (4.5 inches). Total height and crown position (dominant, codominant or intermediate) relative to neighbors' crown positions of twenty five percent of those DBH trees were measured between 2003 and 2011 and were spatially mapped based on azimuth and distance from each plot center. Each plot center coordinate was recorded by a Trimble GeoXH 6000 GPS unit with a Zephyr2 external antenna. At each plot, we collected a minimum of 100 waypoints, and the post correction was carried out using base stations around the PEF. The coordinates were set to UTM NAD83 Zone 19.

Given the differences in dates between tree measurement and acquisition of the LiDAR data in the summer of 2012, the Acadian Variant of the Forest Vegetation Simulator was applied Weiskittel et al. (2012) to project DBH and height to a common year with the number of projections ranging from 1 to 9 annual cycles. Based on simulated DBH and height, total tree volume was estimated using a species-specific taper equation Li et al. (2012); Weiskittel and Li (2012). Total height and stem volume of 1,694 tree-level data in the eleven management units were available for analysis (Table 2.2).

Table 2.2. Examined attributes (mean ± standard deviation) by management unit (MU).

MU	Number of sampled trees (n)	DBH (cm)		Height (m)		Volume (m^3)	
		Mean (SD)	Min Max	Mean (SD)	Min Max	Mean (SD)	Min Max
4	122	18.7 (6.0)	12.3 36.1	11.60 (2.44)	5.37 18.60	0.173 (0.140)	0.040 0.786
15	116	18.5 (4.5)	12.8 32.4	11.13 (2.69)	5.23 1.91	0.156 (0.103)	0.048 0.527
8	173	16.4 (3.3)	11.3 26.9	11.23 (2.15)	2.99 15.98	0.117 (0.061)	0.032 0.380
22	62	19.5 (4.3)	14.5 32.6	10.22 (2.33)	4.08 15.57	0.157 (0.097)	0.063 0.513
9	135	25.0 (10.4)	13.1 72.0	15.67 (4.35)	6.67 30.44	0.487 (0.647)	0.056 5.209
16	241	24.0 (10.3)	11.5 53.7	15.89 (3.84)	6.15 24.6	0.430 (0.450)	0.040 1.853
12	148	23.5 (8.9)	12.0 51.4	15.12 (3.88)	5.01 23.10	0.382 (0.370)	0.044 1.853
17	127	27.1 (9.5)	12.8 61.3	15.97 (4.32)	5.03 27.70	0.530 (0.487)	0.065 3.519
27	176	22.3 (9.2)	12.5 52.3	13.74 (4.14)	3.80 26.51	0.340 (0.382)	0.049 1.965
24	154	22.4 (6.8)	12.9 38.5	15.28 (3.89)	5.15 24.15	0.325 (0.252)	0.043 1.056
28	233	22.1 (8.0)	12.2 44.9	14.50 (3.86)	2.84 23.68	0.324 (0.304)	0.023 1.563
Overall	153	21.8 (7.4)	12.5 45.6	13.67 (3.44)	4.76 21.12	0.311 (0.299)	0.046 1.771

2.2.3 LiDAR System Specifications

In late June 2012, airborne discrete-return LiDAR data were acquired using a VQ-480 (Riegl USA, Orlando, Florida), a component of NASA Goddard's LiDAR, Hyperspectral and Thermal (G-LiHT) airborne imager system Cook et al. (2013). The sensor generated the pulse repetition frequency of 150 kHz, and the laser pulse intensity was 1550 nm with the scan angle of $< 15°$ from the nadir. Mean laser point density was about 3.0 pulses m^{-2} with a footprint of 30 cm, and the sensor collected up to 4 pulse returns.

2.2.4 LiDAR Data Processing and Model Calibration Predictions

All LiDAR data processing including the creation of a digital terrain model and LiDAR metrics were accomplished in FUSION v3.30 McGaughey (2013). FUSION generated a number of potential predictor variables for individual tree height and volume predictions from raw LiDAR data. Ninety eight potential predictor variables were generated for this study. FUSION extracted predictor variables from the raw LiDAR data at each of mapped stem locations in the field. Although horizontal accuracy in geodetic information in LiDAR data is generally controlled within sub-meter accuracy Evans et al. (2009), it is still difficult to assess horizontal accuracy. To account for certain horizontal error and different crown shapes and sizes among individual trees, FUSION metrics were extracted for a 4 m radius circular area around the mapped individual trees locations. Preliminary results indicated that better model fit (greater R^2 values) was achieved based on the LiDAR metrics data extracted from 4 m radius circular plots to compare with other circular plot sizes.

Although understory vegetation heights varied largely depending on silvicultural treatments in each management unit, we disregarded pulse return within 1 m above ground, because preliminary results indicated better model fit during the LiDAR data extraction. A few examples of predictor variables in the LiDAR metrics were maximum height, the number of 1^{st} return pulses in the 90^{th} percentile height, and standard deviation of the number of 1^{st} return pulses.

Predictor variables in the LiDAR metrics tend to be highly correlated with others, and some variables would not meet normal distribution criteria Hudak et al. (2008); Li et al. (2008);

Stone et al. (2011). These issues violate the assumption inherent to linear regression models. In addition, variable selection in high dimensionality metrics is not a simple process and typical data transformations might not be effective for highly skewed or bimodal data. Although Akaike's Information Criteria (AIC) is a popular approach for variable selections in stepwise regression, the developed regression models tend to have model overfit issues, which are generally not stable when used outside of the calibration data. Therefore, the development of inventory prediction models based on simple and multiple linear regressions would not be suitable for this type of dataset.

Alternatively, the random forest technique proposed by Breiman (2001), a nonparametric approach, may be a more effective technique. Random Forest is developed based on the regression trees algorithm, where predictor variables are split to grow a number of nodes to select the best predictor variables. In the random forest approach, the regression tree process is continued multiple times and compared against a bootstrapped validation dataset. A key advantage in random forest is that a greater number of predictor variables of various types (categorical, continuous, and binary) can be handled, and the relative importance of each predictor variable can be estimated during the model calibration process. In this analysis, the random forest algorithm was run iteratively in that the model initially included all covariates, the least influential covariate dropped, and the model reran until there were only 5 covariates, which preliminary analysis had suggested was most effective for prediction accuracy.

To deploy supervised classifications based on the random forest technique in our study, we assumed that overall crown shapes and branch patterns between hardwood and softwood species are different. Among softwood species (spruces (black, white, red), balsam fir, and other softwood) in the PEF, those tree elements are different enough that certain variables in the LiDAR metrics could correlate to shapes of softwood species. Classified species type data and classified softwood species data were used as a covariate for height and volume predictions. Consequently, three different sets of calibration data were used to predict individual tree height and volume: (1) LiDAR metrics only; (2) LiDAR metrics with classified species type (hardwood and softwood); and (3) LiDAR metrics with classified softwood species (spruces (black, white, red), balsam fir and other softwood).

The ‘randomforest’ package Liaw and Wiener (2002) in R v3.0.1 R Development Core Team (2012), was used to calibrate the height and volume prediction models, and to classify species type and softwood species. Each of calibrated models was evaluated using the coefficient of determination (R^2), mean bias and absolute root mean square error (RMSE) between field measured and model predicted individual tree height and volume. To examine the performance of the various models, the bias (observed - predicted) was examined graphically with the use of lowess regression splines. To simplify interpretation of the differences between the original six different silvicultural treatments, the treatments were narrowed down to three broad categories, which included clearcut, diameter-limit and selection (Table 2.1). Classification accuracy for classified species type and softwood species was reported in confusion matrices Congalton and Green (1993).

2.3 Results

A total of 1,694 trees were available for analysis with 82% being softwoods (Table 2.2). Overall, the developed models had a weak agreement with field measured values as the RMSE was relatively large (Table 2.3)

2.3.1 Species Type Classification and Softwood Species Classification

A pulse count-related variable was selected as the most important classification variable for predicting whether trees were hardwoods or softwoods (Table 2.4a). While overall accuracy was about 85%, the Kappa statistic Rosenfield and Fitzpatrick-Lins (1986) was near 0%. This small Kappa statistic indicated that classification results were purely by chance. In particular, the random forest technique did not accurately classify hardwoods when compared with softwood as producer’s and user’s accuracy in hardwood were greatly different (Table 2.4b). Producers accuracy is reflected to omission errors, and users accuracy is reflected to commission errors.

For the 1,394 softwood trees, random forest was again used to classify them into spruces (black, red and white), balsam fir, or other softwood. A height-related variable was selected as the most important classification variable (Table 2.5a). While overall accuracy was about 52%,

Table 2.3. Developed prediction models with the three most important predictor variables. The classified species type and the classified softwood species were derived through supervised classification using the random forest technique based on the LiDAR metrics.

Attribute	Covariates	Key Variables (mean square error)	R^2 (Adj R^2)	MB (SD)	RMSE
Height (m)	LiDAR metrics	Percent 1st returns above mean (13.0) Percent all returns above mean (15.3) 20th percentile height (16.0) 75th percentile height (21.2) 95th percentile height (22.9)	0.269 (0.269)	0.011 (3.47)	3.47
Height (m)	LiDAR metrics + Classified spp type	10th percentile height (10.7) 75th percentile height (11.7) 10th percent height (11.8) 95th percentile height (14.9) Classified species type (17.3)	0.292 (0.291)	0.007 (3.41)	3.41
Height (m)	LiDAR metrics + Classified sw spp	All returns above 1 m (9.0) Mean height (11.3) 90th percentile height (12.9) 95th percentile height (13.4) Classified softwood species (29.7)	0.378 (0.377)	0.018 (3.26)	3.26
Volume (m^3)	LiDAR metrics	30th percentile height (10.8) 70th percent height (10.9) 99th percentile height (10.9) 90th percentile height (11.0) 95th percentile height (12.1)	0.166 (0.166)	0.000 (0.37)	0.37
Volume (m^3)	LiDAR metrics + Classified spp type	80th percentile height intensity (10.7) Elevation variance (10.9) 30th percentile height (12.5) Height standard deviation (12.7) 80th percent height (12.9)	0.165 (0.165)	0.002 (0.37)	0.37
Volume (m^3)	LiDAR metrics + Classified sw spp	Percent 1st returns above mean (6.5) 30th percentile height (9.2) Height standard deviation (11.5) 70th percentile height (11.8) Classified softwood species (31.0)	0.296 (0.295)	0.000 (3.26)	0.36

Kappa statistic was 0%. In particular, the random forest technique did not classify spruces as both producer's and user's accuracy were below 50% (Table 2.5b).

2.3.2 Individual Tree Height Prediction

Using only LiDAR metrics, the model slightly underestimated tree height by 0.01 $\pm$ 3.47 m regardless of silvicultural treatments and species type, while an agreement between field measured and model predicted heights was weak (Table 2.3). Two return count-related variables were selected as the most important predictor variables followed by three height-related variables (Table 2.3). In general, tree heights in the clearcut and diameter-limit units were slightly overestimated, while trees in the selection units were underestimated (Table 2.6 and Figure 2.2a. This model underestimated hardwood to a greater extent than softwood heights (Table 2.6 and Figure 2.2b). Tree heights in the dominant crown position were also underestimated, and overestimated in the codominant and intermediate crown positions (Figure 2.2c). In particular, trees in the intermediate crown position were increasingly overestimated with greater predicted heights.

Based on the LiDAR metrics with the classified species type, the model performed slightly better as it generally underestimated tree height by 0.01 $\pm$ 3.41 m regardless of silvicultural treatment and species type (Table 2.3). The classified species type was selected as the fifth key predictor variable and the other four variables were height-related variables. With respect to the silvicultural treatments, species type and crown position, the general tendencies of accuracy and precision were similar to the result of the model based on the LiDAR metrics only (Table 2.7 and Figures 2.3a,b,c).

Using the LiDAR metrics with the classified softwood species as a covariate produced a model that underestimated the softwood tree height by 0.02 $\pm$ 3.26 m (Table 2.3). After stratification to only softwood species, an agreement between field measured and model predicted heights was improved when compared to the previous models. The classified softwood species type was selected as the fifth key predictor variable, and the other four variables were return count- and height-related variables. As similar to the previous two height prediction models, this model slightly overestimated softwood tree heights in the clearcut and diameter-limit units,

Table 2.4. The results of species type classification (hardwood or softwood) through supervised classification using the random forest technique based on the LiDAR metrics. **(a)** The five most key variables for the species type classification; and **(b)** accuracy assessment in the species type classification.

(a)

Key Variables	Mean Decrease Accuracy	Mean Decrease Gini
Total pulse return counts	0.0201	95.77
Maximum height	0.0331	95.72
Intensity standard deviation	0.0358	107.35
95^{th} percentile height intensity	0.0298	97.50
Percent all returns above mean	0.0145	88.59

(b)

		Observed				
		Hardwood	Softwood	Total	User's accuracy	Commission error
Predicted	Hardwood	101	199	300	0.34	0.66
	Softwood	51	1343	1394	0.96	0.04
	Total	152	1542	1694		
	Producer's accuracy	0.66	0.87			
	Omission error	0.34	0.13			
			Overall accuracy 0.85		Kappa statistics 0.00	

Table 2.5. The results of the softwood classification (spruces, balsam fir and other softwood) through supervised classification using the random forest technique based on the LiDAR metrics. **(a)** The five most key variables for the softwood species type classification; and **(b)** accuracy assessment in the softwood species classification.

(a)

Key Variables	Mean Decrease Accuracy	Mean Decrease Gini
Maximum height	0.0356	164.57
Variance in height	0.0256	175.50
Intensity standard deviation	0.0144	163.97
75^{th} percentile height intensity	0.0649	200.00
Percent all returns above mean	0.0313	178.35

(b)

		Observed					
		Spruces	Balsam fir	Other sw	Total	User's accuracy	Commission error
Predicted	Spruces	76	83	134	293	0.26	0.74
	Balsam fir	47	284	162	493	0.58	0.42
	Other sw	51	138	419	608	0.69	0.31
	Total	174	505	715	1394		
	Producer's accuracy	0.44	0.56	0.59			
	Omission error	0.56	0.44	0.41			
				Overall accuracy 0.56		Kappa statistics 0.00	

Table 2.6. Mean (MB) with standard deviation (SD), and root mean square error (RMSE) by silvicultural treatments, species type (hardwood or softwood), and crown positions. The model was developed based on only LiDAR metrics.

Silvicultural treatment Species type Crown position	Trees (n)	Height MB ± SD (RMSE) (m)	Volume MB ± SD (RMSE) (m^3)
Clearcut	239	-0.35 ± 1.89 (1.92)	-0.0352 ± 0.0887 (0.0952)
Diameter-limit	642	-0.22 ± 3.32 (3.32)	-0.0350 ± 0.2712 (0.2732)
Selection	813	0.30 ± 3.90 (3.91)	0.0375 ± 0.4770 (0.4781)
Hardwood	300	1.12 ± 2.70 (2.92)	-0.0639 ± 0.2050 (0.2144)
Softwood	1394	-0.23 ± 3.57 (3.58)	0.0134 ± 0.3991 (0.3992)
Dominant	751	2.07 ± 2.93 (3.58)	0.2060 ± 0.4238 (0.4710)
Codominant	526	-0.50 ± 2.59 (2.64)	-0.0933 ± 0.2052 (0.2252)
Intermediate	417	-3.05 ± 2.78 (4.13)	-0.2544 ± 0.1942 (0.3199)
All Trees	1694	0.01 ± 3.47 (3.47)	-0.0003 ± 0.3733 (0.3732)

Table 2.7. Mean (MB) with standard deviation (SD), and root mean square error (RMSE) by silvicultural treatments, species type (hardwood or softwood), and crown positions. The model was developed based on LiDAR metrics with classified species type.

Silvicultural treatment Species type Crown position	Trees (n)	Height MB ± SD (RMSE) (m)	Volume MB ± SD (RMSE) (m^3)
Clearcut	239	-0.56 ± 1.85 (1.92)	-0.0394 ± 0.0904 (0.0985)
Diameter-limit	642	-0.29 ± 3.23 (3.24)	-0.0334 ± 0.2732 (0.2750)
Selection	813	0.41 ± 3.83 (3.85)	0.0420 ± 0.4751 (0.4766)
Hardwood	300	0.18 ± 2.68 (2.69)	-0.0622 ± 0.2072 (0.2160)
Softwood	1394	-0.03 ± 3.54 (3.54)	0.0157 ± 0.3985 (0.3987)
Dominant	751	1.98 ± 3.00 (3.60)	0.2041 ± 0.4295 (0.4753)
Codominant	526	-0.46 ± 2.51 (2.55)	-0.0864± 0.1920 (0.2104)
Intermediate	417	-2.96 ± 2.26 (3.95)	-0.2507 ± 0.1988 (0.3200)
All Trees	1694	0.001 ± 3.42 (3.41)	-0.0019 ± 0.3730 (0.3729)

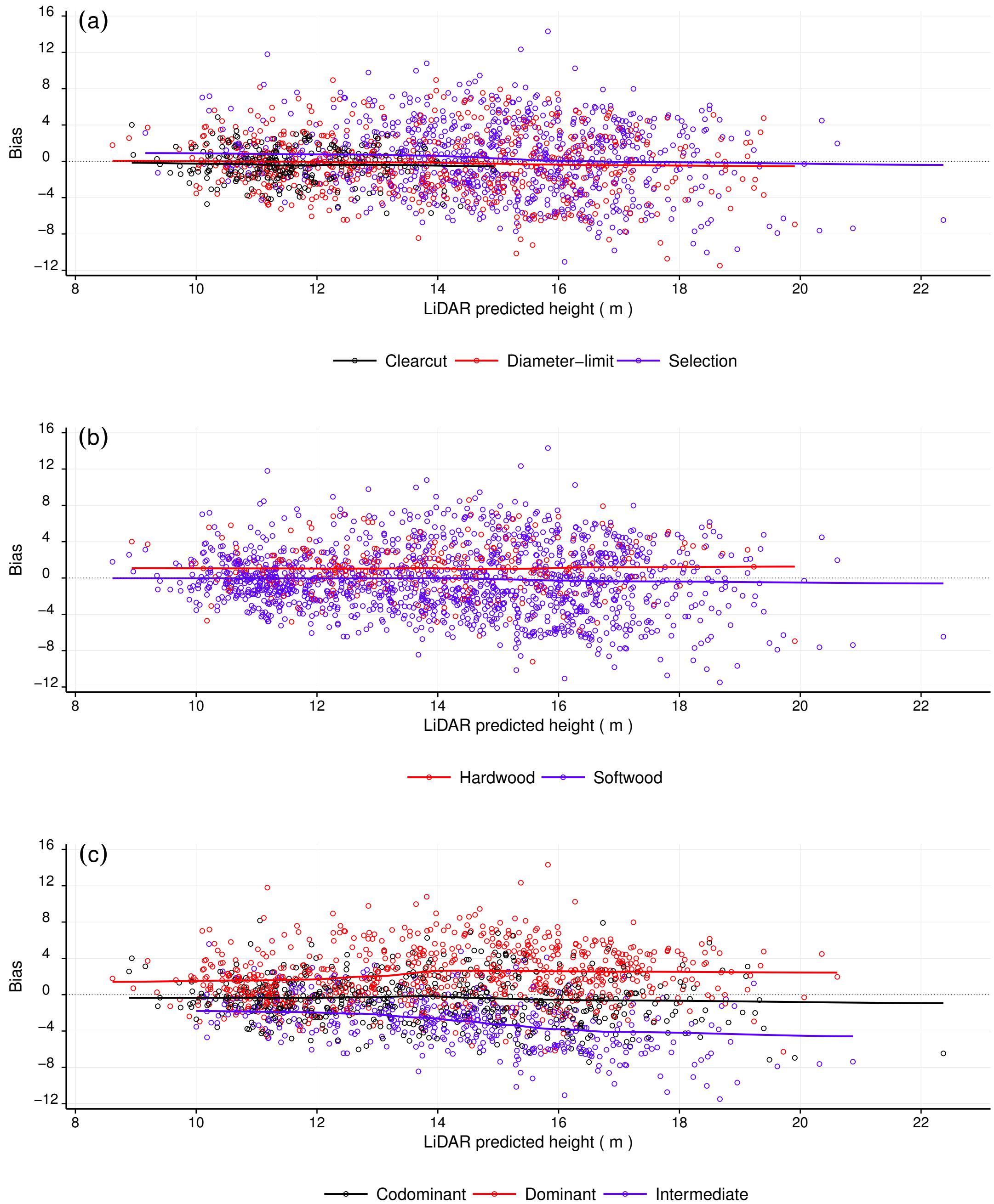

Figure 2.2. Individual tree height prediction model was developed based on LiDAR metrics. Scatterplot of tree height prediction bias (observed - predicted; m) over LiDAR predicted values with lowess regression splines for the different silvicultural treatments **(a)**, species type **(b)**, and crown positions **(c)**.

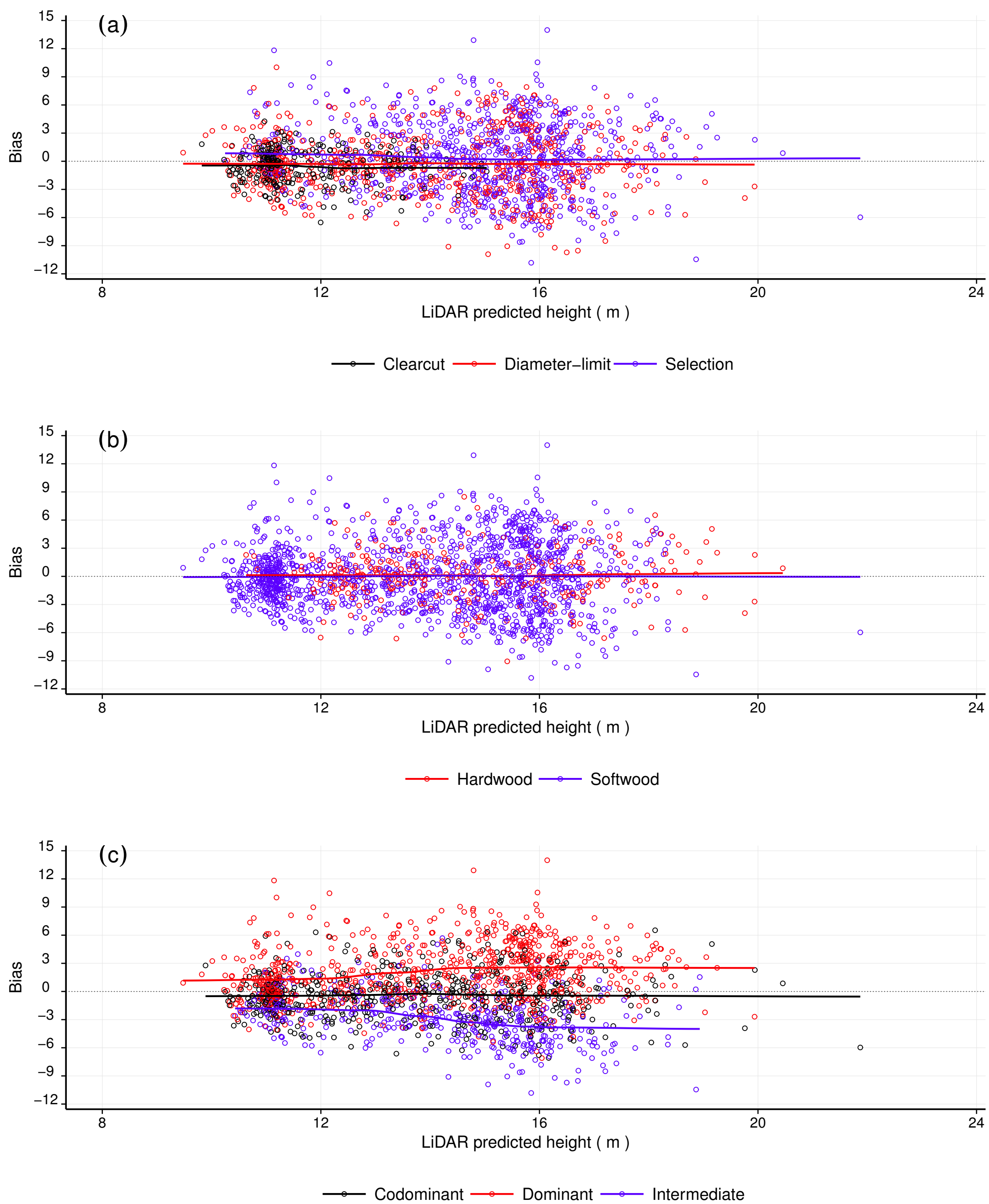

Figure 2.3. Individual tree height prediction model was developed based on LiDAR metrics. Scatterplot of tree height prediction bias (observed - predicted; m) over LiDAR predicted values with lowess regression splines for the different silvicultural treatments **(a)**, species type **(b)**, and crown position **(c)**.

Table 2.8. Mean (MB) with standard deviation (SD), and root mean square error (RMSE) by silvicultural treatments, species type (hardwood or softwood), and crown positions. The model was developed based on LiDAR metrics with classified softwood species.

Silvicultural treatment Softwood species Crown position	Trees (n)	Height MB ± SD (RMSE) (m)	Volume MB ± SD (RMSE) (m^3)
Clearcut	183	-0.42 ± 1.93 (1.97)	-0.0371 ± 0.0962 (0.1029)
Diameter-limit	516	-0.33 ± 3.12 (3.14)	-0.0517 ± 0.2424 (0.2477)
Selection	695	0.40 ± 3.58 (3.60)	0.0484 ± 0.4606 (0.4628)
Spruces	293	0.52 ± 3.53 (3.57)	-0.0344 ± 0.3960 (0.3968)
Balsam fir	493	-0.20 ± 2.15 (2.15)	-0.0269 ± 0.0852 (0.0892)
Other softwood	608	-0.05 ± 3.81 (3.80)	0.0344 ± 0.3960 (0.3968)
Dominant	588	1.95 ± 2.81 (3.42)	0.1989 ± 0.4286 (0.4722)
Codominant	431	-0.28 ± 2.45 (2.46)	-0.0810 ± 0.1899 (0.2062)
Intermediate	375	-2.67 ± 2.67 (3.77)	-0.2183 ± 0.2014 (0.2968)
All Trees	1394	0.02 ± 3.26 (3.26)	-0.0001 ± 0.3619 (0.3618)

while underestimating in the selection unit (Table 2.8 and Figure 2.4a). The model underestimated the heights in the dominant crown position, while overestimating in the codominant and intermediate crown positions (Figure 2.4c). Also, in general, this model slightly underestimated spruces' heights (Table 2.8).

2.3.3 Stem Volume Prediction

Similar to the tree height model, an individual tree volume prediction model was developed based on LiDAR metrics only. All selected predictor variables were height-related variables (Table 2.3). The model had noticeable bias for overestimating volume by $< 0.01 \pm 0.37$ m^3 (Table 2.3 and Figures 2.5a,b,c) . Unlike the tree height prediction model based on LiDAR metrics only, this model underestimated softwood, and overestimated hardwood volumes (Table 2.6). However, like the height prediction model, this model underestimated tree volumes in the dominant crown position, while overestimating in the codominant and intermediate crown positions (Figure 2.5c).

An individual volume prediction model was developed based on the LiDAR metrics with the classified species type (hardwood or softwood), but the classified species type was not

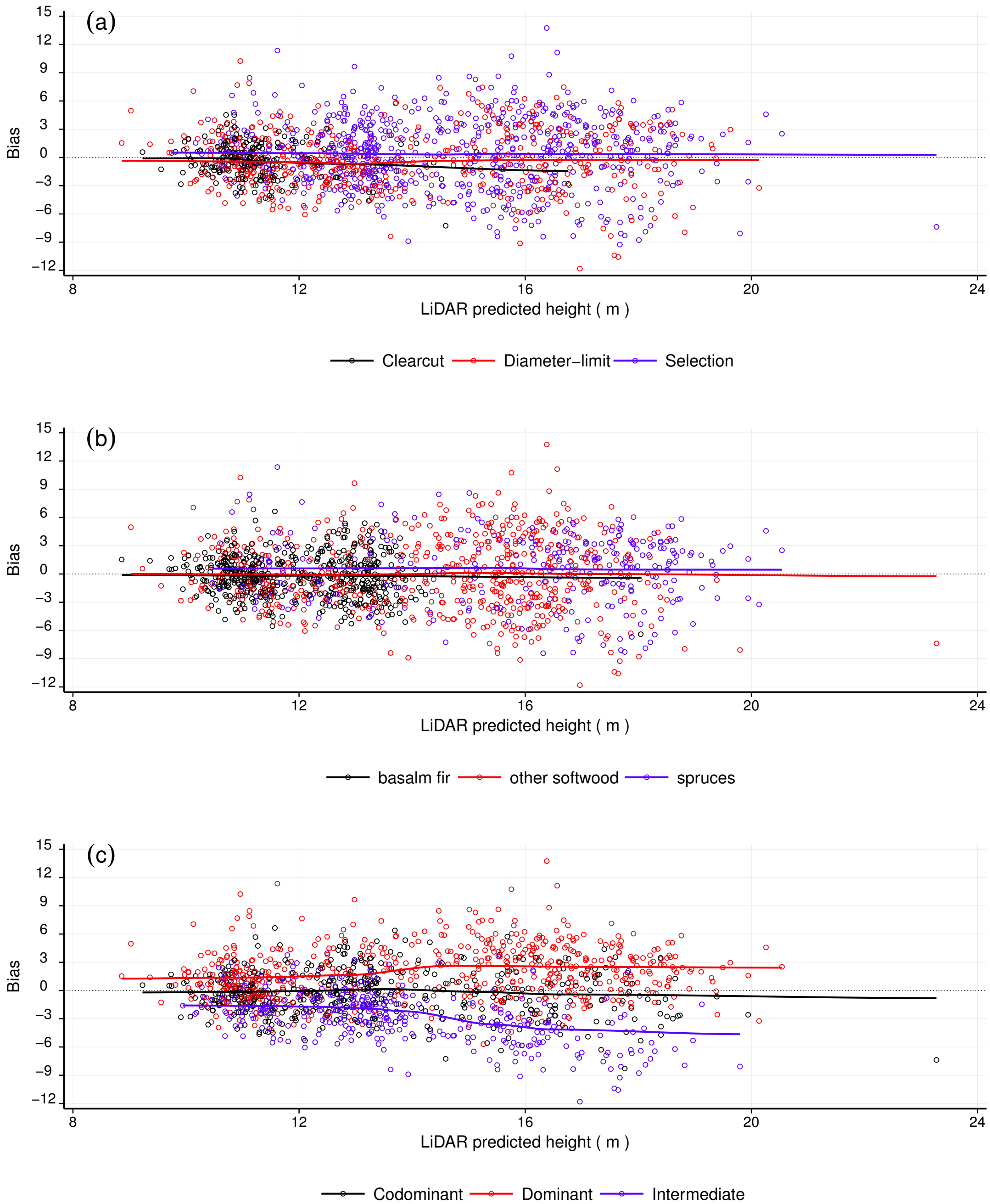

Figure 2.4. Individual tree height prediction model for softwood species was developed based on LiDAR metrics with classified softwood species. Scatterplot of tree height prediction bias (observed - predicted; m) over LiDAR predicted values with lowess regression splines for the different silvicultural treatments **(a)**, softwood species **(b)**, and crown positions **(c)**.

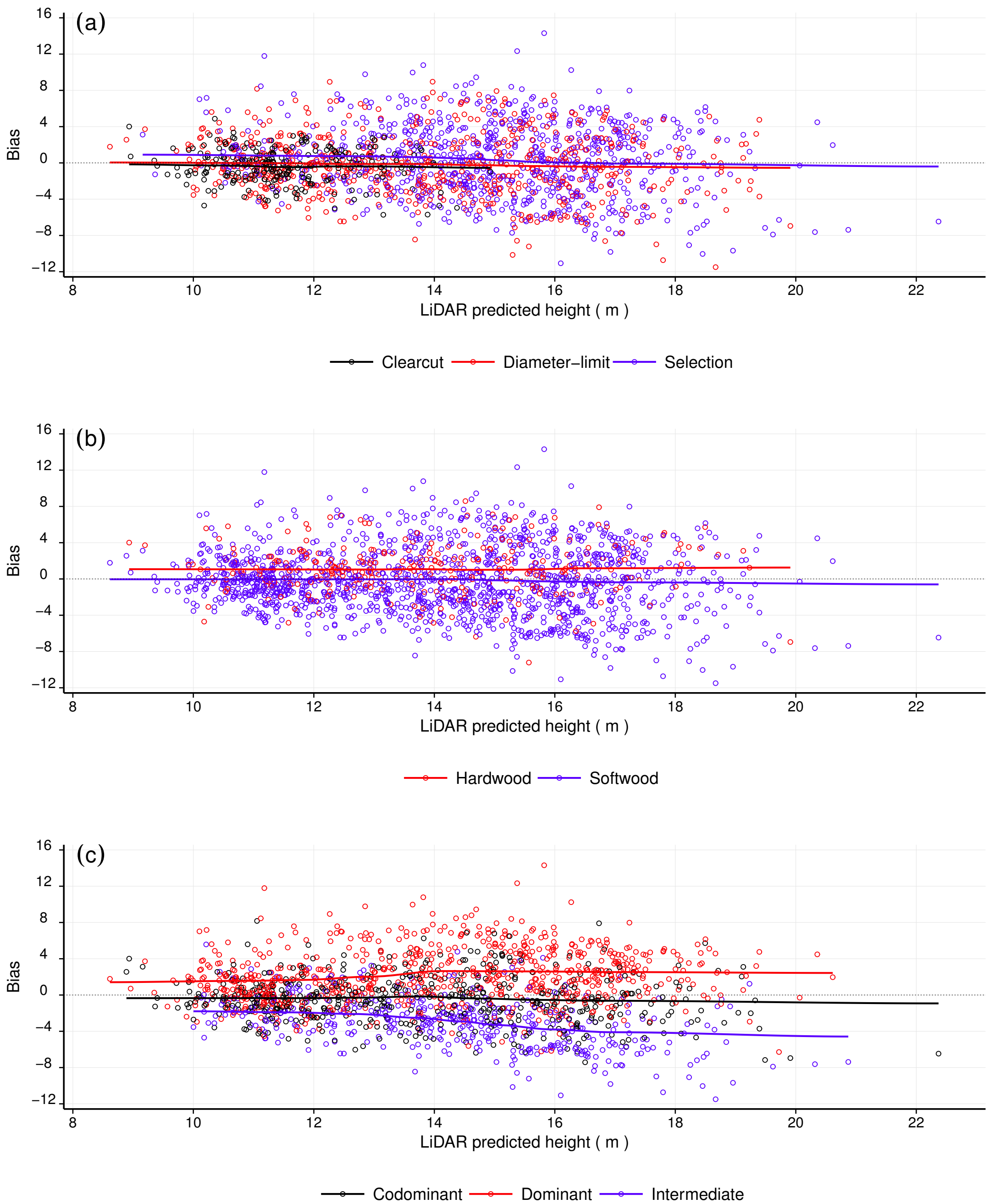

Figure 2.5. Individual tree volume prediction model was developed based on LiDAR metrics. Scatterplot of stem volume prediction bias (observed - predicted; m^3) over LiDAR predicted values with lowess regression splines for the different silvicultural treatments **(a)**, species type **(b)**, and crown positions **(c)**.

selected as an important predictor variable, and there was no appreciably improvement in model predictions (Table 2.7 and Figures 2.6a,b,c).

An individual tree volume prediction model for softwood species (spruces, balsam fir or other softwood) was developed based on the LiDAR metrics with the classified softwood species. This model slightly overestimated the volume by $< 0.01 \pm 0.36$ m^3 (Table 2.3). After stratification to focus on softwood species, an agreement between field measured and model predicted volumes was improved when compared to the previous models. The classified softwood species type was selected as the fifth important predictor variable and the other four variables were pulse return count- and height-related variables. As similar to the previous two volume prediction models, this model slightly overestimated softwood volumes in the clearcut and diameter-limit units, while underestimating in the selection unit (Table 2.8 and Figure 2.7a). The softwood volumes in the dominant crown position were underestimated, while overestimated in the codominant and intermediate crown positions (Figure 2.7c). Also, this model slightly overestimated both spruce and balsam fir volumes (Table 2.8).

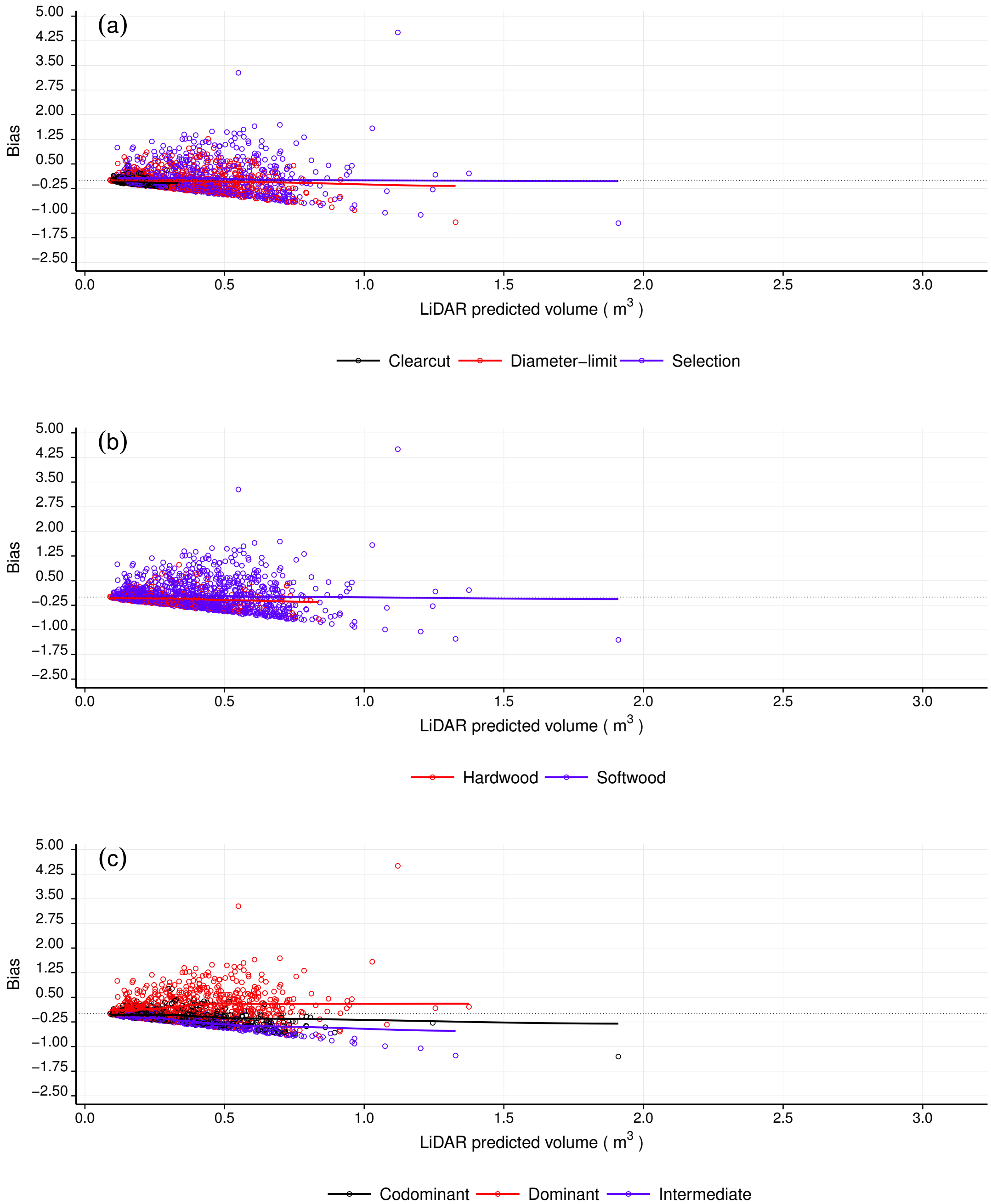

Figure 2.6. Individual tree volume prediction model was developed based on LiDAR metrics with classified species type. Scatterplot of tree volume prediction bias (observed - predicted; m^3) over LiDAR predicted values with lowess regression splines for the different silvicultural treatments **(a)**, species type **(b)**, and crown positions **(c)**.

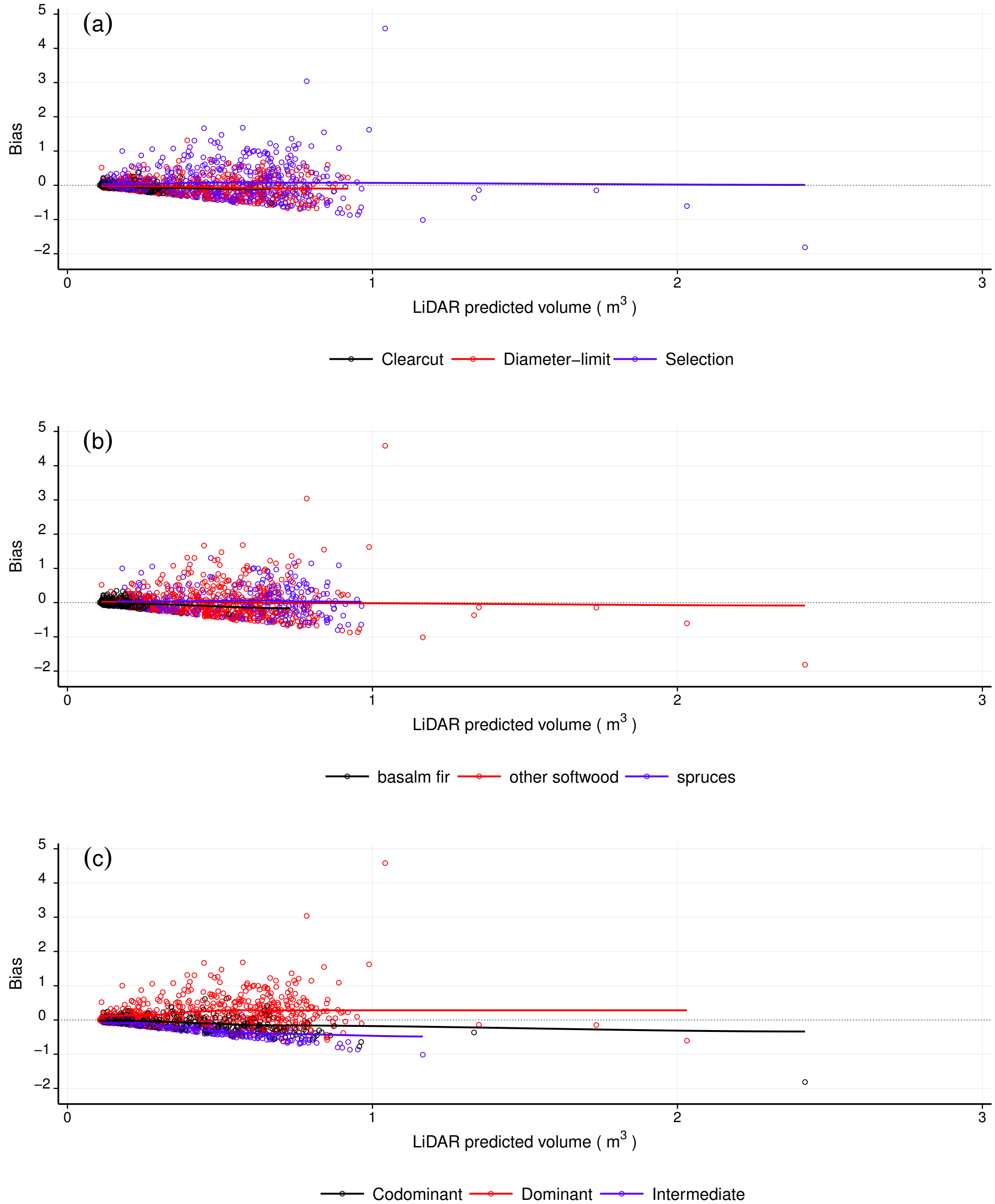

Figure 2.7. Individual tree volume prediction model for softwood species was developed based on LiDAR metrics with classified softwood species. Scatterplot of tree volume prediction bias (observed - predicted; m^3) over LiDAR predicted values with lowess regression splines for the different silvicultural treatments **(a)**, softwood species **(b)**, and crown position **(c)**.

2.4 Discussion

Individual tree height and volume prediction models showed weak correlations between field measured and model predicted values. Although mean bias in each model was relatively small, the RMSE was large (Table 2.3). While some previous area-based method studies only used the first and last LiDAR return data García et al. (2010); Hawbaker et al. (2010); Kim et al. (2009); Magnusson et al. (2007), this study used all available return data in our LiDAR metrics to develop individual tree prediction models. Smaller trees are difficult to discriminate because those trees tend to be found in an intermediate crown position where sufficient amount of pulses would not reach Ørka et al. (2009). Because forest stand structures in the PEF are generally described as a mixed species with multi-age structure, multiple pulse return data would be needed to provide a better depiction of individual tree characteristics.

Also, erroneous georegistration between individual tree locations and LiDAR point cloud seemed to leave a profound effect on the individual tree height and volume predictions in this study. To predict aboveground carbon density, Asner et al. (2009) reported that prediction errors were negligible due to erroneous georegistration between calibration plots and extracted LiDAR metrics plots in an area-based approach. However, Mascaro et al. (2011) reported that prediction error in aboveground carbon density tended to increase with increasing spatial resolution (*e.g.* smaller calibration plots in size).

2.4.1 Species Type Classification and Softwood Species Classification

Although intensity-related variables were available in our LiDAR metrics, we did not have an appropriate tool and other auxiliary data to calibrate for flying attitudes, terrain conditions, and atmospheric conditions for the intensity values. While Korpela et al. (2010) calibrated intensity values based on range-distance, and used the random forest technique to classify Norway spruce, Scots pine and birch, selected important classification variables were all intensity-related variables. Ørka et al. (2009) found that uncalibrated intensity variables were useful in classification between Norway spruce, Scots pine and birch. However, during both species type and softwood species classifications in our study, the random forest technique selected only one intensity-related variable as one of important classification variables, while mainly utilizing height-related variables (Tables 2.4a,b).

In the species type classification, kappa statistic in this classification was near 0%, which indicated that the agreement of correctly classified softwood and hardwood was purely by chance. Hardwood crowns tend to have different shapes depending on species, position in the crown and stem density when compared to softwoods. Thus, producer's accuracy in the hardwood classification was large (Table 2.5b). Korpela et al. (2010) had relatively lower classification accuracy in birch than Scots pine and Norway spruce, and noted that relative height differences within birch influenced in intensity values returned from the uppermost canopy surfaces. Vauhkonen et al. (2009) used intensity- and height-related variables to classify Norway spruce, Scots pine and birch, but a large number of birch tended to be misclassified as Scots pine while a classification between Norway spruce and Scots pine had a better result. Using full waveform LiDAR, Reitberger et al. (2008) reported that hardwood crown surface conditions were varied compared to softwoods, which would have resulted in diverse reflectance from the hardwood crowns. Additionally, Reitberger et al. (2008) and Vauhkonen et al. (2009) found that LiDAR data acquisition under a leaf-off condition had a better classification result in the species type classification because returned intensity-related variables were much different between softwood and hardwood.

Softwood species crown shapes were relatively similar among the species examined; therefore, height-related variables were not effective for classifying softwood species. As kappa statistics was 0%, this classification result was purely by chance. While Holmgren and Persson (2004) mainly used intensity-related variables to classify between Scots pine and Norway spruce, they had a lower classification accuracy for Scots pine because crown shapes of Scots pine varied depending on growth conditions. On the other hand, Suratno et al. (2009) reported that similar pulse return characteristics were observed among different species during a species classification process if those species grow in similar stand conditions such as a crown closure level or stem density; however, pulse intensity characteristics were dissimilar among species. Ørka et al. (2009) and Vauhkonen et al. (2010) found that different height among different species did not much influence in intensity-related variables. Donoghue et al. (2007) reported that intensity value was the most important variable to classify Sitka spruce (*Picea sitchensis*) and lodgepole pine (*Pinus contorta* var. *contorta*) in mixed Sitka spruce-lodgepole pine plots despite the calibration difficulty in intensity values. Thus, for future refinements in our research, the model needs to include appropriately calibrated intensity-related variables. Also, LiDAR systems equip, in general, one spectral wavelength at about 1064 nm, a near infrared region. The LiDAR sensor used in this study, the G-LiHT equips a spectral wavelength of 1550 nm (shortwave region), which may be less suited for species classification. Additionally, Li et al. (2013) reported that greater pulse density improved individual tree classification accuracy, and return pulses should be described in both vertical distribution and horizontal distribution for each individual crown.

2.4.2 Individual Tree Height Prediction

Although an agreement between field measured and model predicted individual tree height in all three models was weak (Table 2.3), one notable result was that predicted individual tree heights were associated with field-assessed crown positions. This association would be improved if we could improve horizontal accuracy between stem mapped trees and LiDAR point cloud. In the tree height prediction models, tree heights in the dominant crown position were constantly underestimated with greater predicted height. Most previous studies reported that LiDAR sensors tended to underestimate tree heights Clark et al. (2004); Næsset (1997) because low pulse density LiDAR likely resulted in insufficient direct hit on treetops Falkowski et al. (2006); Zimble et al. (2003). Although Wang and Glenn (2008) reported that heights of the conical crown shape of softwood trees tended to be underestimated to a greater extent than an ellipsoidal crown shape of hardwood trees, this study observed an opposite result as hardwood heights were generally underestimated. One reason might be that hardwood crown shapes in the PEF might be described as similar to a narrow and rounded shape due to increased crown competition. Another reason might be that the low pulse density LiDAR sensor used in this study could not sufficiently sense individual hardwood trees in the intermediate crown position, which were partially overtopped by trees in the dominant and codominant crown positions. Gonzalez-Ferreiro et al. (2013) noted that some pulses were reflected from the inside of crown rather than crown surfaces. Brandtberg et al. (2003) found that larger trees tended to be underestimated, but smaller trees were overestimated in height predictions. Vauhkonen et al. (2010) reported that height prediction accuracy was better in larger DBH trees than smaller DBH trees. In addition, although some pulses would have been returned from trees below the dominant crown trees, these pulse returns would be difficult to associate with trees in the intermediate crown trees from pulses returned within a dominant or codominant tree crown Brandtberg et al. (2003). Based on our results, those lower canopy LiDAR pulses were returned primarily from dominant or codominant crowns, which resulted in overestimated heights of these smaller trees.

While we added the classified species type as an additional covariate in the height prediction model, it did not much improve the predictions greatly. However, when we compared the field observed species type as a covariate (instead of the classified species type), the R^2 value was again barely improved. Therefore, it is inferred in this study that there was a limited relationship between individual tree height and species type due to the wide range of tree height between and within hardwood and softwood species in the mixed forest environment of the PEF. However, in the individual softwood tree height prediction, an additional covariate of classified softwood species resulted in a slightly better model fit. An explanation could be that certain softwood species such as balsam fir were often observed in the codominant or intermediate crown positions, while spruces tended to be found in the dominant crown position in the PEF.

2.4.3 Stem Volume Prediction

An agreement between field measured and model predicted individual tree volume in all three models was weak (Table 2.3). The field measured volume was derived using a species-specific taper equation, which requires individual tree DBH data besides total height data. Although this study did not report individual tree DBH predictions based on LiDAR metrics, we had low model fits during preliminary analysis. Therefore, due to relatively low accuracy of both height and DBH predictions, our individual tree method would not be an appropriate approach for the individual tree volume prediction. Yu et al. (2011) used similar pulse density LiDAR data as our study to deploy an individual tree-based method for volume prediction in a Scots pine and Norway spruce dominating boreal forest. Based on successfully matched trees between segmented and field located trees, relative RMSE of 21.58% was achieved. Also, Yu et al. (2010) and Yu et al. (2011) reported that omission errors during segmentation process in an individual tree-based method largely affected volume prediction while segmentation accuracy depended on stand structures. For example, segmentation accuracy was higher with lower stem density plots or larger DBH trees; therefore, higher volume prediction accuracy could be achieved for the lower stem density plots or the larger DBH trees in the individual tree-based method. Breidenbach et al. (2010) found that RMSE in volume prediction for birch and trembling aspen (*Populus tremula*) tended to be greater than Norway spruce and Scots pine although 91% of volume were comprised by these softwood species in their study site in Norway. In our study, the volume prediction model for softwood trees had a better model fit than the model for both hardwood and softwood trees. Thus, a high accuracy result in the species type classification would improve the volume predictions to stratify trees between softwood and hardwood.

2.4.4 Conclusion

Based on low density LiDAR data, the individual tree-based method deployed in this study for tree height and volume predictions did not result in high level accuracy and precision when compared to previously reported studies. While we initially hypothesized that the LiDAR metrics and individual tree height would be correlated to some degree, the low density LiDAR data used in this analysis was not sufficient for tree-level predictions. Also, we hypothesized that each tree species would have a rather unique crown shape and branching pattern, but our LiDAR data was not capable of distinguishing between either hardwood or softwood species. One possible explanation is that the mixed species and multi-age forest structure of the PEF promoted high competition for both hardwood and softwood trees, which has resulted in similar crown characteristics between and within a species. For future work, it is important to investigate how horizontal accuracy between LiDAR point cloud and individual trees in the field are matched. Also, it should be compared forest inventory predictions deployed by area- and individual tree-based approaches.

REFERENCES

Akay, A. E., H. Oguz, I. R. Karas, and K. Aruga (2009). Using LiDAR technology in forestry activities. *151*, 117–125.

Anderson, R. S. and P. V. Bolstad (2013). Estimating aboveground biomass and average annual wood biomass increment with airborne leaf-on and leaf-off LiDAR in Great Lakes Forest types. *Northern Journal of Applied Forestry 30*, 16–22.

Asner, G. P., R. F. Hughes, T. A. Varga, D. E. Knapp, and T. Kennedy-Bowdoin (2009). Environmental and biotic controls over aboveground biomass throughout a tropical rain forest. *Ecosystems 12*, 261–278.

Brandtberg, T., T. A. Warner, R. E. Landenberger, and J. B. McGraw (2003). Detection and analysis of individual leaf-off tree crowns in small footprint, high sampling density lidar data from the eastern deciduous forest in North America. *Remote Sensing of Environment 85*, 290–303.

Breidenbach, J., E. Næsset, V. Lien, T. Gobakken, and S. Solberg (2010). Prediction of species specific forest inventory attributes using a nonparametric semi-individual tree crown approach based on fused airborne laser scanning and multispectral data. *Remote Sensing of Environment 114*, 911–924.

Breiman, L. (2001). Random forest. *Machine Learning 45*, 5–32.

Chen, Q., D. Baldocchi, P. Gong, and M. Kelly (2006). Isolating individual trees in a savanna woodland using small footprint lidar data. *Photogrammetric Engineering & Remote Sensing 72*, 923–932.

Clark, M. L., D. B. Clark, and D. A. Roberts (2004). Small-footprint lidar estimation of subcanopy elevation and tree height in a tropical rain forest landscape. *Remote Sensing of Environment 91*, 68–89.

Congalton, R. G. and K. Green (1993). A practical look at the sources of confusion in error matrix generation. *Photogrammetric Engineering & Remote Sensing 59*, 641–644.

Cook, B. D., L. A. Corp, R. F. Nelson, E. M. Middleton, D. C. Morton, J. T. McCorkel, J. G. Masek, K. J. Ranson, V. Ly, and P. M. Montesano (2013). NASA Goddards LiDAR, Hyperspectral and Thermal (G-LiHT) Airborne Imager. *Remote Sensing 5*, 4045–4066.

Donoghue, D. N. M., P. J. Watt, N. J. Cox, and J. Wilson (2007). Remote sensing of species mixtures in conifer plantations using LiDAR height and intensity data. *Remote Sensing of Environment 110*, 509–522.

Evans, J. S., A. T. Hudak, R. Faux, and A. M. S. Smith (2009). Discrete return lidar in natural resources: recommendations for project planning, data processing, and deliverables. *Remote Sensing 1*, 776–794.

Falkowski, M. J., A. M. S. Smith, P. E. Gessler, A. T. Hudak, L. A. Vierling, and J. S. Evans (2008). The influence of conifer forest canopy cover on the accuracy of two individual tree measurement algorithms using lidar data. *Canadian Journal of Remote Sensing 34*, 338–350.

Falkowski, M. J., A. M. S. Smith, A. T. Hudak, P. E. Gessler, L. A. Vierling, and N. L. Crookston (2006). Automated estimation of individual conifer tree height and crown diameter via two-dimensional spatial wavelet analysis of lidar data. *Canadian Journal of Remote Sensing 32*, 153–161.

García, M., D. Riaño, E. Chuvieco, and F. M. Danson (2010). Estimating biomass carbon stocks for a Mediterranean forest in central Spain using LiDAR height and intensity data. *Remote Sensing of Environment 114*, 816–830.

Gobakken, T. and E. Næsset (2009). Assessing effects of positioning errors and sample plot size on biophysical stand properties derived from airborne laser scanner data. *Canadian Journal of Forest Research 39*, 1036–1052.

Goerndt, M. E., V. J. Monleon, and H. Temesgen (2010). Relating forest attributes with area- and tree-based light detection and ranging metrics for western Oregon. *Western Journal of Applied Forestry 25*, 105–111.

Gonzalez-Ferreiro, E., U. Dieguez-Aranda, L. Barreiro-Fernandez, S. Bujan, M. Barbosa, J. C. Suarez, I. J. Bye, and D. Miranda (2013). A mixed pixel- and region-based approach for using airborne laser scanning data for individual tree crown delineation in Pinus radiata D. Don plantations. *International Journal of Remote Sensing 34*, 7671–7690.

Hawbaker, T. J., T. Gobakken, A. Lesak, E. Trømborg, K. Contrucci, and V. Radeloff (2010). Light detection and ranging-based measures of mixed hardwood forest structure. *Forest Science 56*, 313–326.

Heidemann, H. (2012). Lidar base specification version 1.0: U.S. Geological survey techniques and methods. In *Book 11, collection and delineation of spatial data*, pp. 63.

Holmgren, J. and Å. Persson (2004). Identifying species of individual trees using airborne laser scanner. *Remote Sensing of Environment 90*, 415–423.

Hudak, A. T., N. L. Crookston, J. S. Evans, D. E. Hall, and M. J. Falkowski (2008). Nearest neighbor imputation of species-level, plot-scale forest structure attributes from LiDAR data. *Remote Sensing of Environment 112*, 2232–2245.

Hudak, A. T., J. S. Evans, and A. M. S. Smith (2009). LiDAR utility for natural resource managers. *Remote Sensing 1*, 934–951.

Hummel, S., A. T. Hudak, E. H. Uebler, M. J. Falkowski, and K. A. Megown (2011). A Comparison of Accuracy and Cost of LiDAR versus Stand Exam Data for Landscape Management on the Malheur National Forest. *Journal of Forestry 109*, 267–273.

Jaskierniak, D., P. N. J. Lane, A. Robinson, and A. Lucieer (2011). Extracting LiDAR indices to characterise multilayered forest structure using mixture distribution functions. *Remote Sensing of Environment 115*, 573–585.

Jensen, J. L. R., K. S. Humes, T. Conner, C. J. Williams, and J. DeGroot (2006). Estimation of biophysical characteristics for highly variable mixed-conifer stands using small-footprint lidar. *Canadian Journal of Forest Research 36*, 1129–1138.

Jing, L., B. Hu, J. Li, and T. Noland (2012). Automated delineation of individual tree crowns from lidar data by multi-scale analysis and segmentation. *Photogrammetric Engineering & Remote Sensing 78*, 1275–1284.

Kim, Y. S., Z. Q. Yang, W. B. Cohen, D. Pflugmacher, C. L. Lauver, and J. L. Vankat (2009). Distinguishing between live and dead standing tree biomass on the North Rim of Grand Canyon National Park, USA using small-footprint lidar data. *Remote Sensing of Environment 113*, 2499–2510.

Korpela, I., H. O. Orka, M. Maltamo, T. Tokola, and J. Hyyppa (2010). Tree species classification using airborne LiDAR - effects of stand and tree parameters, downsizing of training set, intensity normalization, and sensor type. *SIilva Fennica*, 319–339.

Kwak, D., W. Lee, H. Cho, S. Lee, Y. Son, M. Kafatos, S. Kim, K. DooAhn, L. WooKyun, C. HyunKook, L. SeungHo, S. YoWhan, and K. SoRa (2010). Estimating stem volume and biomass of Pinus koraiensis using LiDAR data. *Journal of Plant Research 123*, 421–432.

Lee, H., K. C. Slatton, B. E. Roth, and W. P. Cropper Jr. (2010). Adaptive clustering of airborne LiDAR data to segment individual tree crowns in managed pine forests. *International Journal of Remote Sensing 31*, 117–139.

Lefsky, M. A., W. B. Cohen, G. G. Parker, and D. J. Harding (2002). Lidar remote sensing for ecosystem studies. *BioScience 52*, 19–30.

Li, J., B. Hu, and T. L. Noland (2013). Classification of tree species based on structural features derived from high density LiDAR data. *Agricultural and Forest Meteorology 171*, 104–114.

Li, R., A. Weiskittel, A. R. Dick, J. A. Kershaw, R. S. Seymour, and J. Kershaw Jr. (2012). Regional stem taper equations for eleven conifer species in the Acadian Region of North America: development and assessment. *Northern Journal of Applied Forestry 29*, 5–14.

Li, W. K., Q. H. Guo, M. K. Jakubowski, and M. Kelly (2012). A new method for segmenting individual trees from the lidar point cloud. *Photogrammetric Engineering & Remote Sensing 78*, 75–84.

Li, Y. Z., H. E. Andersen, and R. McGaughey (2008). A comparison of statistical methods for estimating forest biomass from light detection and ranging data. *Western Journal of Applied Forestry 23*, 223–231.

Liaw, A. and M. Wiener (2002). Classification and regression by randomForest. *R News 2*, 18–22.

Lim, K. S. and P. M. Treitz (2004). Estimation of above ground forest biomass from airborne discrete return laser scanner data using canopy-based quantile estimators. *Scandinavian Journal of Forest Research 19*, 558–570.

Magnussen, S. and P. Boudewyn (1998). Derivations of stand heights from airborne laser scanner data with canopy-based quantile estimators. *Canadian Journal of Forest Research 28*, 1016–1031.

Magnusson, M., J. E. S. Fransson, and J. Holmgren (2007). Effects on estimation accuracy of forest variables using different pulse density of laser data. *Forest Science 53*, 619–626.

Mascaro, J., M. Detto, G. P. Asner, and H. C. Muller-Landau (2011). Evaluating uncertainty in mapping forest carbon with airborne LiDAR. *Remote Sensing of Environment 115*, 3770–3774.

McGaughey, R. (2013). Fusion/LDV: Software for lidar data aalysis and vsualization, 3.30. *USDA Forest Service - Pacific Northwest Research Station*.

McWilliams, W., B. Butler, L. Caldwell, D. Griffith, M. Hoppos, and K. Laustsen (2005). The forests of Maine. *U.S. Department of Agriculture, Forest Service, Northeastern Research Station, Newton Square, Pennsylvania*, pp188.

Means, J. E., S. A. Acker, B. J. Fitt, M. Renslow, L. Emerson, and C. J. Hendrix (2000). Predicting forest stand characteristics with airborne scanning lidar. *Photogrammetric Engineering & Remote Sensing 66*, 1367–1371.

Næsset, E. (1997). Determination of mean tree height of forest stands using airborne laser scanner data. *ISPRS Journal of Photogrammetry and Remote Sensing 52*, 49–56.

Næsset, E. (2002). Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. *Remote Sensing of Environment 80*, 88–99.

Næsset, E. (2004). Practical large-scale forest stand inventory using a small-footprint airborne scanning laser. *Scandinavian Journal of Forest Research 19*, 164–179.

Næsset, E. and T. Økland (2002). Estimating tree height and tree crown properties using airborne scanning laser in a boreal nature reserve. *Remote Sensing of Environment 79*, 105–115.

Nilsson, M. (1996). Estimation of tree heights and stand volume using an airborne lidar system. *Remote Sensing of Environment 56*, 1–7.

Olson, M. G. and R. G. Wagner (2010). Long-term compositional dynamics of Acadian mixedwood stands under different silvicultural regimes. *Canadian Journal of Forest Research 40*, 1993–2002.

Ørka, H. O., E. Næsset, and O. M. Bollandsås (2009). Classifying species of individual trees by intensity and structure features derived from airborne laser scanner data. *Remote Sensing of Environment 113*, 1163–1174.

Parker, R. C. and P. A. Glass (2004). High- versus low-density LiDAR in a double-sample forest inventory. *Southern Journal of Applied Forestry 28*, 205–210.

Persson, Å., J. Holmgren, and U. Söderman (2002). Detecting and measuring individual trees using an airborne laser scanner. *Photogrammetric Engineering & Remote Sensing 68*, 925–932.

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

Popescu, S. C. and R. H. Wynne (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*, 589–604.

R Development Core Team (2012). R: A language and environment for statistical computing. *R Foundation for Statistical Computing: Vienna, Austria*.

Reitberger, J., P. Krzystek, and U. Stilla (2008). Analysis of full waveform LIDAR data for the classification of deciduous and coniferous trees. *International Journal of Remote Sensing 29*, 1407–1431.

Robinson, A. P. and W. R. Wykoff (2004). Imputing missing height measures using a mixed-effects modeling strategy. *Canadian Journal of Forest Research 34*, 2492–2500.

Rosenfield, G. H. and K. Fitzpatrick-Lins (1986). A coefficient of agreement as a measure of thematic classification accuracy. *Photogrammetric Engineering & Remote Sensing 52*, 223–227.

Sendak, P. E., J. C. Brissette, and R. M. Frank (2003). Silviculture affects composition, growth, and yield in mixed northern conifers: 40-year results from the Penobscot Experimental Forest. *Canadian Journal of Forest Research 33*, 2116–2128.

Stone, C., T. Penman, and R. Turner (2011). Determining an optimal model for processing lidar data at the plot level: results for a Pinus radiata plantation in New South Wales, Australia. *New Zealand Journal of Forestry Science 41*, 191–205.

Suratno, A., C. Seielstad, and L. Queen (2009). Tree species identification in mixed coniferous forest using airborne laser scanning. *ISPRS Journal of Photogrammetry and Remote Sensing 64*, 683–693.

Thomas, V., P. Treitz, J. H. McCaughey, and I. Morrison (2006). Mapping stand-level forest biophysical variables for a mixedwood boreal forest using lidar: an examination of scanning density. *Canadian Journal of Forest Research 36*, 34–47.

Treitz, P., K. Lim, M. Woods, D. Pitt, D. Nesbitt, and D. Etheridge (2012). LiDAR sampling density for forest resource inventories in Ontario, Canada. *Remote Sensing 4*, 830–848.

van Aardt, J. A. N., R. H. Wynne, R. G. Oderwald, and J. A. N. van Aardt (2006). Forest volume and biomass estimation using small-footprint lidar-distributional parameters on a per-segment basis. *Forest Science 52*, 636–649.

Vauhkonen, J., I. Korpela, M. Maltamo, and T. Tokola (2010). Imputation of single-tree attributes using airborne laser scanning-based height, intensity, and alpha shape metrics. *Remote Sensing of Environment 114*, 1263–1276.

Vauhkonen, J., T. Tokola, P. Packalén, and M. Maltamo (2009). Identification of Scandinavian Commercial Species of Individual Trees from Airborne Laser Scanning Data Using Alpha Shape Metrics. *Forest Science 55*, 37–47.

Wang, C. and N. F. Glenn (2008). A linear regression method for tree canopy height estimation using airborne lidar data. *Canadian Journal of Remote Sensing 34*, s217–s227.

Weiskittel, A. and R. Li (2012). Development of regional taper and volume equations: Hardwood species. *University of Maine, School of Forest Resources: Orono, ME*, 87–95.

Weiskittel, A., M. Russell, R. Wagner, and R. Seymour (2012). Refinement of the forest vegetation simulator northeast variant growth and yield model: Phase iii. *University of Maine, School of Forest Resources: Orono, ME*, 96–104.

Woods, M., D. Pitt, M. Penner, K. Lim, D. Nesbitt, D. Etheridge, and P. Treitz (2011). Operational implementation of a LiDAR inventory in Boreal Ontario. *87*, 512–528.

Yu, X., J. Hyyppä, M. Holopainen, and M. Vastaranta (2010). Comparison of Area-Based and Individual Tree-Based Methods for Predicting Plot-Level Forest Attributes. *Remote Sensing 2*, 1481–1495.

Yu, X., J. Hyyppä, M. Vastaranta, M. Holopainen, and R. Viitala (2011). Predicting individual tree attributes from airborne laser point clouds based on the random forests technique. *ISPRS Journal of Photogrammetry and Remote Sensing 66*, 28–37.

Zimble, D. A., D. L. Evans, G. C. Carlson, R. C. Parker, S. C. Grado, and P. D. Gerard (2003). Characterizing vertical forest structure using small-footprint airborne LiDAR. *Remote Sensing of Environment 87*, 171–182.

BIOGRAPHY OF THE AUTHOR

Rei Hayashi was born in Koganei City, Tokyo, Japan. He attended Oregon State University in Corvallis, Oregon where he received a Bachelor of Science degree in Forest Management in 2003. Also, he attended University of Georgia in Athens, Georgia where he received a Master of Science degree in Forest Resources in 2006.

Rei Hayashi is a candidate for the Master of Science degree in Forest Resources from University of Maine in June 2010.