Dissertationes Forestales 182

Calibration of the tree size distributions by combining the area-based approach and the individual tree detection using the airborne laser scanning

Qing Xu
School of Forest Sciences
Faculty of Science and Forestry
University of Eastern Finland

Academic dissertation

To be presented, with the permission of the Faculty of Science and Forestry of the University of Eastern Finland, for public criticism in the auditorium BOR 100, University of Eastern Finland, Yliopistokatu 7, Joensuu, on November 28th 2014 at 12 o' clock noon.

Title of dissertation: Calibration of the tree size distributions by combining the area-based approach and the individual tree detection using the airborne laser scanning

Author: Qing Xu

Dissertationes Forestales 182

http://dx.doi.org/10.14214/df.182

Thesis supervisors:
Prof. Timo Tokola
School of Forest Sciences, University of Eastern Finland, Joensuu, Finland

Pre-examiners:
Prof. Juha Hyyppä
Department of Remote Sensing and Photogrammetry, Finnish Geodetic Institute, Masala, Finland
Associate Prof. Michal Zasada
Faculty of Forestry, Warsaw University of Life Sciences, Warsaw, Poland

Opponent:
Prof. Annika Kangas
Department of Forest sciences, University of Helsinki, Helsinki, Finland

ISSN 1795-7389 (online)
ISBN 978-951-651-453-9 (pdf)

ISSN 2323-9220 (print)
ISBN 978-951-651-454-6 (paperback)

2014

Publishers:
Finnish Society of Forest Science
Finnish Forest Research Institute
Faculty of Agriculture and Forestry of the University of Helsinki
School of Forest Sciences of the University of Eastern Finland

Editorial Office:
The Finnish Society of Forest Science
P.O. Box 18, FI-01301 Vantaa, Finland
http://www.metla.fi/dissertationes

Xu, Q. 2014. Calibration of the tree size distributions by combining the area-based approach and the individual tree detection using the airborne laser scanning. Dissertationes Forestales 182. 28 p. Available at
http://dx.doi.org/10.14214/df.182

ABSTRACT

Individual tree detection (ITD) - based forest inventory using the airborne laser scanning (ALS) data suffers from under-estimation problem, which arises mainly from the suppressed trees that are difficult to be detected from the air. Uncertainty of tree-level estimates, like tree height, diameter at breast height (DBH) and the modeled stem volume also contributes to the inaccuracy of the plot-level estimates. The doctoral work tackled the under-estimation problem from the perspectives of both tree and plot levels. At the plot level, suppressed trees were retrieved from the left tail of the tree size distributions derived from the area-based approach. At the tree level, DBH of single trees were predicted using the quantile-based nearest neighbor imputation.

Area-based approach (ABA) - based forest inventory is able to provide accurate and unbiased plot-level estimates of forest attributes, such as total stem volume. K-MSN method is used in the ABA to simultaneously predict the forest attributes of interest. If tree-level field measurements are available in the sample plots, it's possible to apply the k-MSN method to predict tree size (DBH or height) distributions for the sample plots. The combination of the ITD-derived tree size distributions with the ABA-derived distributions makes it possible not only to improve the ABA-derived saw log estimates, but also to retrieve the suppressed trees for the ITD-derived tree size distributions. The replacement and the histogram matching were utilized to calibrate the tree size distributions. The results showed that after the calibration, the RMSE of the predicted total volume decreased by 2 %, and the bias was negligible. The quantile-based nearest neighbor imputation was able to predict the DBH as accurate as the benchmarking method, the k-MSN. It utilized the ALS-measured tree height and crown diameter as the two predictor variables and achieved even better accuracy than the k-MSN method for larger trees with diameters ≥ 16 cm. The improved DBH estimates also benefit stem volume estimation, which led to the improved tree- and plot-level estimations of the stem volume.

Keywords: diameter distribution, height distribution, suppressed trees, nearest neighbor, cut point detection, stem volume

ACKNOWLEDGEMENTS

The doctoral work presented in the thesis was carried out at the School of Forest Sciences, University of Eastern Finland, during the years 2009-2014.

I am thankful to my supervisor, Professor Timo Tokola for his support and guidance over the years. I would also like to express my gratitude to Professor Matti Maltamo for his constructive opinions and discussions during the research work. I would also like to thank my colleagues for the help in these years.

I would thank my husband Zhengyang Hou for his encourage, and my daughter Muru Hou for her close and warm company.

This doctoral work was funded by the Spearhead project of the University of Eastern Finland: Multi-scale geospatial analysis of forest ecosystems, the Finnish Cultural Foundation central fund, the Graduate School of Forest Sciences and the doctoral program of forest sciences at UEF. I appreciate their support very much.

Joensuu,
Oct. 8, 2014

Qing Xu

LIST OF ORIGINAL ARTICLES

The thesis consists of a summary followed by three articles listed below. Articles I and II are reprinted with kind permission of the publishers, while article III is author's version of the submitted manuscript.

I. Xu Q., Hou Z., Maltamo M., Tokola T. (2014). Calibration of area based diameter distribution with individual tree based diameter estimates using airborne laser scanning. ISPRS Journal of Photogrammetry and Remote Sensing 93: 65-75. doi: 10.1016/j.isprsjprs.2014.03.005

II. Xu Q., Hou Z., Maltamo M., Tokola T. (2014). Retrieving suppressed trees from model-based height distribution by combining high and low density airborne laser scanning data. Canadian Journal of Remote Sensing 40 (3): 233-242. doi:10.1080/07038992.2014.935933

III. Xu Q., Hou Z., Vauhkonen J., Maltamo M., Tokola T. (2014). DBH prediction for the ALS-detected trees using the quantile-based nearest neighbor imputation. Manuscript.

Qing Xu is the corresponding author of all three papers and fully responsible for the data analysis and paper writing. Professor Timo Tokola contributed to the thesis in terms of conception of the research idea and supervision. Professor Matti Maltamo provided his expertise in the diameter distribution and valuable comments to the manuscripts. Zhengyang Hou inspired the research through the continuous discussion about the technical details. Dr. Jari Vauhkonen is responsible for DBH estimates using the K-MSN method, which is compared with the proposed quantile-based nearest neighbour method in the Study III.

TABLE OF CONTENTS

1 INTRODUCTION

The purpose of forest inventory is to obtain reliable and unbiased estimates of forest attributes. Forest inventory is conducted at different scales or levels, the most detailed of which is the tree-level inventory for the purpose of long-term management planning (Siitonen 1993). Tree-level description of forest attributes has a greater flexibility in constituting a larger scale description by up-scaling tree-level estimates into plot, stand, and even larger scales. Detailed tree dimensions measured at a given accuracy facilitate the forestry applications such as growth modeling and bucking simulation (Vauhkonen 2010). In nowadays when the automation prevails, remote sensing data acquired from the air are widely used to automatically detect single trees in the forest stands. Studies on the interpretation of aerial photographs for single tree information started in the 1990's (Pollock 1994; Gougeon 1995; Larson 1997; Brandtberg 1999; Pitkänen 2001, Korpela 2004). Individual tree detection (ITD) algorithms and techniques (Hyyppä and Inkinen 1999; Holmgren et al 2003) were also developed to facilitate the tree-level analysis using the high resolution airborne scanning data.

Airborne laser scanning (ALS) data have attracted forest researchers' attention since late 1990's. Since the ALS data are able to depict the three dimensional (3D) information about tree crowns, they are considered to favor forest applications better than other remote sensing data do (Hyyppä and Hyyppä 1999; Magnussen 2006; Maltamo et al. 2006a). High-density ALS data are able to explicitly describe the crown shape of single trees, and they are also able to provide more information about tree stems and lower blanches when the pulses penetrate deeper towards the ground. Segmentation-based tree crown delineation is normally applied to detect single trees, and tree attributes like height and crown variables are obtained directly from the measurement of the segments. The uncertainties of the tree attributes are accumulated when they are used as predictor variables in the allometric equations of DBH (Kalliovirta and Tokola 2005) and of stem volume (laasasenaho 1982). Tree-level uncertainties have a great impact on the plot-level accuracy of the forest attributes, despite of the up-scaling methods. The errors in the single-tree remote sensing propagate through the processing chain of the system (Korpela 2006). Nevertheless, the impact of undetected trees on the plot-level accuracy should not be underestimated.

Suppressed trees beneath the dominant layer are difficult to be detected by the ALS data. Canopy height model (CHM), which is a surface model of the point cloud, is usually used in the segmentation to detect trees (Persson et al. 2002; Vauhkonen et al. 2010). In multi-storied forests, the CHM mainly describe trees in the dominant layer, whereas in the open areas, it may contain some information about the suppressed trees. Detection algorithms based on the CHM inevitably lead to weak detectability of the suppressed trees for vertically complex forests. Yao et al. (2012) integrated CHM-based segmentation with 3D segmentation techniques to detect trees, and improved the detection rate from 48% to 60%. Hyyppä et al. (2012) utilized the last pulse data for tree detection and correctly matched 6% more trees than when the first pulse was used. Lähivaara et al. (2014) utilized the same data as the doctoral work, and developed a Bayesian approach to detect trees by fitting multiple 3D crown height models to point cloud data. They achieved an improvement in detection rate from 53% to 70%. Amelioration in the detection algorithms helps to detect more suppressed trees, but still their detectability is constrained by the density and spatial pattern of trees. Detection of the suppressed trees continues to be a problem, especially for multi-storied forests. Tree-level uncertainties and undetected trees work together, and result in under-estimation for the total forest attributes in the ITD-based forest inventory. This is a systematic under-estimation that stems from the intrinsic mechanism of the ITD system. Contribution of the suppressed trees is not counted in, so that mean forest attributes, such as mean tree height, are over-estimated, while total forest attributes, such as total stem volume, are under-estimated, 12% as the relative bias at worst according to Korpela (2004).

The ALS-based forest inventory has become realistic and feasible since 2002 when Norway started applying the area-based approach (ABA) (Næsset 2002) in the ALS-based forest inventory. Area-based approach is an alternative method to the individual tree detection. It is able to provide accurate and unbiased plot- and stand-level estimates of forest attributes (Næsset 2004; Holmgren 2004; Maltamo et al. 2006b). Non-parametric k-MSN method is usually used in ABA to simultaneously predict forest attributes of interest, because it is not restrained by the normality assumption and the number of predictors. Diameter distributions can also be accurately predicted using the area-based approach (Maltamo and Gobakken 2014). Diameter distribution is an important indicator of the stand vertical and horizontal structures that are related to forest biodiversity (Maltamo and Gobakken 2014). The shape of diameter distributions corresponds to the development class of the stand and it can be uni-modal, multi-modal, irregular or decreasing (Esseen et al. 1997). Traditional studies based on the field measurement of stand attributes (Magnussen 1986; Shiver 1988; Tham 1988) explored the potential of the theoretical probability density functions such as Weibull distribution and Johnson’s SB distribution, in describing the stand dynamics in different types of forests. They found that the theoretical distributions were not capable to describe the multi-modal and irregular shapes of diameter distributions. Cao and Burkhart (1984) developed a segmented approach of the Weibull distribution, and Border et al. (1987) presented a percentile-based method, both for a more flexible description of the diameter distributions. Recently, Zhang et al. (2001), Liu et al. (2002) and Zasada and Cieszewski (2005) utilized the mixture distribution approach to model the irregular and multimodal diameter distributions of the mixed and uneven-aged forest stands.

Among the ALS-based parametric studies for diameter distributions, Gobakken and Næsset (2004, 2005) used the ALS metrics as the predictors, and applied the parameter prediction and parameter recovery methods to predict the Weilbull parameters. Maltamo et al. (2006b) and Holopainen et al. (2010) used the ABA-derived estimates of the mean stand attributes as the predictors, and they applied the existing parameter prediction models to predict the diameter distributions. Non-parametric method is an alternative of the parametric prediction of diameter distributions. Packalén and Maltamo (2008) applied the k-MSN method to predict the species-specific diameter distributions using the ALS and optical data. They concluded that the k-MSN method was able to predict as accurate diameter distributions as the Weibull distribution, but it was favored due to the possibility to predict multi-modal diameter distributions. ABA-derived diameter distribution is accurate enough when taken as a whole in the calculation of other forest attributes, but not accurate enough if details are concerned. This explains why timber assortments are predicted in ABA with lower accuracies than the total stem volume (Holopainen et al. 2010).

The combination of the individual tree detection with the area-based approach inspired some studies, from the perspective of tree size distributions. Maltamo et al. (2004) retrieved the undetected trees from the theoretical Weilbull distributions, the parameters of which were regressed with the stand-level estimates derived from the ITD-detected trees. Mehtätalo (2006) fitted a continuous distribution function to the observed crown areas, in order to recover the suppressed trees. Lindberg et al. (2010) applied the k-NN method to predict the ABA-based target distribution matrix at plot level, according to which the ITD-derived tree list was calibrated. These studies either improved the ITD-derived estimates of total volume and stem number, or resulted in a tree list consistent with the unbiased ABA-derived estimates.

Tree size distribution derived from the ITD is a summation of all trees that are detected from the air. DBH of these trees need to be predicted from the ALS-measured tree attributes, like tree height and crown dimensions. The ITD is able to derive pretty intact and accurate right tail of the distribution, which consists of larger trees in the dominant layer. The left tail is truncated and fragmented, with the smallest trees missing, as well as some trees in the intermediate layer. The combination of the ITD-derived distribution with the ABA-derived distribution makes it possible to improve the ABA-derived saw log estimates. Besides, for the ITD-derived distribution, it enables to retrieve the suppressed trees from the ABA-derived distribution. Namely, the left tail of the calibrated distribution comes from the ABA-derived distribution, whereas the right tail comes from the ITD-derived distribution. The calibrated tree size distribution is assumed to generate more accurate estimates for the total stem volume of forest stands.

The objective of the doctoral work was to improve the plot-level estimates of forest total stem volume by combining the ABA and ITD, the two main ALS-based approaches in forest applications. It tackled the under-estimation problem of the ITD system from two aspects. One was focused on the improvement in tree-level estimates, such as DBH and stem volume, the other worked with the tree size distributions derived from both approaches, and sought the improvement by combining the two distributions. The explicit objectives of the sub-studies were detailed as follows:

I. To automatically detect the proper cut point which divides diameter distribution into left tail and right tail for each sample plot, based on the diameter distributions derived from both approaches. To calibrate the ABA-derived diameter distribution with the ITD-derived diameter estimates, for more accurate diameter distributions and total volume estimates.
II. To improve the calibration scheme based on the height distributions. To examine the calibration effects on the suppressed trees. To explore the necessity of the calibration.
III. To develop a quantile-based nearest neighbor imputation for DBH prediction, using the ALS-measured tree height and crown diameter as predictors.

2 STUDY AREA AND MATERIAL

2.1 Study area and field measurements

All three sub-studies were focused on a managed boreal forest in Kiihtelysvaara, in eastern Finland (62°31’ N, 30°10’ E), where the main tree species are Scots pine (*Pinus sylvestris*), Norway spruce (*Picea abies*) and deciduous trees. The climate is sub-arctic with high humidity.

A total of 5747 trees in 79 square sample plots (Figure 1) were measured in the field during May-June 2010 for tree species, DBH, and height. Size of the sample plot was determined by the mean tree size and stand development class (Table 1), and it varied from

20, 25 to 30 m for the side length of the square plot. All trees were measured if DBH ≥ 4 cm or height ≥ 4 m. ITD-detected trees were first geo-referenced, and undetected trees were registered using the least square adjustment (Korpela et al. 2007) based on the angle and distance from the undetected tree to the detected trees. Stem volumes were calculated by the species-specific equations presented by Laasasenaho (1982) using DBH and tree height as predictors. Forest attributes such as basal area, stem number, diameter and height of the basal area median tree, stem volume, saw log and pulpwood volumes, were calculated and scaled to the hectare level. The average total volume at plot level was 197.59 m^3/ha, and the average stem number was 1258 /ha.

2.2 ALS data

Two separate ALS datasets were utilized in the doctoral work. High density ALS data were collected on June 26th 2009 using an Optech ALTM Gemini laser scanning system, approximately 600 m above the ground level. Along with a field of view of 26 degrees and a pulse repetition frequency of 100 kHz, the scanner setup resulted in a nominal sampling density of about 12 measurements per square meter, and a swath width of approximately 320 m. Side overlap was 55%, to make sure that each location was measured from two flight lines, namely two sides. Multiple echoes were recorded for each pulse. Low density ALS data were collected on July 18th 2009 using the same Optech ALTM Gemini laser scanning system, approximately 2000 m above the ground level. The swath width of approximately 1050 m and the sampling density of about 0.65 measurements per square meter were obtained using a field of view of 30 degrees, side overlap of 20% and pulse repetition frequency of 50 kHz.

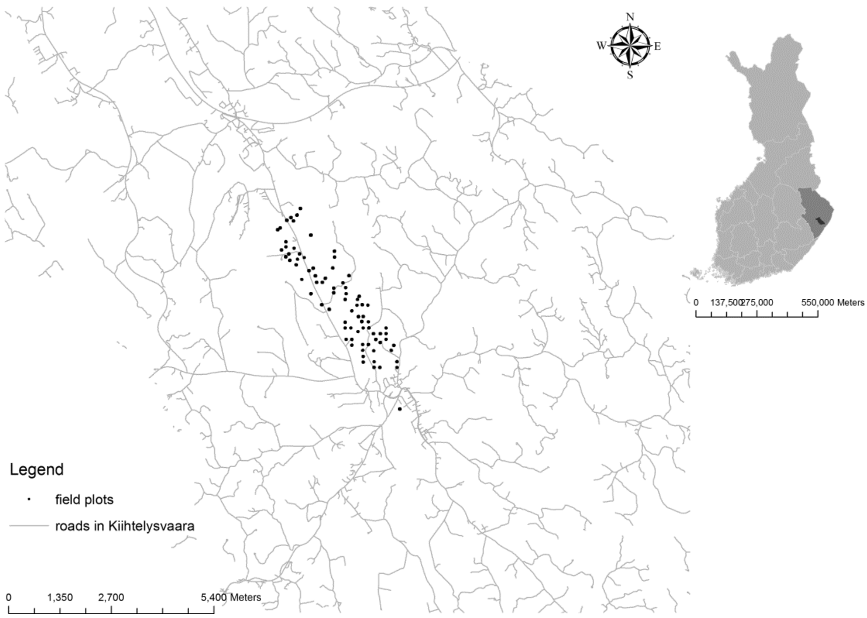

Figure 1. Sample plots in the study area.

Table 1. Range and mean value of the tree attributes of three categories of sample plots.

Plot size (m)	Stem number	Height (m)	DBH (cm)
20	33-74-115	9.06-12.29-17-62	8.78-12.14-19.05
25	32-72-160	8.70-14.11-20.70	8.14-14.88-21.84
30	42-74-118	11.41-17.68-24.11	12.11-19.11-28.40

3 METHODS

The outlines of the methods involved in all papers were presented in Figure 2.

Figure 2. Flowchart of the calibration methods of the ABA-derived tree size distribution with the ITD-derived tree size distribution.

3.1 ABA-based prediction of tree size distributions using the low density ALS data

3.1.1 ALS metrics as predictor variables

Low density ALS data were used to extract the ALS metrics in the area-based approach. After generating the Digital terrain model (DTM) with a pixel size of 0.5 m by taking the mean value of the ground points within the pixel, the aboveground heights of the ALS points were obtained by subtracting the DTM from the orthometric heights of the ALS points. Canopy hits, namely the ALS points with height ≥ 0.5 m, of the first and last returns, were used respectively to calculate the ALS metrics. Canopy height percentiles at 0, 1, 5, 10, 20, 30, 40, 50, 60, 70, 80, 90, 95, 99 and 100 %, as well as proportional canopy densities for these percentiles, were calculated. The ratio of canopy hits to ground hits, the mean height of the canopy hits, and the standard deviation of the canopy hits were also computed. A total of 64 ALS metrics were extracted for variable selection and modeling.

3.1.2 Variable selection

LASSO (Least Absolute Shrinkage and Selection Operator) (Tibshirani 1996) is a method for variable selection. The entire path of LASSO solutions was computed with the least angle regression (LAR) (Efron et al. 2004), using the package LARS in the R statistical computing environment (http://www.r-project.org, 2014). By constraining the sum of the absolute values of the estimated coefficients no bigger than the bound s ($s \geq 0$), some of the coefficients are shrunken to zero. When s varies from zero to infinity, LASSO has different solutions, one of which is the result of the best s that leads to the optimized selection of predictors. Ten-fold-cross-validation was applied in the study to solve the best s for each dependent variable (total and species-specific stem volumes, saw log and pulpwood volumes, stem number, basal area, diameter and height of the basal area median tree). Predictors selected for each dependent variable were integrated in two different ways that resulted in three groups of predictors, including the original set of predictors.

3.1.3 Prediction of tree size distributions using the k-MSN

Two groups of dependent variables, denoted as Y1 and Y2 were tested in the k-MSN imputation (Moeur and Stage 1995; Packalén and Maltamo 2008). Y2 consisted of all Y1 dependent variables except saw log and pulpwood volumes. These two groups were further subdivided into total variables and species-specific variables. The best model was to be selected as the final MSN model to predict tree size distributions. The empirical tree size (DBH or height) distributions (denoted as $\mathrm{df}_{\mathrm{true}}$ and $\mathrm{hf}_{\mathrm{true}}$) of all reference plots are required in the k-MSN imputation. They were specified by the discrete frequency function of the field-measured DBH or height in each sample plot, with class width of 1cm for diameter distributions, and 1m for height distributions. The empirical tree size distributions were expanded to the hectare level by multiplying the area factor, the ratio of one hectare to the plot area. The predicted distribution for the target plot is the distance-weighted mean of the empirical distributions of the most similar neighbor (MSN) plots, which were searched using the R package yaImpute. The distance from the target plot to the reference plots were calculated based on the Canonical Correlation Analysis and three MSNs were utilized in the imputation. The ABA-derived diameter and height distributions were obtained simultaneously and expanded to the hectare level and denoted as $\mathrm{df}_{\mathrm{aba}}$ and $\mathrm{hf}_{\mathrm{aba}}$, with class widths identical with the empirical distributions.

3.2 ITD-based prediction of tree size distributions using the high density ALS data

3.2.1. Tree delineation and measurement of tree height and crown diameters

The segmentation was based on the canopy height model (CHM) produced with the high density ALS data. The CHM was first smoothed using height-based filtering (Pitkänen et al. 2004), and searched for local maxima. Watershed segmentation was applied to delineate tree crowns (segments) around the local maxima. The maximum height in each segment was deemed tree height, and the tree position was obtained from the X and Y coordinates of the pixel that had the maximum height value. The maximum crown diameter and the other that took the direction perpendicular to the former were measured for each segment.

3.2.2 Prediction of height distributions

All ALS trees that located within a certain distance from the field tree were associated with the field tree. The distance was a function of the DBH of the field tree (Olofsson et al. 2008). The linked ALS tree was determined by the shortest Euclidian distance (Olofsson et al. 2008) between tops of the field tree and all associated ALS trees. In case that the same ALS tree was linked to several field trees, tree height difference was calculated to select the pair that had the least difference. All tree pairs that had height difference ≥ 4m were excluded. After the ALS trees were linked with the field trees, the ALS-measured tree heights were calibrated with the field-measured heights using spline regression (Harrell 2001), with 5 knots at the quantiles of 0.05, 0.275, 0.5, 0.725, and 0.95. The ITD-derived height distribution for each sample plot was specified by the discrete frequency function of the calibrated ALS-measured tree heights, with class width of 1m. It was expanded to the hectare level, and denoted as $\mathrm{hf}_{\mathrm{itd}}$.

3.2.3 Prediction of diameter distributions

Since DBH cannot be directly measured from the delineated tree crowns, the ITD-derived diameter distribution depends on the prediction of DBH using the variables that can be directly measured. The allometry-based modelling method used here was a quantile-based nearest neighbor method, in the assumption that a specific quantile of the nearest neighbors represents the target object. The nearest neighbors were determined by the ALS-measured tree height. The probability value that specified the quantile was determined by the ALS-measured crown diameter calibrated with the DBH. All variables used here were in relative magnitude, in order to offer a natural input to the probability value, and also to facilitate the calibration of crown diameter with the DBH. Relative value was defined, for example, as the ratio of the crown diameter of a tree to the maximum crown diameter in the target dataset. After DBH prediction, the ITD-derived diameter distribution was specified by the discrete frequency function of the predicted DBHs of the ITD-detected trees, with class width of 1 cm. It was then expanded to the hectare level, and denoted as $\mathrm{df}_{\mathrm{itd}}$ for later reference.

3.3 Calibration methods that combine the ABA and ITD-derived tree size distributions

3.3.1 Detection of the cut point for diameter and height distribution respectively

The cut point used in this study was a unique integer diameter or height value that distinguishes trees with larger diameters or heights from trees with smaller diameters or heights in each sample plot. The automatic detection of the cut points was based on the ITD-derived tree size distributions, and was realized in a two-step procedure. The first step examined the continuity and reliability of the ITD-derived distributions and selected the candidate cut points for each plot. Reliability was judged with the help of the empirical distribution. Continuity was recognized if the majority of the sequential five classes have trees, starting from a reliable class (including itself). The second step examined whether the candidates can be seem from the air. Sigmoid function (Korpela 2004, Mäkinen et al. 2010) that modeled the detection probability of a tree from the air is a function of the relative height of the tree. The relative height was defined as the ratio of the tree height to the maximum height of the sample plot. For height distributions, candidate cut points was used directly in the sigmoid function, whereas for diameter distributions, tree height was modeled for each candidate cut point, using the nonlinear Korf H-D function fitted to all field-measured trees in the corresponding sample plot. In each sample plot, the smallest candidate that had ≥ 0.95 detection probability was chosen as the final cut point.

3.3.2 Replacement

The purpose of replacement is to constitute the calibrated distribution with the left-tail of the ABA-derived distribution and the right-tail of the ITD-derived distribution. The cut point detected for each sample plot was used to separate the left and right tails, and the cut point is the starting point of the right tail. The intermediate result here was denoted as f_{rep} with a suffix d or h added ahead to indicate diameter or height. Since the ABA-derived distributions obtained trees from the most similar neighbor plots, they sometimes had larger trees than their own. These trees were referred to as the ABA tail in the study. After the replacement, five different ways to deal with the ABA tail resulted in five different calibrated distributions. The total basal area was calculated for all trees in the ABA tail, and was then distributed to the anterior part, namely f_{rep}. Table 2 listed the five differently calibrated distributions and how they were derived.

Table 2. Five differently calibrated tree size distributions after the replacement.

Abbreviation	Description
f_{dis0}	ABA tail was ignored
f_{dis1}	ABA tail was allocated evenly to the right tail of f_{rep}
f_{dis2}	ABA tail was allocated evenly to the whole section of f_{rep}
f_{dis3}	ABA tail was allocated proportionally to the right tail of f_{rep}
f_{dis4}	ABA tail was allocated proportionally to the whole section of f_{rep}

Table 3. Three calibrated diameter distributions following the histogram matching.

Abbreviation	Description
f_{mat1}	Intersection of the target and the reference was used as the turning point
f_{mat2}	Cut point was used as the turning point
$f_{mat2.g5}$	5 cm after the cut point was used as the turning point

3.3.3 Histogram matching

Histogram matching (Gonzalez and Woods 2002) was only used for diameter distribution study. The ABA and ITD derived diameter distributions were converted to the cumulative probability functions first, and were used as the reference $R(j)$ and the target $T(i)$ respectively in the histogram matching. After matching the target with the reference, we obtained the calibrated target which was later converted into the diameter distribution or histogram. For discrete frequency function, it's possible to find two points j and $j + 1$ in the reference for a given $T(i)$ in the target, and it's also possible to utilize the distance between i and j to decide how much the target should resemble the reference. The essential idea was that the target should resemble the reference for the left tail, but it had more liberty in the right tail. The cut point detected in 3.3.1 was first tested here as the turning point, from which the target don't have to resemble the reference. After a careful inspection of the calibrated target of all sample plots, 5 cm after the cut point was further tested. And the intersection of the target and the reference was also tested for this purpose. Table 3 listed the three calibrated diameter distributions.

3.3.4 Height-to-diameter matrix for height distributions

In the height distribution study, to calculate the basal area for trees in the ABA tail and the total stem volume from the height distribution, DBH has to be predicted for trees of the mean heights in all height classes. Instead of predicting DBH within each own system, a standard height-to-diameter matrix that took plots as rows and height classes as columns, was established. The field-measured trees were used to generate each element of the matrix. For each sample plot, all field-measured trees that fall in each height class were extracted; the mean DBH of these trees was filled into the corresponding height class, as the predicted DBH for the mean height tree of the class. The preliminary height-to-diameter matrix had some elements filled with no data because no field trees existed in these height classes. Second, DBH was modeled for these no data elements, as a function of tree height, using all field-measured trees.

3.4 Assessment methods

The accuracies of the predicted distributions were assessed using the Reynolds error index (Reynolds et al. 1988) (Eq. 1), and the RMSE and bias of the predicted total stem volume (Eq. 4 and 5). In diameter distribution study, the sum of powers of diameters (Maltamo 1997) (Eq. 3) was used as the approximation of total stem volume, whereas in height distribution study, species-specific volume equations by Laasasenaho (1982) were applied to predict stem volume for single trees. The entire growing stock, the saw log fraction and the pulpwood fraction were respectively calculated. Diameter value 17 cm and height value 16 m were used to partition the saw log and pulpwood fractions. Packaléns error index (Packaléns and Maltamo 2008) (Eq. 2) was also used to assess the ABA-derived diameter distributions for determining the final MSN model.

$$\mathrm{EI}_{\mathrm{Rey}} = \sum_{c=1}^{m} 100 \times \left|\frac{f_c - \widehat{f_c}}{N}\right| \tag{1}$$

$$\mathrm{EI}_{\mathrm{Pac}} = \sum_{c=1}^{m} 0.5 \times \left|\frac{f_c}{N} - \frac{\widehat{f_c}}{\widehat{N}}\right| \tag{2}$$

where c is the diameter class, and $c = 1, 2, \ldots, m$. f_c and $\widehat{f_c}$ are field-measured and predicted stem frequency in the diameter class c, respectively. N and $\widehat{N}$ are field-measured and predicted total stem number per hectare.

$$D^p = \sum_{c=1}^{m} d_c^p \times f_c \tag{3}$$

where d_c is the mean diameter of the diameter class c. p is the power, $p = 1, 3$. f_c is the stem frequency in the diameter class c, and $c = 1, 2, \ldots, m$.

$$\mathrm{RMSE} = \sqrt{\frac{\sum_{i=1}^{n}(V_i - \widehat{V_i})^2}{n}} \tag{4}$$

where i is the plot index, and $i = 1, 2, \ldots, n$. V_i and $\widehat{V_i}$ are the empirical and predicted total stem volume in plot i in the height distribution study, but in the diameter distribution study, they are D^p of the empirical and predicted diameter distributions.

$$\mathrm{Bias} = \frac{\sum_{i=1}^{n}(V_i - \widehat{V_i})}{n} \tag{5}$$

4 RESULTS

4.1 ABA-derived tree size distributions

For the ABA, the best MSN model that had the least Reynolds and Packalén error indices was selected as the final MSN model to predict both the diameter and height distributions. It had 49 predictor variables and species-specific attributes (Y2) as response variables. Fig. 3 presented an example of the ABA-derived diameter and height distributions, using the 3rd sample plot which was also used to demonstrate the calibrated distributions in the following figures. The ABA-derived tree size distributions were able to depict the general shape of the empirical distributions, such as the number of summits. But they were not accurate enough if the pulpwood and saw log fractions were inspected separately. The loss of trees around 20 cm and 20 m was compensated by trees in the ABA tail, which came from the most similar neighbor plots of this specific sample plot.

4.2 ITD-derived tree size distributions and cut points

Figure 4 presented an example of the ITD-derived diameter and height distributions, using the same sample plot as Figure 3. As assumed, the ITD failed to detect the suppressed trees. However, it did a pretty good prediction for the dominant trees. It's very possible that the diameters of the largest trees were underestimated. Table 4 listed the accuracies of the predicted DBH using the quantile-based nearest neighbor method. Larger trees were predicted with worse accuracy and precision. The heights of the largest trees, after calibration, were very close to the empirical values. The bin sizes used in the study were very small, 1 cm for the diameter distribution and 1 m for the height distribution. If larger bin sizes were used, the visual similarity between the predicted and the empirical would be stronger. Each sample plot had its own cut points. For this plot, the diameter cut point was 17 cm, and the height cut point was 19 m. The cut point indicated the upper bound of the class, that's to say, the calibration started from the classes of 16-17 cm and 18-19 m. The cut points were drawn in the middle of the classes in Figure 4 and 5. There is still room for improvement of the automatic detection algorithm for the cut point. Compared with the ABA-derived tree size distributions, the ITD derived much more accurate diameters and heights for dominant trees, which mostly contributed to the saw log volume.

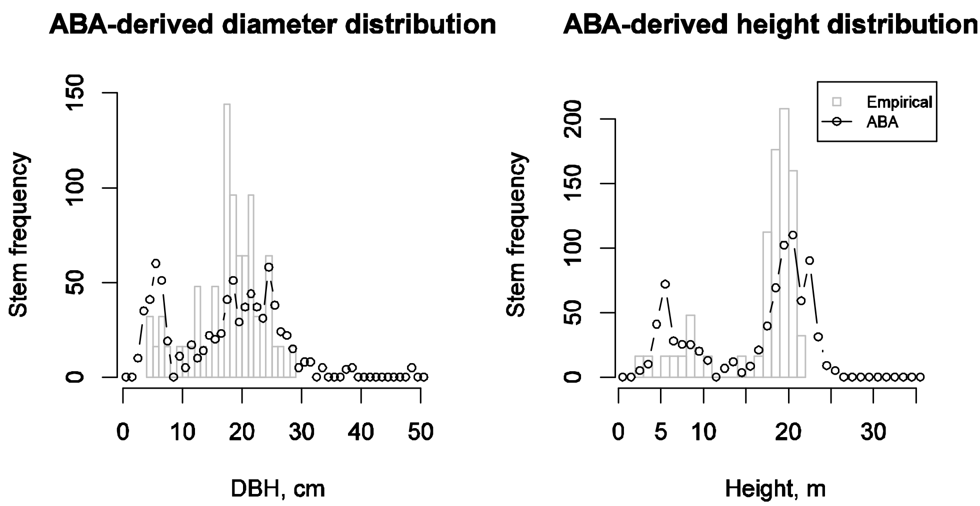

Figure 3. ABA -derived diameter and height distributions.

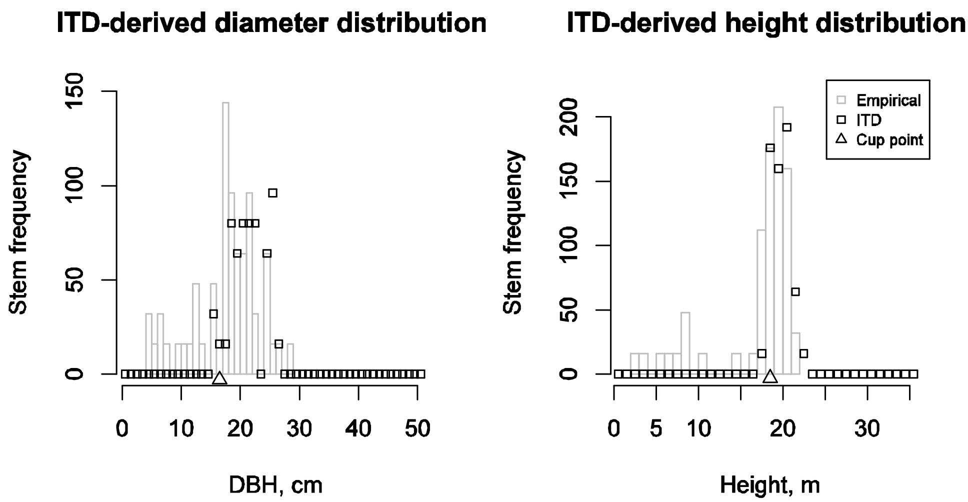

Figure 4. ITD -derived diameter and height distributions.

Table 4. Accuracy of the predicted DBH for all trees and larger trees with DBH ≥16 cm.

	total		pine		spruce		deciduous	
Trees	RMSE (%)	Bias (%)	RMSE	Bias	RMSE	Bias	RMSE	Bias
All	3.01 (15.64)	-0.07 (-0.38)	2.77	0.04	3.20	0.80	3.89	-1.82
≥16 cm	3.10 (13.63)	0.41 (1.82)	2.89	0.40	3.31	1.00	4.20	-0.95

4.3 Calibration of the tree size distributions

Figure 5 presented the optimally calibrated diameter and height distributions for the example plot, according to the Reynolds error index and the total volume estimates. The diameter distribution was the result of f_{mat1} , following the histogram matching. The calibrated distribution resembled the reference distribution, which was the ABA-derived distribution very much. The height distribution was the result of f_{dis0}, following the replacement method.

4.4 Assessment of the tree size distributions

In paper I, the sum of the powers of diameters was calculated as the approximation of the total volume. The relative RMSE and bias of the two basic and eight calibrated diameter distributions for the entire growing stock, pulpwood and saw log fractions are presented in Table 5. For the approximated total volume estimates, $f_{mat2.g5}$ improved the accuracy from 21% (the best between the ABA and ITD) to 19%. Such improvement was brought by the improvement in both the pulpwood and saw log fractions. At the same time, the calibration decreased the bias of the ITD estimates.

In paper II, Laasasenaho (1982)'s volume equations were used to calculate the stem volume. Table 6 presented the accuracies of the total volume and pulpwood volume calculated from the height distributions. All calibrated height distributions lowered the RMSE of the total volume by 2% the most. The improvement in the pulpwood fraction was larger.

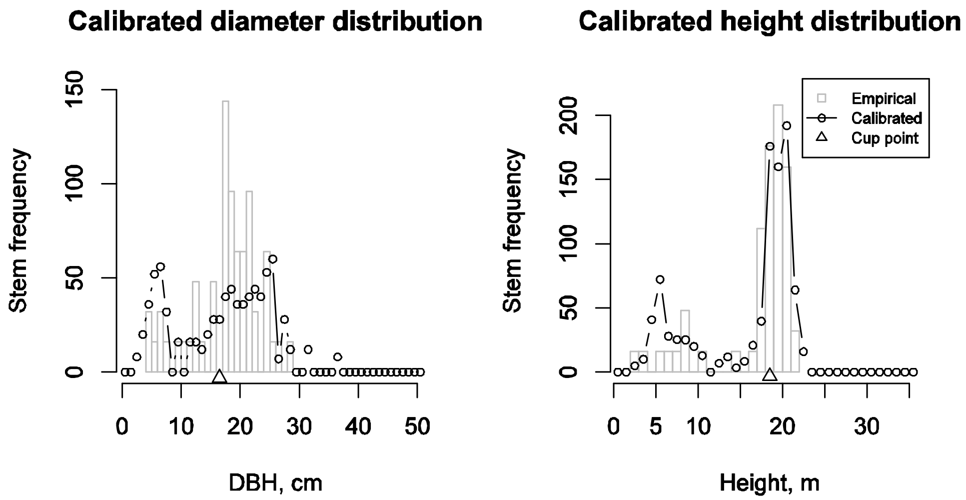

Figure 5. Calibrated tree size distributions.

Table 5. Accuracies of the total volume, pulpwood and saw log fractions generated from the diameter distributions (improvements are underlined).

	Total volume		Pulpwood volume		Saw log volume	
	RMSE %	Bias %	RMSE %	Bias %	RMSE %	Bias %
f_{aba}	21.41	2.38	49.37	-0.93	28.84	3.03
f_{itd}	25.72	16.35	50.32	32.05	26.83	13.24
f_{dis0}	22.97*	6.89	41.64**	0.52	27.43	8.16
f_{dis1}	21.82*	-3.05	42.65	-10.61	26.48	-1.55
f_{dis2}	21.18*	-0.09	44.78	-16.18	25.84	3.10
f_{dis3}	21.98*	-3.23	42.34	-9.03	26.69	-2.08
f_{dis4}	21.75*	-2.43	43.45	-11.68	26.40	-0.60
f_{mat1}	19.79*	4.99	47.42*	-5.36	26.01	7.04
f_{mat2}	21.05	-9.23	49.83	5.23	26.84	-12.09
$f_{mat2.g5}$	18.60*	-1.53	47.69	-2.25	23.78	-1.39

a. * indicates residual of fitd is significantly greater than that of the calibrated distributions (p≤0.05).
b. * indicates residual of faba is significantly greater than that of the calibrated distributions (p≤0.05).

Table 6. Accuracies of the total volume and pulpwood fraction calculated from the height distributions.

	Total volume		Pulpwood volume	
	RMSE %	Bias %	RMSE %	Bias %
f_{aba}	28.84	-6.23	60.21	2.57
f_{itd}	18.66	10.39	61.25	44.60
f_{dis0}	16.70	3.86	47.13	14.70
f_{dis1}	17.41	-0.76	43.74	10.40
f_{dis2}	16.71	0.76	46.18	5.10
f_{dis2}	17.54	-0.93	43.57	10.45
f_{dis4}	17.27	-0.41	45.46	9.01

5 DISCUSSION

The doctoral work focused on the accuracy of the total volume or the approximated total volume in the ITD-based forest inventory, which were explored from two levels. The tree-level study (Paper III) emphasized on the accuracy of the predicted DBH and stem volume of the ALS-detected single trees. The plot-level studies (Paper I and II) paid close attention to the total volume estimates generated from the tree size distributions, which were the union between the ABA and ITD-derived tree size distributions. Since the total volume generated from the tree size distribution is a summation of all trees, tree-level accuracy and the presence of all trees influence the plot-level accuracy of the predicted total volume. In the ITD system, tree detection rate is an essential parameter to indicate the forest vertical structure and the possible prediction accuracy. Single-storied forests, with a large detection rate, are possibly accurately estimated, while the multi-storied forests, with around fifty percent of trees detected, are certainly underestimated. Whether trees are detected and how many are detected by the ALS data has a great impact on the plot-level accuracy.

Two plot-level studies developed the calibration procedures, in which un-detected trees in the ITD system were retrieved from the ABA-derived tree size distributions. The results showed that more accurate tree size distributions and more accurate total volume estimates were obtained. The improvement was contributed mostly by the pulpwood fraction, namely the suppressed trees that were theoretically undetectable from the air. Two calibration methods were tested in the studies, and they have different mechanisms. The calibration result of the replacement method largely depends on the accuracy of the ITD, because all dominant trees detected in the ITD system are kept in the calibrated tree size distributions. The calibration result of the histogram matching is more flexible in that, after the turning point, at least three alternatives are available. They are following the ITD, following the ABA, and going between the ABA and the ITD. Paper I explored the last alternative, which was a compromised scheme for the dominant trees between the ABA and the ITD. The flexibility in the histogram matching could be further explored in the future studies on the species-specific calibrations.

The cut point detection algorithm was an automatic procedure, which examined the reliability and continuity of the ITD-derived tree size distributions first. The reliability was examined with the help of the empirical distributions in the studies. More efforts could be made in the future to realize the same purpose without the help of the empirical distributions, because for wider forest area that is not visited in the field campaign, the empirical tree size distributions are not available. The continuity is examined because the continuous detection of trees is needed all along the distribution, especially from the cut point on. Finally the visibility of the tree of the same size of the cut point from the air was examined. The proposed methods for diameter and height distributions were able to detect good enough cut points for each sample plot, and based on these cut points, the plot-level calibrations led to more accurate tree size distributions.

The tree-level stem volume accuracy depends on the accuracies of tree height and DBH, if they are used as predictors for the stem volume. Tree heights can be derived from the ALS data very accurately, especially after the calibration with the field-measured tree heights. DBH of the larger trees is difficult to be accurately predicted because these trees grow in diameter but not much in height any more. The joining of crown variables in the DBH prediction helps to improve the DBH accuracy. The quantile-based nearest neighbor imputation developed in the Paper III utilized the ALS-measured tree height and crown diameter to predict tree DBH. It achieved the same accuracy as the k-MSN method that was commonly used to predict tree-level attributes. Compared with the k-MSN method, the quantile-based nearest neighbor method used much fewer predictor variables, and obtained slightly better accuracy for trees with diameter $\geq$ 16 cm. Furthermore, based on the predicted DBH and the ALS-measured tree height, the quantile-based method resulted in better accuracy in the tree-level stem volume than the k-MSN method did. Alternatively, the stem volume can be directly predicted in the k-MSN method, together with other tree attributes, such as DBH (Vauhkonen et al. 2013). The direct prediction is assumed to obtain better accuracy for the stem volume, due to the non-presence of the modelling errors. The error propagation in the single-tree remote sensing due to indirect prediction of the stem volume was well studies in Korpela (2006), and they also studies the error contributed by the species recognition, which has not been investigated in this doctoral study.

Previous diameter distribution studies (Maltamo et al. 2000; Kangas and Maltamo 2000) based solely on the field-measured data compared the quantile-based diameter prediction method with the Weibull distribution method, and obtained the predicted total volume at the best accuracy of about 6% as the relative RMSE. They concluded that the quantile-based method was able to describe more diverse forms of the diameter distributions, and predicted the stand characteristics at better accuracy than the Weibull distribution method. Siipilehto (2011), based on the field-measured data again, compared the Weibull distribution with the Johnson's SB distribution in predicting the forest stand structure variables. The RMSE% of the predicted total stem volume was about 2% for the advanced stands, but it varied from 10 to 24% for the young stands. These studies utilized the true stand attributes to construct models, free of errors from measurement, sampling and other sources. They were able to achieve a RMSE usually below 10%. When the models are applied in the real world forest applications where the stand attributes are visually estimated, the RMSE will undoubtedly increase. In the study by Haara and Korhonen (2004), visually assessed stand attributes were used to predict diameter distribution and volume. They reported 24.8% as the RMSE, which is comparable or even higher than our ALS-based studies. Besides, the stand-level studies produce much more accurate estimates than the plot- level studies (Packalén and Maltamo 2007). This doctoral work was made at the plot level, and it’s difficult to have a direct comparison with the stand-level field measurement-based studies.

Compared with the studies focused on developing the tree detection algorithms (Wang et al. 2008; Yao et al. 2012; Lähivaara et al. 2014), This doctoral work aimed at the unbiased estimates for the purpose of the ITD-based forest inventory. Maltamo et al. (2004) manually selected the cut point, which was realized automatically in this doctoral work. Compared with the theoretical Weibull distribution that was used in Maltamo et al. (2004), the ABA-derived distribution is able to describe the multi-modal distributions. This means the left tail of the ABA-derived distribution depicts the possible summit around the smaller trees. Lindberg et al. (2010) obtained a calibrated tree list that was consistent with the ABA-derived estimates. They effectively removed the bias but the increase of accuracy was not guaranteed. This doctoral work, based on the ITD-derived dominant trees, either transplanted the ABA-derived suppressed trees, or made slight adjustment to the left tail of the ABA-derived distributions. It was able to increase the accuracy while decrease the bias. In fact, it also took the necessity of the calibration into account for each sample plot, in paper II. The calibration was abandoned for the sample plot, whose most similar neighbors in the k-MSN imputation didn’t reflect its real structure. A threshold was set here to indicate the unnecessary calibration for the sample plot. The ITD-derived tree size distribution was therefore kept and delivered to the calibrated system. This diminished the obvious side effect of the calibration and navigated the calibration into the direction that increased the accuracy. Breidenbach et al. (2010) developed the semi-ITD, which also tackled the underestimation problem of the ITD. Paper II compared the accuracy of the estimated total volume with the semi-ITD accuracy (Vauhkonen et al. 2013), and concluded that they were comparable.

The developed calibration procedures in the doctoral work combined the two main approaches of analyzing the ALS data for forestry applications. The merits of the ABA and the ITD systems are integrated into the calibrated system that offers more accurate information about the pulpwood and saw log fractions, and produces unbiased estimates for the total stem volume. It can be assured that the calibration procedures also apply to the multistoried forests, because of the ability of the ABA-derived tree size distribution to describe multi-modal distributions. Future works will address the species-specific calibration of the tree size distributions, which requires pretty accurate species recognition that helps to diminish the stem volume modeling error among different species. Stand- level calibration of the tree size distributions are also of future interest.

REFERENCES

Borders B.E., Souter R.A., Bailey R.L., Ware K.D. (1987). Percentile-based distributions characterize forest stand tables. Forest Science 33 (2): 570-576.

Brandtberg T. (1999). Automatic individual tree-based analysis of high spatial resolution remotely sensed data. Doctoral thesis. Acta Universitatis Agriculturae Sueciae, Silvestria 118. Swedish University of Agricultural Sciences, Centre for Image Analysis. Uppsala, Sweden.

Breidenbach J., Næsset E., Lien V., Gobakken T., Solberg S. (2010). Prediction of species specific forest inventory attributes using a nonparametric semi-individual tree crown approach based on fused airborne laser scanning and multispectral data. Remote Sensing of Environment 114 (4): 911-924.
http://dx.doi.org/10.1016/j.rse.2009.12.004

Cao Q.V., Burkhart H.E. (1984). A segmented distribution approach for modelling diameter frequency data. Forest Science 30: 129-137.

Efron B., Hastie T., Johnstone I., Tibshirani R. (2004). Least angle regression. The Annals of Statistics 32 (2): 407-499.
http://dx.doi.org/10.1214/009053604000000067

Esseen P.A., EhnstroÈm B., Ericson L., SjoÈberg K. (1997). Boreal forests. Ecological Bulletins 46: 16-47.

Gobakken T., 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.
http://dx.doi.org/10.1080/02827580410019454

Gobakken T., Næsset E. (2005). Weibull and percentile models for lidar-based estimation of basal area distribution. Scandinavian Journal of Forest Research 20 (6): 490-502.
http://dx.doi.org/10.1080/02827580500373186

Gonzalez R. C., Woods R. E. (2002). Digital image processing, second ed. Upper Saddle River: Prentice Hall, New Jersey, Chapter 3.

Gougeon F.A. (1995). A crown-following approach to the automatic delineation of individual tree crowns of high spatial resolution aerial images. Canadian Journal of Remote Sensing 21(3): 274–284.
http://dx.doi.org/10.1080/07038992.1995.10874622

Haara A., Korhonen K.T. (2004). Kuvioittaisen arvioinnin luotettavuus. Metsätieteen aikakauskirja 4: 489-508.

Holmgren J. (2003). Estimation of forest variables using airborne laser scanning. Doctoral dissertation. Acta Universitatis Agriculturae Sueciae. Silvestria 278. SLU, Umeå, Sweden. 43 p.

Holmgren J. (2004). Prediction of tree height, basal area and stem volume in forest stands using airborne laser scanning. Scandinavian Journal of Forest Research 19 (6): 543-553.
http://dx.doi.org/10.1080/02827580410019472

Holopainen M., Vastaranta M., Rasinmäki J., Kalliovirta J., Mäkinen A., Haapanen R., Melkas T., Yu X., Hyyppä J. (2010). Uncertainty in timber assortment predicted from forest inventory data. European Journal of Forest Research 129: 1131-1142.
http://dx.doi.org/10.1007/s10342-010-0401-4

Hyyppä H., Hyyppä J. (1999). Comparing the Accuracy of Laser Scanner with Other Optical Remote Sensing Data Sources for Stand Attribute Retrieval. The Photogrammetric Journal of Finland 16: 5-15.

Hyyppä J., Inkinen M. (1999). Detecting and estimating attributes for single trees using laser scanner. The Photogrammetric Journal of Finland 16 (2): 27-42.

Hyyppä J., Yu X., Hyyppä H., Vastaranta M., Holopainen M., Kukko A., Kaartinen H., Jaakkola A., Vaaja M., Koskinen J., Alho P. (2012). Advances in forest inventory using airborne laser scanning. Remote Sensing 4: 1190-1207.
http://dx.doi.org/10.3390/rs4051190

Kalliovirta J., Tokola T. (2005). Functions for estimating stem diameter and tree age using tree height, crown width and existing stand database information. Silva Fennica 39 (2): 227-248.
http://dx.doi.org/10.14214/sf.386

Kangas A., Maltamo M. (2000). Percentile based basal area diameter distribution models for Scots pine, Norway spruce and birch species. Silva Fennica 34 (4): 371-380.
http://dx.doi.org/10.14214/sf.619

Korpela I. (2004). Individual tree measurements by means of digital aerial photogrammetry. Silva Fennica, Monograph 3.

Korpela I., Tokola T. (2006). Potential of aerial image-based monoscopic and multiview single-tree forest inventory - a simulation approach. Forest Science 52(2): 136-147.

Korpela I., Tuomola T., Välimäki E. (2007). Mapping forest plots: An efficient method combining photogrammetry and field triangulation. Silva Fennica 41(3): 457-469.
http://dx.doi.org/10.14214/sf.283

Laasasenaho J. (1982). Taper curve and volume function for pine, spruce and birch. Publications of Forest Research Institute in Finland. Communicationes Instituti Forestalis Fenniae 108, 74 p.

Lähivaara T., Seppänen A., Kaipio J.P., Vauhkonen J., Korhonen L., Tokola T., Maltamo M. (2014). Bayesian approach to tree detection based on airborne laser scanning data. IEEE Transaction on Geoscience and Remote Sensing 52 (5): 2690-2699.
http://dx.doi.org/10.1109/TGRS.2013.2264548

Larsen M. (1997). Crown modeling to find tree top positions in aerial photographs. In: Proceedings of the third International Airborne Remote Sensing Conference and Exhibition, Copenhagen, Denmark. ERIM international. Vol 2: 428–435.

Linberg E., Holmgren J., Olofsson K., Wallerman J., 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.
http://dx.doi.org/10.1080/01431160903380649

Liu C., Zhang L., Davis C.J., Solomon D.S., Gove J.H. (2002). A finite mixture model for characterizing the diameter distribution of mixed-species forest stands. Forest Science 48: 653-661.

Magnussen S. (1986). Diameter distributions in Picea abies described by the Weibull model. Scandinavian Journal of Forest Research 1: 493-502.
http://dx.doi.org/10.1080/02827588609382440

Magnusson M. (2006). Evaluation of remote sensing techniques for estimation of forest variables at stand level. Doctoral thesis summary. Swedish University of Agricultural Sciences, Umeå, Sweden. Acta Universitatis Agriculturae Sueciae 2006, 85. 38 p.

Mäkinen A., Holopainen M., Kangas A., Rasinmäki J. (2010). Propagating the errors of initial forest variables through stand- and tree-level growth simulators. European Journal of Forest Research 129 (5): 887-897.
http://dx.doi.org/10.1007/s10342-009-0288-0

Maltamo M. (1997). Comparing basal area diameter distributions estimated by tree species and for the entire growing stock in a mixed stand. Silva Fennica 31 (1): 53-65.
http://dx.doi.org/10.14214/sf.a8510

Maltamo M., Kangas A., Uuttera J., Torniainen T., Saramaki J. (2000). Comparison of percentile based prediction methods and the Weibull distribution in describing the diameter distribution of heterogeneous Scots pine stands. Forest Ecology and Management 133(3): 263-274.
http://dx.doi.org/10.1016/S0378-1127(99)00239-X

Maltamo M., Eerikäinen K., Pitkänen J., Hyyppä J., 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.
http://dx.doi.org/10.1016/j.rse.2004.01.006

Maltamo M., Malinen J., Packalén P., Suvanto A., Kangas J. (2006a). Nonparametric estimation of stem volume using airborne laser scanning, aerial photography, and stand

-register data. Canadian Journal of Forest Research 36: 426–436.
http://dx.doi.org/10.1139/x05-246

Maltamo M., Eerikäinen K., Packalén P., Hyyppä J. (2006b). Estimation of stem volume using laser scanning-based canopy height metrics. Forestry 79 (2): 217-229.
http://dx.doi.org/10.1093/forestry/cpl007

Maltamo M., Gobakken T. (2014). Predicting tree diameter distributions. In Maltamo, M., Næsset, E., Vauhkonen, J. (Eds.), Forestry applications of airborne laser scanning, concepts and case studies (pp. 177-192). Springer.
http://dx.doi.org/10.1007/978-94-017-8663-8_9

Mehtätalo L. (2006). Eliminating the effect of overlapping crowns from aerial inventory estimates. Canadian Journal of Forest Research 36 (7): 1649-1660.
http://dx.doi.org/10.1139/x06-066

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

Næsset E. (2002). Predicting forest stand characteristics with airborne scanning laser using a practical two-stage procedure and field data. Remote Sensing of Environment 80 (1): 88-99.
http://dx.doi.org/10.1016/S0034-4257(01)00290-5

Næsset E. (2004). Practical large-scale forest stand inventory using a small airborne scanning laser. Scandinavian Journal of Forest Research 19 (2): 164-179.
http://dx.doi.org/10.1080/02827580310019257

Olofsson K., Lindberg E., Holmgren J. (2008). A method for linking field-surveyed and aerial-detected single trees using cross correlation of position images and the optimization of weighted tree list graphs. Proc. SilviLaser 2008, 8th International Conference on LiDAR Applications in Forest Assessment and Inventory, Edinburgh, UK, 17-19 September, pp. 95-104.

Packalén P., Maltamo M. (2007). The k-MSN method for the prediction of species-specific stand attributes using airborne laser scanning and aerial photographs. Remote Sensing of Environment 109 (3): 328-341.
http://dx.doi.org/10.1016/j.rse.2007.01.005

Packalén P., Maltamo M. (2008). The estimation of species-specific diameter distributions using airborne laser scanning and aerial photographs. Canadian Journal of Forest Research 38 (7): 1750-1760.
http://dx.doi.org/10.1139/X08-037

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

Pitkänen J. (2001). Individual tree detection in digital aerial images by combining locally adaptive binarization and local maxima methods. Canadian Journal of Forest Research 31: 832-844.
http://dx.doi.org/10.1139/x01-013

Pitkänen J., Maltamo M., Hyyppä J., Yu X. (2004). Adaptive methods for individual tree detection on airborne laser based canopy height model. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 36 (Part 8/W2): 187-191.

Pollock R.J. (1994). A model-based approach to automatically locating individual tree crowns in high-resolution images of forest canopies. Presentation at the first International Airborne Remote Sensing Conference and Exhibition, Strasbourg, France, 11-15 September 1994. 13 p.

R. (2014). The R project for statistic computing, http://www.r-project.org . (Accessed 10 Oct., 2014)

Reynolds M.R., Burk T.E., Huang W.C. (1988). Goodness-of-fit tests and model selection procedures for diameter distribution models. Forest Science 34 (2): 373-399.

Shiver B.D. (1988). Sample sizes and estimation methods for the Weibull distribution for unthinned Slash pine plantation diameter distributions. Forest Science. 34 (3): 809-814.

Siipilehto J. (2011). Methods and applications for improving parameter prediction models for stand structures in Finland. Dissertations Forestales 124.

Siitonen M. (1993). Experiences in the use of forest management planning models. Tiivistelmä: Kokemuksia mallien käytöstä metsätalouden suunnittelussa. Silva Fennica 27(2): 167-178.
http://dx.doi.org/10.14214/sf.a15670

Tham A. (1988). Structure of mixed Picea abies (L.) Karst. And Betula pendula Roth. And Betula Pubenscens Ehrh. stands in South and Middle Sweden. Scandinavian Journal of Forest Research 3: 355-370.
http://dx.doi.org/10.1080/02827588809382523

Tibshirani R. (1996). Regression shrinkage and selection via the lasso. Journal of the Royal Statistical Society Series B 58 (1): 267-288.

Vauhkonen J., Korpela I., Maltamo M., 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.
http://dx.doi.org/10.1016/j.rse.2010.01.016

Vauhkonen J., Packalén P., Malinen J., Pitkänen J., Maltamo M. (2013). Airborne laser scanning-based decision support for wood procurement planning. Scandinavian Journal of Forest Research.

http://dx.doi.org/10.1080/02827581.2013.813063

Wang Y., Weinacker H., Koch B. (2008). A Lidar point cloud based procedure for vertical canopy structure analysis and 3D single tree modeling in forest. Sensors 8 (6): 3938-3951. http://dx.doi.org/10.3390/s8063938

Yao W., Krzystek P., Heurich M. (2012). Tree species classification and estimation of stem volume and DBH based on single tree extraction by exploiting airborne full-waveform LiDAR data. Remote Sensing of Environment 123: 368-380. http://dx.doi.org/10.1016/j.rse.2012.03.027

Zasada M., Cieszewski C.J. (2005). A finite mixture distribution approach for characterizing tree diameter distributions by natural social class in pure even-aged Scots pine stands in Poland. Forest Ecology and Management 204: 145-158. http://dx.doi.org/10.1016/j.foreco.2003.12.023

Zhang L.J., Gove J.H., Liu C., Leak W.B. (2001). A finite mixture of two Weibull distributions for modeling the diameter distributions of rotated-sigmoid, uneven-aged stands. Canadian Journal of Forest Research 31: 1654-1659. http://dx.doi.org/10.1139/x01-086