Dissertationes Forestales 126

Estimating tree size distributions and timber assortment recoveries for wood procurement planning using airborne laser scanning

Jussi Peuhkurinen
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 auditorium BOR100 of the University of Eastern Finland, Yliopistokatu 7, Joensuu on 26th August 2011 at 12 o'clock noon.

Title of dissertation: Estimating tree size distributions and timber assortment recoveries for wood procurement planning using airborne laser scanning

Author: Jussi Peuhkurinen

Dissertationes Forestales 126

Thesis Supervisor:
Prof. Matti Maltamo
School of Forest Sciences, University of Eastern Finland, Joensuu, Finland

Pre-examiners:
Prof. Erik Næsset
Department of Ecology And Natural Resource Management, Norwegian University of Life Sciences, Ås, Norway
Prof. Jori Uusitalo
Finnish Forest Research Institute, Parkano Research Station, Finland

Opponent:
Dr. Johannes Breidenbach
The Norwegian Forest and Landscape Institute, Ås, Norway

ISSN 1795-7389
ISBN 978-951-651-342-6 (PDF)

(2011)

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

Peuhkurinen, J. 2011. Estimating tree size distributions and timber assortment recoveries for wood procurement planning using airborne laser scanning. Dissertationes Forestales 126. 43 p. Available at: http://www.metla.fi/dissertationes/df126.htm

ABSTRACT

ALS-based inventory methods are replacing traditional field inventories in the production of stand-level data for operative and management purposes. Despite the advances made in the species-specific inventories of mean stand variables, ALS has not succeeded in providing accurate growing stock descriptions for operative wood procurement planning, for example, containing information on tree quality, tree size distribution and the distribution of the logs in timber assortment classes. The aim of this thesis is to evaluate and develop ALS-based methods of predicting tree size distributions and timber assortment recoveries.

The experimental work was carried out in two inventory areas both located in Eastern Finland: Matalansalo, representing a typical managed boreal forest, and Koli, located in the Koli National Park. The remote sensing material for Matalansalo consisted of low and high pulse density ALS data and digital aerial images, and the material for the Koli area of high density ALS data only.

The investigated estimation methods were individual tree delineation (ITD) and area-based statistical approach (ABSA), which were also compared within the same test areas. The performance of ITD in estimating tree size distributions and theoretical timber assortment classes was found to be better than that of the compared methods (ABSA and field assessments) in cases where individual trees could be discerned from the ALS data. In the aggregate, the different ALS methods were comparable when estimating volume and basal area, but ITD tended to produce a bias in saw log volumes and tree size distributions because of the errors in tree delineation. It was stated that the errors in both of the methods, ITD and ABSA, were in correlation with the tree size distribution and the spatial distribution of tree locations.

The estimation of theoretical and actual saw log recoveries was investigated using two area-based methods. The results of the linear regression indicated that it is possible to obtain accurate saw log recoveries using an area-based ALS method. The second method employed k-nearest neighbour imputation and harvester-collected stem data bank. The method produced species-specific saw log recoveries although the estimation accuracies were not as good as expected. The method could be improved by using a more representative stem data bank and additional search variables.

The harvester data from final cuttings was found to be suitable material for validating the diameter distributions and theoretical saw log recoveries estimated from ALS data, although there were challenges considering the delineation of the stand, tree positioning accuracy and different bucking preferences. The use of stem data bank as an auxiliary data source was more challenging because the stem data bank did not include reliable information on stand delineation and bucking parameters.

Keywords: Forest inventory; Linear regression; Nearest neighbour; Remote sensing

ACKNOWLEDGEMENTS

I was first introduced to remote sensing of forests in 2001 when I visited Prof. Timo Pukkala to seek for a topic for a seminar work on forest planning and economics. Professor Pukkala asked me in his personal manner: “You do not care too much what the topic is, right?”. I had nothing particular in my mind so I agreed. He then sent me to Mr. Lauri Vesa, who was working as a lecturer in Faculty of Forestry that time. Lauri had some old digitized aerial photographs and field data from inventory by compartments. The idea was to estimate stand level forest characteristics using remote sensing and statistical modeling. From a practical point of view, the results were of “standard remote sensing quality”. That means something between totally useless and of slight academic interest. Although, for me, this exercise was of great importance considering my further career. Soon after this first peek into the world of numerical interpretation of remote sensing material, I found myself working with IKONOS satellite images in my Master's thesis. When Professor Matti Maltamo asked me a couple of years later if I was interested in post-graduate studies and airborne laser scanning I had no other choice than answer him yes. However, at that time I was very sceptical about the usefulness of remote sensing in operative forestry. I have changed my perception of many things from those days, including remote sensing of forests.

Several individuals and organizations have helped me both on my journey towards finishing this thesis and in developing my perceptions on remote sensing of forests among many other things. First of all, I would like to thank my supervisor Prof. Matti Maltamo who has always answered my inquiries and, at the same time, given me the academic freedom to make my own interpretations and choices. I like to thank my co-authors, Prof. Jyrki Kangas, Dr. Jukka Malinen, Dr. Lauri Mehtätalo, Dr. Petteri Packalén, Dr. Lauri Korhonen, Mr. Juho Pitkänen and Mr. Aki Suvanto. Without their expertise this thesis would have never been finished in its current form. In addition, I would like to thank the pre-examiners of this thesis, Prof. Erik Næsset and Prof. Jori Uusitalo for their invaluable comments.

I have had the pleasure to do my work in the best research group I know. Many thanks to my former and current colleagues in the former Faculty of Forest Sciences, currently known as School of Forest Sciences. I hope that the good spirit transfers also to new members of the community. Since Autumn 2007 I have been a member of another work community, too. I would like to thank my colleagues at Arbonaut and especially Dr. Tuomo Kauranne, maybe not for helping with my thesis, but for giving a possibility to put theory in practice.

My project was funded by the Graduate School of Forest Sciences, Finnish Cultural Foundation and Metsämiesten Säätiö. I am very grateful to these financial supporters.

My warmest thanks to my friends and to my family, especially my parents, Ritva and Teuvo, sister Kaisa and brother Lassi and their families. I am very conscious about the fact that the loving and open-minded environment from my very childhood till this day cannot be taken for granted. It has given me strength and confidence without which many good things in my life would not have been possible. Last but not least, I would like to thank Hanna, my dearest partner.

Helsinki, July 2011

Jussi

LIST OF ORIGINAL ARTICLES

The present thesis is based on the following papers, which are referred to in the text by their Roman numerals. Articles are reproduced with the kind permission from the publishers.

I Peuhkurinen, J., Maltamo, M., Malinen, J., Pitkänen, J. & Packalén. P. 2007. Preharvest measurement of marked stands using airborne laser scanning. Forest Science. 53(6): 653–661.
http://www.ingentaconnect.com/content/saf/fs/2007/00000053/00000006/art00005.

II Korhonen, L., Peuhkurinen, J., Malinen, J., Suvanto, A., Maltamo, M., Packalén. P. & Kangas, J. 2008. The use of airborne laser scanning to estimate sawlog volumes. Forestry 81(4): 499-510.
doi:10.1093/forestry/cpn018.

III Peuhkurinen, J., Maltamo, M. & Malinen, J. 2008. Estimating species-specific diameter distributions and saw log recoveries of boreal forests from airborne laser scanning data and aerial photographs: a distribution-based approach. Silva Fennica 42(4): 625–641.
http://www.metla.fi/silvafennica/full/sf42/sf424625.pdf.

IV Peuhkurinen, J., Mehtätalo, L. & Maltamo, M. 2011. Comparing individual tree detection and the area-based statistical approach for the retrieval of forest stand characteristics using airborne laser scanning in Scots pine stands. Canadian Journal of Forest Research 41(3): 583-598.
doi:10.1139/X10-223.

The present author was responsible for all the calculations and analyses of the results in papers I and III, and for most of the calculations and analysis in paper IV except for retrieval of the diameter-height distributions using parameter recovery. In paper II he was responsible for the calculations and analyses related to the harvester-collected data. He was also the principal author of papers I, III and IV.

TABLE OF CONTENTS

LIST OF ABBREVIATIONS

3D	3-dimensional
ABSA	Area-based statistical approach
ALS	Airborne laser scanning
Aseg	Area of the segment
AvgDSeg	Average of MaxDSeg and PerDSeg
CHM	Canopy height model
D6	Diameter at height 6 metres, cm
DBH	Diameter at breast height
Dgm	Diameter (DBH) of the basal area median tree, cm
DTM	Digital terrain model
GPS	Global Positioning System
G	Basal area, m^2/ha
HBF	Height-based filtering
H-D	Height-diameter
Hdom	Dominant height, m
Hgm	Height of the basal area median tree, m
ITD	Individual tree delineation
LiDAR	Light Detection And Ranging
MaxDSeg	Maximum diameter of the segment
MaxDz	Maximum dz value of the segment
MSN	Most similar neighbour
N	Number of trees per hectare
NN	Nearest neighbour
PerDSeg	Diameter perpendicular to the maximum diameter
RMSE	Root mean squared error
SOM	Self-organizing maps
StanForD	Standard for Forestry Data and Communication
V	Volume, m^3/ha
XySeg	XY-coordinates of MaxDz

1 INTRODUCTION

1.1 Airborne laser scanning-based forest inventory methods

Forest inventory methods in the Nordic countries have undergone remarkable changes during the last 10-20 years, chiefly due to a technical breakthrough in remote sensing, namely in the development of airborne laser scanning (ALS). The earliest reported tests concerning the use of airborne LiDAR (Light Detection And Ranging) for forest profiling were conducted in the former Soviet Union in the 1970s (Solodukhin et al. 1977), and after that several studies were carried out mainly in North America (e.g., Nelson et al. 1984, Magnussen et al. 1999) and in Scandinavia (e.g., Nilsson 1996, Næsset 1997, Hyyppä and Inkinen 1999) to explore the potential of ALS for forest inventory purposes. By the beginning of the 21st century ALS-based forest inventory methods were ready to be adopted in practice (Næsset 2002).

The ALS data considered in this thesis represent what is known as discrete return small footprint airborne LiDAR scanning data. Such data are typically captured from an airborne vehicle, either a helicopter or a fixed-wing aircraft flying at an altitude of about 500 to 2000 metres above ground level. The sensors are able to record multiple reflections from each emitted pulse, even though typically only the first and last reflections are used. The size of the footprint varies from about 20 to 60 centimetres on the ground and the nominal pulse density is from 0.5 to just under 10 pulses per square metre. The above ALS data specifications are commonly used in forest applications. A more in-depth account of ALS theory is given by Wehr and Lohr (1999), for example.

ALS data can be understood as constituting a 3-dimensional (3D) point description of the target object. In forestry this point cloud is regarded as a description of the vegetation structure or canopy structure which allows the extraction of various variables that are correlated with tree or forest stand attributes. The estimation methods may be divided into two approaches depending on the unit to be estimated. In the individual tree delineation (ITD) approach the aim is to discern individual trees (Hyyppä and Inkinen 1999, Leckie et al. 2003, Popescu et al. 2003, Maltamo et al. 2004a, Holmgren and Persson 2004), or groups of trees (Breidenbach et al. 2010) based on 3D ALS data. Once the trees have been delineated, tree-level attributes are extracted, or modelled, from the ALS observations regarding each tree. In the area-based statistical approach (later ABSA), also referred to as the canopy height distribution method (Means et al. 2002, Næsset 2002, Packalén and Maltamo 2007, Junttila et al. 2008), the unit is an area of a certain fixed size, typically a few hundred square metres. The mean forest stand variables are estimated for this unit using statistical correlations between explanatory variables derived from the ALS data and forest stand variables. Both approaches can use optical remote sensing material as additional information for species-related attributes (Leckie et al. 2003, Packalén and Maltamo 2007), and both can be used to produce stand-level diameter-height distributions (Gobakken and Næsset 2004, Packalen and Maltamo 2008).

ABSA methods have been shown to provide reliable, unbiased estimates of growing stock. Area-based ALS methods as used in Finland, for example, have achieved stand-level estimation accuracies for species-specific attributes that are comparable with those of conventional field inventories by compartments (Haara and Korhonen 2004), and ALS-based estimates for total characteristics are even more accurate than inventory results based on visual assessment and subjective field measurements; the approach that is now being superseded by ALS-based inventories.

The main challenge in ITD methods has been difficulties in discerning individual trees. Overlapping tree crowns make the correct delineation of adjacent trees virtually impossible in dense, heterogeneous stands, which results in biased stand-level estimates (Maltamo et al. 2004a, 2004b, Koch et al. 2006). If tree delineation is successful, however, ITD can provide valuable information about the tree crown (Popescu and Zhao 2008, Vauhkonen 2010), tree quality (Maltamo et al. 2009a) and tree species (Holmgren & Persson 2004, Vauhkonen et al. 2009, Ørka et al. 2009).

1.2 Estimating stand characteristics for wood procurement planning

1.2.1 Inventory information requirements

ALS-based forest inventories are replacing, or have already replaced, field inventories as the primary method for collecting information about the growing stock in many countries, including Finland, but operational planning still requires additional measurements for checking that the forest management decisions based on the ALS inventory are correct and for collecting information on variables that are not directly evaluated in an ALS inventory, as these variables may affect resource allocation by the wood procurement company and the pricing of the timber. It will be important from the company's point of view, for example, to have precise information on the location of potential stands for final felling, their tree size distributions and tree quality, in order to know where and when to harvest in order to optimize the flow of raw material. On the other hand, forest owners are able to make better decisions on timber sales if they have enough information on the material available for sale.

The information on stand characteristics required for the planning of wood procurement includes the volumes and proportions of the potential timber assortments. The main timber assortments are saw logs and pulpwood, both specified according to size and quality criteria. Species information is also important, since the size and quality criteria are usually species-dependent. Saw logs can be further divided into special assortments, which may differ significantly in their economic value. In addition to saw logs and pulpwood, timber assortments may also include energy wood and some low-value saw timber, e.g. small-diameter saw logs. The value of the wood as a raw material is mainly dependent on the recovery of the most valuable timber assortments, i.e. saw logs. In addition to estimates of total volume and the proportions of the main timber assortments, wood procurement planning requires information on the distribution of saw logs into special assortments. This can be estimated from species-specific tree size distributions, or estimated tree lists. Quality reductions due to branchiness, crookedness, etc. and other defects may have a significant effect on the actual recovery of the timber assortments.

1.2.2 Estimation methods and the content of stand-level forest inventories

The information collected in a stand-level forest inventory in Finland contains details of the growing stock and forest site type together with additional information for use in forest planning. The growing stock variables consist of the basal area or stem number, age and the diameter and height of the basal area median tree for every tree stratum, which is a combination of tree species and tree layer (Solmu maastotyöopas… 2003). The stand volume and timber assortment data are derived from these mean growing stock variables in two steps (Kangas & Maltamo 2002). First, the basal area diameter distribution is estimated based on the mean variables that have been assessed. Second, the volume and timber assortments are predicted using tree-level height and volume models or taper curve functions based on diameters sampled from the predicted diameter distribution. Several theoretical constructs, e.g. Weibull distribution models (Kilkki et al. 1989) and percentile-based diameter distribution models (Kangas & Maltamo 2000a) can be employed for describing the tree size distribution. Maltamo et al. (2000) and Kangas & Maltamo (2000b) have demonstrated that percentile-based distributions describe the diameter distributions slightly more accurately than does the Weibull distribution, especially if the forest structure is heterogeneous. Theoretical saw log volumes were also estimated with better accuracy using percentile-based diameter distributions.

Timber assortments can also be estimated without any predicted diameter distribution. Nyyssönen & Ojansuu (1982) and Päivinen (1983) have formulated saw log percentage models, which can be used for predicting saw log percentages from stand attributes, but these models do not take into account the internal structure of the forest stand and they cannot use a description of the tree size distribution, whereas tree-level saw log reduction models (e.g. Mehtätalo 2002) can be used with estimated diameter distributions to predict actual saw log recovery (i.e. a saw log recovery figure which includes reductions due to bucking constraints (allowable length and diameter combinations) and external technical defects).

1.2.3 Estimation methods developed for wood procurement purposes

The information provided by a stand-level forest inventory alone is not sufficient for optimizing wood procurement. Such information is not usually extensively available, is commonly out of date because of deficient updating or is lacking in reliability for some other reason. The stand delineation in a stand-level forest inventory may also differ considerably from the operative stand delineation, and therefore the information cannot be connected with operative planning units very precisely. Moreover, saw log percentage models and saw log reduction models cannot take into account changes in the bucking constraints. Thus several methods have been developed for providing information for the purposes of wood procurement planning.

The necessary information can be collected using field inventory methods such as the PMP system, which was based on measuring the diameter at breast height (DBH) of every tree in the stand in question (PMP-ohje 1982). In addition, sample trees with more extensive measurements (diameter at 6 metre, tree height, quality information etc.) were performed to complete the survey (PMP-ohje 1982). Extensive field measurements can provide accurate data on the measured stands, but they are laborious and expensive, and it is practically impossible to survey all potential stands. Some sample-based methods have also been proposed in order to reduce the costs of field inventories (e.g. Uusitalo 1995 and Hansson 1999), but these are not widely used in practice since they have been found to be too expensive and laborious relative to the accuracy achieved.

The use of computational methods is one option for obtaining information for wood procurement planning without laborious field measurements. Tommola et al. (1999) studied non-parametric k-nearest neighbour (k-NN) techniques for estimating the characteristics of stands with a potential for harvesting, using stems assessed by means of log measurement instruments installed at sawmills as reference data combined with stand data from an information system (e.g., geographical location, stand area and density, mean tree size etc.) and additional variables that included some accurately measured stem size variables as independent variables. Malinen et al. (2001) and Malinen (2003) examined the k-NN method as a means of estimating stand characteristics for the purposes of wood procurement planning using a harvester-collected stem data bank for calculating the independent variables and extracting the other variables used in the modelling (temperature sum, location, stand age, forest site type etc.) from a field survey. Neural computing has also been examined for its ability to predict stand characteristics in the context of wood procurement planning. Neural network methods such as self-organizing maps (SOM) (Kohonen 1995) have been successfully applied to numerous estimation problems, but the studies of Räsänen et al. (2000) and Lappalainen (2005) indicated that the number of measured stands compared with the number of variables used for describing the stand properties is inadequate for successful SOM training, and maintained that these methods are incapable of competing with NN methods for the estimation of stand characteristics. Computational methods based on NN techniques were found in these studies to be efficient with respect to the accuracies they achieved and the resources they used, they require at least some kind of a-priori information about the potential stands. Common stand-level forest inventory information may be suitable for this purpose if it is available and up-to-date, and if the stand delineation in the stand-level forest inventory does not differ significantly from the operative stand delineation. In practice this means that the use of computational methods may still require a separate stand level forest inventory, although not such a thorough one.

1.3 Diameter distribution and timber assortment estimation with ALS

Diameter distributions can be estimated in several ways in the context of an ALS inventory. In the case of ITD the principle is simple: the stand-level diameter distribution is a composite of the diameters of the individually detected trees in the stand. ITD methods do not directly produce the diameters for the detected trees, however, but rather they are predicted using tree-level allometric models (Hyyppä & Inkinen 1999, Kalliovirta & Tokola 2005), or tree or stand-level variables extracted from ALS data (Vauhkonen et al. 2010, Maltamo et al. 2009a). Furthermore, discerning individual trees is a challenging task and may result in a biased distribution (Maltamo et al. 2004a, Maltamo et al. 2004b, Koch et al. 2006). The effect of bias due to omission error can be diminished by using a theoretical distribution to estimate the diameter distribution of suppressed trees, as was suggested by Maltamo et al. (2004b). As a result, the final distribution is a combination of individually detected trees in the dominant layer and a modelled distribution for the smaller trees in the sub-dominant layer. Lindberg et al. (2010) used a target distribution estimated by an area-based method to calibrate their ITD distributions, and found that this reduced the estimation error relative to the not-calibrated distribution. Other ways of overcoming this bias are to use statistical correction methods as proposed by Flewelling (2008 and 2009) or to use tree clusters instead of individual trees (Breidenbach et al. 2010).

The area-based approach offers possibilities for direct imputation of the empirical diameter distributions of field reference sample plots and several options for using theoretical distributions or percentile-based distributions. One option is to predict mean stand variables using ALS and then estimate the theoretical diameter distribution by predicting parameters for the assumed theoretical distribution using the predicted mean stand variables and existing models for predicting the parameters. This method was tested by Maltamo et al. (2006a) and Holopainen et al. (2010), applying a Weibull function and existing parameter prediction models. The method is analogous to that used in field-based inventories by compartments and differs only in how the mean stand variables are estimated. The process can also be executed using the grid approach, by predicting the parameters of the theoretical distribution using the ALS estimates for all the estimation units (grid cells) separately and treating the stand-level distribution as a combination of the grid cell-level distributions of the stand (Packalen and Maltamo 2008).

Gobakken and Næsset (2004) used regression analysis to relate the parameters of a two-parameter Weibull distribution and Weibull percentiles to variables derived from the ALS data. They modelled both diameter and basal area distributions, and the comparison showed that the basal area distribution produced more accurate volume estimates than diameter distributions scaled with reference to the number of stems. Furthermore, Gobakken & Næsset (2005) compared two approaches for estimating basal area distributions using ALS data, a Weibull distribution and a percentile-based distribution, and found no significant differences in the volume predictions even when comparing plots with small and large diameter variability. Maltamo et al. (2006a) compared two laser scanning-based methods and a field inventory-based method for estimating basal area diameter distributions using laser scanning-based height metrics. The field inventory method and the first ALS method were based on predicting parameters for a Weibull distribution using either field-assessed mean stand characteristics or stand characteristics estimated using ALS data. Another ALS method was based on modelling percentiles of the basal area diameter distribution. Bollandsås & Næsset (2007) studied the estimation of percentile-based basal area distributions in an uneven-sized Norway spruce stand using ALS data.

Maltamo et al. (2007) demonstrated that it is possible to derive diameter distributions directly in the form of stem frequencies instead of from basal area distributions if ALS data are used, without any loss of accuracy in the volume estimates. They also tested a calibration estimation for adjusting the predicted diameter distribution by reference to the ALS-based stand volume estimate.

Mehtätalo et al. (2007) proposed a parameter recovery approach obtaining the theoretical diameter distribution. The parameters of an assumed diameter distribution and height-diameter curve were recovered using the mathematical relationship between the parameters of the theoretical distribution and the stand characteristics as estimated using ALS. The method was found useful in laser scanning approaches where accurate predictions of forest stand characteristics are obtained. One benefit of the parameter recovery approach is that no parameter estimation is needed, since the approach recovers the parameter values without the use of tree-level data.

Breidenbach et al. (2008) presented a one-step procedure for deriving the parameters of a Weibull function to describe diameter distributions. Plotwise height metrics derived from airborne laser scanner data was used as auxiliary variables in the estimation. The method does not require any extensive tree-level observations per sample plot. Thomas et al. (2008) examined the ability to predict Weibull parameters from ALS data using multiple regression analysis, and found that it was possible to characterize young heterogenous stands using unimodal Weibull distributions if the plots were stratified into structurally similar groups based on stem density and canopy top height, whereas the two-parameter Weibull models did not correlate well with the ALS metrics without stratification. They suggested that a better alternative for irregular distributions would be to predict a two-mode Weibull mixture model from the ALS data.

All the above methods for estimating diameter distributions employed with ABSA are based on constructing either theoretical distribution or percentile-based distribution models. Packalen and Maltamo (2008) used k-NN imputed trees to form a diameter distribution, so that the diameter distribution of the target unit was composed of the trees of the k nearest observations. The imputed trees were weighted using the inverse of the most similar neighbour (MSN) distance (see Moeur and Stage 1995). The diameter distribution was then a set of trees, or tree list, classified and ordered by diameter with class frequencies calculated using weights. The method requires the reference plot data to include tree-level data, i.e. the reference data must include a tree list for every plot. Maltamo et al. (2009b) examined the same method using stratified data and also considered the effect of a reduced number of reference plots on estimation accuracies.

If the diameter distribution and diameter-height curve are estimated, the theoretical volumes of timber assortments can be calculated using existing taper curve functions and predefined diameter and height dimension limitations for the different timber assortments based on sample trees from the diameter distribution. Bucking simulations can be used to estimate the length-diameter distributions of timber assortments with different bucking parameters in relation to demand and price matrices (Malinen 2003). If a tree list is available (as in the case of ITD or k-NN-imputed trees) the volume calculations and bucking simulation can be carried out using the trees on the tree list directly.

Timber assortments can also be predicted using direct estimation models. Rooker Jensen et al. (2006) formulated area-based ALS regression models for the volumes of small saw logs, large saw logs and poles, while Maltamo et al. (2009a) estimated the saw log proportion using ITD and k-MSN imputation. The use of direct estimation allows the estimation of actual volumes for timber assortments if the reference data include actual measured volumes, i.e. with the minimum dimensions and quality requirements of timber assortments taken into account in the reference measurements. In Maltamo et al. (2009a), for example, the part of the stem that qualified for saw logs was determined on the basis of field measurements.

Since estimated diameter distributions or tree lists do not necessarily contain information about the technical quality of the timber, estimated saw log volumes do not include the effect of quality reductions and are therefore only theoretical. Average actual saw log recovery can be estimated using saw log reduction models (Mehtätalo 2002), but these models cannot be used with bucking simulations to produce actual length-diameter distributions for timber assortments, since the models do not estimate the position of the defect affecting the quality of the stem.

No previous studies are extensively available that include comprehensive tree quality information in the area-based ALS estimation process. Crown height is possibly the most intensively studied quality-related variable (excluding stem dimensions) in an ALS forest inventory context. ABSA-based estimation of crown height has been studied by Næsset & Økland (2002) Maltamo et al. (2006b), Dean et al. (2008) and Maltamo et al. (2010), whereas vertical crown dimension studies based on individually detected trees are numerous (e.g. Næsset & Økland 2002, Roberts et al. 2005, Maltamo et al. 2006b, Popescu & Zhao 2008, Vauhkonen 2010). Maltamo et al. (2009a) demonstrated that it is also possible to estimate other tree quality attributes such as the height of the lowest dead branch for individually detected trees using ITD. The estimation of some tree crown variables such as crown height is also possible without field calibration (Maltamo et al. 2010). Perhaps the most ambitious research has been the study by Bollandsås et al (2010a). They modeled actual saw log volumes, saw log proportions, mean decrease in millimetres between diameter at 6 metres (D6) and DBH, mean ratio of height and DBH and mean crown height by using ALS data and harvester measured field plots in mature stands.

1.4 Fusion of harvester and remote sensing data

Harvesters equipped with an on-board merchandising or bucking computer are able to collect and record information about the trees that have been felled and processed. In Europe the Standard for Forestry Data and Communication (StanForD) maintained by Skogforsk has become a de-facto standard for the management of merchandising computers and forest communications. This is both a data standard and a file structure standard, and it also includes a Kermit-based communications protocol for connecting a data recorder to the merchandising computer on the harvester (What is StanForD… 2010). StanForD data files are divided into numerous types according to the application concerned, each data file for a specified application being defined by a file name extension. It is the stm files that are the most interesting for forest inventory purposes, as these include individual tree measurement information, i.e. the measured lengths and diameters of the trees that have been felled and processed (Standard for… 2007). By combining stm file data for a certain geographically defined area it is possible to extract information about the tree size distributions, total volumes, tree species proportions etc. of that area. In the case of final fellings, the information provided can be used as a wall-to-wall timber inventory of that area carried out by measuring every single tree larger than a given minimum size. There are certain features of the measurement techniques, however, that make interpretation of the recorded values a somewhat challenging procedure. Firstly, the data apply to trees which fulfil certain minimum dimensions related to the timber assortment classifications used, i.e. trees that do not fulfil certain minimum dimensions or quality attributes are not included. Secondly, there is no explicit information on stump height, so that the first diameter measurement height can differ between machine-software combinations, and the last diameter and height measurement applies to the last cutting point (i.e. information on the top of the tree is missing) (Räsänen et al. 2000). Besides the challenges mentioned above, one element which increases the uncertainty is the spatial delineation of the stand, or the absence of this, for in their absence the area-related variables cannot be used without carrying out a separate stand delineation operation. If the stm file includes recorded coordinates, the stand can be delineated based on this information, but because of the features of the positioning (which are discussed later), separate stand delineation is preferred. The measurement accuracy of the tree dimensions is not in itself the main concern, for in the case of new harvesters about 68 % of all diameter measurements are within error limits of +/- 4 mm (Möller and Arlinger 2007). According to the same study, 84% of the length measurements are within error limits of +/-2 cm and volume is measured with less than a 1.5% error when manual control measurements are used as a reference.

Based on previous studies, fusion of the harvester data remote sensing data can be divided into two applications. The first is to use harvester data for training purposes instead of manually collected field plot data (see Rasinmäki and Melkas 2005, Melkas et al. 2009, Bollandsås et al. 2010a) or as tree-level reference data in the ITD approach (Larsson 2009), although this is problematic in a laser inventory because of the positioning method. The positioning of the field plots used in modelling is an essential part of any area-based inventory method, and it has been demonstrated that a positioning error of up to 2-5 metres in xy location with a plot area of 200-400 square metres will not significantly affect the ALS estimates for mean height, basal area and total volume (Gobakken and Næsset 2009), but the effect of the plot positioning error is dependent on the plot size and forest strata (Gobakken & Næsset 2009) and, presumably, also on the structure of the forest mosaic. It is possible to record x, y and z coordinates on an stm file for every felled tree, but these will be taken from the GPS antenna installed in the cabin of the harvester machine and will not be the coordinates of the harvested tree but those of the machine at the time of felling the tree. Moreover, harvesting operations are rarely optimized for GPS positioning and it is difficult to validate the quality of the recorded coordinates. According to Rasinmäki and Melkas (2005) the location of the harvester is recorded at an accuracy of about 3.2 metres, and without additional information on the harvesting techniques used in a particular stand we can add the harvester's boom length (up to 10 metres) to the positioning error. Bollandsås et al. (2010a) suggested that the limited positioning accuracy of harvester data could be compensated for by using a larger plot size. Furthermore, the correct positioning of reference trees is an even more crucial issue in the ITD approach, since the trees measured in the field must be linked to the trees detected individually in the ALS data. To make this possible, the positions of individual trees should be determined to an accuracy of about 1 metre. At the current state of harvester positioning technology this means that the coordinates of the reference trees would have to be collected with separate GPS measurements before harvesting, the trees numbered and the operator of the harvester machine instructed to record the tree number when processing the tree (Larsson 2009).

The second application is to use harvester data for validation purposes (Maltamo et al. 2010, Holopainen et al. 2010, Bollandsås et al 2010b). Harvester data for a clear-cut stand will provide stand-level wall-to-wall control measurements in an economically feasible way. However, if the harvester data are used for plot or tree-level validation the same positioning challenges apply as in the previous paragraph.

1.5 Objectives

The overall aim of this thesis was to study the use of ALS-based forest inventory methods for retrieving the information needed for wood procurement planning. The information to be considered here comprised diameter distributions and saw log recoveries in addition to the more commonly produced inventory results, i.e. mean stand characteristics. A further aim was to study and develop methods for using harvester-collected data as reference and auxiliary data in ALS-based forest inventory projects. The specific objectives of the thesis were:

1. To investigate the possibility of using an ALS-based individual tree delineation procedure and area based statistical approach to produce pre-harvest measurement information on stands marked for final felling and to compare these with alternative pre-harvest measuring methods (paper I).

2. To investigate the possibility of using ALS-based methods to estimate stand-level diameter distributions (papers I, III and IV).

3. To develop and test ALS-based methods for estimating theoretical and factual saw log recoveries (papers II and III).

4. To examine the possibilities of using harvester-collected data as a validation data (papers I, II and III) and as an auxiliary data source (paper III) in an ALS-based forest inventory.

5. To investigate the estimation accuracies of individual tree delineation and an area-based statistical approach within the same test area (papers I and IV).

2 MATERIALS

The two data sets used in this work applied to the Matalansalo and Koli inventory areas. The Matalansalo data set was used for papers I, II and III and the Koli data set only for paper IV. Both refer to boreal forests, although these are markedly different in many respects with regard to their forest characteristics, field measurements and remote sensing material, so that they will be described separately in the following sections.

2.1 Matalansalo data set

2.1.1 Inventory area and field data

The Matalansalo inventory area is located in the municipality of Varkaus in Eastern Finland (about 28° 29' E, 62° 18' N). The total area is approximately 1,200 hectares and it is dominated by coniferous species (Norway spruce (*Picea abies* (L.) Karst) and Scots pine (*Pinus sylvestris* L.). Deciduous species are found mainly as a minority in mixed species stands. The Matalansalo area can be considered representative of a typical managed boreal forest area in Finland.

The field data were collected in summer 2004. A total of 474 circular plots of radius 9 metres were systematically distributed over the 67 forest stands and positioned with a Global Positioning System (GPS) using differential correction. DBH, species, tree storey and tree class (dead, alive) were recorded for every tree with a DBH of at least 5 cm within each plot. In addition, one tree from every tree storey by species per plot was chosen as a height sample tree. The heights of the rest of the trees were then calculated using the species-specific height models of Veltheim (1987), which were calibrated by plots by means of sample tree measurements. The volumes of individual trees were calculated using the species-specific volume functions of Laasasenaho (1982). Finally, mean stand characteristics, i.e. basal area (G, m^2/ha), number of stems (N, n/ha), volume (V, m^3/ha) and diameter (Dgm, cm), and the height (Hgm, m) of the basal area median tree, were calculated for each plot.

2.1.2 Remote sensing material

The airborne laser scanning (ALS) data, acquired on 3 and 4 August 2004 with an Optech ALTM 2033 laser scanner, arose from two flights at altitudes of 1500 and 380 metres above ground level. The second scan covered only part of the area. The field of view was 30 degrees in both cases. This yielded a swath width of approximately 800 metres for the first scan and 200 metres for the second. The first scan resulted in a nominal sampling density of about 0.7 measurements per square metre and a footprint of 45 cm at ground level. These data was used in papers II and III, and for ABSA (method 6) in paper I. To achieve an adequate density of laser pulses for individual tree detection, the second flight was carried out by scanning the same line of flight four times and combining the data sets from both flights. This produced an overall first pulse density of approximately 6.4 per square metre for the final data set in the ITD methods as used in paper I. A digital terrain model (DTM) for both scans was processed from last pulse data of the first scan with TerraScan software (see www.terrasolid.fi). For this the ground points were first separated from the other points using the method explained by Axelsson (2000) and then a raster DTM was created from the classified ground points by calculating their mean values within each one-metre raster cell. Values for raster cells with no data were derived by Delaunay triangulation and the bilinear interpolation method. The ALS data were further processed by subtracting the DTM from the laser pulses to produce a point cloud with x, y and dz coordinates, where dz is the height above the ground. Only the first pulse data were used to obtain the forest characteristics.

In addition to the ALS data, colour-infrared photographs at a scale of 1:30,000 were acquired on 22nd August 2004 with a Leica RC30 camera having a UAGA-F 13158 objective of focal length 163.18 mm and an anti-vignetting filter (AV525 nm). The films were digitized at a resolution of 14 µm, orthorectified using the DTM generated from the ALS data and re-sampled to a pixel size of 50 cm. Since three aerial photographs were required to cover the area, radiometric calibration was performed in order to make the images comparable. The correction was done by the method presented by Tuominen and Pekkarinen (2004), using a Landsat 7 ETM satellite image of the same area taken in June 2002. The radiometric resolution of the final images was 8 bits.

2.1.3 Harvester-collected data

A total of 14 marked stands located in the Matalansalo inventory area were clear-cut after acquisition of the remote sensing material. These stands were delineated using GPS and differential correction. To avoid the effect of trees left in the clear-cut area, retention tree groups of more than 2 trees were also delineated from the stand polygons. The harvester data on the stands consisted of stm files which included the position of the harvester at the time of felling, diameters of the stem as measured by the harvester in 10 cm intervals from the felling point to the last cutting point, the length of the usable part of the trunk, species, bucking parameters and bucking results for every harvested tree according to the StanForD (Standard for… 2007). The denser scan area (second scan) covered two of the stands, whereas all the stands were inside the area of the first scan. The harvester data were collected independently of the circular sample plot data. Eight out of the 14 stands were such that they contained sample plots or portions of sample plots. A total of 45 complete plots were located inside the test stands.

In addition to the harvester-acquired data, the Matalansalo data set also included stem data bank data on 35 mainly spruce-dominated clear-cut stands originally collected for the Finnish Forest Research Institutes' research project "Value formation of timber stand when targeting for alternative end-products in timber harvesting". The stem data bank stands were located outside the scanned area at distances of up to 250 kilometres north-east, south and west of Matalansalo. The stem data bank included tree data extracted from stm files (species and bucking results for every harvested tree). Unlike the marked stands covered by ALS, the stem data bank stands were not delineated on the ground.

2.2 Koli data set

2.2.1 Inventory area and field data

The Koli inventory area is located in the southern part of the Koli National Park in Eastern Finland (about 29° 55' E, 63° 2' N). Until the foundation of the national park in 1991 the area had been subject to normal forestry management practices. Even though the Koli area cannot be considered representative of managed forests in the region, in view of the long rotation age required in the poor soils of the area, the state of the forests there is close to that in managed semi-natural forests in the boreal zone except for some over-dense young or developing stands.

The field data apply to 14 square plots (30 m by 30 m) located randomly over the Scots pine stands of the area. The plots were further divided into 127 square subplots of 10 by 10 metres. The plots were measured in the field during May and June 2006 and their corners were positioned using differential GPS. DBH, D6, height, species, and xy coordinates were recorded for every tree of DBH over 5 cm. In addition to these characteristics the canopy cover percentage was measured using a Cajanus tube. Stem volumes were calculated using the species-specific stem volume models of Laasasenaho (1982), which employed DBH, D6 and height as independent variables. G, N, V, Dgm and Hgm were calculated for each plot and each subplot. In addition, dominant height (Hdom, m), arithmetic mean of DBH, arithmetic mean of height, standard deviation of DBH, standard deviation of height, the coefficients of variation in DBH and height, and the index of dispersion of counts (Cox and Lewis 1966) were calculated for the plots but not for the subplots.

2.2.2 Remote sensing material

The ALS data were collected on 17th June 2005 with an Optech ALTM 3100 laser scanner. The divergence of the laser beam was 0.26 mrad, scanning angle ±11 degrees and mean operation altitude 900 m above ground level. This resulted in a swath width of approximately 350 m, a nominal pulse density of 4 pulse/m^2 and a footprint of size about 23 cm. The equipment is able to collect up to 4 range measurements: the first, last and two intermediate returns, although only the last and first returns were used here. Any original instances of only returns were duplicated to both the first and last categories. The DTM was generated using same methods as in the case of the Matalansalo data set. Both, the first and last pulse data were used for estimating the forest characteristics.

3 METHODS

Two interpretation approaches were employed here for estimating the stand-level timber variables: ITD and ABSA. The ABSA approach was employed in all the papers, whereas ITD was used only in papers I and IV. The use of both approaches within the same test area, with the same test data set and for the same inventory task enabled comparisons to be made between the methods.

3.1 Area-based statistical approach

ABSA is an empirical method for investigating the relations between forest attributes and ALS height distributions. Estimation methods can be divided into parametric regression methods and non-parametric nearest neighbour methods. Parametric linear regression was used in papers II and IV, and the existing ABSA models used as a reference in paper I were also of the linear regression kind. The non-parametric k-NN method was used in paper III.

3.1.1 Extracting independent variables from remote sensing data

The ABSA methods employed in all the papers used independent variables extracted from the canopy height distribution. The theoretical background to the relationship between ALS observations and canopy height distributions generated from data on individual trees has been elaborated by Magnussen et al. (1999) and Mehtätalo and Nyblom (2009). In theory, assuming that the trees are solid objects, the canopy height distribution of the ALS observations and the actual height distribution of the trees are not completely congruent, since i) trees are sampled in ALS observations with a probability proportional to their crown size as observed from above, ii) if crowns overlap at a given point, only the higher one will be observable, and iii) observed canopy height equals true tree height only at the tip of the tree (Mehtätalo and Nyblom 2009). Thus, large trees are usually overrepresented in an ALS canopy height distribution and its maximum canopy height will underestimate the true maximum height of the canopy (Mehtätalo and Nyblom 2009). Furthermore, the trees are not actually solid objects, and therefore the laser pulses penetrate in the canopy before the discrete return is triggered. This will increase the underestimation of the tree heights, again.

In papers I, II, and IV the canopy height distribution was used in calculating the percentiles (H_1, H_5, H_{10}, H_{20}, H_{30}, …, H_{90}, H_{95}, and H_{99}) and the cumulative proportional canopy densities (P_1, P_5, P_{10}, P_{20}, P_{30}, …, P_{90}, P_{95}, and P_{99}) for 1, 5, 10, 20, 30, …, 90, 95, and 99% heights. The set of independent variables also included the proportion of vegetation hits versus ground hits, and linearizing transformations of all the variables, i.e. logarithmic, second power, square root and inverse transformations. The independent variables were calculated at plot level and for the last and first returns separately.

In paper III the canopy height distributions were used as height histograms, in which the frequency was the proportion of laser returns in that dz class. Histograms were derived only for the first returns. False colour aerial image spectral histograms were also used as independent variables in paper III. Again all the independent variables were calculated separately for each plot.

3.1.2 Estimating mean stand variables

The mean stand variables were estimated using either parametric ordinary least squares linear regression or non-parametric k-NN. Since the data in paper II were hierarchical in nature, the parameters of the model were estimated using the lme function of the R statistical analysis software (R Development… 2010), which fits a linear mixed-effects model allowing for nested random effects. In paper IV the model coefficients were estimated using the lm function. The hierarchical nature was not considered, but instead the observations for the plot to which the model applied were excluded from the modelling data when estimating the coefficients. The predictor variables were initially selected from the total set of independent variables using the stepwise and regsubsets functions of the SPSS and R statistical software, the final selection being based on residual plots and r-squared values for the alternative models.

The k-NN method of estimation as used in paper IV is a non-parametric nearest neighbour method in which the estimates for target objects are produced as weighted averages of nearest neighbours chosen from a set of reference objects. The nearest neighbours are searched for within a multidimensional feature space of independent variables using a chosen set of distance metrics. The distance metrics commonly used in the field of remote sensing for forestry purposes are the MSN distance derived from canonical correlation analysis (see Packalén and Maltamo 2006) and squared euclidean distance (see Tuominen et al. 2003). Since the independent variables in paper III were histograms, a suitable distance metric for defining nearest neighbours was the Minkowski distance of order one applied to the histograms. This was calculated as the sum of the class-wise absolute differences between the proportions of the target and reference histograms:

$$D_{pq} = \sum_{i=1}^{n} \left| p_i - q_i \right| , \qquad (1)$$

where D_{pq} is the calculated distance between objects compared, p and q, p_i is the proportion of observations of class i in the target object's histogram, q_i is the proportion of observations of class i in the reference object's histogram, and n is the number of histogram classes. The distance metric may have values from 0 (the histograms of the target and reference objects are the same) to 2 (the histograms have no observations in the same classes). After computing the distances, the neighbours were sorted and nearest neighbours were assigned for every target object. The final estimates for the target object p were then calculated as weighted averages of nearest neighbours, where the weight for the neighbour q is 2 minus the distance measure D_{pq}.

3.1.3 Estimating diameter distributions

In the case of ABSA, the diameter distributions were estimated by direct k-NN imputation, or alternatively using a theoretical Weibull distribution for which the parameters were predicted on the basis of the mean stand variables. The diameter distribution of the target objects in k-NN imputation is an aggregate of the diameter distributions of the nearest neighbour. Every tree in the nearest neighbour plot is weighted by the inverse of the distance measure and the final distribution is the sum of the weighted trees. The height distribution can be produced in a similar manner. If the distributions and mean stand variables are estimated simultaneously, the resulting distributions are realizations of the estimated mean stand variables.

The second approach, which employed a theoretical Weibull function, was the parameter recovery method, in which a pre-specified functional form is assumed to model the diameter distribution and height-diameter (H-D) curve. This approach searches for parameters for these models that result in given values for the mean stand variables. The method is based on setting the predicted values for certain stand characteristics to be equal to the corresponding values computed using the diameter distribution and H-D curve. This leads to a system of equations that need to be solved numerically (Mehtätalo et al 2007). The variables used in the recovery were V, N, G, Dgm and Hgm. The assumed diameter distribution was a left-truncated two-parameter Weibull(α,β) distribution with a fixed truncation point at the minimum measured diameter. The stand-specific H-D relationship was described by the models of Mehtätalo (2005).

In paper I the parameters of a three-parameter Weibull distribution were estimated using the models presented by Kilkki et al. (1989), which employ G and Dgm as independent variables. Tree heights were predicted using the height curve of Näslund (1936), the parameters for which were predicted with the model presented by Siipilehto (1999).

3.1.4 Estimating saw log recoveries

In paper II the saw log recoveries, both theoretical and actual, were estimated directly using regression models. The theoretical saw log recoveries for the field plots were calculated using the tree-level taper curve models of Laasasenaho (1982) and species-specific predefined minimum dimensions for saw logs, and the actual saw log recoveries were obtained by multiplying the tree-level theoretical saw log volumes by a saw log reduction factor (Mehtätalo 2002). Separate regression models were then formulated for the theoretical and actual saw log recoveries based on ALS-derived independent variables.

Only actual saw log recovery was investigated in paper III, this being estimated by the k-NN method using a stem data bank as a source of reference data. The estimated species-specific stand-level H-D distributions were employed as search variables to find the nearest neighbour stands in the stem data bank, the distance metric used being the same as when estimating the H-D distributions (equation 1). The estimate for the actual saw log volume was then a distance-weighted average of the actual saw log volumes of the nearest neighbours scaled to the target stand by reference to the absolute stand-level species-specific volumes.

The theoretical saw log recovery was calculated in paper I by means of a bucking simulation, the input data for which consisted of the set of individual stems extracted from the estimated diameter distribution, with each stem described in terms of species, DBH, height and the taper curve function of Laasasenaho (1982). The result of the bucking simulation was then an optimal, or near optimal recovery of timber assortments from the input data according to predefined price and demand matrices.

Actual saw log recovery was not considered in paper IV, but instead the potential saw log volume was described as the volume of trees with a DBH over the minimum limit of 15 cm, i.e. the volume of saw log-sized trees. This was estimated using the diameter distribution and a height-diameter curve obtained from the ALS-based estimates for the mean stand variables and the volume models of Laasasenaho (1982).

3.1.5 Estimation unit

The estimation unit was a rectangular grid with a cell size of the same area as the reference plot, except in paper I, in which the mean stand variables were predicted using existing ABSA models in a manner similar to that employed in the original paper by Suvanto et al. (2005), i.e. the models were applied at stand level by calculating the independent variables from the laser returns for the whole stand. In the other papers the stand-level estimates aggregations of the grid cell-level estimates for the stand in question.

3.2 Individual tree detection

The ITD approach is a process which consists of consecutive steps from pre-processing of the ALS observations to delineation of the trees and finally estimation of the tree and stand-level attributes. The following process was used in this thesis to implement the ITD approach: first, a canopy height model (CHM) was created from the ALS data and individual trees were identified from the CHM and segmented. Next, various variables related to the individual trees were derived from the ALS and tree segment data, and the ground-measured height and DBH modelled on the basis of these derived variables. The last step was to predict the heights, DBHs and stem volumes of the individually detected trees and finally to calculate the total volumes, numbers of trees and other stand-level variables in order to form an aggregate of the tree-level information.

3.2.1 Canopy height model

The CHM was obtained by interpolating a raster from the dz values of the first returns by taking the maximum dz value within a circle of a certain radius from the centre of each cell of a rectangular grid laid over the target area. The optimal cell size of the grid and optimal radius of the search circle depended on the pulse density, in that the cell size and search radius were required to be as small as possible to maintain the data from individual laser returns, but large enough to avoid introducing too many gaps into the CHM. The cell sizes and search radii were chosen so that they were both the same and approximated to the nominal spacing of the first returns (40 cm in paper I and 50 cm in paper IV). As the pulse pattern was irregular, the process did not produce a value for every grid cell, so that missing values had to be interpolated by taking the average of a 3 x 3 grid cell window. This interpolation was performed successively until every grid cell had a value.

The resulting raster images were further processed using the height-based filtering (HBF) and segmentation methods introduced by Pitkänen et al. (2004) and Pitkänen (2005). The HBF method takes advantage of a priori information regarding the positive correlation between tree height and the size of the canopy and uses Gaussian filtering, so that the filter size increases with the value of the raster cell being processed. The height ranges and corresponding values for the standard deviations (σ) were 0-8 m σ 0.4, 8-16 m σ 0.6, 16-24 m σ 0.8, 24-32 m σ 1.0 and 32-40 m σ 1.2.

3.2.2 Detection of individual trees

The actual detection of individual trees starts with a search for local maxima which can be considered candidates for treetop locations from the filtered raster image. After that, the image is subjected to watershed segmentation using a drainage direction-following algorithm (Gauch 1999, Pitkänen 2005). A threshold value can be set to mask out raster cells which are probably objects other than tree canopies (e.g. low vegetation or stones). The segmentation results in candidate tree segments with a number of attributes describing them: the maximum dz value of the segment (MaxDz), area of the segment (ASeg), maximum diameter of the segment (MaxDSeg), diameter perpendicular to the maximum diameter (PerDSeg), the average of these diameters (AvgDSeg) and the xy-coordinates of MaxDz (XySeg).

3.2.3 Modelling height and DBH for individually detected trees

The candidate tree segment attributes are not themselves useful for practical forestry, except for MaxDz, which can be related to tree height. MaxDz is not always an observation of the exact tree top, however, as it may be affected by a combination of factors such as pulse density, size of the footprint, scanning angle and the structure of the tree canopy so that it may underestimate the real tree height. Thus, if ground truth measurements of individual trees exist that can be linked to the candidate tree segments, the relation between the actual tree attributes and the candidate tree segment attributes can be modelled. In addition to the candidate tree segment attributes, area-level variables derived from the canopy height distribution or from aerial images can be used to estimate individual tree attributes. Since no ground reference based on individual trees was used in paper I, MaxDz was taken to represent the tree height and DBH was estimated using a local DBH-height regression model formulated from plot-level tree measurements. Individual tree measurements for the area concerned were available for paper IV, however, and therefore local regression models were constructed for tree height and DBH separately. Different local models were employed for estimation purposes, using candidate tree segment attributes and/or area-level variables as independent variables. As an alternative approach to estimating DBH, use was made of the existing species-specific regional models as presented by Kalliovirta and Tokola (2005). These models were formulated using National Forest Inventory (NFI) field data, with tree height and maximum tree crown diameter used as independent variables. The regional models were calibrated for the area concerned in paper IV by reference to local individual tree measurements.

3.2.4 Predicting tree-level and stand-level variables

Tree heights for the candidate tree segments were estimated using either a local model (paper IV) or MaxDz (paper I). DBH was then estimated by means of a regional model or local model. Stem volumes were predicted using the species-specific stem volume models of Laasasenaho (1982), which employ DBH and height as independent variables. Finally, the stand-level mean stand characteristics were formed as aggregates of the individual tree attributes. The saw log recoveries in paper I were estimated using bucking simulation in which the input data consisted of the set of individually detected trees, each described in terms of species, DBH, height and the taper curve function of Laasasenaho. In paper IV the volume of saw log-sized trees was the sum of the volumes of the trees with a DBH over 15 cm.

3.3 Accuracy assessment

The final accuracy assessment was performed at the stand level in all the papers (the “plot” in paper IV corresponds to a stand and “subplot” to a plot). This was possible since the ground reference was measured using wall-to-wall methods, i.e. the whole tree population was measured in the field. The field method concerned in papers I-III was harvester measurement, and that in paper IV a manual field inventory in which all the trees in each stand were tallied. In the case of the ABSA models the estimation accuracy was also assessed at the plot level, this plot-level assessment being performed in papers III and IV by leave-one-out cross-validation with the plots within the same stand as the target plot excluded from the modelling data in each instance. In paper II the parameters of the models were estimated using mixed modelling with the stand as a random parameter.

The estimation accuracies of the mean stand characteristics (V, G, N, Hgm, and Dgm), Hdom and saw log recoveries were assessed by calculating the root mean squared error (RMSE):

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

and bias:

$$bias = \frac{\sum_{i=1}^{n}(y_i - \hat{y}_i)}{n}, \tag{3}$$

where n is the number of plots, y_i is the observed value for the stand characteristic y, and $\hat{y}_i$ is the predicted value. RMSE and bias were also calculated as relative values, i.e. the RMSE and bias values were divided by the observed mean values for stand character y. RMSE and bias were also used to evaluate the tree height and DBH models in paper IV.

The goodness-of-fit of the diameter distribution was investigated using the Kolmogorov-Smirnov goodness-of-fit test and the error index proposed by Reynolds et al. (1988). The equation for the latter is:

$$e = \sum_{i=1}^{k} w_i \,|\, f_i - \hat{f}_i \,|, \tag{4}$$

where k is the number of diameter classes, f_i is the observed number of stems, $\hat{f}_i$ is the predicted number of trees in diameter class i, and w_i is the weight of class i (Reynolds et al. 1988). Basal area was used in weighting.

4 RESULTS

This summary will concentrate on presenting the main results of this work, namely the accuracies achieved in estimating diameter distributions and saw log recoveries.

4.1 Diameter distributions

Diameter distributions were estimated either using ITD, where the diameter distribution is an aggregate of the results for individually detected trees, or by ABSA, where it is formed from k-NN imputed trees or where a theoretical distribution is fitted to estimated mean values (parameter recovery method or parameter prediction method). In paper I the diameter distributions were estimated using field measurement-based techniques as a reference. A summary of the estimation errors observed in the various papers is presented in Table 1.

In paper I the ITD method produced a very well fitting diameter distribution in terms of the calculated Reynolds' error indices, although the inventory by compartments method, which is based on field assessments and is traditionally used for stand-level inventories in Finland, combined with a theoretical Weibull function, produced almost as good a result in this particular case. The largest difference between the two methods appeared in the smaller diameter classes. The ITD method was also able to find smaller trees, whereas fitting a theoretical Weibull function to the estimated mean values produced a diameter distribution function which described the number of large trees exceptionally well but failed to describe the small trees.

The stand in question in paper I was also included in the test stands for paper III. The diameter distribution estimated for that stand using k-NN had of Reynolds' error index value of 8.4, which is almost double that achieved with ITD. This particular stand could be considered as an easy case for ITD; it had a low number of stems/ha (< 500) and had been thinned few years previously, so that the trees were not spatially clustered. The diameter distributions were also estimated by species in paper III, yielding results that were not as accurate as the total distributions.

In paper IV all the stands were assessed using ITD and ABSA, and the basal area-weighted Reynolds' error index values for the diameter distribution estimates indicate that there was no difference in accuracy between the two methods in this data set. It was noticed, however, that several stand-level variables correlate with the estimation accuracy and that some variables have the opposite effects with the two methods. The most notable difference was found in the effect of tree size variation, since a large variation in tree size increased the estimation error attached to the diameter distributions when ITD was used, whereas in the case of ABSA tree size variation had no effect on the estimated diameter distribution. Other variables that decreased the estimation accuracy of the diameter distributions or the overall estimation accuracy with the ITD approach were a large number of trees, a high canopy coverage and a clustered spatial distribution of trees. In the case of ABSA a large mean tree size, small number of trees and more regular distribution of tree locations, which are all typical features of old, managed forests, reduced the accuracy of the diameter distribution estimates when evaluated in terms of the basal area-weighted Reynolds' error index.

Table 1. Basal area-weighted Reynolds' error indices for the estimated diameter distributions. All values are calculated using 2 cm diameter classes, the smallest class being 7-9 cm. The values in parentheses are from species-specific diameter distributions.

Paper	Method of estimation	Mean stand size, ha	Reynolds' error index	
			Average	Stdev
I	ITD	5.9	4.6	-
I	Inventory by compartments, parameter prediction method	5.9	7.2	-
III	ABSA, k-NN	3.1	7.2 (12.0)	2.3 (3.4)
IV	ITD	0.09	17.8	6.8
IV	ABSA, parameter recovery method	0.09	19.2	5.8

Sample diameter distributions for two stands that differ in their attributes are presented in Figures 1 and 2. The stand in Figure 1 is a mature one with a low number of trees and a large average tree size, while that in Figure 2 is a younger stand with large number of trees of a small average size.

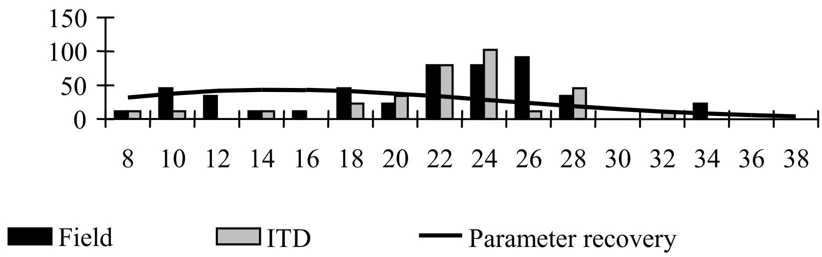

Figure 1. Estimated and field measured diameter distributions for a mature stand. The stand attributes are: total volume 152.2 m^3/ha, basal area 18.8 m^2/ha, number of trees 490/ha, dgm 24.8 cm, hgm 18.6 m, canopy coverage 43.7 %, index of dispersion 0.82, coefficient of variation in DBH 0.31 and coefficient of variation in tree height 0.27. The basal area-weighted Reynolds' error index values for ITD and ABSA are 10.8 and 15.9, respectively.

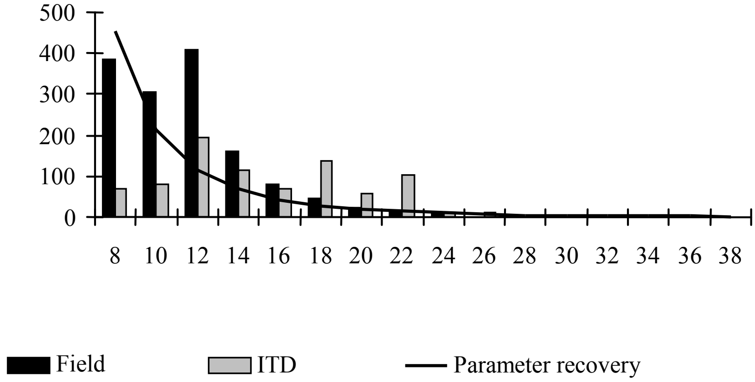

Figure 2. Estimated and field measured diameter distributions for a young stand. The stand attributes are: total volume 94.5 m^3/ha, basal area 17.0 m^2/ha, number of trees 1917/ha, dgm 13.6 cm, hgm 11.5 m, canopy coverage 69.0%, index of dispersion 1.29, coefficient of variation in DBH 0.37, and coefficient of variation in tree height 0.20. The basal area-weighted Reynolds' error index values for ITD and ABSA are 14.2 and 10.8, respectively.

4.2 Saw log recoveries

The saw log recovery estimation results are presented in Table 2. The theoretical saw log recovery was estimated with promising accuracy in papers I and II, but the relative error in estimating the volume of saw log-sized trees in paper IV was about double that in estimating the total volume. Again there was considerable variation between the stands, however. The results presented in paper IV suggest that the estimation accuracy correlates positively with the average mean tree size and negatively with the index of dispersion in tree locations. Thus more accurate results can be expected for mature stands with a regular spatial pattern of tree locations. On the other hand, if we consider the effects of the stand variables on the accuracy of the saw log recovery estimates, there seems to be no difference between the methods, i.e. the stand-level variables seems to affect the results in a similar way regardless of the estimation method. The mechanisms by which they affect the results are different, of course, but the final outcome seems to be same. The ABSA approach performed better in this test on average, the most distinct difference being in the estimation biases, as ITD overestimated the saw log volume by 20 to 30 percent, whereas the bias with ABSA was only a few percent.

Table 2. Accuracy of saw log recovery estimate.

Paper	Method of estimation	Number of stands	Mean stand size, ha	RMSE% of estimate			
				Total theoretical	Total actual	Pine actual	Spruce actual
I*	ITD	1	5.9	0.4** 10.6***			
I*	ABSA, regression + parameter prediction	1	5.9	18.6			
I*	Field assessment + parameter prediction		5.9	4.4			
II	ABSA, direct regression models	14	3.1	9.1	18.0		
III	ABSA, k-NN	14	3.1		18.6	61.8	31.5
IV****	ITD	14	0.09	36.2** 42.1***			
IV****	ABSA, regression + parameter recovery	14	0.09	26.4-27.1			

*estimation error calculated as |field measured value – estimated value| x 100. **local height-dhb model. ***regional height-DBH model. ****estimate is volume of trees with DBH>15cm.

In paper II, where both theoretical and actual saw log recoveries were estimated, the RMSEs in the test data were 19.2 m^3 (9.1%) and 30.6 m^3 (18.0%) at the stand level for the theoretical and actual saw log recovery values, respectively, and the corresponding biases were -10.2 m^3 (-4.8%) and -7.2 m^3 (-4.3%). The actual saw log volume estimation errors were much higher than theoretical ones, but the estimation methods (direct regression modelling in paper II, and k-NN and stem data bank in paper III) resulted in approximately the same accuracies for the total actual saw log volume estimates. The k-NN estimates were slightly less biased than those based on regression models: 1.1% vs. -4.3%. The species-specific saw log recoveries were estimated with a rather poor level of accuracy, the RMSE of the total volume estimate in paper III being less than 10% and that for the saw log estimate about 10 percentage points higher (18.6%). The RMSEs for the species-specific total volumes were 47.7% and 20.3% for pine and spruce, respectively. The RMSEs of the saw log volumes were also about 10 percentage points higher here (61.8% and 31.5 % for pine and spruce respectively), and the biases of the saw log volume estimates were higher than those of the total volumes in all cases.

5 DISCUSSION

Airborne laser scanning-based forest inventories are usually considered an alternative to field inventories. In some sense this is correct, since the timber variables involved are usually the same. Nevertheless, operational ALS-based inventories still lack information on minor species, forest site types, forest habitats, biodiversity issues and forest age, for example, although some attempts have been made to estimate these variables (Breidenbach et al. 2010, Vehmas et al. 2008, Pesonen et al. 2008, Weber & Boss 2009, Maltamo et al. 2009c). One group of variables that are also missing from operational ALS inventories are tree quality characteristics. On the other hand, ALS-based forest inventories have some features that field inventories usually do not have. They measure the forest canopy structure directly, for instance, which provides excellent material for modelling related variables such as spatial, height and diameter distributions. ALS inventories can also be carried out on a wall-to-wall basis for areas of considerable size. Thus both approaches (ABSA and ITD) also allow the examination of forest attributes in terms of spatial xy coordinates, i.e. in a horizontal direction, in addition to inspection of the local (stand or plot-level) forest structure.

The aim of paper I was to test the individual tree detection method for producing information for use in wood procurement planning, validating it against harvester-collected stm data. Individual tree detection can be regarded as having been successful with respect to the accuracies of the estimates for the number of trees and the saw log volumes calculated using bucking simulation. The examination of the diameter distributions and diameter-length distributions of logs similarly pointed to the superiority of the ITD method used here over the other methods tested, i.e. area-based ALS estimation and two field inventories. When the same stand was assessed with k-NN in paper III, ITD again produced a better-fitting diameter distribution. The forest stand may have had some features that could have favoured ITD in this particular case. Its management history as assessed visually, for example, indicated that the stand had been thinned according to the existing forest management guidelines, so that the trees were not located in clusters, imply easier segmentation of the individual tree crowns. The tree segmentation accuracy as such was not examined, however, since the reference data did not include accurate positions for the trees. In addition, the ground truth data favoured the ITD approach, as harvester data include only those trees that are processed by the harvester, i.e. the ground truth data may not include all the small trees in the sub-dominant layer, since they do not fulfil the minimum dimensions for harvested timber assortments. These small trees are still assumed to be included in the stand characteristics as estimated by the methods used for comparison, however, although their effect on the results in this case will be minimal except where the number of trees is concerned. Furthermore, plot level ABSA models were employed at the stand level, which may detract from the estimation results. Since some of the lidar-derived predictors (height percentiles) have non-additive features, lidar-based prediction models are scale-dependent (Zhao et al. 2009), which means that an aggregation of predicted values for individual cells will not be equal to a single prediction based on stand-level predictors. A better alternative would have been to compose stand-level results from the cell-level estimates, as done by Packalén & Maltamo (2008). The importance of DBH estimation was ascertained by comparing a DBH prediction model constructed using local data with regional models. The local model resulted in more accurate estimates for the diameter distributions, saw log volumes and diameter-length distributions of logs. The data available did not support any consideration of tree quality or species recognition issues in this connection, however. The used ALS equipment, Optech ALTM 2033, has two separate electronic circuits receiving the return signal, which are known to require separate calibrations frequently. There was no exact information about the calibration and the differences of the first and last pulse relative heights. The relative difference can be several tens of centimetres (Næsset 2002). Because the DTM was produced from the last pulse data and the CHM from the first pulse data, the relative difference between last and first pulse echo heights affects directly to the estimated tree heights, and furthermore, the estimated diameters, taper curves and volumes. The effect of this was not studied since there were no field observations from individual trees, which would have allowed the inspection of tree height estimation accuracy. Flying altitude affects the properties of the point cloud, as well. Since the ALS data used in ITD was a combination of higher acquired low-density data and lower acquired high-density data, there is a risk of bias due the un-uniform properties of the point data. The risk for this was considered negligible, since the both ALS campaigns covered the whole analysis area, i.e. the data itself was uniform over the whole area.

The starting point in paper II was to take quality reduction into account when estimating stand-level total saw log recoveries. Harvester-collected stm data were employed as ground truth data. The field measured theoretical saw log recovery was calculated using diameter measurements contained in the stm files and predefined minimum dimensions for saw logs, whereas the field measured actual saw log recovery was the sum of the volumes of logs given in the stm files. The result can be considered good by comparison with the estimation accuracies reported by Rooker Jensen et al. (2006) and Bollandsås et al (2010a), although direct comparison is difficult because of differences in the variation in stand characteristics between the areas concerned. Rooker Jensen et al. (2006) achieved RMSEs of 57.4 m^3 and 32.4 m^3 in their validation data for the theoretical volumes of small saw logs and large saw logs, respectively. Bollandsås et al (2010a) reported RMSE of 41.3 m^3/ha for actual saw log volume. Maltamo et al. (2007) modelled basal area and stem frequency distributions using ALS data and estimated the volume of saw-wood-sized trees (trees with dbh > 17 cm) from these diameter distributions. The relative RMSEs of the stand-level volume estimates were quite high, varying between 17.2 and 23.0 percent.

The results of paper II are also quite encouraging when compared with the accuracies of field inventories by compartments, as according to Haara and Korhonen (2004) the RMSE for stand-level saw log volume in an inventory by compartments in Finland was 23.6 m^3 (44.6%). Moreover, the absolute RMSE for mature stands was even higher in absolute terms, 36.9 m^3 (28.2%). The biases for the estimates of saw log volume and the saw log volume of mature stands quoted by Haara and Korhonen (2004) were 1.2 m^3 (2.3%) and 3.1 m^3 (2.4%), respectively. Their RMSEs and biases apply to actual saw log volumes calculated using the saw log reduction models of Mehtätalo (2002) for both the inventory by compartments method and systematic plot sampling, the latter used as validation data, whereas the validation data in paper II consisted of the field measured saw wood recovery recorded contained in the stm files.

Estimated saw timber ratios obtained by computational methods have been between 4 and 26 percent (Tommola et al. 1999, Malinen et al. 2001, Malinen 2003), but the most accurate results presented in those paper were achieved using accurately measured stand data and not taking into account the effect of the defects.

The saw log volumes in paper II were estimated using stratification by species, i.e. the modelling data were stratified by dominant tree species and separate models were constructed for each stratum. In the model application phase the stands were assigned to the correct stratum using a-priori information concerning their species composition. The estimation of actual saw log recoveries was taken one step further in paper III, where the aim was to estimate species-specific diameter distributions and actual saw log recoveries by species. The estimation process introduced a new application of the notion of fusing of harvester and ALS data: the use of harvester data as a stem data bank with ALS-predicted stand variables. The process is analogous to the method proposed by Malinen et al. (2001)

and Malinen (2003) with the exception that in paper III the search variables were estimated using remote sensing. The stand-level harvester data in paper III were related to remotely sensed stand data, and additional information that was not directly estimated using a remote sensing-based inventory (i.e. saw log volumes) was derived from the stem data bank. In the ideal case a stem data bank should consist of stands delineated and covered by the remote sensing data used in an actual remote sensing-based forest inventory. For example, in the ideal situation all the stands in a stem data bank should be submitted to airborne laser scanning with the same specifications as the stands selected for the inventory. This is rarely the case, however, and therefore stem data bank information cannot be related directly to remote sensing data. Instead, a two-phase inventory should be used, in which the variables which can be related to the stand data bank (e.g. tree species proportions, diameter distributions) are estimated for the inventory units in the first phase and then used in the second phase to find stands with similar attributes in the stem data bank.

The two-phase inventory in paper III was carried out using the following process. First the diameter distributions were estimated using k-NN imputation based on the ALS data and aerial image-derived height and spectral histograms. Next, the estimated diameter distributions were used to find the k nearest neighbours in a stem data bank consisting of information extracted from stm files for 35 clear-cut stands located within 250 km of the target stands, and then the actual saw log recoveries for the target stands were imputed from the actual figures for the k nearest stands in the stem data bank.

The diameter distributions were estimated with an accuracy comparable to those reported in earlier ALS studies. The distributions for the dominant species were estimated more accurately than those for the minor species, but it was also possible to obtain multimodal distributions by this method. Exact comparison of the accuracy of the estimated diameter distributions with the results of previous studies is difficult because of differences in the variations in stand characteristics between the areas concerned, and more importantly, because of non-standard accuracy assessment methods. Goodness-of-fit tests such as the KS test, based on Kolmogorov-Smirnov test statistics, and indices such as the error index proposed by Reynolds et al. (1988) are well-known and practicable methods for this purpose. The problem is that changes in the distribution parameters (diameter class width, unit on the y-axis) and weighting of the classes can have considerable effects on the resulting test or index value. The accuracy of the actual saw log recovery estimates can be compared with the results presented in paper II, since the same validation data were used in both papers. The accuracies of the total actual saw log volume estimates were approximately similar, with the direct regression models producing slightly more accurate results, but the species-specific actual saw log volume estimates were not accurate. Haara and Korhonen (2004) reported RMSEs of 16.9 m^3 (52.0%) and 27.2 m^3 (62.3%), respectively, for saw log volumes of pine and spruce, and the accuracy achieved in paper III was comparable in the case of spruce, but the estimation error was larger for pine. The biases for both species were also significantly larger. One possible reason for the poor estimation accuracies could be the low number of reference stands in the stem data bank. The inclusion of test stands in the stem data bank and their evaluation by the leave-one-out cross-validation technique decreased the estimation errors. Thus it can be assumed that a more representative stem data bank consisting of stands located near the inventory area would improve the estimation result. Packalén and Maltamo (2008) estimated theoretical saw log volumes from k-MSN-imputed tree lists and alternatively from predicted theoretical diameter distributions based on ALS data. They did not report total saw log volumes, but the stand-level species-specific RMSEs for their saw log volume estimation accuracies using a Weibull distribution were 20.7 m^3 (41.0%), 32.4 m^3 (61.1%), and 7.0 m^3 (142.9%) for pine, spruce and deciduous species, respectively. The corresponding figures obtained from diameter distributions constructed using k-MSN imputation were 20.2 m^3 (40.0%), 23.1 m^3 (43.6%), and 5.6 m^3 (114.5%). The results were reported to be better than those of an inventory by compartments. In their study of the effects of different sources of error on predicted timber assortments, Holopainen et al. (2010) generated the stem distributions using Weibull distributions and the parameter prediction method and used the harvester machine's stm data for validation. They divided the errors into three parts: 1) errors in stem-form prediction and simulated bucking, 2) errors in generation of the stem distributions, and 3) inventory error. Out of these it was the inventory error that was most significant. When the errors were combined, the RMSEs for ALS-based inventories of pine, spruce and birch saw wood, respectively, were 7.0 m^3 (79.2%), 35.5 m^3 (33.6%), and 6.2 m^3 (78.6%). The field inventory method used for comparison, inventory by compartments, resulted in significantly larger errors in the case of pine and birch saw wood, but similar errors for spruce saw wood.

In paper IV a different data set was used, one which allowed more thorough examination of the performance of the ITD and ABSA methods in retrieving forest stand characteristics and a better comparison of these methods. The research framework was planned so that neither of the methods was favoured. The main findings were that: 1) the ITD method based on segmenting individual trees without calibration underestimates the number of trees, and though the effect of unsuccessful segmentation is not necessarily seen in the estimates for V and G, it is clearly perceived in the errors in estimating Dgm, Hgm and the volume of saw log-sized trees. 2) The tree size distribution and the spatial distribution of tree locations correlates with the estimation accuracies with both methods, ITD and ABSA. The first finding confirms the results presented by Packalen et al. (2008) and Vastaranta et al. (2009a), that both methods provide equally accurate estimates for V and G whereas ITD produces a bias in the estimated number of trees. The effect of errors in tree delineation on the Dgm and Hgm estimates were nevertheless more clearly perceived in paper IV than in earlier studies. Finding 2) confirms earlier observations with regard to ITD, but in the case of ABSA this is a new empirical finding and should be investigated more thoroughly in further studies. If the tree size distribution and the spatial distribution of trees affect the estimation accuracies to a significant extent, and if the predictive variables describing the distributional characteristics are not included in the ABSA models, the variables of a target unit (forest stand or plot) with spatial and tree size distributions that differ from those of a typical plot in the modelling data may be estimated with poor accuracy. Thus the modelling plots should also represent the variation in the tree size distribution and spatial distribution of trees to be found in the whole inventory area. Furthermore, if the inventory area consists of forest stands with varying spatial and size distributions, a predictive variable, or set of variables, correlating with the tree size distribution and spatial distribution should be included in the prediction model.

Since the data set used in paper IV included only observations on pine stands, species recognition issues were not considered. Nevertheless, there are many previous studies dealing with species recognition issues in connection with ALS-based forest inventories. The work of Packalén (2009) is concerned with area-based estimation methods and gives a thorough description of how species-specific growing stock estimates can be produced by combining ALS data with aerial images. Another area-based method for taking species into account involves stratifying the data by forest development class, site type and species composition so that separate models are constructed for each stratum (see Næsset 2002).

The drawback with the stratification method, however, is that it is possible to target only the stand estimates for the main tree species. Species recognition has been studied more widely in the case of ITD, with investigations into the recognition of commercially important species in boreal forests by means of geometrical features derived from laser point clouds (Vauhkonen et al. 2009), laser intensity variables (Korpela et al. 2010), a combination of intensity and laser height distribution variables (Ørka et al. 2009) and a combination of intensity and height distribution variables and geometrical features (Holmgren and Persson 2004, Vauhkonen et al. 2010), for example. Furthermore, Holmgren et al. (2008) identified species by combining predictors derived from ALS data and aerial images. The overall classification accuracies have varied from 70% to over 90% (classification into three classes: pine, spruce and deciduous). The most promising results have been achieved using very dense ALS data and a combination of several predictor variables.

One interesting issue for further research could be to study the ALS-based estimation of tree quality by means of bucking simulation. The estimates of quality characteristics could be produced by either ITD or ABSA using the k-NN method, and the result of the ALS estimation would then be a tree list with information on quality parameters which affect saw log recovery. Applications of this kind would place high requirements on the ground truth data, however, as the quality information would have to be measured for numerous trees that fulfilled the minimum saw log dimensions. Also, it would probably not be realistic to attempt to estimate all the defects affecting the tree quality from ALS data, although it would be theoretically possible using nearest neighbour imputation methods, given that the height of the lowest dead branch, for example, was estimated successfully by Maltamo et al. (2009a).

The estimation of additional variables would place higher requirements on the field plot sample, as well. Several attempts have been made recent to determine what is a sufficient number of field plots. Packalén and Maltamo (2006) suggested that the adequate number of reference plots for predicting species-specific variables in Nordic boreal forests may be as high as several hundred, whereas less than 50 plots might be enough for estimating total variables, e.g. total volume or mean height (Næsset 2002, Lim et al. 2003, Holmgren 2004, Maltamo et al. 2011.). Næsset and Gobakken (2008) reported that the standard deviations of the differences of estimated and field measured biophysical stand properties increased when the number of field plots reduced from 100 to 75 percents, and the increase was even larger when the reduction was from 75 to 50 percents. Total number of plots in their study was 132 and the plots were divided in three strata. They also noted that there was a trade of between plot size and number of plots; larger plot size could to a certain extent compensate for reducing the number of field plots. Maltamo et al. (2009b) noted that less than 100 plots could be sufficient for estimating total diameter distributions (but not species-specific ones) under boreal conditions if the plot sample were to be well representative of different forest types. A representative sample can be obtained using ALS as a priori information in sampling, and an adequate field plot sample can be obtained with a lower number of plots when using ALS-derived height and density than when using random sampling (Hawbaker et al. 2009, Maltamo et al. 2011). The estimation of quality characteristics and the proportions of timber assortments, however, is likely to require several hundred carefully sampled field observations rather than less than one hundred to be accurate enough to give additional information for operative wood procurement.

One possibility for obtaining field plot data could be terrestrial laser scanning (TLS). This can be used for estimating accurate tree-level stem profiles (Thies et al. 2004, Murphy

2008) and branch heights (Henning and Radtke 2006). Thus TLS could be employed for collecting ground truth data on timber assortments and tree quality, replacing laborious manual field measurements. The canopy structure, vegetation density and understorey vegetation all have an effect on the success of detecting trees in TLS data, however (Liang et al. 2009), and terrestrial laser scanners are expensive compared with more traditional field equipment and are not suitable for use under difficult field conditions (Vastaranta et al. 2009b). These factors have prevented the use of TLS for operative field plot surveys, although there has been a great deal of active research into this issue.

A stem data bank offers an opportunity to obtain information on the recovery of timber assortments without significant changes in the field plot measurements. The timber assortments for the target stand can be estimated by searching the stem data bank for stands with similar variables. This approach was tested in paper III without any notable success, for several possible reasons. It is clear that a more representative stem data bank, the inclusion of stand variables such as site index and soil class, and the use of forest density variables (basal area/ha, stem count/ha) could have improved the results. Site index and soil class information are not collected during harvesting, and it is therefore unlikely that such information will exist, or at least be reliable enough, in the near future. Forest density, though, can be estimated from the stem data bank information based on the coordinates recorded for the individual trees, provided these coordinates are recorded. Positional inaccuracy, or missing coordinates for some trees, will not prevent the use of coordinates for delineating rough borders for the stand. For example, if the average stand size is 3 ha, which was the mean test stand area in paper III, and if we assume that every stand is a circle, a 5 metre systematic delineation error (either inward or outward from the real stand border) will result in about a 10% error in the stand area. By investigating the harvester operator’s working methods and taking into account the parameters of the harvester machine (reach and load capacity of the operating boom) it is possible to reduce the positioning error (Rasinmäki & Melkas 2005) and probably eliminate any systematic error. For large inventory areas it may also be possible to obtain harvester data with ALS coverage. If some stands inside the inventory area are “measured” with the harvester just after ALS acquisition the harvester and ALS data can be used jointly for making local models to predict timber assortments without a two-stage estimation process. Such an application would require the time gap between ALS acquisition and harvesting to be short, preferably not more than few months, depending on the growth period.

The timing of ALS acquisition restricts the use of all the methods presented here. Since potential stands for final felling are usually small in area and scattered in location, it may not be economically feasible to collect ALS data only for producing wood procurement planning information every year. In Finland, for example, extensive ALS inventory projects are scheduled in periods of many years at a time, so that information may be collected from the same area every 5 to 10 years. Tree size distributions and other characteristics of the growing stock that are estimated using ALS techniques could be updated by means of growth models, but this would in itself introduce error into the estimates.

6 CONCLUSIONS

The results obtained in this thesis indicate that diameter distributions and timber assortment recoveries can be estimated using either of the ALS approaches, ITD or ABSA. The critical part of the ITD process is delineation of the trees. If this is not successful the estimates may be badly biased. In ABSA the estimates are unbiased, but the bottleneck could be accuracy. The estimation method should be selected according to data needs and inventory area - specific requirements. However, for estimating species-specific diameter distributions and timber assortments for wood procurement planning the field data collected for estimating stand-level inventory information may not be sufficient. The field sample data must be representative of the whole variation to be found in the inventory area. In case of ABSA this means extensive field plot measurements. If a locally representative stem data bank is available, the manual fieldwork related to field plot sampling might be reduced significantly. In case of ITD the field data should cover individually located and precisely measured trees. The amount of fieldwork could then be drastically smaller than in the case of ABSA, because it may be enough to measure only 500 trees instead of 500 plots to provide information that can be used for estimating the variables of interest. The approaches can also be seen as mutually complementary. For example, the basic method used in large-area inventories could be ABSA, complemented in some cases (e.g. potential areas for final felling) with tree quality estimates provided by ITD. The information provided by the ALS-based inventory is of no use if it does not fulfil the quality requirements set by the planning system and, more importantly, if the planning systems cannot use the new information.

The information provided by ALS-based forest inventories has certain benefits over the field inventories. These are the spatially continuous coverage of the information over the area of interest, reliable accuracy estimates and possibility to automate the interpretation process. The information is also always spatial information, i.e. it can be managed and analysed in geographical information systems.

REFERENCES

Axelsson P. 2000. DEM Generation from Laser Scanner Data Using Adaptive TIN Models. International Archives of Photogrammetry and Remote Sensing 33(B4): 110-117.

Bollandsås, O. M. & Næsset, E. 2007. Estimating percentile-based diameter distributions in uneven-sized Norway spruce stands using airborne laser scanner data. Scandinavian Journal of Forest Research 22: 33-47.

—, Maltamo, M., Gobakken, T., Næsset, E. & Lien, V. 2010a. Prediction of timber quality parameters of forest stands by means of small footprint airborne laser scanner data. International Journal of Forest Engineering. Conditionally accepted.

—, Maltamo, M., Gobakken, T. & Næsset, E. 2010b. Comparing parametric and non-parametric modeling of diameter distributions on independent data using airborne laser scanning. Manuscript.

Breidenbach, J., Gläser, C. & Matthias, S. 2008 Estimation of diameter distributions by means of airborne laser scanner data. Canadian Journal of Forest Research 38: 1611-1620.

—, 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: 911–924.

Cox, D. R. and Lewis, P. A. W. 1966. The Statistical Analysis of Series of Events. Methuen's Monographs on Applied Probability and Statistics. John Wiley, London. 285 p.

Dean, T., Cao, Q., Roberts, S. & Evans D. 2009. Measuring heights to crown base and crown median with LiDAR in a mature, even-aged loblolly pine stand Forest Ecology and Management 257: 126–133.

Flewelling, J. W. 2008. Probability models for individually segmented tree crown images in a sampling context. In: Hill R.A., Rosette, J. & Suárez, J. (ed.). Proceedings of SilviLaser 2008: 8th international conference on LiDAR applications in forest assessment and inventory. Edinburgh, UK. p. 284-293.

— 2009. Forest inventory predictions from individual tree crowns: regression modeling within a sample framework. In: McRoberts, R. E., Reams, G. A., Van Deusen, P. C. & McWilliams, W. H. (ed.). Proceedings of the Eighth Annual Forest Inventory and Analysis Symposium 2006. Monterey, California. p. 203-210.

Gauch, J.M. 1999. Image segmentation and analysis via multiscale gradient watershed hierarchies. IEEE Transactions on Image Processing 8(1): 69-79.

Gobakken, T. & Næsset, E. 2008. Assessing effects of laser point density, ground sampling intensity, and field sample plot size on biophysical stand properties derived from airborne laser scanner data. Canadian Journal of Forest Research, 38: 1095-1109.

— & Næsset, E. 2009. Assessing effects of positioning errors and sample plot size on biophysical stand properties derived from airborne laser scanner data. Canadian Journal of Forest Research 39: 1036-1052.

— & 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.

— & Næsset, E. 2005. Weibull and percentile models for lidar-based estimation of basal area distribution. Scandinavian Journal of Forest Research 20(6): 490-502.

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

Hansson, F. 1999. Inventering före avverkning – metodval och resuråtgång. Skogforsk Resultat 15/1999. 4 p.
Hawbaker, T. J., Keuler, N. S., Lesak, A. A., Gobakken, T., Contrucci, K. & Radeloff, V. C. 2009. Improved estimates of forest vegetation structure and biomass with a LiDAR-optimized sampling design. Journal of Geophysical Research 114, G00E04. p. 11.
Henning, J. G. and Radtke, P. J. 2006. Detailed stem measurements of standing trees from ground-based scanning lidar. Forest Science 52(1): 67– 80.
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.
— & Persson, Å. 2004. Identifying species of individual trees using airborne laser scanning. Remote Sensing of Environment 90: 415-423.
—, Persson, Å. & Söderman, U. 2008. Species identification of individual trees by combining high resolution LiDAR data with multi-spectral images. International Journal of Remote Sensing. 29(5):1537-1552.
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(6): 1131-1142.
Hyyppä, J. & Inkinen, M. 1999. Detecting and estimating attributes for single trees using laser scanner. The Photogrammetric Journal of Finland 16(2): 27-42.
Junttila, V., Maltamo, M. & Kauranne T. 2008. Sparse Bayesian Estimation of Forest Stand Characteristics from Airborne Laser Scanning. Forest Science 54(5): 543-552.
Kalliovirta, J. and Tokola, T. 2005. Functions for estimating stem diameter and tree age using tree height, crown width and existing stand data bank information. Silva Fennica 39(2): 227-248.
Kangas, A. & Maltamo, M. 2000a. Percentile based basal area diameter distribution models for Scots pine, Norway spruce and birch species. Silva Fennica 34: 371–380.
— & Maltamo, M. 2000b. Performance of percentile based diameter distribution prediction and Weibull method in independent data sets. Silva Fennica 34(4): 381–398.
— & Maltamo, M. 2002. Anticipating the variance of predicted stand volume and timber assortments with respect to stand characteristics and field measurements. Silva Fennica 36(4): 799–811.
Kilkki, P., Maltamo, M., Mykkänen, R. & Päivinen, R. 1989. Use of the Weibull-function in the estimation of the basal area DBH-distribution. Silva Fennica 23(4): 311-318.
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.
Kohonen, T. 1995. Self-Organizing Maps. Springer. Germany. 349 p.
Korpela, I., Ørka, H.O., Maltamo, M., Tokola, T. & Hyyppä, J. 2010. Tree species classification using airborne LiDAR – effects of stand and tree parameters, downsizing of training set, intensity normalization, and sensor type. Silva Fennica 44(2): 319–339.
Laasasenaho, J. 1982. Taper curve and volume functions for pine, spruce and birch. Communicationes Instituti Forestalis Fenniae 108. The Finnish Forest Research Institute, Helsinki, Finland. 74 p.
Lappalainen, J. 2005. Leimikon tukkijakauman ennustaminen itseohjautuvan kartan tuottamaa luokittelua käyttäen. M.Sc. thesis, Faculty of Forestry, University of Joensuu, Finland. 39 p.
Larsson. H. 2009. Flygburen laserskanning kopplat till skördarmätning för datainsamling till operativ planering. Arbetsrapport 260. Sveriges lantbruksuniversitet. Umeå. 42 p.
Leckie, D., Gougeon, F., Hill, D., Quin, R., Armstrong, L. & Shreenan, R. 2003. Combined high-density lidar and multispectral imagery for individual tree crown analysis. Canadian Journal of Remote Sensing 29(5): 633-649.

Liang, X., Litkey, P. Hyyppä, J., Kaartinen, H., Vastaranta M. & Holopainen, M. 2009. Automatic stem location mapping using TLS for plot-wise forest inventory. In: Popescu, S., Nelson, R., Zhao, K. & Neuenschwander, A. (ed). Proceedings of Silvilaser 2009: The 9th international conference on lidar applications for assessing forest ecosystems. Texas A&M Univerisity College Station, USA. p. 314-323.

Lim. K, Treitz, P., Baldwin, K., Morrison, I., and Green J. 2003. Lidar remote sensing of biophysical properties of tolerant northern hardwood forests. Canadian Journal of Remote Sensing 29(5): 648–678.

Lindberg, 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.

Magnussen, S., Eggermont, P. & LaRiccia, V. N. 1999. Recovering tree heights from airborne laser scanner data. Forest Science 45(3): 407-422.

Malinen, J. 2003. Locally adaptable non-parametric methods for estimating stand characteristics for wood-procurement planning. Silva Fennica 37(1): 109-120.

—, Maltamo, M. and Harstela, P. 2001. Application of most similar neighbor inference for estimating marked stand characteristics using harvester and inventory generated stem data banks. International Journal of Forest Engineering 12(2): 33-41.

Maltamo, M., Kangas, A., Uuttera, J., Torniainen, T. & Saramäki, 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: 263-274.

—, Mustonen, K., Hyyppä, J., Pitkänen, J. and Yu, X. 2004a. The accuracy of estimating individual tree variables with airborne laser scanning in a boreal nature reserve. Canadian Journal of Forest Research 34(9): 1791-1801.

—, Eerikäinen, K., Pitkänen, J., Hyyppä, J. and Vehmas, M. 2004b. 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.

—, Eerikäinen, K., Packalén P. & Hyyppä, J. 2006a. Estimation of stem volume using laser scanning-based canopy height metrics. Forestry 79(2): 217-229.

—, Hyyppä, J. & Malinen J. 2006b. A comparative study of the use of laser scanner data and field measurements in the prediction of crown height in boreal forests. Scandinavian Journal of Forest Research 21: 231-238.

—, Suvanto, A. & Packalén, P. 2007. Comparison of basal area and stem frequency diameter distribution modelling using airborne laser scanner data and calibration estimation. Forest Ecology and Management 247(1): 26-34.

—, Peuhkurinen, J., Malinen, J., Vauhkonen, J., Packalén P. & Tokola, T. 2009a. Predicting tree attributes and quality characteristics of Scots pine using airborne laser scanning data. Silva Fennica 43(3): 507-521.

—, Næsset, E., Bollandsås, O. M., Gobakken, T. & Packalén, P. 2009b. Nonparametric prediction of diameter distributions using airborne laser scanner data. Scandinavian Journal of Forest Research 24(6): 541-553.

—, Packalén, P., A. Suvanto, A., Korhonen, K. T., Mehtätalo, L. & Hyvönen, P. 2009c. Combining ALS and NFI training data for forest management planning: a case study in Kuortane, Western Finland. European Journal of Forest Research 128(3): 305-317.

—, Bollandsas, O.-M., Vauhkonen, J., Breidenbach, J., Gobakken, T. & Næsset, E. 2010. Comparing different methods for prediction of mean crown height in Norway spruce stands using airborne laser scanner data. Forestry 83(3): 257-268.

—, Bollandsås, O. M., Næsset, E., Gobakken, T. & Packalén, P. 2011. Different sampling strategies for field training plots in ALS-assisted forest inventory. Forestry 84(1): 23-31.

Means, J.E., Acker, S.A., Brandon, J.F., Renslow, M., Emerson, L. and Hendrix, C.J. 2000. Predicting forest stand characteristics with airborne scanning lidar. Photogrammetric Engineering & Remote Sensing 66(11): 1367-1371.

Mehtätalo, L. 2002. Valtakunnalliset puukohtaiset tukkivähennysmallit männylle, kuuselle, koivulle ja haavalle. Metsätieteen aikakauskirja 4/2002: 575-591.

— 2005. Height-diameter models for Scots pine and birch in Finland. Silva Fennica 39(1): 55–66.

— & Nyblom, J. 2009. Estimating forest attributes using observations of canopy height: a model-based approach. Forest Science 55(5): 411-422.

—, Maltamo, M., and Packalén, P. 2007. Recovering plot-specific diameter distribution and height-diameter curve using ALS based stand characteristics. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 36(3/W52): 288-293.

Melkas, T., Vastaranta, M., Haapanen, R., Holopainen, M., Yu, X., Hyyppä, J. & Hyyppä, H. 2009. Updating forest resource data by using ALS measurements and information collected with a harvester. In: Popescu, S., Nelson, R., zhao, K. and Neuenschwander, A. (ed). Proceedings of Silvilaser 2009: The 9th international conference on lidar applications for assessing forest ecosystems. Texas A&M Univerisity College Station, USA. p. 128-136.

Moeur, M. & Stage, A.R. 1995. Most Similar Neighbor. An improved sampling inference procedure for natural resource planning. Forest Science 41: 337-359.

Möller, J. & Arlinger, J. 2007. [Online document]. Virkesvärdestest 2006 - mätnoggrannhet. Resultat Nr 5, Skogforsk, 4 p. Available at: http://www.skogforsk.se/PageFiles/48129/Resultat5-2007.pdf.

Murphy, G. 2008. Determining stand value and log product yields using terrestrial lidar and optimal bucking: A case study. Journal of Forestry 106(6): 317-324.

Näslund, M. 1936. Skogsförsökanstaltens gallringsförsök i tallskog. Meddelanden från Statens Skogsförsöksanstalt 29. 169 p.

Næsset, E. 1997. Estimating timber volume of forest stands using airborne laser scanner data. Remote Sensing of Environment 61(2): 246-253.

— 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.

— & Økland, T. 2002. Estimating tree height and tree crown properties using airborne scanning laser in boreal nature reserve. Remote Sensing of Environment 79: 105-115.

Nelson, R.F., Krabill, W.B. & Maclean, G.A., 1984. Determining forest canopy characteristics using airborne laser data. Remote Sensing of Environment 15: 201–212.

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

Nyyssönen, A. Ojansuu, R. 1982. Metsikön puutavaralajirakenteen, arvon ja arvokasvun arviointi. Summary: Assesment of timber assortments, value and value increment of tree stands. Acta Forestalia Fennica 179. 52p.

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

Packalén, P. 2009. Using airborne laser scanning data and digital aerial photographs to estimate growing stock by tree species. Dissertationes Forestales 77. 41 p.

— & Maltamo, M. 2006. Predicting the plot volume by tree species using airborne laser scanning and aerial photographs. Forest Science 52(6): 611–622.

— & 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.

— & Maltamo, M. 2008. Estimation of species-specific diameter distributions using airborne laser scanning and aerial photographs. Canadian Journal of Forest Research 38: 1750–1760.

—, Pitkänen J., and Maltamo, M. 2008. Comparison of individual tree detection and canopy height distribution approaches: a case study in Finland. In: Hill R.A., Rosette, J. & Suárez, J. (ed.). Proceedings of SilviLaser 2008: 8th international conference on LiDAR applications in forest assessment and inventory. Edinburgh, UK. p. 22-29.

Päivinen, R. 1983. Metsikön tukkiosuuden arviointimenetelmä. Summary: a method for estimating the sawlog percentage in Scots pine and Norway spruce stands. Folia Forestalia 564. 16p.

Pesonen, A., Maltamo, M., Eerikäinen, K & Packalèn, P. 2008. Airborne laser scanning-based prediction of coarse woody debris volumes in a conservation area. Forest Ecology and Management 255: 3288-3296.

Pitkänen, J. 2005. A multi-scale method for segmentation of trees in aerial images. In: Hobbelstad, K. (ed.). Forest Inventory and Planning in Nordic Countries, Proceedings of SNS Meeting at Sjusjøen, Norway. Norwegian Institute of Land Inventory. p. 207-216.

—, Maltamo, M., Hyyppä, J., and 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(8/W2): 187–191.

PMP-ohje. 1982. PMP-systeemin hoitokunta.

Popescu, S. & Zhao, K. 2008. A voxel-based lidar method for estimating crown base height for deciduous and pine trees. Remote Sensing of Environment 112: 767-781.

—, Wynne, R. & Nelson. R. 2003. Measuring individual tree crown diameter with lidar and assessing its influence on estimating forest volume and biomass. Canadian Journal of Remote Sensing 29(5): 564–577.

Räsänen, T., Imponen, V., Lindroos, J., Malinen, J. and Sorsa, J-A. 2000. Runkopankki puunhankinnan ohjauksen välineenä. Metsätehon raportti 94. 42 p.

Rasinmäki, J. & Melkas, T. 2005. A method for estimating tree composition and volume using harvester data. Scandinavian Journal of Forest research 20: 85-95.

Reynolds, M. R., Jr., 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.

R Development Core Team. 2010. [Online document] R: A language and environment for statistical computing. R Foundation for Statistical Computing. ISBN 3-900051-07-0. Available at: http://cran.r-project.org/doc/manuals/refman.pdf.

Roberts, S. D., Dean, T. J., Evans, D. L., McCombs, J.W., Harrington, R. L. & Glass, P. A. 2005. Estimating individual tree leaf area in loblolly pine plantations using LiDAR-derived measurements of height and crown dimensions. Forest Ecology and Management 213: 54-70.

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

Siipilehto, J. 1999. Improving the accuracy of predicted basal-area diameter distribution in advanced stands by determining stem number. Silva Fennica 33(4): 281-301.

SOLMU. Maastotyöopas. 2003. Metsätalouden kehittämiskeskus Tapio, Helsinki. 78 p.

Solodukhin, V.I., a Ya Zhukov, I.N. Mazhugin, T.K. Bokova, and V.M. Polezhai. 1977. Vozmozhosti lazernoi aeros emka pro_lei lesa (possibilities of laser aerial photography of forest profiles). Lesn. Khoz 10: 53-58.

Standard for Forest Data and commincations. 2007. [Online document]. Skogforsk. Available at: http://www.skogforsk.se/upload/6867/StanForD_MainDoc_070327.pdf.

Suvanto, A., M. Maltamo, P. Packalén and J. Kangas. 2005. Kuviokohtaisten puustotunnusten ennustaminen laserkeilauksella. Metsätieteen aikakauskirja 4/2005:413-428.

Thies, M., Pfeifer, N., Winterhalder, D. & Gorte, B. 2004. Three-dimensional reconstruction of stems for assessment of taper, sweep and lean based on laser scanning of standing trees. Scandinavian Journal of Forest Research 19(6): 571-581.

Thomas, V., Oliver, R. D., Lim, K. & Woods, M. 2008. Lidar and Weibull modeling of diameter and basal area. The Forestry Chronicle 84(6): 866-875.

Tommola, M., Tynkkynen, M., Lemmetty, J., Harstela P. & Sikanen, L. 1999. Estimating the characteristics of a marked stand using k-nearest-neighbour regression. International Journal of Forest Engineering 10(2): 75-81.

Tuominen, S. & Pekkarinen, A. 2004. Local radiometric correction of digital aerial photographs for multi source forest inventory. Remote Sensing of Environment 89: 72-82.

—, Fish, S. & Poso, S. 2003. Combining remote sensing, data from earlier inventories, and geostatistical interpolation in multisource forest inventory. Canadian Journal of Forest Research 33: 624-634.

Uusitalo, J. 1995. Pre-harvest measurement of pine stands for sawing production planning. University of Helsinki, Department of Forest Resource Management Publications 9, Helsinki, Finland. 96 p.

Vastaranta, M., Holopainen, M., Haapanen, R., Yu, X., Melkas, T., Hyyppä, J., and Hyyppä, H. 2009a. Comparison between an area-based and individual tree detection method for low-pulse density ALS-based forest inventory. International Archives of Photogrammetry, Remote Sensing and Spatial Information Sciences 38(3/W8): 147–151.

Vastaranta, M., Holopainen, M., Kaartinen, H., Hyyppä, H. and Hyyppä, J. 2009b. Uudistuneet metsien maastomittaustarpeet. Metsätieteen aikakauskirja 4/2009: 370-374.

Vehmas, M., Eerikäinen, K., Peuhkurinen, J., Packalén, P. & Maltamo, M. 2008. Airborne laser scanning for the identification of boreal forest site types. In: Hill R.A., Rosette, J. & Suárez, J. (ed.). Proceedings of SilviLaser 2008: 8th international conference on LiDAR applications in forest assessment and inventory. Edinburgh, UK. p. 58-65.

Veltheim, T. 1987. Pituusmallit männylle, kuuselle ja koivulle. M.Sc. thesis, Faculty of Agriculture and Forestry, University of Helsinki, Finland. 59 p.

Vauhkonen, J. 2010. Estimating crown base height for Scots pine by means of the 3-D geometry of airborne laser scanning data. International Journal of Remote Sensing 31(5): 1213-1226.

—, Tokola, T., Packalén, P. and Maltamo, M. 2009. Identification of Scandinavian Commercial Species of Individual Trees from Airborne Laser Scanning Data Using Alpha Shape Metrics. Forest Science 55(1): 37-47.

—, 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.

Weber, T.C. & Boss, D.E. 2009. Use of LiDAR and supplemental data to estimate forest maturity in Charles County, MD, USA. Forest Ecology and Management 258: 2068–2075.

Wehr, A. & Lohr, U. 1999. Airborne laser scanning—an introduction and overview. Photogrammetry & Remote Sensing 54: 68–82.

What is StanForD? 2010. [Internet site]. Skogforsk. Available at: http://www.skogforsk.se/sv/Welcome-to-Skogforsk/About-Skogforsk/Collaboration-/StanForD/ [Cited 8 April 2010].

Zhao, K., Popescu, S. & Nelson, R. 2009. Lidar remote sensing of forest biomass: A scale-invariant estimation approach using airborne lasers. Remote Sensing of Environment 113: 182-196.