AN ABSTRACT OF THE THESIS OF

Joonghoon Shin for the degree of Master of Science in Forest Resources presented on September 17, 2018.

Title: Estimating Forest Inventory Attributes Using Airborne LiDAR in Southwestern Oregon.

Abstract approved:

____

Temesgen Hailemariam

This thesis mainly consists of two parts: (1) comparing statistical modeling methods based on the area-based approach (ABA) for predicting forest inventory attributes using airborne light detection and ranging (LiDAR) data (Chapter 2), and (2) suggesting a new methodology fusing the individual tree detection (ITD) approach and the ABA for generating tree-lists using airborne LiDAR data (Chapter 3).

Chapter 2 compared selected modeling methods used to predict five forest attributes, basal area (BA), stem volume (VOL), Lorey's height (LOR), quadratic mean diameter (QMD), and tree density (DEN), from airborne LiDAR metrics in southwestern Oregon, USA. The selected methods included most similar neighbor (MSN) imputation, gradient nearest neighbor (GNN) imputation, Random Forest (RF) based imputation, BestNN imputation, Ordinary least square (OLS) regression, spatial linear model (SLM), and geographically weighted regression (GWR). Several performances of each method were assessed by 500 simulations with different

numbers of training data. No modeling methods was always superior to the others in prediction of the forest attributes. The best method varied according to response variable, prediction type, and performance measures, even though there was a leading group (SLM, OLS, BestNN, and GWR) that always outperformed the other methods in root mean squared prediction error (RMSPE). Model's performance was quite affected when a small number of training data was used in modeling procedure. The optimal sizes of training data were 100-150 for point prediction and 200-250 for total prediction. SLM showed its applicability to wider conditions in that it produced better performance in most cases. RF imputation produced poorer performances than the other methods, particularly with lower prediction interval coverage. This might be because RF imputation had some bias and smaller prediction standard error; RF's poor performance did not stem from the smaller number of predictor variables.

In Chapter 3, a new approach, combining ITD and ABA, was proposed to generate tree-lists using airborne LiDAR data. ITD based on the Canopy Height Model (CHM) was applied for overstory trees, while ABA based on nearest neighbor (NN) imputation was applied for understory trees. The approach is intended to compensate for the weakness of LiDAR data and ITD in estimating understory trees, keeping the strength of ITD in estimating overstory trees in tree-level. We investigated the effects of three parameters on the performance of our proposed approach: smoothing of CHM, resolution of CHM, and height cutoff (a specific height that classifies trees into overstory and understory). There was no single combination of those parameters that produced the best performance for estimating stems per ha, mean tree height, basal area, diameter distribution and height distribution. The trees in the lowest LiDAR height class yielded the largest relative bias and relative root mean squared error. Although ITD and ABA showed limited explanatory powers to estimate stems per hectare and basal area, there could be improvements from methods such as using LiDAR data with higher density, applying better algorithms for ITD and decreasing distortion of the structure of LiDAR data. Automating the procedure of finding optimal combinations of those parameters is essential to expedite forest management decisions across forest landscapes using remote sensing data.

Estimating Forest Inventory Attributes Using Airborne LiDAR in Southwestern Oregon

by
Joonghoon Shin

A THESIS

submitted to

Oregon State University

in partial fulfillment of
the requirements for the
degree of

Master of Science

Presented September 17, 2018
Commencement June 2019

Master of Science thesis of Joonghoon Shin presented on September 17, 2018

APPROVED:

Major Professor, representing Forest Resources

Head of the Department of Forest Engineering, Resources and Management

Dean of the Graduate School

I understand that my thesis will become part of the permanent collection of Oregon State University libraries. My signature below authorizes release of my thesis to any reader upon request.

Joonghoon Shin, Author

ACKNOWLEDGEMENTS

First of all, I am grateful to Lord God for Your help and love. I praise You for carrying me through life, I will follow you wherever I go. I would like to thank my major professor Dr. Temesgen Hailemariam. He has provided me with support and encouragement since I became an FMBL member. Without his guidance, I could not be here today. Profound thanks to other committee members. Dr. Jay Ver Hoef shared his valuable R codes and offered me his insights. Dr. Thomas Hilker who was my committee member, I will not forget your kindness, and hope you rest in peace in heaven. Dr. Bogdan Strimbu kindly accepted to serve as a committee member although I asked him on such short notice.

Special thanks to Stephen E. Reutebuch and Robert J. McGaughey from USDA Forest Service for their help and support for my research. I also appreciate Dr. Jim Flewelling for his insights and review of an earlier draft for chapter 2. And thanks to FMBL members, Dr. Krishna Poudel, Dr. Francisco Mauro, Dr. Jacob Strunk, Karin Kralicek, Bryce Frank, Al Pancoast, Ty Nietupski, Mike Shettles, Lacey Jeroue, Elijah Allensworth, Steve Huff, and Michael Hoe. It was a great pleasure working with the FMBL members. Especially, I wish Dr. Krishna Poudel's success in his new position. Also thanks to Madison Dudley in the FERM office. She has helped me a lot on everything related to academic affairs. Finally, I would like to thank my father, mother and younger brother for their love and understanding.

This work was funded by BLM and USDA Forest Service.

CONTRIBUTION OF AUTHORS

Dr. Temesgen Hailemariam, Dr. Jacob L. Strunk and Dr. Thomas Hilker provided valuable comments and contributions to the analysis and writing of Chapter 2. Dr. Temesgen Hailemariam provided extensive input on Chapter 3.

TABLE OF CONTENTS

LIST OF FIGURES

LIST OF FIGURES (Continued)

LIST OF TABLES

LIST OF TABLES (Continued)

Chapter 1: General Introduction

Forestry has been increasingly using remote sensing data for various goals of forest management. Among various types of remote sensing data, airborne light detection and ranging (LiDAR), which can provide the profiles of forest structure by penetrating forest canopies, has reached operational applications in forestry; typical forest inventory, habitat and biodiversity assessment, canopy structure and forest fuel estimation (Vauhkonen et al., 2014).

In scaling forest inventory information from field surveys or remotely sensed data, statistical modeling is essential to obtain wall-to-wall information from discretely sampled observations. There has been various modeling methods for using either LiDAR or other remotely sensed data: ordinary least square (OLS) regression (Gobakken et al. 2012; Strunk et al. 2008), nonlinear least squares regression (Packalén et al. 2011), most similar neighbor (MSN) imputation (Kankare et al. 2013; Muinonen et al. 2001), gradient nearest neighbor (GNN) imputation (Hudak et al. 2008; Hudak et al. 2014; Pierce et al. 2009; Temesgen and Ver Hoef 2015), Random-Forest (RF)-based imputation (Falkowski et al. 2010; Hudak et al. 2008; Hudak et al. 2012; Temesgen and Ver Hoef 2015), geographically weighted regression (GWR) technique (Chen et al. 2011; Gagliasso et al. 2014), Artificial Neural Networks (ANN) techniques (Niska et al. 2010), and others.

Corona et al. (2014) and Junttila et al. (2015) showed that dealing with spatial autocorrelation in predicting forest attributes could improve the performance of the prediction. Thus, we might obtain more insight to analyzing a model's performance by considering spatial autocorrelation in modeling process. Spatial linear model (SLM) has shown potential as a modeling method for forestry applications (Temesgen and Ver Hoef 2015; Ver Hoef and Temesgen 2013). SLM is similar to OLS regression in that it assumes a linear model. But in contrast to OLS regression, SLM assumes that its model error has spatial dependence that is described by a variance–covariance matrix, which means that it can explain spatial autocorrelation in given data.

Despite the growing research on GWR, GNN, and RF and their wide use for mapping and estimating stand density, basal area, and volume per hectare for the Pacific Northwest forests, detailed analyses that compare the performance, efficiency, and suitability of GWR, RF, GNN, and SLM for predicting (or mapping) LOR, QMD, DEN, basal area, and volume per hectare, at point and block-level are lacking. The first part of this thesis compared the performances of the selected modeling methods for predicting 5 forest attributes: basal area (BA), stem volume (VOL), Lorey's height (LOR), quadratic mean diameter (QMD), and tree density (DEN) using airborne LiDAR and ground data. The selected modeling methods are MSN with k = 1 and 5, RF-based imputation with k = 1 and 5, GNN with k = 1 and 5, BestNN—a modified k-NN method by Ver Hoef and Temesgen (2013)—OLS, SLM, and GWR.

A tree-list can provide detailed data such as tree species, diameter at breast height (DBH), tree height (HT), basal area (BA) and stem volume required for forest management and planning. As applied to typical forest inventories, airborne LiDAR can be applied to generating tree-lists. For tree-list generation, however, it is necessary to use a different approach from the first part of this thesis.

There are two major approaches in using airborne LiDAR in forestry – the area-based approach (ABA) used in the first part of this thesis, and the individual tree detection (ITD) approach (Vauhkonen et al., 2014). ABA assumes that the vertical height distribution of laser point clouds is related to variables of interest in an area. A host of summary statistics derived from the point cloud are used to predict many forest inventory attributes. In contrast, ITD identifies individual trees and provides estimates of forest attributes based on the identified individual trees. A rasterized canopy height model (CHM) is commonly used to segment individual trees with horizontal location of treetop and height across the CHM area. Therefore, ITD would be more suitable for tree-level inventories like tree-lists than ABA.

However, information on understory vegetation is likely to be missed when applying ITD (Koch et al., 2014). Additionally, it is well known that LiDAR has weaknesses for detecting or estimating understory vegetation regardless of the approach used because LiDAR data lack information on understory vegetation (lower proportion of point clouds in understory) (Takahashi et al., 2006).

Many approaches have been proposed to overcome the limitations above. Maltamo et al. (2004) combined a theoretical probability distribution function with the tree height distribution estimated from ITD to detect small and suppressed trees. ITD first estimated the height distribution and the number of large trees, and then the Weibull distribution was applied to estimate small trees. Lindberg et al. (2010)

proposed a methodology to generate a tree-list combining a CHM-based ITD and ABA estimation. To better detect trees that are close to each other or small; 1) the number of trees per segment was estimated using a training dataset in which the number of field-measured trees for each tree crown segment was known, and 2) a candidate tree-list from the ITD was calibrated using the target distributions of HT and DBH estimated by a *k*-Nearest Neighbor (NN) approach. Other than those, a 3 D clustering method (Lindberg et al., 2013) and the use of vertical stratification of point cloud and LiDAR data with high point cloud density (50 points / m^2) (Hamraz et al., 2017) were proposed to improve detecting understory trees.

There are many parameters that affect the performance of ITD. As biological parameters, forest structure, stand density, and tree clustering affected more than tree detection techniques (Vauhkonen et al., 2012). On the contrary to this, Kaartinen et al. (2012) reported that the methods for ITD were found to affect the performance of ITD.

According to a typical ITD method (Yu et al., 2010), the performance of ITD is affected by the smoothing and resolution for CHM, and the algorithm used for tree segmentation. In addition to those parameters, Wiggins (2017) reported that excluding trees below a specific height (minimum height cutoff) improved ITD’s accuracy for overstory trees. Maltamo et al. (2003) noted that a proper value of the truncation parameter of Weibull for DBH distribution, which can be considered the same as a height cutoff, should be further studied.

Other than ITD, detailed information on forest resources, such as a tree-list or stand table, has been mainly estimated by diameter distribution modeling or imputation. In diameter distribution modeling, parameters of some theoretical distributions are estimated to describe the distribution of tree diameters. Imputation methods directly substitute measured values from sample locations (references) for locations for which a prediction is desired (targets). Temesgen et al. (2003) used a set of proxy variables to represent a tree-list in NN imputations because there is no single variable to represent the tree-list. On the other hand, Strunk et al. (2017) used plot identities as a response variable in NN imputations in evaluating NN strategies to impute a tree-list.

In the second part of this thesis, we combined ABA and ITD to estimate tree-list using airborne LiDAR data inspired by the ideas from Maltamo et al. (2003), Maltamo et al. (2004) and Wiggins

(2017). This was for overcoming the weakness of LiDAR data and the ITD method in identifying understory trees, and utilizing the strength of ITD over ABA. ITD with watershed segmentation (Vincent and Soille, 1991) was used for estimating overstory trees (trees taller than a height cutoff) and ABA by NN ($k = 1$) imputation was used for estimating understory trees (trees shorter than the height cutoff). We examined the effects of the combination of the three parameters, smoothing of CHM, resolution of CHM and the height cutoff, as well as LiDAR height classification of field plots on estimating tree-lists via ITD. The explanatory power of our approach was also investigated. We evaluated the performance of generating tree-lists in terms of BA, mean HT, stems per hectare (SPH), and distributions of DBH and HT.

Chapter 2: Comparing Modeling Methods for Predicting Forest Attributes Using LiDAR Metrics and Ground Measurements

Joonghoon Shin, Hailemariam Temesgen, Jacob L. Strunk, and Thomas Hilker

Canadian Journal of Remote Sensing

Canadian Aeronautics and Space Institute

350 Terry Fox Drive, Suite 104

Kanata, Ontario K2K 2W5, CANADA

Volume 42, Issue 6, 2016

Introduction

Forest inventory is fundamental to forest planning and forest policy, as it provides information about the quality and quantity of forest resources. Traditional forest inventory has used field surveys for collecting ground data, and has used remotely sensed data mainly for stand delineation and creation of strata, however this method is expensive and time-consuming. Recent advances in remote sensing technologies such as image spectroscopy, high-resolution satellite imagery, microwave radars, and laser scanning allow surveying large areas at reasonable costs (Holopainen and Kalliovirta, 2006; McRoberts et al., 2010). Airborne light detection and ranging (LiDAR), as an independent measurement source, in particular has demonstrated ability to improve the accuracy of forest inventory parameters, such as height and volume estimates. Although wall-to-wall LiDAR information across large areas is not currently available due to high costs and data volumes (Wulder et al., 2012), LiDAR supported forest inventory is an important and active topic in forest research and other related fields.

Scaling forest inventory information acquired from field surveys or remotely sensed data requires statistical modeling to obtain wall-to-wall information from discretely sampled observations. Various methods have been implemented for using either LiDAR or other remote sensed data such as ordinary least square (OLS) regression (Gobakken et al., 2012; Strunk et al., 2008), non-linear least squares regression (Packalén et al., 2011), most similar neighbor (MSN) imputation (Kankare et al., 2013; Muinonen et al., 2001), gradient nearest neighbor (GNN) imputation (Hudak et al., 2008; Hudak et al., 2014; Pierce et al., 2009; Temesgen and Ver Hoef, 2015), Random Forest (RF) based imputation (Falkowski et al., 2010; Hudak et al., 2008; Hudak et al., 2012; Temesgen and Ver Hoef, 2015), geographically weighted regression (GWR) technique (Chen et al., 2011; Gagliasso et al., 2014), Artificial Neural Networks (ANN) techniques (Niska et al., 2010) etc.

Among the methods above, arguably MSN and GNN have been widely used in the Pacific Northwest while linear regression modeling such as OLS has been used widely with LiDAR data and biomass estimation by most remote sensing analysts (Næsset and Gobakken, 2008). For instance, MSN methods have been used for tree volume estimation, quadratic mean diameter and stand density (LeMay et al., 2008), stand table (Temesgen et al., 2003), forest aboveground biomass and stem density (Kankare et al., 2013), and cavity tree abundance (Temesgen et al., 2008). Application examples for GNN are wildland fuel and forest structure attributes such as canopy cover, basal area, etc. (Pierce et al., 2009), and tree species composition or forest structure attributes such as total basal area, stand density, etc. (Ohmann and Gregory, 2002). Similarly, RF-based imputation methods have been used increasingly (Eskelson et al., 2009b; Hudak et al., 2008; Hudak et al., 2014; Latifi and Koch, 2012; Latifi et al., 2010). One advantage of RF methods is that it does not require a variable selection procedure to handle full sets of variables (Latifi et al., 2010; Penner et al., 2013). However, some potential issues of RF imputation have been reported. Breidenbach et al. (2010a) described that RF *k*-NN consistently produced larger biases than other NN methods. RF imputation with large number of predictor variables did not improve the modeling performance compared to RF imputation with small number of predictors (Breidenbach et al., 2010a; Vauhkonen et al., 2010). OLS regression is a common modeling method and shows competitiveness in estimation of forest biomass and carbon stock using LiDAR data compared with other modeling methods such as RF, boosted regression trees, support vector regression, etc. (Li et al., 2014). GWR is a spatial predictor that assumes spatial non-stationarity in the relationship between response and predictor variable (Chen et al., 2011). In Gagliasso et al. (2014), the lowest root mean squared error (RMSE) for estimating basal area were produced by GWR using ground measurements, LiDAR metrics, satellite image data and climate data in eastern Oregon, USA.

As an alternative to the methods described, spatial linear model (SLM) has emerged and showed potential for forestry applications including inventory (Temesgen and Ver Hoef, 2015; Ver Hoef and Temesgen, 2013). SLM is similar to OLS regression in that it assumes a linear model. But in contrast to OLS regression, SLM assumes that its model error has spatial dependence that is described by a variance-covariance matrix. SLM has shown improved results compared to RF imputation and GNN (Temesgen and Ver Hoef, 2015), and achieved better performance than OLS regression, NN using Mahalanobis distance, MSN, and BestNN (see method section for BestNN) for estimation of potential forest productivity and biomass (Ver Hoef and Temesgen, 2013).

Environmental measurements are often autocorrelated when measurements are taken near in space to each other (Hoeting, 2009). Spatial autocorrelation is not, however, typically considered in modeling methods such as OLS, nearest neighbor imputation, or ANN for forest attributes (Brosofske et al., 2014). Leveraging the observed spatial correlation can improve the quality of inferences from the data (Schabenberger and Gotway, 2004). In Corona et al. (2014), universal kriging and co-kriging showed better performance than locally weighted regression and k-nearest neighbor imputation when there was strong spatial autocorrelation of forest variables. Junttila et al. (2015) suggested handling spatial correlation of model residuals to further improve their Bayesian linear model for estimating forest attributes with LiDAR data. Therefore, dealing with spatial autocorrelation in predicting forest attributes might give more insight to analyzing model's performance.

Temesgen and Ver Hoef (2015) compared the suitability and performance of *k*-NN, RF, GNN, and SLM empirically, and reported that SLM is a better option for point and total prediction of forest biomass and potential productivity. Despite the growing research on GWR, GNN and RF and their wide use for mapping and estimating stand density, basal area, and volume per hectare for the Pacific Northwest forests, detailed analyses that compare the performance, efficiency and suitability of GWR, RF, GNN and SLM for predicting (or mapping) LOR, QMD, DEN, basal area and volume per hectare, at point and block-level are lacking. The overall goal of this article is to make similar comparisons of SLM to GWR, GNN and RF that were missing from Temesgen and Ver Hoef (2015).

The goal of the study was to compare the performances of the selected modeling methods for predicting five forest attributes, basal area (BA), stem volume (VOL), Lorey's height (LOR), quadratic mean diameter (QMD), and tree density (DEN), using Bureau of Land Management LiDAR and ground data. The statistical performances of the modeling methods from simulations are assessed. The selected modeling methods are MSN with *k*=1 and 5, RF-based imputation with *k*=1 and 5, GNN with *k*=1 and 5, BestNN: a modified *k*-NN method by Ver Hoef and Temesgen (2013), OLS, SLM, and GWR.

Methods

Study Area

Our study area covers four counties in southwestern Oregon: Coos, Curry, Douglas, and Lane counties (Figure 2.1). The extent of the area is 647,951 hectares with an elevation range of approximately 20 – 1,000 m above sea level. Douglas-fir (*Pseudotsuga menziesii*) is the main tree species in the study area. Other important species are western hemlock (*Tsuga heterophylla*), red alder (*Alnus rubra*), Oregon myrtle (*Umbellularia californica*), bigleaf maple (*Acer macrophyllum*), tanoak (*Lithocarpus*), western redcedar (*Thuja plicata*), and grand fir (*Abies grandis*).

Airborne LiDAR data collection specifications and processing

An airborne LiDAR survey was conducted in the study area between April 27th, 2008 and April 5th, 2009 using Leica ALS50 Phase II instrumentation. The average pulse density (the average number of pulses returned from surfaces) was 8.10/m^2 for the study area. Further specifications of the LiDAR survey are shown in Table 2.1. LiDAR metrics were computed over the entire acquisition area with grid cells of 22.86 m by 22.86 m using all returns by FUSION software (McGaughey, 2010). Points with elevations above ground level less than 1 m and larger than 91.44 m were excluded from the computation.

Field plot sampling

Field plots for this study were collected by stratifying the study area using LiDAR height metrics (Hawbaker et al., 2009). This strategy has been shown to improve the predictive accuracy of resultant models by increasing the variability observed in the explanatory variables.

A stratified sampling procedure was implemented as follows. First, a large random sample was taken of the grid cells in the study area described in the section 2.2 to obtain an idea of the range of the two chosen LiDAR metrics, 80th height percentile (P80) and standard deviation of LiDAR heights (SDH). Based on the sample cells, the P80 was subdivided into 10 classes of 6.10 m each. The maximum height of the 10th P80 class bin was increased up to 83.52 m to cover the values of the gird cells in the full dataset. And SDH within each P80 class was subdivided into three equal width classes. The minimum value of all the lowest SDH classes was set to zero, and the maximum value of all of the highest SDH classes was set to the highest SDH value in the full dataset, 35.36 m. Bins are larger than or equal to minimum values, and are less than maximum values. Consequently, thirty bins (10 x 3) were developed.

Fifty potential plot locations were randomly sampled from each of the 30 bins shown in Figure 2.2. Field crews obtained 30 plots in each bin (Figure 2.2). Eight plots were found to have no live trees, and another one in bin number 12 was missed; these nine plots were removed, leaving a total of 891 plots.

Ground measurement

A field survey was implemented between May 25th, 2010 and May 10th, 2011. Nested plots with two plot sizes were used to measure trees at each plot center. For the plot size of 12.65 m radius (505.9 m^2), all live trees with a DBH of 14.0 cm and larger were measured. Five attributes were calculated for each plot and used as response variables for the statistical models described below: BA (m^2/ha), VOL (m^3/ha), LOR (m), QMD (cm), and DEN (trees/ha). Tree-level volume was predicted using the USDA Forest Service National Volume Estimator Library Excel Volume Functions from: http://www.fs.fed.us/fmsc/measure/volume/nvel/index.php. LOR was computed as basal area weighted mean height (Husch et al., 2002). A summary of plot-level forest attributes is provided in Table 2.2. The plot center locations were measured using GPS receiver and recorded in UTM 10 NAD83 CONUS meters.

Modeling methods

The modeling methods used in this study fall into two categories.

- Non-parametric methods: do not depend on theoretical probability distribution, and are implicit (can predict only values that are within range of training data). e.g. *k*-nearest neighbor (NN) imputation methods such as MSN (*k*=1, 5), RF (*k*=1, 5), GNN (*k*=1, 5), BestNN.
- Parametric methods: depend on theoretical probability distribution, and are explicit (can predict values that are outside of range of training data). e.g. linear modeling methods such as OLS, SLM, and GWR.

k-NN methods work by direct substitution (imputation) of measured values from sample locations (references) for locations for which we desire a prediction (targets). In this strategy, key considerations include the distance metric that we use to identify suitable references, and the number of references (k) that are used in a single imputation (prediction) (Eskelson et al., 2009a). We examined both k=1 and k=5 neighbors for each of the following distance metrics that we used. The first distance, canonical correlation, was termed – MSN when proposed by Moeur and Stage (1995). Ohmann and Gregory (2002) suggested the GNN method using distance based on canonical correspondence analysis (CCA). CCA is considered to work well in identifying relationships which occur along a gradient. BestNN, a modified k-NN by Ver Hoef and Temesgen (2013), adopts Mahalanobis distance and MSN procedure as distance metrics, and tries k = 1, 2, ..., 30. Then the distance metric and k with the smallest root mean squared prediction error (RMSPE) were selected using a cross-validation approach based on the observed data. The final k-NN method, Random Forest (Breiman, 2001) based NN, measures the distances between observations by RF algorithm. RF constructs multiple classification (or regression) trees with bootstrap samples of training data, while selecting predictors randomly to find the best split at each node in the trees. In RF based NN, the distance is computed as one minus the proportion of classification trees where a target observation is in the same terminal node as a reference observation (Hudak et al., 2008).

A key assumption under which OLS is a best linear unbiased estimator is that the residuals are independent. A more efficient estimator is feasible when this assumption is violated by estimating values in the variance-covariance matrix using restricted maximum likelihood and a fixed model for spatial dependence (Ver Hoef and Temesgen, 2013). The last linear modeling method used in this study, GWR was originally proposed by Brunsdon et al. (1998). GWR allows model coefficients to vary across geographical space, whereas OLS regression assumes that the coefficients are constant across the space. This means that GWR considers spatial non-stationarity in the relationship between response variable and predictor variables (Chen et al., 2011). And advantages and disadvantages of the selected modeling methods are given in Table 2.3. Modeling strategies were implemented in R (R Core Team, 2013). The k-NN imputation methods were implemented by the '*yaImpute*' package (Crookston and Finley, 2008) in combination with the '*randomForest*' package (Liaw and Wiener, 2002). We also used the '*stats*' package (R Core Team, 2013) and the '*spgwr*' package (Bivand et al., 2013) for OLS and GWR, respectively. The '*bbslm*' package (Ver Hoef 2012; not published yet) was used for SLM.

LiDAR metrics and variable selection

We used FUSION (McGaughey, 2010) to generate the candidate LiDAR metrics in Table 2.4. We removed lower percentile height (less than 50^{th}) because they are not representative for the upper-story vegetation. Other metrics that are hard to biologically interpret were also excluded from consideration.

ʹBest subset' variable selection procedure was implemented with the candidate LiDAR metrics of full dataset to obtain a final set of predictor variables for every modeling method. Because there is no variable selection technique that can be applied commonly to all the selected methods, the 'best subset' procedure based on OLS regression was used. The 'best subset' procedure was implemented by the function '*regsubsets*' in R package '*leaps*' (Lumley, 2009). For model's parsimony, the maximum number of predictor variables in a model was set as three through the 'best subset' variable selection procedure. The best 15 selections of predictor variables were produced for each of models with one, two, and three predictor variables, respectively. The preferred models were selected based on Bayesian information criteria (BIC). Then, we selected a model in which each predictor variable has high significance considering interaction terms, and multi-collinearity by variance inflation factor by the function '*vif*' in R package '*faraway*' (Faraway, 2011). The final set of predictor variables obtained through the variable selection is given in Table 2.5.

Even though there is no common variable selection techniques for every modeling method in this study, one might wonder if the 'best subset' based on linear modeling method favors linear models over imputation methods. To investigate this issue, another variable selection procedure by the functions '*varSelection*' and '*bestVars*' in R package '*yaImpute*' (Crookston and Finley, 2008) was also implemented as an imputation variable selection technique. The final set of predictor variables via the imputation-based variable selection is in Table 2.6.

Prediction types

We assume that the sampled ground data is the population of the response variable, and the population consists of two parts that are at the i^{th} and the j^{th} locations, respectively ($i \neq j$). One is the observed set (denoted as **O**) at the *i*th location with *n* elements, and the other is unobserved set (denoted as **U**) at the *j*th location with *m* elements. In the observed set, all response and predictor variables are observed. In the unobserved set, predictor variables are observed, but response variable is not.

There are two types of prediction for response variable in this study. The first is point prediction at the *j*th location in the unobserved set. The second is total prediction, which is the sum of true values of the response variable at the *i*th location in the observed set and the predicted values of the response variable at the *j*th location in the unobserved set. That is, the total prediction is for the whole population set. The total prediction, $\hat{T}$, can be mathematically expressed as,

$$\hat{T} = \sum_{i \in \boldsymbol{O}} y_i + \sum_{j \in \boldsymbol{U}} \hat{y}_j \tag{2.1}$$

where y_i is true value of the response variable, and $\hat{y}_i$is the predicted value (using one of the modeling methods that were described earlier) of the response variable.

Performance Measure (PM)

Simulation by cross-validations with some different numbers of training data (*n*) was used to assess the performance of the prediction methods for the two prediction types and the effect of *n* using the predictor variables via 'best subset' variable selection. The whole dataset (*N*=891) was randomly split into two groups 500 times for each number of training data. One group was the observed set (*n*=30, 50, 100, 150, 200, 250, 300, 350, 400, 446, 500) of plot observations as the training (reference) data. The other group was the unobserved set (*m*=861, 841, 791, 741, 691, 641, 591, 541, 491, 445, 397) of plot observations as the validation (target) data. Simulation by 2-fold cross-validation (*m*=446 and *n*=445) using the predictor variables via imputation variable selection was also implemented to examine if 'best subset' variable selection favors the linear modeling methods.

For each simulation run, we computed several performance measures of the prediction methods by response variables and the two prediction types based on Ver Hoef and Temesgen (2013). The performance measures are:

- o Percentage bias for point prediction is calculated by

$$\%Bias_P(\%) = \frac{\left[\frac{1}{mR}\sum_{r=1}^{R}\sum_{j=1}^{m}\left(\hat{y}_{j,r} - y_{j,r}\right)\right]}{\bar{y}} \times 100 \tag{2.2}$$

for point prediction, where *R* is the number of simulations, *m* is the number of point predictions for the r^{th} simulation, $\hat{y}_{j,r}$ is the prediction at the j^{th} location for the r^{th} simulation, $y_{j,r}$ is the true value at the j^{th} location for the r^{th} simulation, and $\bar{y}$ is the true average value of a particular response variable. And percentage bias of total prediction is

$$\%Bias_T(\%) = \frac{\left[\frac{1}{R}\sum_{r=1}^{R}\left(\hat{T}_r - T\right)\right]}{T} \times 100 \tag{2.3}$$

where $\hat{T}_r$ is the predicted total value for the r^{th} simulation, and *T* is the true total value. We performed a two-sided t-test to examine if the estimated bias of the modeling methods was significantly different from zero reporting the p-values.

- RMSPE measures the difference between true values and predicted values. For point prediction, the RMSPE is

$$RMSPE_P = \frac{1}{R}\sum_{r=1}^{R}\left(\sqrt{\frac{1}{m}\sum_{j=1}^{m}\left(\hat{y}_{j,r} - y_{j,r}\right)^2}\right) \tag{2.4}$$

RMSPE for total prediction is calculated as

$$RMSPE_T = \sqrt{\frac{1}{R}\sum_{j=1}^{R}\left(\hat{T}_r - T\right)^2} \tag{2.5}$$

- 90% Prediction Interval Coverage (PIC90) measures how frequently a prediction interval with a 0.9 probability contains a true value. PIC90 for point prediction is as

$$PIC90_P(\%) = \frac{1}{mR}\sum_{r=1}^{R}\sum_{j=1}^{m} I\left(\left(\hat{y}_{j,r} - 1.645\widehat{se}\left(\hat{y}_{j,r}\right)\right) < y_{j,r} \;\&\; y_{j,r} < \left(\hat{y}_{j,r} + 1.645\widehat{se}\left(\hat{y}_{j,r}\right)\right)\right) \times 100 \tag{2.6}$$

where $I(x) = \begin{cases} 1, \text{if x is true} \\ 0, \text{otherwise} \end{cases}$, and $\widehat{se}(\hat{y}_{j,r})$ is the estimated standard error of the predicted value. For total prediction,

$$PIC90_T(\%) = \frac{1}{R}\sum_{r=1}^{R} I\left(\left(\hat{T}_r - 1.645\widehat{se}(\hat{T}_r)\right) < T \;\&\; T < \left(\hat{T}_r + 1.645\widehat{se}(\hat{T}_r)\right)\right) \times 100 \tag{2.7}$$

where $\widehat{se}(\hat{T}_r)$ is the estimated standard error of the total prediction. If these prediction intervals are correctly estimated, then PIC90 should be close to 0.90. Note that we follow Ver Hoef and Temesgen (2013) to estimate standard errors of *k*-NN methods for both point and total prediction.

Variance explained measures such as R^2 is commonly used to evaluate model performance but they are not chosen in this study. Because the error components of NN method of imputation are different from the ones from regression (Stage and Crookston, 2007), it might not be an appropriate way to compare the selected modeling methods using variance explained measures.

We also performed simulation 500 times again with 2-fold cross validation (in the case of *n*= 446 and *m*=445) using all the pre-selected metrics (totaling 15 metrics) listed in Table 2.4, using only the RF *k*-NN methods to see whether the poor performance comes from the small number of predictor variables.

Measures of spatial autocorrelation

Moran's I is one of the common indices for quantifying spatial autocorrelation. This study adopted the generalized one with modified weights suggested by Cliff and Ord (1981). The generalized Moran's I is

$$I = \frac{n\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}(x_i - \bar{x})(x_j - \bar{x})}{\sum_{i=1}^{n}\sum_{j=1}^{n} w_{ij}\sum_i (x_i - \bar{x})^2} \tag{2.8}$$

where $x_i$and x_j are the values of a response variable on locations *i* and *j*, respectively, $\bar{x}$ is its mean, and w_{ij} is the inverse Euclidean distance weight between locations *i* and *j*. The function '*Moran.I*' in R package '*ape*' (Paradis et al., 2004) was used to compute Moran's I.

An alternative measure for spatial autocorrelation is a semivariogram. A semivariogram is defined as an half of the variance of the difference between observations from two locations, which is expressed as

$$\gamma(\mathbf{h}) = \frac{1}{2}\, Var[Z(\mathbf{s}+\mathbf{h}) - Z(\mathbf{s})] \tag{2.9}$$

where $\gamma(\mathbf{h})$ is the semivariance, h is a separation distance, $Z(\mathbf{s}+\mathbf{h})$and $Z(\mathbf{s})$ are the observed values at locations $\mathbf{s}+\mathbf{h}$ and $\mathbf{s}$, respectively. The theoretical semivariogam is estimated by empirical semivariogram using sampled data:

$$\hat{\gamma}(\mathbf{h}) = \frac{1}{2|N(\mathbf{h})|}\sum_{N(\mathbf{h})}[Z(\mathbf{s}+\mathbf{h}) - Z(\mathbf{s})]^2 \tag{2.10}$$

where $\hat{\gamma}(\mathbf{h})$ is the estimated semivariance, $N(\mathbf{h})$ is the set of the distinct sample pairs lagged by $\mathbf{h}$, and $|N(\mathbf{h})|$ is the number of elements in $N(\mathbf{h})$. While Moran's I shows the presence of spatial autocorrelation in data, a semivariogram characterizes the structure of spatial autocorrelation through three parameters, nugget, sill and range. Nugget effect occurs if there are micro-scale variabilities in data that cannot be captured due to large minimum sampling interval or measurement errors. Sill represents the semivariance value that semivariogram reaches at the range. Range indicates the distance at which data are no longer autocorrelated (Cressie, 1993). Semivariogram analysis was done by the function '*variogram*' and '*fit.variogram*' in R package '*gstat*' (Pebesma, 2004).

Results

Relationships Between Responses and Predictors

891 plots were measured by filed crews and airborne LiDAR. Plots' LOR for those plots ranged from 6.71m to 76.05 m, and plots' BA varied between 0.3 m^2/ha and 214.4 m^2/ha. Generally, every response variable except DEN had at least one predictor variable with high correlation coefficient to the corresponding response. Table 2.7 shows the correlation coefficients between each response and corresponding selected predictor variables. The higher correlation coefficient between response and predictor variables, the better prediction performance. The height mean showed high correlation

coefficients 0.787, 0.843 and 0.857 with BA, VOL, and LOR, respectively. The height 75^{th} percentile had a correlation coefficient of 0.803 with QMD. The highest coefficient between the selected predictor variables and DEN, 0.343, came from the percentage of all returns above height of 2 m. In addition to correlation, the scatter plots between responses and the corresponding predictors are in Figure 2.3. All the response variables except DEN show the relationships close to linearity with 'height mean' or 'height 75^{th} percentile.

Performance Measures

Table 2.8 shows the performance measures in 2-fold cross-validation with the predictor variables via 'best subset' variable selection. There are 10 cases that come from the combination of 5 response variables (BA, VOL, LOR, QMD, and DEN) and 2 prediction types (point and total). In most cases, BestNN, OLS, SLM, and GWR showed better performances than the other model types in terms of RMSPE and bias, and were comparable among each other. For *k*-NN methods, *k*=5 neighbors had better performance than with *k*=1 based on RMSPE.

For BA, BestNN showed the best RMSPE in both point and total prediction. BestNN reduced point RMSPE by 28.2% over MSN1 and GNN1 respectively. The amount of decrease in total RMSPE by BestNN was 26.5% over MSN1 (GNN1). For VOL, SLM and GWR were the best in point RMSPE. SLM and GWR decreased RMSPE by 27.7% over MSN1 (GNN1), respectively. OLS provided the lowest RMSPE in total prediction for VOL. The amount was 22.8% that OLS decreased over MSN1 (GNN1) in total RMSPE for VOL. For LOR, SLM provided the lowest RMSPE in point prediction. It reduced RMSPE by 28.7% over MSN1 (GNN1). In total RMSPE for LOR, OLS was the best method with the lowest RMSPE decreasing 23.5% of total RMSPE over MSN1 (GNN1). GWR, the best one in QMD point RMSPE, decreased RMPSE by 27.8% over MSN1 (GNN1). And SLM produced the best RMSPE in QMD total RMPSE, which was improved by 27.3% over RF1. For DEN, SLM showed the best RMPSE in both point and total prediction. The amounts of reduced RMSPE by SLM were 29.1% over MSN1 (GNN1) and 20.2% over RF1 for point and total prediction, respectively. Of the non-leading group, MSN1, RF1, and GNN1 showed poorer performances with larger RMPSEs for both point and total prediction. The remaining methods of this group, MSN5, RF5, and GNN5, performed better than k-NN methods with *k*=1, but not as good as the leading group for most cases.

Underestimation was found in most cases giving negative bias. RF5, SLM, and OLS always underestimated the forest attributes of interest. RF1, GWR, and BestNN had only one overestimation for VOL, QMD, and LOR, respectively. In contrast, MSN1 and GNN1 showed overestimation for BA, VOL, and DEN. The p-values of the two sided t-test for testing the unbiasedness of the biases are under the bias values in the parentheses in Table 2.8. Several cases had p-values less than 0.05, which means that they are statistically different from zero. For example, RF5 and BestNN were biased for every case. And RF1 and GWR showed many biased cases. Note that BestNN and GWR showed larger biases in the most cases even though they are in the leading group. OLS appeared biased only in DEN total prediction. However, the magnitudes of the biases were small compared to the mean (or total) of response variables. The percentages of biases to the mean (or total) of response variables were at most 1.2% in modulus.

Most methods showed valid PIC90 that are close to the nominal level 0.9. However, RF1 and RF5 had lower PIC90s for both point and total predictions; those PIC90 were generally around or less than 0.85. The PIC90s of MSN1 and GNN1 showed values of more than 0.93 for LOR, QMD, and DEN total prediction. Note that RF5 showed the smallest prediction standard errors for every response variable from the 500 simulations (Figure 2.4). The prediction standard errors of RF1 were less than MSN1 and GNN1. RF5 had comparable RMSPE to MSN5 and GNN5, but relatively more biased than other methods in point and total prediction. The larger bias and smaller prediction standard error might lead to the poor PIC90 performance for RF5. But RF1's poor PIC90 cannot be explained by the larger bias and smaller prediction standard error for every case. For example, BestNN showed PIC90s close to 0.9 for VOL and LOR with larger bias than RF1 and comparable prediction standard error.

No single modeling technique worked best for all the cases by forest attributes of interest, prediction type, and performance measures. Though there was a leading group with better performances than others, no one in the group is always the best for each case. Nevertheless, it is worthwhile to report that SLM and OLS produced the best result in RMSPE in both point and total prediction.

Influence of Variable Selection Technique on Model Performance

The performance measures of imputation methods with the predictor variables via imputation variable selection is shown in Table 2.9. Compared with the results in Table 2.8, imputation variable selection did not improve the performance measures of imputation methods in terms of RMSPE and bias for many cases. Some remarkable changes were 1) improvement in bias of BestNN except for DEN, 2) improvement in RMSPE of every imputation method for VOL prediction (only even for BestNN in VOL point prediction), and 3) decline in PIC90 of RF imputation methods (only except for DEN point prediction). But none of the improvements in terms of RMSPE and bias gave imputation methods superiority to the linear modeling methods.

Influence of Number of Training Data on Model Performance

Figure 2.5 shows RMSPE in point and total prediction by response variables and modeling methods with different numbers of training data. Note that BestNN did not produce predictions with 30 training data, and GWR also did not produce predictions with n = 30 and 50 for every response and n = 100 for QMD. For point prediction, BestNN, SLM, OLS and GWR consisted of a leading group with most numbers of training data. For total prediction, SLM and OLS were also in the leading group in RMSPE while BestNN and GWR sometimes were replaced by MSN5, GNN5 or RF5. Generally, RMSPE in point prediction got stabilized with 100 - 150 training data. In contrast, RMSPE in total prediction consistently decreased as n increased. However, rate of the decrease for RMSPE in total prediction also slowed down, and it might be considered as approximately stabilized when n is between 200 and 250.

Percentage bias with different n is given in Figure 2.6. Increasing n in point prediction does not necessarily accompany advancement of bias while increasing n in total prediction generally brought improvement of bias. Influence of n on PIC90 is shown in Figure 2.7. For point prediction, PIC90 for every modeling method appeared to be stabilized with more than 100 or 150 training data. Regardless of the number of training data, RF imputation methods consistently had lower PIC90 than the nominal level 90%. The trend of PIC90 in total prediction looked more variable than in point prediction. As with the case in point prediction, RF imputation methods showed generally lower PIC90 in total prediction than 90% or other methods' PIC90 in total prediction.

Range of Prediction

Boxplots of prediction values from 500 simulations and observations as reference are shown in Figure 2.8. The two horizontal dotted lines in Figure 2.8 represent the first and third quartiles of observations of response variables. While the linear modeling methods such as SLM, OLS, and GWR mostly outperformed the other methods except BestNN in global performance measures, the linear modeling methods produced some negative predictions at the locations with smaller observation values, which do not make sense for forest inventory attributes. Those negative predictions were found in BA, VOL, QMD (only in GWR), and DEN.

The Low PIC90 of Random Forest Based Imputation

A prediction interval consists of a point estimate (prediction value) and a margin of error (critical value × prediction standard error). The closer to the true value the prediction value gets, the more likely the prediction interval covers the true value. And the wider prediction interval width we have, the more likely the prediction interval covers true value. So if we have smaller biases in modulus and larger prediction standard errors, we would have higher prediction interval coverage. However, as mentioned earlier in the section 4.2, the bias and prediction standard error were not enough to explain such lower PIC90s of RFs for every case.

Influence of Number of Predictors in Random Forest Based Imputation

RF-imputation has been reported to handle a large number of predictor variables that are highly correlated each other. Table 2.10 presents the results of the re-simulation. We found that using all the metrics does not guarantee improvement of performance measures. Particularly, PIC90 in every case greatly decreased with the correlated metrics. Even though there were some improvements in performance, RF *k*-NN methods with all of the metrics was not ranked best in any performance measures. From our re-simulation, it does not seem that the poor performances of RF imputation come from the number of predictor variables in our study.

Autocorrelation in Response Variables

The semivariograms for each response variable are given in Figure 2.9. From the four fitted models on the semivariograms, the five response variables have more than 80% relative nugget effects (nugget effect / sill). The relative nugget effects for each response variable were as follows: BA (89.1%), VOL (93.6%), LOR (82.6%), QMD (87.2%), DEN (89.4%). These high relative nugget effects indicate high individual variation (not related to spatial dependence) in the response variables, or there is a large minimum sampling interval that cannot capture enough spatial dependence.

Table 2.11 includes statistics for Moran correlation coefficient (Moran's I) for each response variable and each selected predictor variable from all data. According to the Moran's I, all the response and predictor variables displayed significant spatial clustering. The Moran's I of response variables in training data and validation data from the 500 simulations is shown in Table 2.12. The training and validation data also show spatial autocorrelation for most of the simulations.

Figure 2.10 shows the distributions of Z-scores of Moran's I from prediction errors by modeling methods. The horizontal lines are on the values of 1.96 and -1.96, respectively. If a Z-score is outside of the interval between -1.96 and 1.96, then it is considered as a case of autocorrelation. The autocorrelation in prediction errors was observed in a small number of cases for most modeling methods except for BA and DEN. In general, k-NN methods with k=1 had smaller number of autocorrelation cases than the other methods. Only for DEN, SLM generated fairly less cases of autocorrelation.

Discussion

Performance Measures

Several previous studies asserted that one modeling method showed superior performance for all the cases they investigated. SLM was always the best method in RMSPE for both artificial data and real forest data in Ver Hoef and Temesgen (2013). Temesgen and Ver Hoef (2015) found that SLM produced the smallest RMSPE for artificial and forestry dataset except lognormal dataset. RF imputation was the best for basal area and tree density estimation in root mean square distance (RMSD) in Hudak et al. (2008), and was the best for nine forest attributes in RMSD in Hudak et al. (2014). Eskelson et al. (2009b) indicated that RF imputation showed the best results in terms of RMSE for both mean annual change in basal area, stems per ha, volume, and biomass and those current forest attributes using climate, topography and satellite data. In our study, however, no single method was found to be superior. This result matches up with the findings from Brosofske et al. (2014) that each modeling method has its own strengths and weaknesses, and there is no best one for all cases. The best modeling method varied by response variable, performance measure and prediction type. Pierce et al. (2009) indicated that the best modeling method among GNN, OLS, kriging, and universal kriging (mathematically the same as SLM) for vegetation and fuel variables varied by response variables and regions where the models were applied. Corona et al. (2014) concluded that spatial autocorrelation in response variables and correlation between response variables and auxiliary variables affects the performance of a modeling method.

We found that BestNN and RF5 consistently showed biasedness, and RF1 and GWR had many biased cases as well. Ver Hoef and Temesgen (2013) claimed that, in the resampling of real forest data, there appeared to be some bias for total prediction by *k*-NN methods with Mahalanobis distance (*k*=1 or *k*=5) and BestNN, except with spatially balanced sampling. BestNN also appeared biased in total prediction for artificially generated Poisson and lognormal data as well as forest productivity and biomass data (Temesgen and Ver Hoef, 2015). In Eskelson et al. (2009b), RF imputation was biased for prediction of mean annual change in basal area, volume and biomass using climate, topography and satellite data. Bias was also found from RF imputation for current basal area, volume and biomass estimation with the same dataset (Gagliasso et al., 2014). As stated earlier, RF imputation consistently produced larger biases than other *k*-NN methods (Breidenbach et al., 2010a).

From our study, *k*-NN methods with *k*=5 had lower RMSPE than *k*-NN methods with *k*=1. Similar results have been reported by several other studies. Muinonen et al. (2001) reported that MSN had lower RMSE% while *k* increased up to 5 for plot-level tree volume estimation. In Ver Hoef and Temesgen (2013), MSN5 always had lower RMSPE than MSN1 for forest data, and for artificial data except for count data in total prediction. In Temesgen and Ver Hoef (2015), GNN5 had smaller RMSPE than GNN1 for artificial and forest dataset. MSN with *k*=5 had the smallest RMSE for biomass and basal area estimation followed by MSN with *k*=3 and MSN in Gagliasso et al. (2014). In terms of bias, however, *k*-NN methods with larger *k* does not always provide less bias than *k*-NN methods with smaller *k*. As with our study, Temesgen and Ver Hoef (2015) and Ver Hoef and Temesgen (2013) reported that the effect of *k* to bias of *k*-NN varied by response variable or dataset used. MSN with *k*=3 showed the lowest bias for biomass and basal area estimation compared to MSN with *k*=5 and MSN (Gagliasso et al., 2014).

k-NN methods provided predictions within the biologically reasonable bounds, not giving negative values of forest attributes examined while the linear modeling methods such as SLM, OLS, and GWR had some negative predictions. This is one of the advantages of *k*-NN methods (Eskelson et al., 2009a). And *k*-NN methods, particularly with *k*=1, had higher upper ranges of predictions than SLM and OLS except for LOR as shown in Figure 2.9. This indicated that linear methods smoothed the highest predictions. Due to this smoothing, the predictions for the extreme high values by the linear models were not able to be very accurate. On the other hand, GWR, one of the linear modeling methods, had higher upper ranges of predictions for BA, VOL and DEN, and also had smaller lower ranges of predictions for BA, VOL, QMD and DEN. Cho et al. (2009) reported that GWR tends to produce extreme local regression coefficients with small number of training data. These extremes regression coefficients might cause the large prediction ranges of GWR.

Contrary to some previous studies indicating advantages or superiority of RF imputation for estimating forest attributes (Hudak et al., 2008; Hudak et al., 2014; Latifi et al., 2010), RF imputation had consistently lower PIC90s than the nominal level 0.9 in our study. Temesgen and Ver Hoef (2015) also reported poor prediction interval coverage of RF imputation, and asserted that RF's high bias is one of the reasons for the poor performance. As we saw, RF showed biasedness in our study as well. But it is not enough to explain RF's poor PIC90 with only bias because the modeling methods with larger bias such as BestNN or GWR in VOL and LOR prediction had PIC90s close to the nominal level. The lowest prediction standard error of RF5, representing narrower width of a prediction interval, could be one of the reasons for the poor PIC90 performance. However, it is not enough to explain RF1's poor PIC90 with bias and prediction standard error.

The linear modeling methods generally showed superiority though they had the deficiency of producing negative predictions unrealistic for forest attributes. But the relatively poor performance of *k*-NN methods might be partially from the 2-fold cross-validation used in our simulation. In order to improve accuracy of imputation, firstly, predictor variables in the randomly sampled training data in a simulation should cover the entire joint ranges of predictor variables, and secondly, response variable in the training data in a simulation should represent all values of the response variable (Eskelson et al., 2009a). However, randomly splitting the dataset into two sets for every simulation run would not always guarantee the condition mentioned above for every simulation run. The performance of GWR might be influenced as well due to our simulation approach. If the random sample does not represent spatial nonstationarity, then GWR would not show its advantage. Temesgen and Ver Hoef (2015) concluded that a sample's representativeness and the closeness between training and validation datasets affect the accuracy of modeling methods they examined.

Transformations of independent variables by square, cube, square root, or exponent could improve performance of the OLS method though the transformations were not applied to our analysis. RMSE of OLSs from the 15^{th} simulated training and validation dataset for BA, VOL and DEN were improved by from 0.45% to 1.52%. These transformation might offer some r improvement for some of the other modeling methods.

Influence of Variable Selection Technique on Model Performance

Other than the imputation variable selection method proposed by Crookston and Finley (2008) used in this study, there are several approaches to select predictor variables for NN imputation method. An algorithm to minimize relative root mean square error for NN imputation using transformation and stepwise optimization was proposed by Maltamo et al. (2006). Hudak et al. (2008) repeated RF to discard the least important predictor variable, which is similar to backward stepwise variable selection in multiple regression model. Latifi et al. (2010) claimed that the prediction performance depended on the combinations of NN imputation methods and variable selection procedure. With NN imputation using Euclidean and Mahalanobis distance metrics, genetic-algorithm based variable selection produced better prediction performance. By contrast, full dataset, i.e. without variable selection, outperformed genetic-algorithm based variable selection when using MSN and RF imputation. The backwards stepwise variable selection gave the least performance compared to genetic-algorithm based variable selection and full dataset. Garcia-Gutierrez et al. (2014) compared three variable selection techniques (stepwise, best subset and genetic-algorithm selections) for estimating some forest stand variables using LiDAR. Genetic-algorithm was reported to perform better than other techniques based on BIC of three regression models. Packalén et al. (2012) concluded that variable selection is an essential part of NN imputation. An algorithm using optimization to minimize relative RMSE was the best variable selection strategy compared with full dataset, another using canonical correlation analysis and the other using RF importance. But the full dataset surpassed the algorithm with canonical correlation analysis and the one using RF importance.

The results comparing the two variable selection techniques in present study showed that ‘best subset’ variable selection generally gave better performance than the imputation variable selection proposed by Crookston and Finley (2008). However, our results cannot be generalized compared with the previous researches because the combination of modeling methods and variable selection techniques in our study were different from others’ one. It would be more meaningful and practical to determine which combination of variable selection technique, and size of *n* provides the best performance for NN imputation, because identifying the effective sample size of filed plots is an important issue in LiDAR-based forest inventory. In the point of this view, our comparison is not enough to determine the optimal combination because the comparison was implemented only using two variable selection techniques and one training data size of 446. This should be further studied in the future.

Influence of Number of Training Data on Model Performance

Junttila et al. (2015) proposed an alternative model to overcome overfitting and multicollinearity when using small number of filed sample plots for LiDAR-based forest inventory. To that end, the effects of number of predictor variables and number of training data were investigated along with different field plot sampling designs. Junttila et al. (2015) reported that 50 field sample plots yielded only 5-15% larger relative error than several hundred field sample plots with their proposed model. Maltamo et al. (2011) examined several plot selection strategies for field training data in LiDAR-assisted forest inventory with different field training data sizes. More than 150 training field plots gave similar prediction accuracy regardless of plot selection methods. Maltamo et al. (2011) concluded that the number of field training plots can be less than 100 plots if those training plots capture the population's variation.

RMSPEs for every modeling method from smaller *n* had slightly higher values than the ones from lager *n* in our study. This trend was clearer in total prediction than in point prediction. The minimum *n* with relatively good accuracy in terms of RMSPE might be around 100-150 for point prediction, and might be around 200-250 for total prediction based on our results. Thus, to determine the effective *n* in terms of RMSPE, it would be necessary to consider point prediction and total prediction separately. Because the field plots were sampled by LiDAR-assisted stratified random sampling proposed by Hawbaker et al. (2009), training data of relatively smaller *n* from most of simulation runs might keep the variability of population of interest. That might be why those smaller *n*, larger than about 100 in point prediction and lager than about 200 in total prediction, gave fairly good RMSPE compared to RMSPE with more than 400 training data. Though percentage biases varied along the numbers of training data, their absolute values for every modeling method were less than 1.5 % except with the smaller $n = 30$ or 50. In terms of *n*, bias in modeling forest inventory using LiDAR data might be avoided if one use $n \geq 100$. The effect of *n* to PIC90 was less than the one to other performance measures. With only the smallest $n = 30$, PIC90 for most modeling methods had lower PIC90s than the ones at lager *n*. The factor that mainly affected PIC90 was modeling method. Regardless of *n*, PIC90 of RF imputation methods showed lower PIC90 in most prediction cases. Although PIC90 in total prediction looked more variable, most modeling methods were close to the nominal level 90%.

According to Chen et al. (2011), GWR tended to give poorer performance in lower field data sampling density than in larger field data sampling density for estimating forest canopy height. But in our study, that trend in GWR's performance was not able to be checked because GWR did not provide predictions with $n = 30$ and 50 for every response and $n = 100$ for QMD. Instead of mentioning

GWR's performance, we could say about the sensitivity of GWR itself to a small number of training data. This might be related to sparse training data in a regression point to fit a local model, but the exact reason for the GWR's sensitivity was not considered here.

According to Junttila et al. (2015), the number of predictor variables affected a model's performance depending on both how much variability of response variable is captured by predictor variables and how much overfitting arises in a model by predictor variables. Relative RMSPE increased with either a small number of predictors or a too large number of predictors, especially in a small number of training data. However, because we limited the maximum number of predictors in a model as three, our study might lack evaluations at least for the models with larger number of predictor variables.

Autocorrelation in response variables

Although SLM was a member of the leading group with good performance measures, it was not best predictive method. The ranges of the variograms in Figure 2.9 are too short compared to the size of study area, except for LOR. This is a potential weakness of SLM, as studies use much fewer than 891 field observation plots. As seen in Table 2.11, all the response and predictor variables had spatial autocorrelations of similar magnitudes. This also might decrease the benefit from using spatial predictor such as SLM handling autocorrelation.

Figure 2.10 shows that autocorrelations were well captured by most modeling methods for VOL, LOR and QMD. Interestingly, *k*-NN with *k*=1 generally showed better property on autocorrelation in prediction errors than the other modeling methods even though it does not consider spatial structure of given data. From this result, we would not say that the better performances of SLM for BA, VOL, LOR and QMD come from dealing with autocorrelation. Under strong correlation between response and predictors such as those responses, the merit of handling autocorrelation might diminish. Only for DEN, SLM was much less autocorrelation in prediction errors than other methods. In DEN prediction, the correlations between response and predictor variables were lower than other response variables. The situation where correlations between response and predictors are weak might give a merit to handling spatial autocorrelation. Although the performance of the imputation methods might be decreased due to the simulation approach that we adopted, SLM might have applicability to wider conditions in that it did provide better performance for most cases in present study.

Influence of number of predictors in Random Forest based imputation

In the research of Latifi and Koch (2012), RF imputation was more accurate with the full dataset than the selected dataset by the genetic algorithm. Penner et al. (2013) claimed that the RF *k*-NN method takes full advantage of the dataset, which means that RF *k*-NN methods can include highly correlated predictor variables. But Hastie et al. (2009) pointed out that if the number of predictors is large and the fraction of important variables small, then Random Forest might perform poorly with small number of predictors for splitting a node. In Vauhkonen et al. (2010), it was reported that including more predictor variables did not improve model's performance. RF *k*-NN methods with 1846 predictor variables produced more relative bias than RF with 130 and 24 predictors. Breidenbach et al. (2010a) asserted that RF *k*-NN methods with the most important variables were slightly better than the models with all predictor variables. Murphy et al. (2010) also found that using metrics correlated one another in RF regression decreased explanatory power. Our result partially supports the latter as shown in the result section. Generally, the improvements were not sizable, and the decreases in PIC90 were substantial.

Conclusion

From our study, no modeling technique was always superior to the others for prediction of selected forest attributes. This result agrees with Brosofske et al. (2014). Even though there was a leading group (BestNN, SLM, OLS, and GWR) that always outperformed the other methods in RMSPE for both point and total prediction, the best method varied according to response variables, prediction types, and performance measures. The effective size of training data depended on the prediction type. About 100-150 training data showed comparable performance in point prediction, whereas about 200-250 training data showed comparable performance in total prediction. Therefore, selecting a modeling technique for forest attributes should be carefully determined based on the objectives, conditions and scales at which that researchers or forest managers face.

OLS appeared to have very good performances. BestNN produced comparable performances to the linear models. Despite of its biasedness, as a member of *k*-NN techniques, BestNN has the advantage that its prediction values are within the biologically reasonable bounds compared to the linear models. SLM showed its potential to estimate forest attributes for broader conditions in our study. As a more generalized approach than OLS, it could have good performance in various situations in terms of diverse combinations of relationships between responses and predictors. GWR also produced better performance but showed sensitivity with a small number of training data.

RF imputation did not perform well, particularly in PIC90, compared to other methods. From the previous studies, bias has been reported as one of the reasons for the poor performance of RF imputation technique. In addition to bias, we observed that smaller prediction standard error also could impact PIC90. And the smaller number of predictor variables was not the reason for the RF's poor PIC90.

References

Bivand, R., Bivand, M.R., Brunsdon, M.C., and Fortheringham, S. 2013. *spgwr: Geographically weighted regression*. R package version 0.6-24.

Breidenbach, J., Nothdurft, A., and Kändler, G. 2010a. Comparison of nearest neighbour approaches for small area estimation of tree species-specific forest inventory attributes in central Europe using airborne laser scanner data. *European Journal of Forest Research* 129(5): 833-846. doi: 10.1007/s10342-010-0384-1.

Breiman, L. 2001. Random forests. *Machine Learning* 45(1): 5-32. doi: 10.1023/a:1010933404324.

Brosofske, K.D., Froese, R.E., Falkowski, M.J., and Banskota, A. 2014. A Review of Methods for Mapping and Prediction of Inventory Attributes for Operational Forest Management. *Forest Science* 60(4): 733-756. doi: 10.5849/forsci.12-134.

Brunsdon, C., Fotheringham, S., and Charlton, M. 1998. Geographically weighted regression. *Journal of the Royal Statistical Society: Series D (The Statistician)* 47(3): 431-443. doi: 10.1111/1467-9884.00145.

Chen, G., Zhao, K., McDermid, G.J., and Hay, G.J. 2011. The influence of sampling density on geographically weighted regression: a case study using forest canopy height and optical data. *International Journal of Remote Sensing* 33(9): 2909-2924. doi: 10.1080/01431161.2011.624130.

Cho, S., Lambert, D.M., Kim, S.G., and Jung, S. 2009. Extreme coefficients in Geographically Weighted Regression and their effects on mapping. In 2009 Annual Meeting. Agricultural and Applied Economics Association, Milwaukee, Wisconsin.

Cliff, A.D., and Ord, J.K. 1981. *Spatial processes: models & applications*. Pion, London.

Corona, P., Fattorini, L., Franceschi, S., Chirici, G., Maselli, F., and Secondi, L. 2014. Mapping by spatial predictors exploiting remotely sensed and ground data: A comparative design-based perspective. *Remote Sensing of Environment* 152(0): 29-37. doi: http://dx.doi.org/10.1016/j.rse.2014.05.011.

Cressie, N. 1993. *Statistics for Spatial Data* (rev. ed.). Wiley, New York, NY, USA. pp. 69-70.

Crookston, N.L., and Finley, A.O. 2008. yaImpute: An R package for kNN imputation. *Journal of Statistical Software* 23(10): 16.

Eskelson, B.N., Temesgen, H., Lemay, V., Barrett, T.M., Crookston, N.L., and Hudak, A.T. 2009a. The roles of nearest neighbor methods in imputing missing data in forest inventory and monitoring databases. *Scandinavian Journal of Forest Research* 24(3): 235-246.

Eskelson, B.N.I., Barrett, T.M., and Temesgen, H. 2009b. Imputing Mean Annual Change to Estimate Current Forest Attributes. *Silva Fennica* 43(4): 649-658. doi: 10.14214/sf.185.

Falkowski, M.J., Hudak, A.T., Crookston, N.L., Gessler, P.E., Uebler, E.H., and Smith, A.M.S. 2010. Landscape-scale parameterization of a tree-level forest growth model: a k-nearest neighbor imputation approach incorporating LiDAR data. *Canadian Journal of Forest Research* 40(2): 184-199. doi: 10.1139/X09-183.

Faraway, J. 2011. *faraway: Functions and datasets for books by Julian Faraway*. R package version 1.0.5.

Gagliasso, D., Hummel, S., and Temesgen, H. 2014. A comparison of selected parametric and non-parametric imputation methods for estimating forest biomass and basal area. *Open Journal of Forestry* 4(01): 42.

Garcia-Gutierrez, J., Gonzalez-Ferreiro, E., Riquelme-Santos, J.C., Miranda, D., Dieguez-Aranda, U., and Navarro-Cerrillo, R.M. 2014. Evolutionary feature selection to estimate forest stand variables using LiDAR. International Journal of Applied Earth Observation and Geoinformation 26: 119-131. doi: http://dx.doi.org/10.1016/j.jag.2013.06.005.

Gobakken, T., Næsset, E., Nelson, R., Bollandsås, O.M., Gregoire, T.G., Ståhl, G., Holm, S., Ørka, H.O., and Astrup, R. 2012. Estimating biomass in Hedmark County, Norway using national forest inventory field plots and airborne laser scanning. *Remote Sensing of Environment* 123(0): 443-456. doi: http://dx.doi.org/10.1016/j.rse.2012.01.025.

Hastie, T., Tibshirani, R., Friedman, J., Hastie, T., Friedman, J., and Tibshirani, R. 2009. *The elements of statistical learning* (2nd ed.). Springer-Verlag New York. pp. 745.

Hawbaker, T.J., Keuler, N.S., Lesak, A.A., Gobakken, T., Contrucci, K., and Radeloff, V.C. 2009. Improved estimates of forest vegetation structure and biomass with a LiDAR-optimized sampling design. *Journal of Geophysical Research: Biogeosciences* 114(G2): G00E04. doi: 10.1029/2008JG000870.

Hoeting, J.A. 2009. The importance of accounting for spatial and temporal correlation in analyses of ecological data. *Ecological Applications* 19(3): 574-577. doi: 10.1890/08-0836.1.

Holopainen, M., and Kalliovirta, M.S.J. 2006. Modern data acquisition for forest inventories. In *Forest Inventory*. Edited by A. Kangas and M. Maltamo. Springer. pp. 343-362.

Hudak, A.T., Crookston, N.L., Evans, J.S., Hall, D.E., and Falkowski, M.J. 2008. Nearest neighbor imputation of species-level, plot-scale forest structure attributes from LiDAR data. *Remote Sensing of Environment* 112(5): 2232-2245. doi: http://dx.doi.org/10.1016/j.rse.2007.10.009.

Hudak, A.T., Haren, A.T., Crookston, N.L., Liebermann, R.J., and Ohmann, J.L. 2014. Imputing forest structure attributes from stand inventory and remotely sensed data in Western Oregon, USA. *Forest Science* 60(2): 253-269. doi: 10.5849/forsci.12-101.

Hudak, A.T., Strand, E.K., Vierling, L.A., Byrne, J.C., Eitel, J.U.H., Martinuzzi, S., and Falkowski, M.J. 2012. Quantifying aboveground forest carbon pools and fluxes from repeat LiDAR surveys. *Remote Sensing of Environment* 123(0): 25-40. doi: http://dx.doi.org/10.1016/j.rse.2012.02.023.

Husch, B., Beers, T.W., and Kershaw Jr, J.A. 2002. *Forest mensuration*. John Wiley & Sons. pp. 176-177.

Junttila, V., Kauranne, T., Finley, A.O., and Bradford, J.B. 2015. Linear Models for Airborne-Laser-Scanning-Based Operational Forest Inventory With Small Field Sample Size and Highly Correlated LiDAR Data. *IEEE Transactions on Geoscience and Remote Sensing* 53(10): 5600-5612. doi: 10.1109/TGRS.2015.2425916.

Kankare, V., Vastaranta, M., Holopainen, M., Räty, M., Yu, X., Hyyppä, J., Hyyppä, H., Alho, P., and Viitala, R. 2013. Retrieval of forest aboveground biomass and stem volume with airborne scanning LiDAR. *Remote Sensing* 5(5): 2257-2274.

Latifi, H., and Koch, B. 2012. Evaluation of most similar neighbour and random forest methods for imputing forest inventory variables using data from target and auxiliary stands. *International Journal of Remote Sensing* 33(21): 6668-6694. doi: 10.1080/01431161.2012.693969.

Latifi, H., Nothdurft, A., and Koch, B. 2010. Non-parametric prediction and mapping of standing timber volume and biomass in a temperate forest: application of multiple optical/LiDAR-derived predictors. *Forestry* 83(4): 395-407.

LeMay, V., Maedel, J., and Coops, N.C. 2008. Estimating stand structural details using nearest neighbor analyses to link ground data, forest cover maps, and Landsat imagery. *Remote Sensing of Environment* 112(5): 2578-2591. doi: http://dx.doi.org/10.1016/j.rse.2007.12.007.

Li, M., Im, J., Quackenbush, L.J., and Liu, T. 2014. Forest biomass and carbon stock quantification using airborne LiDAR data: a case study over Huntington Wildlife Forest in the Adirondack Park. *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of* 7(7): 3143-3156. doi: 10.1109/JSTARS.2014.2304642.

Liaw, A., and Wiener, M. 2002. Classification and Regression by randomForest. *R news* 2(3): 18-22.

Lumley, T. 2009. *leaps: regression subset selection*. R package version 2.9. http://cran.r-project.org/web/packages/leaps/index.html.

Maltamo, M., Bollandsås, O., Næsset, E., Gobakken, T., and Packalén, P. 2011. Different plot selection strategies for field training data in ALS-assisted forest inventory. *Forestry* 84(1): 23-31.

Maltamo, M., Malinen, J., Packalén, P., Suvanto, A., and Kangas, J. 2006. Nonparametric estimation of stem volume using airborne laser scanning, aerial photography, and stand-register data. *Canadian Journal of Forest Research* 36(2): 426-436. doi: 10.1139/x05-246.

McGaughey, R.J. 2010. *FUSION/LDV: Software for LIDAR data analysis and visualization*, Forest service. Pacific Northwest research station, United States department of agriculture.

McRoberts, R.E., Cohen, W.B., Naesset, E., Stehman, S.V., and Tomppo, E.O. 2010. Using remotely sensed data to construct and assess forest attribute maps and related spatial products. *Scandinavian Journal of Forest Research* 25(4): 340-367.

Moeur, M., and Stage, A.R. 1995. Most similar neighbor: an improved sampling inference procedure for natural resource planning. *Forest Science* 41(2): 337-359.

Muinonen, E., Maltamo, M., Hyppänen, H., and Vainikainen, V. 2001. Forest stand characteristics estimation using a most similar neighbor approach and image spatial structure information.

Remote Sensing of Environment 78(3): 223-228. doi: http://dx.doi.org/10.1016/S0034-4257(01)00220-6.

Murphy, M.A., Evans, J.S., and Storfer, A. 2010. Quantifying Bufo boreas connectivity in Yellowstone National Park with landscape genetics. *Ecology* 91(1): 252-261. doi: 10.1890/08-0879.1.

Næsset, E., and Gobakken, T. 2008. Estimation of above- and below-ground biomass across regions of the boreal forest zone using airborne laser. *Remote Sensing of Environment* 112(6): 3079-3090. doi: http://dx.doi.org/10.1016/j.rse.2008.03.004.

Niska, H., Skon, J., Packalen, P., Tokola, T., Maltamo, M., and Kolehmainen, M. 2010. Neural Networks for the Prediction of Species-Specific Plot Volumes Using Airborne Laser Scanning and Aerial Photographs. *IEEE Transactions on Geoscience and Remote Sensing,* 48(3): 1076-1085. doi: 10.1109/TGRS.2009.2029864.

Ohmann, J.L., and Gregory, M.J. 2002. Predictive mapping of forest composition and structure with direct gradient analysis and nearest-neighbor imputation in coastal Oregon, USA. *Canadian Journal of Forest Research* 32(4): 725-741.

Packalén, P., Mehtätalo, L., and Maltamo, M. 2011. ALS-based estimation of plot volume and site index in a eucalyptus plantation with a nonlinear mixed-effect model that accounts for the clone effect. *Annals of Forest Science* 68(6): 1085-1092. doi: 10.1007/s13595-011-0124-9.

Packalén, P., Temesgen, H., and Maltamo, M. 2012. Variable selection strategies for nearest neighbor imputation methods used in remote sensing based forest inventory. *Canadian Journal of Remote Sensing* 38(5): 557-569. doi: 10.5589/m12-046.

Paradis, E., Claude, J., and Strimmer, K. 2004. APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics* 20(2): 289-290. doi: 10.1093/bioinformatics/btg412.

Penner, M., Pitt, D.G., and Woods, M.E. 2013. Parametric vs. nonparametric LiDAR models for operational forest inventory in boreal Ontario. *Canadian Journal of Remote Sensing* 39(5): 426-443. doi: 10.5589/m13-049.

Pierce, K.B., Ohmann, J.L., Wimberly, M.C., Gregory, M.J., and Fried, J.S. 2009. Mapping wildland fuels and forest structure for land management: a comparison of nearest neighbor imputation and other methods. *Canadian Journal of Forest Research* 39(10): 1901-1916. doi: 10.1139/X09-102.

R Core Team. 2013. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

Schabenberger, O., and Gotway, C.A. 2004. *Statistical methods for spatial data analysis*. CRC press.

Stage, A.R., and Crookston, N.L. 2007. Partitioning error components for accuracy-assessment of near-neighbor methods of imputation. *Forest Science* 53(1): 62-72.

Strunk, J.L., Reutebuch, S.E., and Foster, J.R. 2008. LIDAR inventory and monitoring of a complex forest. In ASPRS Annual Conference, Portland, Oregon.

Temesgen, B.H., LeMay, V.M., Froese, K.L., and Marshall, P.L. 2003. Imputing tree-lists from aerial attributes for complex stands of south-eastern British Columbia. *Forest Ecology and Management* 177(1–3): 277-285. doi: http://dx.doi.org/10.1016/S0378-1127(02)00321-3.

Temesgen, H., Barrett, T.M., and Latta, G. 2008. Estimating cavity tree abundance using nearest neighbor imputation methods for western Oregon and Washington forests. *Silva Fennica* 42(3): 337-354. doi: 10.14214/sf.241.

Temesgen, H., and Ver Hoef, J.M. 2015. Evaluation of the spatial linear model, random forest and gradient nearest-neighbour methods for imputing potential productivity and biomass of the Pacific Northwest forests. *Forestry* 88(1): 131-142. doi: 10.1093/forestry/cpu036.

Vauhkonen, J., Korpela, I., Maltamo, M., and Tokola, T. 2010. Imputation of single-tree attributes using airborne laser scanning-based height, intensity, and alpha shape metrics. *Remote Sensing of Environment* 114(6): 1263-1276.

Ver Hoef, J.M., and Temesgen, H. 2013. A comparison of the spatial linear model to nearest neighbor (k-NN) methods for forestry applications. *PloS one* 8(3): e59129.

Wulder, M.A., White, J.C., Nelson, R.F., Næsset, E., Ørka, H.O., Coops, N.C., Hilker, T., Bater, C.W., and Gobakken, T. 2012. Lidar sampling for large-area forest characterization: A review. *Remote Sensing of Environment* 121: 196-209. doi: http://dx.doi.org/10.1016/j.rse.2012.02.001.

Table 2.1: LiDAR survey specifications

Attribute	Description
Sensor	Leica ALS50 Phase II
Survey altitude	900 m (flown at 900 meters above ground level)
Pulse rate	> 105 kHz (> 105,000 laser pulse per second)
Pulse mode	Single
Mirror scan rate	52.5 Hz
Field of view	28° (±14° from nadir*)
Roll compensated	Up to 20°
Overlap	100 % (50 % side-lap)

* Point on the ground vertically beneath the laser sensor on the aircraft.

Table 2.2: Summary statistics of forest attributes from ground measurement as response variables

Response	Total*	Minimum	Maximum	Median	Mean	SD†
BA (m^2/ha)	48,853	0.3	214.4	45.9	54.8	41.3
VOL (m^3/ha)	712,045	1.2	3,899.9	599.2	799.2	707.5
LOR (m)	32,549	6.7	76.0	36.0	36.5	16.0
QMD (cm)	43,162	14.0	177.2	42.8	48.4	23.8
DEN (trees/ha)	292,157	19.8	1,324.5	276.8	327.9	207.1

*Combined values for all the ground plots.

†Standard deviation.

Table 2.3: Advantage and disadvantage of the selected modeling methods

Method	Advantage	Disadvantage
NN	Easy to implement and intuitive to understand Fewer assumptions required Do not rely on any probability distribution Predictions are always within the biologically reasonable bounds	Need to determine the optimal number of reference observations (*k*) and the type of distance metric that are used in a single imputation Accuracy could be affected if reference dataset do not well cover the distributions of predictor variables Neither extrapolate values out of the ranges of reference dataset nor interpolate when k=1
SLM	Robust to misspecification of covariance model Estimate a spatial covariance matrix automatically in very general conditions Robust to non-normal data Can extrapolate	Data are used twice for covariance parameters and best linear unbiased prediction Might produce nonsense prediction
OLS	Most familiar method Strong theoretical background Can extrapolate	Many assumptions required Need to check interaction between predictor variables Might produce nonsense prediction
GWR	Explore spatial variations in relationships between response variable and predictor variables Consider spatial structure of data Can extrapolate	Multi-collinearity among predictor variables in each local regression model (Wheeler and Tiefelsdorf, 2005) Tend to produce extreme coefficients using less dense training data (Cho et al., 2009) Need to determine the optimal kernel bandwidth in each local regression model Might produce nonsense prediction

Table 2.4: Fifteen LiDAR metrics as candidate predictor variables

LiDAR metrics	Min*	Max†	Mean	SD‡	Description
elev_max (m)	6.3	92.6	52.0	19.7	Height maximum
elev_mean (m)	2.4	59.0	25.0	12.8	Height mean
elev_std (m)	1.0	26.3	11.7	5.9	Height standard deviation
elev_cv	0.2	1.7	0.5	0.3	Height coefficient of variation
elev_iq (m)	0.8	55.6	16.1	11.3	Height interquartile range
elev_p50 (m)	1.5	66.4	25.1	14.7	Height 50^{th} percentile
elev_p60 (m)	1.7	70.4	28.1	15.9	Height 60^{th} percentile
elev_p75 (m)	2.0	76.6	32.9	17.3	Height 75^{th} percentile
elev_p80 (m)	2.3	78.0	34.8	17.8	Height 80^{th} percentile
elev_p90 (m)	3.7	82.3	39.9	18.7	Height 90^{th} percentile
elev_p95 (m)	4.2	85.3	43.6	19.0	Height 95^{th} percentile
CRR	0.1	0.8	0.5	0.2	$\text{Canopy relief ratio} = \frac{\text{height mean} - \text{height minimum}}{\text{height maximum} - \text{height minimum}}$
p_1th_retn_6.56 (%)	8.3	100.0	90.4	15.6	Percentage of first returns above height of 2 m
p_all_retn_6.56 (%)	8.2	99.9	86.4	15.0	Percentage of all returns above height of 2 m
p_all6.56_all1th (%)	8.4	204.8	126.3	34.1	$\frac{\text{Number of all returns above height of 2 m}}{\text{Number of total first returns}} \times 100$

*Minimum value.

†Maximum value.

‡Standard deviation.

Table 2.5: List of LiDAR metrics as predictors for each response variable via 'best subset' variable selection

Response	Adj. R^2	RMSE	Predictors
BA (m^2/ha)	0.658	24.1	elev_mean, elev_std, p_all_retn_6.56, elev_mean:p_all_retn_6.56*
VOL (m^3/ha)	0.736	364.6	elev_mean, elev_std, CRR, elev_mean:CRR*, elev_mean:elev_std:CRR*
LOR (m)	0.810	7.0	elev_mean, elev_std
QMD (cm)	0.675	13.6	elev_cv, elev_p75
DEN (trees/ha)	0.394	160.8	elev_p75, CRR, p_all_retn_6.56, CRR:p_all_retn_6.56*

*A multiplication interaction between the pair of listed variables.

Table 2.6: List of LiDAR metrics as predictors for each response variable via imputation variable selection

Response	Predictors
BA (m^2/ha)	elev_mean, elev_max, elev_p90, p_1th_retn_6.56, p_all_retn_6.56
VOL (m^3/ha)	elev_mean, elev_p60, p_all_retn_6.56
LOR (m)	elev_p90, elev_cv, elev_p80, p_1th_retn_6.56
QMD (cm)	elev_p80, elev_p95, elev_p90
DEN (trees/ha)	elev_p90, elev_p75, elev_cv, elev_mean, elev_p80

Table 2.7: Correlation coefficients between each response variable and corresponding selected predictor variables

BA

	BA	elev_mean	elev_std
elev_mean	0.787		
elev_std	0.485	0.682	
p_all_retn_6.56	0.471	0.518	0.288

VOL

	VOL	elev_mean	CRR
elev_mean	0.843		
CRR	0.558	0.696	
elev_std	0.549	0.682	0.082

LOR

	LOR	elev_mean
elev_mean	0.857	
elev_std	0.785	0.682

QMD

	QMD	elev_p75
elev_p75	0.803	
elev_cv	-0.114	-0.349

DEN

	DEN	p_all_retn_6.56	CRR
p_all_retn_6.56	0.343		
CRR	0.260	0.503	
elev_p75	-0.186	0.481	0.609

Table 2.8: Performance measures (PM) by response variable and prediction type (PT) via ‘best subset’ variable selection

Response	PT	PM	MSN1	MSN5	RF1	RF5	GNN1	GNN5	BestNN	SLM	OLS	GWR
BA (m^2/ha)	Point	%Bias	0.082	0.062	-0.897	-0.483	0.082	0.062	-0.583	-0.085	-0.087	-0.778
			(0.5263)	(0.5367)	(<0.0001)	(<0.0001)	(0.5263)	(0.5340)	(<0.0001)	(0.3592)	(0.3558)	(<0.0001)
		RMSPE	33.44	25.79	31.84	26.25	33.44	25.79	24.02	24.04	24.24	24.08
		PIC90	89.6	89.7	82.6	83.5	89.6	89.7	89.4	89.6	89.8	89.9
	Total	%Bias	0.041	0.031	-0.448	-0.241	0.041	0.031	-0.291	-0.043	-0.043	-0.388
			(0.6438)	(0.6529)	(<0.0001)	(0.0005)	(0.6438)	(0.6508)	(<0.0001)	(0.5207)	(0.5173)	(<0.0001)
		RMSPE	965.7	746.8	853.0	761.8	965.7	746.8	709.3	723.7	729.0	750.9
		PIC90	90.6	90.8	82.6	80.8	90.6	91.0	87.6	89.0	89.0	86.8
VOL (m^3/ha)	Point	%Bias	0.085	-0.123	0.190	-0.365	0.085	-0.127	-1.069	-0.082	-0.102	-0.692
			(0.5296)	(0.2411)	(0.1517)	(0.0006)	(0.5296)	(0.2235)	(<0.0001)	(0.3994)	(0.2968)	(<0.0001)
		RMSPE	508.4	394.3	499.1	400.2	508.4	394.3	371.2	367.8	368.1	367.8
		PIC90	89.4	89.4	83.7	84.3	89.4	89.4	88.8	89.3	89.3	89.2
	Total	%Bias	0.042	-0.061	0.095	-0.182	0.042	-0.064	-0.534	-0.041	-0.051	-0.346
			(0.6313)	(0.3832)	(0.2547)	(0.0107)	(0.6313)	(0.3649)	(<0.0001)	(0.5496)	(0.4551)	(<0.0001)
		RMSPE	14,046	11,179	13,250	11,379	14,046	11,170	11,249	10,920	10,837	11,260
		PIC90	91.6	91.4	85.0	83.6	91.6	91.4	88.2	88.2	88.6	88.0
LOR (m)	Point	%Bias	-0.045	-0.122	-0.068	-0.225	-0.045	-0.124	-0.334	-0.047	-0.046	0.017
			(0.4286)	(0.0059)	(0.2137)	(<0.0001)	(0.4286)	(0.0051)	(<0.0001)	(0.2492)	(0.2538)	(0.6742)
		RMSPE	9.81	7.63	9.42	7.65	9.81	7.63	7.18	6.99	7.00	7.02
		PIC90	89.7	90.6	86.5	87.3	89.7	90.6	90.4	90.5	90.6	90.6
	Total	%Bias	-0.023	-0.061	-0.034	-0.112	-0.023	-0.062	-0.167	-0.023	-0.023	0.009
			(0.5455)	(0.0428)	(0.3164)	(0.0002)	(0.5455)	(0.0392)	(<0.0001)	(0.4157)	(0.4166)	(0.7657)
		RMSPE	271.0	218.8	246.9	221.0	271.0	218.8	218.5	208.8	207.3	208.8
		PIC90	93.0	91.4	87.0	83.2	93.0	91.4	88.6	90.8	91.2	91.4
QMD (cm)	Point	%Bias	-0.227	0.015	-1.080	-0.648	-0.227	0.017	-0.174	-0.051	-0.040	-0.215
			(0.0058)	(0.8118)	(<0.0001)	(<0.0001)	(0.0058)	(0.7893)	(0.0042)	(0.3897)	(0.5084)	(0.0003)
		RMSPE	18.79	14.67	18.47	14.95	18.79	14.67	13.86	13.63	13.64	13.55
		PIC90	89.8	90.5	85.9	86.9	89.8	90.5	90.7	91.0	90.9	91.6
	Total	%Bias	-0.113	0.008	-0.539	-0.324	-0.113	0.009	-0.087	-0.026	-0.020	-0.108
			(0.0342)	(0.8616)	(<0.0001)	(<0.0001)	(0.0342)	(0.8448)	(0.0350)	(0.5225)	(0.6253)	(0.0082)
		RMSPE	518.2	421.6	530.9	431.8	518.2	421.6	398.8	386.1	388.6	393.7
		PIC90	93.2	91.0	84.2	84.8	93.2	91.0	89.6	90.0	90.2	89.2
N (trees/ha)	Point	%Bias	0.129	-0.078	-1.064	-1.200	0.129	-0.079	0.409	-0.191	-0.244	-0.694
			(0.3695)	(0.4873)	(<0.0001)	(<0.0001)	(0.3695)	(0.4804)	(0.0001)	(0.0605)	(0.0195)	(<0.0001)
		RMSPE	222.5	173.0	217.8	177.0	222.5	173.0	164.6	157.7	161.6	161.0
		PIC90	89.7	90.5	82.8	83.0	89.7	90.5	90.6	91.4	91.4	91.5
	Total	%Bias	0.065	-0.039	-0.531	-0.599	0.065	-0.039	0.204	-0.096	-0.122	-0.347
			(0.4639)	(0.5943)	(<0.0001)	(<0.0001)	(0.4639)	(0.5886)	(0.0056)	(0.1736)	(0.0907)	(<0.0001)
		RMSPE	5,748	4,757	5,908	5,209	5,748	4,759	4,831	4,589	4,712	4,744
		PIC90	94.4	92.4	83.4	80.0	94.4	92.2	88.4	90.4	91.2	91.2

Table 2.9: Performance measures (PM) for the imputation methods with a subset of predictor variables via imputation variable selection

Response	PT	PM	MSN1	MSN5	RF1	RF5	GNN1	GNN5	BestNN
BA (m^2/ha)	Point	%Bias	-0.164	-0.136	-1.827	-1.120	-0.164	-0.133	-0.281
			(0.2109)	(0.1856)	(<0.0001)	(<0.0001)	(0.2109)	(0.1941)	(0.0034)
		RMSPE	33.79	26.46	32.05	26.06	33.79	26.46	24.73
		PIC90	89.3	89.4	82.1	82.4	89.3	89.4	89.2
	Total	%Bias	-0.082	-0.068	-0.912	-0.559	-0.082	-0.066	-0.140
			(0.3441)	(0.3415)	(<0.0001)	(<0.0001)	(0.3441)	(0.3508)	(0.0441)
		RMSPE	942.4	776.4	956.6	809.7	942.4	776.6	760.9
		PIC90	90.8	90.8	77.0	75.8	90.8	90.8	87.6
VOL (m^3/ha)	Point	%Bias	-0.066	-0.143	0.101	-0.273	-0.066	-0.144	-0.343
			(0.6119)	(0.1636)	(0.4208)	(0.0076)	(0.6119)	(0.1606)	(0.0005)
		RMSPE	493.6	386.4	473.1	385.1	493.6	386.5	371.2
		PIC90	89.7	89.5	83.4	84.7	89.7	89.5	89.0
	Total	%Bias	-0.033	-0.071	0.051	-0.136	-0.033	-0.072	-0.171
			(0.7015)	(0.3069)	(0.5306)	(0.0472)	(0.7015)	(0.3035)	(0.0150)
		RMSPE	13,775	11,112	12,808	10,946	13,775	11,116	11,217
		PIC90	91.8	91.2	82.8	82.4	91.8	91.2	88.6
LOR (m)	Point	%Bias	-0.084	-0.190	-0.153	-0.096	-0.084	-0.193	-0.297
			(0.1586)	(<0.0001)	(0.0068)	(0.0354)	(0.1586)	(<0.0001)	(<0.0001)
		RMSPE	10.25	7.99	9.75	7.86	10.25	7.99	7.47
		PIC90	89.5	90.7	85.5	85.9	89.5	90.7	90.8
	Total	%Bias	-0.042	-0.095	-0.077	-0.048	-0.042	-0.096	-0.148
			(0.2864)	(0.0031)	(0.0357)	(0.1201)	(0.2864)	(0.0027)	(<0.0001)
		RMSPE	285.9	234.4	265.5	224.6	285.9	234.7	226.8
		PIC90	92.0	91.2	82.0	82.2	92.0	91.0	88.6
QMD (cm)	Point	%Bias	0.018	-0.101	-0.159	-0.627	0.018	-0.103	0.133
			(0.8340)	(0.1266)	(0.0544)	(<0.0001)	(0.8340)	(0.1186)	(0.0332)
		RMSPE	19.24	15.13	18.89	15.23	19.24	15.13	14.21
		PIC90	89.7	90.4	83.7	85.3	89.7	90.4	90.6
	Total	%Bias	0.009	-0.051	-0.080	-0.313	0.009	-0.052	0.066
			(0.8733)	(0.2671)	(0.1469)	(<0.0001)	(0.8733)	(0.2569)	(0.1197)
		RMSPE	533.4	439.4	528.3	447.0	533.4	439.4	411.5
		PIC90	93.2	89.4	82.2	82.6	93.2	89.4	91.2
DEN (trees/ha)	Point	%Bias	0.068	-0.120	-2.369	-1.964	0.068	-0.117	1.875
			(0.6716)	(0.3335)	(<0.0001)	(<0.0001)	(0.6716)	(0.3467)	(<0.0001)
		RMSPE	246.8	191.5	217.6	180.3	246.8	191.5	174.2
		PIC90	90.0	90.3	81.6	84.0	90.0	90.3	91.1
	Total	%Bias	0.340	-0.060	-1.183	-0.981	0.034	-0.058	0.936
			(0.7485)	(0.4862)	(<0.0001)	(<0.0001)	(0.7485)	(0.4980)	(<0.0001)
		RMSPE	6,881.9	5,604.3	6,869.5	6,002.2	6,881.9	5,604.3	5,836.6
		PIC90	91.8	90.4	75.0	76.6	91.8	90.4	84.6

Table 2.10: Comparison of RF's performance measures (PM) between using the selected variables and using all the reviewed variables

Response	Prediction	PM	RF1			RF5		
			Selected	All	Status	Selected	All	Status
BA	Point	%Bias	-0.897 (<0.0001)	-1.024 (<0.0001)	D	-0.483 (<0.0001)	-0.625 (<0.0001)	D
		RMSPE	31.84	32.40	D	26.25	26.47	D
		PIC90	82.6	80.7	D	83.5	80.1	D
	Total	%Bias	-0.448 (<0.0001)	-0.511 (<0.0001)	D	-0.241 (0.0005)	-0.312 (<0.0001)	D
		RMSPE	853.0	902.0	D	761.8	777.0	D
		PIC90	82.6	76.4	D	80.8	75.6	D
VOL	Point	%Bias	0.190 (0.1517)	-0.961 (<0.0001)	D	-0.365 (0.0006)	-0.352 (0.0005)	I
		RMSPE	499.1	469.3	I	400.2	382.0	I
		PIC90	83.7	80.5	D	84.3	81.3	D
	Total	%Bias	0.095 (0.2547)	-0.480 (<0.0001)	D	-0.182 (0.0107)	-0.176 (0.0117)	I
		RMSPE	13,250	13,142	I	11,379	11,125	I
		PIC90	85.0	76.4	D	83.6	75.4	D
LOR	Point	%Bias	-0.068 (0.2137)	-0.424 (<0.0001)	D	-0.225 (<0.0001)	-0.164 (0.0003)	I
		RMSPE	9.42	9.62	D	7.65	7.78	D
		PIC90	86.5	83.6	D	87.3	83.5	D
	Total	%Bias	-0.034 (0.3164)	-0.212 (<0.0001)	D	-0.112 (0.0002)	-0.082 (0.0063)	I
		RMSPE	246.9	273.0	D	221.0	219.0	I
		PIC90	87.0	75.0	D	83.2	77.2	D
QMD	Point	%Bias	-1.080 (<0.0001)	-0.716 (<0.0001)	I	-0.648 (<0.0001)	-0.677 (<0.0001)	D
		RMSPE	18.47	17.90	I	14.95	14.40	I
		PIC90	85.9	81.2	D	86.9	81.9	D
	Total	%Bias	-0.539 (<0.0001)	-0.358 (<0.0001)	I	-0.324 (<0.0001)	-0.338 (<0.0001)	D
		RMSPE	530.9	500.0	I	431.8	428.0	I
		PIC90	84.2	77.6	D	84.8	77.2	D
DEN	Point	%Bias	-1.064 (<0.0001)	-1.983 (<0.0001)	D	-1.200 (<0.0001)	-0.745 (<0.0001)	I
		RMSPE	217.8	205.0	I	177.0	168.2	I
		PIC90	82.8	78.1	D	83.0	79.5	D
	Total	%Bias	-0.531 (<0.0001)	-0.991 (<0.0001)	D	-0.599 (<0.0001)	-0.372 (<0.0001)	I
		RMSPE	5,908	6,192	D	5,209	4,937	I
		PIC90	83.4	74.6	D	80.0	76.2	D

Note: ‘I’ in the column ‘Status’ represents ‘improved performance’ when using all the reviewed predictor variables gave better performance than using the selected predictor variables, and ‘D’ in the column ‘Status’ dose ‘declined performance’ when using all the reviewed predictor variables gave worse performance than using the selected predictor variables.

Table 2.11: Autocorrelation statistics from whole data

	Response variable				
	BA	VOL	LOR	QMD	DEN
Moran's I	0.0401	0.0550	0.0790	0.0674	0.0195
Z score	8.82	12.02	17.14	14.67	4.42

	Predictor variable					
	elev_mean	elev_std	elev_cv	elev_p75	CRR	p_all_retn_6.56
Moran's I	0.0972	0.0782	0.0404	0.0941	0.0493	0.0578
Z score	21.03	16.96	8.89	20.35	10.79	12.64

Table 2.12: Autocorrelation statistics from 500 simulations

	Training data					Validation data				
	Average	Minimum	Maximum	SD*	PA†	Average	Minimum	Maximum	SD*	PA†
Moran's I										
BA	0.0361	0.0007	0.0701	0.0108	-	0.0377	0.0106	0.0711	0.0103	-
VOL	0.0505	0.0114	0.0866	0.0126	-	0.0526	0.0227	0.0947	0.0122	-
LOR	0.0744	0.0360	0.1195	0.0140	-	0.0763	0.0390	0.1183	0.0138	-
QMD	0.0631	0.0297	0.1109	0.0140	-	0.0648	0.0333	0.1166	0.0140	-
DEN	0.0173	-0.0051	0.0419	0.0084	-	0.0175	-0.0042	0.0504	0.0091	-
Z score										
BA	4.47	0.33	8.25	1.23	0.986	4.66	1.52	8.58	1.19	0.994
VOL	6.16	1.55	10.14	1.45	0.998	6.41	2.27	11.33	1.41	1.000
LOR	8.93	4.50	13.76	1.59	1.000	9.15	4.74	14.05	1.58	1.000
QMD	7.63	3.78	12.75	1.63	1.000	7.83	4.57	13.54	1.62	1.000
DEN	2.29	-0.34	5.37	0.99	0.620	2.31	1.20	5.86	1.07	0.622

*Standard deviation

†Proportion of having autocorrelation in the dataset from 500 simulations, i.e. proportion of cases where Z-score is greater than 1.96.

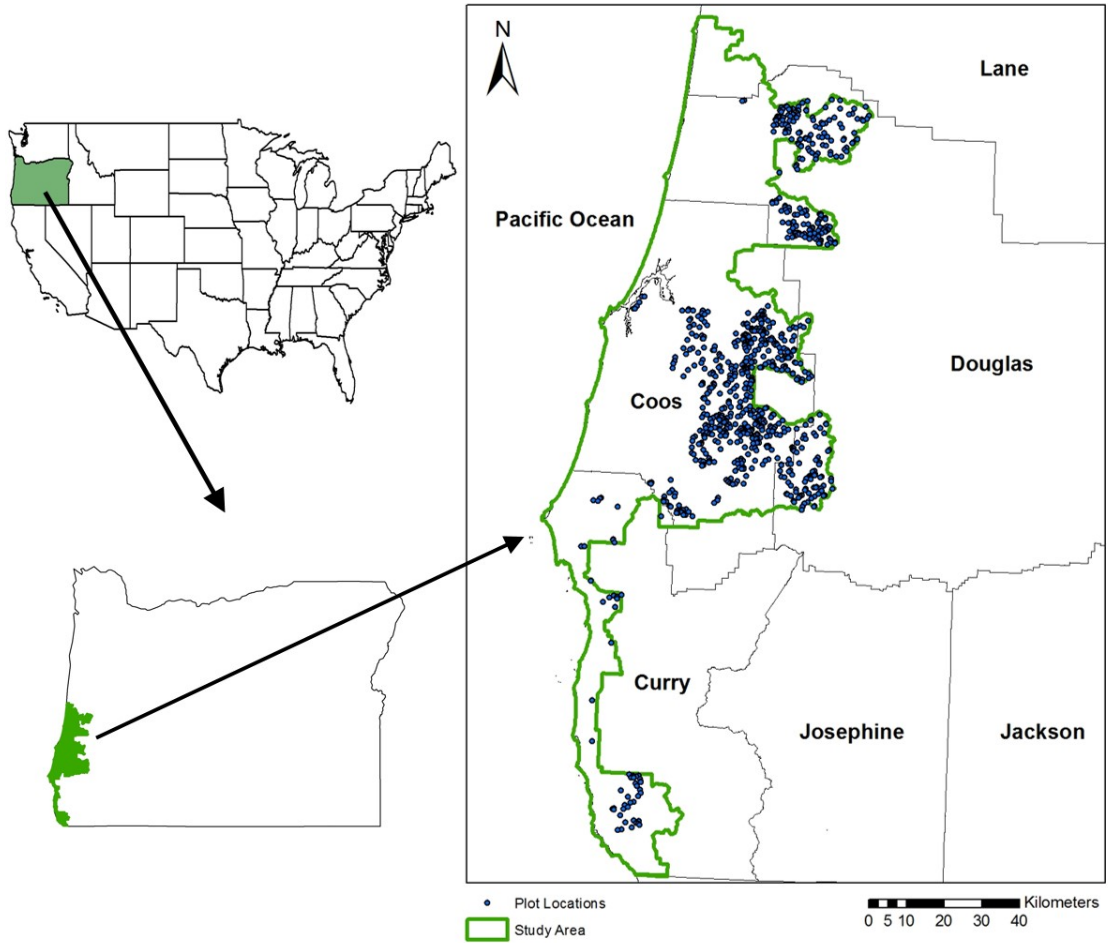


Figure 2.1: Location map of study area and plots

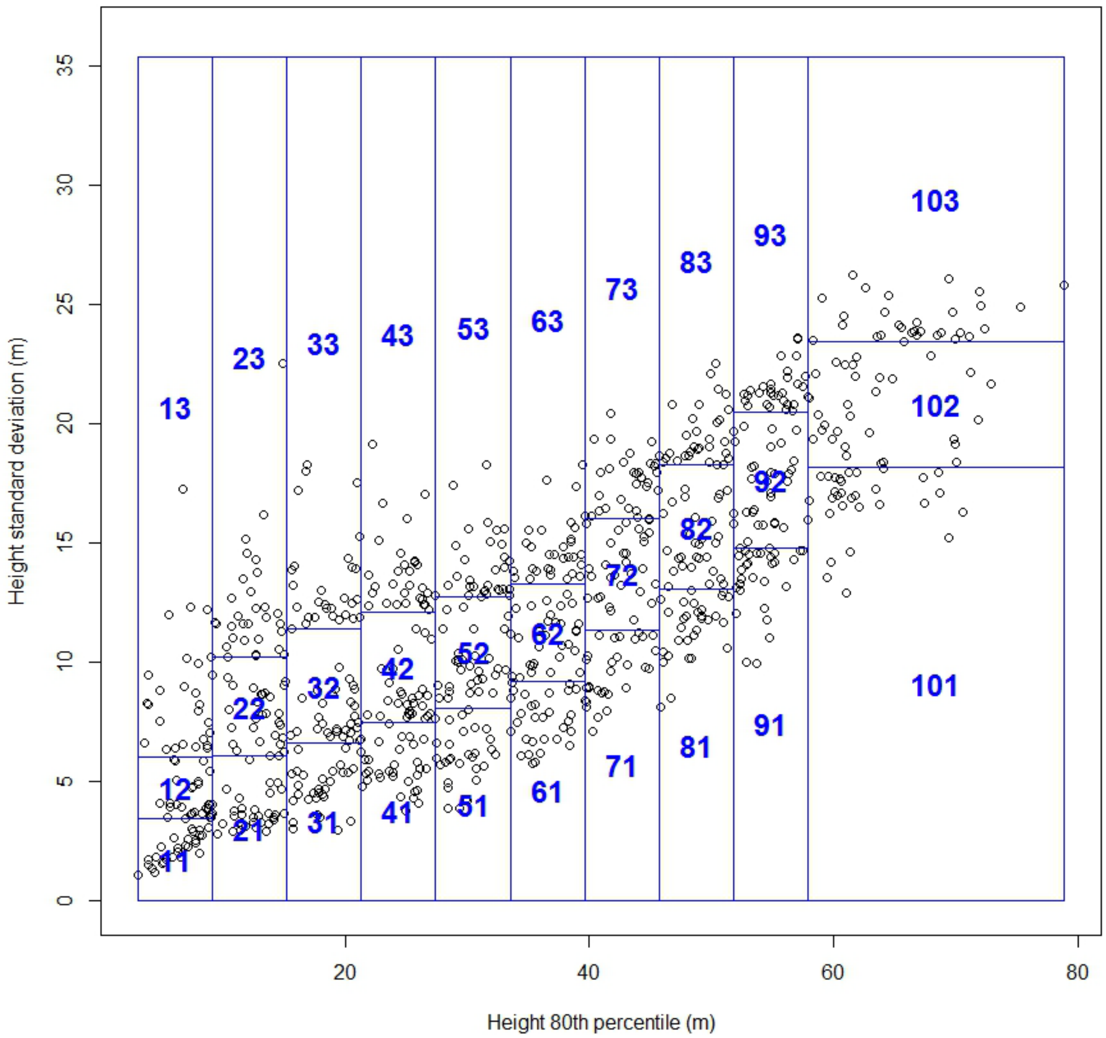


Figure 2.2: Plot bins divided by height 80th percentile and height standard deviation (Note that the numbers in each box represents the bin numbers.)

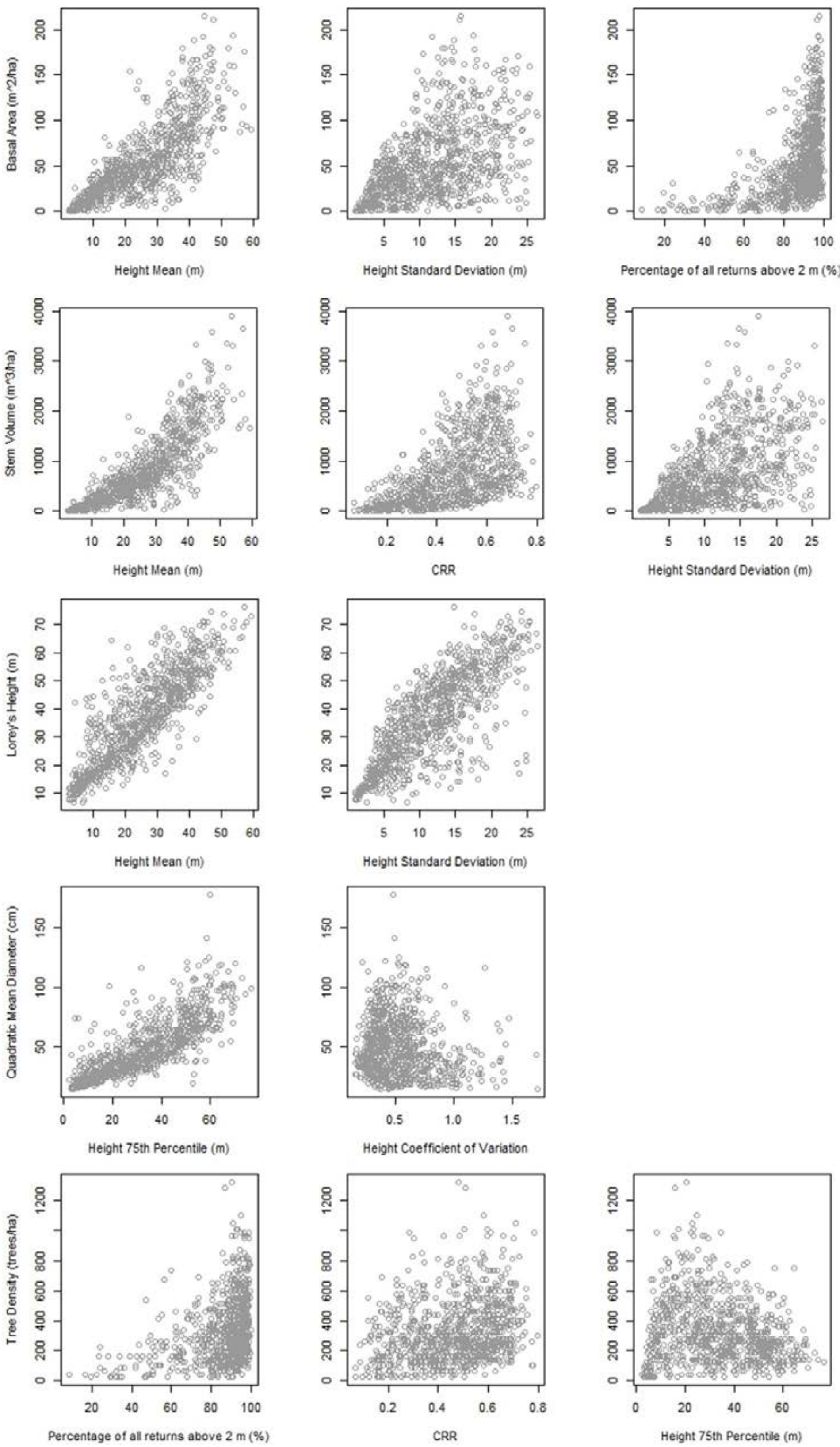


Figure 2.3: Scatter plots of response variables by predictor variables

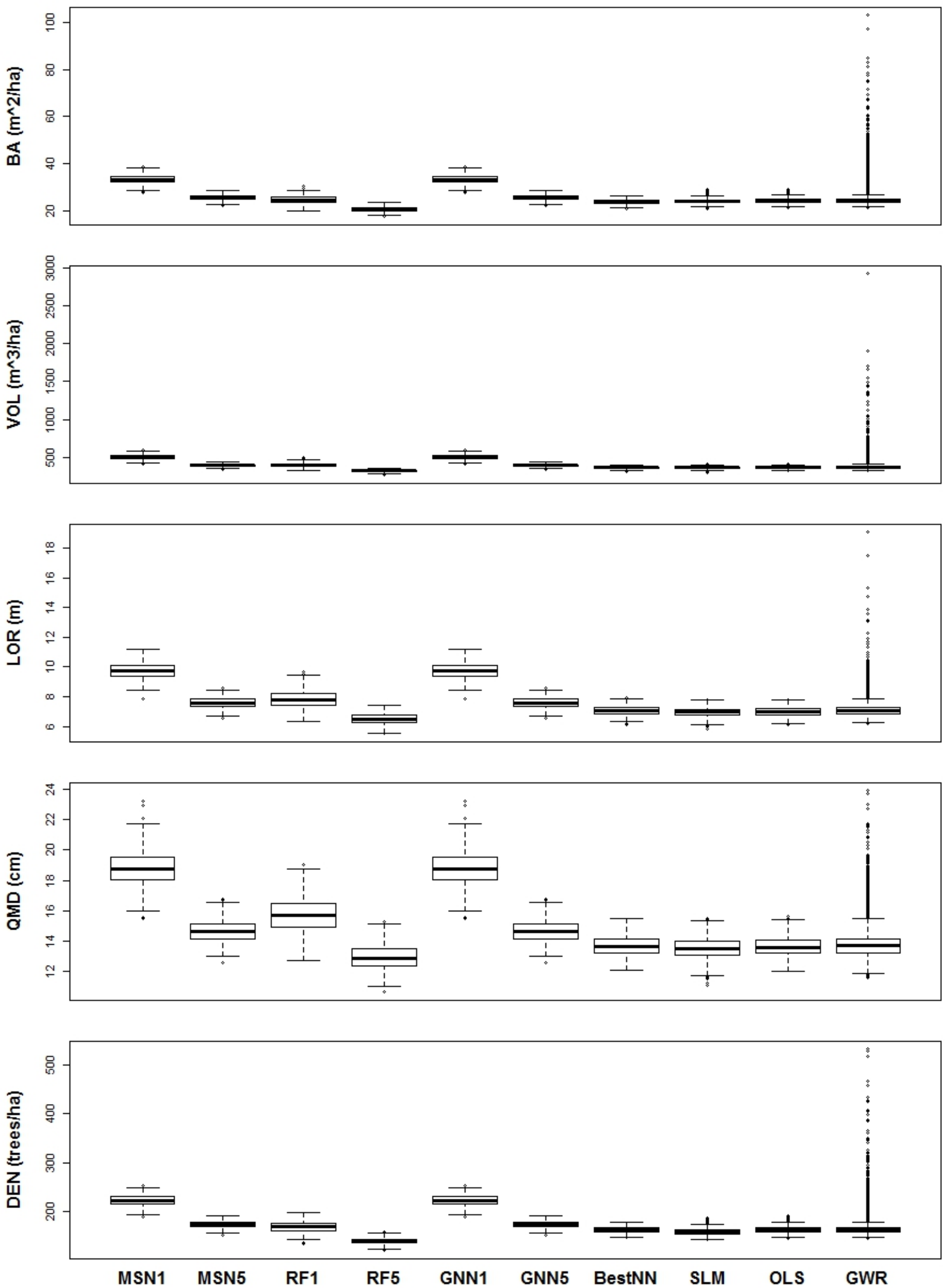


Figure 2.4: Boxplots of prediction standard error by modeling methods for each response variable

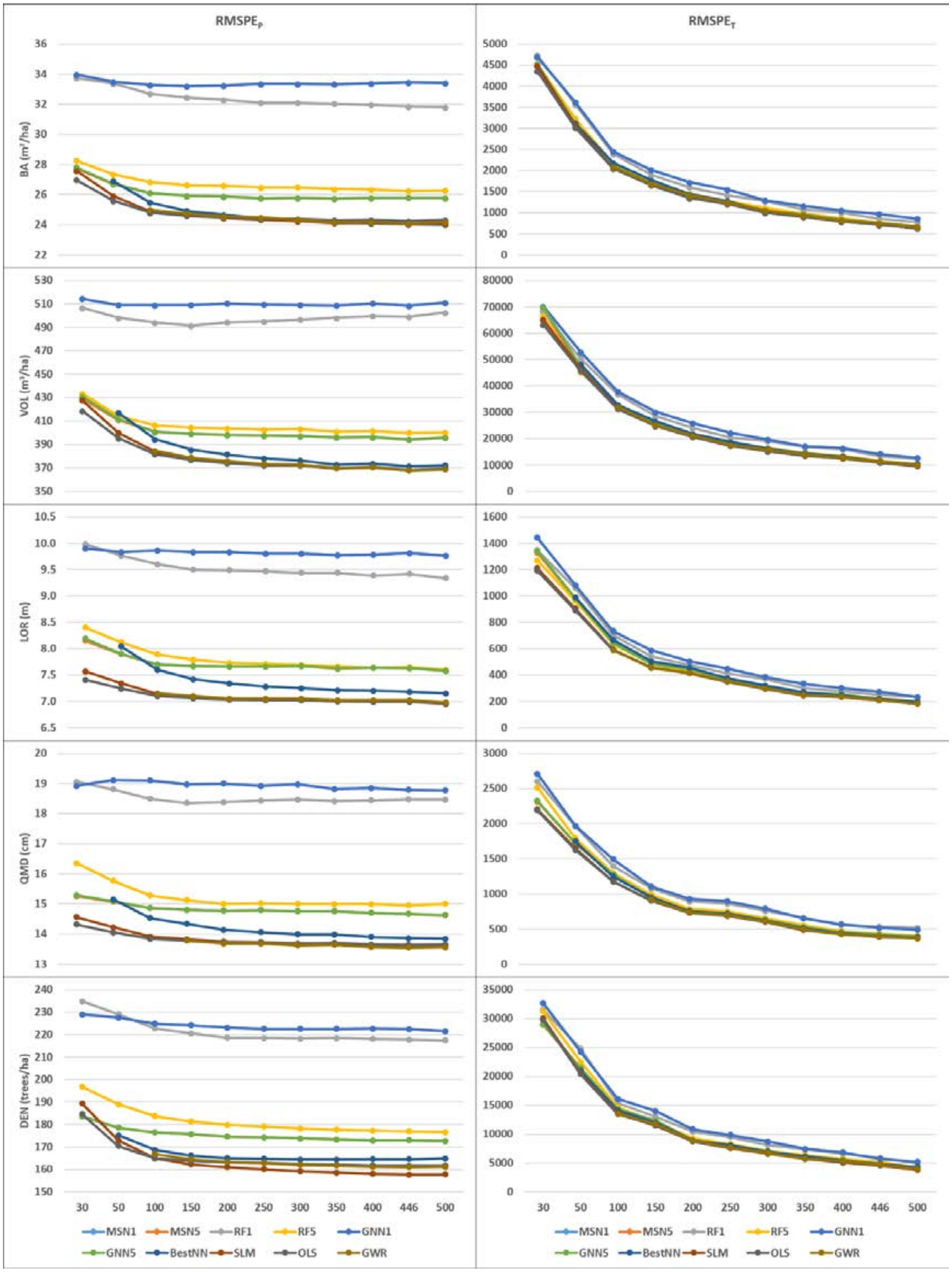


Figure 2.5: (a) RMSPE in point and (b) total prediction by response variables and modeling methods with different numbers of training data.

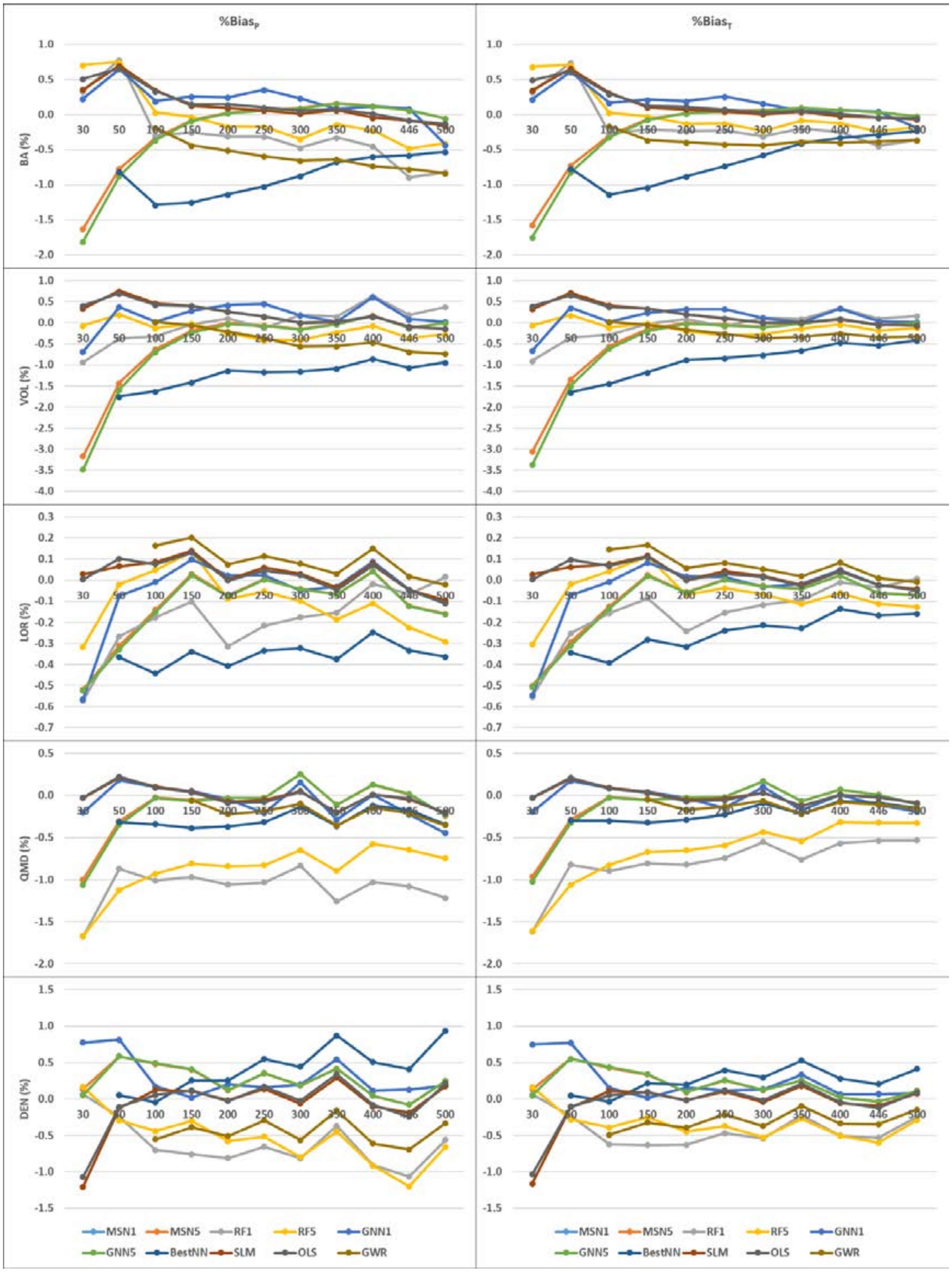


Figure 2.6: (a) Percentage bias in point and (b) total prediction by response variables and modeling methods with different numbers of training data.

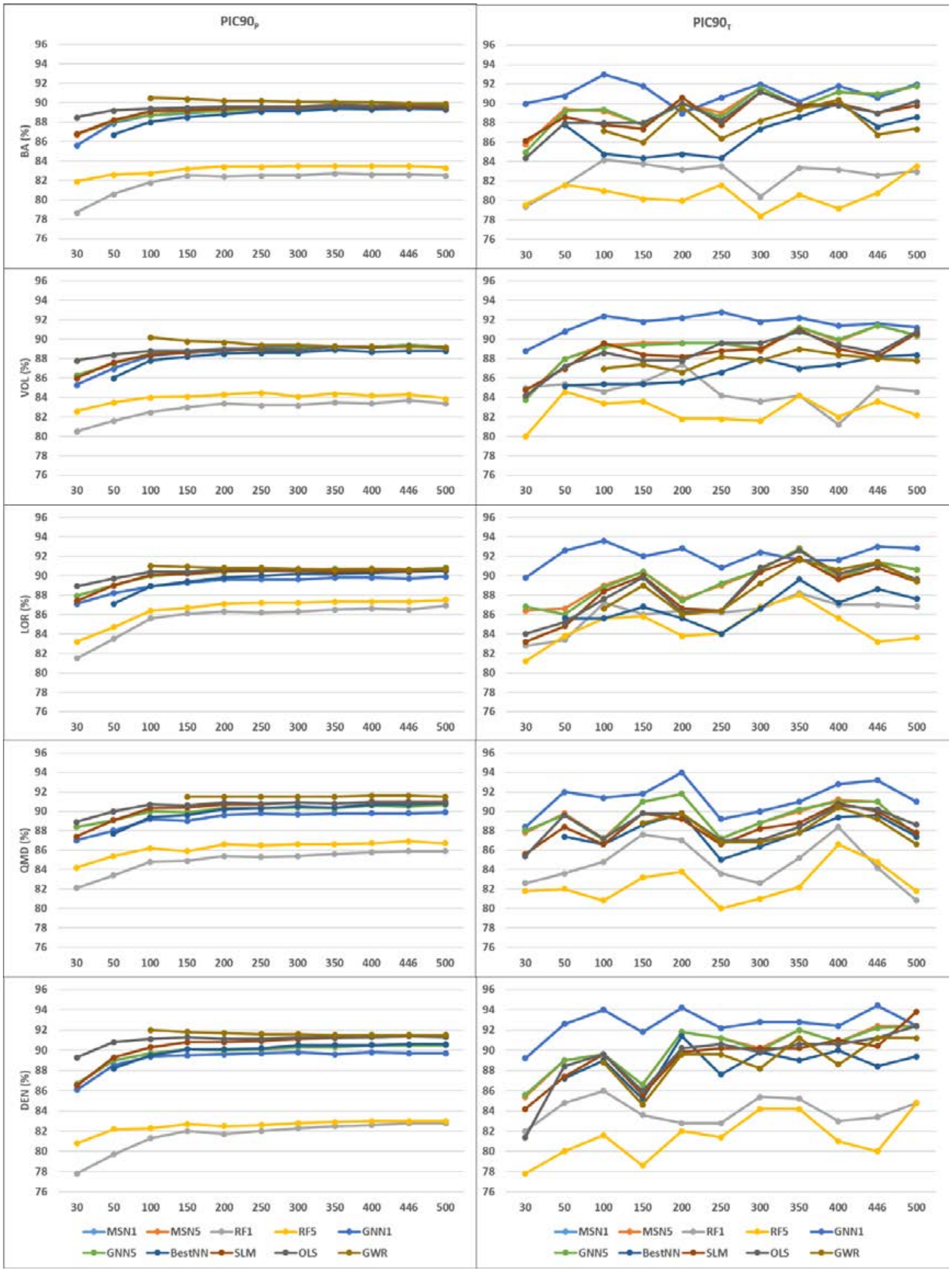


Figure 2.7: (a) PIC90 in point and (b) total prediction by response variables and modeling methods with different numbers of training data.

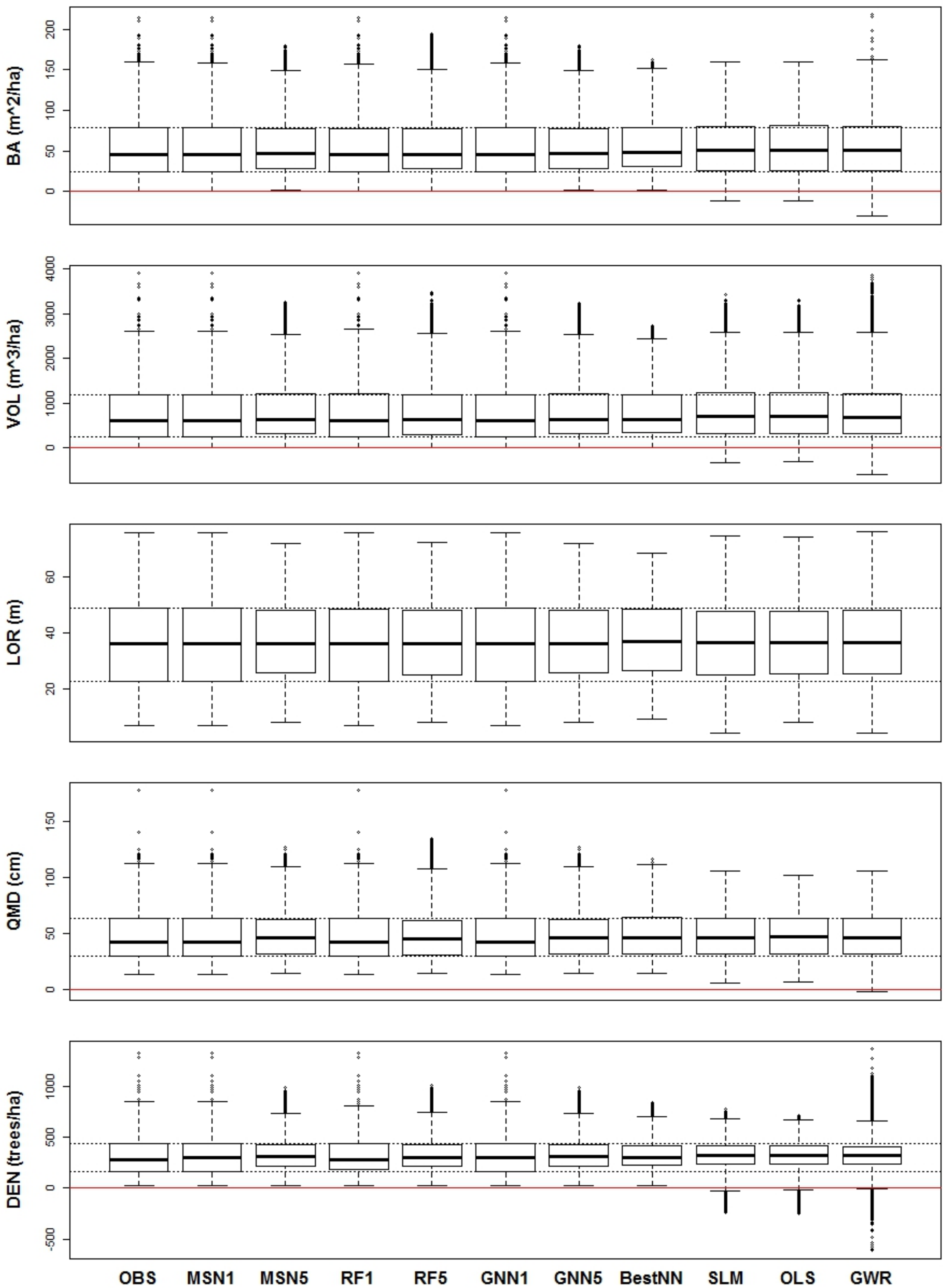


Figure 2.8: Boxplots of observation and prediction value by modeling methods, for each response variable.

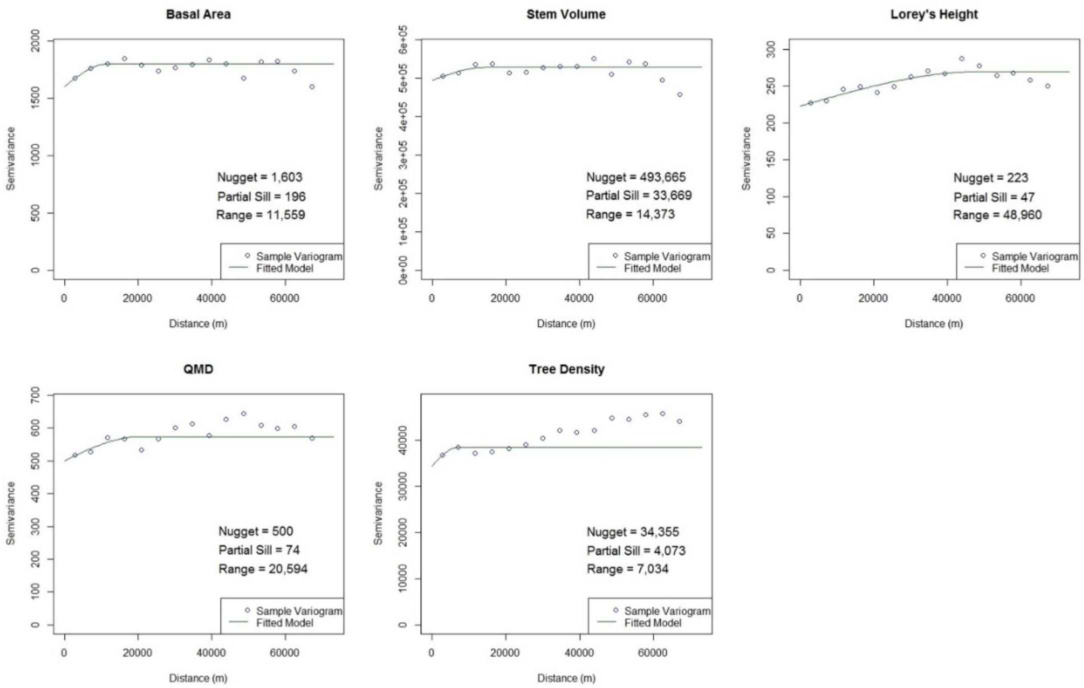


Figure 2.9: Semivariograms for each response variable.

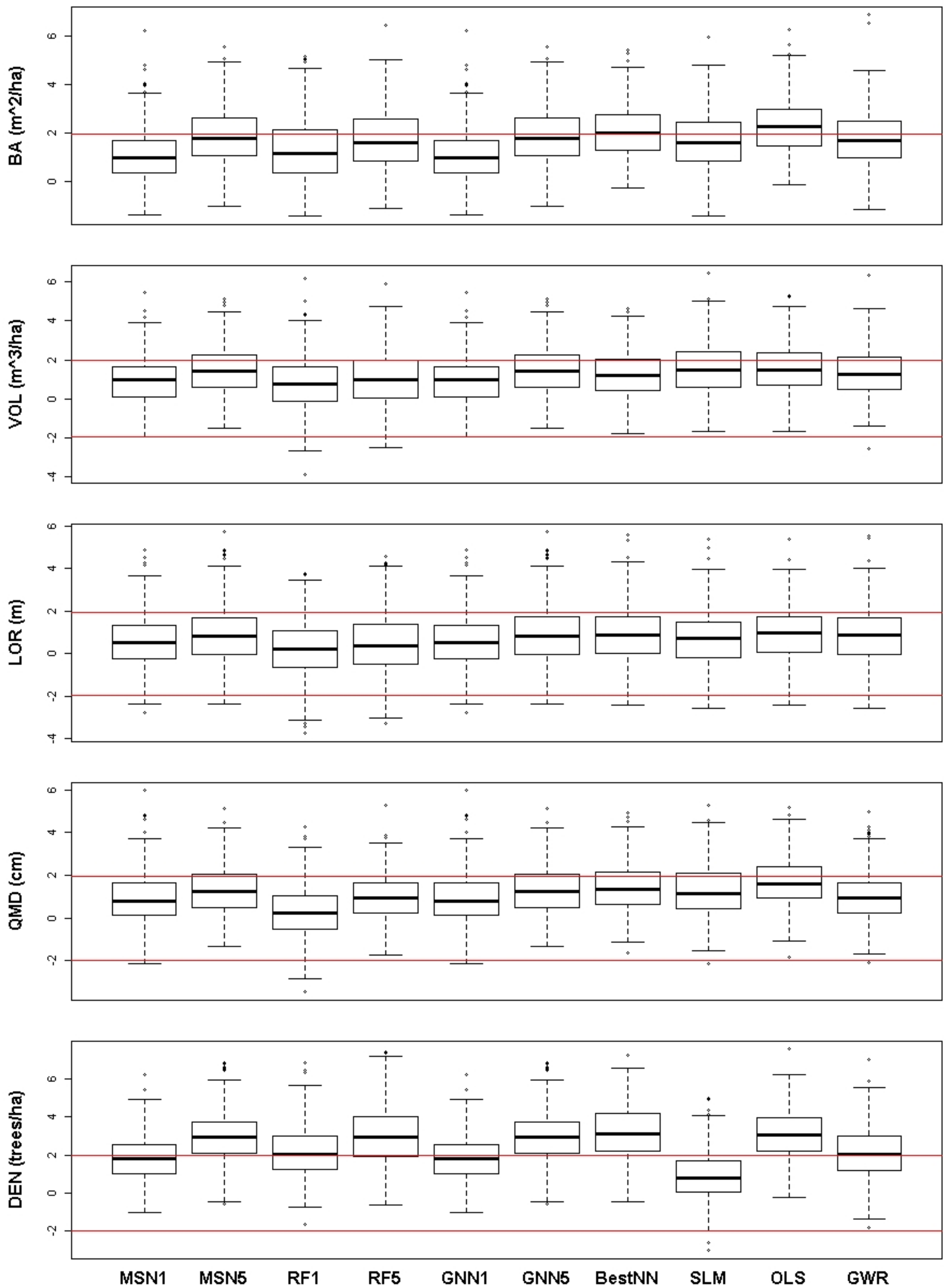


Figure 2.10: Z-score of Moran's I from prediction errors for each response variable by modeling methods.

Chapter 3: Generating Tree-lists by Fusing Individual Tree Detection and Nearest Neighbor Imputation Using Airborne LiDAR Data

Introduction

A tree-list provides detailed data foresters often desire for management and planning such as tree species, diameter at breast height (DBH), tree height (HT), basal area (BA) and stem volume. Field cruising has been commonly used to obtain such data. Field cruising is costly, however, and remote sensing data can be used as auxiliary information to improve the accuracy and precision of estimates in forest inventory.

Among various remote sensing techniques, airborne light detection and ranging (LiDAR) has been increasingly used in forestry applications during the last decade. LiDAR has performed well in estimating forest attributes such as biomass (Næsset and Gobakken, 2008), diameter distribution (Gobakken and Næsset, 2004), volume and BA (Lindberg and Hollaus, 2012). Tree-lists have also been estimated by LiDAR (Lindberg et al., 2010, 2013) or aerial photographs (Temesgen et al., 2003).

In general, there are mainly two approaches using LiDAR data in forestry, the area-based approach (ABA) and the individual tree detection (ITD) approach (Vauhkonen et al., 2014). ABA assumes that the vertical height distribution of laser point clouds is related to variables of interest in an area. A host of summary statistics derived from the point cloud are used to predict many forest inventory attributes. Information on the LiDAR point cloud is not fully utilized in ABA, i.e., most of the studies have focused on vertical height distribution in a sample plot and only a few studies using horizontal information obtained from the LiDAR point cloud. Pippuri et al. (2012) found horizontal texture metrics from a canopy height model (CHM) could be used to predict the spatial pattern of trees, and horizontal landscape metrics from a CHM used to predict the need for first thinning.

In contrast, ITD identifies individual trees and provides estimates of forest attributes based on the identified individual trees. Although many variations exist, ITD commonly uses a rasterized CHM to segment individual trees with horizontal location of treetop and height across the CHM area. Thus, ITD has apparent advantages over ABA regarding utilization of horizontal information in LiDAR point clouds and can be more suitable for tree-level forest inventories than ABA. However, information on understory vegetation is likely to be missed when using ITD (Koch et al., 2014). This is because rasterizing LiDAR point clouds into CHM means that there is a rounding effect of summarizing all the point clouds within a range of cells into one cell height value mainly focusing on higher point clouds making it difficult to detect or estimate understory vegetation. Additionally, it is well known that LiDAR has weaknesses for detecting or estimating understory vegetation regardless of the approach used because LiDAR data lack information on understory vegetation (lower proportion of point clouds in understory) (Takahashi et al., 2006).

Many approaches have been proposed to overcome the limitations above. Maltamo et al. (2004) combined a theoretical probability distribution function with the tree height distribution estimated from ITD to detect small and suppressed trees. ITD first estimated the height distribution and the number of large trees. For small trees, two approaches have been used including - the complete Weibull distribution with the parameter prediction method and the left-truncated Weibull distribution with estimation of parameters from the estimated height distribution by ITD. These approaches were tested for the estimation of the height distribution and the number of trees. DBHs for large and small trees were then predicted using the relationship between DBH and LiDAR metrics. Total timber volume and stem density were finally determined by summing the estimates from the two approaches for large and small trees. Lindberg et al. (2010) proposed a methodology to generate a tree-list combining a CHM-based ITD and ABA estimation. To better detect trees that are close to each other or small; 1) the number of trees per segment was estimated using a training dataset in which the number of field-measured trees for each tree crown segment was known, and 2) a candidate tree-list from the ITD was calibrated using the target distributions of HT and DBH estimated by a k-Nearest Neighbor (NN) approach. The combined approach improved the estimation of distributions for DBH and HT, and produced unbiased estimates of forest attributes. In addition to ITD based on CHM, Lindberg et al. (2013) utilized a 3D clustering method to model a tree crown using a priori information on the shape and proportions of tree crowns. The 3D clustering method identified more trees below the tallest canopy layer and with a DBH < 20 cm than ITD based on CHM. Hamraz et al. (2017) proposed the use of vertical stratification of point clouds and LiDAR data with high point cloud density (50 points / m2), which would have more information on understory vegetation than the one with low density, to detect understory trees. The proposed approach improved detecting understory trees without affecting the overall quality of segmentation for overstory trees.

Many parameters affect the performance of tree segmentation by ITD; these can be classified into two parameters, biological and technical. For the biological parameter, Vauhkonen et al. (2012) claimed that the performance of ITD methods depends more on forest structure, stand density, and tree clustering than on detection techniques. For example, an estimated tree segment by ITD could have no, one, or several trees in it (Breidenbach et al., 2010b), and trees in an understory under a dense upper canopy are hard to detect with LiDAR (Maltamo et al., 2004). On the other hand, the methods for ITD were reported as the primary parameter affecting the performance of ITD by Kaartinen et al. (2012). Substantial differences in the percentage of matched and missed trees, and commission error were found among the ITD methods. Also, the accuracy of determining tree location, tree height, and crown delineation changed according to the ITD methods. In contrast, pulse density showed less impact on ITD.

A typical ITD method consists of the following two steps: 1) generating a rasterized CHM with appropriate smoothing and resolution using normalized LiDAR point cloud data, and 2) tree segmentation using a segmentation technique on the rasterized CHM (finding local maxima as treetops and delineating tree crowns) (Yu et al., 2010). Therefore, the performance of ITD is affected by the parameters (smoothing and resolution for CHM, and the algorithm used for tree segmentation). In addition to these parameters, Wiggins (2017) reported that excluding trees below a specific height (minimum height cutoff) improved ITD's accuracy for overstory trees. Maltamo et al. (2003) noted that a proper value of the truncation parameter of Weibull for DBH distribution, which can be considered the same as a height cutoff, should be further studied. According to McGaughey (2016) and Wiggins (2017), there might be an optimal parameterization that balances the smoothing of the CHM, resolution of the CHM, and the height cutoff to best identify individual trees, although Koch et al. (2014) and McGaughey (2016) pointed out that the optimal parameterization can vary over large forest areas with diverse and complicated structure. To offset the variation of the optimal parameters, Koch et al. (2006) proposed applying different intensities of smoothing according to HT. This method would prevent under- and over-representation of local HT maxima.

Other than ITD, detailed information on forest resources, such as a tree list or stand table, has been estimated by several methods that can be mainly classified into two categories: 1) diameter distribution modeling, and 2) imputation. In diameter distribution modeling, parameters of some theoretical distributions are estimated to describe the distribution of tree diameters. Three approaches that are commonly used are the parameter prediction method, parameter recovery method, and quantile prediction method (Temesgen et al., 2003). Imputation methods directly substitute measured values from sample locations (references) for locations for which a prediction is desired (targets). The distance metric used to identify suitable references and the number of references used in a single imputation (k) are the key considerations to classify the imputation methods such as most similar neighbor, gradient nearest neighbor, or Random Forest NN (RF NN hereafter) (Eskelson et al., 2009). Temesgen et al. (2003) used a set of proxy variables to represent a tree-list in NN imputations because there is no single variable to represent the tree-list. On the other hand, Strunk et al. (2017) used plot identities as a response variable in NN imputations in evaluating NN strategies to impute a tree-list.

In our study, we combined ABA and ITD to estimate tree-list using LiDAR data inspired by the ideas from Maltamo et al. (2003), Maltamo et al. (2004) and Wiggins (2017). This was for overcoming the weakness of LiDAR data and the ITD method in identifying understory trees, and utilizing the strength of ITD over ABA. Maltamo et al. (2003) combined pattern recognition of single trees with the truncated Weibull distribution to estimate forest characteristics using digital video imagery. Trees were grouped into large (DBH > 17cm) and small (DBH ≤ 17cm) trees. The cutoff DBH value (17 cm) was the minimum size of trees that could be detected by the pattern recognition method. The value of 17 cm in DBH was used as a truncation parameter of the left-truncated Weibull. Pattern recognition was applied to large trees (DBH > 17cm), and the diameter distribution modeling to small trees (DBH ≤ 17cm), respectively. This idea was improved upon by Maltamo et al. (2004), who combined ITD based on CHM for large trees and diameter distribution modeling for small trees. HT distribution was modeled using LiDAR metrics as auxiliary variables. Wiggins (2017) examined the effect of height cutoff on the accuracy of LiDAR data for estimating forest structure of taller trees and found that a 12 m height cutoff produced better results in estimating forest structure and spatial pattern.

For ITD, we used watershed segmentation (Vincent and Soille, 1991) for overstory trees (trees taller than a height cutoff) and ABA by NN (k = 1) imputation for understory trees (trees shorter than the height cutoff). While the performances of diameter distribution modeling depended on the results from large tree estimation by the single tree pattern recognition in Maltamo et al. (2003) or the ITD based on CHM in Maltamo et al. (2004), in this study, we used ITD and ABA independently. They were only linked by a height cutoff when generating a complete tree-list. Whereas Lindberg et al. (2010) estimated a tree-list for all trees by an ITD method and calibrated it, our approach separated a forest stand into overstory and understory trees, then applied different methods to the overstory and understory trees, respectively. We examined the effects of the combination of the three parameters, smoothing of CHM, resolution of CHM and the height cutoff, as well as LiDAR height classification of field plots on estimating tree-lists via ITD. The explanatory power of our approach was also investigated. We evaluated the performance of generating tree-lists in terms of BA, mean HT, stems per hectare (SPH), and distributions of DBH and HT.

Methods

Study Area

The study area is located in southwestern Oregon with the extent of 647,951 hectares (Figure 3.1). The elevation of the area ranges approximately from 20 m to 1,000 m above sea level in elevation. The range of slopes in the area is 0° to 89.97°. Douglas-fir (*Pseudotsuga menziesii*) is the dominant tree species in the study area, and other important species are western hemlock (*Tsuga heterophylla*), red alder (*Alnus rubra*), Oregon myrtle (*Umbellularia californica*), bigleaf maple (*Acer macrophyllum*), tanoak (*Notholithocarpus densiflorus*), western redcedar (*Thuja plicata*), and grand fir (*Abies grandis*).

Airborne LiDAR

Airborne LiDAR data were collected between April 27th, 2008 and April 5th, 2009 using Leica ALS50 Phase II instrumentation. The average pulse density (the average number of pulses returned from surfaces) was 8.10/m^2 for the study area. Table 3.1 shows the specifications for the LiDAR survey. Laser points with elevations above ground level lower than 1 m and higher than 91.44 m (300 feet) were excluded from the computation because they did not likely represent vegetation of interest (the maximum tree height measured in the field data was 88.4 m).

Field Data

Stratified sampling based on the LiDAR metrics (Hawbaker et al., 2009) was used for field data collection. Only the lands owned by the BLM or the Coquille Tribe in the study area were considered. Then, the non-forested areas were removed. Within this pre-selected area, a set of LiDAR grid metrics (22.86 m by 22.86 m) were calculated from the LiDAR point clouds. Using the principal component analysis, the 80th percentile and standard deviation of the LiDAR height were selected as describing best the variation in forest structure in the pre-selected area. Two thousand cells were randomly selected from the cells with the pre-selected area. Based on these random samples, the range of 80th percentile heights was subdivided into ten classes with a length of 6.10 m, and the range of standard deviations within each height class into three equal-width classes. The maximum height of the uppermost 80th percentile class was increased to 83.52 m to cover the values of the grid cells in the full dataset. A total of 30 bins (10 × 3) were created.

Every grid cell in the pre-defined area was assigned to the bins. Then, 30 primary and 20 alternate plot locations from each bin were randomly selected from the grid cells. 30 plot locations from each bin were measured by field crews from those 50 locations using the primary plot locations unless inconsistencies were found between the LiDAR measured structure and the actual state of the forest. Such inconsistencies were caused by disturbances, such as timber harvesting, fires, or windthrow that occurred after the LiDAR data acquisition. In that case, the next available alternate plot would replace the primary plot. Plot locations overlapping roads, and in tall shrub vegetation near the coast were discarded.

Field sampling was conducted between May 25, 2010 and May 10, 2011. Nested plots with two plot sizes (12.68 m and 5.09 m) were used to measure large (all live and dead with DBH larger than 14 cm) and small (only live with heights taller than 1.4 m and DBH less than 14 cm) trees, respectively.

Note that only the large tree data were used for this analysis. There was one missing plot, resulting in a total of 899 plots. Table 3.2 and Table 3.3 provide a plot-level and tree-level summary of the field measurements. The ten 80th percentile classes for the stratification sampling were used as LiDAR height classes in the current study (from '1' to '10' as height increases) to investigate the effect of LiDAR height classification of field plots on the performance of our proposed approach.

Generating Tree-lists

The general steps of our approach are shown in Figure 3.2. Trees taller than a specified height (a height cutoff) were estimated by ITD using LiDAR data yielding the number and HT of the taller trees. DBHs for the taller trees were predicted based on the estimated HT using the relationship between DBH and HT from field data. For estimating the trees shorter than the height cutoff, tree-lists for target plots were first imputed with the tree-list from reference plots by RF NN imputation using both LiDAR and field data. Then, the shorter trees were selected from the imputed tree-lists. A complete tree-list can be generated by combining those estimated taller and shorter trees. The variables in the complete tree-list were the tree ID, HT, and DBH.

Individual Tree Detection

ITD was implemented by the function '*TreeSeg*' in the FUSION software (McGaughey, 2016) with the argument '*ht_threshold*' to estimate the tree-list for large trees. This function applies a generalized watershed segmentation algorithm by Vincent and Soille (1991) to a CHM. It should be noted that over-segmentation, known as one of the disadvantages of the watershed algorithm, may be produced with noisy imagery (Romero-Zaliz and Reinoso-Gordo, 2018). Conceptually, the CHM is inverted, so tree crowns appear as basins. Water fills the basins from local height minima in the CHM by the algorithm, and the basins fill and join with adjacent basins, then watershed edges are established (McGaughey, 2016). This also can be explained at the pixel level on the CHM. In every CHM pixel above a height threshold, a path is placed by iteratively moving to the neighboring pixel with the largest height value until a local height maximum is reached. A tree crown segment is defined by cells that reach the same local height maximum (Lindberg and Holmgren, 2017). The '*ht_threshold*' sets minimum height (height cutoff) for tree segmentation. Fractions of CHM below this height cutoff were excluded in the segmentation process. The other two parameters, the amount of smoothing and the resolution of the CHM, were applied in generating the CHM implemented by the function '*CanopyModel*' in FUSION. We set levels of those three parameters as follows: 1) 3 levels of smoothing of CHM - no smoothing, median filter using a 3 by 3 neighbor window and median filter using a 5 by 5 neighbor window, 2) 24 resolutions of CHM - 0.2, 0.3, …, 2.4, and 2.5 m, 3) 9 percentile height cutoffs on the LiDAR height for each plot - 10th, 20th, …, 90th. Because the range of HT is extensive, the LiDAR height percentiles were used as height cutoffs instead of absolute heights as in Wiggins (2017).

After implementing ITD, we obtained a tree-list above a height cutoff including information on individual tree count, HT, a location of tree, and a number of CHM cells within a tree crown at a combination of smoothing, resolution of CHM, and height cutoff. To predict the DBHs of trees in the estimated tree-lists, an RF regression model for DBH was fitted with the HTs from the field data (16,200 trees). With this model, the DBHs of trees in the estimated tree-lists were predicted using the HTs of those trees. Then, those predicted DBHs were added to the estimated tree-lists. The model was fitted in R version 3.3.3 (R Core Team, 2017) using the R package '*randomForest*' (Liaw and Wiener, 2002).

Nearest Neighbor Imputation

To estimate tree-lists for understory trees, we used RF NN imputation instead of diameter distribution modeling because there were many sample plots with multimodal or irregular shapes in diameter distribution and some plots had a small number of trees. NN imputation directly substitutes measured values from references for targets. The type of NN imputation is determined mainly by the distance metric and number of neighbors (k) (Eskelson et al., 2009). The distance metric measures the similarity between target and reference observations, and the k indicates how many reference observations are used in a single imputation (prediction). Four distance metrics, Euclidean, Mahalanobis, most similar neighbor and RF (Breiman, 2001), were tested. RF appeared the best for BA, SPH and error index (EI; will be defined in the following section), and Euclidean showed the best for HT (this result is not presented in this manuscript). Thus, we selected the RF algorithm as the distance metric and chose $k = 1$. RF builds multiple classification (or regression) trees, called forests, with bootstrap samples of training data, while selecting predictors randomly for the best split at each node in the trees. Distance in RF NN is computed as one minus the proportion of classification trees where a target observation is in the same terminal node as a reference observation (Crookston and Finley, 2008). To estimate tree-lists by RF NN imputation, we imputed plot identities as in Strunk et al. (2017).

To fit an NN model, it is necessary to define response and predictor variables. Predictor variables were derived from LiDAR point clouds at each filed plot location. It is not clear which a single response variable or multiple response variables should be used for estimating tree-lists because many attributes can be extracted from a tree-list. For example, Temesgen et al. (2003) used a set of 22 proxy variables to represent a tree-list. We considered several forest inventory attributes (basal area, stem volume, Lorey's height, quadratic mean diameter, stems per ha) simultaneously to select appropriate predictor variables for estimating tree-lists via RF NN imputation. "Best subsets" was used as a variable selection method producing the best three predictors for each forest inventory attribute.

From the best predictors for each forest inventory attribute, we obtained a total of 11 predictors after removing duplicates. The selected predictors are shown in Table 3.4. Like leave-one-out validation, the target plot was excluded from training data when modeling. Nine different tree-lists for each height cutoff were generated from the estimated tree-lists by subtracting trees above the corresponding height cutoff. The variable selection and imputation modeling were implemented in R version 3.3.3 (R Core Team, 2017) using R packages '*yaImpute*' (Crookston and Finley, 2008) and '*randomForest*' (Liaw and Wiener, 2002).

Performance Measures

Bias and root mean squared error (RMSE) for mean HT, BA and SPH were computed as follows:

$$bias = \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)}{n} \qquad (3.1)$$

$$RMSE = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{n}} \tag{3.2}$$

where $\hat{y}_i$ is the prediction at the i^{th} plot, y_i is the field-measured value at the j^{th} plot, and n is the number of total sample plots.

Large trees would produce greater uncertainty in estimation than small ones because the larger trees have greater values of HT, DBH, etc. To see the effect of several parameters on tree-list estimation free from the influence of greater value, relative bias (RBias) and relative RMSE (RRMSE) were also calculated for each LiDAR height class by the equations below:

$$RBias\ (\%) = \frac{\sum_{i=1}^{n}(\hat{y}_{ih} - y_{ih})}{n_h} \times \frac{100}{\bar{y}_h} \tag{3.3}$$

$$RRMSE\ (\%) = \sqrt{\frac{\sum_{i=1}^{n}(\hat{y}_{ih} - y_{ih})^2}{n_h}} \times \frac{100}{\bar{y}_h} \tag{3.4}$$

where $\hat{y}_{ih}$ is the prediction at the i^{th} plot in the h^{th} LiDAR height class, y_{ih} is the field-measured value at the i^{th} plot in the h^{th} LiDAR height class, $\bar{y}_h$ is the average of field-measured values at in the h^{th} LiDAR height class, h is the number of LiDAR height classes, and n_h is the number of sample plots in the h^{th} LiDAR height class.

The error index (EI) (Reynolds et al., 1988) was used to evaluate the size distributions of DBH and HT, respectively. EI measures the proportions of absolute deviation between the predicted and field-measured number of trees to the total number of field-measured trees over the entire distribution. EI for a plot was computed as:

$$EI(\%) = \frac{\sum_{i=1}^{k}\left|n_{pi} - n_{oi}\right|}{N} \times 100 \tag{5}$$

where n_{pi} and n_{oi} are the predicted and observed numbers of trees, respectively, in DBH or HT class *i*. *k* is the number of DBH or HT classes. *N* is the total number of field-measured trees. The bin widths for classifying DBH and HT were 10 cm and 5 m, respectively.

The coefficient of determination measures (R^2) the proportion of variance in a response variable that is explained by predictor variables. It shows that how well a model's predictions fit the observed values of the response variable, which means the actual explanatory power of the model. The R^2 is calculated as:

$$R^2 = 1 - \frac{\sum_{i=1}^{n}(\hat{y}_i - y_i)^2}{\sum_{i=1}^{n}(y_i - \bar{y})^2} \tag{6}$$

where $\hat{y}_i$ is the prediction at the i^{th} plot, y_i is the field-measured value at the i^{th} plot, $\bar{y}$ is the average of field-measured values of the total sample plots, and n is the number of total sample plots.

Results

Effects of Smoothing, Resolution, and Height Cutoff on Tree-list Estimation

All the resolutions with pixel size less than 1 m produced too large of estimates of SPH and yielded unreasonable estimates of other attributes regardless of the amount of smoothing and the height cutoff. Hence, resolutions with pixel sizes less than 1 m were dropped from the analysis. The amount of smoothing in CHM had a relatively small effect on tree-list estimation compared to the other parameters. The smoothing generally decreased the variability of estimation among the resolutions at a given height cutoff or the height cutoffs at a given resolution. For this reason, we show the performance only from the smoothing of 3 by 3 neighbor window.

Most cases of the combinations of resolution and height cutoff resulted in the underestimation of SPH (Figure 3.3.A). Unbiased SPH estimations were found around 1.1 m to 2.0 m in CHM resolution with the various height cutoffs. Generally, a higher cutoff had a smaller absolute bias compared with the absolute bias from a lower cutoff. In terms of precision, Figure 3.3.B shows that a higher cutoff had a relatively consistent RMSE along with resolutions in CHM, which means that higher cutoffs were less affected by resolution for SPH estimation than lower cutoffs as also shown in Figure 3.3.A. The combinations of the finer resolutions (1.2 ~ 1.3 m) and the lower height cutoffs (p20 and p30) provided the lowest RMSEs. For overstory trees, the patterns of performance measures were similar to the patterns from the combined approach, but the best RMSEs were always found at height cutoff p90. For understory trees, bias and RMSE increased as height cutoff increased except for the bias at height cutoff p90.

For BA estimation, bias decreased as resolution decreased as shown in Figure 3.4.A. Unbiased BA estimation was achieved for the combination of several cutoffs from p10 to p60 and resolutions with pixel sizes larger than 1.7 m. RMSE in BA estimation also decreased as resolution decreased

(Figure 3.4.B). Lower cutoffs yielded lower RMSE. The lowest RMSEs appeared for resolutions around 1.8 ~ 2.0 m. For overstory trees, the differences in RMSE between height cutoffs at a given resolution were smaller than the differences for the combined approach except for 1.0 m resolution. For understory trees, bias and RMSE increased as height cutoff increased, and all height cutoffs overestimated SPH.

HT estimation had better performance than the other attributes. The pattern for HT estimation was different from the other attributes. The best accuracy in HT estimation was found with the cutoff at p50 or p60 for any resolution. The poorest accuracy in HT estimation appeared only for the cutoff p10, which had a worse bias for HT estimation as resolution decreased. HT estimation became unbiased as resolution decreased except with cutoffs p10 and p20 (Figure 3.5.A). Height cutoffs showing better RMSEs were p50 and p60 with the middle and higher resolutions, and p80 in the lower resolutions at any smoothing level. RMSE increased as resolution decreased especially for cutoffs p10, p20, and p30 (Figure 3.5.B). Bias and RMSE of HT estimation for overstory trees only by ITD increased as the resolution decreased. For understory trees, Bias and RMSE for HT estimation also increased as height cutoff increased.

For the lower resolutions, the lower cutoffs showed better DBH distribution estimation than the higher cutoffs, while it was the opposite with the higher resolutions (Figure 3.6.A). This pattern was also observed in HT distribution estimation. The best DBH distribution was found with cutoffs p30 and p40 and lower resolutions while cutoff p90 had the best DBH distribution for the higher resolutions. The HT distribution estimation, in most cases, had the better result with the lower cutoffs than the higher cutoffs (Figure 3.6.B). The cutoff p50 had the best performance in most cases, except p90 for 1 and 1.1 m resolutions, and p30 for 1.3 ~ 1.5 m resolutions. The resolutions with medium pixel sizes were better for estimating the HT distribution. For overstory trees, EI for DBH decreased as resolution decreased, and the lowest height cutoff p10 always yielded the best DBH distribution estimation at every resolution. DBH distribution estimation for overstory trees was poorer than for both overstory and understory trees. The best height cutoff for HT distribution of overstory trees estimation increased as resolution decreased. For understory trees, EIs for DBH and HT were reduced as height cutoff increased except for cutoff p80. Contrary to HT estimation for both overstory and understory trees by the combined approach, the best height cutoffs in the estimation of the HT distribution for overstory trees by ITD was for higher cutoffs from p60 to p80 except for resolutions higher than 1.4 m.

Compared to the combined approach for all trees or the ITD for overstory trees, NN imputation produced much lower biases for understory trees' SPH, BA, and HT (Figures 3.3, 3.4, and 3.5). The smallest biases for understory trees for SPH, BA and HT estimation were found at cutoffs p10, p20, and p40, respectively. The smallest RMSEs in the three attributes were observed only at cutoff p10.

Effects of Classification of Field Plots by LiDAR Height on Tree-list Estimation

The absolute and relative performance measures separated by LiDAR height class were calculated for each forest attribute estimated. The smallest group, class 1, showed distinct properties in those performances. For the absolute measures, such as bias and RMSE, lower LiDAR height classes, especially the lowest class, generally yielded comparable or better performances for BA and SPH than the higher classes. However, based on the relative measures, the lowest class had much poorer results. Similar patterns were found in EIs for DBH and HT as well. The effect of the amount of smoothing in CHM by LiDAR height class was relatively small. The performances by height cutoff in a given resolution were averaged for this section because it is better to show the general effect of height class on tree-list estimation performance. For SPH estimation (Figure 3.7), bias decreased as resolution decreased for every height class, but the resolutions showing unbiasedness varied among height classes. Lower height classes had larger variability in bias among resolutions than higher height class. Height class 1 had much larger RBias at higher resolutions than the other height classes. Larger RMSE occurred in height classes 1 through 6, and the largest RMSE was found in height class 3. RRMSE in height class 1 was largest at every resolution. Relatively larger RRMSEs at higher resolutions were observed in the taller height classes.

In BA estimation (Figure 3.8), biases in the taller height classes were generally larger than biases in the shorter height classes. This pattern was similar for RBias except for height class 1. RBias in height class 1 was larger than the other height classes at resolutions less than or equal to 2.3 m. Lower height classes generally had smaller RMSE than higher height classes, but height class 1 had a much larger RRMSE than the other height classes. Figure 3.9 shows the performance measures for HT estimation by height class. The pattern of HT estimation among height classes was different from the pattern of SPH and BA estimation. Height class 1 had comparable or better performance in bias, RBias, and RMSE. The primary difference in bias and RBias between class 1 and the other classes was that class 1 mainly underestimated HT while the other classes overestimated. RRMSE for HT in height class 1 had slightly larger values than RRMSE from other height classes. Estimated distributions of DBH and HT for height class 1 were much poorer than the distributions for the other classes. Except class 1, lower height classes showed better performance in EIs for both DBH and HT than higher height classes. Lower resolution generally had lower EIs (Figure 3.10).

Explanatory Power of Individual Tree Detection for Overstory Trees and Random Forest Nearest Neighbor Imputation for Understory Trees

Tables 3.5 through 3.7 show R^2s for SPH, BA and HT estimation for trees over a given height cutoff (overstory trees) via ITD by resolution of CHM and height cutoff with smoothing using a 3 by 3 window. For SPH estimation (Table 3.5), the best R^2 was found at resolutions between 1.2 m and 1.7 m for each height cutoff. Height cutoff p90 yielded the largest R^2, 0.501, and the best R^2 decreased as the height cutoff decreased. The lowest height cutoff p10 had negative R^2 at all the resolutions. BA estimation by ITD showed poor explanatory power for overstory (Table 3.6). Most combinations of resolutions and height cutoffs had negative R^2s, and the best R^2 was 0.338 with the resolution 2.0 m and the height cutoff p10. Larger height cutoffs, from p70 to p90 provided negative R^2 at every resolution. In HT estimation (Table 3.7), the 1.0 m resolution yielded the best R^2 at every height cutoff except p90. The middle height cutoffs, p50 or p60, had better R^2 than the other height cutoffs. Inferior explanatory power was found at height cutoffs p10 and p90. The explanatory power for HT estimation generally decreased as resolution increased.

Table 3.8 shows the explanatory power of RF NN imputation for trees under a given height cutoff (understory trees). For HT estimation, the R^2s were around 0.5. However, the R^2s for BA and SPH estimation were much poorer than the R^2s for HT estimation or even had negative values. For understory trees for each forest inventory attribute, the scatter plots of observed vs. predicted via RF NN imputation did not show any anomaly. The lower height cutoff we used, the more observations with zero values we had. The prediction results for those observations with zero values were inferior for every height cutoff.

Discussion

No single combination of smoothing, resolution and height cutoff was found to produce the best results for all performance measures (Table 3.9). Koch et al. (2014) and McGaughey (2016) also reported similar findings. Similarly, ITD's performance varied depending on the algorithm used to delineate trees in the CHM (Kaartinen et al., 2012). Differences in performance between the lowest LiDAR height class and the other classes were found based on both absolute and relative performance measures. Kaartinen et al. (2012) reported that the HT class did not generally impact the accuracy of HT estimation, but greater uncertainty was observed for ITD methods capable of finding small trees. According to Hopkinson et al. (2005), vegetation classes with short height, such as low shrub and aquatic vegetation, yielded the largest relative errors in canopy height estimation, whereas tall vegetation classes showed the largest absolute errors. The low level of penetration of LiDAR returns into the sub-canopy surface might be an essential reason for the high relative bias for low shrub and aquatic vegetation. For aquatic vegetation, it was also believed that the weak laser backscatter from the saturated ground caused the high relative bias. These results were very similar to ours although the smallest height class in our research almost exclusively consisted of trees.

As we reported above, resolutions with pixel sizes less than 1.0 m were dropped in the analysis because it yielded unreasonably large SPH estimations. Pouliot et al. (2002) claimed that in high-resolution imagery, tree detection and crown delineation became more complicated. This is because high-resolution imagery can display very detailed objects such as branches causing tree crowns to deviate from the conic shape. Thus, more tree crowns could be estimated at higher image resolutions. Conversely, in low-resolution imagery, it is more challenging to identify crown boundaries because they become less distinct. Another reason for our large SPH estimation might be data pits, which are height irregularities in a CHM. The function '*CanopyModel*' in FUSION used for generating CHMs in our study fills pixels without LiDAR point clouds using an eight-way search and a distance-weighted average (McGaughey, 2016). However, it might be difficult to avoid irregularities in height on a CHM if laser pulses used for our LiDAR data acquisition penetrated deeply into tree crowns causing large height variations within individual tree crowns (Persson et al., 2002). Image smoothing with various filters using mean, median, or Gaussian approaches have been applied to reduce data pits (Persson et al., 2002; Yu et al., 2011). In our study, the smoothing did not work well at resolutions with pixel sizes larger than 1.0 m although the smoothing using a 5 by 5 window showed smaller SPH estimation than no smoothing and the smoothing with a 3 by 3 window. A pit-free CHM proposed by Khosravipour et al. (2014) was found to improve the accuracy of tree detection based on either high or low-density LiDAR data; however, this approach could help solve our large SPH estimation at the finer resolutions.

The ratio of average crown diameter to image pixel size was proposed as a guide to determine an optimal image resolution for tree detection and crown delineation using digital camera imagery (Pouliot et al., 2002). With a small crown diameter to pixel size ratio, it is hard to have distinct crown boundaries in an image, resulting in under-segmentation. However, a large crown diameter to pixel size ratio might cause high variability within a crown in an image resulting in over-segmentation. Although our data from field surveys do not have information on crown diameter, there might be significant variations in the tree crowns considering the diversity of forest stands in our study area. This might be one of the reasons why the high CHM resolutions overestimated SPH in our study. Barnes et al. (2017) found that no single CHM resolution produced the best performance of ITD for both healthy and diseased larch trees, claiming that not only the tree crown size but also the maximum tree height governed an optimal size of CHM resolution. The performance of ITD with high-resolution CHMs (0.15 m) was best for plots with low maximum height (< 20 m), and the performance with low-resolution CHMs (0.5 m) was best for plots with high maximum height (> 30 m).

LiDAR point cloud density might be related to the optimal CHM resolution as with the tree crown diameter. With LiDAR data of high point cloud density, high CHM resolution could yield high within-crown variations on a CHM. Inversely, with LiDAR data of low point cloud density, low CHM resolution could produce less distinct crown boundaries making it difficult to identify tree crowns. The high CHM resolutions should have yielded good performance in that the LiDAR data used for this study had low point cloud density. However, the high diversity of forest stands in the study area might add more within-crown variations. Even though an optimal resolution of CHM was set based on the crown diameter to CHM resolution ratio, it should be noted that the performance of ITD was still affected by LiDAR point cloud density for trees with small DBH (< 20cm) as reported in Khosravipour et al. (2014).

The results of large SPH estimation are quite different from previous studies. Stereńczak et al. (2008) reported that the 0.25 and 0.5 m resolutions in CHMs were better than the 1.0 m resolution for estimating SPH through individual tree delineation based on a similar method to Heurich and Weinacker (2004). It was found that the number of detected trees decreased as the resolution of CHM decreased (Stereńczak et al., 2008), and this was also observed in our work, excluding height cutoffs p10 and p20. Smreček et al. (2018) showed very similar results to ours for SPH estimation based on ITD. At the highest resolution (0.5 m), the number of trees identified was hugely overestimated; the number of trees identified decreased as the CHM resolution decreased from 0.5 m to 2.0 m, as was the case in our study. The optimal resolutions for tree identification were 1.0 and 1.5 m depending on the sample plot. Smreček et al. (2018) claimed that this was because the CHM with 0.5 m resolution was too detailed. We observed many estimated trees from ITD with extreme small areas compared to their estimated heights. Those trees should have been removed from the estimated tree-list using an appropriate criterion. With this filtering process, overestimation at high resolutions would be decreased.

Most combinations of parameters resulted in underestimating SPH. According to Lindberg et al. (2010), ITD underestimates SPH because ITD often misses trees below dominant trees or recognizes trees close to each other as one tree. It was expected that there would be more underestimation as pixel size increased. The larger pixel size we have, the more aggregated information we would get, so lower resolution also would result in underestimating SPH. For this reason, estimates of BA decreased as resolution decreased. The approach of Lindberg et al. (2010) could give an improvement for estimating overstory trees for our study. Considering that most of the combinations for overstory trees by ITD produced negative biases in SPH estimation in our study (Figure 3.3.A), estimating of the number of trees per segment would improve the negative biases in SPH estimation by increasing the number of detected trees.

Performance measures for HT estimation were better than measures for the other variables tested. This might be because LiDAR directly measures heights of target objects, so there is less uncertainty in height estimation than other attribute estimation. According to Stereńczak et al. (2008), there was no difference between the three resolutions (0.25, 0.5, and 1.0 m) in CHM for HT estimation. For understory trees, biases in HT estimation less than 0.15 m in absolute value were produced by RF NN at every height cutoff. The higher the height cutoff applied, the larger the RMSE obtained. This is attributed to the fact that RF NN will have more and larger trees to estimate with higher height cutoffs.

While RF NN imputation showed better performance in estimating SPH than the combined approach and tree segmentation (Figure 3.3 to Figure 3.6), this does not mean that RF NN imputation is better than the combined approach or ITD. It is because the target trees for those two methods are different from each other (tall trees above a height cutoff for ITD and short trees below the height cutoff for RF NN imputation). Therefore, the values dealt with in RF NN imputation were smaller than ITD. Based on relative measures not included on this manuscript, RF NN imputation was generally better in RBias, comparable in EIs, and worse in RRMSE.

The errors for BA and mean HT estimation in taller height classes were larger than in shorter height classes contradicting the fact that airborne LiDAR has difficulty in detecting understory vegetation. This might be because large trees have larger DBH and HT than small trees. To offset this potential issue, relative performance measures such as RBias and RRMSE were calculated. These relative measures revealed that the performance of estimation in shorter height classes was poorer than for the trees in taller height classes. Stereńczak et al. (2008) found a similar phenomenon for young stands.

There was also no single combination of the three parameters tested for explanatory powers that proved best overall. While HT estimation was good, estimation of BA and SPH were poor. Especially, BA estimation was very poor. BA estimation for overstory trees via ITD had more uncertainty sources than the other attributes, including SPH estimation and subsequent prediction of DBH for each detected individual tree (estimated HT used to predict DBH provided additional uncertainty source to the DBH prediction). These uncertainty sources might partially explain the poor performance in BA estimation. Utilizing the limited information in LiDAR data might affect the poor performance for the explanatory powers. We used CHM-based ITD; this method has limitation summarizing LiDAR point clouds within a range of cell into one cell height value regardless of generating a pit-free CHM. Instead, 3D ITD methods have been recently studied using information in LiDAR as much as possible (Kandare et al., 2016). However, the 3D ITD methods required more complex algorithms to implement, and also processing time could be a new parameter to consider (Pirotti et al., 2017).

It is well known that it is difficult to estimate characteristics of understory vegetation. Eskelson et al. (2011) used beta regression to estimate percent shrub cover, and it yielded poor explanatory power. Rahman and Gorte (2008) developed a tree filtering technique to separate dominant tree and undergrowth vegetation, but it was found difficult to separate undergrowth vegetation very close to a tree using the filtering. Liu et al. (2013) suggested a method to extract individual tree crowns from airborne LiDAR in residential areas showing promising applications, but also reported that small trees were omitted if there were an only small number of points representing them in the dataset. Our results for understory trees via RF NN were not good (Table 3.8). To improve NN estimation with LiDAR data having low point cloud density, we investigated many LiDAR metrics such as metrics from LiDAR point clouds under several height cutoffs as Wing et al. (2012) proposed to estimate understory vegetation cover with airborne LiDAR. Some of the metrics from understory point clouds were selected for NN imputation (Table 3.4). However, it did not greatly improve the performance of NN imputation compared to NN imputation without those metrics (not presented here). This might fundamentally be because our LiDAR data lacked information on understory vegetation.

In NN imputation, one of the critical parameters is the selection of a number of neighbors for imputation modeling or distance metrics used to measure the similarity between the reference and target plot using auxiliary variables (Eskelson et al., 2009). While their result varied among different forest types, Strunk et al. (2017) reported that k = 3 and Mahalanobis distance metric produced better performance over other NN strategies in estimating tree-lists. In this study, we only used k = 1 and RF as a distance metric in NN modeling. Combination of the two parameters needs to be examined for understory vegetation. In addition to these two factors, implementing variable selection procedure for NN imputation to each LiDAR height class could have the potential to improve NN modeling performance.

Compared to the results of HT estimation, results related to DBH estimation such as BA and EI for DBH showed poorer performance. It is known that predicting tree-level DBH from height-derived metrics has considerable variability (Maltamo and Gobakken, 2014). Kaartinen et al. (2012) reported that estimation of DBH based on HT and crown size would have considerable uncertainty because allometric equations used for estimating DBH are sensitive to errors in input data such as the size of tree crown or HT. Another potential reason is the dead trees in the field data. The Pearson's correlation coefficients between the field-measured HT and DBH for live and dead trees are 0.771 and 0.212, respectively. Even though the dead trees account for only 8.8 % of a total number of field-measured trees, appropriate handling for dead trees would give opportunities to improve estimating tree-lists.

The scanning angle is another parameter to consider for LiDAR projects (Gatziolis and Andersen, 2008). If the scanning angle increases, it facilitates changes in pulse propagation direction and increases the distance the pulse moves through the canopy. The change in pulse direction and the increased distance are related to LiDAR data artifacts such as returns below the ground. Therefore, with a wide scanning angle, LiDAR data might have more data artifacts than with a narrow-angle. Additionally, these data artifacts could increase when data acquisition is carried out on a slope, as an off-nadir scanning angle increases on the slope (Gatziolis and Andersen, 2008).

39.4 % of our field plots had slopes more than 30° based on digital terrain models from the study site. Khosravipour et al. (2015) showed that normalized LiDAR point clouds could distort tree locations detected from CHM and height estimation depending on the steepness of slope and crown shape. For the slope of more than 30°, 44.6 % of correctly detected trees with wider and irregular crown shapes were affected by the horizontal and vertical displacements. They suggested using a non-normalized CHM to avoid the adverse effect of the distortion by steep slopes, especially in a heterogeneous forest with multiple species. The slope was also found to affect the ABA approach by distorting heights of LiDAR point clouds (Hansen et al., 2017). They proposed two methods, Procrustean transformation and histogram matching, to counter the distortion of LiDAR point clouds on slope terrain for extracting LiDAR metrics. These point cloud distortions by slope terrain could worsen our results for both overstory and understory estimations.

Another issue is that there was the time lag between LiDAR acquisition and field surveys. This might have the potential source of error, particularly for younger fast-growing stands. Also, there were seasonal differences in the LiDAR acquisition dates (e.g., April through June in the spring, June through August in the summer, and September and October in the fall). According to Gatziolis and Andersen (2008), the seasonal differences can induce considerable variability in canopy penetrability by LiDAR pulses especially for deciduous forests (e.g., leaf-on and leaf-off conditions) and weather-related limitations. The variability in canopy penetrability might increase uncertainty in modeling forest attributes, and the weather-related limitations could make it difficult to keep the quality of LiDAR data consistent over our whole study area. Time windows, part of LiDAR data acquisition considerations in Gatziolis and Andersen (2008), should be carefully planned according to project objectives.

Conclusion

We proposed an approach to combine ITD and ABA to generate a tree-list using airborne LiDAR data and field measured data. The approach aimed to compensate for the disadvantage of LiDAR data and ITD in estimating understory trees, and to keep the strength of ITD in estimating overstory trees in tree-level. The selected parameters, smoothing, resolution and height cutoff, were examined to determine how they affected the performance of the proposed approach. There was no single combination of the three parameters that provides the best estimation results for all the forest attributes in this study. For each attribute, the best results depended on different combinations of those parameters. This is concurrent with what Koch et al. (2014) and McGaughey (2016) reported. It would be practical and useful to determine how to automatically find the optimal combinations of those parameters across the forest landscape using remote sensing data. In addition to the three parameters tested in the present study, the automation for the optimal combinations would require considering additional parameters such as forest types, tree species, tree-size parameters (tree crown width or maximum tree height) and topography.

There are several topics for further study to improve the combined approach. A denser point cloud data would have more information on both overstory and understory vegetation in a forest, thus could increase the combined approach's performance. The algorithm used to generate a CHM and to delineate trees on the CHM is another critical parameter in ITD. Comparison of different algorithms for processing the CHM is an active area of research. Estimating the number of trees per crown segment would help obtain unbiased SPH estimation. A point cloud based ITD method could lead to improvement by utilizing more information in LiDAR data. A minimum crown area by ITD should be examined so that tiny crown would not degrade the quality of the predicted tree-lists. The effect of slope on CHM generation and LiDAR metrics extraction need to be considered for better estimation. Fusing ITD and ABA to predict overstory and understory vegetation shown in this research indicates that forest analysts can benefit from the predictive abilities of the imputation approach and the quality information provided by LiDAR. In that, the approach presented herein can be sufficient for strategic inventory purposes.

References

Barnes, C., Balzter, H., Barrett, K., Eddy, J., Milner, S., and Suárez, J. 2017. Individual Tree Crown Delineation from Airborne Laser Scanning for Diseased Larch Forest Stands. *Remote Sensing* 9(3): 231.

Breidenbach, J., Næsset, E., Lien, V., Gobakken, T., and Solberg, S. 2010b. 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(4): 911-924. doi: http://dx.doi.org/10.1016/j.rse.2009.12.004.

Breiman, L. 2001. Random forests. *Machine Learning* 45(1): 5-32. doi: 10.1023/a:1010933404324.

Crookston, N.L., and Finley, A.O. 2008. yaImpute: An R package for kNN imputation. *Journal of Statistical Software* 23(10): 16.

Eskelson, B.N.I., Madsen, L., Hagar, J.C., and Temesgen, H. 2011. Estimating Riparian Understory Vegetation Cover with Beta Regression and Copula Models. *Forest Science* 57(3): 212-221.

Eskelson, B.N.I., Temesgen, H., Lemay, V., Barrett, T.M., Crookston, N.L., and Hudak, A.T. 2009. The roles of nearest neighbor methods in imputing missing data in forest inventory and monitoring databases. *Scandinavian Journal of Forest Research* 24(3): 235-246.

Gatziolis, D., and Andersen, H.-E. 2008. A guide to LIDAR data acquisition and processing for the forests of the Pacific Northwest. US Department of Agriculture, Forest Service, Pacific Northwest Research Station. 32 p, Gen. Tech. Rep. PNW-GTR-768. Portland, OR. p. 32.

Gobakken, T., and Næsset, E. 2004. Estimation of diameter and basal area distributions in coniferous forest by means of airborne laser scanner data. *Scandinavian Journal of Forest Research* 19(6): 529-542. doi: 10.1080/02827580410019454.

Hamraz, H., Contreras, M.A., and Zhang, J. 2017. Vertical stratification of forest canopy for segmentation of understory trees within small-footprint airborne LiDAR point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing* 130: 385-392. doi: https://doi.org/10.1016/j.isprsjprs.2017.07.001.

Hansen, E., Ene, L., Gobakken, T., Ørka, H., Bollandsås, O., and Næsset, E. 2017. Countering Negative Effects of Terrain Slope on Airborne Laser Scanner Data Using Procrustean Transformation and Histogram Matching. *Forests* 8(10): 401.

Hawbaker, T.J., Keuler, N.S., Lesak, A.A., Gobakken, T., Contrucci, K., and Radeloff, V.C. 2009. Improved estimates of forest vegetation structure and biomass with a LiDAR-optimized sampling design. *Journal of Geophysical Research: Biogeosciences* 114(G2): G00E04. doi: 10.1029/2008JG000870.

Heurich, M., and Weinacker, H. 2004. Automated tree detection and measurement in temperate forests of central Europe using laser scanning data. In *ISPRS working group on Laser-Scanners for Forest and Landscape Assessment*. Edited by M. Thies and B. Koch and H. Spiecker and H. Weinacker, Freiburg, Germany. pp. 198-203.

Hopkinson, C., Chasmer, L.E., Sass, G., Creed, I.F., Sitar, M., Kalbfleisch, W., and Treitz, P. 2005. Vegetation class dependent errors in lidar ground elevation and canopy height estimates in a boreal wetland environment. *Canadian Journal of Remote Sensing* 31(2): 191-206. doi: 10.5589/m05-007.

Kaartinen, H., Hyyppä, J., Yu, X., Vastaranta, M., Hyyppä, H., Kukko, A., Holopainen, M., Heipke, C., Hirschmugl, M., Morsdorf, F., Næsset, E., Pitkänen, J., Popescu, S., Solberg, S., Wolf, B.M., and Wu, J.-C. 2012. An International Comparison of Individual Tree Detection and Extraction Using Airborne Laser Scanning. *Remote Sensing* 4(4): 950.

Kandare, K., Ørka, H.O., Chan, J.C.-W., and Dalponte, M. 2016. Effects of forest structure and airborne laser scanning point cloud density on 3D delineation of individual tree crowns. *European Journal of Remote Sensing* 49(1): 337-359. doi: 10.5721/EuJRS20164919.

Khosravipour, A., Skidmore, A.K., Isenburg, M., Wang, T., and Hussin, Y.A. 2014. Generating pit-free canopy height models from airborne lidar. *Photogrammetric Engineering & Remote Sensing* 80(9): 863-872.

Khosravipour, A., Skidmore, A.K., Wang, T., Isenburg, M., and Khoshelham, K. 2015. Effect of slope on treetop detection using a LiDAR Canopy Height Model. *ISPRS Journal of*

Photogrammetry and Remote Sensing 104: 44-52. doi: https://doi.org/10.1016/j.isprsjprs.2015.02.013.

Koch, B., Heyder, U., and Weinacker, H. 2006. Detection of Individual Tree Crowns in Airborne Lidar Data. *Photogrammetric Engineering & Remote Sensing* 72(4): 357-363. doi: 10.14358/PERS.72.4.357.

Koch, B., Kattenborn, T., Straub, C., and Vauhkonen, J. 2014. Segmentation of Forest to Tree Objects. In *Forestry Applications of Airborne Laser Scanning: Concepts and Case Studies*. Edited by M. Maltamo and E. Næsset and J. Vauhkonen. Springer Netherlands, Dordrecht. pp. 89-112.

Liaw, A., and Wiener, M. 2002. Classification and regression by randomForest. *R news* 2(3): 18-22.

Lindberg, E., and Hollaus, M. 2012. Comparison of methods for estimation of stem volume, stem number and basal area from airborne laser scanning data in a hemi-boreal forest. *Remote Sensing* 4(4): 1004-1023.

Lindberg, E., and Holmgren, J. 2017. Individual Tree Crown Methods for 3D Data from Remote Sensing. *Current Forestry Reports*: 1-13. doi: 10.1007/s40725-017-0051-6.

Lindberg, E., Holmgren, J., Olofsson, K., Wallerman, J., and Olsson, H. 2010. Estimation of tree lists from airborne laser scanning by combining single-tree and area-based methods. *International Journal of Remote Sensing* 31(5): 1175-1192. doi: 10.1080/01431160903380649.

Lindberg, E., Holmgren, J., Olofsson, K., Wallerman, J., and Olsson, H. 2013. Estimation of Tree Lists from Airborne Laser Scanning Using Tree Model Clustering and k-MSN Imputation. *Remote Sensing* 5(4): 1932.

Liu, J., Shen, J., Zhao, R., and Xu, S. 2013. Extraction of individual tree crowns from airborne LiDAR data in human settlements. *Mathematical and Computer Modelling* 58(3): 524-535. doi: https://doi.org/10.1016/j.mcm.2011.10.071.

Maltamo, M., Eerikäinen, K., Pitkänen, J., Hyyppä, J., and Vehmas, M. 2004. Estimation of timber volume and stem density based on scanning laser altimetry and expected tree size distribution functions. *Remote Sensing of Environment* 90(3): 319-330. doi: http://dx.doi.org/10.1016/j.rse.2004.01.006.

Maltamo, M., and Gobakken, T. 2014. Predicting Tree Diameter Distributions. In *Forestry Applications of Airborne Laser Scanning: Concepts and Case Studies*. Edited by M. Maltamo and E. Næsset and J. Vauhkonen. Springer Netherlands, Dordrecht. pp. 177-191.

Maltamo, M., Tokola, T., and Lehikoinen, M. 2003. Estimating Stand Characteristics by Combining Single Tree Pattern Recognition of Digital Video Imagery and a Theoretical Diameter Distribution Model. *Forest Science* 49(1): 98-109.

McGaughey, R.J. 2016. *FUSION/LDV: Software for LIDAR Data Analysis and Visualization*, Forest Service. Pacific Northwest research station, United States department of agriculture.

Næsset, E., and Gobakken, T. 2008. Estimation of above- and below-ground biomass across regions of the boreal forest zone using airborne laser. *Remote Sensing of Environment* 112(6): 3079-3090. doi: http://dx.doi.org/10.1016/j.rse.2008.03.004.

Persson, A., Holmgren, J., and Soderman, U. 2002. Detecting and measuring individual trees using an airborne laser scanner. *Photogrammetric Engineering and Remote Sensing* 68(9): 925-932.

Pippuri, I., Kallio, E., Maltamo, M., Peltola, H., and Packalén, P. 2012. Exploring horizontal area-based metrics to discriminate the spatial pattern of trees and need for first thinning using airborne laser scanning. *Forestry* 85(2): 305-314. doi: 10.1093/forestry/cps005.

Pirotti, F., Kobal, M., and Roussel, J.R. 2017. A COMPARISON OF TREE SEGMENTATION METHODS USING VERY HIGH DENSITY AIRBORNE LASER SCANNER DATA. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci*. XLII-2/W7: 285-290. doi: 10.5194/isprs-archives-XLII-2-W7-285-2017.

Pouliot, D.A., King, D.J., Bell, F.W., and Pitt, D.G. 2002. Automated tree crown detection and delineation in high-resolution digital camera imagery of coniferous forest regeneration. *Remote Sensing of Environment* 82(2): 322-334. doi: https://doi.org/10.1016/S0034-4257(02)00050-0.

R Core Team. 2017. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rahman, M.Z.A., and Gorte, B. 2008. Tree filtering for high density Airborne LiDAR data. In *Proceedings of SilviLaser 2008: 8th International Conference on LiDAR Applications in Forest Assessment and Inventory*, Edinburgh, UK. pp. 544-553.

Reynolds, M.R., Burk, T.E., and Huang, W.-C. 1988. Goodness-of-Fit Tests and Model Selection Procedures for Diameter Distribution Models. *Forest Science* 34(2): 373-399.

Romero-Zaliz, R., and Reinoso-Gordo, J.F. 2018. An Updated Review on Watershed Algorithms. In *Soft Computing for Sustainability Science*. Edited by C. Cruz Corona. Springer International Publishing, Cham. pp. 235-258.

Smreček, R., Michnová, Z., Sačkov I, Danihelová, Z., Levická, M., and Tuček, J. 2018. Determining basic forest stand characteristics using airborne laser scanning in mixed forest stands of Central Europe. *iForest - Biogeosciences and Forestry* 11(1): 181-188. doi: 10.3832ifor2520-010.

Stereńczak, K., Będkowski, K., and Weinacker, H. 2008. Accuracy of crown segmentation and estimation of selected trees and forest stand parameters in order to resolution of used DSM and nDSM models generated from dense small footprint LIDAR data. In *ISPRS Congress*, Beijing, China. pp. 27-32.

Strunk, J., Gould, P., Packalen, P., Poudel, K., Andersen, H.-E., and Temesgen, H. 2017. An Examination of Diameter Density Prediction with k-NN and Airborne Lidar. *Forests* 8(11): 444.

Takahashi, T., Yamamoto, K., Miyachi, Y., Senda, Y., and Tsuzuku, M. 2006. The penetration rate of laser pulses transmitted from a small-footprint airborne LiDAR: a case study in closed canopy, middle-aged pure sugi (Cryptomeria japonica D. Don) and hinoki cypress (Chamaecyparis obtusa Sieb. et Zucc.) stands in Japan. *Journal of Forest Research* 11(2): 117-123. doi: 10.1007/s10310-005-0189-0.

Temesgen, H., LeMay, V.M., Froese, K.L., and Marshall, P.L. 2003. Imputing tree-lists from aerial attributes for complex stands of south-eastern British Columbia. *Forest Ecology and Management* 177(1–3): 277-285. doi: http://dx.doi.org/10.1016/S0378-1127(02)00321-3.

Vauhkonen, J., Ene, L., Gupta, S., Heinzel, J., Holmgren, J., Pitkänen, J., Solberg, S., Wang, Y., Weinacker, H., Hauglin, K.M., Lien, V., Packalén, P., Gobakken, T., Koch, B., Næsset, E., Tokola, T., and Maltamo, M. 2012. Comparative testing of single-tree detection algorithms under different types of forest. *Forestry: An International Journal of Forest Research* 85(1): 27-40. doi: 10.1093/forestry/cpr051.

Vauhkonen, J., Maltamo, M., McRoberts, R.E., and Næsset, E. 2014. Introduction to Forestry Applications of Airborne Laser Scanning. In *Forestry Applications of Airborne Laser Scanning: Concepts and Case Studies*. Edited by M. Maltamo and E. Næsset and J. Vauhkonen. Springer Netherlands, Dordrecht. pp. 1-16.

Vincent, L., and Soille, P. 1991. Watersheds in digital spaces: an efficient algorithm based on immersion simulations. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 13(6): 583-598. doi: 10.1109/34.87344.

Wiggins, H.L. 2017. The influence of tree height on LiDAR's ability to accurately characterize forest structure and spatial pattern across reference landscapes. In *Department of Ecosystem and Conservation Sciences*. University of Montana.

Wing, B.M., Ritchie, M.W., Boston, K., Cohen, W.B., Gitelman, A., and Olsen, M.J. 2012. Prediction of understory vegetation cover with airborne lidar in an interior ponderosa pine forest. *Remote Sensing of Environment* 124(Supplement C): 730-741. doi: https://doi.org/10.1016/j.rse.2012.06.024.

Yu, X., Hyyppä, J., Holopainen, M., and Vastaranta, M. 2010. Comparison of Area-Based and Individual Tree-Based Methods for Predicting Plot-Level Forest Attributes. *Remote Sensing* 2(6): 1481.

Yu, X., Hyyppä, J., Vastaranta, M., Holopainen, M., and Viitala, R. 2011. Predicting individual tree attributes from airborne laser point clouds based on the random forests technique. *ISPRS Journal of Photogrammetry and Remote Sensing* 66(1): 28-37. doi: http://dx.doi.org/10.1016/j.isprsjprs.2010.08.003.

Table 3.1: LiDAR survey specifications

Attribute	Description
Sensor	Leica ALS50 Phase II
Survey altitude	900 m (flown at 900 meters above ground level)
Pulse rate	> 105 kHz (> 105,000 laser pulse per second)
Pulse mode	Single
Mirror scan rate	52.5 Hz
Field of view	28° (±14° from nadir*)
Roll compensated	Up to 20°
Overlap	100 % (50 % side-lap)

* Point on the ground vertically beneath the laser sensor on the aircraft.

Table 3.2: Plot-level summary statistics of attributes from the field measurements

Attribute	Minimum	Maximum	Median	Mean	SD*
BA (m^2/ha)	0.0	236.5	50.3	61.9	45.9
HT (m)	0.0	63.3	23.4	24.6	10.1
SPH (stems/ha)	0.0	1,462.9	316.3	354.1	222.1

*Standard deviation.

Table 3.3: Tree-level summary statistics from the field measurements

Attribute	Minimum	Maximum	Median	Mean	SD*
DBH (cm)	14.0	266.2	26.9	37.4	28.7
HT (m)	0.3	88.4	19.51	23.5	14.2

*Standard deviation.

Table 3.4: Selected predictor variables for RF NN imputation

Metrics	Min	Max	Mean	SD	Description
sqrt_mean (m)	2.4	63.1	27.7	13.5	LiDAR height quadratic mean
CHM_SD (m)	1.1	30.1	10.2	6.2	Height standard deviation of rasterized CHM
Vol_3D (m^3)	768.9	30,258.3	12,461.5	6,645.6	Volume of the region between rasterized CHM and ground
AShape.4 (m^3)	792.7	20,231.1	8,627.4	3,938.8	3D alpha shape with alpha value of 4
mode_30th (m)	1.0	53.7	12.8	11.9	LiDAR height mode from the point clouds less than LiDAR height 30th percentile
SD_30th (m)	0.1	17.6	4.9	3.7	LiDAR height standard deviation from the point clouds less than LiDAR height 30th percentile
sqrt_10 (m)	1.8	8.8	5.6	1.2	LiDAR height quadratic mean from the point clouds under 10 m
p.a.2 (%)	8.4	204.8	125.6	34.9	Percentage of first returns above height of 2 m
p.u.5 (%)	0.0	98.6	15.5	20.5	Percentage of first returns under height of 5 m
p.a.15 (%)	0.0	99.6	64.0	32.7	$\frac{\text{Number of total first returns above 15 m}}{\text{Number of total first returns above 2 m}} \times 100$
p.a.10th (%)	12.0	100.0	86.0	16.2	$\frac{\text{Number of total first returns above LiDAR height } 10^{\text{th}} \text{ percentile}}{\text{Number of total first returns above 2m}} \times 100$

Table 3.5: Explanatory power (R^2) of SPH via ITD for trees taller than given height cutoffs with the smoothing of 3 by 3 neighbor window

Height cutoff	Resolution (m)															
	1.0	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0	2.1	2.2	2.3	2.4	2.5
p10	-0.100	-0.020	-0.009	-0.039	-0.104	-0.199	-0.267	-0.343	-0.418	-0.514	-0.566	-0.644	-0.710	-0.779	-0.832	-0.904
p20	-0.115	0.048	0.106	0.107	0.069	-0.008	-0.059	-0.124	-0.192	-0.280	-0.326	-0.399	-0.458	-0.523	-0.574	-0.641
p30	-0.134	0.084	0.170	0.196	0.180	0.122	0.086	0.029	-0.027	-0.110	-0.147	-0.216	-0.267	-0.332	-0380	-0.439
p40	-0.139	0.106	0.215	0.262	0.262	0.223	0.196	0.156	0.104	0.033	0.000	-0.062	-0.103	-0.167	-0.210	-0.261
p50	-0.144	0.124	0.253	0.311	0.319	0.297	0.285	0.252	0.207	0.142	0.113	0.054	0.018	-0.044	-0.081	-0.129
p60	-0.131	0.138	0.274	0.347	0.364	0.356	0.357	0.333	0.296	0.234	0.211	0.158	0.127	0.066	0.039	-0.011
p70	-0.057	0.198	0.321	0.386	0.409	0.403	0.418	0.400	0.369	0.319	0.301	0.252	0.227	0.171	0.148	0.101
p80	0.048	0.264	0.371	0.421	0.452	0.448	0.464	0.460	0.434	0.396	0.390	0.345	0.325	0.276	0.255	0.217
p90	0.236	0.376	0.431	0.464	0.497	0.480	0.500	0.501	0.483	0.454	0.455	0.423	0.417	0.372	0.353	0.330

Table 3.6: Explanatory power (R^2) of BA via ITD for trees taller than given height cutoffs with the smoothing of 3 by 3 neighbor window

Height cutoff	Resolution (m)															
	1.0	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0	2.1	2.2	2.3	2.4	2.5
p10	-2.421	-1.008	-0.443	-0.197	0.033	0.178	0.240	0.283	0.326	0.284	0.338	0.294	0.289	0.322	0.271	0.280
p20	-2.947	-1.320	-0.665	-0.378	-0.102	0.078	0.168	0.220	0.279	0.242	0.314	0.270	0.273	0.318	0.263	0.286
p30	-3.593	-1.717	-0.972	-0.628	-0.300	-0.065	0.049	0.118	0.197	0.166	0.265	0.217	0.232	0.287	0.229	0.266
p40	-4.306	-2.159	-1.303	-0.904	-0.516	-0.219	-0.088	0.011	0.106	0.084	0.203	0.161	0.182	0.256	0.197	0.253
p50	-5.072	-2.663	-1.706	-1.244	-0.794	-0.432	-0.273	-0.141	-0.035	-0.052	0.095	0.055	0.089	0.181	0.132	0.202
p60	-6.058	-3.323	-2.242	-1.740	-1.201	-0.752	-0.561	-0.387	-0.253	-0.257	-0.068	-0.108	-0.051	0.060	0.011	0.098
p70	-6.510	-3.724	-2.598	-2.104	-1.525	-1.037	-0.832	-0.616	-0.488	-0.488	-0.254	-0.295	-0.233	-0.118	-0.133	-0.040
p80	-7.105	-4.372	-3.169	-2.752	-2.116	-1.574	-1.377	-1.092	-0.990	-0.928	-0.641	-0.707	-0.615	-0.479	-0.501	-0.354
p90	-6.662	-4.746	-3.638	-3.239	-2.737	-2.125	-2.020	-1.748	-1.648	-1.613	-1.263	-1.386	-1.245	-1.045	-1.084	-0.872

Table 3.7: Explanatory power (R^2) of HT via ITD for trees taller than given height cutoffs with the smoothing of 3 by 3 neighbor window

Height cutoff	Resolution (m)															
	1.0	1.1	1.2	1.3	1.4	1.5	1.6	1.7	1.8	1.9	2.0	2.1	2.2	2.3	2.4	2.5
p10	0.073	0.022	-0.052	-0.156	-0.223	-0.298	-0.347	-0.427	-0.504	-0.621	-0.639	-0.733	-0.794	-0.844	-0.952	-1.040
p20	0.514	0.481	0.447	0.385	0.364	0.300	0.283	0.231	0.192	0.113	0.110	0.054	0.013	-0.014	-0.070	-0.131
p30	0.700	0.681	0.661	0.629	0.619	0.578	0.567	0.536	0.523	0.469	0.463	0.434	0.407	0.389	0.349	0.307
p40	0.784	0.770	0.760	0.738	0.736	0.718	0.709	0.695	0.685	0.650	0.639	0.625	0.620	0.597	0.569	0.554
p50	0.804	0.797	0.788	0.772	0.771	0.762	0.751	0.744	0.733	0.710	0.708	0.694	0.686	0.676	0.655	0.652
p60	0.775	0.768	0.756	0.748	0.751	0.745	0.738	0.728	0.723	0.713	0.702	0.699	0.687	0.691	0.667	0.671
p70	0.676	0.673	0.670	0.659	0.663	0.658	0.652	0.641	0.637	0.628	0.621	0.611	0.608	0.613	0.592	0.598
p80	0.453	0.450	0.432	0.424	0.414	0.419	0.411	0.405	0.393	0.392	0.387	0.382	0.383	0.379	0.367	0.357
p90	0.028	0.033	0.010	0.014	-0.001	-0.001	-0.008	-0.015	-0.001	-0.021	-0.017	-0.003	-0.039	-0.006	-0.038	-0.024

Table 3.8: Performance measures of RF NN imputation by inventory attributes for trees shorter than given height cutoffs

Height cutoff	SD_{SPH}*	$Bias_{SPH}$	$RMSE_{SPH}$	R^2_{SPH}	SD_{BA}*	$Bias_{BA}$	$RMSE_{BA}$	R^2_{BA}	SD_{HT}*	$Bias_{HT}$	$RMSE_{HT}$	R^2_{HT}
p10	89.20	0.73	87.99	0.03	9.13	-0.25	10.56	-0.34	6.09	-0.15	4.15	0.54
p20	115.32	1.72	110.63	0.08	11.35	0.08	13.26	-0.37	7.46	-0.05	4.88	0.57
p30	133.73	5.01	129.52	0.06	13.77	0.38	15.41	-0.25	8.11	0.10	5.69	0.51
p40	146.14	6.07	145.21	0.01	15.95	0.51	17.79	-0.25	8.58	0.04	6.12	0.49
p50	156.35	7.83	158.25	-0.03	18.42	1.00	20.59	-0.25	8.89	0.05	6.23	0.51
p60	166.25	8.42	168.72	-0.03	22.11	1.22	24.12	-0.19	9.05	-0.10	6.32	0.51
p70	177.05	9.04	179.29	-0.03	25.18	1.41	26.86	-0.14	9.20	0.08	6.40	0.52
p80	187.56	10.80	193.97	-0.07	29.47	2.16	29.78	-0.02	9.24	0.05	6.84	0.45
p90	197.84	8.78	204.64	-0.07	35.78	2.20	34.17	0.09	9.55	0.08	7.27	0.42

* Standard deviation of field-measured inventory attribute under given height cutoffs

Table 3.9: Best performance for each assessment by estimation method

Method	Target	$Bias_{SPH}$	$RMSE_{SPH}$	$Bias_{BA}$	$RMSE_{BA}$	$Bias_{HT}$	$RMSE_{HT}$	EI_{DBH}	EI_{HT}
Combined	All	0.3079	212.2541	0.0233	35.0225	1.0967	8.4800	91.0333	91.2693
		3/2.1/p90*	No/1.3/p30	5/2.0/p30	3/2.0/p10	No/2.3/p50	No/2.0/p50	No/2.3/p40	No/1.9/p50
ITD	Overstory	0.5497	89.4945	0.0189	29.7892	1.7553	7.7367	96.6311	83.4631
		5/1.2/p50*	3/1.7/p90	No/1.8/p20	No/2.4/p60	No/1.0/p60	No/1.0/p40	No/2.3/p10	No/1.9/p70
NN	Understory	0.7256	87.9857	0.0829	10.5628	0.0420	4.1457	98.2339	101.8002
		p10†	p10	p20	p10	p40	p10	p90	p90
NN	All	-0.3958	217.6176	0.1438	36.2711	-0.4132	8.3843	91.6049	96.4568

* The first argument indicates the amount of smoothing, the second resolution in CHM, and the third percentile height cutoff for the combined method.

† This represents percentile height cutoff.

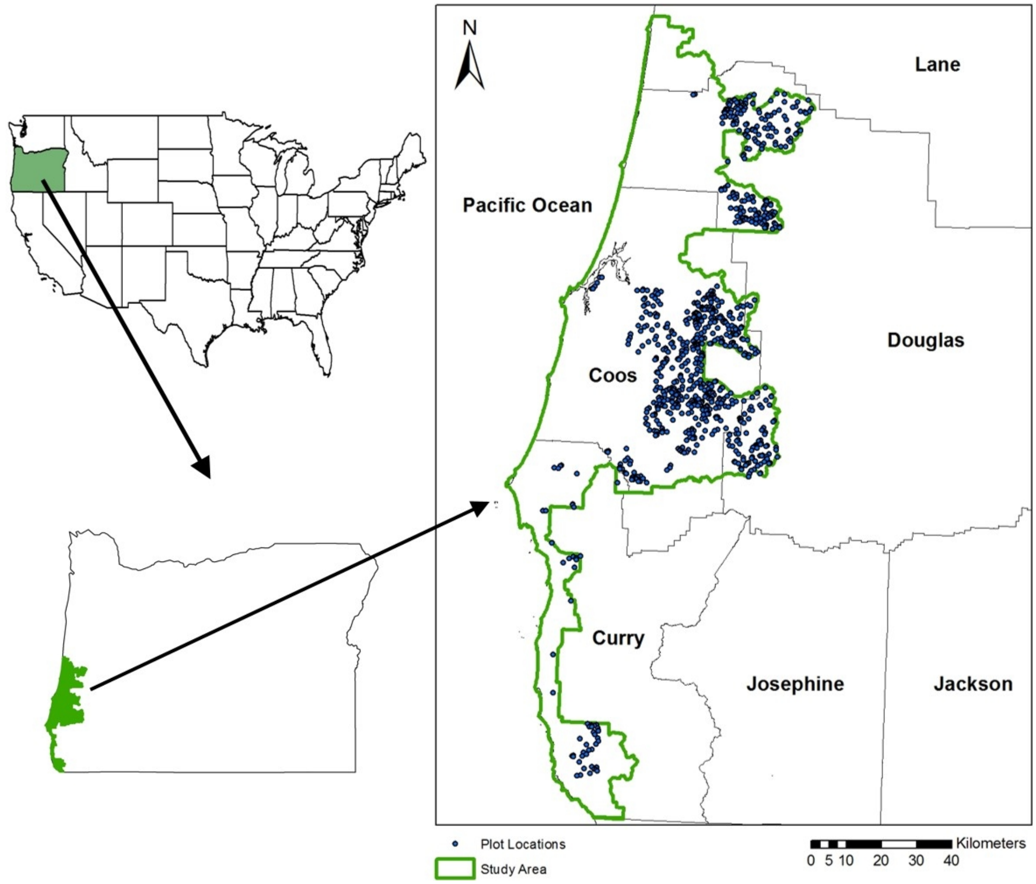


Figure 3.1. Map of study area and plots

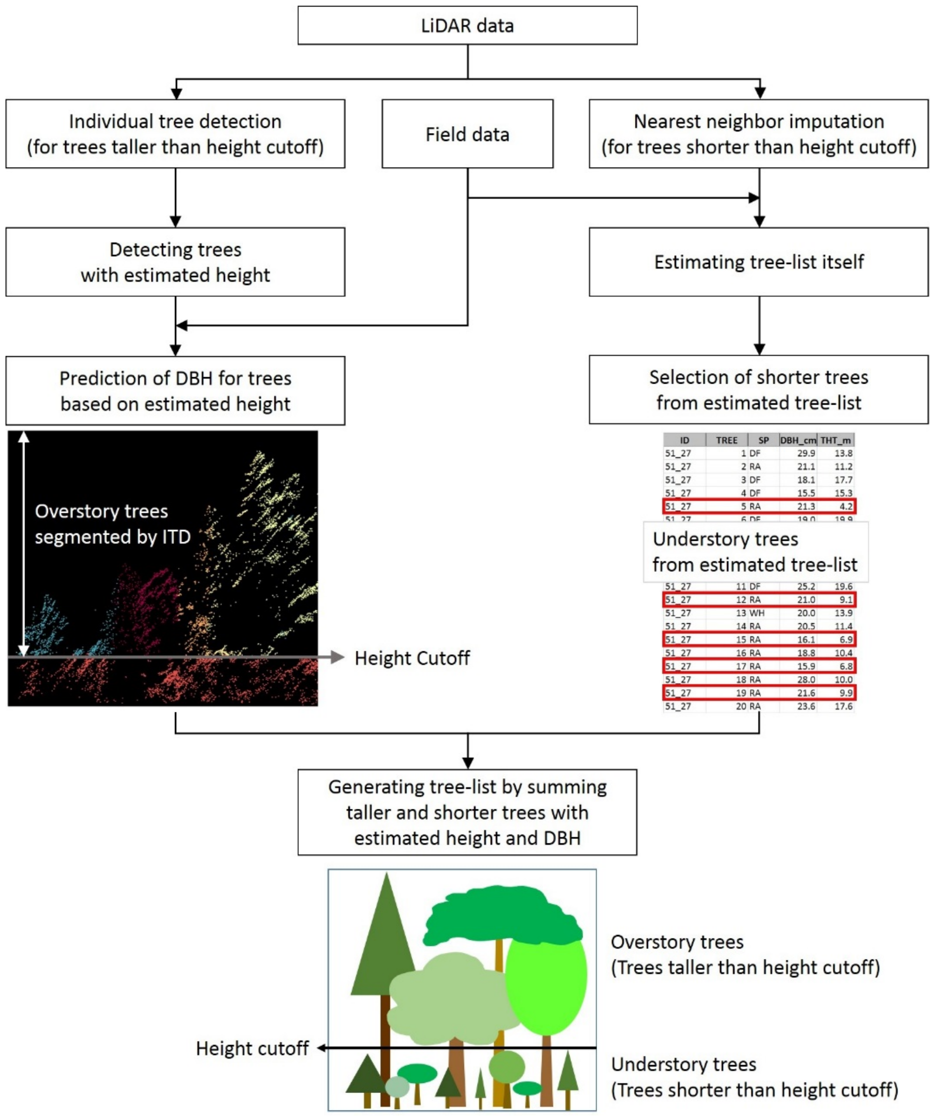


Figure 3.2. Flowchart of the approach

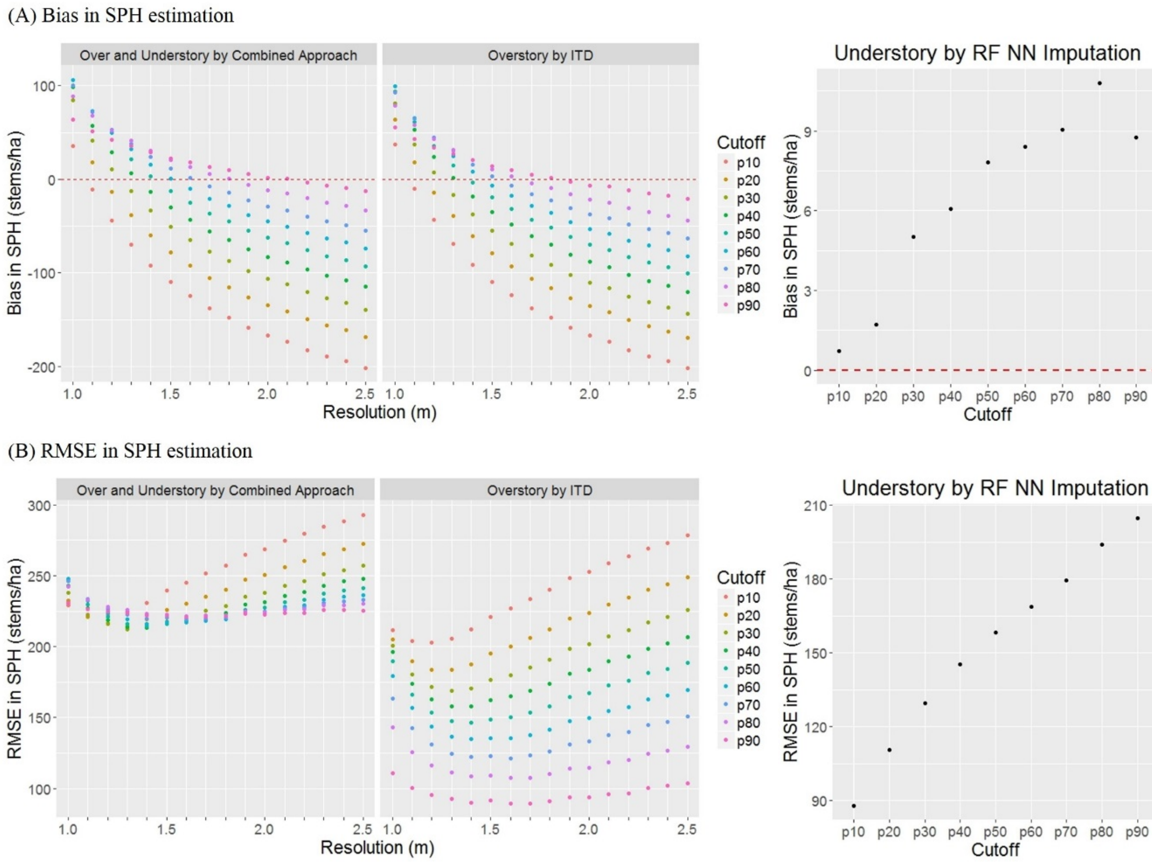


Figure 3.3. (A) Bias and (B) RMSE in SPH estimation: the left graph is for overstory and understory trees via the combined approach by resolution of CHM and height cutoff; the middle graph is for overstory trees via ITD by amount of smoothing, resolution of CHM and height cutoff; and the right graph is for understory trees via RF NN by height cutoff (the horizontal dashed line indicates unbiased estimates).

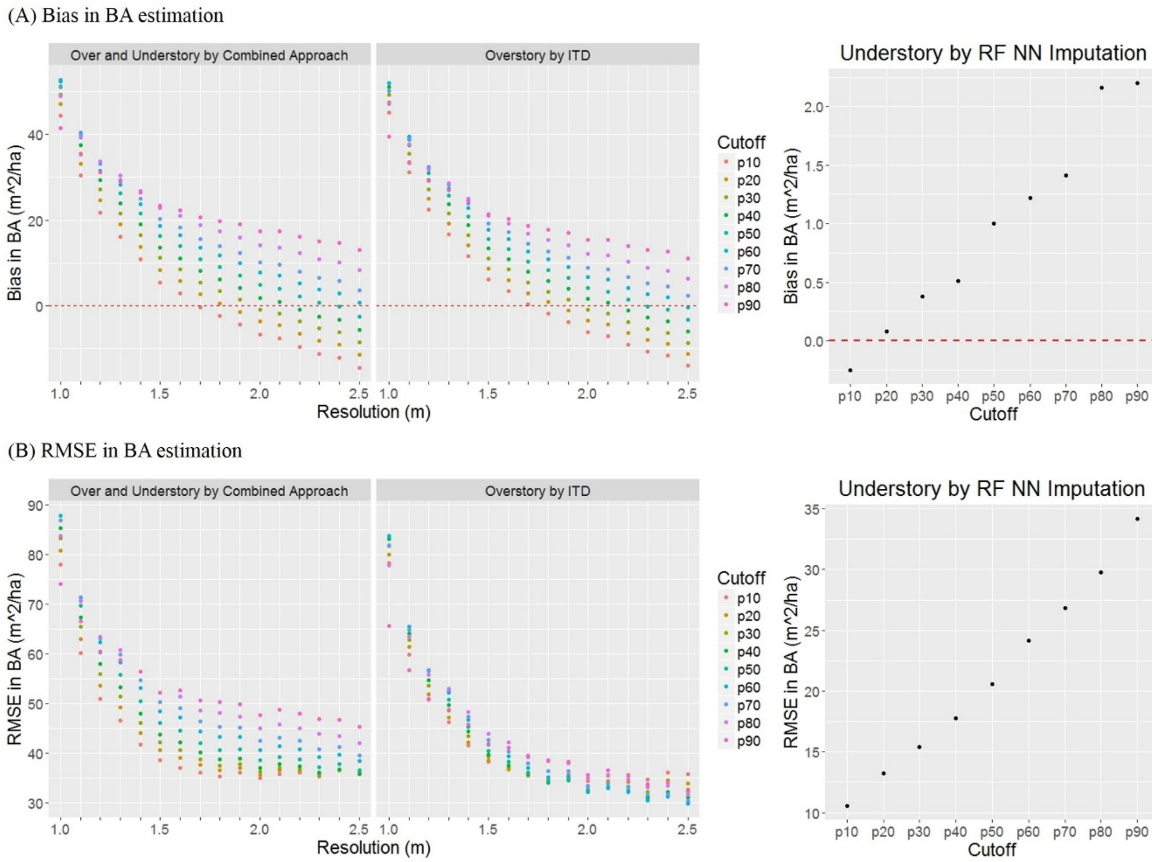
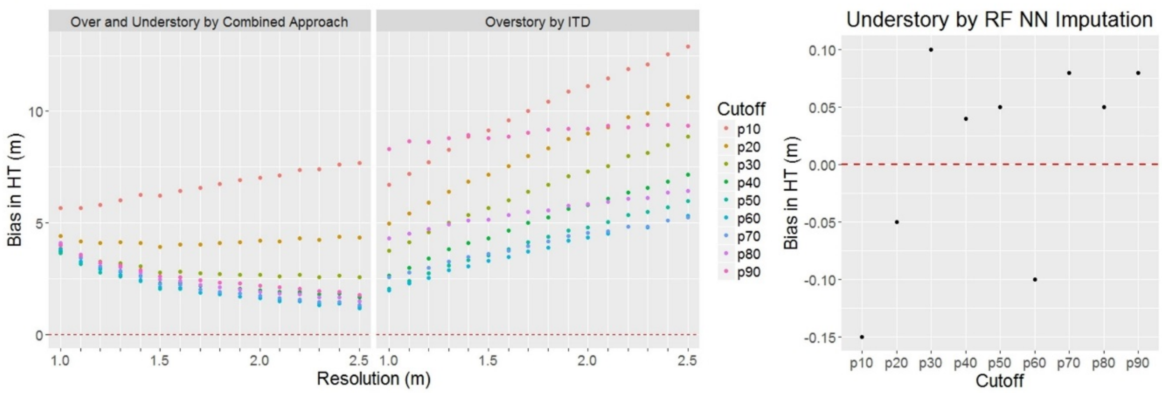


Figure 3.4. (A) Bias and (B) RMSE in BA estimation: the left graph is for overstory and understory trees via the combined approach by resolution of CHM and height cutoff; the middle graph is for overstory trees via ITD by amount of smoothing, resolution of CHM and height cutoff; and the right graph is for understory trees via RF NN by height cutoff (the horizontal dashed line indicates unbiased estimates).

(A) Bias in HT estimation



(B) RMSE in HT estimation

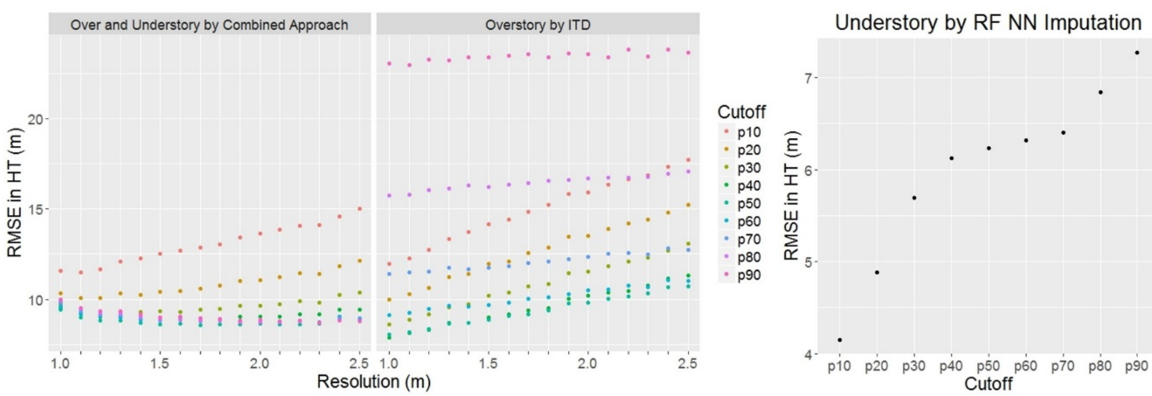


Figure 3.5. (A) Bias and (B) RMSE in HT estimation: the left graph is for overstory and understory trees via the combined approach by resolution of CHM and height cutoff; the middle graph is for overstory trees via ITD by amount of smoothing, resolution of CHM and height cutoff; and the right graph is for understory trees via RF NN by height cutoff (the horizontal dashed line indicates unbiased estimates).

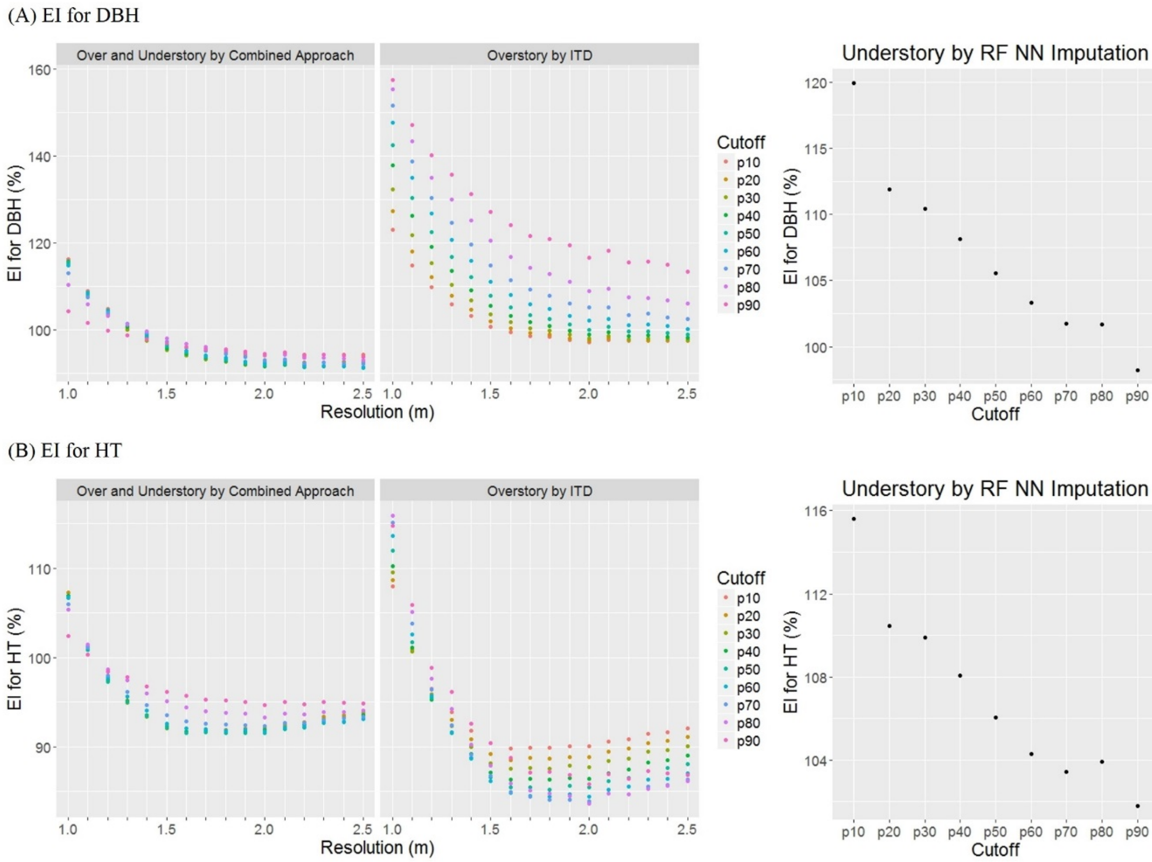


Figure 3.6. (A) EI for DBH and (B) EI for HT: the left graph is for overstory and understory trees via the combined approach by resolution of CHM and height cutoff; the middle graph is for overstory trees via ITD by amount of smoothing, resolution of CHM and height cutoff; and the right graph is for understory trees via RF NN by height cutoff.

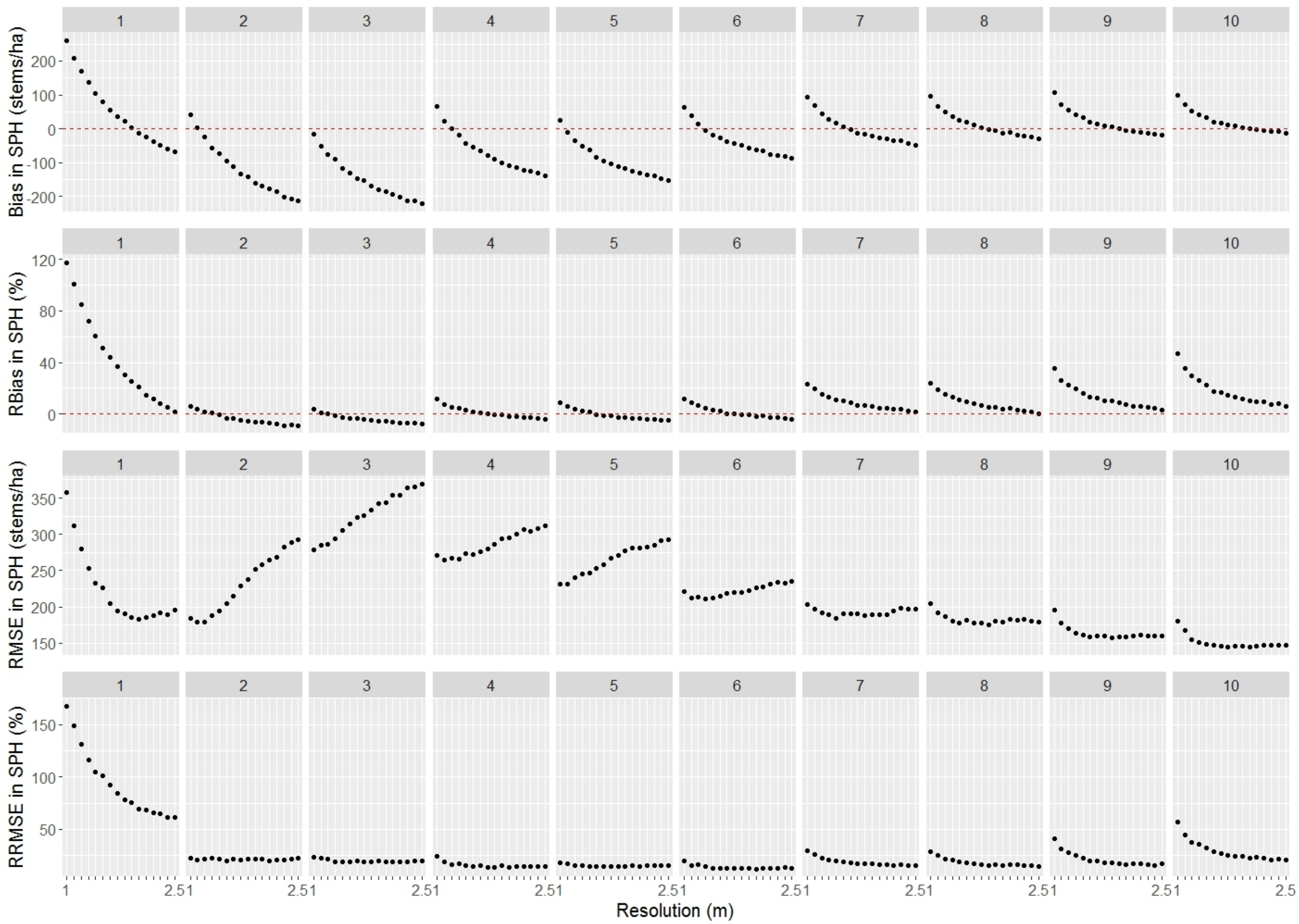


Figure 3.7. Bias, RBias, RMSE, and RRMSE for SPH estimation via the combined approach by LiDAR height class and resolution of CHM: the values of each performance by height cutoff in a given resolution are averaged.

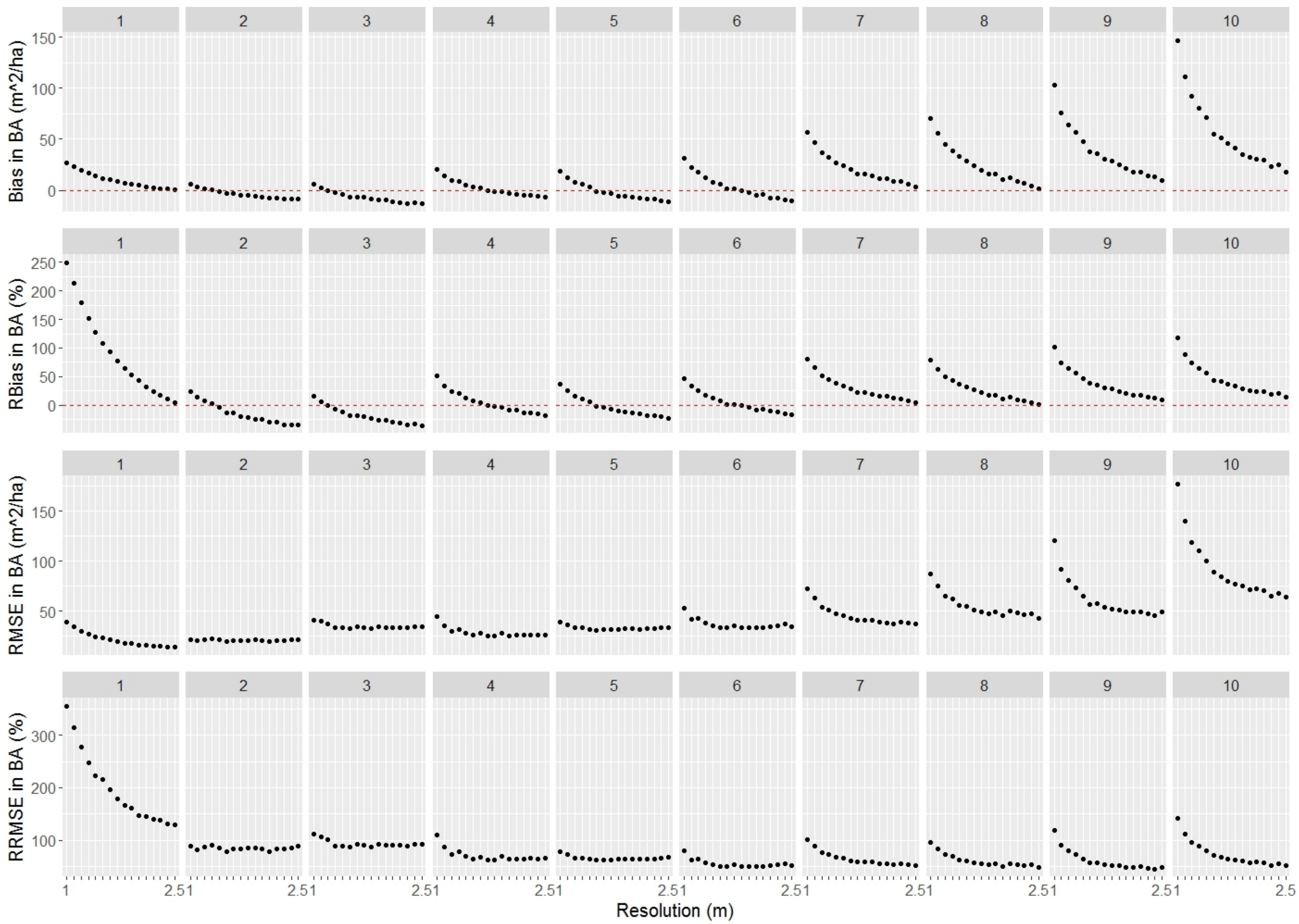


Figure 3.8. Bias, RBias, RMSE, and RRMSE for BA estimation via the combined approach by LiDAR height class and resolution of CHM: the values of each performance by height cutoff in a given resolution are averaged.

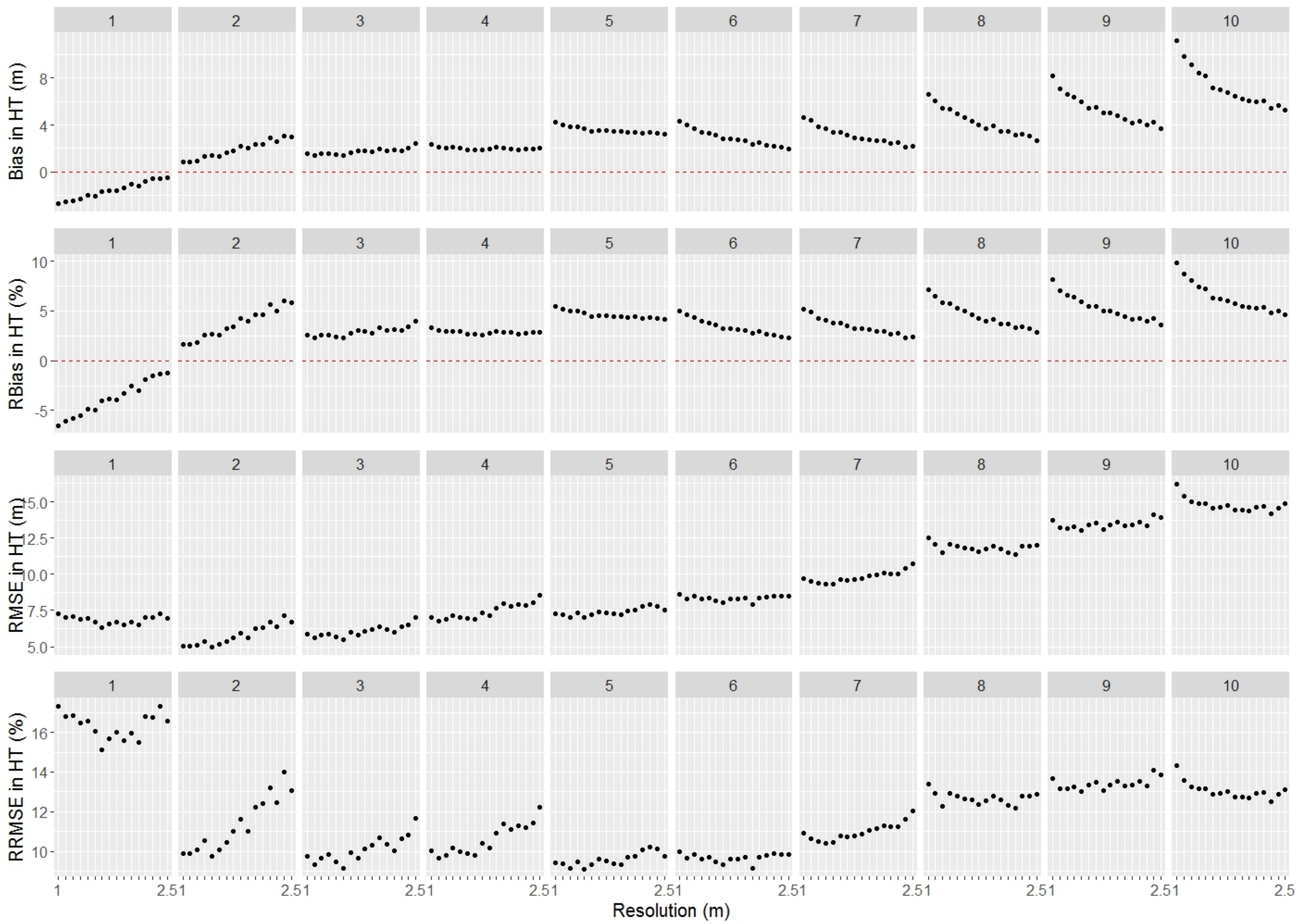


Figure 3.9. Bias, RBias, RMSE, and RRMSE for HT estimation via the combined approach by LiDAR height class and resolution of CHM: the values of each performance by height cutoff in a given resolution are averaged.

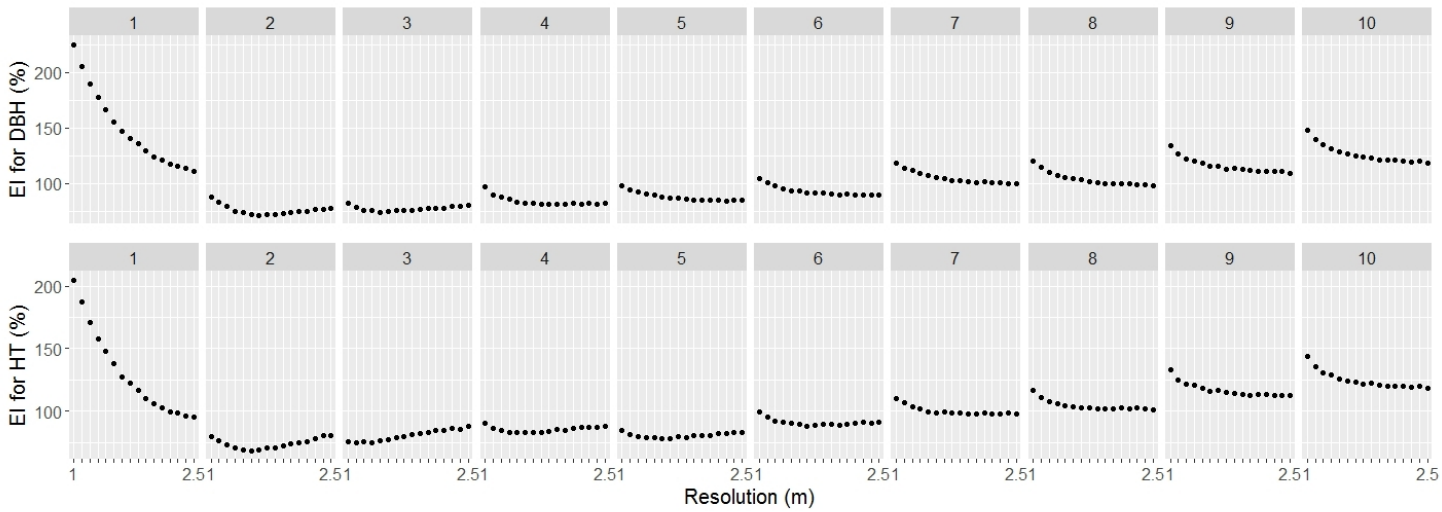


Figure 3.10. EIs for DBH and HT estimation via the combined approach by LiDAR height class and resolution of CHM: the values of each performance by height cutoff in a given resolution are averaged.

Chapter 4: General Conclusion

This study consists of two parts: (1) comparing statistical modeling methods for predicting forest inventory attributes using LiDAR data, and (2) generating tree-lists by fusing ITD with CHM and crown segmentation and ABA with NN imputation using LiDAR data. From the first part, no modeling technique was found to be the best for every case. The best method varied according to response variables, prediction type, and performance measures. The effective size of training data depended on the prediction type. About 100-150 training data showed comparable performance in point prediction, whereas about 200-250 training data showed comparable performance in total prediction. Thus, it should be carefully determined to select a modeling technique

Within the leading group of the modeling methods (BestNN, SLM, OLS, and GWR), each method had its own properties. OLS generally appeared to have very good performance. BestNN yielded comparable performances to the linear models. It was found that there was biasedness in BestNN prediction, but the prediction values of BestNN were within the observed biological bounds. SLM showed its potential to robustly estimate forest attributes. As a more generalized model than OLS regression, it could have good performance in various conditions in terms of diverse combinations of relationships between responses and predictors. GWR produced better performance but GWR’s performance was sensitive to a small number of training data in estimating forest attributes.

RF imputation had poor performance, particularly in PIC90. It was previously reported that bias is one of the reasons for the poor performance of RF imputation. In our study, we observed that smaller prediction standard error also could impact PIC90. And increasing the number of predictor variables did not guarantee improvement in RF’s PIC90.

In the second part of our study, we proposed an approach to combine ITD and ABA to generate a tree-list using airborne LiDAR and field measured data. The approach aimed to compensate for the disadvantage of LiDAR data and ITD in estimating understory trees, and to keep the strength of ITD in estimating overstory trees in tree-level. Smoothing, resolution and height cutoff were investigated to observe how they affected the performance of the proposed approach. No single combination of the three parameters was found to produce the best estimation results for every forest attribute. The best result for each attribute varied according to different combinations of the parameters. It would be operational and useful to determine how to automatically find the optimal combinations of those parameters across the forest landscape. Additional parameters such as forest types, tree species, tree-size parameters (tree crown width or maximum tree height) and topography should be considered to automatically determine the optimal combinations.

There are several topics for further study to improve the combined approach. The performance of the combined approach could be improved by using a denser point cloud data. There is an active area of research in comparing of different algorithms for processing a CHM. Estimating the number of trees per delineated crown segment would help provide unbiased SPH estimation. A point cloud based ITD utilizing more information in LiDAR data could lead to improvement. There were unreasonably tiny crowns that degraded the quality of the predicted tree-lists. Setting a minimum crown area according to the corresponding height could be an option for our approach’s improvement. The effect of slope on both CHM generation and LiDAR metrics extraction needs to be considered for better estimation. Fusing ITD and ABA to predict overstory and understory vegetation shown in this research indicates that forest analysts can benefit from the predictive abilities of the imputation approach and the quality information provided by LiDAR. In that, the approach presented herein can be sufficient for strategic inventory purposes.

Bibliography

Barnes, C., Balzter, H., Barrett, K., Eddy, J., Milner, S., and Suárez, J. 2017. Individual Tree Crown Delineation from Airborne Laser Scanning for Diseased Larch Forest Stands. *Remote Sensing* 9(3): 231.

Bivand, R., and Yu, D. 2013. *spgwr: Geographically weighted regression*. R package version 0.6-24..

Breidenbach, J., Næsset, E., Lien, V., Gobakken, T., and Solberg, S. 2010a. 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(4): 911-924. doi: http://dx.doi.org/10.1016/j.rse.2009.12.004.

Breidenbach, J., Nothdurft, A., and Kändler, G. 2010b. Comparison of nearest neighbour approaches for small area estimation of tree species-specific forest inventory attributes in central Europe using airborne laser scanner data. *European Journal of Forest Research* 129(5): 833-846. doi: 10.1007/s10342-010-0384-1.

Breiman, L. 2001. Random forests. *Machine Learning* 45(1): 5-32. doi: 10.1023/a:1010933404324.

Brosofske, K.D., Froese, R.E., Falkowski, M.J., and Banskota, A. 2014. A Review of Methods for Mapping and Prediction of Inventory Attributes for Operational Forest Management. *Forest Science* 60(4): 733-756. doi: 10.5849/forsci.12-134.

Brunsdon, C., Fotheringham, S., and Charlton, M. 1998. *Geographically weighted regression. Journal of the Royal Statistical Society: Series D (The Statistician)* 47(3): 431-443. doi: 10.1111/1467-9884.00145.

Chen, G., Zhao, K., McDermid, G.J., and Hay, G.J. 2011. The influence of sampling density on geographically weighted regression: a case study using forest canopy height and optical data. *International Journal of Remote Sensing* 33(9): 2909-2924. doi: 10.1080/01431161.2011.624130.

Cho, S., Lambert, D.M., Kim, S.G., and Jung, S. 2009. Extreme coefficients in Geographically Weighted Regression and their effects on mapping. In 2009 Annual Meeting. Agricultural and Applied Economics Association, Milwaukee, Wisconsin.

Cliff, A.D., and Ord, J.K. 1981. *Spatial processes: models & applications*. Pion, London.

Corona, P., Fattorini, L., Franceschi, S., Chirici, G., Maselli, F., and Secondi, L. 2014. Mapping by spatial predictors exploiting remotely sensed and ground data: A comparative design-based perspective. *Remote Sensing of Environment* 152(0): 29-37. doi: http://dx.doi.org/10.1016/j.rse.2014.05.011.

Cressie, N. 1993. *Statistics for Spatial Data* (rev. ed.). Wiley, New York, NY, USA. pp. 69-70.

Crookston, N.L., and Finley, A.O. 2008. yaImpute: An R package for kNN imputation. *Journal of Statistical Software* 23(10): 16.

Eskelson, B.N., Temesgen, H., Lemay, V., Barrett, T.M., Crookston, N.L., and Hudak, A.T. 2009a. The roles of nearest neighbor methods in imputing missing data in forest inventory and monitoring databases. *Scandinavian Journal of Forest Research* 24(3): 235-246.

Eskelson, B.N.I., Barrett, T.M., and Temesgen, H. 2009b. Imputing Mean Annual Change to Estimate Current Forest Attributes. *Silva Fennica* 43(4): 649-658. doi: 10.14214/sf.185.

Eskelson, B.N.I., Madsen, L., Hagar, J.C., and Temesgen, H. 2011. Estimating Riparian Understory Vegetation Cover with Beta Regression and Copula Models. *Forest Science* 57(3): 212-221.

Falkowski, M.J., Hudak, A.T., Crookston, N.L., Gessler, P.E., Uebler, E.H., and Smith, A.M.S. 2010. Landscape-scale parameterization of a tree-level forest growth model: a k-nearest neighbor imputation approach incorporating LiDAR data. *Canadian Journal of Forest Research* 40(2): 184-199. doi: 10.1139/X09-183.

Faraway, J. 2011. *faraway: Functions and datasets for books by Julian Faraway*. R package version 1.0.5.

Gagliasso, D., Hummel, S., and Temesgen, H. 2014. A comparison of selected parametric and non-parametric imputation methods for estimating forest biomass and basal area. *Open Journal of Forestry* 4(01): 42.

Garcia-Gutierrez, J., Gonzalez-Ferreiro, E., Riquelme-Santos, J.C., Miranda, D., Dieguez-Aranda, U., and Navarro-Cerrillo, R.M. 2014. Evolutionary feature selection to estimate forest stand variables using LiDAR. *International Journal of Applied Earth Observation and Geoinformation* 26: 119-131. doi: http://dx.doi.org/10.1016/j.jag.2013.06.005.

Gatziolis, D., and Andersen, H.-E. 2008. A guide to LIDAR data acquisition and processing for the forests of the Pacific Northwest. US Department of Agriculture, Forest Service, Pacific Northwest Research Station. 32 p, Gen. Tech. Rep. PNW-GTR-768. Portland, OR. p. 32.

Gobakken, T., and Næsset, E. 2004. Estimation of diameter and basal area distributions in coniferous forest by means of airborne laser scanner data. *Scandinavian Journal of Forest Research* 19(6): 529-542. doi: 10.1080/02827580410019454.

Gobakken, T., Næsset, E., Nelson, R., Bollandsås, O.M., Gregoire, T.G., Ståhl, G., Holm, S., Ørka, H.O., and Astrup, R. 2012. Estimating biomass in Hedmark County, Norway using national forest inventory field plots and airborne laser scanning. *Remote Sensing of Environment* 123(0): 443-456. doi: http://dx.doi.org/10.1016/j.rse.2012.01.025.

Hamraz, H., Contreras, M.A., and Zhang, J. 2017. Vertical stratification of forest canopy for segmentation of understory trees within small-footprint airborne LiDAR point clouds. *ISPRS Journal of Photogrammetry and Remote Sensing* 130: 385-392. doi: https://doi.org/10.1016/j.isprsjprs.2017.07.001.

Hansen, E., Ene, L., Gobakken, T., Ørka, H., Bollandsås, O., and Næsset, E. 2017. Countering Negative Effects of Terrain Slope on Airborne Laser Scanner Data Using Procrustean Transformation and Histogram Matching. *Forests* 8(10): 401.

Hastie, T., Tibshirani, R., Friedman, J., Hastie, T., Friedman, J., and Tibshirani, R. 2009. *The elements of statistical learning* (2nd ed.). Springer-Verlag New York. pp. 745.

Hawbaker, T.J., Keuler, N.S., Lesak, A.A., Gobakken, T., Contrucci, K., and Radeloff, V.C. 2009. Improved estimates of forest vegetation structure and biomass with a LiDAR-optimized sampling design. *Journal of Geophysical Research: Biogeosciences* 114(G2): G00E04. doi: 10.1029/2008JG000870.

Heurich, M., and Weinacker, H. 2004. Automated tree detection and measurement in temperate forests of central Europe using laser scanning data. In *ISPRS working group on Laser-Scanners for Forest and Landscape Assessment*. Edited by M. Thies and B. Koch and H. Spiecker and H. Weinacker, Freiburg, Germany. pp. 198-203.

Hoeting, J.A. 2009. The importance of accounting for spatial and temporal correlation in analyses of ecological data. *Ecological Applications* 19(3): 574-577. doi: 10.1890/08-0836.1.

Holopainen, M., and Kalliovirta, M.S.J. 2006. Modern data acquisition for forest inventories. In *Forest Inventory*. Edited by A. Kangas and M. Maltamo. Springer. pp. 343-362.

Hopkinson, C., Chasmer, L.E., Sass, G., Creed, I.F., Sitar, M., Kalbfleisch, W., and Treitz, P. 2005. Vegetation class dependent errors in lidar ground elevation and canopy height estimates in a boreal wetland environment. *Canadian Journal of Remote Sensing* 31(2): 191-206. doi: 10.5589/m05-007.

Hudak, A.T., Crookston, N.L., Evans, J.S., Hall, D.E., and Falkowski, M.J. 2008. Nearest neighbor imputation of species-level, plot-scale forest structure attributes from LiDAR data. *Remote Sensing of Environment* 112(5): 2232-2245. doi: http://dx.doi.org/10.1016/j.rse.2007.10.009.

Hudak, A.T., Haren, A.T., Crookston, N.L., Liebermann, R.J., and Ohmann, J.L. 2014. Imputing forest structure attributes from stand inventory and remotely sensed data in Western Oregon, USA. *Forest Science* 60(2): 253-269. doi: 10.5849/forsci.12-101.

Hudak, A.T., Strand, E.K., Vierling, L.A., Byrne, J.C., Eitel, J.U.H., Martinuzzi, S., and Falkowski, M.J. 2012. Quantifying aboveground forest carbon pools and fluxes from repeat LiDAR surveys. *Remote Sensing of Environment* 123(0): 25-40. doi: http://dx.doi.org/10.1016/j.rse.2012.02.023.

Husch, B., Beers, T.W., and Kershaw Jr, J.A. 2002. *Forest mensuration*. John Wiley & Sons. pp. 176-177.

Junttila, V., Kauranne, T., Finley, A.O., and Bradford, J.B. 2015. Linear Models for Airborne-Laser-Scanning-Based Operational Forest Inventory With Small Field Sample Size and Highly Correlated LiDAR Data. *IEEE Transactions on Geoscience and Remote Sensing* 53(10): 5600-5612. doi: 10.1109/TGRS.2015.2425916.

Kaartinen, H., Hyyppä, J., Yu, X., Vastaranta, M., Hyyppä, H., Kukko, A., Holopainen, M., Heipke, C., Hirschmugl, M., Morsdorf, F., Næsset, E., Pitkänen, J., Popescu, S., Solberg, S., Wolf, B.M., and Wu, J.-C. 2012. An International Comparison of Individual Tree Detection and Extraction Using Airborne Laser Scanning. *Remote Sensing* 4(4): 950.

Kandare, K., Ørka, H.O., Chan, J.C.-W., and Dalponte, M. 2016. Effects of forest structure and airborne laser scanning point cloud density on 3D delineation of individual tree crowns. *European Journal of Remote Sensing* 49(1): 337-359. doi: 10.5721/EuJRS20164919.

Kankare, V., Vastaranta, M., Holopainen, M., Räty, M., Yu, X., Hyyppä, J., Hyyppä, H., Alho, P., and Viitala, R. 2013. Retrieval of forest aboveground biomass and stem volume with airborne scanning LiDAR. *Remote Sensing* 5(5): 2257-2274.

Khosravipour, A., Skidmore, A.K., Isenburg, M., Wang, T., and Hussin, Y.A. 2014. Generating pit-free canopy height models from airborne lidar. *Photogrammetric Engineering & Remote Sensing* 80(9): 863-872.

Khosravipour, A., Skidmore, A.K., Wang, T., Isenburg, M., and Khoshelham, K. 2015. Effect of slope on treetop detection using a LiDAR Canopy Height Model. *ISPRS Journal of Photogrammetry and Remote Sensing* 104: 44-52. doi: https://doi.org/10.1016/j.isprsjprs.2015.02.013.

Koch, B., Heyder, U., and Weinacker, H. 2006. Detection of Individual Tree Crowns in Airborne Lidar Data. *Photogrammetric Engineering & Remote Sensing* 72(4): 357-363. doi: 10.14358/PERS.72.4.357.

Koch, B., Kattenborn, T., Straub, C., and Vauhkonen, J. 2014. Segmentation of Forest to Tree Objects. In *Forestry Applications of Airborne Laser Scanning: Concepts and Case Studies*. Edited by M. Maltamo and E. Næsset and J. Vauhkonen. Springer Netherlands, Dordrecht. pp. 89-112.

Latifi, H., and Koch, B. 2012. Evaluation of most similar neighbour and random forest methods for imputing forest inventory variables using data from target and auxiliary stands. *International Journal of Remote Sensing* 33(21): 6668-6694. doi: 10.1080/01431161.2012.693969.

Latifi, H., Nothdurft, A., and Koch, B. 2010. Non-parametric prediction and mapping of standing timber volume and biomass in a temperate forest: application of multiple optical/LiDAR-derived predictors. *Forestry* 83(4): 395-407.

LeMay, V., Maedel, J., and Coops, N.C. 2008. Estimating stand structural details using nearest neighbor analyses to link ground data, forest cover maps, and Landsat imagery. *Remote Sensing of Environment* 112(5): 2578-2591. doi: http://dx.doi.org/10.1016/j.rse.2007.12.007.

Li, M., Im, J., Quackenbush, L.J., and Liu, T. 2014. Forest biomass and carbon stock quantification using airborne LiDAR data: a case study over Huntington Wildlife Forest in the Adirondack Park. *Selected Topics in Applied Earth Observations and Remote Sensing, IEEE Journal of* 7(7): 3143-3156. doi: 10.1109/JSTARS.2014.2304642.

Liaw, A., and Wiener, M. 2002. Classification and Regression by randomForest. *R news* 2(3): 18-22.

Lindberg, E., and Hollaus, M. 2012. Comparison of methods for estimation of stem volume, stem number and basal area from airborne laser scanning data in a hemi-boreal forest. *Remote Sensing* 4(4): 1004-1023.

Lindberg, E., and Holmgren, J. 2017. Individual Tree Crown Methods for 3D Data from Remote Sensing. *Current Forestry Reports*: 1-13. doi: 10.1007/s40725-017-0051-6.

Lindberg, E., Holmgren, J., Olofsson, K., Wallerman, J., and Olsson, H. 2010. Estimation of tree lists from airborne laser scanning by combining single-tree and area-based methods. *International Journal of Remote Sensing* 31(5): 1175-1192. doi: 10.1080/01431160903380649.

Lindberg, E., Holmgren, J., Olofsson, K., Wallerman, J., and Olsson, H. 2013. Estimation of Tree Lists from Airborne Laser Scanning Using Tree Model Clustering and k-MSN Imputation. *Remote Sensing* 5(4): 1932.

Liu, J., Shen, J., Zhao, R., and Xu, S. 2013. Extraction of individual tree crowns from airborne LiDAR data in human settlements. *Mathematical and Computer Modelling* 58(3): 524-535. doi: https://doi.org/10.1016/j.mcm.2011.10.071.

Lumley, T. 2009. *leaps: regression subset selection*. R package version 2.9. http://cran.r-project.org/web/packages/leaps/index.html.

Maltamo, M., Bollandsås, O., Næsset, E., Gobakken, T., and Packalén, P. 2011. Different plot selection strategies for field training data in ALS-assisted forest inventory. *Forestry* 84(1): 23-31.

Maltamo, M., Eerikäinen, K., Pitkänen, J., Hyyppä, J., and Vehmas, M. 2004. Estimation of timber volume and stem density based on scanning laser altimetry and expected tree size distribution functions. *Remote Sensing of Environment* 90(3): 319-330. doi: http://dx.doi.org/10.1016/j.rse.2004.01.006.

Maltamo, M., and Gobakken, T. 2014. Predicting Tree Diameter Distributions. In *Forestry Applications of Airborne Laser Scanning: Concepts and Case Studies*. Edited by M. Maltamo and E. Næsset and J. Vauhkonen. Springer Netherlands, Dordrecht. pp. 177-191.

Maltamo, M., Malinen, J., Packalén, P., Suvanto, A., and Kangas, J. 2006. Nonparametric estimation of stem volume using airborne laser scanning, aerial photography, and stand-register data. *Canadian Journal of Forest Research* 36(2): 426-436. doi: 10.1139/x05-246.

Maltamo, M., Tokola, T., and Lehikoinen, M. 2003. Estimating Stand Characteristics by Combining Single Tree Pattern Recognition of Digital Video Imagery and a Theoretical Diameter Distribution Model. *Forest Science* 49(1): 98-109.

McGaughey, R.J. 2010. *FUSION/LDV: Software for LIDAR data analysis and visualization*, Forest service. Pacific Northwest research station, United States department of agriculture.

McGaughey, R.J. 2016. *FUSION/LDV: Software for LIDAR Data Analysis and Visualization*, Forest Service. Pacific Northwest research station, United States department of agriculture.

McRoberts, R.E., Cohen, W.B., Naesset, E., Stehman, S.V., and Tomppo, E.O. 2010. Using remotely sensed data to construct and assess forest attribute maps and related spatial products. *Scandinavian Journal of Forest Research* 25(4): 340-367.

Moeur, M., and Stage, A.R. 1995. Most similar neighbor: an improved sampling inference procedure for natural resource planning. *Forest Science* 41(2): 337-359.

Muinonen, E., Maltamo, M., Hyppänen, H., and Vainikainen, V. 2001. Forest stand characteristics estimation using a most similar neighbor approach and image spatial structure information. *Remote Sensing of Environment* 78(3): 223-228. doi: http://dx.doi.org/10.1016/S0034-4257(01)00220-6.

Murphy, M.A., Evans, J.S., and Storfer, A. 2010. Quantifying Bufo boreas connectivity in Yellowstone National Park with landscape genetics. *Ecology* 91(1): 252-261. doi: 10.1890/08-0879.1.

Næsset, E., and Gobakken, T. 2008. Estimation of above- and below-ground biomass across regions of the boreal forest zone using airborne laser. *Remote Sensing of Environment* 112(6): 3079-3090. doi: http://dx.doi.org/10.1016/j.rse.2008.03.004.

Niska, H., Skon, J., Packalen, P., Tokola, T., Maltamo, M., and Kolehmainen, M. 2010. Neural Networks for the Prediction of Species-Specific Plot Volumes Using Airborne Laser Scanning and Aerial Photographs. *IEEE Transactions on Geoscience and Remote Sensing*, 48(3): 1076-1085. doi: 10.1109/TGRS.2009.2029864.

Ohmann, J.L., and Gregory, M.J. 2002. Predictive mapping of forest composition and structure with direct gradient analysis and nearest-neighbor imputation in coastal Oregon, USA. *Canadian Journal of Forest Research* 32(4): 725-741.

Packalén, P., Mehtätalo, L., and Maltamo, M. 2011. ALS-based estimation of plot volume and site index in a eucalyptus plantation with a nonlinear mixed-effect model that accounts for the clone effect. *Annals of Forest Science* 68(6): 1085-1092. doi: 10.1007/s13595-011-0124-9.

Packalén, P., Temesgen, H., and Maltamo, M. 2012. Variable selection strategies for nearest neighbor imputation methods used in remote sensing based forest inventory. *Canadian Journal of Remote Sensing* 38(5): 557-569. doi: 10.5589/m12-046.

Paradis, E., Claude, J., and Strimmer, K. 2004. APE: Analyses of Phylogenetics and Evolution in R language. *Bioinformatics* 20(2): 289-290. doi: 10.1093/bioinformatics/btg412.

Penner, M., Pitt, D.G., and Woods, M.E. 2013. Parametric vs. nonparametric LiDAR models for operational forest inventory in boreal Ontario. *Canadian Journal of Remote Sensing* 39(5): 426-443. doi: 10.5589/m13-049.

Persson, A., Holmgren, J., and Soderman, U. 2002. Detecting and measuring individual trees using an airborne laser scanner. *Photogrammetric Engineering and Remote Sensing* 68(9): 925-932.

Pierce, K.B., Ohmann, J.L., Wimberly, M.C., Gregory, M.J., and Fried, J.S. 2009. Mapping wildland fuels and forest structure for land management: a comparison of nearest neighbor imputation and other methods. *Canadian Journal of Forest Research* 39(10): 1901-1916. doi: 10.1139/X09-102.

Pippuri, I., Kallio, E., Maltamo, M., Peltola, H., and Packalén, P. 2012. Exploring horizontal area-based metrics to discriminate the spatial pattern of trees and need for first thinning using airborne laser scanning. *Forestry* 85(2): 305-314. doi: 10.1093/forestry/cps005.

Pirotti, F., Kobal, M., and Roussel, J.R. 2017. A COMPARISON OF TREE SEGMENTATION METHODS USING VERY HIGH DENSITY AIRBORNE LASER SCANNER DATA. *Int. Arch. Photogramm. Remote Sens. Spatial Inf. Sci*. XLII-2/W7: 285-290. doi: 10.5194/isprs-archives-XLII-2-W7-285-2017.

Pouliot, D.A., King, D.J., Bell, F.W., and Pitt, D.G. 2002. Automated tree crown detection and delineation in high-resolution digital camera imagery of coniferous forest regeneration. *Remote Sensing of Environment* 82(2): 322-334. doi: https://doi.org/10.1016/S0034-4257(02)00050-0.

R Core Team. 2013. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

R Core Team. 2017. *R: A language and environment for statistical computing*. R Foundation for Statistical Computing, Vienna, Austria.

Rahman, M.Z.A., and Gorte, B. 2008. Tree filtering for high density Airborne LiDAR data. In *Proceedings of SilviLaser 2008: 8th International Conference on LiDAR Applications in Forest Assessment and Inventory*, Edinburgh, UK. pp. 544-553.

Reynolds, M.R., Burk, T.E., and Huang, W.-C. 1988. Goodness-of-Fit Tests and Model Selection Procedures for Diameter Distribution Models. *Forest Science* 34(2): 373-399.

Romero-Zaliz, R., and Reinoso-Gordo, J.F. 2018. An Updated Review on Watershed Algorithms. In *Soft Computing for Sustainability Science*. Edited by C. Cruz Corona. Springer International Publishing, Cham. pp. 235-258.

Schabenberger, O., and Gotway, C.A. 2004. *Statistical methods for spatial data analysis*. CRC press.

Smreček, R., Michnová, Z., Sačkov I, Danihelová, Z., Levická, M., and Tuček, J. 2018. Determining basic forest stand characteristics using airborne laser scanning in mixed forest stands of Central Europe. *iForest - Biogeosciences and Forestry* 11(1): 181-188. doi: 10.3832ifor2520-010.

Stage, A.R., and Crookston, N.L. 2007. Partitioning error components for accuracy-assessment of near-neighbor methods of imputation. *Forest Science* 53(1): 62-72.

Stereńczak, K., Będkowski, K., and Weinacker, H. 2008. Accuracy of crown segmentation and estimation of selected trees and forest stand parameters in order to resolution of used DSM and nDSM models generated from dense small footprint LIDAR data. In *ISPRS Congress*, Beijing, China. pp. 27-32.

Strunk, J., Gould, P., Packalen, P., Poudel, K., Andersen, H.-E., and Temesgen, H. 2017. An Examination of Diameter Density Prediction with k-NN and Airborne Lidar. *Forests* 8(11): 444.

Strunk, J.L., Reutebuch, S.E., and Foster, J.R. 2008. LIDAR inventory and monitoring of a complex forest. In *ASPRS Annual Conference*, Portland, Oregon.

Takahashi, T., Yamamoto, K., Miyachi, Y., Senda, Y., and Tsuzuku, M. 2006. The penetration rate of laser pulses transmitted from a small-footprint airborne LiDAR: a case study in closed canopy, middle-aged pure sugi (Cryptomeria japonica D. Don) and hinoki cypress (Chamaecyparis obtusa Sieb. et Zucc.) stands in Japan. *Journal of Forest Research* 11(2): 117-123. doi: 10.1007/s10310-005-0189-0.

Temesgen, B.H., LeMay, V.M., Froese, K.L., and Marshall, P.L. 2003. Imputing tree-lists from aerial attributes for complex stands of south-eastern British Columbia. *Forest Ecology and Management* 177(1–3): 277-285. doi: http://dx.doi.org/10.1016/S0378-1127(02)00321-3.

Temesgen, H., Barrett, T.M., and Latta, G. 2008. Estimating cavity tree abundance using nearest neighbor imputation methods for western Oregon and Washington forests. *Silva Fennica* 42(3): 337-354. doi: 10.14214/sf.241.

Temesgen, H., and Ver Hoef, J.M. 2015. Evaluation of the spatial linear model, random forest and gradient nearest-neighbour methods for imputing potential productivity and biomass of the Pacific Northwest forests. *Forestry* 88(1): 131-142. doi: 10.1093/forestry/cpu036.

Vauhkonen, J., Ene, L., Gupta, S., Heinzel, J., Holmgren, J., Pitkänen, J., Solberg, S., Wang, Y., Weinacker, H., Hauglin, K.M., Lien, V., Packalén, P., Gobakken, T., Koch, B., Næsset, E., Tokola, T., and Maltamo, M. 2012. Comparative testing of single-tree detection algorithms under different types of forest. Forestry: *An International Journal of Forest Research* 85(1): 27-40. doi: 10.1093/forestry/cpr051.

Vauhkonen, J., Korpela, I., Maltamo, M., and Tokola, T. 2010. Imputation of single-tree attributes using airborne laser scanning-based height, intensity, and alpha shape metrics. *Remote Sensing of Environment* 114(6): 1263-1276.

Vauhkonen, J., Maltamo, M., McRoberts, R.E., and Næsset, E. 2014. Introduction to Forestry Applications of Airborne Laser Scanning. In *Forestry Applications of Airborne Laser Scanning: Concepts and Case Studies*. Edited by M. Maltamo and E. Næsset and J. Vauhkonen. Springer Netherlands, Dordrecht. pp. 1-16.

Ver Hoef, J.M., and Temesgen, H. 2013. A comparison of the spatial linear model to nearest neighbor (k-NN) methods for forestry applications. *PloS one* 8(3): e59129.

Vincent, L., and Soille, P. 1991. Watersheds in digital spaces: an efficient algorithm based on immersion simulations. *IEEE Transactions on Pattern Analysis and Machine Intelligence* 13(6): 583-598. doi: 10.1109/34.87344.

Wiggins, H.L. 2017. The influence of tree height on LiDAR's ability to accurately characterize forest structure and spatial pattern across reference landscapes. In *Department of Ecosystem and Conservation Sciences*. University of Montana.

Wing, B.M., Ritchie, M.W., Boston, K., Cohen, W.B., Gitelman, A., and Olsen, M.J. 2012. Prediction of understory vegetation cover with airborne lidar in an interior ponderosa pine forest. *Remote Sensing of Environment* 124(Supplement C): 730-741. doi: https://doi.org/10.1016/j.rse.2012.06.024.

Wulder, M.A., White, J.C., Nelson, R.F., Næsset, E., Ørka, H.O., Coops, N.C., Hilker, T., Bater, C.W., and Gobakken, T. 2012. Lidar sampling for large-area forest characterization: A review. *Remote Sensing of Environment* 121: 196-209. doi: http://dx.doi.org/10.1016/j.rse.2012.02.001.

Yu, X., Hyyppä, J., Holopainen, M., and Vastaranta, M. 2010. Comparison of Area-Based and Individual Tree-Based Methods for Predicting Plot-Level Forest Attributes. *Remote Sensing* 2(6): 1481.

Yu, X., Hyyppä, J., Vastaranta, M., Holopainen, M., and Viitala, R. 2011. Predicting individual tree attributes from airborne laser point clouds based on the random forests technique. *ISPRS*

Journal of Photogrammetry and Remote Sensing 66(1): 28-37. doi:
http://dx.doi.org/10.1016/j.isprsjprs.2010.08.003.

Appendix: Acronyms

ABA: Area-based approach

ANN: Artificial neural networks

BA: Basal area

BIC: Bayesian information criteria

BestNN: a modified *k*-NN

CHM: Canopy height model

DBH: Diameter at breast height

DEN: tree density

EI: Error index

GNN: Gradient nearest neighbor

GWR: Geographically weighted regression

HT: Tree height

ITD: Individual tree detection

LiDAR: light detection and ranging

LOR: Lorey's height

MSN: Most similar neighbor

NN: Nearest neighbor

OLS: Ordinary least square

P80: 80^{th} height percentile

PIC90: 90 % prediction interval coverage

PM: Performance measure

PT: Prediction type

QMD: Quadratic mean diameter

R^2: Coefficient of determination

RBias: Relative bias

RF: Random forest

RMSE: Root mean squared errors

RMSPE: Root mean squared prediction error

RRMSE: Relative root mean squared error

SDH: Standard deviation of LiDAR heights

SLM: Spatial linear model

SPH: Stems per hectare

VOL: Stem volume